Acta Cryst. (1971). B27, 2063

A Generalization of the Tangent Formula

BY J. KARLE

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

(*Received* 27 *January* 1971)

By making use of the higher order inequalities which were derived on the basis that the electron density function in the unit cell of a crystal is non-negative, a generalization is obtained of the formula presently used for the tangent of the phase angle of a structure factor. Also derived are probability measures in the form of probability distributions and variances to accompany the generalized tangent formula. The generalization is based on the fact that because of the non-negativity of the electron density, a structure factor is bounded in the complex plane by a circle centered at a point in the complex plane. The location of this center is of particular significance, representing the expected value of the structure factor of interest, given the values of the structure factors required to compute the central point. Both the radius of the bounding circle and its center are expressible as ratios of determinants involving the structure factors. Theoretical considerations imply that the restrictiveness of the bound on the structure factor of interest increases as the order of the determinants involved increases. Corresponding sign determining formulas and probability measures are presented for the centrosymmetric case.

Introduction

The investigation of the mathematical consequences of the non-negativity of the electron density function led to an infinite set of inequalities of increasing order among the structure factors (Karle & Hauptman, 1950). It has been pointed out (Goedkoop, 1950; Hauptman & Karle, 1950) that for structure factors representing point atoms, the determinants of sufficiently high order are equal to zero and therefore become equalities among the structure factors. It is also known from the mathematical literature, as pointed out by Tsoucaris (1970), that the values of these non-negative determinants, scaled to contain unitary structure factors as elements, monotonically approach zero as the order of the determinants increases.

It is a remarkable fact that the main phase determining formulas in current use, *e.g.* $\sum_2$, the sum of angles formula and the tangent formula are contained in the inequality of only the third order. In view of the greater constraints implied by the higher order inequalities, it is desirable to consider the development of calculations for facilitating their use. Recent investigations concerned with deriving phase information directly from the high-order determinants have been carried out by Tsoucaris (1970). This type of calculation involves the use of a large number of known phases in a high-order determinant in order to determine a most probable set of phase values for a relatively small number of reflections. So far the applications have been limited to centrosymmetric crystals. In this note, an alternative form for the inequalities, one in which a single structure factor is bounded within a circle centered in the complex plane, will be used to arrive at tangent formulas, based on higher order inequalities, which appear to use phase information in a more sophisticated fashion than the one presently used. To aid in facilitating their use, appropriate probability measures will be associated with the tangent formulas.

Mathematical development

The quasi-normalized structure factor is defined

$$\mathscr{E}_{\mathbf{h}} = \left(\sum_{j=1}^{N} f_{j\mathbf{h}}^2\right)^{-1/2} \sum_{j=1}^{N} f_{j\mathbf{h}} \exp\left(2\pi i \mathbf{h} \cdot \mathbf{r}_j\right) \qquad (1)$$

or

$$\mathscr{E}_{\mathbf{h}} \simeq \sigma_2^{-1/2} \sum_{j=1}^{N} Z_j \exp\left(2\pi i \mathbf{h} \cdot \mathbf{r}_j\right) \qquad (2)$$

where

$$\sigma_n = \sum_{j=1}^{N} Z_j^n . \qquad (3)$$

The quantity $f_{j\mathbf{h}}$ is the atomic scattering factor for the jth atom in a unit cell containing N atoms, Z_j is its atomic number and the components of $\mathbf{h}$ are the Miller indices for a particular reflection.

