Identification of Systematically Aberrant Phase Relationships Arising from Structural Regularity*

By W. E. Thiessen and W. R. Busing

Chemistry Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37830, U.S.A.

In many crystals the asymmetric unit is composed largely of repeating structural fragments all having the same orientation. This situation occurs in crystals of organic compounds containing fused benzene ring systems, for example. In the direct determination of such crystal structures approximate values for the phase-angle sums of certain triplets of strong reflections whose indices sum to zero can be predicted from the Fourier transform of a single fragment in its proper orientation. We describe a method for determining this orientation by fitting the square of the transform of the fragment to the intensity data and indicate how information derived from the oriented transform may be applied to the direct solution of crystal structures. An application of the method to the determination of the crystal structure of the plant pigment methoxydalrubone ($C_{20}H_{20}O_5$; $P\bar{1}$, $Z=2$) is given.

I. Introduction

To a good approximation, direct methods of crystal structure analysis may be viewed as an exercise in the repeated application of the formula

$$\varphi(\mathbf{h}) \simeq \langle \varphi(\mathbf{k}) + \varphi(\mathbf{h} - \mathbf{k}) \rangle_{\mathbf{k}_r}, \qquad (1)$$

(Karle & Karle, 1966), where the subscript r refers to the restriction of the scheme to those Bragg reflections with the largest values of $|\mathscr{E}(\mathbf{k})|$, the quasi-normalized structure factor magnitude (Karle & Hauptman, 1959).

Since equation (1) derives unknown phase angles only from other phase angles whose values are known, structure solution usually begins from a small set of reflections whose phases may be arbitrarily assigned or given symbolic values (Karle & Karle, 1966). In the early stages of this 'bootstrap' procedure the number of vectors $\mathbf{k}_r$ over which the average is taken is sharply limited and may, in fact, be one. In this case, equation (1) reduces to

$$\varphi(\mathbf{h}) \simeq \varphi(\mathbf{k}) + \varphi(\mathbf{h} - \mathbf{k}) \qquad (2)$$

or, since

$$\varphi(\mathbf{h}) = -\varphi(-\mathbf{h})$$
$$\varphi(-\mathbf{h}) + \varphi(\mathbf{k}) + \varphi(\mathbf{h} - \mathbf{k}) \simeq 0 . \qquad (3)$$

The chances of equation (3) being a good approximation are greatest for those triples with large $|\mathscr{E}(-\mathbf{h})\mathscr{E}(\mathbf{k})\mathscr{E}(\mathbf{h}-\mathbf{k})|$, but there will always be some triples which grossly violate equation (3) or combine to violate equation (1) (Fisher, Hancock & Hauptman, 1970a). The cosine invariant cos $[\varphi(-\mathbf{h}) + \varphi(\mathbf{k}) + \varphi(\mathbf{h} - \mathbf{k})]$ ranges from $+1$, if equation (3) holds exactly, through smaller values to -1 if it does not. Several

statistical methods of obtaining estimates of this quantity from the entire set of normalized structure factor magnitudes have been suggested (Hauptman, 1970; Karle, 1970a; Hauptman, 1972).

In this paper we will show that aberrant triples (those whose phase-angle sums are near π) which are not correctly predicted by statistical methods can be identified for crystals whose molecules are dominated by repeating structural units, such as condensed systems of benzene or *trans*-fused cyclohexane rings. We will describe a method for verifying the presence of such repeating units, and finally we will make some recommendations as to how this information can be used in developing the correct set of phases from the observed intensities.

II. A review of some properties of Fourier transforms

Let us approximate a single benzene ring by an array of six point scatterers at the corners of a regular hexagon. The Fourier transform of this array consists of columns of intensity normal to the plane of the ring. If the origin is taken at the center of the hexagon, the transform will be everywhere real and can be calculated according to the formula

$$T(\mathbf{h}) = 2 \sum_{j=1}^{N/2} \cos [2\pi(\mathbf{h} \cdot \mathbf{x}_j)] , \qquad (4)$$

where $\mathbf{x}_j$ is the position vector of one of a centrosymmetrically related pair of atoms. The components of the vector $\mathbf{h}$ are not necessarily integral. A section of the transform parallel to the plane of the molecule is displayed in Fig. 1. In a triclinic crystal containing one of these hexagonal molecules per unit cell, the effect of the three-dimensional translational symmetry is to sample the transform at the reciprocal-lattice points (Lipson & Taylor, 1958). Thus the value of the transform at the end of the reciprocal-lattice vector $\mathbf{h}$,

scaled by the factor $N^{-1/2}$, is the quasi-normalized structure factor $\mathscr{E}(\mathbf{h})$.

