Acta Cryst. (1958). **11**, 515

Joint Probability Distributions of Structure Factors and the Phase Problem

BY A. KLUG

Birkbeck College Crystallography Laboratory, (University of London), Torrington Square, London, W.C. 1, England

(*Received* 16 *December* 1957)

A method is derived for calculating the joint probability distribution of any desired set of structure factors. The expressions are obtained in the form of a strictly asymptotic series in powers of $N^{-\frac{1}{2}}$ in the case of N equal atoms in the unit cell, and in terms of related quantities in the more general case. The results are thus more accurate than those of Hauptman & Karle, Cochran & Woolfson and Bertaut, and are cast in such a way that it is easy to make a connexion between the statistical approach to the phase problem and the underlying physical basis of the sign relations that have been used. In particular, the dependence of the reliability of the latter on the number of atoms in the unit cell is brought out directly.

The general mathematical theory is given in §§ 1 and 2. In § 1 the probability distribution of a structure factor is calculated by using a relatively simple and powerful method for expressing the probability distribution of a sum of random variables in terms of the cumulants of the distribution of one variable. In § 2 this method has been extended to develop formulae for the multivariate case, and a symbolism is derived for writing down the distribution of a set of structure factors to the desired degree of approximation.

In § 3 the method is used to make a detailed study of the sign relation between $E_{2\mathbf{h}}$ and $|E_{\mathbf{h}}|$ in space group $P\overline{1}$, since this is the prototype of all relations in which signs are determined from intensities. This example is also used to illustrate a general method for determining the mean and higher moments of a structure factor as a function of the values of other related structure factors.

§ 4 is devoted to a thorough study of the sign relationship $S(\mathbf{h}) \approx S(\mathbf{h}')S(\mathbf{h}+\mathbf{h}')$ and to Sayre's equation. In the course of the work a new inequality relating $U(\mathbf{h})$, $U(\mathbf{h}')$ and $U(\mathbf{h}+\mathbf{h}')$ was developed. This gives an expression for the bounding contour of the probability distribution of the three structure factors, and the latter is seen to pass smoothly to the limiting case of the inequality. It is pointed out that this connexion between a probability distribution and an inequality (or equality) can always be made for any chosen set of structure factors.

In § 5 the considerations of the earlier sections are applied to give a critical assessment of the value of various relations or formulae that have been proposed for sign determination. The idea of an *order* of a sign relation is introduced and related to the variance of the estimate of a sign. The advantages claimed for using a formula having strict, rather than merely statistical, validity are disputed and illustrated by a study of the relation between $S(2\mathbf{h}+\mathbf{h}')$ and the values of $|E_{\mathbf{h}}|$ and $E_{\mathbf{h}'}$. A new and more realistic formula is proposed which includes the dependence of $S(2\mathbf{h}+\mathbf{h}')$ on the known value of $|E_{2\mathbf{h}}|$. Finally, it is concluded that there is as yet no routine solution of the phase problem, and that the high-order formulae are not useful for anything but the simplest structures.

Notation

Where possible the notation of Hauptman & Karle (1953) and Cochran & Woolfson (1955) has been used. Note particularly that:

$\overline{X_{\mathbf{h}}}^{\mathbf{h}}$ is the average of X over a range of values of **h**.
$\langle X\rangle$ is the expected value of X.
$\langle X\rangle_Y$ is the expected value of X, when Y is held constant.
$\xi(\mathbf{h})$ is the contribution of an atom to the geometrical structure factor.
N = number of atoms in the unit cell = nt.
t = number of independent atoms in the cell.
n = symmetry number.

$$\varphi_j = f_j \Big/ \sqrt{\sum_{j=1}^{N} f_j^2} \text{ (Bertaut, 1955}a, c\text{).}$$

$$\zeta_\nu = \sum_{j=1}^{t} \varphi_j^\nu .$$

$$z_\nu = n\zeta_\nu = \sum_{j=1}^{N} \varphi_j^\nu = \frac{\varepsilon_\nu}{\varepsilon_2^{\nu/2}} \text{ (Cochran \& Woolfson, 1955).}$$

m_ν = is the νth order moment of ξ.
k_ν = is the νth order cumulant of ξ.
$\lambda_\nu = k_\nu/n^{\nu/2}$ = standardized cumulant.
M = moment generating function (m.g.f.).
C = characteristic function (c.f.).
K = cumulant generating function (c.g.f.).

$$\left.\begin{array}{l}\mathscr{M} = \text{m.g.f.}\\ \mathscr{C} = \text{c.f.}\\ \mathscr{K} = \text{c.g.f.}\end{array}\right\} \text{ of a set of structure factors.}$$

$\mathscr{E}$ is the mathematical expectation operator.
$S(\mathbf{h})$ or $S(E_{\mathbf{h}})$ is the sign of $E_{\mathbf{h}}$.

Introduction

It is now some years since the appearance of the monograph of Hauptman & Karle (1953), who attempted to develop a routine procedure, which did not require previous knowledge of any signs, for determining signs of structure factors in the space group $P\overline{1}$ and other centrosymmetric space groups. That this procedure could lead to a routine solution of the phase problem for crystals of more than moderate complexity was, however, severely disputed by several authors (for references see Cochran & Woolfson (1955)). The importance of Hauptman & Karle's work lies in the introduction of the idea of a joint probability distribution of structure factors which enables one to calculate the probability that a certain structure factor (or set of structure factors) has a certain sign when certain other related structure factors have their observed values. Some relations involve only the magnitudes of the latter related structure factors, while others require that the signs be known as well. The latter are usually known as sign relations. We may call the first category sign determining relations.

The distribution functions that have been obtained by Hauptman & Karle, and others, notably Cochran & Woolfson (1955), are not in closed form, and it is not at all clear what degree of approximation is involved in the results. The statistical problem in crystallography is, however, quite clearly one involving linear sums (structure factors) of random variables (the atomic contributions to a structure factor), and there are more or less standard methods in the theory of probability for dealing with problems of this kind. In this paper we shall use (and extend) these methods to set up accurate expressions (to any desired degree of approximation) for the joint probability distribution of any selected set of structure factors. We shall then apply these results to calculate in detail probability distributions in two distinct examples, one involving the estimation of a sign from intensities ($P(U_{2\mathbf{h}}, U_{\mathbf{h}})$), and the other the estimation of signs from other signs ($P(U_{\mathbf{h}}, U_{\mathbf{h}'}, U_{\mathbf{h}+\mathbf{h}'})$).

Finally we shall discuss the bearing of this work on the solution of the phase problem.

1. Linear sum of random variables and the probability distribution of a single structure factor

The concepts and methods of probability theory which we shall use are summarized in Appendix A.

In this section we shall use these methods to rederive the results of Hauptman & Karle (1953, p. 8, equation (1·29)) and of Bertaut (1955*c*, equation (II–1)) for the probability distribution of a single structure factor. This will serve as a preparation for the more complicated case of distributions involving more than one structure factor. It should be stated, however, that there is also nothing essentially new in the method or the results. We shall only sketch the derivation since what is involved is the well-known theory of the Gram-Charlier and Edgeworth series, and a very good account is given in Cramér's book (1946).

Let us assume that we have a sum E of t random variables ξ_i each of weight φ_i:

$$E = \varphi_1\xi_1+\varphi_2\xi_2+\ldots+\varphi_t\xi_t\,. \qquad (1\cdot1)$$

The connexion with the crystallographic case is obvious, for, in the usual notation, a normalized structure factor is given by

$$E(\mathbf{h}) = \sum_{i=1}^{t} \varphi_i\xi_i(\mathbf{h})\,, \qquad (1\cdot2)$$

where φ_i is the normalized scattering factor introduced by Bertaut (1955*a*, *c*)

$$\varphi_j = f_j\Big/\Big(\sum_{1}^{N} f_j^2\Big)^{\frac{1}{2}} \qquad (1\cdot3)$$

(we shall in what follows ignore the dependence of φ and f on $\mathbf{h}$). $\xi(\mathbf{h})$ is the trigonometric factor for the space group involved and t is the number of independent atoms in the unit cell. The total number of atoms in the cell is $N = nt$, where n is the symmetry number. For space group $P\overline{1}$, $n = 2$,

$$\xi(\mathbf{h}) = 2\cos 2\pi\mathbf{h}.\mathbf{r} \qquad (1\cdot4)$$

and $N = 2t$.

It also follows that

$$\overline{E^2} = \sum_{j=1}^{t} \varphi_j^2\,\overline{\xi_j^2} = \sum_{j=1}^{N/n} n\varphi_j^2 = \sum_{j=1}^{N} \varphi_j^2 = 1\,. \qquad (1\cdot5)$$

When nothing is known *a priori* about the distribution of atoms in a unit cell, the position vector $\mathbf{r}$ may be assumed to be a random variable and hence $\xi(\mathbf{h})$ will also be a random variable. Note that, in what follows, the distribution of $\mathbf{r}$ and of $\xi(\mathbf{h})$ is not explicitly introduced, and in fact the general theory holds for any distribution one might consider. In practice, when one comes to calculate a result, it is simplest to assume that the atoms are distributed uniformly* over the unit cell.

While the analysis can be carried out quite generally, it is instructive for the moment to assume that all the atoms are equal and have the same distribution. Then equation (1) becomes

$$E = \varphi\sum_{1}^{t} \xi_i\,, \qquad (1\cdot6)$$

where

$$\varphi = N^{-\frac{1}{2}} = (nt)^{-\frac{1}{2}}\,. \qquad (1\cdot6a)$$

All the components ξ_i have the same distribution, with mean zero and standard deviation σ_1 given by

$$\sigma_1^2 = \overline{\xi^2} = n, \text{ the symmetry number.} \qquad (1\cdot6b)$$

The sum E has a zero mean and a standard deviation

---

* The effect of imposing stronger restrictions on the possible positions of atoms or pairs of atoms will be considered elsewhere.

of unity. If $C_1(u)$ is the characteristic function of the single random variable ξ, and $C(u)$ is the c.f. of the sum E, then, from equation (A·14),

$$C(u) = [C_1(\varphi u)]^t \,,$$

which, using (A·21), becomes

$$C(u) = \exp\,[tK_1(i\varphi u)] \,, \tag{1·7}$$

where $K_1(u)$ is the c.g.f. of the single random variable ξ. That is,

$$K_1(iu) = \sum_{\nu=1}^{\infty} k_\nu \frac{(iu)^\nu}{\nu!} \,, \tag{1·8}$$

where the k_ν are the cumulants of the distribution of ξ. So, combining (7) and (8),

$$\begin{aligned} C(u) &= \exp\left[t\sum_{\nu=1}^{\infty} k_\nu \frac{(i\varphi u)^\nu}{\nu!}\right] \\ &= \exp\left[t\sum_{\nu=1}^{\infty} \frac{k_\nu}{n^{\nu/2}\nu!}\left(\frac{iu}{t^{\frac{1}{2}}}\right)^\nu\right], \end{aligned} \tag{1·9}$$

from (1·6a).

At this stage it is clearly convenient to transform the distribution of ξ to standard measure by dividing the cumulants k_ν by $\sigma_1^\nu = (\sqrt{n})^\nu$, that is, we introduce the standardized cumulants

$$\lambda_\nu = k_\nu/\sigma_1^\nu = k_\nu/n^{\nu/2} \,. \tag{1·10}$$

(Note that the distribution of E is already in standard measure; had we initially worked with $F(\mathbf{h})$ rather than $E(\mathbf{h})$ we should have had to transform to the latter at this point.)

Hence we may write equation (9) as

$$C(u) = \exp\left[t\sum_{\nu=1}^{\infty} \frac{\lambda_\nu}{\nu!}\left(\frac{iu}{t^{\frac{1}{2}}}\right)^\nu\right]. \tag{1·11}$$

Now $\lambda_1 = 0$ and $\lambda_2 = 1$, and taking the term in $\nu = 2$ to the left-hand side, we write the last equation as

$$\exp\left(\frac{u^2}{2}\right) C(u) = \exp\left[t\sum_{\nu=3}^{\infty} \frac{\lambda_\nu}{\nu!}\left(\frac{iu}{t^{\frac{1}{2}}}\right)^\nu\right]. \tag{1·12}$$

To derive the Gram-Charlier series, the right hand side is now expanded in powers of u, giving (see Cramér, 1946, § 17·6)

$$\exp\left(\frac{u^2}{2}\right) C(u) = 1 + A_3\frac{(iu)^3}{3!} + A_4\frac{(iu)^4}{4!} + \ldots \,, \tag{1·13}$$

or

$$C(u) = \exp\left(-\frac{u^2}{2}\right) + \frac{A_3}{3!}(iu)^3 \exp\left(-\frac{u^2}{2}\right) + \ldots \,, \tag{1·14}$$

where

$$\begin{aligned} A_3 &= \lambda_3/t^{1/2}, & A_7 &= \frac{\lambda_7}{t^{5/2}} + \frac{35\lambda_3\lambda_4}{t^{3/2}} \,, \\ A_4 &= \lambda_4/t, & A_8 &= \frac{\lambda_8}{t^3} + \frac{1}{t^2}\begin{pmatrix}\text{terms in } \lambda_3\lambda_5 \\ \text{and } \lambda_4^2\end{pmatrix}, \\ A_5 &= \lambda_5/t^{3/2}, & A_9 &= \frac{\lambda_9}{t^{7/2}} + \frac{280\lambda_3^3}{t^{3/2}} + \ldots \,. \\ A_6 &= \frac{\lambda_6}{t^2} + \frac{10\lambda_3^2}{t} \,, \end{aligned} \tag{1·15}$$

To obtain the probability distribution of E, $P(E)$, that we seek, we now take the Fourier transform of equation (14). The transform of each term is obtained by using the result

$$\begin{aligned} \frac{1}{2\pi}\int_{-\infty}^{\infty} (iu)^\nu \exp\,[-\tfrac{1}{2}u^2] \exp\,[-iux]du \\ = (-1)^\nu \varphi^{(\nu)}(x) = H_\nu(x)\varphi(x) \,, \end{aligned} \tag{1·16}$$

where $\varphi^{(\nu)}(x)$ is the νth derivative of the standardized Gaussian function

$$\varphi(x) = (2\pi)^{-\frac{1}{2}} \exp\,(-\tfrac{1}{2}x^2) \tag{1·17*}$$

and $H_\nu(x)$ is the Hermite polynomial of the νth order, defined by the second member of equation (16). Some of the $H_\nu(x)$ are given and tabulated in Appendix F.

