

New Approaches to Structure Analysis

BY J. KARLE

Laboratory for the Structure of Matter, Naval Research Laboratory, Washington, DC 20375-5341, USA

(Received 1 July 1994; accepted 3 January 1995)

Abstract

There are relationships among the phases and magnitudes of the structure factors that have not been extensively studied regarding their potential for enhancing procedures for structure determination. These relationships arise from a special way of writing the determinantal inequalities that form the necessary and sufficient conditions for a Fourier series to be non-negative. This particular form also lends itself readily to the development of probability measures by the use of the central limit theorem. Higher-order determinants are of interest. The relationships among the phases and magnitudes of the structure factors are algebraic relationships and the focus is on those which retain their reliability, even though the magnitudes of the structure factors contain experimental errors. The future utility of the algebraic relationships depends upon the development of suitable algorithms for solving them to obtain values for the unknown phases. One approach concerns a method for extending the range of least-squares calculations by modifying the defining equations without changing the global minima and by further altering the nature of the minimization function from time to time during the course of the least-squares calculation, while still preserving the global minima. The objective is to smooth the minimization function and alter the remaining false minima from time to time so that the minimization function is not trapped in a false minimum. Some calculations have been made that indicate the nature of the algebraic relationships among the phases and magnitudes and how the results of the calculations are benefited by having large values for the structure-factor magnitudes in the determinants. It is also noted that large numbers of lower-order determinants can be spawned from one large order one. A possible virtue of this is that the structure factors which appear in the one large determinant appear with great redundancy in the lower-order ones. A question to be answered is whether it is better to use many lower-order determinants for the algebraic equations or one larger determinant, or perhaps some combinations. There are very many alternatives to pursue and probably extensive experimentation will be needed to determine what is optimal.

Introduction

There is reason to believe that structure analysis may continue to improve in terms of the speed and facility

with which increasingly complex structures can be analyzed. This follows not only from the general overdeterminacy of the data for crystal structure problems, but also from the existence of a variety of powerful phase relationships that have received little or no attention. The latter relationships can be rather complex, accounting at least in part for the lack of attention. Modern computing facilities, however, can readily handle much of the complexity. Additional help may derive from a method for finding global minima in systems of nonlinear simultaneous equations (Karle, 1991).

The phase relationships to be discussed here are based on the determinants that arise from the non-negativity of a Fourier series (Karle & Hauptman, 1950). There are many ways in which the determinants may be used, *e.g.* as systems of simultaneous equations and as the components of generalized tangent formulae (Karle, 1971). The determinants also appear as elements of probability functions (Tsoucaris, 1970; Karle, 1971, 1978). The probabilistic implications of the determinant inequality theory may be realized by applying the central limit theorem to a particular form for writing the inequalities, namely (30) of Karle & Hauptman (1950).

There are a very large number of possible variations in the selection and mode of application of the formulae. The determination of their potential and the evaluation of their relative efficacies, therefore, could occupy much future effort.

In many circumstances, the formulae to be discussed are quite accurate. Although they are based on the determinants that characterize non-negative Fourier series, the mathematical relationships are not based on the maximum determinant rule, given in special form (Tsoucaris, 1970) and discussed in more general form (Heinerman, Kroon & Krabbendam, 1979; Karle, 1978). They are, instead, algebraic relationships that derive from special properties of the determinants. Some of the special properties can be used to derive joint probability distributions of the elements in the determinants that are consistent with the general form of the maximum determinant rule (Karle, 1978). The algebraic relationships are, nevertheless, quite distinct from the rule.

Progress has been made in the application of the general form of the maximum determinant rule as described by de Gelder, de Graaff & Schenk (1993). Although the maximum determinant rule has not been proven rigorously, it is likely to be correct and that a

correct structure is to be found among those atomic arrangements that make sets of determinants with values close to their maximum. As noted, however, this article offers an alternative approach.