If the molecule is translated relative to the origin by a vector $\mathbf{t}$, the phase angle is modified but the magnitude of the transform at any point $\mathbf{h}$ remains unchanged:

$$T'(\mathbf{h}) = T(\mathbf{h}) \exp(-2\pi i \mathbf{h} . \mathbf{t}) . \qquad (5)$$

Consider now a structure which replaces the original six-membered ring by two such rings, one displaced from the original position by a vector $\mathbf{t}$ and the other by a vector $-\mathbf{t}$. Since the transform of the sum is the sum of the transforms, the transform of the resulting structure will be

$$\begin{aligned} T'(\mathbf{h}) &= [T(\mathbf{h}) \exp(-2\pi i \mathbf{h} . \mathbf{t}) + T(\mathbf{h}) \exp(2\pi i \mathbf{h} . \mathbf{t})] \\ &= 2T(\mathbf{h}) \cos(2\pi \mathbf{h} . \mathbf{t}) . \qquad (6) \end{aligned}$$

That is, the transform of the two molecules is the same as that of one molecule but rescaled and multiplied by a cosine function whose value depends only on the component of $\mathbf{h}$ in the $\mathbf{t}$ direction. The transform of the naphthalene molecule shown in Fig. 2 illustrates this result because the naphthalene molecule
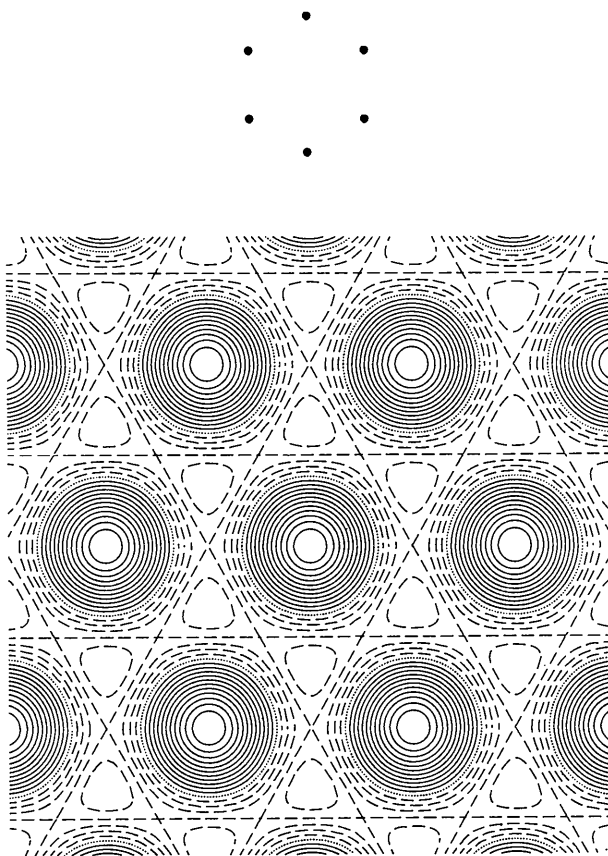

Fig. 1. A planar hexagonal array of point scatterers representing the carbon atoms of a benzene molecule and a parallel section through its Fourier transform. The contours of the transform are drawn at equal intervals with the zero level dotted and the negative levels dashed. The origin is at the central peak.
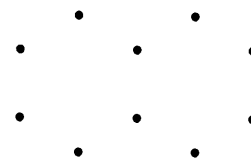
Fig. 2. A planar ten-atom array of point scatterers approximating the carbon atoms of a naphthalene molecule and a parallel section through its Fourier transform. The contours of the transform are drawn at equal intervals with the zero level dotted and the negative levels dashed. The origin is at the central peak.

approximates two benzene rings translated in equal and opposite directions. It has columns of large magnitude in the same positions as the benzene transform, but some of the signs have been changed by the cosine function. This characteristic benzenoid transform was first discussed by Knott (1940) in a pioneering paper. Studies in which rather complex benzenoid structures were solved by recognizing this pattern have been reported (Stadler, Bolton & Maitland, 1963; Mackay, 1962, and references therein), and a review article is available (Chaudhuri, 1972).

III. Phase relationships in structures with parallel repeating units

Let us now turn our attention to phase relationships among strong reflections in benzenoid systems. We have seen that the strongest Bragg reflections for our hypothetical triclinic crystal containing one hexagonal molecule per unit cell will fall on or near the centers of the high-intensity columns of Fig. 1. These columns are of two types: the principal columns have central lines of maximal intensity equal to that at the origin

and are of positive sign, while the secondary columns are negative in sign and have peak magnitudes which are half those of the primary columns. The columns are schematically represented by the plus and minus signs (to indicate phase angles of 0 and π radians respectively) in Fig. 3 along with some possible vector triples which add to zero and thus satisfy the condition imposed on the three reflections in equation (3). Note that some triples have positive sign products in accordance with equation (3); others have negative sign products corresponding to the replacement of the 0 in equation (3) by π.* Each of the arrangements shown in Fig. 3 is representative of a family of triples lying in various planes passing through the origin. Another member of the family of triples represented in Fig. 3(c) is shown in Fig. 4. Note that each member of the families represented by 3(a) and 3(b) is normal; each member of 3(c) and 3(d) is aberrant.

Now consider any two benzene rings related by an inversion center at the origin. Since the rings are themselves centrosymmetric, they are related by equal and opposite translations from the origin and equation (6) will hold. The product of three quasi-normalized structure factors whose reciprocal vectors (indices) sum to zero is related to the product of three values of the transform of the single-ring structure by

$$\mathscr{E}(-\mathbf{h})\mathscr{E}(\mathbf{k})\mathscr{E}(\mathbf{h}-\mathbf{k}) = N^{-3/2}T'(-\mathbf{h})T'(\mathbf{k})T'(\mathbf{h}-\mathbf{k})$$
$$= 8N^{-3/2}T(-\mathbf{h})T(\mathbf{k})T(\mathbf{h}-\mathbf{k}) . m \qquad (7)$$

where

$$m = \cos\,[2\pi(-\mathbf{h}\,.\,\mathbf{t})]\,\cos\,[2\pi\mathbf{k}\,.\,\mathbf{t}]\,\cos\,[2\pi(\mathbf{h}-\mathbf{k})\,.\,\mathbf{t}]\,. \qquad (8)$$

Evaluation of the product m for all possible values of the arguments reveals that $-1/8 \le m \le 1$, so that if $|m| > 1/8$ then $m > 0$. Therefore if the magnitude of the triple product for the two-ring structure is large, it is expected to have the same sign as that for a single ring.

