Acta Cryst. (1968). **B24**, 182

Partial Structural Information Combined with the Tangent Formula for Noncentrosymmetric Crystals

By J. Karle
U.S. Naval Research Laboratory, Washington, D.C. 20390, *U.S.A.*

(*Received* 13 *March* 1967)

A method is presented for extending the knowledge of a partial structure to a complete structure determination by means of the tangent formula. Phases are computed from the partial structure and a criterion is presented for accepting certain of these to form a basic set for expansion by the tangent formula. Tests have been carried out on two initially unknown structures, one with almost equal atoms, and one with some moderately heavy atoms. It was found for the structure containing almost equal atoms that a knowledge of the positions of the two strongest peaks, as they appeared in an initial E map obtained from the symbolic addition procedure, could be readily developed into a complete determination of the seventeen atom structure. After location of the moderately heavy atoms in the second structure by use of a Patterson map, the remaining 21 of 23 atoms were quite readily located by use of the tangent formula.

Introduction

It is a common experience in structure investigations of equal atom noncentrosymmetric crystals to observe that Fourier series calculations of the electron density are often inadequate for improving limited structural information. This difficulty manifests itself in the failure to add the proper additional atoms to a partially known structure or to remove badly placed atoms from an approximate structure. Recent studies have indicated however that the tangent formula has the capacity to overcome these problems, leading to improved procedures for structure determination.

There are many types of partial structural information. In addition to often having knowledge of chemical composition, there is the vast store of accumulating background information concerning bond lengths, bond angles, atomic groupings and molecular configurations and packings. Partial information may also consist of the knowledge of the approximate coordinates of a portion of the atoms in the unit cell. It is this latter type of information to which we wish to apply the tangent formula directly, recognizing that in the course of the application, the investigator may also explicitly employ some additional background information.

The capacity of the tangent formula to yield the complete structure from a partial structure has been investigated in the course of several structure investigations. In equal atom or almost equal atom structures, it is of interest to obtain an estimate of how small a fraction of the structure can be used for starting purposes. For structures containing some somewhat heavier atoms, Fourier methods are more amenable, once the heavier atoms have been located. In this case it is of interest to determine whether there are advantages in using the tangent formula to help complete the structure determination.

The results of the experimental studies to be described have implications with respect to the procedural details for structure determination. For example, instead of initially attempting to locate the approximate positions of all the atoms in an equal atom or almost equal atom structure, it is only necessary to locate a portion of them. This is generally easier to do and it is expected that structure determinations will be thereby facilitated. This matter will be discussed in some detail.

A method for overcoming the limitations of least-squares and Fourier methods for improving approximate or partially known structures has been suggested previously by Hoppe (1963). He termed it a method of phase refinement and it is based upon improving approximate phases by means of a least-squares refinement associated with the improvement of the fit to a phase determining relation. A procedure for phase correction using distorted electron density functions has also been suggested by Hoppe, Huber & Gassmann (1963). A further progress report has been given by Gassmann (1966). The use of the tangent formula in connection with partial structural information in isomorphous replacement has also been discussed (Coulter, 1965; Karle, 1966).

Method

The tangent formula is (Karle & Hauptman, 1956)

$$\tan \varphi_{\mathbf{h}} \simeq \frac{\Sigma\,|E_{\mathbf{k}}E_{\mathbf{h-k}}|\sin(\varphi_{\mathbf{k}}+\varphi_{\mathbf{h-k}})}{\Sigma\,|E_{\mathbf{k}}E_{\mathbf{h-k}}|\cos(\varphi_{\mathbf{k}}+\varphi_{\mathbf{h-k}})}\,, \tag{1}$$

where $\varphi_{\mathbf{h}}$ is the phase associated with the normalized structure factor $E_{\mathbf{h}}$. It is apparent that in order to apply (1) some initial phases need to be known. The sum is carried out over the known pairs $(\varphi_{\mathbf{k}},\varphi_{\mathbf{h-k}})$ and it has been observed that good accuracy is obtained from relatively few contributors to the sum so long as the $|E_{\mathbf{h}}|$, $|E_{\mathbf{k}}|$ and $|E_{\mathbf{h-k}}|$ are among the largest magnitudes (Karle & Karle, 1966). The question of interest here is whether a sufficiently accurate set of preliminary phases among the largest $|E|$ values could be obtained from the knowledge of the approximate positions of a portion of a structure in order ultimately to complete the structure determination by means of (1).