The objective of this paper is to provide new mathematical opportunities with the potential to enhance capabilities in evaluating phases. New developments in structure determination evolve from new mathematical opportunities. This article describes some of the special features of the determinants related to the phase problem, including the determinants of higher and higher order, associated generalized tangent formulae based on them and corresponding probability distributions. Also discussed is a way of extending the utility of the least-squares method. There is the potential here for useful analytical algorithms. In developing such algorithms in the future, progress may also stem from the discovery of new techniques in or new applications of numerical analysis.

Examination of the mathematical presentation will indicate that there is a vast number of ways to generate the almost endless number of determinants and tangent formulae. There is, therefore, a vast number of choices that need to be considered and evaluated in the course of developing useful algorithms.

Determinants

The determinants associated with non-negative Fourier series can be written in terms of structure factors, unitary structure factors or quasi-normalized structure factors. It is not accurate to write them in terms of normalized structure factors. We choose here to present the relevant mathematics in terms of quasi-normalized structure factors. Quasi-normalized structure factors \mathcal{E}_h are defined by

$$\mathcal{E}_h = \left(\sum_{j=1}^N f_{jh}^2 \right)^{-1/2} \sum_{j=1}^N f_{jh} \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_j), \quad (1)$$

where the sum on the right side defines F_h , the structure factor associated with the reciprocal vector \mathbf{h} , and f_{jh} represents the atomic scattering factor for the j th atom in the unit cell containing N atoms. A general form for the determinants of interest is

$$D_{m,p} = \begin{vmatrix} \mathcal{E}_0 & \mathcal{E}_{-\mathbf{k}_1} & \mathcal{E}_{-\mathbf{k}_2} & \cdots & \mathcal{E}_{-\mathbf{k}_{m-1}} \\ \mathcal{E}_{\mathbf{k}_1} & \mathcal{E}_0 & \mathcal{E}_{\mathbf{k}_1-\mathbf{k}_2} & \cdots & \mathcal{E}_{\mathbf{k}_1-\mathbf{k}_{m-1}} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \mathcal{E}_{\mathbf{k}_{m-2}} & \mathcal{E}_{\mathbf{k}_{m-2}-\mathbf{k}_1} & \cdots & \cdots & \mathcal{E}_{\mathbf{k}_{m-2}-\mathbf{k}_{m-1}} \\ \mathcal{E}_{\mathbf{k}_{m-1}} & \mathcal{E}_{\mathbf{k}_{m-1}-\mathbf{k}_1} & \cdots & \cdots & \mathcal{E}_0 \end{vmatrix} \geq 0 \text{ (rank } N), \quad (2)$$

where m is the order of the determinant and p represents a particular set of $\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_{m-1}$. Rank N indicates

$D_{m,p} = 0$ when $m \geq N$, where N is the number of atoms in the unit cell. There is the assumption that \mathcal{E} truly represents a point atom structure (Hauptman & Karle, 1950; Goedkoop, 1950). This is strictly true when f_{jh} in (1) all have the same shape as a function of scattering angle. The necessary and sufficient condition for having a non-negative Fourier series is the non-negativity of an infinite set of determinants of type (2) (Karle & Hauptman, 1950). Alternatively, the determinantal inequalities can be written based on (30) of Karle & Hauptman (1950). A comparable expression is

$$|\mathcal{E}_{\mathbf{k}_q} - \delta_{m,p}| \leq r_{m,p}, \quad (3)$$

which is somewhat more general in form, although not in content, since $\mathcal{E}_{\mathbf{k}_q}$ refers to any element in $D_{m,p}$, not just the particular element in the last row and first column. (3) states that the quasi-normalized structure factor $\mathcal{E}_{\mathbf{k}_q}$ is bounded by a circle in the complex plane whose center is $\delta_{m,p}$ and radius is $r_{m,p}$. The quantities $\delta_{m,p}$ and $r_{m,p}$ are defined by

$$\delta_{m,p} = \Delta'_{m,p} / \Delta_{m,p} \quad (4)$$

and

$$r_{m,p} = \Delta_{1,m,p}^{1/2} \Delta_{2,m,p}^{1/2} / \Delta_{m,p}, \quad (5)$$

where:

(a) $\Delta'_{m,p}$ is formed from $D_{m,p}$ by omitting the row and column that contain $\mathcal{E}_{-\mathbf{k}_q}$, replacing the element $\mathcal{E}_{\mathbf{k}_q}$ by zero, and multiplying by $(-1)^{r+s}$ where r is the row and s is the column in which $\mathcal{E}_{\mathbf{k}_q}$ occurs. $\mathbf{k}_q \equiv \mathbf{k}_{r-1} - \mathbf{k}_{s-1}$, where $\mathbf{k}_0 = (0, 0, 0)$.

(b) $\Delta_{m,p}$ is formed from $D_{m,p}$ by omitting the rows and columns in which $\mathcal{E}_{\mathbf{k}_q}$ and $\mathcal{E}_{-\mathbf{k}_q}$ occur.

(c) $\Delta_{1,m,p}$ is formed from $D_{m,p}$ by omitting the row in which $\mathcal{E}_{\mathbf{k}_q}$ occurs and the column in which $\mathcal{E}_{-\mathbf{k}_q}$ occurs.

(d) $\Delta_{2,m,p}$ is formed from $D_{m,p}$ by omitting the column in which $\mathcal{E}_{\mathbf{k}_q}$ occurs and the row in which $\mathcal{E}_{-\mathbf{k}_q}$ occurs.

The assumption is made (Karle, 1971) that the expected value \mathcal{E} is given by the center of the bounding circle in (3) so that

$$\bar{\mathcal{E}}_{\mathbf{k}_q} = \delta_{m,p} \quad (6)$$

and

$$\mathcal{E}_{\mathbf{k}_q} \simeq \langle \delta_{m,p}(\mathbf{k}_q) \rangle_{m,p}. \quad (7)$$

A variety of simultaneous equations can be generated from (7). The accuracy of such equations is of considerable interest. In particular, with the vast choice of determinants, optimal selection and the development of procedural algorithms that are both accurate and efficient are important considerations. The ensuing mathematical discussions will characterize more specifically the nature of the equations involved and suggest some paths to investigate for their application that may, in time, become fruitful.

The generation of determinants

As (2) indicates, a determinant is formed by selection of the elements for the first column. The first row is formed from the corresponding complex conjugates of the elements in the first column. The subscripts of elements within the remaining part of the determinant are formed by summing the subscripts in the corresponding first row and first column. A relatively large number of large $|\mathcal{E}|$ populating the determinants is desirable, since such a circumstance enhances the accuracy of the information obtained from them. There are various ways to approach this problem, as, for example, described by Taylor, Woolfson & Main (1978) and de Gelder, de Graaff & Schenk (1990). One method we have found useful is to select large $|\mathcal{E}|$ values for the first column of a high-order determinant that is of much higher order than that desired, and then eliminate the smallest magnitudes by crossing out the rows and columns that preserve principal minors. There are other procedures that could be considered, based in part, for example, on the average values of the various rows in the determinants as rows and columns are removed.

Numerous lower-order determinants can be composed as principal minors of a higher-order determinant. Given a determinant of order m , the number of determinants of order n which are principal minors of the original determinant is

$$N_{m,n} = m! / [(m - n)!n!], \tag{8}$$

and the number of those determinants that all contain one specific element, \mathcal{E}_{k_q} , which exists in the original determinant, is

$$N_{m,n,k_q} = (m - 2)! / [(m - n)!(n - 2)!]. \tag{9}$$

As an illustration, consider forming seventh-order determinants from a twelfth-order one. There are 792 determinants of seventh-order formed from the one twelfth-order determinant and each individual element, if distinct in the original determinant, occurs in 252 determinants. A potentially valuable feature of the determinants formed in this way is the fact that the only elements that occur in the determinants are those in the initial determinant and they all occur in a uniform fashion. Of course, if the initial determinant has some duplicate elements, they will continue to exist in the principal minors to a disproportionate extent.