Returning now to the transforms displayed in Figs. 1 and 2, readers may satisfy themselves that although the phases of individual columns for naphthalene have been modified by the cosine modulating function, the sign products indicated for benzene in Figs. 3 and 4 are unchanged for napthalene.

The transform of two naphthalene molecules related by vectors t and −t is shown in Fig. 5. The vector t is arbitrarily chosen except that it lies in the plane of the ring for convenience in displaying the transform. Again the characteristic pattern of the benzenoid transform appears and is modulated by the cosine function with nodal planes perpendicular to the vector t. The three vectors −h, k, and h−k which are illustrated in Fig. 5 represent a triple which is aberrant for this structure and which would also have been aberrant for the single benzene ring [Fig. 3(c)].

Our conclusion, then, is as follows. To the extent that any crystal structure is dominated by a system of similarly oriented benzene rings occupying the same or parallel planes, its phase relationships with large values of $|\mathscr{E}(-\mathbf{h})\mathscr{E}(\mathbf{k})\mathscr{E}(\mathbf{h}-\mathbf{k})|$ will be normal or aberrant depending on whether the corresponding relationships in the correctly oriented benzene transform are normal or aberrant. Further, although we have used the benzene ring as our example, the same reasoning can be applied to any repeating unit which is itself centrosymmetric, such as the chair form of cyclohexane.

IV. Determining the orientation of the repeating unit

In order to utilize the foregoing generalization in the solution of a given crystal structure, it is necessary to
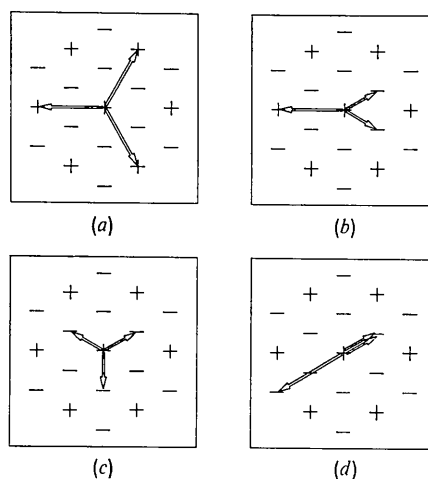

(a) (b)

(c) (d)

Fig. 3. Some phase relationships among triples of vectors which sum to zero and which end in high-intensity regions in the basal plane of the benzene transform. Relationships (a) and (b) are normal; relationships (c) and (d) are aberrant.
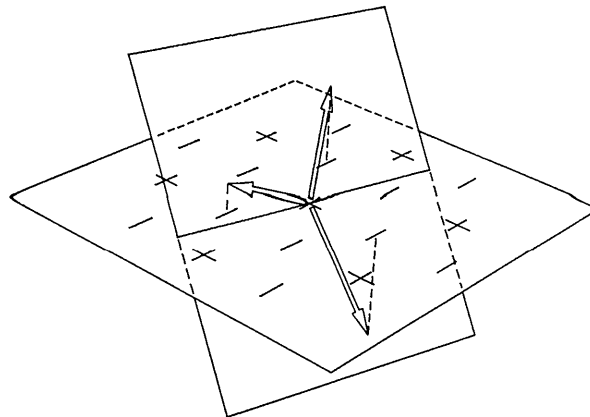