The procedure to be studied involves the computation of structure factors from the partial structure and to accept those phases for a basic set for which there is a correspondence between the largest computed and observed structure factor magnitudes. The basic set of phases so obtained is then used with (1) to improve this initial set and to generate additional phases among the largest observed normalized structure factor magnitudes. A Fourier map of the electron density is then made from the newest set of phases with the expectation that additional atoms can be added to the initial approximate structure. Conceivably too, improved positions may also be found for the initial set of atoms. Using the improved partial structure thus obtained, the process is to be repeated until a complete structure is obtained. A final refinement would still employ the standard least-squares technique.

Several questions arise concerning the implementation of this procedure:

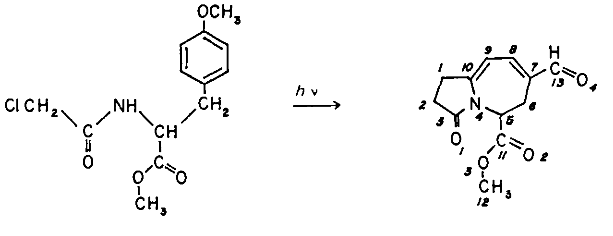
1. How large a fraction of a structure is required to be known initially and how accurately in order to carry out the above procedure?
2. What are the criteria for accepting phases calculated from partial structures?
3. How many additional phases are to be computed from the tangent formula using a starting set of phases calculated from a partial structure?

Although experimental experience with the generation of successive partial structures using the tangent formula has been rather limited to date, it is possible, on the basis of these studies, to offer some tentative answers to the above questions.

Experimental tests

A structure with almost equal atoms

The structure determination of a rearrangement product obtained from the ultraviolet irradiation of a derivative of tyrosine (Yonemitsu, Witkop & Karle, 1967)



crystallizing in space group $P2_12_12_1$ has afforded a good test of the use of partial structures in combination with the tangent formula. The sequence of steps is outlined in Table 1. A detailed description of the structure has been given elsewhere (Karle, Karle & Estlin, 1967).

The structure determination was initiated by using the symbolic addition procedure (Karle & Karle, 1966). As noted in Table 1, this gave 63 initial independent phases which were expanded to 238 by use of the tangent formula. From the resulting E map, five atoms were chosen from among the six largest peaks ranging in value from 124 to 392. The acceptance of the five peaks in this case was based on the criteria that the peaks made a sensible chemical configuration and that they were spaced by bonded distances. The sixth peak was spaced by more than a bonded distance and, in fact, turned out to be spurious. Spurious peaks of significant magnitude did not appear in subsequent cycles.

It should be noted that although the configuration involving the five chosen atoms was well represented in the initial E map obtained from the symbolic addi-

Table 1. *Details of the successive cycles for developing the complete structure of the photocyclization product from the initial partial structure obtained by use of the symbolic addition procedure*

Cycle	Basis	Initial phases accepted	Minimum $\lvert E\rvert$	Phases developed by tangent formula	Minimum $\lvert E\rvert$	R
I	sym. add.	63	1·6	238	1·2	
II	5 atoms	58	1·4	233	1·2	64·4
III	10 atoms	81	1·4	233	1·2	52·3
IV	14 atoms	128	1·4	229	1·2	42·3
V	16 atoms	140	1·4	227	1·2	36·9
	17 atoms					28·1

Table 2. *Successive values of the phases for a typical set of reflections as developed in the series of cycles listed in Table* 1