The equations

The phase equations are formed from sets of determinants, such as described above, with the use of (4) and (7). They are

$$|\mathcal{E}_h| \cos \varphi_h \simeq \langle RP \Delta'_{h,s} / \Delta_{h,s} \rangle_s \tag{10}$$

$$|\mathcal{E}_h| \sin \varphi_h \simeq \langle IP \Delta'_{h,s} / \Delta_{h,s} \rangle_s \tag{11}$$

$$\sin \varphi_h \langle RP \Delta'_{h,s} / \Delta_{h,s} \rangle_s \simeq \cos \varphi_h \langle IP \Delta'_{h,s} / \Delta_{h,s} \rangle_s \tag{12}$$

$$|\mathcal{E}_h|^2 \simeq \langle RP \Delta'_{h,s} / \Delta_{h,s} \rangle_s^2 + \langle IP \Delta'_{h,s} / \Delta_{h,s} \rangle_s^2, \tag{13}$$

where s labels a contributor to the average, $RP \Delta'_h = (\Delta'_h + \Delta_h^*)/2$, $IP \Delta'_h = (\Delta'_h - \Delta_h^*)/2i$, and $*$ denotes the complex conjugate. (12) and (13) can be derived from (10) and (11). As will be seen, although their information content is similar, the redundant equations can play a useful role in the way the equations are handled. (12) is a form of the tangent formula and (13), evidently, can compute magnitudes.

Probability measures have been developed that could be relevant to the use of (10)–(13). Following an earlier derivation (Karle, 1971), we have for the probability distribution for φ_{k_q} from the use of the central limit theorem

$$P(\varphi_{k_q}) = [2\pi I_0(\alpha)]^{-1} \exp[\alpha \cos(\varphi_{k_q} - \beta)], \tag{14}$$

where

$$\alpha = \left\{ \left[\sum_{m,p} \kappa_{m,p}(\mathbf{k}_q) \cos \theta_{m,p}(\mathbf{k}_q) \right]^2 + \left[\sum_{m,p} \kappa_{m,p}(\mathbf{k}_q) \sin \theta_{m,p}(\mathbf{k}_q) \right]^2 \right\}^{1/2} \tag{15}$$

$$\kappa_{m,p}(\mathbf{k}_q) = 2|\mathcal{E}_{k_q} \delta_{m,p}(\mathbf{k}_q)| / \sigma_{m,p}^2(\mathbf{k}_q) \tag{16}$$

$$\delta_{m,p}(\mathbf{k}_q) = |\delta_{m,p}(\mathbf{k}_q)| \exp[i\theta_{m,p}(\mathbf{k}_q)] \tag{17}$$

$$\sigma_{m,p}^2(\mathbf{k}_q) \simeq [\mathcal{E}_{000}^{-1} r_{m,p}(\mathbf{k}_q)]^2 \tag{18}$$

and

$$\tan \beta \simeq \frac{\sum_{m,p} |\delta_{m,p}(\mathbf{k}_q)| \sin \theta_{m,p}(\mathbf{k}_q)}{\left[\sum_{m,p} |\delta_{m,p}(\mathbf{k}_q)| \cos \theta_{m,p}(\mathbf{k}_q) \right]}. \tag{19}$$

The variance of φ_{k_q} , given a set of $\delta_{m,p}(\mathbf{k}_q)$ and a set of $\kappa_{m,p}(\mathbf{k}_q)$, is given by

$$V = (\pi^2/3) + [I_0(\alpha)]^{-1} \sum_{n=1}^{\infty} I_{2n}(\alpha) / n^2 - 4[I_0(\alpha)]^{-1} \sum_{n=1}^{\infty} I_{2n+1}(\alpha) / (2n + 1)^2, \tag{20}$$

where the I are Bessel functions of imaginary argument. Except for the new more general definition of α , this formula compares with (3.33) of Karle & Karle [(1966) after a correction of the sign of the last term]. The variance formula in the latter reference refers to the case $m = 3$.