A typical inequality representing the infinite set appropriate to point atoms may be written as the Toeplitz form (Karle & Hauptman, 1950),

$$D_{m,p}(\mathbf{h}) = \begin{vmatrix} \mathscr{E}_{000} & \mathscr{E}_{-\mathbf{k}_1}\mathscr{E}_{-\mathbf{k}_2} \cdots \mathscr{E}_{-\mathbf{h}} \\ \mathscr{E}_{\mathbf{k}_1} & \mathscr{E}_{000}\ \mathscr{E}_{\mathbf{k}_1-\mathbf{k}_2} \cdots \mathscr{E}_{\mathbf{k}_1-\mathbf{h}} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \mathscr{E}_{\mathbf{k}_{m-2}}\mathscr{E}_{\mathbf{k}_{m-2}-\mathbf{k}_1} \cdots \mathscr{E}_{\mathbf{k}_{m-2}-\mathbf{h}} \\ \mathscr{E}_{\mathbf{h}} \quad \mathscr{E}_{\mathbf{h}-\mathbf{k}_1} \cdots\cdots \mathscr{E}_{000} \end{vmatrix} \geq 0 \quad (\text{rank } N) \qquad (4)$$

The determinant $D_{m,p}(\mathbf{h})$ of order m is formed by composing the first column with $\mathscr{E}_{000}$ as the first element, followed by $m-2$ arbitrarily chosen structure factors and then finally $\mathscr{E}_{\mathbf{h}}$, the structure factor of interest. The subscript p labels a particular set of $m-2$ vectors $\mathbf{k}_1, \ldots, \mathbf{k}_{m-2}$ such that $D_{m,p}(\mathbf{h}) \equiv D_{m,\mathbf{k}_1 \cdots \mathbf{k}_{m-2}}(\mathbf{h})$. In practice the magnitudes of the structure factors in the first column are generally chosen to be large. Once the

first column is specified, the remainder of the determinant is readily constructed by first forming the first row with the complex conjugates of the elements in the first column in the same sequence. Each element in the body of the determinant is a structure factor whose subscript is the sum of the subscripts of the elements in the first row and first column corresponding to the row and column of the element of interest.

The structure factor $\mathscr{E}_{\mathbf{h}}$ for a noncentrosymmetric reflection is bound by a circle of radius $r_{m,p}(\mathbf{h})$ in the complex plane centered about the point $\delta_{m,p}(\mathbf{h})$ (Karle & Hauptman, 1950),

$$|\mathscr{E}_{\mathbf{h}}-\delta_{m,p}(\mathbf{h})|\leq r_{m,p}(\mathbf{h}) \tag{5}$$

where

$$\delta_{m,p}(\mathbf{h})=\Delta'_{m,p}(\mathbf{h})/\Delta_{m,p} \tag{6}$$

and

$$r_{m,p}(\mathbf{h})=\Delta^{1/2}_{1,m,p}\Delta^{1/2}_{2,m,p}(\mathbf{h})/\Delta_{m,p}\,. \tag{7}$$

The determinant Δ' is formed from D by omitting the first row and last column of D, replacing the element $\mathscr{E}_{\mathbf{h}}$ by zero and multiplying by $(-1)^{m-1}$,

$$\Delta'_{m,p}(\mathbf{h})=(-1)^{m-1}\begin{vmatrix} \mathscr{E}_{\mathbf{k}_1} & \mathscr{E}_{000} & \mathscr{E}_{\mathbf{k}_1-\mathbf{k}_2} & \cdots & \mathscr{E}_{\mathbf{k}_1-\mathbf{k}_{m-2}} \\ \mathscr{E}_{\mathbf{k}_2} & \mathscr{E}_{\mathbf{k}_2-\mathbf{k}_1} & \mathscr{E}_{000} & \cdots & \mathscr{E}_{\mathbf{k}_2-\mathbf{k}_{m-2}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathscr{E}_{\mathbf{k}_{m-2}} & \mathscr{E}_{\mathbf{k}_{m-2}-\mathbf{k}_1} & \cdots & \cdots & \mathscr{E}_{000} \\ 0 & \mathscr{E}_{\mathbf{h}-\mathbf{k}_1} & \cdots & \cdots & \mathscr{E}_{\mathbf{h}-\mathbf{k}_{m-2}} \end{vmatrix} \tag{8}$$