So, finally, we have

$$\begin{aligned} P(E) &= \varphi(E) - \frac{A_3}{3!}\varphi^{(3)}(E) + \frac{A_4}{4!}\varphi^{(4)}(E) + \ldots \,. \\ &= \varphi(E)\left[1 + \frac{A_3}{3!}H_3(E) + \frac{A_4}{4!}H_4(E) + \ldots\right], \end{aligned} \tag{1·18}$$

where the A_ν's are given by equation (15).

This result is perfectly general. To arrive at the results of Hauptman & Karle (1953, p. 8) for the space group $P\bar{1}$, all one has to do is to substitute from (15) for the A_ν's in terms of the cumulants, and then for the latter in terms of the moments of the probability distribution of ξ (see (A·19)). As stressed above, it is only at this stage in the derivation that the nature of the assumed probability distribution enters into the theory.

From equations (15) and (18), we see that the correction terms to the simple Gaussian distribution for E depend on powers of $1/t^{\frac{1}{2}}$, where t is the number of random variables contributing to E, or, in the crystallographic case, on powers of $1/N^{\frac{1}{2}}$ where N is the number of atoms in the unit cell. Since in practice it is necessary to take only a finite number of terms in the series, we require that the terms tend *regularly* to zero so that a term is negligible compared with the preceding one, i.e. that we have a true asymptotic series. This is not, however, the case for the above series since terms in $1/N$, for instance, occur in A_6 as well as in A_4. Furthermore, we see from (15) that contributions up to order $1/N^{\frac{1}{2}}$ contain no cumulants of order ≥ 3, those up to order $1/N$ no cumulants of order > 4, and so on. Indeed it is easy to show that if we wish to consider only cumulants (and moments) up to order ν we need not take into account any terms of power greater than $1/N^{\frac{1}{2}\nu-1}$. The form of the Gram-Charlier series (18), in which the individual terms A_ν arise as coefficients of powers of the carrying variable u, obscures this simple and important fact.

---

* No confusion ought to arise between the two standard usages of φ in (1·3) and in (1·17), since the two meanings never occur together.

It was to overcome this objection that Edgeworth devised a series (see Cramér, 1946, Chap. 7) which gives a straightforward expansion in powers of $1/N^{\frac{1}{2}}$ and which is such that the calculation of terms up to a desired order does not require knowledge of any moments that are not really necessary. We cannot here go over this well-known section of random variable theory; but the point is raised here so that we might, in what follows, aim at proper asymptotic expansions in which the remainder is of the same order as the first term neglected.

The method given here for calculating the probability distribution of a structure factor is a precise formulation of the original approach of Hauptman & Karle (1953) and of Bertaut's first papers (1955*a*, *b*). In a later paper Bertaut (1955*c*, equations (I·16) and (II·1)) makes use of a rather dubious expansion of the Dirac δ-function to arrive at what is essentially the Gram-Charlier series given above. Since he has made great use of this method, perhaps it is worth while making the following comments on the form of his equations.

When one is dealing with a probability distribution $f(x)$ that is approximately Gaussian, it is convenient to express it in the form of a series in terms of the normal error function and its derivatives, thus:

$$f(x) = \varphi(x)+\frac{c_3}{3!}\varphi^{(3)}(x)+\frac{c_4}{4!}\varphi^{(4)}(x)+\dots, \quad (1{\cdot}19)$$

which, by (16), becomes

$$f(x) = \varphi(x)\left[1+\frac{A_3}{3!}H_3(x)+\frac{A_4}{4!}H_4(x)+\dots\right], \quad (1{\cdot}20)$$

where

$$A_\nu = (-1)^\nu c_\nu = \int_{-\infty}^{\infty} H_\nu(x) f(x)\,dx \quad (1{\cdot}21)$$

because of the orthogonal property of the Hermite polynomials. Using angular brackets to denote averages, we may write

$$A_\nu = \langle H_\nu(x)\rangle\,, \quad (1{\cdot}21a)$$

and then (20) becomes

$$f(x) = \varphi(x)\left[1+\sum_{\nu=3}^{\infty}\frac{\langle H_\nu(x)\rangle H_\nu(x)}{\nu!}\right], \quad (1{\cdot}22)$$

which is exactly the expression that Bertaut's expansion of the δ-function was designed to give. While formally correct, we have seen that, for the problems we are dealing with, x is itself a sum of random variables, and the form of equation (22) hides all reference to the orders of the terms of the series, as shown above for (18). It should, however, be mentioned that in his first approach, subsequently abandoned, Bertaut (1955*a*, equation (III·41)) gives the dependence of the terms on powers of $1/N^{\frac{1}{2}}$.

Finally we may note, for completeness, that curves of $P(E)$ (equation (18)) have been calculated and plotted by Slack (1946) in another context.

2. Joint probability distribution of structure factors: general expression

Consider a set of m normalized structure factors $E(\mathbf{h}_i)$, which, when there is no ambiguity, we shall write freely as E_i (here again as in § 1 it is convenient to normalize from the outset):

$$\begin{aligned} E(\mathbf{h}_1) &= \varphi_1\xi_1(\mathbf{h}_1)+\varphi_2\xi_2(\mathbf{h}_1)+\dots+\varphi_t\xi_t(\mathbf{h}_1)\,,\\ E(\mathbf{h}_2) &= \varphi_1\xi_1(\mathbf{h}_2)+\dots\dots\dots\dots+\varphi_t\xi_t(\mathbf{h}_2)\,, \quad (2{\cdot}1)\\ &\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\ E(\mathbf{h}_m) &= \varphi_1\xi_1(\mathbf{h}_m)+\dots\dots\dots\dots+\varphi_t\xi_t(\mathbf{h}_m)\,. \end{aligned}$$

The definitions and notation we shall use for a multivariate probability distribution are given in Appendix B.

We wish to develop an expression for the joint probability distribution of the $E(\mathbf{h}_i)$, analogous to equations (1·18) and (1·15) for a single structure factor. Mathematically, we have to derive the Gram-Charlier, or rather Edgeworth, series for a multivariate distribution. A search through the statistical literature failed to reveal the formal expressions, but it is fairly easy to generalize the method of § 1 and Appendix A to derive them. That is, the characteristic function of the multivariate distribution is first evaluated, and then the probability distribution is found by taking the Fourier transform. For brevity, however, we work with the entirely equivalent moment-generating function to save writing complex i's, and, as will be seen, the mathematical operations become largely symbolic ones, very easy to manipulate.

For each $E(\mathbf{h}_i)$ we introduce a carrying variable u_i (note that it is the E's that are from this point of view the random variables, not simply the ξ's as might at first have been thought). Then the m.g.f. of the E's is, from (2·1),

$$\begin{aligned} \mathscr{M}(u_1, u_2, \dots u_m) &\equiv \mathscr{E}\exp\left[u_1E(\mathbf{h}_1)+\dots+u_mE(\mathbf{h}_m)\right]\\ &= \mathscr{E}\exp\left[u_1\{\varphi_1\xi_1(\mathbf{h}_1)+\dots+\varphi_t\xi_t(\mathbf{h}_1)\}\right.\\ &\qquad +u_2\{\varphi_1\xi_1(\mathbf{h}_2)+\dots+\varphi_t\xi_t(\mathbf{h}_2)\}\\ &\qquad +\dots\dots\dots\dots\dots\dots\\ &\qquad \left.+u_m\{\varphi_1\xi_1(\mathbf{h}_m)+\dots+\varphi_t\xi_t(\mathbf{h}_m)\}\right]\,, \end{aligned}$$

which we re-arrange as

$$\begin{aligned} &=(\mathscr{E}\exp\left[u_1\varphi_1\xi_1(\mathbf{h}_1)+u_2\varphi_1\xi_1(\mathbf{h}_2)+\dots+u_m\varphi_1\xi_1(\mathbf{h}_m)\right])\\ &\times(\mathscr{E}\exp\left[u_1\varphi_2\xi_2(\mathbf{h}_1)+u_2\varphi_2\xi_2(\mathbf{h}_2)+\dots+u_m\varphi_2\xi_2(\mathbf{h}_m)\right])\\ &\times\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\ &\times(\mathscr{E}\exp\left[u_1\varphi_t\xi_t(\mathbf{h}_1)+u_2\varphi_t\xi_t(\mathbf{h}_2)+\dots+u_m\varphi_t\xi_t(\mathbf{h}_m)\right])\\ &=M_1(\varphi_1u_1, \varphi_1u_2, \dots, \varphi_1u_m)M_2(\varphi_2u_1, \varphi_2u_2, \dots, \varphi_2u_m)\\ &\times\dots\dots\dots\dots\dots\dots\times M_t(\varphi_tu_1, \varphi_tu_2, \dots, \varphi_tu_m)\,, \end{aligned} \quad (2{\cdot}2)$$

where $M_i(\varphi_iu_1, \varphi_iu_2, \dots, \varphi_iu_m)$ is the m.g.f. of the joint probability distribution of the contributions of the ith atom to the structure factors $E_1, \ldots, E_m$. If all the atoms are assumed to be distributed independently through the unit cell in the same way, then all the ξ_i have the same distribution and we can drop the suffix i. Then we have

$$\mathscr{M}(u_1, \ldots, u_m) = M(\varphi_1 u_1, \ldots, \varphi_1 u_m) \ldots M(\varphi_t u_1, \ldots, \varphi_t u_m) \,. \tag{2·3}$$

We shall in this section confine ourselves to the case where all the atoms have the same scattering factor, i.e. $\varphi_1 = \ldots = \varphi_t = \varphi$, and shall find that, just as in the univariate case, the coefficients in the series finally obtained will be powers of $1/N^{\frac{1}{2}}$. The case of unequal atoms is treated in Appendix C.

For equal atoms, (2·3) becomes

$$\mathscr{M}(u_1, \ldots, u_m) = [M(\varphi u_1, \ldots, \varphi u_m)]^t, \tag{2·4}$$

where $M(u_1, \ldots, u_m)$

$$= 1 + \frac{m_{200\ldots}}{2!\,0!\,0!\ldots} u_1^2 + \frac{m_{020\ldots}}{0!\,2!\,0!\ldots} u_2^2 + \ldots + \frac{m_{pq\ldots w}}{p!\,q!\ldots w!} u_1^p u_2^q \ldots u_m^w + \ldots \qquad (2·5)^*$$

and the mixed moments $m_{pqr\ldots}$ are defined as

$$m_{pqr} = \langle [\xi(\mathbf{h}_1)]^p [\xi(\mathbf{h}_2)]^q [\xi(\mathbf{h}_3)]^r \ldots \rangle \tag{2·6}$$

$$= \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} [\xi(\mathbf{h}_1)]^p [\xi(\mathbf{h}_2)]^q \ldots \times \psi(\xi(\mathbf{h}_1), \ldots, \xi(\mathbf{h}_m)) d\xi(\mathbf{h}_1) \ldots d\xi(\mathbf{h}_m) \,,$$

where $\psi(\xi(\mathbf{h}_1) \ldots \xi(\mathbf{h}_m))$ is the *a priori* probability that $\xi(\mathbf{h})$ has the value $\xi(\mathbf{h}_1)$ for $\mathbf{h} = \mathbf{h}_1$, the value $\xi(\mathbf{h}_2)$ for $\mathbf{h} = \mathbf{h}_2$, etc. The simplest assumption, which has been so far used by almost all authors on this subject, is that in which the atoms are distributed *uniformly* throughout the cell (i.e. they may be considered to perform a random walk). Then the distribution of $\xi(\mathbf{h})$ is independent of $\mathbf{h}$, and the mixed moments may be calculated very easily (Hauptman & Karle, 1953; see also Appendix D). In this paper we shall make this assumption as well, but it should be noted that the general theory of this section does not require a specification of the distribution; only when we come to carry out computations do we have to insert values for the mixed moments $m_{pqr\ldots}$.

As indicated earlier, we shall not carry out the general analysis in terms of moments, but now introduce the multivariate cumulant generating functions (see Appendix B)

$$\mathscr{K}(u_1, \ldots, u_m) = \log \mathscr{M}(u_1, \ldots, u_m)$$

and

$$K(u_1, \ldots, u_m) = \log M(u_1, \ldots, u_m) \,, \tag{2·7}$$

where $K(u_1, \ldots, u_m)$

$$= \frac{k_{200\ldots}}{2!\,0!\,0!\ldots} u_1^2 + \frac{k_{020\ldots}}{0!\,2!\,0!\ldots} u_2^2 + \ldots + \frac{k_{pq\ldots w}}{p!\,q!\ldots w!} u_1^p u_2^q \ldots u_m^w + \ldots \,. \qquad (2·8)^*$$

In this expansion there will be as many terms of the νth degree, as there are partitions of

$$p+q+r+\ldots = \nu \,, \tag{2·9}$$

and we may refer to the set of corresponding cumulants $k_{pqr\ldots}(p+q+r\ldots = \nu)$ as νth order cumulants. The relations between the cumulants and moments are given in Appendix B.

