

A Multivariate Joint Probability Distribution for Phase Determination

BY E. E. CASTELLANO,* A. D. PODJARNY AND J. NAVAZA

Departamento de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata, República Argentina

(Received 28 November 1972; accepted 20 April 1973)

A multivariate conditional joint probability distribution of a set of K normalized structure factors has been developed using a novel approach. The covariance matrix of the distribution is calculated for all the space groups in terms of linear combinations of specific unitary structure factors. It is shown that if any number of off-diagonal elements are set arbitrarily to zero, an approximation is obtained to the covariance matrix which corresponds to a particular set of *a priori* conditions. The importance of this result in practical phase-determining methods is pointed out. Group theory is used to obtain results valid for all space groups. The multivariate distribution is used to calculate more general versions of the Cochran and Woolfson sign probability and the Karle and Hauptman tangent formulae.

Introduction

The joint probability distribution of a set of K quasi-normalized structure factors was first introduced by Hauptman & Karle (1953) and developed further by several authors. In particular, Klug (1958) employed an original manipulative technique which summarizes all previous results. Although, as pointed out by Karle & Hauptman (1959), Klug's formalism is mathematically equivalent to the previous one developed by these authors, it has the advantage that the joint probability is explicitly written in the form of a strictly asymptotic series in powers of $N^{-1/2}$ for the case of N equal atoms in the unit cell. The first term of this series corresponds to a K -dimensional Laplace–Gauss distribution law. From Klug's work it follows that neglecting terms of order higher than the first is a good approximation if:

(a) The number of atoms in the asymmetric unit is big.

(b) The moments of order higher than the second are very small. This implies that the distribution of the E 's is normal, which is equivalent to assuming the atoms to be randomly distributed throughout the asymmetric unit.

The conditions (a) and (b), taken simultaneously, are equivalent to assuming the Central Limit Theorem to hold.

The covariance matrix for the Gauss–Laplace distribution law was given by Klug in terms of the second-order cumulants (mixed moments). He further stated that in most practical applications its off-diagonal elements are zero. However, it turns out that when some specific *a priori* information is available, that is, when the concept of *conditional* joint probability is used, the elements of the covariance matrix can be proved not to be necessarily zero, and their values, fixed by the particular conditions introduced into the problem, may be calculated.

This was first done by Tsoucaris (1970) who started directly by writing down a Laplace–Gauss distribution law for a set of normalized structure factors (E 's), calculating then the covariance matrix elements on the basis of certain hypotheses, and assuming that a specific set of unitary structure factors (U 's) was given as conditions to the problem.

The elements of the covariance matrix of the distribution are the second-order mixed moments defined by

$$\langle E_i E_j^* \rangle = \int \int_{-\infty}^{\infty} \dots \int f(E_1, \dots, E_K) E_i E_j^* d^K E, \quad (1)$$

where $f(E_1, \dots, E_K)$ is the conditional joint probability distribution function of the set of K structure factors E_1, \dots, E_K . Since f is not known (it is precisely the function we are seeking), the moments cannot be calculated from (1) directly. The way we overcome this problem is by considering the atomic coordinates \mathbf{r}_j as the independent variables, because some *a priori* distribution of them can be assumed on physical grounds.

Any function of several structure factors can be considered as a function of the atomic coordinates \mathbf{r}_j and reciprocal vectors \mathbf{h} . In X-ray diffraction experiments the reciprocal vectors \mathbf{h} are known parameters while the \mathbf{r} 's are unknowns. For this reason, in statistical interpretations of the phase problem it is convenient to consider the \mathbf{r} 's as random variables giving rise to certain distributions of the structure-factor functions. But this, in general, will not lead to a distribution mathematically equivalent to that obtained by assuming fixed atomic coordinates and the reciprocal vectors \mathbf{h} to be the random variables. An exception to this is the case in which the function of structure factors is such that it depends symmetrically on the \mathbf{r} 's and the reciprocal vector \mathbf{h} , as it is the case for one structure factor (Hauptman & Karle, 1953).

This consideration will be expanded in detail in § 1, taking space group $P\bar{1}$ as an example, since it forms the basis for the main conclusions of the present work.

In § 2, the way of calculating the elements of the covariance matrix for all the space groups is developed

* Member of the Consejo Nacional de Investigaciones Científicas y Técnicas.

with the aid of group theory. Specific examples for some particular space groups are given.

The concept of similar reflexions is introduced in § 3 to discuss the mutual dependence of certain types of second-order mixed moments.

Finally, in § 4, the conditional joint probability distribution function is used to calculate improved versions of the Cochran & Woolfson sign probability and the Karle & Hauptman tangent formulae.

Notation

N = number of atoms in the unit cell.

n = symmetry number.

$t = N/n$, number of atoms in the asymmetric unit.

\mathbf{r} = vector of coordinates of the j th atom.

\mathbf{h} = reciprocal vector [row matrix (h, k, l)].

$\hat{G}_\alpha = (G_\alpha/\mathbf{g}_\alpha)$, each of the symmetry elements of a given space group expressed as an affine transformation where: G_α is the three-by-three matrix of the point group and \mathbf{g}_α defines the translation vector of the transformation.