Fig. 4. A triple of vectors which sum to zero and which end in high-intensity columns of the benzene transform. The plane of these vectors passes through the origin but is not parallel to the plane of the ring.
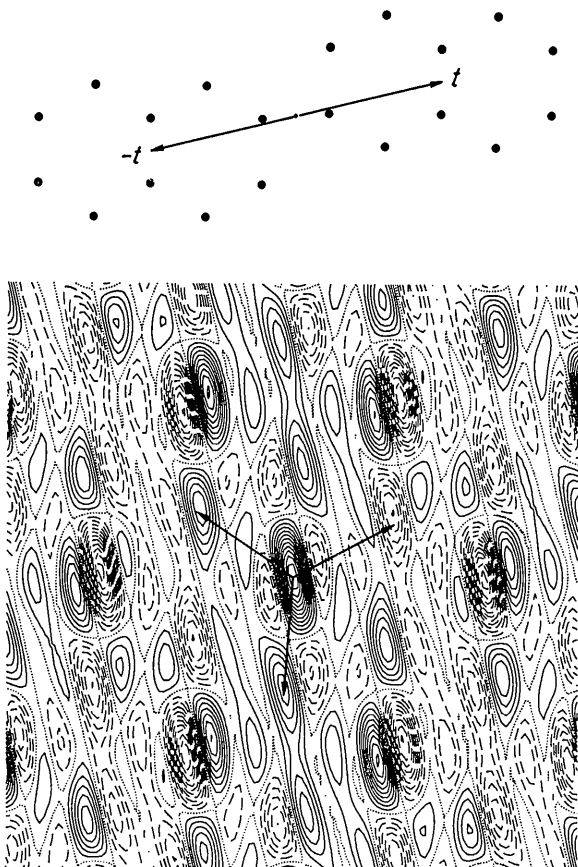
Fig. 5. A planar 20-atom array representing two naphthalene molecules translated from the origin by vectors t and −t and a section through its Fourier transform. A triple of vectors whose sum is zero is superimposed on the transform. This relationship is aberrant and would also have been aberrant for a single benzene ring. [See Fig. 3(c)].
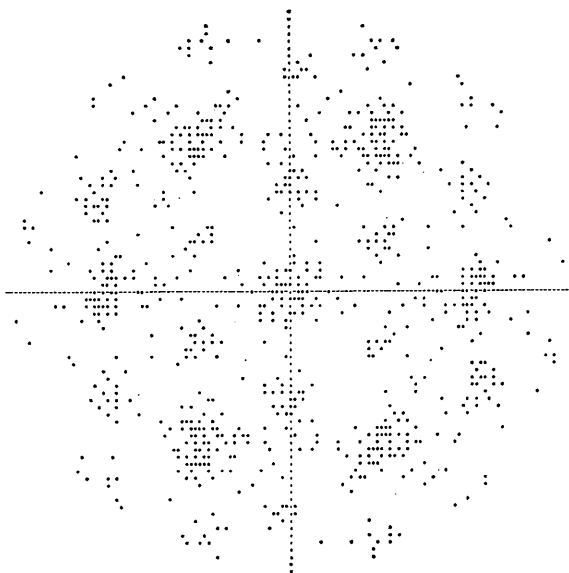
recognize whether structural regularity is present and, if so, to discover the orientation of the regularly repeating unit. Intensity statistics provide a convenient early warning of the presence of structural regularity; for example, centrosymmetric molecules or fragments arranged centrosymmetrically will give rise to the 'hypercentric' distribution described by Lipson & Woolfson (1952). More generally, any parallelism will cause the statistics of the set of normalized structure factor magnitudes to be affected in a known way (Rogers & Wilson, 1953).

If the entire molecular structure or a large fraction thereof is known with certainty, Patterson search techniques (Nordman, 1970) may be used to establish not only the orientation of the molecule in the unit cell but also its location with respect to the symmetry elements. More diffuse structural knowledge, for example, the suspicion that benzene rings are present, is best utilized by a direct search of the $|\mathscr{E}|$-weighted reciprocal lattice for the Fourier transform of the suspected fragment. We will now describe a computer program which we have prepared to carry out this search.

This program accepts as input the indices and normalized structure-factor magnitudes of those reflections with the largest $|\mathscr{E}|$ values, the lattice parameters, and the Cartesian coordinates of the atoms of the suspected fragment in some standard orientation. The program performs an orientational search using Lattman's (1972) rotation angles in the way described by Williams (1973). The limits of the search are established according to the symmetry of the fragment and of the crystal using the methods of Hirshfeld (1968). For each orientation of the fragment a figure of merit

$$S = \sum_i |E(\mathbf{h}_i)|^2 |T(\mathbf{h}_i)|^2 \tag{9}$$

is computed (Lattman & Love, 1970) where the summation is over the reflections included and

$$T(\mathbf{h}_i) = \sum_j \exp (2\pi i \mathbf{h}_i . \mathbf{x}'_j) . \tag{10}$$

Here $\mathbf{x}'_j$ is the position of an atom of the fragment after rotation. A specified number of the orientations having the highest figures of merit are subjected to a finer rotational scan and finally refined to maximize S by adjusting the three orientation angles and a single scale factor that is applied to the atomic coordinates of the fragment. The program then produces a printer plot of the locations of the strong reflections in reciprocal space viewed along a specified axis of the oriented fragment. Fig. 6 displays such a plot in which the typical pattern of the benzene transform is clearly visible.

Finally, the program produces an output on punched cards or auxiliary storage for use in phase determination. This consists of the Miller indices $\mathbf{h}$ and $|\mathscr{E}(\mathbf{h})|$ along with the corresponding transform value $T(\mathbf{h})$. The latter is a signed real number if the fragment

Fig. 6. A printer plot produced by program ORTRAN of the Bragg reflections from methoxydalrubone having $|E| \geq 1\cdot5$. The view is along the columns of the transform.

is centrosymmetric or a complex quantity if it is noncentrosymmetric. This program, called *ORTRAN*, is available from the authors on request.

## V. Taking predictably aberrant relationships into account in phase determination

As mentioned in § I, most direct-method procedures are based on the assumption that all triplet relationships among strong reflections (reflections having large values of $|\mathscr{E}|$) are normal. In order to select the result most likely to be correct in the multisolution approaches (Long, 1965; Main, Woolfson & Germain, 1971) or in the symbolic-addition method (Karle & Karle, 1966), self-consistency indexes based on this assumption are calculated, although Karle (1970b) has argued persuasively against the 'highest self-consistency' approach in deciding which of several possible sets of phases is likely to be the correct one. He notes that, in space group $P\bar{1}$, the most self-consistent solution has all signs positive but is unlikely to be correct for a crystal structure with equal atoms in general positions. In fact, the fraction of relationships of a given magnitude $|\mathscr{E}(-\mathbf{h})\mathscr{E}(\mathbf{k})\mathscr{E}(\mathbf{h}-\mathbf{k})|$ which are aberrant is predictable (Fisher, Hancock & Hauptman, 1970a). Proceeding on the assumption that all relationships are normal succeeds in solving many structures because the aberrant relationships are distributed among most or all of the strong reflections, so that any one reflection is involved in many normal relationships and only a few aberrant ones. In cases of structural regularity of the type described in this paper, however, a relatively small fraction of the strong reflections (those falling in the secondary columns of the benzene transform, for example) are involved in all of the predictably aberrant relationships so that in application of equation (1) these aberrant relationships may easily outweigh normal ones at some stage of the phase determination. Once an incorrect phase is assigned to one reflection, its normal relationships with other reflections will give false phases, leading quickly to disaster.