Then equation (2·4) for the m.g.f. may be written

$$\begin{aligned}\mathscr{M}(u_1, \ldots, u_m) &= \{\exp [K(\varphi u_1, \ldots, \varphi u_m)]\}^t \\ &= \exp [tK(\varphi u_1, \ldots, \varphi u_m)] \\ &= \exp \left[t \sum_{p+q\ldots=2}^{\infty} \frac{k_{pqr\ldots} \varphi^{p+q+r\ldots}}{p!\,q!\,r!\ldots} \times u_1^p u_2^q u_3^r \ldots \right] . \end{aligned} \tag{2·10}$$

Now, for equal atoms,

$$\varphi = 1/(nt)^{\frac{1}{2}} \,, \tag{2·11}$$

where n is the symmetry number of the space group. Furthermore, by (1·6b) the standard deviation of the random variable ξ is $n^{\frac{1}{2}}$, and hence

$$k_{200\ldots} = k_{020\ldots} = \langle \xi^2 \rangle = n \,.$$

We transform our notation to that of standard measure (as in § 1) by introducing the standardized cumulants

$$\begin{aligned}\lambda_{pqr\ldots} &= k_{pqr\ldots} / k_{200\ldots}^{p/2} k_{020\ldots}^{q/2} k_{002\ldots}^{r/2} \ldots \\ &= k_{pqr\ldots} / (\sqrt{n})^{p+q+r\ldots} \,. \end{aligned} \tag{2·12}$$

Then from (2·11) and (2·12), (2·10) becomes

$$\mathscr{M}(u_1, \ldots, u_m) = \exp \left[t \sum_{p+q\ldots=2}^{\infty} \frac{\lambda_{pq\ldots}}{p!\,q!\ldots} \left(\frac{u_1}{t^{\frac{1}{2}}}\right)^p \left(\frac{u_2}{t^{\frac{1}{2}}}\right)^q \ldots \right], \tag{2·13}$$

where $\lambda_{200\ldots} = \lambda_{020\ldots} = \ldots = 1$.

We now proceed just as in the univariate case considered in § 1. Taking the terms in $p+q+r\ldots = 2$ to the left-hand side, the last equation can be written as

$$\exp [-\tfrac{1}{2}(u_1^2 + u_2^2 + \ldots + 2\lambda_{110\ldots} u_1 u_2 + 2\lambda_{101\ldots} u_1 u_3 + \ldots)] \times \mathscr{M}(u_1, \ldots, u_m)$$

$$= \exp \left[t \sum_{p+q\ldots=3}^{\infty} \frac{\lambda_{pqr\ldots}}{p!\,q!\,r!\ldots} \left(\frac{u_1}{t^{\frac{1}{2}}}\right)^p \left(\frac{u_2}{t^{\frac{1}{2}}}\right)^q \left(\frac{u_3}{t^{\frac{1}{2}}}\right)^r \ldots \right] . \tag{2·14}$$

To simplify writing out the expansion of the right-hand side of this equation, it is convenient to introduce the notation

---

* We are working here in standard measure where the means $m_{100\ldots} = k_{100\ldots} = 0$, $m_{010\ldots} = k_{010\ldots} = 0$, etc.

Λ_ν = set of all terms with $p+q+r\ldots = \nu$

$$= \mathop{\mathsf{S}}_{p+q+\ldots=\nu} \frac{\lambda_{pqr\ldots}}{p!\,q!\,r!\ldots} u_1^p u_2^q u_3^r \ldots, \qquad (2\cdot15)$$

where the sum S is taken over all partitions of $p+q+r\ldots = \nu$. The right-hand side of (2·14) is, in this notation,

$$\exp\left[t \sum_{\nu=3}^{\infty} \frac{\Lambda_\nu}{t^{\nu/2}}\right] = \exp\left[\sum_{\nu=3}^{\infty} \frac{\Lambda_\nu}{t^{\nu/2-1}}\right],$$

and, on expanding it, we obtain

$$\begin{aligned} &1 \\ &+\frac{1}{t^{\frac{1}{2}}}\Lambda_3 \\ &+\frac{1}{t}\left(\Lambda_4+\frac{1}{2!}\Lambda_3^2\right) \\ &+\frac{1}{t^{3/2}}\left(\Lambda_5+\frac{1}{2!}2\Lambda_3\Lambda_4+\frac{1}{3!}\Lambda_3^3\right) \qquad (2\cdot16) \\ &+\ldots. \end{aligned}$$

In this last expression we have, as foreshadowed in § 1, collected terms according to powers of $1/t^{\frac{1}{2}}$, where t is the number of random variables involved (and which, we recall, is proportional to N). The notation (2·15) has been chosen to bring out just this similarity between the multivariate and univariate cases, and (2·16) should be compared with (1·15).

The inversion of the m.g.f. $\mathscr{M}(u_1, u_2, \ldots, u_m)$ to give the joint probability distribution of the E's, $P(E_1, E_2, \ldots, E_m)$ now follows just as in § 1. The second-order terms are brought back to to the right-hand side to give

$$\begin{aligned} &\mathscr{M}(u_1, u_2, \ldots, u_m) \\ &= \exp\left[\tfrac{1}{2}(u_1^2+u_2^2+\ldots+2\lambda_{110}u_1u_2+\ldots)\right] \\ &\times\left\{1+\frac{1}{t^{\frac{1}{2}}}\Lambda_3+\frac{1}{t}(\Lambda_4+\tfrac{1}{2}\Lambda_3^2)+\frac{1}{t^{3/2}}(\Lambda_5+\Lambda_3\Lambda_4+\tfrac{1}{6}\Lambda_3^3)+\ldots\right\}. \end{aligned} \qquad (2\cdot17)$$

In the inversion (cf. equation (1·16)) a term

$$\exp\left[\tfrac{1}{2}(u_1^2+u_2^2+2\lambda_{110\ldots}u_1u_2+\ldots)\right]u_1^p u_2^q u_3^r\ldots$$

will give rise to the term

$$H_p(E_1)H_q(E_2)H_r(E_3)\ldots\varphi_0(E_1, E_2, \ldots, E_m),$$

where $\varphi_0(E_1\ldots E_m)$ is the standardized multivariate normal distribution function

$$\begin{aligned} &\varphi_0(E_1, \ldots, E_m) \\ &\quad = (|D|)^{\frac{1}{2}}(2\pi)^{-m/2}\exp\left[-\tfrac{1}{2}(E_1^2+E_2^2+2a_{12}E_1E_2+\ldots)\right], \end{aligned} \qquad (2\cdot18)$$

$|D|$ being the determinant of the matrix

$$\mathbf{D} = \left\| \begin{matrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \ldots & a_{mm} \end{matrix} \right\|, \qquad (2\cdot18a)$$

where

$$a_{ij} = \frac{1}{|M|}\frac{\partial \mathbf{M}}{\partial \lambda_{ij}}. \qquad (2\cdot18b)$$

Here $\mathbf{M}$ is the moment matrix

$$\mathbf{M} = \left\| \begin{matrix} \lambda_{200\ldots} & \lambda_{110\ldots} & \cdots & \lambda_{100\ldots1} \\ \lambda_{110\ldots} & \lambda_{020\ldots} & \cdots & \lambda_{010\ldots1} \\ \ldots & \ldots & \ldots & \ldots \\ \lambda_{100\ldots1} & \ldots & \ldots & \lambda_{00\ldots2} \end{matrix} \right\| \qquad (2\cdot18c)$$

and λ_{ij} in (2·18b) refers to the λ in the ith row and jth column of the matrix $\mathbf{M}$.

In all* practical applications $\lambda_{110\ldots} = \lambda_{011\ldots} = \ldots = 0$, and $\lambda_{200\ldots} = \lambda_{0200} = \ldots = 1$, so that (2·18) simply becomes

$$\varphi(E_1\ldots E_m) = (2\pi)^{-m/2}\exp\left[-\tfrac{1}{2}(E_1^2+\ldots+E_m^2)\right]. \qquad (2\cdot19)$$

We thus see that it is a comparatively simple matter to derive the expression for the joint probability distribution, and no purpose would be served in writing down a formal answer. It is enough to write down the moment-generating function and invert term by term, using the rules just given. The mathematics thus becomes merely a matter of carrying out rather simple symbolic operations. As an example we shall give the result for one of the possible terms in the set $\frac{1}{2t}\Lambda_3^2$ of the m.g.f. (2·17). For instance, the term

$$\begin{aligned} &\frac{1}{2t}\frac{\lambda_{300\ldots}}{3!\,0!\,0!\ldots}u_1^3 \times \frac{\lambda_{1110\ldots}}{1!\,1!\,1!\ldots}u_1u_2u_3 \\ &= \frac{1}{2t}\frac{1}{6}\lambda_{300\ldots}\lambda_{111\ldots}u_1^4u_2u_3 \end{aligned}$$

gives rise, on inversion, to the term

$$\frac{1}{2t}\frac{1}{6}\lambda_{300\ldots}\lambda_{100\ldots}H_4(E_1)H_1(E_2)H_2(E_3),$$

apart from a Gaussian factor which is the same for all terms. The problem of inversion to obtain the complete asymptotic series is largely one of choosing a good notation; the process should become clearer when we deal with practical examples in the next two sections.

3. Relations between the sign of a structure factor and the intensities

3·1. *Joint probability distribution of* $U_{2\mathbf{h}}$ *and* $U_{\mathbf{h}}$

We shall calculate the example in detail for the space group $P\bar{1}$. The results will carry over to the case of higher space groups, and indeed the example is the prototype of all calculations of signs of a selected group of structure factors from the intensities of others related to them, e.g. in $P2_1/a$ the signs of $U_{2h,0,2l}$ from

---

* Except when certain restrictions are made on the distribution of ξ; see remarks at the end of Appendix B.

the values of $U^2_{h,k,l}$ where k is a free index (Hauptman & Karle, 1953, p. 50; Cochran, 1954).

For brevity, we use the notation

$$E_1 = E_{2\mathbf{h}},\quad E_2 = E_{\mathbf{h}}\,.$$

The subscripts on a mixed moment or cumulant are then to be taken as referring to E_1 and E_2 in that order

$$m_{pq} = \langle (2\cos 2\pi 2\mathbf{h}.\mathbf{r})^p (2\cos 2\pi\mathbf{h}.\mathbf{r})^q\rangle\,. \tag{3·1}$$

The only mixed moments that do not vanish are those that are even in both suffixes, and those that are odd in the first and even in the second. Their values may be found directly by evaluating the average (3·1) or more simply by the method of Bertaut (1955*a*, *b*) (see Appendix D). The mixed cumulants are then calculated by means of the equations (B·7).

From (2·15) and (2·17) the m.g.f. function is given by

$$\begin{aligned}\mathscr{M}(u_1,u_2) = \exp\left[\tfrac{1}{2}(u_1^2+u_2^2)\right]\Bigg\{1 &+ \frac{1}{t^{\frac{1}{2}}}\frac{\lambda_{12}}{1!\,2!}u_1u_2^2\\ &+\frac{1}{t}\left[\frac{\lambda_{40}}{4!\,0!}u_1^4+\frac{\lambda_{04}}{0!\,4!}u_2^4\right]+\frac{1}{2t}\left[\frac{\lambda_{12}}{1!\,2!}u_1u_2^2\right]^2\\ &+\frac{1}{t^{3/2}}\left[\frac{\lambda_{32}}{3!\,2!}u_1^3u_2^2+\frac{\lambda_{14}}{1!\,4!}u_1u_2^4\right]\\ &+\frac{1}{t^{3/2}}\left[\frac{\lambda_{12}}{1!\,2!}u_1u_2^2\right]\left[\frac{\lambda_{40}}{4!\,0!}u_1^4+\frac{\lambda_{04}}{0!\,4!}u_2^4\right]\\ &+\frac{1}{6t^{3/2}}\left[\frac{\lambda_{12}}{1!\,2!}u_1u_2^2\right]^3+\ldots\Bigg\}.\end{aligned} \tag{3·2}$$

Inverting according to the rules enunciated in the last section, we obtain

$$\begin{aligned}P(E_1,E_2) = \frac{1}{2\pi}\exp\left[-\tfrac{1}{2}(E_1^2+E_2^2)\right]\Bigg\{1 &+ \frac{1}{t^{1/2}}\frac{\lambda_{12}}{1!\,2!}\\ &\times H_1(E_1)H_2(E_2)+\frac{1}{t}\left[\frac{\lambda_{40}}{4!\,0!}H_4(E_1)+\frac{\lambda_{04}}{0!\,4!}H_4(E_2)\right]\\ &+\frac{1}{2t}\left(\frac{\lambda_{12}}{1!\,2!}\right)^2 H_2(E_1)H_4(E_2)\\ &+\frac{1}{t^{3/2}}\left[\frac{\lambda_{32}}{3!\,2!}H_3(E_1)H_2(E_2)+\frac{\lambda_{14}}{1!\,4!}H_1(E_1)H_4(E_2)\right]\\ &+\frac{1}{t^{3/2}}\left[\frac{\lambda_{12}\lambda_{40}}{1!\,2!\,4!\,0!}H_5(E_1)H_2(E_2)\right.\\ &\left.+\frac{\lambda_{12}\lambda_{04}}{1!\,2!\,0!\,4!}H_1(E_1)H_6(E_2)\right]\\ &+\frac{1}{6t^{3/2}}\left(\frac{\lambda_{12}}{1!\,2!}\right)^3 H_3(E_1)H_6(E_2)+\ldots\Bigg\}.\end{aligned} \tag{3·3}$$