$h\ k\ l$	$\lvert E\rvert$	I	II	III	IV	V	Final value
8 5 2	2·59	3·14	2·96	2·86	3·03	3·01	−3·03
5 4 14	2·23	0·78	1·00	0·79	0·80	0·60	0·50
2 1 1	2·11	0	−0·23	−0·47	−0·74	−1·01	−0·98
4 6 9	1·94	1·57	0·47	0·24	0·61	0·58	0·68
5 2 6	1·82	−2·36	−2·84	3·02	2·98	2·90	−3·10
8 6 2	1·71	3·14	1·60	1·48	1·44	1·29	1·48

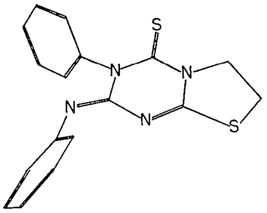
tion procedure, most of the remainder of the structure was not. This was due to fairly large errors in some of the initial phases. Since this may often occur in structures having equal or almost equal atoms, the development of the present procedure utilizing a partial structure initially is a matter of practical importance.

Throughout the procedure until all 17 atoms were located, the atoms were all assumed to be carbon atoms. Structure factors were computed at first from the five initially chosen atoms employing an isotropic temperature constant, $B=3{\cdot}5$. Phases so obtained were accepted if the criterion $|F_{\mathbf{h}}|_{\text{calc}} \geq 0{\cdot}25|F_{\mathbf{h}}|_{\text{obs}}$ was satisfied and if $|F_{\mathbf{h}}|_{\text{obs}}$ had a corresponding normalized structure factor $|E_{\mathbf{h}}|_{\text{obs}} \geq 1{\cdot}4$. In this way 58 phases were obtained, as shown in Table 1, and with the use of the tangent formula, they were expanded to 233 phases. From an E map computed with these phases, five additional atomic positions were chosen from among the strong peaks. The criteria for accepting these additional atomic positions were again that all atoms should make a sensible chemical configuration and be connected together by bonded distances. The bonded distance criterion was adhered to throughout the procedure, since the configuration of the molecule was unknown. As expected, other correct atoms at nonbonded distances also appeared in the E maps at intermediate stages. They were ignored, however, because it was thought that they might be spurious. Apparently this led to no difficulty and, as seen in Tables 1 and 2, the structure was obtained in five steps. Had the large peaks at nonbonded distances been accepted, the structure would have been obtained in no more than three steps. As more atoms were added, the criterion for accepting a phase from the computation of structure factors was changed, *e.g.* for a partial structure composed of 10 atoms, the criterion used was, $|F_{\mathbf{h}}|_{\text{calc}} \geq 0{\cdot}33\, |F_{\mathbf{h}}|_{\text{obs}}$ and for 14 atoms and more, $|F_{\mathbf{h}}|_{\text{calc}} \geq 0{\cdot}5|F_{\mathbf{h}}|_{\text{obs}}$. Even in the map showing 16 atoms, the weights were rather disparate, ranging from 108 to 396. However, all sixteen peaks were larger than any spurious one.

The R values for successive cycles based on 990 independent data are shown in Table 1. No attempt was made to optimize the agreement. The $|F|_{\text{obs}}$ were used as they were obtained from the data reduction program and had a scaling error of about 10%, which was subsequently corrected in the least-squares refinement. Also, in the computation of $|F|_{\text{calc}}$ all atoms were treated as if they were carbon atoms and an isotropic temperature constant of $B=3{\cdot}5$ was employed. Nevertheless the trend of improvement in succeeding cycles is apparent. The differences between the values of the atomic coordinates obtained in successive cycles and those obtained from the final least-squares refinement were found rarely to exceed 0·01 of a fractional coordinate and often much less. Therefore the continued improvement is for the most part due to the addition of new atomic positions rather than the improvement of earlier ones. The variation of the phases for a representative sample of reflections is seen in Table 2. In the course of the procedure some phases changed only slightly, some, *e.g.* for $\mathbf{h}=8,6,2$, changed as much as 1·66 radians, and a few others changed even more.

