

other half, and then expand the composite function as a Fourier series. The principal difficulty caused by the inclusion is that the expansion coefficients do not decrease as rapidly as in the case without an inclusion. The method will be applied to a physical three-dimensional case in a future publication.

We wish to thank Mr M. Woodcock for advice and assistance with the programming.

Acta Cryst. (1972). A28, 270

Limits of Phase Expansion in Direct Methods

BY J. GASSMANN AND K. ZECHMEISTER

Abteilung für Röntgenstrukturforschung am Max-Planck-Institut für Eiweiss- und Lederforschung, Schillerstrasse 46, 8 München 15, Germany (BRD)

(Received 22 November 1971 and in revised form 31 January 1972)

Phase expansion starting from a few initial phases is investigated with reference to the size of the starting set, phase errors in the starting set, the lower limit of the E value in expansion and the different phase-determining formulae. The results stress the need for a sufficient size of the initial phase set with small phase errors for subsequent application of a phase-expanding procedure. The common basis of phase-expansion procedures is shown to consist of a cyclic modification of the preliminary structure and explains the impossibility of correcting substantial errors in already known phases associated with larger E values by subsequent phase determination for smaller E values. The phenomenon of losing structure information by careless application of the tangent formula and consequently the appearance of partial structures is pointed out. This information-destroying phase expansion is shown not to exist in the 'phase-correction' procedure.

Introduction

Several direct methods of crystal structure analysis, differing in their theoretical foundation and practical application, have been developed. The most widely used approaches are the symbolic addition procedure and the multiresolution method, several different versions of which have been programmed. Whereas the symbolic addition requires some manual intervention, the multiresolution method may be done automatically. (Karle & Karle, 1966; Germain & Woolfson, 1968).

Although many structures have been determined by direct methods, in some cases, for no obvious reasons, there occur difficulties in finding a structure. Sometimes a different initial phase set is successful. In other examples only partial structures have evolved. This investigation has been set up to find reasons for this behaviour of direct methods. As test examples we have chosen structures solved by the symbolic addition procedure, multiresolution and structure invariant method (Hauptman, Fisher, Hancock & Norton, 1969).

Theoretical considerations

Apart from the initial stages of phase determination in the above-mentioned methods, the addition of new

References

- GRADSHTEYN, I. S. & RYZHIK, J. M. (1965). *Table of Integrals, Series and Products*. New York and London: Academic Press.
- NELDER, J. A. & MEAD, R. (1965). *Comput. J.* 7, 308.
- STUHRMANN, H. B. (1970a). *Acta Cryst.* A26, 297.
- STUHRMANN, H. B. (1970b). *Z. Phys. Chem.* 72, 177.
- STUHRMANN, H. B. (1970c). *Z. Phys. Chem.* 72, 185.
- WU, H.-I. & SCHMIDT, P. W. (1970). *J. Appl. Cryst.* 4, 224.

phases is done by the tangent formula (Cochran 1955; Karle & Hauptman, 1956). A general form of this formula is

$$U_h = \langle E_k E_{h-k} \rangle_k, \quad (1)$$

where U = unitary structure factor, and E = normalized structure factor.

The index k indicates a summation over reflexions. Considering a single contribution with large $E_k E_{h-k}$ this leads to the triple product relation for phases

$$\varphi_h = \varphi_k + \varphi_{h-k}. \quad (1a)$$

The summation taken over the largest products $E_k E_{h-k}$ leads to the tangent formula

$$\varphi_h = \text{phase of } \langle E_k \cdot E_{h-k} \rangle_k, \\ \text{where } k = \text{limited set of reflexions}. \quad (1b)$$

If the summation in equation (1) is taken over all reflexions this is called the Hughes formula and is equivalent to squaring the electron density in direct space (Sayre, 1952).

The initial situation for the addition of new phases in direct methods is as follows. A restricted number of phases associated with large E values has been determined. These phases, together with their observed E values constitute an 'electron density' ϱ_m of a 'prelimi-

nary structure'. The only possible way to calculate new phases for reflexions not yet included is by modification of this preliminary structure. In direct space this modification is a change of the electron density ϱ_m to a new electron density ϱ_m^* . Fourier transformation of ϱ_m^* allows the calculation of new phases φ_m^* and E values E_{calc} . For eventual success the new electron density ϱ_m^* ought to progress towards the final correct electron density ϱ . This means that the newly calculated Fourier coefficients should agree better with the observed information. To improve convergence to the correct structure only those Fourier coefficients are kept whose modulus E_{calc} is large. The phases φ_m^* of these Fourier coefficients are combined with the observed E values E_{obs} to give a new preliminary structure:

$$E_{\text{new}} = |E_{\text{obs}}| \cdot \exp(i\varphi_m^*),$$

where φ_m^* = phase of modified electron density ϱ_m^* .

The process is then repeated. Modification is made by routine procedures which do not take into account the properties of the structure in question. These properties are only reflected by the observed E values.

The most commonly used modification is squaring, which is formulated in reciprocal space as the tangent formula. In the tangent formula the selection of the Fourier coefficients with large values of E_{calc} is done automatically by including only large terms $E_k E_{h-k}$ in the summation of equation (1b).

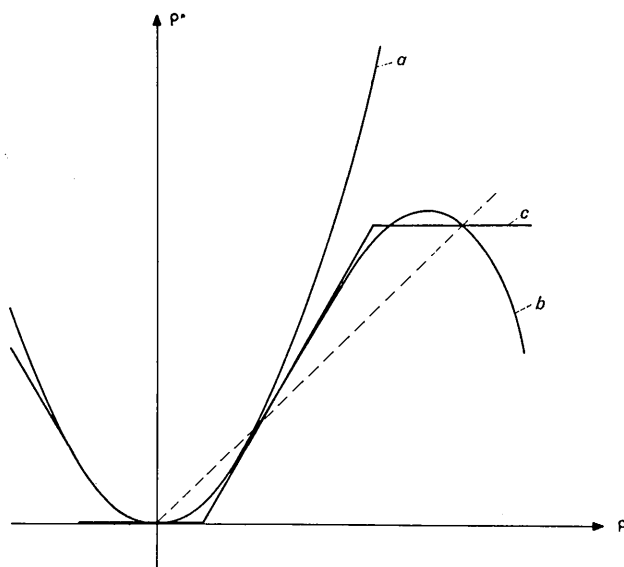


Fig. 1. Different forms of phase expansion represented in direct space. Modification after the parabolic curve (a) is equivalent to phase expansion with the tangent formula. The polynomial form (b) is the equivalent of equation 2 (see text). The linear form (c) involves higher convolutions in reciprocal space. The preliminary electron density ϱ is changed into ϱ^* corresponding to the different modification curves. The squaring effect is only exhibited by the tangent formula.