Substituting numerical values for the λ's (Appendix D), and putting $N = 2t$, we finally obtain the asymptotic expansion

$$\begin{aligned}P(E_1,E_2) = \frac{1}{2\pi}\exp\left[-\tfrac{1}{2}E_1^2-\tfrac{1}{2}E_2^2\right]\Bigg\{1 &+ \frac{1}{2N^{\frac{1}{2}}}H_1(E_1)H_2(E_2)\\ &-\frac{1}{8N}\left[H_4(E_1)+H_4(E_2)\right]+\frac{1}{8N}H_2(E_1)H_4(E_2)^*\\ &-\frac{1}{N^{3/2}}\left[\tfrac{1}{4}H_3(E_1)H_2(E_2)+\tfrac{1}{3}H_1(E_1)H_4(E_2)\right]\\ &-\frac{1}{N^{3/2}}\left[\tfrac{1}{16}H_5(E_1)H_2(E_2)+\tfrac{1}{16}H_1(E_1)H_6(E_2)\right]^*\\ &+\frac{1}{48N^{3/2}}H_3(E_1)H_6(E_2)^*+\ldots\Bigg\}.\end{aligned} \tag{3·4}$$

In this expression we have marked by an asterisk those terms which would not be given by Bertaut's symbolic expansion of the δ-function (1955*c*, equation (III·1)) up to terms of the fifth order. These terms could only be retrieved by carrying his expansion as far as the ninth-order terms. The other terms in (3·4) are identical with those that would be obtained from Bertaut's method since it may be shown (Cochran, private communication; cf. Cochran & Woolfson, 1955, Appendix II) that the averages that would occur in his expressions reduce to

$$\begin{aligned}\langle E_1(E_2^2-1)\rangle &= N^{-\frac{1}{2}}\,,\\ \langle E_1(E_2^4-6E_2^2)\rangle &= -8N^{-3/2}\,,\\ \langle (E_1^3-3E_1)(E_2^2-1)\rangle &= -3N^{-3/2}\,.\end{aligned} \tag{3·5}$$

We have described the steps leading up to equation (3·4) in some detail since all probability distributions of interest can be derived in the same way. The only new work that has to be done in the case of a different set of structure factors is the calculation of the non-vanishing mixed moments and cumulants.

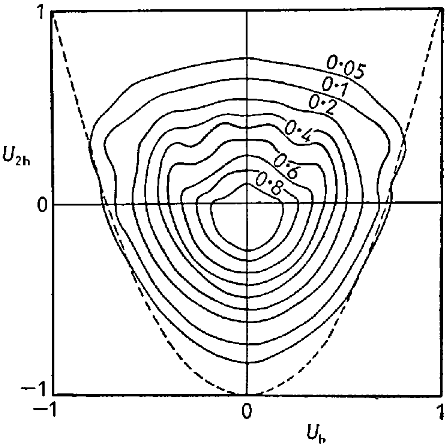

Fig. 1. Joint probability distribution of $U_{2\mathbf{h}}$ and $U_{\mathbf{h}}$ for space group $P\bar{1}$ calculated from equation (3·4) for the particular case $N = 10$.

To obtain the normalized distribution the numbers with which the contours are labelled should be divided by 2π.

The broken line shows the bounding contour of the range permitted by the Harker–Kasper inequality $U_{2\mathbf{h}} \geq 2U^2_{\mathbf{h}}-1$, and therefore corresponds to $P = 0$ for the probability distribution.

We have carried the expansion (3·4) as far as terms in $1/N^{3/2}$ so that we might gain some insight into the importance of the various orders of terms in practical calculations of the distribution functions. We have computed $P(E_{2\mathbf{h}}, E_{\mathbf{h}})$ from (3·4) for the case of $N = 10$ and the results are plotted in Fig. 1. This particular value of N has been chosen so that a comparison might be made with Fig. 2 of Vand & Pepinsky (1953), which was computed for this special case by numerical convolution and is therefore strictly accurate. The contours including terms up to power $1/N^{3/2}$ agree with those of Vand & Pepinsky down to about $P = 0{\cdot}3$ and show that the theory is correct. The fact that the contours beyond this range do not join quite smoothly into the limits set by the Harker–Kasper inequality, $U_{2\mathbf{h}} \geq 2U_{\mathbf{h}}^2-1$, arises from the neglect of higher-order terms in the asymptotic expansion. By including more terms the curves could be made to resemble the true ones to any desired degree. The effect of including only the terms up to $1/N^{\frac{1}{2}}$ in the computation may be seen by looking at Fig. 1 of Cochran & Woolfson (1955); this was computed from their expression (3·12) for $P(E_{2\mathbf{h}}, E_{\mathbf{h}})$, which is correct to the order $1/N^{\frac{1}{2}}$. Only the area near the origin is well reproduced in this approximation.

3·2. *The conditional probability distribution* $P(E_{2\mathbf{h}}|E_{\mathbf{h}})$

A useful procedure for obtaining some insight into the nature of the probability distribution is to evaluate certain characteristics of the conditional distribution $P(E_{2\mathbf{h}}|E_{\mathbf{h}})$ of $E_{2\mathbf{h}}$ on $E_{\mathbf{h}}$. The expression for $P(E_{2\mathbf{h}}|E_{\mathbf{h}})$ is, of course, identical with $P(E_{2\mathbf{h}}, E_{\mathbf{h}})$, but the new notation emphasizes that we shall be concerned with the probability distribution of $E_{2\mathbf{h}}$ as a function of $E_{\mathbf{h}}$.

We first make some remarks about the mathematical form of the conditional distribution, and about notation.

3·21.—If $F(x, y)$ is a bivariate probability distribution we shall denote the conditional distribution of x on y by $F(x|y)$ or by $F_y(x)$, which may be thought of as a set of univariate distributions depending on the parameter y. More specifically, if we wish to emphasize that the various moments and characteristics of $F_y(x)$ depend on y, we include y as a subscript, e.g. the conditional mean and variance of x are written $\langle x\rangle_y$ and $\sigma_y^2(x)$ respectively. More generally we shall denote the sth moment of x by $\mu_s'(x)_y$ and the sth moment about the mean by $\mu_s(x)_y$:

$$\mu_s'(x)_y \equiv \langle x^s\rangle_y = \frac{\int_{-\infty}^{\infty} x^s F(x, y)\,dx}{\int_{-\infty}^{\infty} F(x, y)\,dx}\,, \tag{3·6}$$

$$\mu_s(x)_y \equiv \langle (x-\langle x\rangle_y)^s\rangle\,. \tag{3·7}$$

We note that

$$\sigma^2 \equiv \mu_2 = \mu_2'-\langle x\rangle^2\,, \tag{3·8}$$

$$\mu_3 = \mu_3'-3\mu_2\langle x\rangle-\langle x\rangle^3\,. \tag{3·9}$$

If any of the characteristics is plotted as a function of y, we shall speak of the curve of regression of the mean of x on y, the regression of the variance, etc.

When the conditional probability $F_y(x)$ is not Gaussian, it may be expanded as a Gram-Charlier series, as in § 1:

$$F_y(x) = \varphi_y(x)+\frac{1}{3!}\mu_3(x)_y\frac{\partial^3\varphi_y(x)}{\partial x^3}+\ldots\,, \tag{3·10}$$

where

$$\varphi_y(x) = \frac{1}{\sigma_y(2\pi)^{\frac{1}{2}}}\exp\left[-\frac{(x-\langle x\rangle_y)^2}{2\sigma_y^2}\right]. \tag{3·11}$$

The third moment in (3·10) determines the skewness of the distribution. The higher-order moments (or, more strictly, the cumulants; see equations (1·15) and (1·18)) reflect the more subtle deviations from the Gaussian curve. When these can be neglected the position of the mode (most probable value) is given by

$$\langle x\rangle+\mu_3/2\sigma^2\,. \tag{3·12}$$

All the probability distributions of interest in this and later sections are adequately described by the mean, variance and third moment, since it will turn out that the latter are of order $1/N^{\frac{1}{2}}$, $1+O(1/N)$, and $1/N^{3/2}$ respectively. This follows from the fact that the random variables we are dealing with (i.e. the structure factors) are themselves sums of N random variables (the atom contributions). Furthermore, because we have represented the probability distributions by a truly asymptotic series, the actual values we shall obtain for the characteristics will be strictly correct to the degree of approximation we choose.

The probability distribution (3·4) is of the form

$$F(x, y) = \varphi(x, y)+\sum_{i+j\geq 3}\beta_{ij}\frac{\partial^{i+j}}{\partial x^i\,\partial y^j}\varphi(x, y) \tag{3·13}$$

$$= \varphi(x, y)\left[1+\sum_{i+j\geq 3}(-1)^{i+j}\beta_{ij}H_i(x)H_j(y)\right], \tag{3·14}$$

where

$$\varphi(x, y) = \varphi(x)\varphi(y) = (2\pi)^{-1}\exp\left[-\tfrac{1}{2}(x^2+y^2)\right] \tag{3·15}$$

is the bivariate normal distribution. The moments (3·6) will then be given by

$$\mu_s'(x)_y = \frac{\int_{-\infty}^{\infty} x^s\left[\varphi(x, y)+\sum\beta_{ij}\dfrac{\partial^{i+j}\varphi(x, y)}{\partial x^i\,\partial y^j}\right]dx}{\int_{-\infty}^{\infty}\left[\varphi(x, y)+\sum\beta_{ij}\dfrac{\partial^{i+j}\varphi(x, y)}{\partial x^i\,\partial y^j}\right]dx} \tag{3·16}$$

$$= \frac{I_{00}^s+\sum_{i+j\geq 3}\beta_{ij}I_{ij}^s}{I_{00}^0+\sum_{i+j\geq 3}\beta_{ij}I_{ij}^0}\,, \tag{3·17}$$

where

$$I_{ij}^s \equiv \int_{-\infty}^{\infty} x^s\frac{\partial^{i+j}\varphi(x, y)}{\partial x^i\,\partial y^j}\,dx \tag{3·18}$$

(note that on the left-hand side s is a superfix, not a power). The I^s_{ij} may be calculated once and for all, as in Appendix E. The β_{ij} will be known (in terms of the cumulants) for the particular probability distribution of interest by comparing the latter with (3·14). They may then be substituted in (3·17) to give the desired moment. The moments about the mean are then found by using equations (3·8) and (3·9). The calculations are rather lengthy but straightforward.

3·22.—In this way we have determined the characteristics of the distribution $P(E_{2\mathbf{h}}|E_{\mathbf{h}})$. Writing $E_{2\mathbf{h}} = E_1$, $E_{\mathbf{h}} = E_2$, the results to the order $1/N^{3/2}$ are:

Mean: $\langle E_{2\mathbf{h}}\rangle_{E_\mathbf{h}}$

$$= \frac{\dfrac{1}{2N^{\frac{1}{2}}}(E^2_{\mathbf{h}}-1)-\dfrac{1}{N^{3/2}}\left[\frac{1}{3}H_4(E_{\mathbf{h}})+\frac{1}{16}H_6(E_{\mathbf{h}})\right]}{1-\dfrac{1}{8N}H_4(E_{\mathbf{h}})} . \quad (3·19)$$

$$\text{Variance: } \sigma^2_{E_\mathbf{h}}(E_{2\mathbf{h}}) = \frac{1-\dfrac{1}{4N}H^2_2(E_{\mathbf{h}})}{1-\dfrac{1}{4N}H_4(E_{\mathbf{h}})} \quad (3·20)$$

$$\fallingdotseq 1+1/2N-E^2_{\mathbf{h}}/N . \quad (3·20a)$$

(Equation (3·20a) is obtained by expanding the denominator by the binomial theorem, which will be justified over the range of E we shall be concerned with).

Third moment about the mean:

$$\mu_3(E_{2\mathbf{h}})_{E_\mathbf{h}} = \frac{1}{N^{3/2}}\left[-\tfrac{3}{2}H_2(E_{\mathbf{h}})+H_4(E_{\mathbf{h}})+\tfrac{5}{16}H_6(E_{\mathbf{h}})\right] . \quad (3·21)$$

In Fig. 2 we have plotted the regression curves of the various characteristics for the case $N = 10$, computed from the above equations.

If we include only terms up to the order $1/N^{\frac{1}{2}}$, we see that $E_{2\mathbf{h}}$ has a mean of $\dfrac{1}{2N^{\frac{1}{2}}}(E^2_{\mathbf{h}}-1)$, a variance of unity and a third moment of zero. (The regression curves to this approximation are also plotted in Fig. 2.) Hence, taking account of (3·10) and (3·11), it is clear that the distribution is a simple Gaussian one, and exactly equivalent to Cochran & Woolfson's (1955) equation (3·8) when the latter is rewritten in terms of E's rather than U's. We thus see that Cochran & Woolfson's result is an approximation correct to the order $1/N^{\frac{1}{2}}$, and, as already mentioned, will give a good representation of the probability distribution so long as $E_{\mathbf{h}}$ and $E_{2\mathbf{h}}$ are small compared with their maximum value, $N^{\frac{1}{2}}$, as is commonly the case in practice when N is large (see § 5·2 below). When $E_{\mathbf{h}}$ is large, the approximation is inadequate, as may be seen from Fig. 2, and indeed entirely misleading, mathematically speaking. Thus when $|E_{\mathbf{h}}| \sim N^{\frac{1}{2}}$, the variance is no longer unity but tends to zero, as required by the Harker–Kasper inequality. Our equation (3·20), which is the correct approximation to the order $1/N^{3/2}$, does indicate this effect correctly, as will be obvious from (3·20a).