The determinant Δ is formed from D by omitting the first and last rows and columns of D,

$$\Delta_{m,p}=\begin{vmatrix} \mathscr{E}_{\mathbf{k}_1-\mathbf{k}_1} & \mathscr{E}_{\mathbf{k}_1-\mathbf{k}_2} & \cdots & \mathscr{E}_{\mathbf{k}_1-\mathbf{k}_{m-2}} \\ \mathscr{E}_{\mathbf{k}_2-\mathbf{k}_1} & \mathscr{E}_{\mathbf{k}_2-\mathbf{k}_2} & \cdots & \mathscr{E}_{\mathbf{k}_2-\mathbf{k}_{m-2}} \\ \cdots & \cdots & \cdots & \cdots \\ \mathscr{E}_{\mathbf{k}_{m-2}-\mathbf{k}_1} & \mathscr{E}_{\mathbf{k}_{m-2}-\mathbf{k}_2} & \cdots & \mathscr{E}_{\mathbf{k}_{m-2}-\mathbf{k}_{m-2}} \end{vmatrix} \tag{9}$$

The diagonal elements of (9) are evidently $\mathscr{E}_{000}$. The determinants Δ_1 and Δ_2 and formed from D by omitting the last row and column of D and omitting the first row and column of D, respectively,

$$\Delta_{1,m,p}\Delta_{2,m,p}(\mathbf{h})=\begin{vmatrix} \mathscr{E}_{000} & \mathscr{E}_{-\mathbf{k}_1} & \cdots & \mathscr{E}_{-\mathbf{k}_{m-2}} \\ \mathscr{E}_{\mathbf{k}_1} & \mathscr{E}_{000} & \cdots & \mathscr{E}_{\mathbf{k}_1-\mathbf{k}_{m-2}} \\ \mathscr{E}_{\mathbf{k}_2} & \mathscr{E}_{\mathbf{k}_2-\mathbf{k}_1} & \cdots & \mathscr{E}_{\mathbf{k}_2-\mathbf{k}_{m-2}} \\ \cdots & \cdots & \cdots & \cdots \\ \mathscr{E}_{\mathbf{k}_{m-2}} & \mathscr{E}_{\mathbf{k}_{m-2}-\mathbf{k}_1} & \cdots & \mathscr{E}_{000} \end{vmatrix} \times \begin{vmatrix} \mathscr{E}_{000} & \mathscr{E}_{\mathbf{k}_1-\mathbf{k}_2} & \cdots & \mathscr{E}_{\mathbf{k}_1-\mathbf{h}} \\ \mathscr{E}_{\mathbf{k}_2-\mathbf{k}_1} & \mathscr{E}_{000} & \cdots & \mathscr{E}_{\mathbf{k}_2-\mathbf{h}} \\ \mathscr{E}_{\mathbf{k}_3-\mathbf{k}_1} & \mathscr{E}_{\mathbf{k}_3-\mathbf{k}_2} & \cdots & \mathscr{E}_{\mathbf{k}_3-\mathbf{h}} \\ \cdots & \cdots & \cdots & \cdots \\ \mathscr{E}_{\mathbf{h}-\mathbf{k}_1} & \mathscr{E}_{\mathbf{h}-\mathbf{k}_2} & \cdots & \mathscr{E}_{000} \end{vmatrix} \tag{10}$$

A simple manipulation of the inequality (5) can demonstrate its relationship to a general tangent formula. The assumption is made that

$$\mathscr{E}_{\mathbf{h}}\propto\langle\delta_{m,p}(\mathbf{h})\rangle_p\,. \tag{11}$$