The phases calculated with the tangent formula are the phases of the squared preliminary structure ($\varrho_m^* = \varrho_m^2$).[†] This preliminary structure ϱ_m and also its square ϱ_m^* normally consist of peaks with different heights. This range of peak heights is not just a result of phase errors but occurs even with exact phases due to the limited number of known phases. This range of peak heights is increased by squaring ϱ_m , thereby suppressing weak features and overemphasising strong features present in the former electron density ϱ_m . If these suppressed features contain indications of the correct structure, then the application of the tangent formula destroys structure information.

The common outcome of such an information-destroying process is a Fourier synthesis with one or a few large peaks. In terms of equation (1b) this means that many products $E_k E_{h-k}$ have about the same phase and 'line up' in the complex plane. There exist less extreme stages of this phase-destroying use of the tangent formula, where a few correct atomic peaks are left and the identification of a partial structure may be possible. If phase determination is expanded[‡] to small E values, the range of the peak heights will be larger and consequently the Fourier interpretation will be more difficult. The possibility of losing structure information is enhanced if the initial phase set already contains considerable phase errors. This is of special concern in the symbolic addition procedure, at the point when actual phases are substituted for symbols. Compared with the multisolution approach where equation (1b) can be applied from the very beginning, the equivalent phase set of the symbolic addition procedure in many cases contains larger phase errors, thereby increasing the danger of finding only a partial structure after phase expansion with the tangent formula, due to the squaring effect.

To suppress this squaring effect various formulae have been given (Hoppe & Gassmann, 1968; Hoppe, Gassmann & Zechmeister, 1970). One of these formulae has a form similar to the tangent formula,

$$\varphi_h = \text{phase of } \{ \langle E_k E_{h-k} \rangle_k + a \langle \langle E_l \cdot E_{k-l} \cdot E_{h-k} \rangle \rangle_{k,l} \} \quad (2)$$

[†] There exists a minor difference between squaring and the tangent formula with respect to the redetermined phases of already included reflexions. The terms $E_o E_h$ in equation (1b) are normally ignored. This is equivalent in direct space to:

$$\varrho_m^* = (\varrho_m - \varrho_o)^2 \text{ with } \varrho_o = E_o/V; V = \text{cell volume.}$$

A contribution occurs only for the already determined, known phases which tend back to their former values. With an increasing number of known phases ($\varrho_m \gg \varrho_o$) the contributions from ϱ_o get less important:

$$\varphi_m^* = \text{phase of } \{ \varrho_m^2 - 2\varrho_o\varrho_m + \varrho_o^2 \}$$

$$= \text{phase of } \{ \varrho_m^2 - 2\varrho_o\varrho_m \} \simeq \text{phase of } \{ \varrho_m^2 \}.$$

[‡] Phase expansion in this context is defined as the determination of new phases associated with E values above a certain limit. Normally the lowest limit of phase expansion is accepted as $E=1.0$. This phase expansion is different from the phase extension with respect to $\sin \theta/\lambda$ (Hoppe, 1962; Hoppe & Gassmann, 1964; Kartha, 1969), where new phases are calculated to get better resolution.

where the coefficient a depends on the size of the structure. More efficient phase determining formulae including higher convolutions may be written down more easily in direct space. A very convenient type is the so called linear form. Fig. 1 shows the squaring tangent formula, the polynomial form after equation (2) and the linear form in direct space.

Phase determination with direct-space treatment of electron density then consists of three steps: calculation of a Fourier synthesis to get the initial electron density, modification of this electron density and a reciprocal Fourier transformation to get the new phases. The new determined phases, the changed initial phases and the observed E values constitute the starting point for the next cycle of the above mentioned phase determination, which was called 'phase correction'.

Another approach for correcting phases has been given by Karle (1968). In this procedure steps two and three of 'phase correction' are replaced by interpretation of the Fourier synthesis and subsequent structure factor calculation. The calculated phases meeting certain criteria (*e.g.* ratio of $F_{\text{calc}}/F_{\text{obs}}$ large) are used as input to the tangent formula. The effect of this procedure may be seen as follows. The Fourier synthesis interpretation and selection of possible atoms suppresses the peak range present in the preliminary structure. Smaller peaks are eliminated in order to increase the correct portion of the Fourier synthesis. This procedure ignores any other information besides large 'reasonable' peaks present in the Fourier synthesis as well as the different peak heights of the remaining 'atomic' peaks. The phases determined in the structure factor calculation together with the observed E values represent a preliminary structure with different peak heights, especially between postulated and unknown atomic peaks. The subsequent use of the tangent formula again enforces the increase in the range of peak heights. The continuation of the proce-

sure is only possible if the resulting phases lead to a Fourier synthesis with more interpretable information. This may not always be the case.

A similar situation of finding only a partial structure may occur if the starting phase set is too restricted.* To determine additional phases more interactions with less probable phase indications have to be taken into account. Therefore, phases of large E values are determined with larger errors, propagating these errors to the phases of smaller E values. This is especially important if a whole subset of phases is dependent on one single phase which has been determined wrongly. The equivalent situation in the symbolic addition procedure is the incorrect choice of the phase of a symbol associated with a large E value. In most cases the resulting Fourier synthesis will not be interpretable. In structures with high internal symmetry there may result partially shifted structures (Bürgi & Dunitz, 1971). As discussed above it will not be possible to correct these phase errors by extending the phase determination to phases associated with smaller E values. This is independent of the choice of any particular phase-determining relation, which is always based on the squaring or modification of the preliminary structure. A multisolution approach with variation of starting phases leads definitely to the correct structure in addition to partially shifted structures, although a choice of the correct solution, based on the phase determination criteria (*e.g.* consistency of phases), might be difficult. In general, several Fourier syntheses will have to be calculated. Their interpretation will be

* We refer only to the restriction on the number of initial phases. The case of a restricted number of reflexions with high E values, due to limited experimental data, might also lead to difficulties but is not considered here. Our experience has shown that structures could be solved by direct methods using data out to $\sin \theta/\lambda = 0.60$, whereas the same structures could not be solved with limited data to $\sin \theta/\lambda = 0.45$.