$\hat{G}_1 = [I/(u_1, u_2, u_3)]$ is the identity transformation, where

I is the unitary matrix and u_i are arbitrary integers.

$E(\mathbf{h}_j) = E_j$ is a quasi-normalized structure factor = $|E_j| \exp(i\Phi_j)$.

$U(\mathbf{h}) = N^{-1/2} E(\mathbf{h})$

$$C^j(\mathbf{h}_k) = C_k^j = \sum_{\alpha=1}^n \exp(2\pi i \mathbf{h}_k \hat{G}_\alpha \mathbf{r}_j).$$

Note that

$$E(\mathbf{h}_k) = N^{-1/2} \sum_{j=1}^t \sum_{\alpha=1}^n \exp(2\pi i \mathbf{h}_k \hat{G}_\alpha \mathbf{r}_j) = N^{-1/2} \sum_{j=1}^t C_k^j.$$

1. Calculus of the second-order mixed moments.

Space group $P\bar{1}$ as an example

The random character of normalized structure factors, E 's, arises from the fact that they are continuous and single-valued functions of the atom coordinates \mathbf{r}_j , which are assumed to be randomly and uniformly distributed throughout the asymmetric unit.

The cumulative distribution function H of a set of K normalized structure factors may then be written as a $3N/n$ -dimensional integral:

$$H(E_1, \dots, E_K) = \iiint_{\Omega} \dots \int f_r d^{3N/n} \mathbf{r} \\ = \iiint_{\text{ASYM.}} \dots \int_{\text{UNIT}} \theta(E_1 - \hat{E}_1) \dots \theta(E_K - \hat{E}_K) f_r d^{3N/n} \mathbf{r} \quad (1.1)$$

where θ is the Heaviside step function, Ω is a region of the asymmetric unit such that for every \mathbf{k} , $\hat{E}_k(\mathbf{r}) \leq E_k$ and $f_r = f(\mathbf{r}_1, \dots, \mathbf{r}_t)$ is the joint probability distribution of the atomic positions. The assumption of atoms randomly and uniformly distributed implies this function to be unity. \hat{E}_k is considered as an explicit function of

the \mathbf{r}_j 's to distinguish it from the fixed value of $E(\mathbf{h}_k)$ denoted here simply as E_k .

Taking all the partial derivatives in (1.1) we obtain

$$f_E(E_1, \dots, E_K) = \partial^K H / \partial E_1 \dots \partial E_K \\ = \iiint_{\text{ASYM.}} \dots \int_{\text{UNIT}} \delta(E_1 - \hat{E}_1) \dots \delta(E_K - \hat{E}_K) f_r d^{3N/n} \mathbf{r},$$

where f_E is the joint probability distribution function of E_1, \dots, E_K and δ is the Dirac delta function.

The mean value of any function of the E 's, L say, is

$$\langle L \rangle_E = \iiint_{-\infty}^{\infty} \dots \int L(E_1, \dots, E_K) \\ \times \left\{ \iiint_{\text{ASYM.}} \dots \int_{\text{UNIT}} \delta(E_1 - \hat{E}_1) \dots \delta(E_K - \hat{E}_K) f_r d^{3N/n} \mathbf{r} \right\} d^K E.$$

Changing the order of integration we finally obtain

$$\langle L \rangle_E = \iiint_{\text{ASYM.}} \dots \int_{\text{UNIT}} f_r L(E_1, \dots, E_K) d^{3N/n} \mathbf{r} = \langle L \rangle_r. \quad (1.2)$$

We shall now use this formula to calculate the second-order mixed moments $m_{\dots, i, \dots, j, \dots}$ in space group $P\bar{1}$. We shall later generalize the method to all the space groups.

We have already pointed out that the covariance element σ_{pq} is the second-order mixed moment $m_{\dots, i, p, \dots, i, q, \dots}$ (Klug, 1958). So, by definition:

$$\sigma_{pq} = \langle E_p E_q^* \rangle = \iiint_{\Omega} \dots \int f_E E_p E_q^* d^K E \\ = \int_{\text{ASYM.}} \dots \int_{\text{UNIT}} f_r E_p E_q^* d^{3N/n} \mathbf{r}.$$

If we assume all equal atoms, $E_p(\mathbf{r}) = 2(N)^{-1/2} \times \sum_i \cos(2\pi \mathbf{h}_i \mathbf{r}_p)$. Then by (1.2)

$$\sigma_{pq} = 4(N)^{-1} \iiint \dots \int \sum_{ij} \cos(2\pi \mathbf{h}_p \mathbf{r}_i) \cos(2\pi \mathbf{h}_q \mathbf{r}_j) d^{3N/n} \mathbf{r}.$$

Separating the terms with $i=j$

$$\langle E_p E_q^* \rangle = \iiint \dots \int f_r (U(\mathbf{h}_p - \mathbf{h}_q) + U(\mathbf{h}_p + \mathbf{h}_q)) d^{3N/n} \mathbf{r} \\ + 4(N)^{-1} \iiint \dots \int f_r \left[\sum_{i \neq j} \cos(2\pi \mathbf{h}_p \mathbf{r}_i) \right. \\ \left. \times \cos(2\pi \mathbf{h}_q \mathbf{r}_j) \right] d^{3N/n} \mathbf{r}. \quad (1.3)$$

The function between square brackets in the second term of the right-hand side averages zero because of the cosine periodicity. We assume that the introduction of *a priori* conditions does not alter essentially the constant character of f_r in the sense that the term

$$\iiint \dots \int f_r \sum_{i \neq j} \cos(2\pi \mathbf{h}_p \mathbf{r}_i) \cos(2\pi \mathbf{h}_q \mathbf{r}_j) d^{3N/n} \mathbf{r}, \quad (1.3')$$

always vanishes.