Algebraic analysis involving expansion of the determinants and use of the structure factor equations (1) or (2) can, in fact, verify this assumption. If we express

$$\mathscr{E}_{\mathbf{h}}=|\mathscr{E}_{\mathbf{h}}|\exp(i\varphi_{\mathbf{h}}) \tag{12}$$

and

$$\delta_{m,p}(\mathbf{h})=|\delta_{m,p}(\mathbf{h})|\exp[i\theta_{m,p}(\mathbf{h})]\,, \tag{13}$$

then, from (11) the generalized tangent formula may be written,

$$\tan\varphi_{\mathbf{h}}\simeq\frac{\sum\limits_p|\delta_{m,p}(\mathbf{h})|\sin\theta_{m,p}(\mathbf{h})}{\sum\limits_p|\delta_{m,p}(\mathbf{h})|\cos\theta_{m,p}(\mathbf{h})}\,. \tag{14}$$

It is easily determined from equation (6) that the tangent formula presently used corresponds to the case when $m=3$.

It is desirable to associate probability measures with the tangent formula (14). In doing so the central limit theorem is employed in which it is assumed that the expected values of the real and imaginary parts of $\mathscr{E}_{\mathbf{h}}$ are given by the real and imaginary parts of $\delta_{m,p}(\mathbf{h})$. It is also assumed that the variance of both the real and imaginary parts of $\mathscr{E}_{\mathbf{h}}$ is equal to 0·5. With these assumptions, the probability distribution for $\varphi_{\mathbf{h}}$ can be expressed as

$$P(\varphi_{\mathbf{h}})=[2\pi I_0(\alpha)]^{-1}\exp[\alpha\cos(\varphi_{\mathbf{h}}-\beta)] \tag{15}$$

where

$$\alpha=\{[\sum_p\kappa_{m,p}(\mathbf{h})\cos\theta_{m,p}(\mathbf{h})]^2+[\sum_p\kappa_{m,p}(\mathbf{h})\sin\theta_{m,p}(\mathbf{h})]^2\}^{1/2}\,, \tag{16}$$

$$\kappa_{m,p}(\mathbf{h})=2|\mathscr{E}_{\mathbf{h}}\delta_{m,p}(\mathbf{h})| \tag{17}$$

and $\tan\beta$ is given by the right side of equation (14). Equation (15) is identical in form with equation (3·25) of Karle & Karle (1966) and for the case that $m=3$, the equations are the same except for the unimportant replacement of $\sigma_3/\sigma_2^{3/2}$ by $\mathscr{E}_{000}^{-1}=\sigma_2^{1/2}/\sigma_1$. For equal atoms both functions become $N^{-1/2}$. For this special case of $m=3$, equation (15) is also equivalent to the probability formulas (6), (7) and (10) of Cochran (1955). The variance formula (3·33) and Fig. 2 of Karle & Karle (1966) are applicable with the new definition of α in equation (16). (Note that a typographical error occurs in equation (3·33). The sign before the last term should be minus. Fig. 2, however, was computed correctly.)

For the case of a centrosymmetric crystal, inequality (5) implies a bound on the real axis with

$$s\mathscr{E}_{\mathbf{h}}\simeq s\sum_p\delta_{m,p}(\mathbf{h}) \tag{18}$$

where s means 'sign of'. The probability that the sign of $\mathscr{E}_{\mathbf{h}}$ is positive is

$$P_+(\mathbf{h})\simeq\tfrac{1}{2}+\tfrac{1}{2}\tanh|\mathscr{E}_{\mathbf{h}}|\sum_p\delta_{m,p}(\mathbf{h})\,. \tag{19}$$

It is readily seen that for $m=3$, and replacement of the quasi-normalized structure factors $\mathscr{E}$ by the normalized structure factors E, equation (18) becomes the $\sum_2$

relation. For $m=3$, equation (19) corresponds to the one of Cochran & Woolfson (1955).