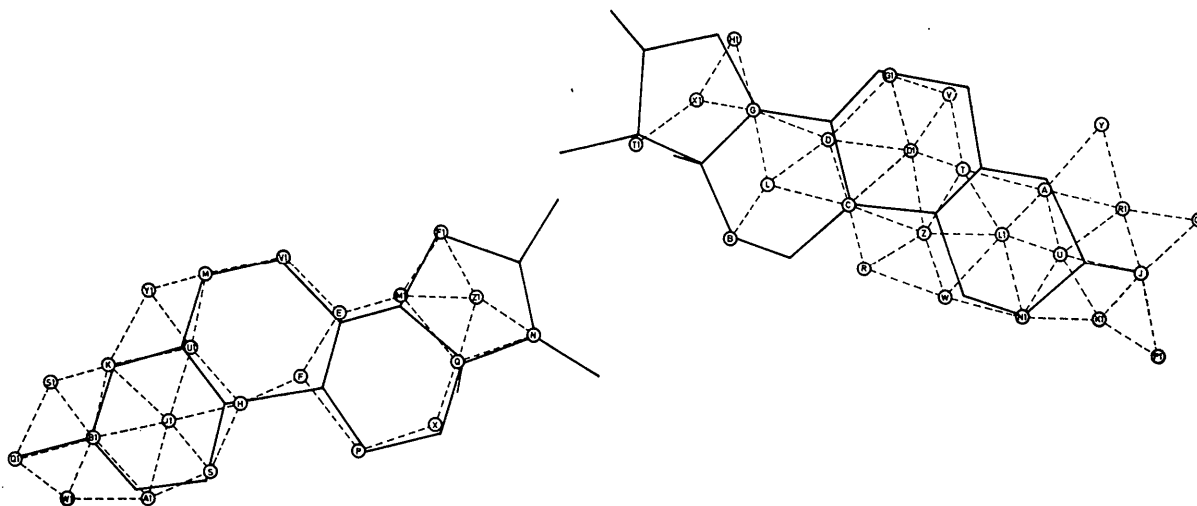


Fig. 2. Fourier synthesis interpretation of estriol. The Fourier synthesis calculated with 41 phases out of the 50 highest E values is shown by dashed lines. Phase expansion with the linear form of phase correction led to the total structure (full lines) with no interspersed spurious peaks.

facilitated by automatic peak searches combined with calculation of intra- and intermolecular distances.

Experimental results

Actual examples of structure determinations have been used to demonstrate the influence of the effects described above and their importance in the different direct-method approaches.

Estriol

The crystal structure of estriol ($C_{18}H_{24}O_3$)₂ has been solved by Cooper, Norton & Hauptman (1969) by the determination of structure invariants followed by phase determination using the tangent formula.

We used estriol as a test structure to answer the following questions:

- is it possible to solve the structure with the multi-solution method;
- is it possible to identify the structure without phase expansion after the tangent formula;
- what is the effect of phase expansion;
- what result occurs in using 'phase correction' instead of 'tangent formula'?

We used the following starting set (space group $P2_1$, angles in cycles):

h	k	l	E	Phase
$\bar{1}$	14	4	3.27	0
$\bar{1}$	15	4	3.13	0.125
$\bar{3}$	13	5	3.07	0.125
0	4	6	3.30	0.125/0.375
2	0	0	5.78	0/0.5

The three-dimensional reflexions and reflexion 046 were used to fix the origin and enantiomorph. The four possibilities were calculated deliberately confining the determination to 41 phases out of the 50 highest E values. All other E values were assumed to be zero in the process of phase determination.

Fig. 2 shows the resulting peaks in the Fourier synthesis (dashed lines) with the best statistical criteria. An interpretation of the structure at this stage would have been difficult. Comparison with the correct structure showed the presence of 32 atoms which were displaced on the average by 0.33 Å.

Changing the above restriction on the number of E values to 100 and 300 resulted in Fourier syntheses where 35 and 36 atoms respectively appeared, together with 34 to 38 spurious interspersed peaks. There was no essential change in the Fourier synthesis calculated with 300 phases instead of 100 phases. The main effect of using 300 phases was an increase in the range of atomic peak heights (60–13) compared to the former range of 26–10.

To demonstrate this tendency, namely to generate high peaks, we expanded the 41 original phases to reflexions with E values ≥ 1.0 using all contributing terms in the tangent formula. This is equivalent to a repeated squaring of the electron density of the preli-

minary structure in direct space (see Fig. 1). About 1000 phases could be determined by 'squaring'. The resulting Fourier synthesis showed one high peak; the structure was lost.

A second trial with the additional condition, that phases were accepted only if $E_{\text{calc}} \geq 0.2E_{\text{obs}}$ left 10 partially connected atomic peaks in the Fourier synthesis which could hardly be interpreted (see Table 1). The squaring effect on the first high peaks makes the smaller atomic peaks disappear in the background.

Table 1. Comparison of phase expansion in estriol using the tangent formula and phase correction

The tendency to generate high peaks and to lose structure information in using the tangent formula is demonstrated. Starting position for tangent formula and phase correction were 41 phases. The 32 atoms present in the Fourier synthesis phased with these 41 phases (see Fig. 2) get partially lost after applying the tangent formula. The total structure is recognizable after 30 cycles of phase correction.

Phase correction linear function (30 cycles)		Tangent formula squaring $q \rightarrow q^2$ (15 cycles)	
Value	Name	Value	Name
75.9	C(29)	215.9	C(10)
74.8	O(3)	108.6	C(8)
73.3	C(25)	54.9	C(2)
72.5	C(23)	54.7	C(13)
70.0	C(4)	46.3	C(6)
69.2	C(8)	45.3	C(15)
66.3	C(10)	45.0	C(25)
64.6	C(30)	41.9	C(4)
63.6	C(11)	40.8	C(11)
60.0	C(21)	35.0	C(7)
59.8	C(15)	33.4	C(29)
58.7	C(13)	31.6	—
57.4	C(17)	29.8	—
57.3	C(5)	29.0	—
56.7	O(16)	28.2	—
55.9	O(23)	27.0	—
55.8	C(24)	26.2	—
55.5	C(6)	24.5	—
55.5	C(22)	23.7	C(14)
55.0	C(14)	23.4	—
54.2	C(27)	23.4	—
53.7	C(2)	23.2	—
53.2	O(37)	22.4	C(21)
53.0	C(3)	22.4	—
52.2	O(17)	22.2	—
51.8	C(37)	21.3	—
50.2	C(34)	20.9	—
50.0	O(36)	20.9	O(23)
49.7	C(26)	20.7	—
49.5	C(9)	20.7	—
47.9	C(12)	20.5	C(9)
47.4	C(36)	20.5	—
45.4	C(1)	20.1	—
44.6	C(7)	19.9	C(22)
44.0	C(16)	19.8	—
42.7	C(33)	19.7	—
40.7	C(32)	19.6	—
40.6	C(28)	19.6	—
40.0	C(31)	19.2	—
35.7	C(38)	19.1	—
33.7	C(35)	19.0	—
20.9	C(18)	18.9	—
20.8	—	18.9	—
18.5	—	18.8	—

In a third trial we expanded the 41 phases by 'phase correction' using the linear from (see Fig. 1) down to a minimum E value of 1.0. After 30 cycles all 42 atoms appeared with no spurious peaks between (see Table 1 and Fig. 2).