The first integral, however, can take on different values according to the conditions imposed on the

problem. If, for instance, $U(\mathbf{h}_p - \mathbf{h}_q)$ is known [that is, if f_r is such that $\hat{U}(\mathbf{h}_p - \mathbf{h}_q)$ always equals $U(\mathbf{h}_p - \mathbf{h}_q)$], it is a constant that can be taken out of the integral and the mixed moment resulting is precisely $U(\mathbf{h}_p - \mathbf{h}_q)$. More generally, the value of the mixed moment depends upon whether $U(\mathbf{h}_p - \mathbf{h}_q)$ and/or $U(\mathbf{h}_p + \mathbf{h}_q)$ are given or not. There exist four possibilities:

- Neither $U(\mathbf{h}_p - \mathbf{h}_q)$ nor $U(\mathbf{h}_p + \mathbf{h}_q)$ is known. All integrals vanish because of the cosine periodicity.
- $U(\mathbf{h}_p + \mathbf{h}_q)$ is known. All the terms average zero, except the second one in the single sum. The result is $U(\mathbf{h}_p + \mathbf{h}_q)$.
- $U(\mathbf{h}_p - \mathbf{h}_q)$ is known. All the terms average zero, except the first one in the single sum. The result is $U(\mathbf{h}_p - \mathbf{h}_q)$.
- Both $U(\mathbf{h}_p - \mathbf{h}_q)$ and $U(\mathbf{h}_p + \mathbf{h}_q)$ are known. Then, only the double sum averages zero and the final result is $U(\mathbf{h}_p - \mathbf{h}_q) + U(\mathbf{h}_p + \mathbf{h}_q)$.

It is interesting to compare this result with the one of Hughes (1953). This author obtained what are essentially the second-order mixed moments, by averaging in reciprocal space under the condition that certain specific vectors are kept fixed during the procedure. He assumed the vector $\mathbf{h}_p + \mathbf{h}_q$ to be fixed, obtaining a result coincident with our result (b). Had he fixed the vector $\mathbf{h}_p - \mathbf{h}_q$, or none, he would have obtained either result (c) or result (a) respectively. But his formalism would not allow him to obtain our result (d) since fixing $\mathbf{h}_p - \mathbf{h}_q$ and $\mathbf{h}_p + \mathbf{h}_q$ simultaneously fixes completely \mathbf{h}_p and \mathbf{h}_q .

This example shows that, in the way they have been carried out, averages in direct and reciprocal space are not equivalent. The physical hypothesis implied by the fact that a specific vector has been kept fixed during the averaging procedure is by no means apparent. In fact, our result (d) shows those hypotheses to be too restrictive in the particular case we have just analysed.

The second-order mixed moments (correlation coefficients) are calculated by Tsoucaris with the aid of Sayre's equation

$$\langle E_p E_q^* \rangle = \langle E(\mathbf{L} + \mathbf{H}_p) E^*(\mathbf{L} + \mathbf{H}_q) \rangle_{\mathbf{L}} = U(\mathbf{H}_p - \mathbf{H}_q). \quad (1.4)$$

This result coincides with (1.3) in space group $P1$ only, if $U(\mathbf{H}_p - \mathbf{H}_q)$ is supposed to be known. Equation (1.4) is symmetry independent, that is, it holds independently of the space group of the structure. In a personal communication we received while the present paper was in proof (Tsoucaris, 1973), it is claimed that full symmetry can be taken into account if, and only if, all numbers of each symmetry-related set are included in the set of E values whose probability law is to be found.

2. Calculus of the second-order mixed moments in all the space groups

We have just discussed the form of the covariance matrix for the particular case of space group $P\bar{1}$. We

shall now generalize this result to all space groups by a straightforward application of group theory.