A structure with moderately heavy atoms

A sulfur-containing compound having the structural formula



was found to crystallize in space group $P2_12_12_1$. A detailed description of the structure will be given elsewhere (Karle, Flippen & Karle, 1967). The sulfur atoms could be readily located from a Patterson map, and it was of interest to use the partial structure consisting of these sulfur atoms in combination with the tangent formula to see how readily the complete structure might be obtained. This study also afforded an opportunity to make a comparison with the symbolic addition procedure, the heavy atom method and the modified heavy atom method of Sim (1960).

Table 3 shows the initial values for the positions of the sulfur atoms as obtained from the Patterson map and their final values obtained from a least-squares refinement. Structure factors were computed from these initial positions employing an isotropic temperature constant, $B=3{\cdot}5$. The criterion for accepting a phase was $|F_{\mathbf{h}}|_{\text{calc}} \geq 0{\cdot}25|F_{\mathbf{h}}|_{\text{obs}}$, where $|F_{\mathbf{h}}|_{\text{obs}}$ is associated with an $|E_{\mathbf{h}}|_{\text{obs}} \geq 1{\cdot}43$. In this way 90 independent phases were obtained which were expanded to 547 by means of the tangent formula. The minimum $|E|$ value associated with a phase in the expanded set was 1·0. A Fourier map was computed from these phases combined with the corresponding $|E_{\mathbf{h}}|$ as coefficients, an E map. It was possible to locate all 23 atoms heavier than hydrogen atoms in the E map whose positions could then be refined by means of the least-squares method. The range of heights for the light atoms was 126–289, except for two weak ones which were 82 and 90. There were three spurious peaks of heights 132, 151 and 161 which did not confuse the interpretation since their positions did not make chemical sense. It was found that if the number of phases was extended to 656, where the minimum value of E was 0·90, the

Table 3. *Initial and final positions of the sulfur atoms*

		Patterson	Least-squares
S(1)	*x*	0·512	0·510
	y	0·329	0·322
	z	0·750	0·762
S(2)	*x*	0·312	0·312
	y	0·657	0·655
	z	0·000	0·989

weak peaks and the spurious peaks were not materially affected in value.

For comparison a structure determination was carried out by means of the symbolic addition procedure employing $|E| \geq 1{\cdot}6$, initially. It was necessary to consider a few smaller $|E|$ values in order to specify and establish the enantiomorph. The structure was readily obtained. The classical heavy atom method was also employed. This is distinguished from the first procedure described in that a Fourier map is computed directly from phases associated with the partial structure of heavy atoms, omitting the intermediate step of expanding a selected set of phases by means of the tangent formula. Seven successive F maps were computed to complete the structure.

A modification of the heavy atom method for centrosymmetric crystals has been described by Woolfson (1956) and for noncentrosymmetric crystals by Sim (1960). In these modifications the Fourier coefficients are given weights obtained from statistical considerations which are a measure of the reliability with which the phases are obtained from the partial structure. Such a procedure was also employed and although the process was not continued to a final answer, it was apparent that fewer successive F maps would have been required than for the standard heavy atom method.

Discussion

The experimental tests described in the previous section provide partial answers to the questions raised near the end of the section on *Method.* In the examples given, the tangent formula (1) readily developed phases ultimately to complete the structure when coupled with structural fragments comprising about 30% of the scattering power in the photocyclization product and 40% in the compound containing sulfur.

To test further how small the partial structure might be, the method was tested with a two-atom fragment of the photocyclization product. The peak positions were those corresponding to the two largest peaks appearing in the initial E map obtained from applying the symbolic addition procedure. This two-atom fragment, constituting less than 15% of the scattering power, generated a map from the phases, developed by the tangent formula, in which the seven largest peaks all corresponded to proper atomic locations. In this case, knowledge of the positions of two atoms could be routinely developed into a complete structure determination for a molecule containing seventeen atoms. It was not possible to proceed with the structure analysis when only the single atomic position corresponding to the largest peak appearing in the initial E map was used as the initial structural fragment. When an F map was computed from phases based on the two-atom partial structure and accepted if $|F|_{\text{calc}} \geq 5$ and $|F|_{\text{obs}} \geq 10$, instead of making use of the tangent formula, the resulting map showed only the two initial atoms and two others which had peak values similar to many of the spurious peaks. The two additional atoms would not be readily identifiable in an unknown structure. Apparently, the tangent formula plays an essential role.