It is, however, in just these cases that we can predict which relationships are aberrant on the basis of the values of $T(\mathbf{h}_i)$ which have been associated with each strong reflection. As each triplet relationship $(-\mathbf{h},\mathbf{k},\mathbf{h}-\mathbf{k})$ among the strong reflections is found, the quantity $P = T(-\mathbf{h})T(\mathbf{k})T(\mathbf{h}-\mathbf{k})$ is computed and aberrant relationships are recognized as those for which P is negative (for a centrosymmetric fragment) or has a phase angle near π (in the case of a noncentrosymmetric fragment). For structures dominated by a repeating structural unit, the aberrant relationships are included with an assumed phase of π and the self-consistency index is computed accordingly. Notice that for space group $P\bar{1}$ the 'all-plus' solution is no longer necessarily the most nearly self-consistent and that the probability of the most nearly self-consistent set being the correct one is increased.

For cases in which the transform of a fragment can be found in the $|\mathscr{E}|$-weighted reciprocal lattice but is not particularly dominant, it may be preferable simply to omit the relationships suspected of being aberrant from the phase determination, or at least from its early stages.

VI. An application in space group $P\bar{1}$

As an example of the application of these methods we will describe the solution of the crystal structure of methoxydalrubone, $C_{20}H_{20}O_5$, a plant pigment (Dreyer & Thiessen, 1974). Its rather regular benzenoid molecular structure, (I) was not known until revealed by the crystal structure analysis.
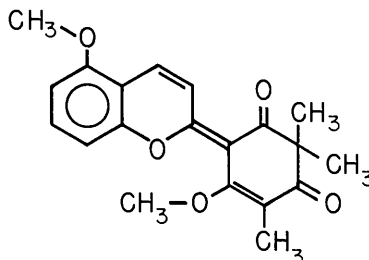
The space group is $P\bar{1}$ ($Z=2$), so that all benzene rings are parallel. The intensity data for 3641 reflections with $\sin\theta/\lambda < 0\cdot642$ were reduced to normalized structure factor magnitudes using program *ORESTES** (Thiessen & Levy, 1973). The cumulative distribution of reflections with $|E|^2 > z$ for methoxydalrubone falls between curves 1 and 2 of Fig. 3 in Rogers & Wilson's (1953) paper, indicating considerable hypersymmetry. Note that in the triclinic system the normalized structure factor E is identical to the quasi-normalized structure factor $\mathscr{E}$.

The printer plot of Fig. 6 is a representation of the location in reciprocal space of the 430 Bragg reflections (and their Friedel mates) from methoxydalrubone with $|\mathscr{E}| \geq 1\cdot5$. The pattern of the benzene transform is clearly visible (compare Fig. 6 with Fig 1) with many reflections in the primary columns [$T(\mathbf{h})$ positive] but with a considerable number in the secondary columns [$T(\mathbf{h})$ negative].

Standard methods for solution of the phase problem, such as Long's (1965) program for reiterative application of Sayre's equation, failed uniformly in the case of methoxydalrubone, even with starting sets chosen by the convergence mapping method of Germain, Main & Woolfson (1971). E maps resulting from these attempts displayed largely planar arrays of peaks on a hexagonal lattice, *i.e.* fused benzene rings with additional atoms at their centers, an arrangement

---

* This program estimates overall absolute scale and thermal parameters (which may be anisotropic if desired) by a least-squares technique, calculates and sorts a set of $|E|$'s, and computes relevant statistics of the set. It is available from the authors on request.

which cannot correspond to chemical reality. The structure was actually solved by expanding the starting set of reflections to include a planar hexagonal array of strong reflections which was recognized in the data set and given signs corresponding to those of the benzene transform. In an *ex post facto* computer experiment, Long's program was modified as described in the preceding section to treat as aberrant all relationships among three strong reflections each of which had a calculated value of $T(\mathbf{h}) \leq -2\cdot2$ for the benzene transform [$T(000) = 6\cdot0$]. The correct solution then had the highest self-consistency index of the 16 sign sets produced.

The totals at the bottom of Table 1 classify the 5568 triplet relationships among the 430 strong reflections ($|E| \geq 1\cdot5$) of methoxydalrubone. The true phase of the product $E(-\mathbf{h})E(\mathbf{k})E(\mathbf{h}-\mathbf{k})$ is based on the final crystal structure, and the predicted phase is obtained from the sign of $P = T(-\mathbf{h})T(\mathbf{k})T(\mathbf{h}-\mathbf{k})$ if $|T(\mathbf{h})| > 2\cdot2$ for each individual transform and is unpredicted otherwise. Of the 177 relationships predicted to be aberrant, only three are normal.