The covariance between the normalized structure factors E_i and E_j can be written

$$\sigma_{ij} = \langle E_i E_j^* \rangle = (N)^{-1} \left\langle \sum_{l=1}^{N/n} C_l^i (C_l^j)^* \right\rangle_{\mathbf{r}} + (N)^{-1} \left\langle \sum_{l \neq m} C_l^i (C_m^j)^* \right\rangle_{\mathbf{r}}. \quad (2.1)$$

The second term averages zero because of the exponential periodicity while the first one takes different forms for the different space groups. To show this, we shall first prove that the product $C_l^i (C_l^j)^*$ may be written in the form

$$C_l^i (C_l^j)^* = \sum_{\alpha=1}^n \exp(2\pi i \mathbf{h}_i \mathbf{g}_\alpha) C^l(\mathbf{h}_i G_\alpha - \mathbf{h}_j). \quad (2.2)$$

From its definition we can write

$$C_l^i (C_l^j)^* = \sum_{\alpha, \beta} \exp(2\pi i \mathbf{h}_i \hat{G}_\alpha \mathbf{r}_l) \exp(-2\pi i \mathbf{h}_j \hat{G}_\beta \mathbf{r}_l).$$

Applying the rearrangement theorem in the sum over α we can write

$$C_l^i (C_l^j)^* = \sum_{\alpha, \beta} \exp(2\pi i (\mathbf{h}_i \hat{G}_\alpha - \mathbf{h}_j) \hat{G}_\beta \mathbf{r}_l)$$

which reduces to (2.2).

Replacing this relation into (2.1) we obtain

$$\sigma_{ij} = \sum_{\alpha} \exp(2\pi i \mathbf{h}_i \mathbf{g}_\alpha) \langle (N)^{-1} \sum_{l=1}^{N/n} C^l(\mathbf{h}_i G_\alpha - \mathbf{h}_j) \rangle_{\mathbf{r}}.$$

The expression averaged over \mathbf{r} is the unitary structure factor $U(\mathbf{h}_i G_\alpha - \mathbf{h}_j)$ so that

$$\sigma_{ij} = \sum_{\alpha} \exp(2\pi i \mathbf{h}_i \mathbf{g}_\alpha) \langle U(\mathbf{h}_i G_\alpha - \mathbf{h}_j) \rangle_{\mathbf{r}}. \quad (2.3)$$

In this formula the mean value will be conditioned by the *a priori* information given to the problem. It will take on different values according to which of the $U(\mathbf{h}_i G_\alpha - \mathbf{h}_j)$ are given as a *priori* information.

As $U(\mathbf{h}_i G_\alpha - \mathbf{h}_j)$ depends only on the matrices G_α of the point group and $\exp(2\pi i \mathbf{h}_i \mathbf{g}_\alpha)$ on the translation vectors, we can write (2.3) in the form

$$\sigma_{ij} = \sum_s \langle U(\mathbf{h}_i G_s - \mathbf{h}_j) \rangle_{\mathbf{r}} \sum_{\mathbf{t}} \exp(2\pi i \mathbf{h}_i \mathbf{g}_{s\mathbf{t}}), \quad (2.4)$$

where G_s are the different matrices of the space group, and $\mathbf{g}_{s\mathbf{t}}$ are the different translations of each matrix.

The sum over the translation vectors can be readily calculated with the aid of *International Tables for X-ray Crystallography* (1952). It can be seen that this sum is the same as accounts for systematic extinctions.

As examples we give below the explicit form of the covariance matrix elements for space groups $P2_1$, $B2$, $P2_1/b$, $P2_12_12_1$.

Space group $P2_1$, unique axis c

$$\sigma_{ij} = \langle U(\mathbf{h}_i - \mathbf{h}_j) \rangle_{\mathbf{r}} + (-1)^{l_i} \langle U(\bar{h}_i + \bar{h}_j, \bar{k}_i + \bar{k}_j, l_i, \bar{l}_j) \rangle_{\mathbf{r}}.$$

Space group B_2

$$\sigma_{ij} = [1 + (-1)^{h_j + l_j}] \langle U(\mathbf{h}_i - \mathbf{h}_j) \rangle_{\mathbf{r}} + \langle U(\bar{h}_i + \bar{h}_j, \bar{k}_i + \bar{k}_j, l_i + \bar{l}_j) \rangle_{\mathbf{r}}.$$

Space group $P2_1/b$, unique axis c

$$\sigma_{ij} = \langle U(\mathbf{h}_i - \mathbf{h}_j) + U(\mathbf{h}_i + \mathbf{h}_j) + (-1)^{k_i + l_i} [U(\bar{h}_i + \bar{h}_j, \bar{k}_i + \bar{k}_j, l_i + \bar{l}_j) + U(\bar{h}_i + \bar{h}_j, k_i + \bar{k}_j, \bar{l}_i + \bar{l}_j)] \rangle_{\mathbf{r}}.$$

Space group $P2_12_12_1$

$$\sigma_{ij} = \langle U(\mathbf{h}_i - \mathbf{h}_j) + (-1)^{h_i + k_j} U(\bar{h}_i + \bar{h}_j, \bar{k}_i + \bar{k}_j, l_i + \bar{l}_j) + (-1)^{h_j + k_i} U(\bar{h}_i + \bar{h}_j, \bar{k}_i + \bar{k}_j, \bar{l}_i + \bar{l}_j) + (-1)^{k_i + j_i} \times U(\bar{h}_i + \bar{h}_j, k_i + \bar{k}_j, \bar{l}_i + \bar{l}_j) \rangle_{\mathbf{r}}.$$

We wish to thank the referee for drawing our attention to the fact that the preceding formalism was first used by Goedkoop (1950) in connexion with the introduction of symmetry into Karle & Hauptman determinantal inequalities. It is also shown in that paper that the matrix given by (2.3) or (2.4) is positive definite.