For the case of a centrosymmetric crystal, inequality (3) implies a bound on the real axis with

$$S \mathcal{E}_{k_q} \simeq S \sum_{m,p} \delta_{m,p}(\mathbf{k}_q), \tag{21}$$

where S means sign of. The probability that the sign of

Table 1. Results of computations of phase errors and quasi-normalized structure-factor magnitudes by the use of single determinants of increasing order and (10) and (11) with $s = 1$

Element	$ \mathcal{E} $	Magnitude of error in φ_{calc} (radians)				$ \mathcal{E} _{\text{calc}} (m = 45)$
		$m = 15$	$m = 25$	$m = 35$	$m = 45$	
2,1	3.099	0.340	0.193	0.010	0.009	2.928
3,1	1.250	0.201	0.139	0.030	0.020	1.250
4,1	2.852	0.256	0.100	0.072	0.007	2.804
5,1	2.662	0.050	0.011	0.026	0.006	2.634
6,1	1.649	0.249	0.148	0.027	0.004	1.601
7,1	1.602	0.237	0.088	0.025	0.006	1.606
8,1	1.158	1.522	1.079	0.334	0.059	1.109
9,1	1.684	0.352	0.219	0.057	0.043	1.803
10,1	1.002	0.616	0.488	0.386	0.040	0.985
11,1	1.694	0.244	0.146	0.079	0.011	1.730
12,1	1.105	0.233	0.232	0.206	0.023	1.160
13,1	1.607	0.129	0.043	0.115	0.008	1.633
14,1	1.746	0.294	0.212	0.054	0.015	1.697
15,1	1.074	0.921	0.242	0.029	0.022	1.068

Exact phases and magnitudes, for use in the determinant calculations, were computed from a 50-atom structural model in space group $P1$. The magnitudes of the phase errors are listed for determinants of order 15, 25, 35 and 45, obtained as subdeterminants of an initial determinant of order 50. Structure-factor magnitudes computed from the determinants are listed only for $m = 45$. The elements considered are in the first column of the determinants. The correct values of $|\mathcal{E}|$ are given in the second column and may be compared with those calculated from the determinant of order 45, listed in the last column. The magnitudes of the phase errors show a correlation with the corresponding magnitudes of $|\mathcal{E}|$ for $m = 15$. They evidently become very small as the order of the determinant increases.

$\mathcal{E}_{\mathbf{k}_q}$ is positive is

$$p_+(\mathbf{k}_q) \simeq \frac{1}{2} + \frac{1}{2} \tan h |\mathcal{E}_{\mathbf{k}_q}| \sum_{m,p} \delta_{m,p}(\mathbf{k}_q) / \sigma_{m,p}^2(\mathbf{k}_q). \quad (22)$$

For $m = 3$, (22) corresponds to the probability formula of Cochran & Woolfson (1955).

Some calculations

A few calculations have been made that indicate some of the characteristics of the higher-order determinants that occur in (2) and (4). The determinants were prepared by placing in the first column a selection of elements with indices not exceeding 3 and whose magnitudes were among the larger ones. The low indices for the first column assured that all the elements in the determinant would be easily found in the range of the data at hand and that the determinant would contain many elements of low magnitude. The model structure was based on coordinates for the zervamicin IIA analog (Karle, Sukumar & Balaram, 1986). The first 50 sets of atomic coordinates listed in Table 1 of the latter reference were treated as if they all belonged to the same type of atom and were placed in a unit cell of space group $P1$. Actually, the atoms were C, N and O.

In actual phase determination procedures, it is very important to use determinants that have as high a population of large $|\mathcal{E}|$ values as possible. In general, determinants that have the largest $|\mathcal{E}|$ values are associated with the most accurate phase evaluations.