Table 1. *Classification of the* 5568 *triplet relationships of methoxydalrubone*

The totals listed at the bottom divide the relationships into categories according to the actual sum of phases and that predicted from the benzene transform. The remainder of the table subdivides these totals according to the value of $M(D-KS)$ for each relationship.

$M(D-KS)$	Predicted 0 Actual sum of phases		Not predicted Actual sum of phases		Predicted π Actual sum of phases	
	0	π	0	π	0	π
3·5						
	2		1			1
3·0						
	18		1			3
2·5						
	98		15	1		8
2·0						
	541		68	3		28
1·5						
	1861	2	234	5		60
1·0						
	1071	9	388	13		61
0·5						
	222	45	405	27		13
0·0						
	19	72	158	55	3	
−0·5						
	2	10	13	32		
−1·0						
Totals	3834	138	1283	136	3	174

In general, it may be more effective and just as convenient to predict a relationship as aberrant if the product P is less than some negative limit rather than placing a limit on each individual $T(\mathbf{h})$. This point is still under investigation.

VII. Comparison with other methods

It seemed desirable to test whether the method of identifying aberrant relationships based on the recognition of structural regularity is equivalent to or different from the predictions based on statistical methods. We therefore calculated the quantity $M(D-KS)$ in the way described by Fisher, Hancock & Hauptman (1970b) for the 5568 relationships among the strong reflections for methoxydalrubone.

The results are shown in the main body of Table 1 in which we have classified the relationships according to the actual phases of the products, the phases predicted from the benzene transform, and also according to the values of $M(D-KS)$. Large values of this quantity are expected to be associated with normal relationships and indeed they are for those which we predict to be positive. However, the situation seems to be reversed for those relationships which we predict to be aberrant; most of these are also associated with substantial positive values of $M(D-KS)$. Indeed, the only three relationships predicted to be aberrant which are in fact normal are also the only three in this group having *negative* values of $M(D-KS)$. We conclude that our prediction of aberrant relationships based on structural regularity is complementary to the predictions based on statistical methods, and that the two methods employed together may be more powerful than either one alone.

For benzenoid systems at least, predictable aberrancy based on the transform seems to be related to the correction term for 'chance interactions' proposed by Hauptman (1964). The correction term $C(\mathbf{h}_1, -\mathbf{h}_3)$ will be important for just those triples in which each of the vectors ends in a secondary column. Of course, if the orientation of the benzene rings can be determined as we have described here, replacing the correction term by its value averaged over all orientations, as Hauptman proposes, becomes unnecessary.

VIII. Applicability to crystals of higher symmetry

The conclusions of the present paper are applicable when repeated structural units are parallel. This situation occurs most commonly in space group $P1$ with one molecule in the unit cell or in space group $P\bar{1}$ with one or two molecules per unit cell, but can also occur for certain molecular orientations in space groups of higher symmetry. When repeating moieties occur in more than one orientation, two questions arise: (1) are our methods for predicting aberrant relationships still effective, and (2) do these structures have as large a fraction of predictably aberrant relationships as those we have considered above? We have made some preliminary calculations for both real and hypothetical structures in an attempt to answer these questions.

There seems to be little doubt that program *ORTRAN*, given the point-group symmetry of the

crystal, will seek out the superimposed transforms of a given fragment with similar effectiveness for one, two, or four equivalent positions not related by inversion centers. We believe (see Appendix) that the quantity which predicts whether a triple is normal or aberrant in the monoclinic case is

$$\mathscr{P} = P + P' = T(-\mathbf{h})T(\mathbf{k})T(\mathbf{h}-\mathbf{k})$$
$$+ T(-\mathbf{h}')T(\mathbf{k}')T(\mathbf{h}'-\mathbf{k}') \qquad (11)$$

where, if an unprimed vector represents a set of indices hkl, the corresponding primed vector is the symmetry-related reflection ($h\bar{k}l$ for the monoclinic system, b axis unique). Relationships with $|\mathscr{P}|$ larger than some suitably chosen limit are predicted to have the phase of $\mathscr{P}$. It may be desirable to omit from the sum terms with one or more $|T(\mathbf{h})|$ below a specified limit. The corresponding quantity for orthorhombic symmetry would have four terms, etc.

Our tentative answer to the second question is that nonparallel arrays of benzenoid systems seem to produce a considerably smaller fraction of predictably aberrant triples. We take this to indicate that such structures will cause less difficulty in structure solution by standard methods.

We wish to express our appreciation to H.A. Levy for many helpful discussions.

APPENDIX

Equation (11) can be used to predict the phase of a relationship for a structure in which the unit cell contains several fragments with known orientations but in unknown positions. Some of these fragments may be single atoms. We will present here a qualitative derivation of this equation; we hope to develop a quantitative treatment for future publication.

Let

$$\mathscr{E}(\mathbf{k}) = \sum_{i=1}^{n} Z_i(\mathbf{k}) \qquad (12)$$

where

$$Z_i(\mathbf{k}) = N^{-1/2} T_i(\mathbf{k}) \exp(2\pi i \mathbf{k} . \mathbf{t}_i) . \qquad (13)$$

Here n is the number of fragments in a unit cell, $T_i(\mathbf{k})$ is the known value of the complex transform for one fragment in its correct orientation, and $\mathbf{t}_i$ is the unknown translation vector for that fragment. N is the number of atoms in the unit cell, and for this derivation we will assume them all to be of the same kind.