3. Similar reflexions and correlation factors

The reflexions of the set $\{G\}$

$$\{E(\mathbf{h}_i G_\alpha), \alpha = 1, \dots, n\}$$

will be called similar to one another to distinguish them from those of the set of equivalent reflexions defined by

$$\{E(\mathbf{h}_i L_\beta)\} = \{L\}$$

where L_β are the symmetry matrices of the Laue group.

The property of the similar reflexions which is relevant to later discussion is that they always equal one another except for a phase factor independent of the atomic coordinates, as can be readily seen starting from

$$E_i = (N)^{-1/2} \sum_{i=1}^{N/n} \sum_{\alpha=1}^n \exp(2\pi i \mathbf{h}_i \hat{G}_\alpha \mathbf{r}_j)$$

and applying the rearrangement theorem

$$E_i = (N)^{-1/2} \sum_j \sum_\alpha \exp(2\pi i \mathbf{h}_i \hat{G}_\beta \hat{G}_\alpha \mathbf{r}_j) = \exp(2\pi i \mathbf{h}_i \mathbf{g}_\beta) E(\mathbf{h}_i G_\beta).$$

The difference between equivalent and similar reflexions arises in the case in which non-centrosymmetrical space groups are considered. In this case the sets $\{G\}$ and $\{L\}$ differ in that $\{L\}$ contains the inversion matrix (accounting for Friedel's law) giving rise to a complex conjugation operation which is never present in similar reflexions.

This can be seen by noting that if $E(\mathbf{h}_i L_\beta)$ is a reflexion equivalent but not similar to $E(\mathbf{h}_i)$ then $L_\beta = -IG_\alpha$ for some G_α , and

$$E(\mathbf{h}_i L_\beta) = E^*(\mathbf{h}_i G_\alpha) = \exp(2\pi i \mathbf{h}_i \mathbf{g}_\alpha) E^*(\mathbf{h}_i).$$

We are now in the position to show that the elements of a conventional Karle & Hauptman matrix are terms of covariance elements rather than covariances, since the latter would imply the contradiction that the second-order mixed moments between some reflexion and each of two other similar reflexions could have different absolute values. In $P\bar{1}$, for example, the covariance between $E(\mathbf{h})$ and $E(\mathbf{h}')$ is $U(\mathbf{h} - \mathbf{h}') + U(\mathbf{h} + \mathbf{h}')$ and coincides with the covariance between $E(\mathbf{h})$ and $E(-\mathbf{h}')$, $E(\mathbf{h}')$ and $E(-\mathbf{h}')$ being similar reflexions in this space group. If the Karle & Hauptman matrix had been considered, the same covariances would have been $U(\mathbf{h} - \mathbf{h}')$ and $U(\mathbf{h} + \mathbf{h}')$ respectively, which have in general different values.

An important property of similar reflexions is that they can be arbitrarily replaced by one another in the distribution function. To show this, we recall the result, formally identical with that of Tsoucaris,

$$E^*(\sigma)^{-1} \mathbf{E} = N \begin{pmatrix} D_K & & & \\ & -\Delta_{K+1} & & \\ & & & \end{pmatrix} / D_K$$

where D_K is the determinant of the covariance matrix and Δ_{K+1} is defined by

$$\Delta_{K+1} = N \begin{vmatrix} N/n & E_1^* & \dots & E_j^* & \dots & E_K^* \\ E_1 & \langle E_1^* E_1 \rangle & \dots & \langle E_j^* E_1 \rangle & \dots & \langle E_K^* E_1 \rangle \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ E_j & \langle E_1^* E_j \rangle & \dots & \langle E_j^* E_j \rangle & \dots & \langle E_K^* E_j \rangle \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ E_K & \langle E_1^* E_K \rangle & \dots & \langle E_j^* E_K \rangle & \dots & \langle E_K^* E_K \rangle \end{vmatrix}.$$

This determinant is identical with the Δ_{m+1} of Tsoucaris in space group $P1$. Multiplying the j th column by the factor $\exp(2\pi i \mathbf{h}_j \mathbf{g}_\alpha)$ and the j th row by $\exp(-2\pi i \mathbf{h}_j \mathbf{g}_\alpha)$ we get

$$\Delta_{K+1} = \frac{1}{N} \begin{vmatrix} N/n & \dots & \exp(2\pi i \mathbf{h}_j \mathbf{g}_\alpha) E_j^* & \dots & E_K^* \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ \exp(-2\pi i \mathbf{h}_j \mathbf{g}_\alpha) E_j & \dots & \langle E_j^* E_j \rangle & \dots & \exp(-2\pi i \mathbf{h}_j \mathbf{g}_\alpha) \langle E_K^* E_j \rangle \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ E_K & \dots & \exp(2\pi i \mathbf{h}_j \mathbf{g}_\alpha) \langle E_j^* E_K \rangle & \dots & \langle E_K^* E_K \rangle \end{vmatrix}.$$

which is formed with a reflexion similar to $E(\mathbf{h}_j)$ replacing $E(\mathbf{h}_j)$. A similar reasoning applies to D_K .