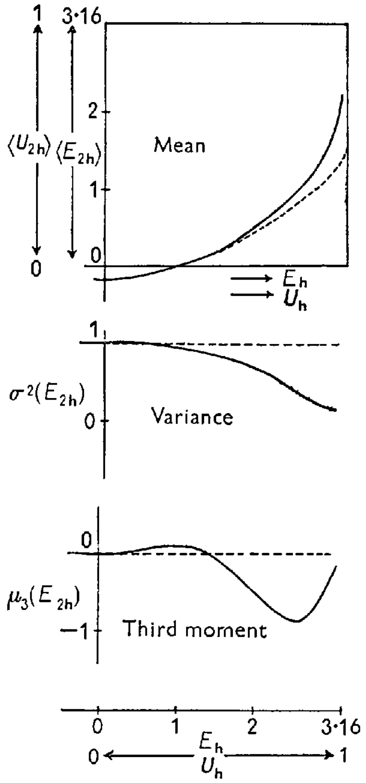

Fig. 2. Regression of the mean, variance and third moment of $E_{2\mathbf{h}}$ on $E_{\mathbf{h}}$ for $P\bar{1}$, $N = 10$.

The full lines are calculated from the expressions (3·19), (3·20) and (3·21) respectively, which are approximations correct to the order $1/N^{3/2}$. The broken lines represent the corresponding approximations to the order $1/N^{\frac{1}{2}}$, and would follow, for instance, from the theory of Cochran & Woolfson (1955).

This means, in turn, that Cochran & Woolfson's (1955) expression (3·10) for the probability $P_+(E_{2\mathbf{h}})$ that $E_{2\mathbf{h}}$ is positive gives an underestimate for this probability when $E_{\mathbf{h}}$ or $E_{2\mathbf{h}}$ is large, since it is based on their equation (3·8) in which the variance has been taken as a constant (unity) not depending on the magnitude of $E_{\mathbf{h}}$. The correct expression for $P_+(E_{2\mathbf{h}})$ may be found from our equation (3·4) for $P(E_{2\mathbf{h}}, E_{\mathbf{h}})$ by using the very simple but powerful result of Bertaut (1955a, 1956b) that

$$P_{\pm} = \tfrac{1}{2}\pm\tfrac{1}{2}P_o/P_e , \quad (3·22)$$

where P_o and P_e refer to the terms in the probability expression which are respectively odd and even in the structure factor involved ($E_{2\mathbf{h}}$ in this case). This gives

$$P_{\pm}(E_{2\mathbf{h}}) = \tfrac{1}{2}\pm\tfrac{1}{2}$$
$$\times\left[\frac{\dfrac{1}{2N^{\frac{1}{2}}}E_{2\mathbf{h}}(E_{\mathbf{h}}^2-1)+\text{terms in }\dfrac{1}{N^{3/2}}\text{ in (3·4)}+\ldots}{1+\text{terms in }1/N\text{ in (3·4)}+\ldots}\right]. \quad (3·23)$$

When all terms beyond that in $1/N^{\frac{1}{2}}$ are dropped, the result of Hauptman & Karle (1953, p. 40) is obtained.

We have not computed $P_{+}(E_{2\mathbf{h}})$ for the example $N = 10$ since its behaviour is quite clear from the general expression $P(E_{2\mathbf{h}}, E_{\mathbf{h}})$ plotted in Fig. 1. For large values of $E_{\mathbf{h}}$ we should expect the probability distribution to pass smoothly into the range of the Harker–Kasper inequality (where $P_{+}(E_{2\mathbf{h}}) = 1$), as shown in our Fig. 1. It is thus clear that when a sign is 'almost but not quite' determined by the inequality we can quite safely take it as so determined. These remarks are made here because they are also applicable to the more important case of the Sayre–Cochran–Zachariasen sign relation (see (4·1) below)

$$S(\mathbf{h}) \approx S(\mathbf{h}')S(\mathbf{h}+\mathbf{h}') . \quad (3·24)$$

Cochran & Woolfson's (1955) expression (4·16) for the probability that this holds is again an underestimate when the structure factors involved are large.

3·3. *The regression of $E_{\mathbf{h}}$ on $E_{2\mathbf{h}}$*

The result (3·23) gives the probability of the truth of the sign-determining relation

$$S(2\mathbf{h}) \approx S(E_{\mathbf{h}}^2-1) \quad (3·25)$$

and would formally close this section. However, as an introduction to the wider discussion of the phase problem, it is instructive to derive the characteristics of the conditional distribution $P(E_{\mathbf{h}}|E_{2\mathbf{h}})$. $\langle E_{\mathbf{h}}\rangle_{E_{2\mathbf{h}}}$ is of course zero, and we find that, to the order $1/N^{3/2}$,

$$\langle E_{\mathbf{h}}^2\rangle_{E_{2\mathbf{h}}} = \frac{1+\dfrac{1}{N^{\frac{1}{2}}}E_{2\mathbf{h}}-\dfrac{1}{8N}H_4(E_{2\mathbf{h}}) -\dfrac{1}{2N^{3/2}}H_3(E_{2\mathbf{h}})-\dfrac{1}{8N^{3/2}}H_5(E_{2\mathbf{h}})}{1-\dfrac{1}{8N}H_4(E_{2\mathbf{h}})}$$
$$= 1+\frac{1}{N^{\frac{1}{2}}}E_{2\mathbf{h}} , \quad (3·26)$$

where, in the second step, we have made use of the result (cf. equation (4·20*a*) below)

$$\frac{x-A\,\{4H_3(x)+H_5(x)\}}{1-AH_4(x)} = x . \quad (3·27)$$

Hence we have shown that

$$\langle E_{\mathbf{h}}^2-1\rangle = \frac{1}{N^{\frac{1}{2}}}E_{2\mathbf{h}} \quad (3·28)$$

is correct not only to order $1/N^{\frac{1}{2}}$, as might at first be thought, but to order $1/N^{3/2}$. We would surmize that it may be true to an even higher order of approximation. The average in (3·28) means an average over an ensemble of all possible structures having the same number of atoms. Statistically, we might say (cf. Bullough & Cruickshank, 1955) that $N^{\frac{1}{2}}(E_{\mathbf{h}}^2-1)$ provides an estimate of $E_{2\mathbf{h}}$. A measure of the reliability of such an estimate is provided by the variance of $E_{\mathbf{h}}^2$ about its mean. By the methods of § 3·21 and Appendix *E*, we can calculate $\langle E_{\mathbf{h}}^4\rangle_{E_{2\mathbf{h}}}$, and we find that the required variance

$$\langle E_{\mathbf{h}}^4\rangle-\langle E_{\mathbf{h}}^2\rangle^2 = 2+\frac{4}{N^{\frac{1}{2}}}E_{2\mathbf{h}}-\frac{2}{N}(3-E_{2\mathbf{h}}^2) . \quad (3·29)$$

Hence the estimate $N^{\frac{1}{2}}(E_{\mathbf{h}}^2-1)$ of $E_{2\mathbf{h}}$ has a variance of

$$2N+4N^{\frac{1}{2}}E_{2\mathbf{h}}-2(3-E_{2\mathbf{h}}^2) . \quad (3·30)$$

When $E_{2\mathbf{h}}$ is small, the dominating term in the last expression will be the first, and we see that in this case Bullough & Cruickshank's approximate trigonometric method for calculating the variance is adequate. However, when $E_{2\mathbf{h}}$ is large and positive, the variance is much greater than would appear from the first term $2N$. In this case, we should not use $N^{\frac{1}{2}}(E_{\mathbf{h}}^2-1)$ as an estimate of $E_{2\mathbf{h}}$ but the accurate formula of Cochran (1954) given later (equation (5·3)). When $E_{2\mathbf{h}}$ is large and negative, we see that $N^{\frac{1}{2}}(E_{\mathbf{h}}^2-1)$ is a very good estimate of $E_{2\mathbf{h}}$, as is to be expected, since in this case $E_{\mathbf{h}}$ is likely to be small.

3·4. *The determination of signs from intensities in space groups of higher symmetry*

In space groups other than $P\bar{1}$, the moment m_{12} will be non-vanishing for other structure factors besides the pair $E_{2\mathbf{h}}$ and $E_{\mathbf{h}}$. Thus in $P2_1/a$, $m_{12} \neq 0$ for the pair $E_{2h,0,2l}$ and E_{hkl} (Hauptman & Karle, 1953, Table 10), and, to the order $1/N^{\frac{1}{2}}$,

$$P_{+}(E_{2h,0,2l}) = \tfrac{1}{2}+\frac{(-1)^{h+k}}{2\sqrt{2}N^{\frac{1}{2}}}|E_{2h,0,2l}|(E_{h,k,l}^2-1) . \quad (3·31)$$

Since *k* is a free index one may sum over it (Hauptman & Karle, 1953, equation (4·15)), so that

$$S(E_{2h,0,2l}) \approx S\left\{\sum_{k}(-1)^{h+k}(E_{hkl}^2-1)\right\}. \quad (3·32)$$

Cochran (1954) has shown that, quite irrespective of probability theory, the deep origin of (3·32) is to be found in the exact formula expressing the relation between the Patterson–Harker section at $y = \frac{1}{2}$ and the projected electron density on $y = 0$:

$$E_{2h,0,2l} = N^{\frac{1}{2}}\overline{(-1)^{h+k}(E_{hkl}^2-1)}^{k} \quad (3·33)$$
$$= N^{\frac{1}{2}}\frac{1}{K}\sum_{k=0}^{K}(-1)^{h+k}(E_{hkl}^2-1) , \quad (3·34)$$

where *K* is the number of terms in the summation.

(Vand & Pepinsky (1953, 1954) had earlier pointed out the connexion of the results of the probability theory with the Harker section.) Cochran has discussed the physical basis of the conditions under which (3·34) gives a reliable estimate of the sign of $E_{2h,0,2l}$.

From the point of view of statistical theory, by analogy with the discussion in § 3·3, one of the terms $(-1)^{h+k}N^{\frac{1}{2}}(E^2_{hkl}-1)$ provides an estimate of $E_{2h,0,2l}$ with a variance of the form

$$aN+N^{\frac{1}{2}}f(E_{2h,0,2l})+b\ , \tag{3·35}$$

where a and b are constants. Except when $E_{2h,0,2l}$ is large, the first term will dominate. Now on reasonable assumptions (see § 5·2 below), the number of terms K in the summation in (3·34) will be proportional to $N^{\frac{1}{3}}$, so that the variance of the estimate of the $E_{2h,0,2l}$ through the formula (3·34) will be of order $N^{2/3}$. Hence, as the number of atoms in the structure increases, the reliability of the formula will decrease. The physical reason is the increasing number of 'non-Harker' peaks to be found in the Patterson section at $y=\frac{1}{2}$. Only when $E_{2h,0,2l}$ is large and the second and third terms in (3·35) are important can we expect the sign indication to be reliable.

The argument in the last paragraph may appear to be rather condensed, but the reader will easily be able to fill in the details for himself after reading § 5·2 below, where other formulae involving a summation over a set of contributing terms are discussed at some length.

4. Relations between the signs of structure factors

4·1. *Probability distribution of* $E_{\mathbf{h}}$, $E_{\mathbf{h}'}$ *and* $E_{\mathbf{h}+\mathbf{h}'}$

It is well known that when the unitary structure factors $U_{\mathbf{h}}$, $U_{\mathbf{h}'}$ and $U_{\mathbf{h}+\mathbf{h}'}$ are large enough their signs are probably related by the Sayre–Cochran–Zachariasen (S.C.Z.) relation

$$S(\mathbf{h})\approx S(\mathbf{h}')S(\mathbf{h}'+\mathbf{h}) \tag{4·1}$$

or

$$S(\mathbf{h})S(\mathbf{h}')S(\mathbf{h}'+\mathbf{h})>0\ . \tag{4·1a}$$

In this section we shall find an exact expression for the probability that this relation is true.

The notation above is in common usage but hides the symmetry between the three structure factors entering into relation (4·1). We shall accordingly denote the indices by $\mathbf{h}_1$, $\mathbf{h}_2$ and $\mathbf{h}_3$, where

$$\mathbf{h}_1\pm\mathbf{h}_2\pm\mathbf{h}_3=0\ . \tag{4·2}$$

Remembering that, for $P\bar{1}$, $F(\mathbf{h})=F(-\mathbf{h})$, equation (4·2) means that if $\mathbf{h}_1$ and $\mathbf{h}_2$ are identified with $\mathbf{h}$ and $\mathbf{h}'$ respectively, then $\mathbf{h}_3$ may be identified either with $\mathbf{h}+\mathbf{h}'$ as in (4·1) or else $\mathbf{h}-\mathbf{h}'$. It is then easy to show that the joint probability distribution of $U_{\mathbf{h}}$, $U_{\mathbf{h}'}$, $U_{\mathbf{h}+\mathbf{h}'}$ and $U_{\mathbf{h}-\mathbf{h}'}$ factorizes into a product of the two probability distributions $P(U_{\mathbf{h}}, U_{\mathbf{h}'}, U_{\mathbf{h}+\mathbf{h}'})$ and $P(U_{\mathbf{h}}, U_{\mathbf{h}'}$ and $U_{\mathbf{h}-\mathbf{h}'})$, which have exactly the same form. Thus, while the probability of the truth of relation (4·1) is formally independent of $U_{\mathbf{h}-\mathbf{h}'}$, in the practical sign determination of $U_{\mathbf{h}}$ one considers both $U_{\mathbf{h}+\mathbf{h}'}$ and $U_{\mathbf{h}-\mathbf{h}'}$, and the larger of the two will dominate. In what follows $\mathbf{h}_3$ is to be understood as referring to either $\mathbf{h}+\mathbf{h}'$ or $\mathbf{h}-\mathbf{h}'$, without prejudice.