Photolysis product ($C_{12}H_{13}NO_4$)

This crystal structure has been solved by Karle, Karle & Estlin (1967). A partial structure was obtained by the application of the symbolic addition procedure (Karle & Karle, 1966). The complete structure evolved by the use of the tangent formula in a recycling procedure (Karle, 1968).

We used this structure to clarify several points:

- does the multisolution method, like symbolic addition, lead to a partial structure;
- what is the comparison between the use of the tangent formula in the recycling procedure and phase correction;
- what are the most favourable requirements and parameters for using the tangent formula?

We took the same starting phase set as Karle, Karle & Estlin (1967) and calculated 74 phases for reflexions

with $E \geq 1.6$. These initial phases were expanded to $E \geq 1.2$, including 217 phases in all. Since our procedure was identical with the method applied in the course of solving the structure we expected to find a partial structure.*

Our resulting Fourier synthesis showed 16 atoms with 9 interspersed spurious peaks and the structure interpretation was possible unambiguously; the first two spurious peaks occurred after seven atomic peaks (see Table 2, 6th column). To get a partial structure we confined our data to $E \geq 1.8$, calculating 35 phases. Eleven of the phases deviated by more than one radian from the correct values. The Fourier synthesis showed eight atomic peaks with 14 interspersed spurious peaks (Table 2, 1st column). Four of the eight atomic peaks were connected by interatomic distances. The phases of this artificial partial structure were extended down to $E \geq 1.3$. This resulted in a Fourier synthesis calculated, with 171 phases, which could not be inter-

* We used all the data given in the above mentioned paper, including the observed structure factors. Deviations in scaling due to the Weissenberg recording technique were not taken into account.

Table 2. Comparison of peak heights and identified atoms for photolysis product ($C_{12}H_{13}NO_4$) using tangent formula expansion or phase correction

Asterisks indicate peaks identified as atoms. An expansion to an E value of 1.4, determining 128 phases shows the best result for the tangent formula. Further expansion results in a very high peak for the first atom (squaring effect). Application of phase correction does not show the squaring effect and leads to a stable solution where no spurious peaks occur among the atomic peaks.

Determined phases	Tangent formula								Squaring	Phase correction	
	35	58	74	128	173	217	245	331	300	270	300
Minimum E value of expansion	1.8	1.7	1.6	1.6-1.4	1.6-1.3	1.6-1.2	1.6-1.15	1.6-1.0	1.6-1.0	1.6-1.0	1.8-1.0
Identified atoms	8	10	10	16	14	16	15	16	13	17	17
Peak heights	15*	22	34*	53*	70*	82*	100*	108*	132*	61*	67*
	15	21*	21*	43*	58*	68*	56*	83*	51*	58*	67*
	15	19*	21	32*	36*	39*	39*	50*	37*	57*	64*
	14*	18	20*	32*	36*	35*	33*	37*	34*	52*	60*
	14	18	20	30*	31*	34*	31	36*	30	49*	54*
	14*	17	19*	26*	31*	33*	31*	34*	26	48*	54*
	14*	16*	19	24	25	29*	28	29*	24*	48*	52*
	14	16*	18*	23*	23*	23	24	28*	23	47*	51*
	14*	16	16*	22*	22*	22	23*	28*	23*	46*	50*
	13	15	15*	20*	22*	21*	23*	27*	22*	45*	50*
	13	15	15*	19*	21*	21*	22*	26	22*	41*	49*
	13	14*	14	19*	21	21*	22*	25	21	40*	45*
	12*	14	14	18*	20*	21*	20*	25*	20*	35*	37*
	12	14	14	18*	19	20	19	24*	19*	33*	34*
	12	14*	14*	18*	18*	20*	19*	23*	19*	32*	30*
	12	13	14	17	17	19*	19*	20	19	31*	29*
	12	13*	14	17*	17*	19*	19	20	18	28*	27
	12	13	13	16	16	19	18*	20*	18	25	25*
	12	13	13	16	16	19	18*	18	18*	20	16
	12	13*	13	16	16	18	18	18	18	19	16
	11*	13	13	15*	16	18	17	18*	18	16	
	11*	12*	13*	15	16	18*	17*	18*	17		
	11	12	13	14	16	17	17	17	17		
	11	12	13	14	16*	17	17	17	17		
	11	12*	13	14	16	17*	17	17	17*		
	10	12	13	14	16	17	16	17	16		

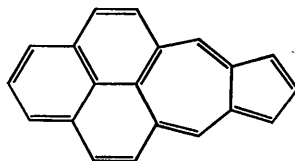
puted. If only those terms in the tangent formula contributed, which were above a given probability limit, the Fourier synthesis showed all 17 atoms with two interspersed spurious peaks. This shows that a careless application of the tangent formula destroys valuable phase information. Further experiments indicated that phase expansion below $E \geq 1.4$ does not contribute new structural information (see Table 2). For this purpose 74 phases ($E \geq 1.6$) were chosen as starting set. Expansions by the tangent formula to $E \geq 1.0$ did not improve the Fourier synthesis. On the contrary the range of peak heights increased (squaring effect) as can be seen from Table 2.

Whereas the simple application of the tangent formula starting from 35 phases ($E \geq 1.8$) did not yield the correct structure, phase correction from the same starting point proceeded to the structure. The Fourier synthesis contained only one interspersed spurious peak (see Table 2, last column). The range of peak heights was considerably lower than the comparable case ($E \geq 1.0$) with application of the tangent formula (see Table 2, 4th-last column).

In all cases of tangent formula expansion only the fifty largest terms $E_k E_{h-k}$ contributed to the calculation of the phase φ_h . If all terms $E_k E_{h-k}$ were taken into account the squaring effect inherent in the tangent formula was much more clearly pronounced as can be seen in the third last column of Table 2. A further expansion below $E \geq 1.0$ would lead to the same 'large-peak' Fourier synthesis already encountered in the case of estriol. In all the Fourier syntheses at different expansions the first four atomic peaks located were identical, indicating the squaring effect on the initial structure information. This initial structure naturally did not consist of atoms with equal peak height as can be seen in Table 2, column 3. Whereas the range of peak heights in the initial structure is $34:13 \approx 3$, this ratio is increased to $108:18 \approx 6$. To keep a small ratio of peak heights is of essential importance for the interpretation of a Fourier synthesis, since small atomic peaks get lost in the background.