We note that in the preceding proof the fact that \mathbf{g}_a is the translation of a symmetry operation has not been used. In fact the result holds for every arbitrary translation. This implies that the quadratic (hermitic) form

$$\mathbf{E}(\sigma)^{-1}\mathbf{E}$$

is a structure invariant, that is, does not depend on the particular origin selected.

4. Derivation of probability formulae for centrosymmetrical and non-centrosymmetrical space groups

A manipulation of the multivariate distribution function can demonstrate its relationship to the Cochran & Woolfson (1955) sign probability and to the Karle & Hauptman tangent formula. This can be done without assuming statistical independence among the triple products involved. It only requires the formulation of the problem with some specific *a priori* information. We have already pointed out that Tsoucaris interpretation of Sayre's equation as the correlation coefficient between two normalized structure factors leads to the conclusion, in space groups other than $P1$, that similar reflexions can have different estimates of the correlation coefficient. The same applies to an equivalent interpretation of the Karle & Hauptman tangent formula. This situation does not appear in the derivation used in the present paper.

Let us choose the *a priori* conditions to be such that the covariance matrix of the distribution law is estimated by

$$U = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1J} & \cdots & \sigma_{1K} \\ \sigma_{21} & \sigma_{22} & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \sigma_{J1} & 0 & \cdots & \sigma_{JJ} & \cdots & 0 \\ \sigma_{K1} & 0 & \cdots & 0 & \cdots & \sigma_{KK} \end{pmatrix} \quad (4.1)$$

It must be realized that the appearance of null off-diagonal elements in the matrix (4.1) makes its character not necessarily positive definite. This means that in this case U is not strictly a covariance matrix, but must rather be interpreted as an approximation to the true covariance matrix. Rollett & Townsend (1973) have pointed out that the use of a matrix with missing elements is related to analysis of the electron density corresponding to a Fourier summation lacking some terms. In suitable cases such series can be expected to display the main features of the true electron density, but they need not be non-negative.

We shall now proceed to expand the quadratic form

$$\mathbf{E}^*U^{-1}\mathbf{E}$$

for the particular covariance matrix given by (4.1). We shall start from the relation

$$\mathbf{E}^*U^{-1}\mathbf{E} = N \frac{D_K}{n} \frac{-\Delta_{K+1}}{D_K} \quad (4.2)$$

with Δ_{K+1} and D_K already defined in § 3. Only the terms independent of Φ_1 , where Φ_1 is the phase of E , the first component of \mathbf{E} , will be taken into account in expanding (4.2), since all the others will cancel out when we impose normalization conditions.

In the Appendix it is shown that

$$\mathbf{E}^*U^{-1}\mathbf{E} = \frac{-2|E_1| \sum_{j=2}^K \frac{|E_j| |\sigma_{1j}|}{\sigma_{jj}} \cos(\Phi_1 - \Phi_j - \Phi_{1j})}{\sigma_{11} - \sum_{j=2}^K \frac{|\sigma_{1j}|}{\sigma_{jj}}} \quad (4.3)$$

In the centrosymmetrical case we can write

$$\cos(\varphi_1 - \varphi_j - \varphi_{1j}) = \text{sg}(E_1) \text{sg}(E_j) \text{sg}(E_{1j}).$$

With this relation, plus the normalizing condition

$$p^+(E_1) + p^-(E_1) = 1$$

we obtain, following Cochran & Woolfson (1955)

$$p^+(E_1) = \frac{1}{2} + \frac{1}{2} \tanh \left\{ \frac{|E_1| \sum_{j=2}^K \frac{E_j \sigma_{1j}}{\sigma_{1j}}}{\sigma_{11} - \sum_{j=2}^K \frac{|\sigma_{1j}|^2}{\sigma_{1j}}} \right\} \quad (4.4)$$

Equation (4.4) is a modified version of the Cochran & Woolfson hyperbolic tangent formula. Note that the mixed moments $\langle E_i E_j^* \rangle$ as given by (2.4) do appear in the original version of the formula, when not only the independent but all the reflexions are used to form the triple-product interactions. In the original derivation, however, it is assumed that all the 'interactions' are statistically independent. This approximation may not be very realistic, especially when similar reflexions are considered. In our present derivation it is only assumed that a particular set of *a priori* conditions is given. The statistical independence among triple products is automatically taken into account. (4.4) is expected, for the above mentioned reasons, to give more reliable probability indications.

In the non-centrosymmetrical case, recalling that

$$p(\Phi_1) = |E_1| p[|E_1| \exp(i\Phi_1)]$$

and using formula (4.3), we see that the expression for $p(\Phi_1)$ is formally identical with (3.21) of Karle & Karle (1966). Then, following these authors, we end with the tangent formula

$$\tan \Phi_1 = \frac{\left[\left(\sum_{j=2}^K \frac{|E_j| |\sigma_{1j}|}{\sigma_{jj}} \sin(\Phi_j + \Phi_{1j}) \right) \right]}{\left[\sum_{j=2}^K \frac{|E_j| |\sigma_{1j}|}{\sigma_{jj}} \cos(\Phi_j + \Phi_{1j}) \right]} \quad (4.5)$$

In space group $P1$, (4.5) differs from the original version only in the weighting factor σ_{jj} . Of course, in higher-symmetry space groups, σ_{ij} is no longer given by $U(\mathbf{h}_i - \mathbf{h}_j)$.