The probability distribution $P(E_1, E_2, E_3)$ may be easily written down, following the rules given in § 2. The mixed moments

$$m_{pqr}=\langle\xi(\mathbf{h}_1)^p\xi(\mathbf{h}_2)^q\xi(\mathbf{h}_3)^r\rangle \tag{4·3}$$

subject to $\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3=0$

are found by Bertaut's method in Appendix D for $P\bar{1}$. The moment-generating function to the order $1/N^{3/2}$ is given by (2·15) and (2·17):

$$\begin{aligned}\exp\,&[-\tfrac{1}{2}(u_1^2+u_2^2+u_3^2)]\,\mathscr{M}(u_1, u_2, u_3)\\ &=1+\frac{1}{t^{\frac{1}{2}}}\left[\frac{\lambda_{111}}{1!\,1!\,1!}\,u_1u_2u_3\right]\\ &+\frac{1}{t}\left[\frac{\lambda_{400}}{4!\,0!\,0!}\,u_1^4+\text{cyc.}\right]+\frac{1}{2t}\left[\frac{\lambda_{111}}{1!\,1!\,1!}\,u_1u_2u_3\right]^2\\ &+\frac{1}{t^{3/2}}\left[\frac{\lambda_{113}}{1!\,1!\,3!}\,u_1u_2u_3^3+\text{cyc.}\right]\\ &+\frac{1}{t^{3/2}}\left[\frac{\lambda_{111}}{1!\,1!\,1!}\,u_1u_2u_3\right]\left[\frac{\lambda_{400}}{4!\,0!\,0!}\,u_1^4+\text{cyc.}\right]\\ &+\frac{1}{6t^{3/2}}\left[\frac{\lambda_{111}}{1!\,1!\,1!}\,u_1u_2u_3\right]^3;\end{aligned} \tag{4·4}$$

here 'cyc.' stands for the cyclic permutations of the indices; for example in the bracket in $1/t$ the terms are

$$\frac{1}{4!}\,(\lambda_{400}u_1^4+\lambda_{040}u_2^4+\lambda_{004}u_3^4)\ .$$

Inverting, and substituting values for λ_{pqr} from (D·11), we find for the joint probability distribution of E_1 $(=E_{\mathbf{h}})$, E_2 $(=E_{\mathbf{h}'})$, E_3 $(=E_{\mathbf{h}+\mathbf{h}'})$

$$\begin{aligned}P(E_1, E_2, E_3)&=\frac{1}{(2\pi)^{3/2}}\exp\,[-\tfrac{1}{2}(E_1^2+E_2^2+E_2^3)]\\ &\times\Big\{1+\frac{1}{N^{\frac{1}{2}}}\,E_1E_2E_3-\frac{1}{8N}\,[H_4(E_1)+\text{cyc.}]\\ &+\frac{1}{2N}\,H_2(E_1)H_2(E_2)H_2(E_3)\\ &-\frac{1}{2N^{3/2}}\,[H_1(E_1)H_1(E_2)H_3(E_3)+\text{cyc.}]\\ &-\frac{1}{8N^{3/2}}\,[H_1(E_1)H_1(E_2)H_5(E_3)+\text{cyc.}]\\ &+\frac{1}{6N^{3/2}}\,H_3(E_1)H_3(E_2)H_3(E_3)+\dots\Big\}\ ,\end{aligned} \tag{4·5}$$

where 'cyc.' now stands for the cyclic permutation of the arguments of the Hermite polynomials, so that the expression is symmetrical in E_1, E_2 and E_3, as required. To obtain a general idea of the form of the distribution we have computed the sections of $P(E_1, E_2, E_3)$ at $E_1 = 0$ and $E_1 = 2$ for the case of $N = 10$ equal atoms. The results are plotted in Fig. 3.

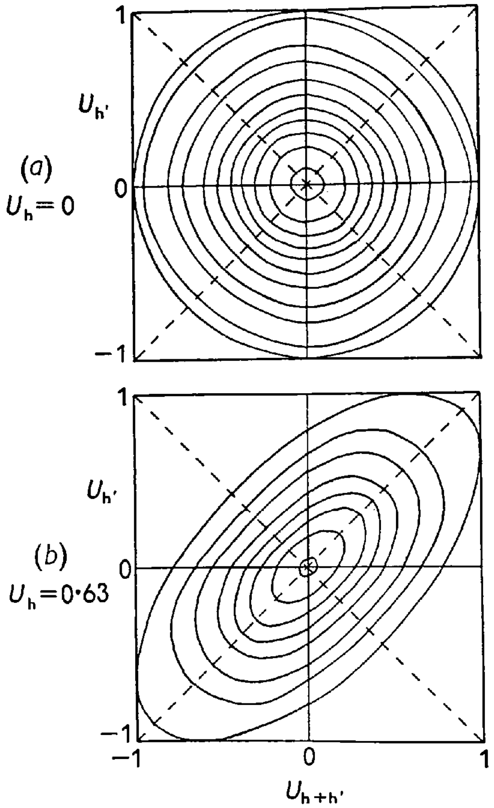

Fig. 3. Two sections of the joint probability distribution of $U_{\mathbf{h}}$, $U_{\mathbf{h}'}$ and $U_{\mathbf{h}+\mathbf{h}'}$ for space group $P\bar{1}$, calculated from equation (4·5) for the particular case of $N = 10$. (*a*) and (*b*) show the distribution of $U_{\mathbf{h}'}$ and $U_{\mathbf{h}+\mathbf{h}'}$ for $U_{\mathbf{h}} = 0$ and $U_{\mathbf{h}} = 0{\cdot}63$ (i.e., $E_{\mathbf{h}} = 2$) respectively.

The contours in (*a*), reading from outwards in, are 0, 0·01, 0·05, 0·1, and then in steps of 0·1 to 0·8; in (*b*) they are 0, 0·010, 0·025, and then in steps of 0·025 to 0·150. These values have to be divided by $(2\pi)^{3/2}$ to obtain normalized values.

The outermost zero contour is obtained not from the probability theory, but from the inequality (4·7).

The contours of P fit very smoothly into the outermost surface $P = 0$, which is obtained quite independently of the probability theory by means of an inequality to be described in the next sub-section. This shows that our theory (including terms up to $1/N^{3/2}$) gives the correct form of the probability distribution, even for relatively high values of E_1, E_2 and E_3. The approximation obtained by including the terms only as far as $1/N^{\frac{1}{2}}$ gives an accurate description only in a relatively small range near the origin $E_1, E_2 \ll N^{\frac{1}{2}}$. This will be clear when we study the form of (4·5) in more detail later.

4·2. *An inequality relating* $U_{\mathbf{h}}$, $U_{\mathbf{h}'}$ *and* $U_{\mathbf{h}+\mathbf{h}'}$

Karle & Hauptman (1950) have described a very general method of constructing inequalities among a set of m structure factors $U_{\mathbf{h}_i}$ whose indices are linearly related. The relevance of this to the calculation of the probability distribution of such a set, is that the limiting case of the inequality (i.e. equality) demarcates a forbidden region in the m-dimensional space formed by the axes $U_{\mathbf{h}_i}$, irrespective of the number of atoms in the unit cell. An example of this has already been given in the case of $P(U_{2\mathbf{h}}, U_{\mathbf{h}})$, where the curve $U_{\mathbf{h}}^2 = \frac{1}{2}(1+U_{2\mathbf{h}})$ gives the outermost boundary of the possible range of the structure factors.

For the particular case in hand, we can apply Karle & Hauptman's (1950) equation (34), to obtain

$$|NU_{\mathbf{h}+\mathbf{h}'} - NU_{\mathbf{h}}U_{\mathbf{h}'}| \le \frac{1}{N}(N^2-N^2U_{\mathbf{h}}^2)^{\frac{1}{2}}(N^2-N^2U_{\mathbf{h}'}^2)^{\frac{1}{2}}\,. \tag{4·6}$$

Cancelling the factor N, squaring, and writing in a form symmetrical in the three structure factors, we obtain finally the desired inequality

$$2U_{\mathbf{h}}U_{\mathbf{h}'}U_{\mathbf{h}+\mathbf{h}'} \ge U_{\mathbf{h}}^2 + U_{\mathbf{h}'}^2 + U_{\mathbf{h}+\mathbf{h}'}^2 - 1$$

or

$$U_{\mathbf{h}}U_{\mathbf{h}'}U_{\mathbf{h}+\mathbf{h}'} \ge \tfrac{1}{2}(U_{\mathbf{h}}^2 + U_{\mathbf{h}'}^2 + U_{\mathbf{h}+\mathbf{h}'}^2) - \tfrac{1}{2}\,. \tag{4·7}$$

If

$$U_{\mathbf{h}}^2 + U_{\mathbf{h}'}^2 + U_{\mathbf{h}+\mathbf{h}'}^2 \ge 1\,, \tag{4·8}$$

then $U_{\mathbf{h}}U_{\mathbf{h}'}U_{\mathbf{h}+\mathbf{h}'} \ge 0$, which means that the S.C.Z. relation (4·1) must certainly be true in the region outside the unit sphere. This condition (4·8) does not appear to have been noticed before (but compare Cochran (1952)). It would seem to be of practical use only for very simple crystals. Not all the region outside the unit sphere is permissible, only that contained within the limits set by (4·7).

For brevity, writing

$$U_{\mathbf{h}} = x,\quad U_{\mathbf{h}'} = y,\quad U_{\mathbf{h}+\mathbf{h}'} = z\,, \tag{4·9}$$

we see that the allowed region of variation of x, y and z is bounded by the surface

$$xyz = \tfrac{1}{2}(x^2+y^2+z^2) - \tfrac{1}{2}\,. \tag{4·10}$$

Fig. 4 shows a photograph of a model of this surface inscribed in the cube $-1 \le \left\{\begin{matrix} x \\ y \\ z \end{matrix}\right\} \le 1$. It looks like a tetrahedron with curved faces, and in fact we might call it a 'parabolic tetrahedron' since a section containing a face diagonal of the cube is a parabola (in contrast to a triangle in the case of a tetrahedron). It is easy to show that sections of the surface parallel to a cube face are ellipses. For a section at $x > 0$, say, the major semi-axis has a value $(1+x)^{\frac{1}{2}}$, and lies in the direction $z = y$; the length of minor semi-axis is $(1-x)^{\frac{1}{2}}$ and lies in the direction $z = -y$ (see outermost contours in Fig. 3). When x is negative, the directions of

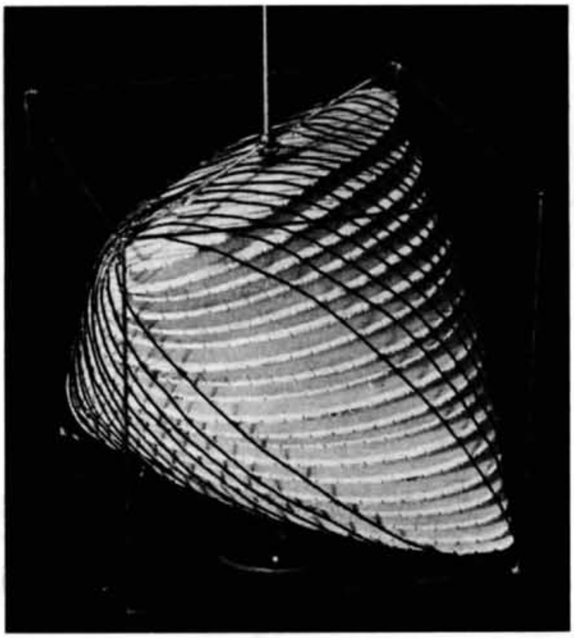

Fig. 4. A model of the surface $xyz = \frac{1}{2}(x^2+y^2+x^2)-\frac{1}{2}$ set in the cube $-1 \leq x \leq 1$, $-1 \leq y \leq 1$, $-1 \leq z \leq 1$.

Sections of the surface parallel to a cube face are ellipses, and two of the three such sets of sections are indicated, one by pieces of cardboard and the other by wires.

The surface encloses the allowed domain of variation of $U_{\mathbf{h}} = x$, $U_{\mathbf{h}'} = y$, $U_{\mathbf{h}+\mathbf{h}'} = z$.

the major and minor axes are interchanged. It is also easy to show that a x section meets the sides of the bounding square at ($y = 1$, $z = x$) and ($y = x$, $z = 1$).

Kitaigorodski (1954) has proved that the product $U_{\mathbf{h}}U_{\mathbf{h}'}U_{\mathbf{h}+\mathbf{h}'}$ has a minimum* value of $-\frac{1}{8}$. It is easy to derive this result from the more general inequality (4·7). Taking a section at $x > 0$, as in Fig. 3(b), the minimum value of yz occurs at the points of intersection of the bounding ellipse and the minor semi-axis $y = -z$, i.e. at $\pm y = \mp z = \{\frac{1}{2}(1-x)\}^{\frac{1}{2}}$. Hence, for fixed x, the minimum value of yz is $-\frac{1}{2}(1-x)$ and that of xyz is $-\frac{1}{2}(x-x^2)$. It then follows that xyz will have an absolute minimum at $x = \frac{1}{2}$ of value $-\frac{1}{8}$. By symmetry, minima will occur at the cyclic permutations of $(\frac{1}{2}, \frac{1}{2}, -\frac{1}{2})$.