Table 2. Results of computations of phase errors and quasi-normalized structure-factor magnitudes by the use of single determinants of increasing order

Element	$ \mathcal{E} $	Magnitude of error in φ_{calc} (radians)				$ \mathcal{E} _{\text{calc}} (m = 45)$
		$m = 15$	$m = 25$	$m = 35$	$m = 45$	
1,2	3.099	0.340	0.193	0.010	0.009	2.928
3,2	0.805	0.053	0.309	0.126	0.032	0.826
4,2	0.390	2.886	2.751	0.169	0.039	0.357
5,2	2.852	0.066	0.026	0.037	0.010	2.885
6,2	2.662	0.113	0.005	0.009	0.009	2.571
7,2	1.649	0.243	0.116	0.031	0.003	1.629
8,2	0.544	0.449	0.370	0.241	0.058	0.594
9,2	0.753	1.688	1.360	0.053	0.012	0.655
10,2	0.539	1.419	1.195	0.546	0.048	0.467
11,2	0.563	0.021	0.208	0.007	0.032	0.537
12,2	0.118	1.879	1.583	1.294	0.173	0.175
13,2	0.728	2.835	2.380	0.277	0.001	0.691
14,2	0.185	2.542	2.539	0.924	0.210	0.201
15,2	0.787	0.202	0.003	0.108	0.028	0.786

Similar to Table 1, but concerning elements in the second column of the determinants. Again, it is seen that the large errors for the phases are correlated with low values for $|\mathcal{E}|$ and the lower values for m . For $m = 45$, there are no significant errors either for the calculated phases or magnitudes.

The purposes of the illustrations here, however, are served by the manner in which the determinants were formed.

Table 1 concerns 14 elements in the first column of determinants of the order 15, 25, 35 and 45. Similarly for Table 2, which concerns 14 elements in the second column of the determinants of increasing order. Using (10) and (11), the phases were computed from the selected determinants in which the correct magnitudes and phases were inserted. This illustrates the accuracy of the determinant formulae for these particular determinants. It is apparent from examining the tables that the largest errors in the computed phases are associated with smaller values of $|\mathcal{E}|$. It is also seen that the errors decrease significantly as the order of the determinant increases. The $|\mathcal{E}|_{\text{calc}}$ values from the 45th-order determinant agree well with the correct answers.

Least-squares methods

Least-squares procedures require knowledge of the partial derivatives of the determinants with respect to the unknown phases. The partial derivative of a determinant D with respect to the $\varphi_{\mathbf{k}}$ phase when, in general, similar and different functions of $\varphi_{\mathbf{k}}, f(\varphi_{\mathbf{k}})$ constitute some of the elements of the determinant is

$$\frac{\partial D}{\partial \varphi_{\mathbf{k}}} = \sum_{ij \in \varphi_{\mathbf{k}}} \frac{\partial f_{ij}(\varphi_{\mathbf{k}})}{\partial \varphi_{\mathbf{k}}} M_{ij}, \quad (23)$$

where $ij \in \varphi_{\mathbf{k}}$ indicates that the sum is taken over all ij possessing $f(\varphi_{\mathbf{k}})$ and M_{ij} is the minor of the ij element, i.e. $(-1)^{i+j}$ times the determinant formed by deleting the i th row and j th column.

With the use of (23), the derivatives required for the application of least-squares analysis to (10)–(13) may be obtained. Evidently Δ and Δ' are functions of many of

the φ appearing in the determinant from which Δ and Δ' are formed. The least-squares treatment of highly nonlinear systems of equations such as (10)–(13) would normally have a very small convergence range. The next section is devoted to a discussion of how the convergence range may possibly be extended.

Function modification and rolling

A method for extending the range of convergence for the least-squares minimization technique in nonlinear systems has been described (Karle, 1991). The basic concepts are as follows.