We recall that our expression (4.3) is not an approximate one in the sense that the statistical dependence among the different 'addition pairs' has not been neglected.

We can still find a more general form of the tangent formula if no restriction is imposed on the amount of *a priori* information. This implies, however, working with the inverse of the covariance matrix defined as $(\sigma^{-1})_{ij} = \sigma^{ij}$; we have:

$$\mathbf{E}^* \sigma^{-1} \mathbf{E} = E_i^* \sum_{j \neq 1} \sigma^{ij} E_j + E_1 \sum_{j \neq 1} \sigma^{ji} E_j^* + C$$

where C is a constant independent of Φ_1 .

Noting that

$$(\sigma^{-1})_{ji} = \sigma^{ji} = (\sigma^{ij})^*$$

we get

$$\mathbf{E}^* \sigma^{-1} \mathbf{E} = 2|E_1| \left[\sum_{j=2}^K |\sigma^{1j}| |E_j| \cos(\Phi_1 - \Phi_j - \Phi^{1j}) \right] + C$$

from which it follows

$$\tan \Phi_1 = \frac{\left[\sum_{j=2}^K |\sigma^{1j}| |E_j| \sin(\Phi_j - \Phi^{1j}) \right]}{\left[\sum_{j=2}^K |\sigma^{1j}| |E_j| \cos(\Phi_j - \Phi^{1j}) \right]}. \quad (4.6)$$

This 'generalized tangent formula' bears no resemblance to the one obtained by Karle (1971). Again the present one has been obtained in a more exact fashion, the assumption $E_h \simeq \langle \delta_{m,p}(\mathbf{h}) \rangle$ [formula (11) of Karle (1971)] being not required.

Formula (4.6) can concentrate a great amount of information and may perhaps prove to be useful in cases where other phasing procedures fail. It has the computational disadvantage of implying matrix inversions but this, however, is not a fundamental objection if modern computing facilities are available.

Concluding remarks

A method for calculating the conditional joint probability distribution of a set of normalized structure factors has been discussed. It has been shown to contain, to a third-order approximation, the formulae which have proved to be the most successful in crystal-structure determination. The success of these formulae lies, to a great extent, in the way information can be introduced in a step-by-step procedure. The more powerful higher-order formulae of the Karle & Hauptman-Tsoucaris type lack this facility for handling the information because, in general, too many structure factors are needed to be known in each step. In the present approach this difficulty can be partially overcome by the fact that the covariance matrix can be approximated

by setting some of its elements (or terms of its elements) to zero, as discussed in the text.

This approach offers also the possibility of introducing in the calculus of the covariances, information concerning the physical fact that real atoms cannot adopt any random configuration in the asymmetric unit since any two of them cannot be closer than a known fixed distance. Since the single fact that the positive character of the electron-density function gives rise to most of the useful direct methods known at present, it is interesting to speculate on the possibility that the inclusion of the set of inequalities of the form $|\mathbf{r}_i - \mathbf{r}_j| \geq b$, where b is the average interatomic bond distance, may lead to formulae powerful enough for attacking bigger structures. Investigation on these lines is in progress.

APPENDIX

Calculus of the positive definite Hermitian form $\mathbf{E}^* U^{-1} \mathbf{E}$ for the U matrix defined by (4.1)

Expanding the Δ_{K+1} determinant along the first two rows we get

$$\begin{aligned} \Delta_{K+1} = & \frac{1}{N} \sum_{j=2}^K \left\{ \frac{N/n E_j^*}{E_1 \sigma_{1j}^*} (-1)^{(1+2)+[1+(j+1)]} (-1)^j \prod_{l=2}^K \frac{\sigma_{ll}}{\sigma_{jj}} \sigma_{jl} \right. \\ & \left. + \frac{E_1^* E_j^*}{\sigma_{11} \sigma_{1j}^*} (-1)^{(1+2)+[2+(j+1)]} (-1)^j \prod_{l=2}^K \frac{\sigma_{ll}}{\sigma_{jj}} E_j \right\} \\ & + \text{terms independent of } \Phi_1. \end{aligned}$$

Noting that $\sigma_{ji} = \sigma_{ij}^*$ we finally find

$$\begin{aligned} \Delta_{K+1} = & 2 \frac{|E_1|}{N} \prod_{j=2}^K \sigma_{ll} \sum_{j=2}^K \frac{E_j \sigma_{1j}}{\sigma_{jj}} \cos(\Phi_1 - \Phi_j - \Phi_{1j}) \\ & + \text{terms independent of } \Phi_1. \end{aligned}$$

Expanding the determinant of the covariance matrix along the first row

$$\begin{aligned} D_K = & \prod_{l=1}^K \sigma_{ll} + \sum_{j=2}^K \sigma_{1j} (-1)^{j+1} (-1)^j \sigma_{j1} \prod_{l=2}^K \frac{\sigma_{ll}}{\sigma_{jj}} \\ D_K = & \prod_{l=2}^K \sigma_{ll} \left\{ \sigma_{11} - \sum_{j=2}^K \frac{|\sigma_{1j}|^2}{\sigma_{jj}} \right\}. \end{aligned}$$

Finally, replacing the values of Δ_{K+1} and D_K into the relation

$$N \frac{D_K}{D_K} = - N \frac{\Delta_{K+1}}{D_K} + \text{terms independent of } \Phi_1$$

formula (4.3) is obtained.