Azuleno phenalene, ($C_{20}H_{12}$)₂

The crystal structure has been solved by Brandl, Graziani & Zechmeister (1971). The structure of azuleno [5,6,7-*cd*] phenalene (I) was proposed from chemical evidence by Jutz & Kirchlechner (1966).



(I)

This structure could not be solved by the routine application of the multisolution approach which resulted in a partial structure. We examined the reasons for getting the partial structure and reached the total structure from different initial phase (sign) sets.

After measuring the intensity data it became immediately apparent that to determine this structure by direct methods would give difficulties. 2507 reflexions out to $\theta = 70^\circ$ were collected on a diffractometer. Of these, 1386 reflexions were too weak to be measured, leaving 1121 observed reflexions. Though Wilson statistics indicated a centre of symmetry, it turned out that the correct space group was $Pna2_1$ with two molecules in the asymmetric unit. These two molecules could be approximately superposed on each other by a non-crystallographic screw axis. This pseudo screw axis was also indicated by the partial extinction of the corresponding reflexions. Using the normal multisolution method, the resulting Fourier syntheses could not be interpreted. Observing the centric distribution of the intensities and the form of the molecule, it was decided to start with signs instead of phases. This procedure, which leads to the centrosymmetric supergroup $Pnam$ of the non-centrosymmetric space group $Pna2_1$, has been proposed by Burzlaff (1970). The choice of an initial sign set requires the determination of signs instead of phases, corresponding to a superposition of the structure and its enantiomorph. The actual sign set used was:

<i>h</i>	<i>k</i>	<i>l</i>	<i>E</i>	sign	} origin defining
2	2	1	5.46	+	
7	6	0	4.39	+	
1	13	2	3.80	+	
1	5	18	3.86	+/-	
1	6	15	3.70	+/-	

which gave four solutions. The highest peaks of the Fourier synthesis calculated with 200 signs of the most consistent phase set are shown in Fig. 3. It would be very difficult to unravel the structure and its enantiomorph even though the molecular configuration was known approximately and, as it appeared later, 28 peaks were in the vicinity of atomic locations.

To separate the structure from its enantiomorph, phase correction using the linear form was applied, including all reflexions with $E \geq 1.0$ and $E_{calc} \geq 0.2E_{obs}$. To reach a non-centrosymmetric configuration, and consequently phases instead of signs, we had to break the artificial centre of symmetry. This was achieved by subtracting the highest peak 'A' from the Fourier synthesis. After 30 cycles of phase correction all 40 atoms appeared and showed the structure in Fig. 4. The first spurious peak appeared after the 29th atomic peak. As can be seen from Fig. 4 subtracted peak 'A' of Fig. 3 was actually an atom, but phase correction proceeded quite normally.

Re-examining the earlier Fourier syntheses calculated with 200 actual phases, a partial structure of the molecular arrangement could be found (Fig. 5). The third and sixth largest peaks were spurious ones, indicating large phase errors. Starting with this structural information application of phase correction showed 39 atoms (Fig. 6) with four interspersed spurious peaks.

Both calculations show that application of phase correction leads to the correct structure starting from the centrosymmetric as well as from the non-centrosymmetric partial structure information. The use of the tangent formula for determining the initial 200

phases raised the largest peaks which were already very large compared with the smaller ones. Further application of the tangent formula to extend the phase set would not give new information, but enhance the 'squaring' effect.

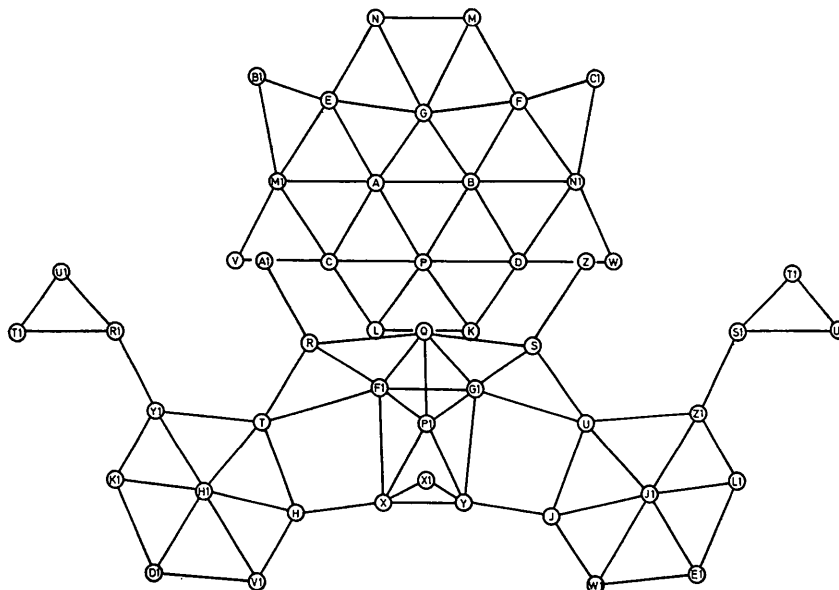


Fig. 3. Partial structure of azuleno phenalene obtained with multisolution approach in centric super space group $Pnam$, using 200 signs. Actual space group is acentric, $Pna2_1$. The structure and its enantiomorph picture are superposed and difficult to unravel, although peaks occur in the vicinity of 32 atoms. For application of phase correction the highest peak A was eliminated to break the pseudo-symmetry.

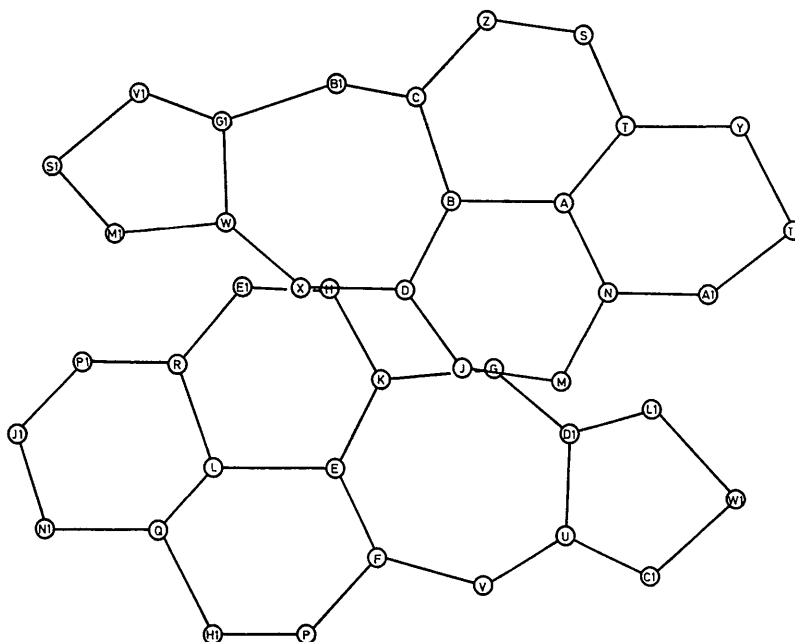


Fig. 4. The structure of azuleno phenalene after 30 cycles of phase correction starting with the centric partial structure of Fig. 3. Only five spurious peaks ($F1$, $K1$, $Q1$, $R1$, $U1$) appeared between all 40 atomic peaks.