The criterion which is suggested for accepting phases computed from a partial structure and which has been approximately applied in the foregoing tests employing the tangent formula is $|F_{\mathbf{h}}|_{\text{calc}} \geq p|F_{\mathbf{h}}|_{\text{obs}}$, where p is the fraction of the total scattering power contained in the fragment and where $|F_{\mathbf{h}}|_{\text{obs}}$ is associated with an $|E_{\mathbf{h}}| \geq 1{\cdot}5$. An additional approximate rule is if $p < 0{\cdot}25$ it is replaced by 0·25 and if $p > 0{\cdot}6$, it is replaced by 0·6. Alternative and additional criteria may be readily developed and may turn out to be preferable. For example, in the first application of the tangent formula employing only phases obtained from the partial structure, it may be advantageous to replace $|E_{\mathbf{k}}|$ and $|E_{\mathbf{h-k}}|$ by $W_{\mathbf{k}}|E_{\mathbf{k}}|$ and $W_{\mathbf{h-k}}|E_{\mathbf{h-k}}|$, respectively, where $W_{\mathbf{k}}$ is the appropriate W defined by Sim (1960). In subsequent steps, when new phases determined solely by the tangent formula are introduced, this type of weighting would become meaningless and would be omitted. Weighting could also be based on the variance as given in Fig. 2 of a previous paper (Karle & Karle, 1966). It is not clear at this point however that a significant practical advantage would accrue from complicating the tangent formula with weighting functions.

The extent to which the initial phases computed from a partial structure are to be expanded by means of the tangent formula may be set in terms of the lowest $|E|$ value considered. As seen in Table 1, the minimum $|E|$ value used for the photocyclization product was 1·2. This is apparently a suitable value.

All four methods for determining the structure of the sulfur compound in the second test example, the use of the partial structure of sulfur atoms combined with the tangent formula, the symbolic addition procedure, the standard heavy atom method and the modified heavy atom method all led to a determination of the structure. It was found that the first method employing the partial structure was very simple to apply and seems to be preferable.

It appears from the limited tests carried out so far that once a partial structure constituting a relatively small fraction of the scattering power is properly located in the unit cell, the remaining structure can be readily obtained from application of the tangent formula. The symbolic addition procedure has afforded the needed partial structural information in several essentially equal atom structures studied recently. Besides the example in this paper, this includes reserpine, digitoxiginen and hydroxycrinamine. It should be noted that rather sophisticated techniques are being developed for locating and orienting known partial structural configurations by means of the Patterson function (Hoppe, 1957; Hoppe & Will, 1960; Huber & Hoppe, 1965; Nordman & Nakatsu, 1963; Nordman, 1966). This constitutes another way of locating a structural fragment for use with the tangent formula.

It may be worthwhile to apply the procedure described in this paper to heavy atom derivatives of complex materials such as proteins. The relative scattering power of the heavy atoms and the collection of as many data as possible are important considerations.

Dr Isabella Karle and Mrs Judith Flippen have collaborated in carrying out the experimental tests. Their fine cooperation is very much appreciated. We are grateful to Mr. Stephen Brenner for programming the calculation of the recycling procedure.