We note, incidentally, that it can be shown in a similar way from the Harker–Kasper inequality $U_{2\mathbf{h}} \geq 2U_{\mathbf{h}}^2-1$, that the minimum value of $U_{2\mathbf{h}}U_{\mathbf{h}}^2$ is also equal to $-\frac{1}{8}$.

4·3. *Characteristics of the probability distribution* $P(E_{\mathbf{h}}, E_{\mathbf{h}'}, E_{\mathbf{h}+\mathbf{h}'})$

To get some insight into the nature of the distribution expressed by (4·5), and also to obtain an expression for the probability that the S.C.Z. relation (4·1) holds, it is necessary to hold either one or two of the variables $E_{\mathbf{h}} = E_1$, $E_{\mathbf{h}'} = E_2$ and $E_{\mathbf{h}+\mathbf{h}'} = E_3$ fixed, and to calculate the corresponding conditional distributions.

The conditional probability $P(E_1|\,E_2, E_3)$ may be expanded in the form of a Gram-Charlier series (cf. (3·10)).

$$P_{E_2, E_3}(E_1) = (\sqrt{(2\pi)}\sigma_{E_2, E_3})^{-1} \times \exp\left[-(E_1-\langle E_1\rangle_{E_2, E_3})^2/2\sigma^2_{E_2, E_3}\right] + \text{term in } \mu_3 + \ldots\ldots \quad (4\cdot11)$$

In what follows the third moment about the mean is very small or zero (as the reader may check for himself by drawing the curve corresponding to a line of constant E_2 or E_3 in Fig. 3). We shall thus omit the corresponding term completely, so that we are dealing only with Gaussian distributions. It is only necessary to calculate the mean $\langle E_1\rangle_{E_2, E_3}$ and the variance $\sigma^2_{E_2, E_3}(E_1)$ as functions of E_2 and E_3. This may be done by means of the operations outlined in Appendix E.

We find, to the order $1/N^{3/2}$, that

$$\langle E_1\rangle_{E_2, E_3} = \frac{\frac{1}{N^{\frac{1}{2}}}E_2E_3+\frac{1}{2N^{3/2}}[-H_3(E_2)H_1(E_3)-H_1(E_2)H_3(E_3)] + \frac{1}{8N^{3/2}}[-H_5(E_2)H_1(E_3)-H_1(E_2)H_5(E_3)]}{1-\frac{1}{8N}[H_4(E_2)+H_4(E_3)]} \quad (4\cdot12)$$

and

$$\sigma^2_{E_2, E_3}(E_1) = \frac{1-\frac{1}{4N}[H_4(E_2)+H_4(E_3)]+\frac{1}{N}(1-E_2^2-E_3^2)}{1-\frac{1}{4N}[H_4(E_2)+H_4(E_3)]}, \quad (4\cdot13)$$

$$\risingdotseq 1+\frac{1}{N}-\frac{1}{N}(E_2^2+E_3^2), \quad (4\cdot13a)$$

expanding the denominator by the binomial theorem. These functions are plotted for the case of $N = 10$ equal atoms in Figs. 5(a) and 5(b). They are, of course, symmetric in E_2 and E_3, but cannot be expressed only in terms of the product E_2E_3 and its higher powers. The expressions involve products of mixed powers of E_2 and E_3. This is illustrated by a comparison of Figs. 5(a) and 5(c). In the latter we have simply plotted $\frac{1}{N^{\frac{1}{2}}}E_2E_3$, which is the correct approximation of order $1/N^{\frac{1}{2}}$ to the mean $\langle E_1\rangle$. To the same order the variance is unity, and if these approximations are substituted in (4·11) we obtain (with a slight reversion in notation)

$$P(E_{\mathbf{h}}) = (2\pi)^{-\frac{1}{2}}\exp\left[-\tfrac{1}{2}(E_{\mathbf{h}}-\frac{1}{N^{\frac{1}{2}}}E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'})^2\right], \quad (4\cdot14)$$

which is the expression originally obtained by Woolfson (1954). It will now be clear that his results are

---

* 'Maximum negative' would give the correct emphasis.

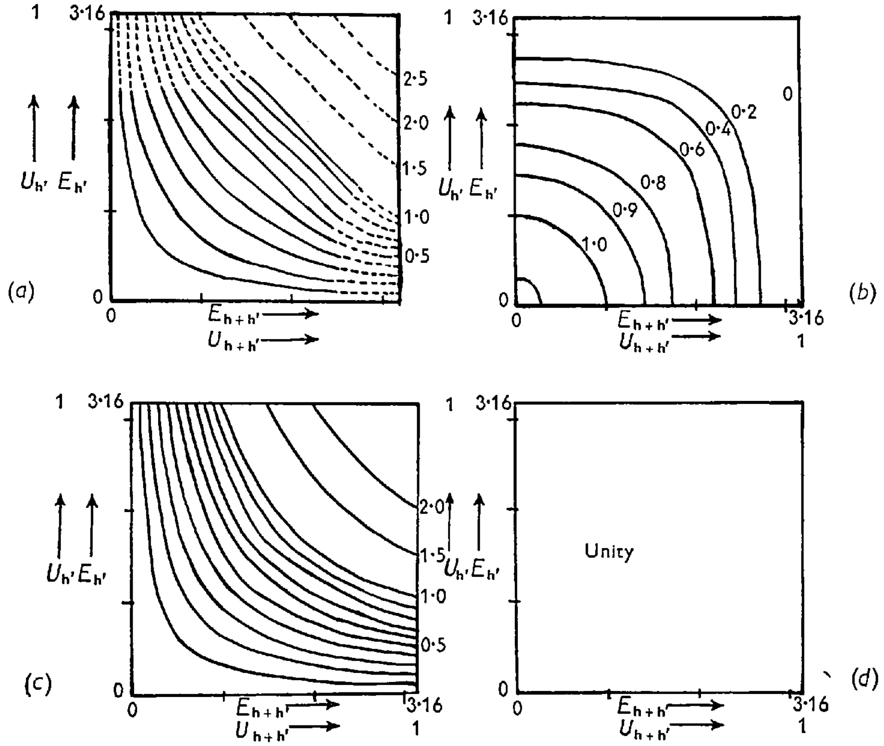

Fig. 5. Regression of (*a*) the mean, and (*b*) the variance, of $E_{\mathbf{h}}$ on $E_{\mathbf{h}'}$ and $E_{\mathbf{h}+\mathbf{h}'}$, for $P\bar{1}$, $N = 10$, calculated from equations (4·12) and (4·13) respectively. The broken lines in (*a*) are obtained by extrapolation, since the approximation used (of order $1/N^{3/2}$) breaks down at very large values of E. (*c*) and (*d*) show the mean ($= 1/N^{\frac{1}{2}}\, E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'}$) and variance ($= 1$) respectively for the approximation to the order $1/N^{\frac{1}{2}}$ (Cochran & Woolfson, 1955).

accurate only to the order $1/N^{\frac{1}{2}}$, and will apply only over the range of values of E small compared with $N^{\frac{1}{2}}$. The approximation correct to order $1/N^{3/2}$ is given by substituting (4·12) and (4·13) in (4·11). We have not plotted out the distribution $P(E_{\mathbf{h}})$ for our example of 10 atoms in the cell, since, as already stressed, it is not simply a function of the product $E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'}$. We should have to plot a family of distributions for various combinations of $E_{\mathbf{h}'}$ and $E_{\mathbf{h}+\mathbf{h}'}$. The form of this distribution, however, is clear from Figs. 5(*a*) and 5(*b*). As $E_{\mathbf{h}'}$ and $E_{\mathbf{h}-\mathbf{h}'}$ increase, the mean increases rather less slowly at first than $\frac{1}{N^{\frac{1}{2}}}\, E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'}$ along the diagonal, but the variance falls fairly rapidly below the value unity, the contours being approximately spherical (cf. equation (4·13*a*)). This means that the distribution becomes much sharper around the mean that Woolfson's approximate formula (4·14) would suggest.

To show the general character of the distribution, we have used (4·12) and (4·13) to compute the probability that $E_{\mathbf{h}}$ has the same sign as the product $E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'}$ irrespective of the value of $E_{\mathbf{h}}$. This probability is easily shown to be equal to $\Phi\left(\frac{\text{mean}}{\text{variance}^{\frac{1}{2}}}\right)$, where $\Phi(x)$ is the standard statistical integral

$$\Phi(x) = \frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^{x} \exp\left(-\tfrac{1}{2}u^2\right)du\,. \qquad (4·15)$$

The results are plotted in Fig. 6, and it is satisfying that the computed contour $P = 1$ falls so close to the unit circle $E_2^2+E_3^2 = N$, as is required by the condition

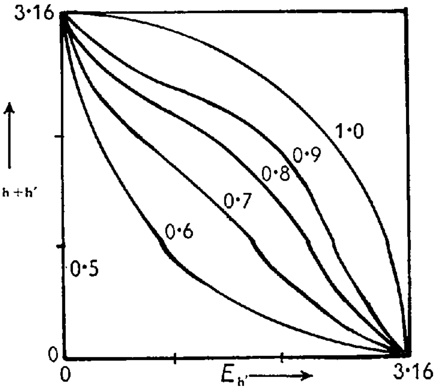

Fig. 6. The probability that $E_{\mathbf{h}}$ has a positive value, irrespective of its magnitude, as a function of $E_{\mathbf{h}'}$ and $E_{\mathbf{h}+\mathbf{h}'}$; $P\bar{1}$, $N = 10$. Computed from equations (4·12), (4·13) and (4·15).

A plot of the approximation to order $1/N^{\frac{1}{2}}$ would give a set of rectangular hyperbolae similar in appearance to Fig. 5(*c*).

(4·8). The plot is intended to be only illustrative in character, since, in any application of the theory, the magnitude of E_1 will always be known.

The probability $P_+(E_1)$ that E_1 is positive, taking into account the (known) value of E_1, is calculated as follows. When the form of the probability distribution of a random variable is Gaussian, as is that of E_1 in (4·11), then it follows quite generally from the reasoning of Cochran & Woolfson (1955, § 4·3) that the probability that it has a positive sign is given by

$$P_+(E_1) = \tfrac{1}{2}+\tfrac{1}{2}\tanh\left\{\frac{|E_1|\times\text{mean}}{\text{variance}}\right\}. \qquad (4{\cdot}16)$$

For our problem, the mean and variance to be substituted in this are given by equation (4·12) and (4·13) respectively. If the terms of order higher than $1/N^{\frac{1}{2}}$ are dropped, (4·16) reduces to the result of Cochran & Woolfson (1955) (their equation (4·16)). The effect of retaining the higher-order terms in (4·16) will be obvious from the example plotted in Fig. 5 and discussed in the last paragraph. At high values of E_2 and E_3, $P_+(E_1)$ is higher than would be suggested by the simple Cochran & Woolfson equation. The effect is entirely analogous to that discussed in § 3.22. Cochran & Woolfson's (1955) equation (4·16) thus gives an underestimate of the probability when the structure factors involved are large. This should be borne in mind in deciding on weightings for the S.C.Z. relations in a method for sign determination such as, say, that of Cochran & Douglas (1955) or related methods. In practice it would be very tedious to use the correct weightings given by equations (4·16), (4·12) and (4·13) above, but the theory does suggest that S.C.Z. relations involving very large structure factors should be kept in a separate class, or classes (cf. Cochran & Douglas), rather than attempt to use the simple approximate formula uniformly for all the sign relations.

Since in (4·12) E_2 and E_3 occur only as products of odd powers and in (4·13) only as even powers, $P_+(E_1)$ as calculated from (4·16) also gives the probability $P_+(E_1E_2E_3)$ that the sign relation (4·1) is true, when values of $|E_2|$ and $|E_3|$ are substituted, in (4·12) and (4·13). It will be noticed, however, that the expression is not completely symmetrical in E_1, E_2 and E_3. This asymmetry arises from the manner of derivation of (4·16) from (4·11) in which one of the three structure factors is singled out. The correct expression for $P_+(E_1E_2E_3)$ may be calculated from the joint distribution $P(E_1, E_2, E_3)$, equation (4·5), by using Bertaut's equation (3·22). In the latter P_o is to be interpreted as the terms in (4·5) which are odd in E_1, E_2 and E_3. This is obvious from the form of (4·1a) and the fact that the only powers $E_1^pE_2^qE_3^r$ occurring in (4·5) are those with p, q, r either all odd or all even. (It may be proved quite generally by the methods of Appendix D that moments m_{pqr} with mixed odd and even indices all vanish.)* The result is

$$P_+(E_1E_2E_3) = \tfrac{1}{2}+\tfrac{1}{2}$$
$$\times\left[\frac{\dfrac{1}{N^{\frac{1}{2}}}\,|E_1E_2E_3|+\text{terms in }\dfrac{1}{N^{3/2}}\text{ in }(4{\cdot}5)}{1-\dfrac{1}{8N}\,[H_4(E_1)+\text{cyc.}]+\dfrac{1}{2N}\,H_2(E_1)H_2(E_2)H_2(E_3)}\right], \qquad (4{\cdot}17)$$

The last equation may be compared with that obtained after substituting in (4·16) from (4·12) and (4·13). It will be found that some of the $1/N^{3/2}$ terms in E_1 are missing in the latter. The departure from complete symmetry, however, is only slight and the general arguments given in earlier paragraphs are quite unaffected. On neglecting all terms beyond those in $1/N^{\frac{1}{2}}$, equation (4·17) reduces to the result of Hauptman & Karle (1953, equation (3·30)).