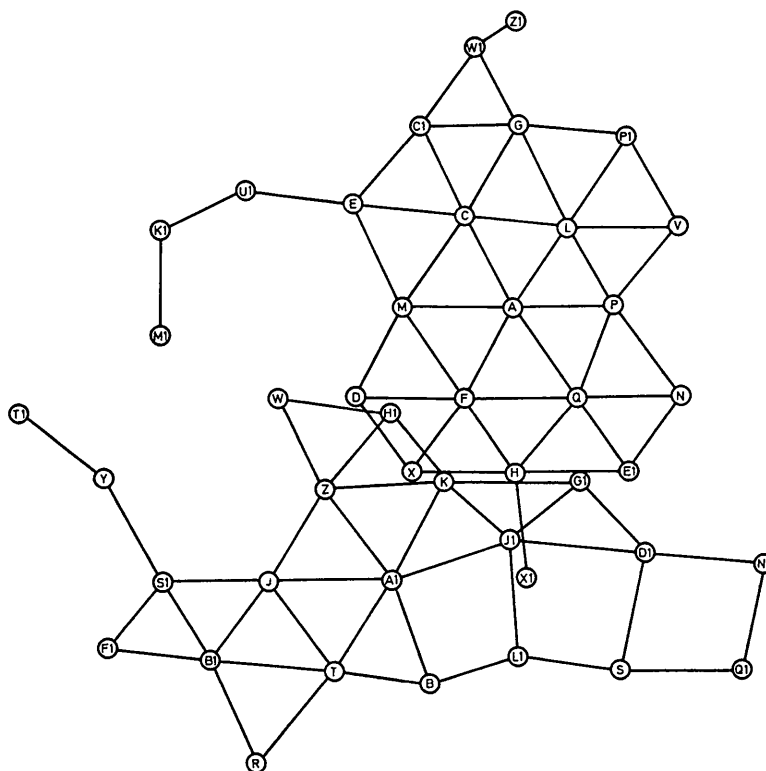


Fig. 5. Partial structure of azuleno phenalene obtained with multisolution approach in acentric space group $Pna2_1$, using 200 phases determined with the tangent formula with $E \geq 1.60$.

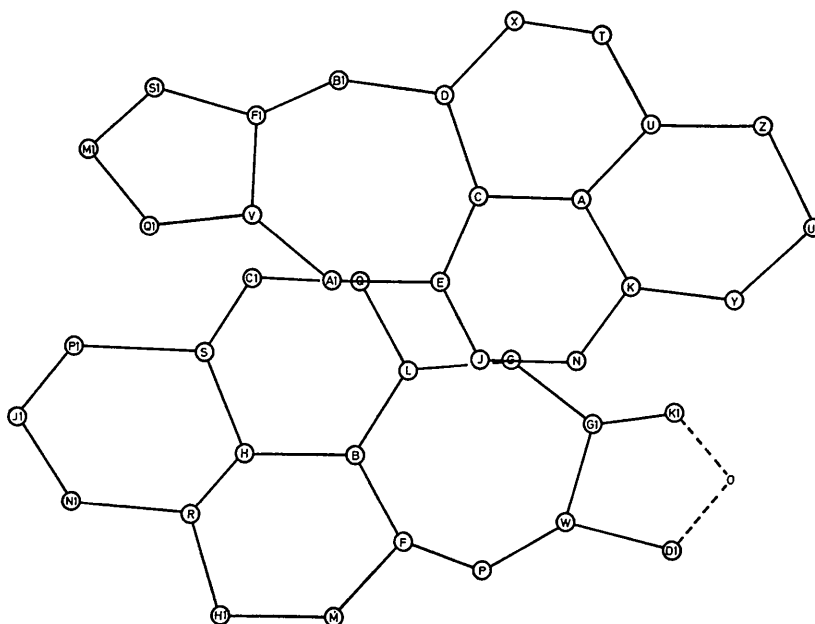


Fig. 6. Structure of azuleno phenalene after 40 cycles of phase correction starting with the acentric partial structure of Fig. 5. Only four ($E1$, $L1$, $R1$, $T1$) spurious peaks appeared between 39 atoms. The expansion of phases included all E values to 1.0.

Diphenyl-Δ²-pyrazoline (DPP), C₁₅N₂H₁₄

This structure with centrosymmetric space group $P2_1/c$ has been solved by Duffin (1968). The symbolic addition procedure led to an incorrect Fourier synthesis with partially shifted images of the structure (Bürgi & Dunitz, 1971). We re-examined the structure determination by direct methods to show the following:

- The wrong image found initially results from a too-restricted starting sign set.
- The whole subset of wrong signs depends on a single sign (which is incorrectly indicated by a Σ_1 relation).
- Selection of the most probable symbols in the symbolic addition procedure leads to a wrong solution. The selection criteria in the multisolution approach also indicate this wrong solution as the best one. The correct solution, which is lost in the symbolic addition procedure, gives the second best criterion.
- It is impossible to correct phase errors of the initial phase set (*i.e.* the initial structure) by expanding phase determination to reflexions with smaller E values.

To obtain a true comparison with the original structure determination we used the same starting conditions*. In the list of reflexions with descending E values one recognizes the original sign set as reflexions number 1,2,6,11 (Table 3). Comparing with the correct solution one recognizes that the sign given to the symbol was chosen correctly.

Nevertheless, sign expansion to reflexions satisfying $E \geq 2.0$ produced a Fourier synthesis phased with 69 signs, which we would claim as not being interpretable. Besides many spurious peaks, the 2nd, 6th, 8th, 12th, 21st and 22nd highest peak occurred in the vicinity of atomic locations. After expanding signs to E values greater than 1.5 we obtained a similar Fourier synthesis, phased with 156 signs, with only minor changes relative to the Fourier synthesis for $E \geq 2.0$. The highest peaks again were at the same locations indicating the squaring effect of the tangent formula which raises the highest peaks to the largest extent.