References

COULTER, C. L. (1965). *J. Mol. Biol.* **12**, 292.
GASSMANN, J. (1966). *Acta Cryst.* **21**, A6.
HOPPE, W. (1957). *Z. Elektrochem.* **61**, 1076.
HOPPE, W. (1963). *Z. Kristallogr.* **118**, 121.
HOPPE, W., HUBER, R. & GASSMANN, J. (1963). *Acta Cryst.* **16**, A4.
HOPPE, W. & WILL, G. (1960). *Z. Kristallogr.* **113**, 104.
HUBER, R. & HOPPE, W. (1965). *Chem. Ber.* **98**, 2403.
KARLE, I. L., KARLE, J. & ESTLIN, J. A. (1967). *Acta Cryst.* **23**, 494.
KARLE, J. (1966). *Acta Cryst.* **21**, 273.
KARLE, J., FLIPPEN, J. & KARLE, I. L. (1967). *Z. Kristallogr.* To be published.
KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635.
KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849.
NORDMAN, C. E. & NAKATSU, K. (1963). *J. Amer. Chem. Soc.* **85**, 353.
NORDMAN, C. E. (1966). *Trans. Amer. Cryst. Assn.* **2**, 29.
SIM, G. A. (1960). *Acta Cryst.* **13**, 511.
WOOLFSON, M. M. (1956). *Acta Cryst.* **9**, 804.
YONEMITSU, O., WITKOP, B. & KARLE, I. L. (1967). *J. Amer. Chem. Soc.* **89**, 1039.

Acta Cryst. (1968). B**24**, 186

Atomic Ordering in Binary *A*15-Type Phases

BY E.C. VAN REUTH*
Naval Ship Research and Development Center, Annapolis, Maryland, U.S.A.

AND R.M. WATERSTRAT
Dental Research Section, National Bureau of Standards, Washington, D.C., U.S.A.

(*Received* 29 *November* 1966 *and in revised form* 10 *June* 1967)

The degree of long-range order has been determined for 20 binary *A*15-type phases containing various transition elements. A tendency toward a lower degree of order was noted as the component elements were chosen successively from columns in the periodic table approaching the Mn column. A comparison of the ordering in the *A*15-type phases with the ordering previously reported for various binary σ phases suggests that the remarkable stability of these phases may result from an interdependence between the electronic structure and the ability of the atoms to undergo deformations in conforming to geometrical packing requirements.

The *A*15-type structure may be regarded as belonging to a group of crystal structures which have common crystallographic features and occur in many alloy phases formed by the transition elements. These structures have been described in terms of four characteristic atomic coordination polyhedra (Kasper polyhedra) possessing coordination numbers of 12, 14, 15 and 16 respectively (Kasper, 1956; Shoemaker, Shoemaker & Wilson, 1957; Frank & Kasper, 1958, 1959; Komura, Sly & Shoemaker, 1960; Shoemaker & Shoemaker, 1963). If atomic packing considerations are an important factor in stabilizing these phases, one might expect an atomic ordering to occur such that the icosahedral 12-coordinated polyhedra would contain the smaller atoms while the larger atoms enter sites with 14, 15 or 16 coordinations. Such an ordering of atoms has been observed in many of these structures [a summary is given in a recent paper by Shoemaker, Shoemaker & Mellor (1965)] but attempts to measure the extent of atomic ordering at each atomic site have often been restricted by the necessity of simultaneously determining the atomic position parameters. In the *A*15 structure type, however, the atomic position parameters are fixed by the symmetry requirements of space group *Pm*3*n* so that for the 'ideal' (A_3B) stoichiometric composition the B atoms are in positions 2(*a*) (0,0,0) and the A atoms are in positions 6(*c*) $(0,\frac{1}{4},\frac{1}{2})$. Consequently, measurements of atomic ordering in these phases can be made with greater accuracy and with less ambiguity since there are only two crystallographic lattice sites in this structure.

Geller, Matthias & Goldstein (1955) have reported a high degree of atomic ordering in Nb_3Os, Nb_3Ir, Nb_3Pt and V_3Sn, but only a partial ordering in Ta_3Sn.

---

* On leave of absence to Kamerlingh Onnes Laboratory, University of Leiden, Leiden, The Netherlands, for the 1966–67 academic year.