Both the magnitude and phase of each $T_i(\mathbf{k})$ are known, but only the magnitude of each $Z_i(\mathbf{k})$ is available. The magnitude $|\mathscr{E}(\mathbf{k})|$ is also known, and it is clear from (12) that

$$|\mathscr{E}(\mathbf{k})| \le \sum_{i=1}^{n} |Z_i(\mathbf{k})| . \qquad (14)$$

Let us assume temporarily that we are interested only in reflections which are so strong that the equal sign in equation (14) holds. For these reflections

$$\varphi(\mathbf{k}) = \varphi\{Z_i(\mathbf{k})\} \qquad (15)$$

for every fragment i. That is, each contribution to $\mathscr{E}(\mathbf{k})$ in equation (12) has the same phase and that phase is the phase of $\mathscr{E}(\mathbf{k})$.

Now let us consider a triple of three such strong reflections:

$$\mathscr{E}(-\mathbf{h})\mathscr{E}(\mathbf{k})\mathscr{E}(\mathbf{h}-\mathbf{k})$$

$$= \{\sum_{i=1}^{n} Z_i(-\mathbf{h})\} \{\sum_{i=1}^{n} Z_i(\mathbf{k})\} \{\sum_{i=1}^{n} Z_i(\mathbf{h}-\mathbf{k})\}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{m=1}^{n} Z_i(-\mathbf{h})Z_j(\mathbf{k})Z_m(\mathbf{h}-\mathbf{k}) . \qquad (16)$$

It follows from (15) that

$$\varphi(-\mathbf{h}) + \varphi(\mathbf{k}) + \varphi(\mathbf{h}-\mathbf{k}) = \varphi\{Z_i(-\mathbf{h})Z_j(\mathbf{k})Z_m(\mathbf{h}-\mathbf{k})\} , \qquad (17)$$

so that each of the n^3 terms in equation (16) has the same phase, namely the phase of the relationship.

Each of these terms is given by

$$Z_i(-\mathbf{h})Z_j(\mathbf{k})Z_m(\mathbf{h}-\mathbf{k}) = N^{-3/2} T_i(-\mathbf{h})T_j(\mathbf{k})T_m(\mathbf{h}-\mathbf{k})$$
$$\times \exp\{2\pi i[-\mathbf{h} . \mathbf{t}_i + \mathbf{k} . \mathbf{t}_j + (\mathbf{h}-\mathbf{k}) . \mathbf{t}_m]\} \qquad (18)$$

and in general we cannot calculate its phase directly because the vectors $\mathbf{t}_i$ are unknown. There are, however, n terms for which $i = j = m$ and for these

$$Z_i(-\mathbf{h})Z_i(\mathbf{k})Z_i(\mathbf{h}-\mathbf{k}) = N^{-3/2} T_i(-\mathbf{h})T_i(\mathbf{k})T_i(\mathbf{h}-\mathbf{k})$$
$$\qquad (19)$$

because the exponent is zero. Then for these very strong reflections the phase of the relationship is given by

$$\varphi(-\mathbf{h}) + \varphi(\mathbf{k}) + \varphi(\mathbf{h}-\mathbf{k}) = \varphi\{T_i(-\mathbf{h})T_i(\mathbf{k})T_i(\mathbf{h}-\mathbf{k})\} \qquad (20)$$

where each of the n values of i should yield the same phase.

We will now show that in practical problems, for which the magnitudes of $\mathscr{E}(-\mathbf{h})$, $\mathscr{E}(\mathbf{k})$, and $\mathscr{E}(\mathbf{h}-\mathbf{k})$ do not have their maximum allowable values, better results can be obtained by rewriting equation (20) as

$$\varphi(-\mathbf{h}) + \varphi(\mathbf{k}) + \varphi(\mathbf{h}-\mathbf{k}) = \varphi\{\mathscr{P}(\mathbf{h},\mathbf{k})\} \qquad (21)$$

where

$$\mathscr{P}(\mathbf{h},\mathbf{k}) = \sum_{i=1}^{n} T_i(-\mathbf{h})T_i(\mathbf{k})T_i(\mathbf{h}-\mathbf{k}) . \qquad (22)$$

This is essentially equation (11) of the text.

If $|\mathscr{E}(\mathbf{k})|$ is large but less than $\sum_{i=1}^{n} |Z_i(\mathbf{k})|$ then the

complex quantities $Z_i(\mathbf{k})$ need not have exactly the phase of $\mathscr{E}(\mathbf{k})$. Let the phase differences be

$$\Delta\varphi_i(\mathbf{k}) = \varphi\{Z_i(\mathbf{k})\} - \varphi(\mathbf{k}) . \quad (23)$$

These differences require us to modify the simple discussion presented above. Incorporating them in equation (20) we obtain

$$\varphi(-\mathbf{h}) + \varphi(\mathbf{k}) + \varphi(\mathbf{h}-\mathbf{k}) = \varphi\{T_i(-\mathbf{h})T_i(\mathbf{k})T_i(\mathbf{h}-\mathbf{k})\}$$
$$-\Delta\varphi_i(-\mathbf{h}) - \Delta\varphi_i(\mathbf{k}) - \Delta\varphi_i(\mathbf{h}-\mathbf{k}) , \quad (24)$$

an equation which does not yield the phase of the relationship because the phase differences are unknown. The effects of these differences tend to be reduced in equation (21), however, for two reasons. First, the values of $\Delta\varphi_i(\mathbf{k})$ tend to be smaller for larger $|Z_i(\mathbf{k})|$ and therefore the larger terms which dominate equation (22) will tend to have the smaller errors.