Relative to the $E \geq 2.0$ sign expansion none of the 69 signs with highest E values changed in extending signs to $E \geq 1.5$. Since a structure with 17 atoms in a centrosymmetric space group ought to be determined in its configuration by the 69 largest reflexions, it is quite obvious that expansion to $E \geq 1.5$ cannot bring any configurational changes.

Such a wrong sign determination may originate at the very beginning of the process. In the present case it is due to the Σ_1 relation (Hauptmann & Karle, 1953):

$$s_{13}(\bar{2}08) = s_6(\bar{1}64) \cdot s_6(\bar{1}\bar{6}4) = \text{positive}$$

This relation has a probability of $P_+ = 0.97$.

* We used the observed structure factors published by Duffin (1968). Scaling deviations due to film data originally used for direct method approach were not taken into account. Our origin corresponds to the calculated structure factors given in the paper by Duffin.

Even if Σ_1 relations are excluded from the sign determination, the interrelation between the remaining reflexions leads to the same wrong result of determining the sign of $\bar{2}08$ as positive by the following sequence: (lower indices refer to position of E value)

$$\begin{aligned} s_{31}(\bar{3}25) &= s_{24}(\bar{1}46) \cdot s_{29}(\bar{2}21) \\ s_{50}(\bar{4}49) &= s_6(\bar{1}64) \cdot s_{31}(\bar{3}25) \\ s_5(\bar{3},6,12) &= s_1(123) \cdot s_{50}(\bar{4}49) + s_{31}(\bar{3}25) \cdot s_{46}(087) \\ s_{13}(\bar{2}08) &= s_5(\bar{3},6,12) \cdot s_6(\bar{1}\bar{6}4). \end{aligned}$$

The first indication is wrong, leading to wrong signs for reflexions with numbers 31, 50, 5 and, ultimately, 13.

There appear to be two ways out of this difficulty. One way is to use a different starting set, so that the wrong sign indication occurs later and can be cancelled out by other competing terms in the tangent formula. The new determined sign then has a lower consistency, but is determined correctly and used correctly in the further process. This correcting influence on a sign determination through the use of a different term sequence is totally different from the correction of a once-determined sign by reflexions with smaller E values which themselves may be dependent on this sign. Since the crucial wrong sign indication in most cases is unknown, this is a method of trial and error, which does not necessarily lead in all cases to the correct solution.

The second way is to enlarge the basic sign (phase) set. This basic phase set must contain as a preliminary structure all features which distinguish it from any wrong solution. A correct expansion of this phase set or as one may call it the 'modification of the preliminary structure' must be possible. Whereas with a small starting set, phases with low probability must be used to continue phase determination, indications with higher probability occur with larger starting sets. Such a change in sequence may be recognized manually or with computer programs (*e.g.* program *CONVERGE*, Germain, Main & Woolfson, 1970).

In the present case of DPP such an analysis shows the following sequence of sign determination using reflexions with numbers 1, 2, 6, 11 (Table 3):

$$\begin{aligned} s_4(\bar{3},7,10) &= s_1(\bar{1}2\bar{3}) \cdot s_2(\bar{2},5,13) \\ s_7(\bar{2},3,15) &= s_1(\bar{1}2\bar{3}) \cdot s_{11}(\bar{1},1,18) \\ s_3(268) &= s_4(\bar{3},7,10) \cdot s_{11}(\bar{1},1,18). \end{aligned}$$

This leaves reflexion number 5 ($\bar{3},6,12$) with an undetermined phase. Adding this reflexion to enlarge the initial sign set (see Table 3) and using the multisolution approach we obtained four solutions. Sign expansion to $E \geq 2.0$ gave 69 phases with the following consistency values Q ($Q=0$ means that for every reflexion all double product contributions $E_k E_{h-k}$ to the tangent formula are consistent and indicate the same phase value, whereas $Q=1$ means a random distribution of all contributions to the tangent formula):

$$Q = 1 - \left\{ \sum_h |E_h| \cdot \left| \sum_k E_k E_{h-k} / \sum_k |E_k E_{h-k}| \right| \right\} / \sum_h |E_h|.$$

Q values for different expansions are:

	$E \geq 2.0$ (69 signs)	$E \geq 1.5$ (156 signs)
1. Wrong solution	0.028	0.162
2. Wrong solution (symbolic addition)	0.0005	0.095
3. Wrong solution	0.058	0.207
Correct solution	0.035	0.132

This comparison shows that the wrong solution selected by the symbolic addition procedure gives the best consistency of all solutions. The consistency value of the correct solution has the second best value for the expansion to $E \geq 1.5$. Since in the multisolution approach it is accepted that the correct solution is not always the most consistent one, the situation for DPP is quite normal. Fourier syntheses will be calculated until the correct structure is encountered. Rapid interpretation of these Fourier syntheses with available computer programs is readily achievable. The Fourier syntheses phased correctly showed the total structure with no interspersed spurious peaks either with 69 reflexions ($E \geq 2.0$) or with 156 reflexions ($E \geq 1.5$). The remaining two wrong solutions when expanded to $E \geq 1.5$ gave uninterpretable Fourier syntheses similar to those with 69 reflexions ($E \geq 2.0$). The signs of the larger E values remained unchanged when expanding with the tangent formula.

Table 3. Comparison of different initial sign sets for DPP

Although the restricted sign set with only one symbol is correct, sign determination leads to a wrong solution. The same sign set enlarged by one reflexion using all sign possibilities in a multisolution approach develops the correct structure,

Number	h	k	l	E	Original sign set	Wrong solution	Enlarged sign set	Correct solution
1	1	2	3	3.91	+	+	+	+
2	2	5	13	3.79	-	-	-	-
3	2	6	8	3.69		-		-
4	3	7	10	3.67		+		+
5	3	6	12	3.55		+	+/-	+
6	1	6	4	3.43	+	+	+	+
7	2	3	15	3.40		+		+
8	1	1	3	3.26		+		-
9	2	4	6	3.21		+		+
10	5	0	4	3.12		-		+
11	1	1	18	3.00	-	-	+/-	+
12	4	2	7	2.97		-		+
13	2	0	8	2.81		+		-
14	1	8	5	2.76		+		+
15	0	10	3	2.74		-		-

Conclusions

The symbolic addition procedure and the multisolution approach both work in centrosymmetric and non-centrosymmetric space groups. The symbolic addition procedure searches for single relations between phases (triple products) and indicates the most consistent choice for these symbols. This most consistent symbol set may not be the correct one since other

solutions are deleted when substituting actual phases for symbols. Thereby the correct solution may be lost. This will especially occur in the following cases.