(i) Modification of the character of the defining equations so as to reduce the number of false minima in the minimization function, while preserving the global minima.

(ii) Alteration of the shape of the minimization function from time to time during the least-squares calculation, while keeping the modified character of the defining equations constant. The false minima that are present may thus change in their number and location, while the global minima remain the same.

Operation (i) is called function modification and (ii) is rolling. They both preserve the global minima but alter the minimization function. Operation (i) does so by changing the defining equations and operation (ii) does so in many other ways as a function of time. For example, sets of data and particular equations used for the minimization function may change during the course of the computation. Of course, the form of the defining equations could also be modified from time to time during the least-squares procedure. In some cases, operation (ii) suffices to bring a problem to rapid convergence with the use of the original defining equations and variation of the data sets over time. Examples of the application of concepts (i) and (ii) have been presented (Karle, 1991) in some initial studies of the problem of the direct calculation of atomic positions from known structure-factor magnitudes.

Concluding remarks

An objective for introducing the least-squares technique that involves modification of the defining equations, or alteration of the minimization function, or both, is to avoid the necessity for making many initial starts of the phase determination. There are many instances in which it is not necessary to modify the defining equations, merely alter the minimization function from time to time. This occurs when there are not too many false minima. A

precise statement concerning how many is too many cannot be made. There are too many, however, when attempts to alter the minimization function do not result in ultimate convergence to a global minimum.

There are a number of matters that merit further attention. One concerns the further improvement of the least-squares technique in order to facilitate convergence. Much experimentation is also required to discover optimal ways to use the phase-determining relations. For example, is it preferable to use one large-order determinant or the lower-order determinants that can be prepared from it, or some combination of determinants including low- and high-order ones? In addition, it may be worthwhile to explore the numerous ways of selecting the elements in the first column of the determinants that can generate closely related determinants, which might be valuable when used in combination. The sensitivity of the formulae to errors in the data also needs to be evaluated.

As a final remark, we note that the mathematical formulae currently in general use for phase determination are rather special cases of the mathematics discussed here. Studies of the general formulation can be approached in their algebraic context or with the application of the maximum determinant rule, or some combination.

My thanks to Mr Stephen Brenner for writing the appropriate programs and making the computations. This research was supported in part by the Office of Naval Research and in part by Public Health Service Grant GM30902.

References

- COCHRAN, W. & WOOLFSON, M. M. (1955). *Acta Cryst.* **8**, 1–12.
 GELDER, R. DE, DE GRAAFF, R. A. G. & SCHENK, H. (1990). *Acta Cryst.* **A46**, 688–692.
 GELDER, R. DE, DE GRAAFF, R. A. G. & SCHENK, H. (1993). *Acta Cryst.* **A49**, 287–293.
 GOEDKOOP, J. A. (1950). *Acta Cryst.* **3**, 374–378.
 HAUPTMAN, H. & KARLE, J. (1950). *Phys. Rev.* **80**, 244–248.
 HEINERMAN, J. J. L., KROON, J. & KRABBENDAM, H. (1979). *Acta Cryst.* **A35**, 105–107.
 KARLE, J. (1971). *Acta Cryst.* **B27**, 2063–2065.
 KARLE, J. (1978). *Proc. Natl Acad. Sci. USA*, **75**, 2545–2548.
 KARLE, J. (1991). *Proc. Natl Acad. Sci. USA*, **88**, 10099–10103.
 KARLE, J. & HAUPTMAN, H. (1950). *Acta Cryst.* **3**, 181–187.
 KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849–859.
 KARLE, I. L., SUKUMAR, M. & BALARAM, P. (1986). *Proc. Natl Acad. Sci. USA*, **83**, 9284–9288.
 TAYLOR, D. J., WOOLFSON, M. M. & MAIN, P. (1978). *Acta Cryst.* **A34**, 870–883.
 TSOUCARIS, G. (1970). *Acta Cryst.* **A26**, 492–499.