References

COCHRAN, W. & WOOLFSON, M. M. (1955). *Acta Cryst.* **8**, 1-12.

- GOEDKOOP, J. A. (1950). *Acta Cryst.* **3**, 374–378.
 HAUPTMAN, H. & KARLE, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetrical Crystal*. A.C.A. Monograph No. 3.
 HUGHES, E. W. (1953). *Acta Cryst.* **6**, 871.
International Tables for X-ray Crystallography (1952). Vol. I. Birmingham: Kynoch Press.
 KARLE, J. (1971). *Acta Cryst.* **B27**, 2063–2065.
 KARLE, J. & HAUPTMAN, H. (1959). *Acta Cryst.* **12**, 404–410.
 KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849–859.
 KLUG, A. (1958). *Acta Cryst.* **11**, 515–543.
 ROLLETT, J. S. & TOWNSEND, S. P. (1973). Private communication.
 TSOUCARIS, G. (1970). *Acta Cryst.* **A26**, 492–499.
 TSOUCARIS, G. (1973). Personal communication.

Acta Cryst. (1973). **A29**, 615

Molecular Distortion in Orthorhombic Sulphur; a Calculation

BY JYRKI KURITTU

Department of Physics, University of Helsinki, Siltavuorenpenger 20, 00170 Helsinki 17, Finland

AND G. S. PAWLEY

Physics Department, Edinburgh University, King's Buildings, EH9 3JZ, Edinburgh, Scotland

(Received 30 March 1973; accepted 24 April 1973)

A function for the intermolecular forces, based on atom-atom interactions, is used to calculate the distortion forces on the molecule of octasulphur in the orthorhombic structure. By making use of the internal force field resulting from spectroscopic studies the behaviour of the molecule under these forces is found. The calculated molecular distortion agrees very well with that found by performing constrained refinements, except that the distortions calculated are systematically too small. This calculation shows that measurable distortions are to be expected for molecules with fairly low internal-mode frequencies, and that simple models for intermolecular forces predict the observed effect. The energy involved in the molecular distortion is 0.0837 kcal/mole (5.81×10^{-22} joule/molecule).

Introduction

There are many examples in X-ray and neutron diffraction experiments indicating static distortions of molecules in the crystalline environment compared with their free-state shape. In some cases the shifts in the atomic positions are several times larger than the experimental errors.

The static distortions are due to the intermolecular forces in which the repulsive terms between non-bonded atoms must predominate. If we have a potential model for the interaction between these non-bonded atoms we are then able to calculate the force on every atom of each molecule. The response of the molecule depends on the force field between the atoms in the molecule itself. This force field can be found by spectroscopic methods, as described by Wilson, Decius & Cross (1955) or by Cyvin (1972).

We have made a calculation for the distortion of octasulphur molecules in the orthorhombic structure. The potential model from the lattice-statics calculations of Rinaldi & Pawley (1973) is used, along with the molecular force field of Cyvin (1970).

The calculation method

The potential between two non-bonded sulphur atoms located at positions \mathbf{r}_i and \mathbf{r}_j is assumed to be

$$V(r_{ij}) = -\frac{A}{r_{ij}^6} + B \exp(-\alpha r_{ij}) \quad (1)$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ and A , B and α are constants determined from thermodynamic data and lattice statics calculations (Rinaldi & Pawley, 1973):

$$\begin{aligned} A &= 2149 \text{ kcal } \text{\AA}^6/\text{mole} \\ B &= 199900 \text{ kcal/mole} \\ \alpha &= 3.49 \text{ \AA}^{-1}. \end{aligned}$$

By summing over j we get the contribution to the crystal potential of atom i . Only interactions for interatomic distances smaller than 5 Å are taken into account, as the contribution to the potential from more distant contacts depends very little on the detailed atomic arrangement.

The static force on the atom i is the negative gradient of the atom potential, written

$$\mathbf{f}_i = -\text{grad}_i V_i = -\text{grad}_i \sum_j V(r_{ij}) \quad (2)$$

where the gradient is taken with respect to the coordinates of atom i .

The total force on the eight-atom molecule can be represented by one 24-component vector

$$\mathbf{f} = \begin{pmatrix} \mathbf{f}_1 \\ \vdots \\ \mathbf{f}_i \\ \vdots \\ \mathbf{f}_8 \end{pmatrix}. \quad (3)$$

This force must be balanced by the internal force field, which is considered to be harmonic. This leads to the equation

$$\mathbf{f} = \phi \mathbf{u} \quad (4)$$