Second, every positive deviation $\Delta\varphi_i(\mathbf{k})$ must be accompanied by a negative deviation $\Delta\varphi_j(\mathbf{k})$ in some other term of equation (12). These deviations will affect different terms of equation (22) in opposite ways and tend to cancel each other.

Using equation (21) and following the procedure of Karle & Karle (1966), we can readily derive a revised tangent formula:

$$\tan \varphi(\mathbf{h})$$
$$= \frac{\langle w(\mathbf{h},\mathbf{k}) \sin [\varphi(\mathbf{k}) + \varphi(\mathbf{h}-\mathbf{k}) - \varphi\{\mathscr{P}(\mathbf{h},\mathbf{k})\}]\rangle_{\mathbf{k}}}{\langle w(\mathbf{h},\mathbf{k}) \cos [\varphi(\mathbf{k}) + \varphi(\mathbf{h}-\mathbf{k}) - \varphi\{\mathscr{P}(\mathbf{h},\mathbf{k})\}]\rangle_{\mathbf{k}}} . \quad (25)$$

Here $w(\mathbf{h},\mathbf{k})$ is an appropriate weighting factor that reflects the precision with which $\varphi\{\mathscr{P}(\mathbf{h},\mathbf{k})\}$ represents the phase of each relationship. We hope to present a more complete discussion of this quantity in a future publication.

References

CHAUDHURI, B. (1972). In *Optical Transforms*, chap 3, pp. 71–113. Edited by H. LIPSON. London: Academic Press.

DREYER, D. L. & THIESSEN, W. E. (1974). Paper in preparation

FISHER, J. E., HANCOCK, J. H. & HAUPTMAN, H. A. (1970*a*). The Conditional Probability Distribution of Crystal Structure Invariants, NRL Report 7132, Naval Research Laboratory, Washington, D. C.

FISHER, J. E. ,HANCOCK, J. H. & HAUPTMAN, H. (1970*b*). Computer Program for the Calculation of the Crystal Structure Invariants Based on Certain Conditional Distributions, NRL Report 7157, Naval Research Laboratory, Washington, D. C.

GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971). *Acta Cryst.* A27, 368–376.

HAUPTMAN, H. (1964). *Acta Cryst.* 17, 1421–1433.

HAUPTMAN, H. A. (1970). In *Crystallographic Computing*, chap A5, pp. 41–51. Edited by F. R. AHMED, S. R. HALL & C. P. HUBER. Copenhagen: Munksgaard.

HAUPTMAN, H. (1972). *Crystal Structure Determination. The Role of the Cosine Seminvariants*, pp. 192–197. New York: Plenum.

HIRSHFELD, F. L. (1968). *Acta Cryst.* A24, 301–311.

KARLE, J. (1970*a*). *Acta Cryst.* B26, 1614–1617.

KARLE, J. (1970*b*). In *Crystallographic Computing*, chap. A7, pp. 75–77. Edited by F. R. AHMED, S. R. HALL and C. P. HUBER. Copenhagen: Munksgaard.

KARLE, J. & HAUPTMAN, H. (1959). *Acta Cryst.* 12, 404–410.

KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* 21, 849–859.

KNOTT, G. (1940). *Proc. Phys. Soc.* 52, 229–238.

LATTMAN, E. E. (1972). *Acta Cryst.* B28, 1065–1068.

LATTMAN, E. E. & LOVE, W. E. (1970). *Acta Cryst.* B26, 1854–1857.

LIPSON, H. & TAYLOR, C. A. (1958). *Fourier Transforms and X-Ray Diffraction*. London: Bell.

LIPSON, H. & WOOLFSON, M. M. (1952). *Acta Cryst.* 5, 680–682.

LONG, R. E. (1965). Ph. D. dissertation, Univ. of California, Los Angeles.

MACKAY, K. J. H. (1962). *Acta Cryst.* 15, 157–160.

MAIN, P., WOOLFSON, M. M. & GERMAIN, G. (1971). The report: *MULTAN, A Computer Program for the Automatic Solution of Crystal Structures*. University of York, York, England.

NORDMAN, C. E. (1970). In *Crystallographic Computing*, Topic B, pp. 81–123, containing contributions from R. A. JACOBSEN, N. C. SEEMAN, P. TOLLIN, R. HUBER, J. HORNSTRA, C. E. NORDMAN, and J. W. SCHILLING. Edited by F. R. AHMED, S. R. HALL and C. P. HUBER. Copenhagen: Munksgaard.

ROGERS, D. & WILSON, A. J. C. (1953). *Acta Cryst.* 6, 439–449.

STADLER, H. P., BOLTON, W. & MAITLAND, R. (1963). *Acta Cryst.* 17, 1025–1028.

THIESSEN, W. E. & LEVY, H. A. (1973). Program Accession No. 247 in *World List of Crystallographic Computer Programs*, 3rd ed. *J. Appl. Cryst.* 6, 309–346.

WILLIAMS, D. E. (1973). *Acta Cryst.* A29, 408–414.