Concluding remarks

In deriving the probability measures, the variances used for the real and imaginary parts of the structure factor were based on the assumption that the arguments of the trigonometric terms in the structure factor equation were uniformly and independently distributed on the interval $-\pi$ to π. This is a quite accurate assumption when the determinant $\delta_{m,p}(\mathbf{h})$ is of fairly low order, since the known values of the elements in a single determinant of low order usually do not limit significantly the possible positions of the atoms in the unit cell. In view of the fact that the theory presented here does not take into account the important questions concerning the effects of inaccuracies in the data, it is probably desirable to have conservative estimates of the variance in the probability formulas (15) and (19), especially when m is small. However, the known values for the elements of $\delta_{m,p}(\mathbf{h})$ do constrain the distributions of the trigonometric terms in the structure factor equations. When m is large, values of 0·5 for the variance of the real and imaginary parts of the structure factor for noncentrosymmetric crystals and 1·0 for the variance of the structure factor for centrosymmetric ones are too large. A more accurate measure of the variance can be based on the radius of the bounding circle $r_{m,p}(\mathbf{h})$. If we multiply $r_{m,p}(\mathbf{h})$ by $\mathscr{E}_{000}^{-1}$, we have the bounding radius which would be obtained if the inequalities were based on unitary structure factors, in other words the bounding radius for $U_{\mathbf{h}}$. This radius is essentially unity when m is small, except when *e.g.* $U_{\mathbf{h}}$ and $U_{\mathbf{h}-\mathbf{k}}$ are very large for $m=3$, and approaches zero monotonically, with the addition of new structure factor information, as m approaches $N+2$. It apparently possesses the correct properties to represent the variance which reasonably should be associated with the bounding radius squared. We therefore set

$$\sigma_{m,p}^2(\mathbf{h}) \simeq (\mathscr{E}_{000}^{-1} r_{m,p}(\mathbf{h}))^2 \qquad (20)$$

for centrosymmetric crystals and half this value for noncentrosymmetric ones. If the variance $\sigma_{m,p}^2(\mathbf{h})$ were to be introduced into the probability formulas (15) and (19), we would have the following replacement,

$$\delta_{m,p}(\mathbf{h}) \rightarrow \delta_{m,p}(\mathbf{h})/\sigma_{m,p}^2(\mathbf{h}) \,. \qquad (21)$$

The sums in the generalized tangent formula (14) and the generalized sign formula (18) are taken over p, each p representing a particular set of vectors $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_{m-2}$. It is conceivable that there may be some advantage to summing also over m.

Experience with formulas arising from $m=3$, has shown that it is worthwhile in practice to replace the quasi-normalized structure factors $\mathscr{E}$ by the normalized structure factors E. This replacement may also be worthwhile in working with the higher order determinants.

The inequality (4) would be equally valid if the elements were all changed to $|\mathscr{E}|^2$ or $|\mathscr{E}|^2-1$. The rank however would change to $N(N-1)+1$ or $N(N-1)$, respectively. In view of the fact that the quantities δ and r in inequality (5) are given by the ratio of determinants that differ in order only by unity, it may be possible to compute determinants of high enough order accurately and rapidly enough to afford a practical method for extrapolating experimental intensity data. The possibility of extrapolation is also applicable to the magnitude and phase information available for poorly resolved structures.

Computer technology has developed by several orders of magnitude since 1950 when the general inequalities appeared in the literature. Inasmuch as most of the major features of the present direct methods of structure analysis are contained in the third order inequality $D_{3,\mathbf{k}}(\mathbf{h})$, it seems worthwhile to take advantage of the advances in computer technology to determine what additional analytical facility the more restrictive higher order determinants may offer.

References

COCHRAN, W. (1955). *Acta Cryst.* **8**, 473.
COCHRAN, W. & WOOLFSON, M. M. (1955). *Acta Cryst.* **8**, 1.
GOEDKOOP, J. A. (1950). *Acta Cryst.* **3**, 374.
HAUPTMAN, H. & KARLE, J. (1950). *Phys. Rev.* **60**, 244.
KARLE, J. & HAUPTMAN, H (1950). *Acta Cryst.* **3**, 181.
KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849.
TSCOUCARIS, G. (1970). *Acta Cryst.* A**26**, 492.