- A very strong phase relation is wrong and through this relation a whole subset of phases is determined incorrectly.
- The initial number of symbols is too small and weak phase relations with corresponding small probabilities must be used to determine new phases.
- The phase values replacing the symbols in acentric space groups contain too large deviations from the correct phases. The whole phase set at that point is relatively inconsistent.

Consequences of this may be uninterpretable Fourier syntheses or Fourier syntheses displaying partial structures.

In the multisolution approach actual phase values are used from the beginning. Variation of the phase values produces a scan through all possible solutions. Since all contributing terms are used from the beginning with their corresponding probability to determine a new phase, the internal consistency will be better. Several weak indications with low probability may combine to give a new phase more accurately. The points (b) and (c) mentioned for symbolic addition can be disposed of more easily in a multisolution approach, whereas similar consequences arise if a strong phase relation fails. There is however the possibility that a sum of weaker indications supercedes the strong indication which is wrong, thereby determining the phase value correctly.

Even though several solutions have to be calculated using multisolution methods compared to the single solution in symbolic addition, computer time requirements are only slightly higher and manual intervention and interpretation is less. After accepting actual phase values in symbolic addition the new determination of additional phases uses the same principles as the multisolution approach. In most structure determinations this phase expansion occurs by use of the tangent formula. The phases calculated with this formula are the phases of the squared preliminary structure represented by the limited number of large E values associated with the corresponding known phases. Squaring tends to increase the range of peak heights. Small peaks get lost in the background and structural information may be lost, leaving only a partial structure or no structure at all in the Fourier synthesis. The squaring and consequent loss of structure information is more likely to occur in non-centrosymmetric space groups. In these space groups the continuous change of phase values is possible, approaching the phases of the squared preliminary structure. In centric space groups a change of sign is less likely to occur and the structure tends back to the unsquared preliminary structure.

Structures using only a limited number of phases associated with E values greater than 1.8 can normally be identified. Phase expansion with the tangent formula

to E values of about 1.4 contributes only limited configurational information. Expansion to E values of 1.0 adds no new information and, due to the squaring effect, may even destroy valuable phase information. The use of the different phase expansion procedure, 'phase correction', eliminates the squaring tendency and allows phase expansion to E values of 1.0. Compared with the tangent formula the phases are more accurate resulting in a narrow range of atomic peak heights, few or no interspersed spurious peaks, elimination of wrong peaks, enhancement of unknown atoms and reduction of background. Phase sets treated with the tangent formula which do not converge to an interpretable or partially recognizable structure may be improved by phase correction to show the complete structure.

We are grateful to Professor W. Hoppe for valuable discussions and his continuous support of this work. This research has been supported by the Deutsche Forschungsgemeinschaft, the Fonds der Chemischen Industrie and the Badische Anilin- & Soda-Fabrik. We should like to express our sincere thanks.

References

BRANDL, F., GRAZIANI, A. & ZECHMEISTER, K. (1972). In preparation.

- BÜRGI, H. B. & DUNITZ, J. D. (1971). *Acta Cryst.* **A27**, 117.
 BURZLAFF, H. (1970). *Z. Kristallogr.* **132**, 423.
 COCHRAN, W. (1955). *Acta Cryst.* **8**, 473.
 COOPER, A., NORTON, D. A. & HAUPTMANN, H. (1969). *Acta Cryst.* **B25**, 814.
 DUFFIN, B. (1968). *Acta Cryst.* **B24**, 1256.
 GERMAIN, G. & WOOLFSON, M. M. (1968). *Acta Cryst.* **B24**, 91.
 GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1970). *Acta Cryst.* **B26**, 274.
 HAUPTMANN, H., FISHER, J., HANCOCK, H. & NORTON, D. A. (1969). *Acta Cryst.* **B25**, 811.
 HAUPTMANN, H. & KARLE, J. (1953). *Acta Cryst.* **6**, 136.
 HOPPE, W. (1962). *Acta Cryst.* **15**, 13.
 HOPPE, W. & GASSMANN, J. (1964). *Ber. Bunsenges. phys. Chem.* **68**, 808.
 HOPPE, W. & GASSMANN, J. (1968). *Acta Cryst.* **B24**, 97.
 HOPPE, W., GASSMANN, J. & ZECHMEISTER, K. (1970). In *Crystallographic Computing*, p. 26, edited by F. R. AHMED. Copenhagen: Munksgaard.
 JUTZ, CH. & KIRCHLECHNER, R. (1966). *Angew. Chem.* **78**, 493.
 KARLE, I. L., KARLE, J. & ESTLIN, J. A. (1967). *Acta Cryst.* **23**, 494.
 KARLE, J. (1968). *Acta Cryst.* **B24**, 182.
 KARLE, J. & HAUPTMANN, H. (1956). *Acta Cryst.* **9**, 635.
 KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849.
 KARTHA, G. (1969). *Acta Cryst.* **A25**, S87.
 SAYRE, D. (1952). *Acta Cryst.* **5**, 60.

Acta Cryst. (1972). **A28**, 280

Classification of Magnetic Structures in Some Orthorhombic Space Groups

BY E. GUREWITZ AND H. SHAKED

Physics Department, Nuclear Research Centre-Negev, P.O.B. 9001, Beer Sheva, Israel

(Received 19 April 1971)

The magnetic structures belonging to magnetic space groups isomorphic with D_{2n}^i , $i=1,2, \dots, 16$, are classified for the general and special positions of these space groups.

Introduction

Given a crystal with ionic magnetic moments S_1, S_2, \dots, S_n on a set of n -fold positions, the $3n$ dimensional space

$$S = S_1 \times S_2 \times \dots \times S_n \quad (1)$$

has subspaces invariant under the paramagnetic symmetry group of the crystal. The bases of all these subspaces, known as modes* (Bertaut, 1968), may serve to classify the possible magnetic structures in the crystal. A classification of this sort for special positions in D_{2n}^i is given in the literature (Bertaut, 1968). We pre-

sent an extension of the classification to 2-, 4- and 8-fold positions in D_{2n}^i , $i=1,2, \dots, 16$, as a part of an idea to extend the classification to all space groups. Tables of limiting conditions for allowed reflexions in these space groups are given elsewhere (Gurewitz & Shaked, 1971). An example of the use of the classification scheme and the tables of allowed reflexions in the analysis of neutron diffraction from a polycrystalline sample of $KFeCl_3$, is given in the present paper.

Classification of the magnetic structures in crystals belonging to D_{2n}^i , $i=1,2, \dots, 16$

The point group of the space groups D_{2n}^i is mmm . This is a commutative group of order 8. It has: (a) seven

* In analogy with normal modes of the theory of vibrations.