4·4. *The statistical theory of Sayre's equation*

For the case of equal atoms, Sayre (1952) proved that

$$\frac{1}{N}\,U_{\mathbf{h}} = \frac{1}{H}\sum_{\mathbf{h}'}^{H} U_{\mathbf{h}'}U_{\mathbf{h}+\mathbf{h}'}\,, \qquad (4{\cdot}18a)$$

where H is the number of terms in the summation. In other words, as shown by Hughes (1953) and Cochran (1953),

$$U_{\mathbf{h}} = N\,\overline{U_{\mathbf{h}'}U_{\mathbf{h}+\mathbf{h}'}}^{\,\mathbf{h}'}$$

or

$$E_{\mathbf{h}} = N^{\frac{1}{2}}\,\overline{E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'}}^{\,\mathbf{h}'}, \qquad (4{\cdot}18b)$$

Sayre's equation is strictly true only over an infinite range of $\mathbf{h}'$. Cochran & Woolfson (1955) have studied the probability distribution of the values of 'a Sayre product' $E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'}$ for a fixed value of $E_{\mathbf{h}}$, and their expression (4·1) will be accurate to the order $1/N^{\frac{1}{2}}$ since their derivation is essentially based upon Woolfson's equation (our equation (4·14)). It is possible to carry out a derivation on similar lines but based upon the more accurate formulae given earlier in this section. This is, however, very tedious and we shall not give the calculation here. Instead, we shall content ourselves with calculating the mean value and variance of a Sayre product $E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'}$ as a function of $E_{\mathbf{h}}$ by the methods of Appendix E. Note that, while Sayre's equation (4·18) is the average over a range of values of $\mathbf{h}'$, our whole theory is based upon averages over the atomic positions $\mathbf{r}$. From the symmetry of the trigonometric form of $\xi(\mathbf{h})$ (equation (1·4)) in $\mathbf{h}$ and $\mathbf{r}$

---

* This statement is true only in the general case. There can be exceptions, however: e.g. if $E_1 \equiv E_{\mathbf{h}}$, $E_2 \equiv E_{3\mathbf{h}}$, $E_3 \equiv E_{4\mathbf{h}}$, then $m_{221} = 2$.

I am indebted to the referee for this remark, and also for some other suggestions for making the paper clearer.

we expect that the two averages will yield the same result (see Hauptman & Karle, 1953, equation (3·17)). We shall soon see that this is indeed the case.

Writing $E_{\mathbf{h}'} = E_1$, $E_{\mathbf{h}'+\mathbf{h}} = E_2$, $E_{\mathbf{h}} = E_3$, we find from equation (E·14), together with (E·13) and (E·18), that

$$\langle E_1E_2\rangle_{E_3} = \frac{\dfrac{1}{N^{\frac{1}{2}}}E_3 - \dfrac{1}{2N^{3/2}}H_3(E_3) - \dfrac{1}{8N^{3/2}}H_5(E_3)}{1-\dfrac{1}{8N}H_4(E_3)}$$

$$= \frac{\dfrac{1}{N^{\frac{1}{2}}}\left[E_3+\dfrac{1}{8N}Q(E_3)\right]}{1-\dfrac{1}{8N}H_4(E_3)}, \qquad (4{\cdot}19)$$

where

$$Q(x) = -\{4H_3(x)+H_5(x)\}\ . \qquad (4{\cdot}20)$$

Now $\langle E_1E_2\rangle_{E_3}$ must by (4·18b) be strictly equal to $\dfrac{1}{N^{\frac{1}{2}}}E_3$. It is clearly so to the order of approximation $1/N^{\frac{1}{2}}$, as was first shown by Cochran & Woolfson (1955, § 4·1). Despite the appearance of (4·19), Sayre's equation is also obeyed to the order $1/N^{3/2}$ since it may be proved that

$$\frac{x+AQ(x)}{1-AH_4(x)} = x\ , \qquad (4{\cdot}20a)$$

so that (4·19) reduces to

$$\langle E_1E_2\rangle_{E_3} = \frac{1}{N^{\frac{1}{2}}}E_3,\ \text{to the order}\ \frac{1}{N^{3/2}}\ . \qquad (4{\cdot}21)$$

The result (4·21) would seem to provide a demonstration of the general correctness of our statistical methods. We would in fact expect that if we carried the approximation to terms of order higher than $1/N^{3/2}$, we would nevertheless always obtain the result (4·21), because our expansions are truly asymptotic.

From (E·14), (E·13) and (E·18) we may also calculate the second moment $\langle E_1^2E_2^2\rangle_{E_3}$ of the Sayre product E_1E_2, and hence the variance. We find

$$\sigma^2_{E_3}(E_1E_2) = \frac{1-\dfrac{1}{4N}H_4(E_3)+\dfrac{1}{N}(E_3^2-2)}{1-\dfrac{1}{4N}H_4(E_3)}\ , \qquad (4{\cdot}22)$$

$$\fallingdotseq 1+\frac{1}{N}(E_3^2-2) \qquad (4{\cdot}22a)$$

on expanding by the binomial theorem. As may be seen from the approximate form (4·22a) and from Table 1, the variance remains of the order unity for all values of E_3, and indeed increases with E_3.

Table 1. *Variance of a Sayre product* $E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'}$ *as a function of* $E_{\mathbf{h}}$

$P\bar{1}$, $N = 10$

$E_{\mathbf{h}}$	0	0·5	1·0	1·5	2·0	2·5
Variance	0·79	0·82	0·91	1·02	1·18	1·48

This behaviour is quite the contrary of any we have found before for the variance. A moment's reflection, however, will show that this is just what we should expect in the case of Sayre's equation (4·18). When $E_{\mathbf{h}}$ is large we may expect that some of the Sayre products will be much larger than the average, namely those products for which the probability of the truth of the S.C.Z. sign relation (4·1) is high. When $E_{\mathbf{h}}$ is small there will tend to be a smaller number of large terms on the right hand side of (4·18a). The reader can easily convince himself of the correctness of this argument by imagining curves of constant E_2E_3 (e.g. Fig. 5(c)) to be laid on the top ($E_1 = 0$) and bottom ($E_1 = 2$) halves of Fig. 3. Larger values of E_2E_3 are possible in the case of the higher value of E_1. (Note, however, that in both cases the most probable value of E_2E_3 would still appear to be zero; and this would account for the empirical results found by Woolfson (1954) and embodied in his Fig. 1.)

It should perhaps be emphasized that the variance given in (4·22) is that of a single term in the summation involved in Sayre's equation. Strictly, we should say from (4·21) and (4·22) that $N^{\frac{1}{2}}E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'}$ is an estimate of $E_{\mathbf{h}}$ with variance $N+(E_{\mathbf{h}}^2-2) \fallingdotseq N$. Hence the estimate of $E_{\mathbf{h}}$ through the Sayre equation, including, say, s terms,

$$E_{\mathbf{h}} = N^{\frac{1}{2}}\sum_{\mathbf{h}'}^{s}E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}}$$

has a variance N/s. Ideally s is infinite, so Sayre's equation is strictly true. In practice, on certain reasonable assumptions (see below, § 5·2), $s \propto N^2$, so that the variance of the estimate through Sayre's equation is $\sim 1/N$, which is still very small. More precisely, to the order $1/N^{\frac{1}{2}}$, the probability of the truth of Sayre's equation is (cf. (4·16))

$$P_+(E_{\mathbf{h}}) = \tfrac{1}{2}+\tfrac{1}{2}\tanh\{N^{-\frac{1}{2}}|E_{\mathbf{h}}|\sum_{\mathbf{h}'}E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'}\}\ ,$$

a result which has already been obtained by Cochran & Woolfson (1955, equation (4·20)). We have however couched the argument in terms of the variance of an estimate in order to foreshadow § 5.

4·5. *Probability distribution of* $E_{\mathbf{h}}$, $E_{\mathbf{h}'}$ *and* $E_{\mathbf{h}+\mathbf{h}'}$ *when the atoms are not equal*

The general form of the distribution is given in (C·8). The required non-vanishing cumulants k_{pqr} for space group $P\bar{1}$ are given in (D·11). For comparison with the results of other workers we express the results in terms of

$$z_\nu = \sum_{j=1}^{N}\varphi_j^\nu \qquad \text{(Bertaut, 1955}a, b\text{)} \qquad (4{\cdot}23)$$

$$= \varepsilon_\nu/\varepsilon_2^{\nu/2} \qquad \text{(Cochran \& Woolfson, 1955)} \qquad (4{\cdot}23a)$$

instead of the ζ_ν of our equation (C·4). Note that

$$\zeta_\nu = z_\nu/n$$
$$= z_\nu/2 \quad \text{for } P\bar{1}\,. \tag{4·24}$$

We obtain

$$P(E_1, E_2, E_3) = \frac{1}{(2\pi)^{3/2}} \exp\left(-\tfrac{1}{2}E_1^2 - \tfrac{1}{2}E_2^2 - \tfrac{1}{2}E_3^2\right)$$
$$\times \Big\{1 + z_3E_1E_2E_3$$
$$-\frac{z_4}{8}[H_4(E_1) + \text{cyc.}] + \frac{z_3^2}{2}H_2(E_1)H_2(E_2)H_2(E_3)$$
$$-\frac{z_5}{2}[H_1(E_1)H_1(E_2)H_3(E_3) + \text{cyc.}]$$
$$-\frac{z_3z_4}{8}[H_1(E_1)H_1(E_2)H_5(E_3) + \text{cyc.}]$$
$$+\frac{z_3^3}{6}H_3(E_1)H_3(E_2)H_3(E_3) \tag{4·25}$$
$$+ \text{higher order terms}\Big\}\,.$$

In this equation we have kept the same ordering of terms as in (4·5), to which (4·25) reduces when the atoms are all made equal, since then

$$z_\nu = 1/N^{\frac{\nu}{2}-1}\,. \tag{4·26}$$

The term in z_3 in (4·25) has been given by Hauptman & Karle (1953, (3·30)) and by Bertaut (1955*b*).

We may now apply the methods of Appendix E to obtain various characteristics of the distribution. Comparing (4·25) with (E·8), we have

$$\begin{aligned} \beta_{111} &= -z_3, & \beta_{004} &= -z_4/8, & \beta_{222} &= z_3^2/8\,, \\ \beta_{113} &= z_5/2, & \beta_{115} &= z_3z_4/8, & \beta_{333} &= -z_3^3/6\,. \end{aligned} \tag{4·27}$$

The generalizations of (4·12) and (4·13) are

$$\langle E_1\rangle_{E_2,E_3} = \frac{z_3E_2E_3 - \frac{z_5}{2}[H_3(E_2)H_1(E_3) + H_1(E_2)H_3(E_3)] - \frac{z_3z_4}{8}[H_5(E_2)H_1(E_3) + H_1(E_2)H_5(E_3)]}{1 - \frac{z_4}{8}[H_4(E_2) + H_4(E_3)]} \tag{4.28}$$

and

$$\sigma^2_{E_2,E_3}(E_1) = \frac{1 - \frac{z_4}{4}[H_4(E_2) + H_4(E_3)] + z_3^2[1 - E_2^2 - E_3^2]}{1 - \frac{z_4}{4}[H_4(E_2) + H_4(E_3)]} \tag{4·29}$$
$$\doteqdot 1 + z_3^2(1 - E_2^2 - E_3^2)\,. \tag{4·29a}$$

Comparing (4·13*a*) and (4·29*a*), the variance decreases faster from unity for large values of E_2 and E_3 when the atoms are unequal, since in this case

$$z_3 > 1/N^{\frac{1}{2}}\,. \tag{4·30}$$

The probability of the truth of the S.C.Z. relation (4·1) is now

$$P_+(E_1E_2E_3) = \tfrac{1}{2} + \tfrac{1}{2}\left[\frac{z_3|E_1E_2E_3| + \text{terms in } z_5,\, z_3z_4 \text{ and } z_3^3 \text{ in (4·25)}}{1 - \frac{z_4}{8}[H_4(E_1) + \text{cyc.}] + \frac{z_3^2}{2}H_2(E_1)H_2(E_2)H_2(E_3)}\right]. \tag{4·31}$$

To the order z_3, this result agrees with the expansion of the closed expression obtained by Cochran & Woolfson (1955, equation (4·16)). Comparing (4·31) and (4·17), and noting (4·30), we see that the efficiency of sign relationships is improved when the atoms are unequal, a fact already noticed and explained by Cochran (1952) and Sayre (1953).

When the atoms are not equal, Sayre's equation (4·18*b*) is no longer strictly true. However, Cochran & Woolfson (1955) have derived a statistical form of Sayre's equation

$$V_{\mathbf{h}} = \overline{U_{\mathbf{h}'}U_{\mathbf{h}+\mathbf{h}'}}^{\mathbf{h}'} \doteqdot \frac{\varepsilon_3}{\varepsilon_2}U_{\mathbf{h}} \tag{4·32}$$

or

$$\overline{E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'}}^{\mathbf{h}'} \doteqdot z_3E_{\mathbf{h}} \tag{4·33}$$

which is true on the average (in the sense of least squares) over all **h**. In our theory the generalization of (4·19) is

$$\langle E_1E_2\rangle_{E_3} = \frac{z_3E_3 - \frac{z_5}{2}H_3(E_3) - \frac{z_3z_4}{8}H_5(E_3)}{1 - \frac{z_4}{8}H_4(E_3)}\,, \tag{4·34}$$

and, although the factorization which led to (4·21) in the case of equal atoms is no longer possible, (4·34) may be simplified by writing it as

$$\langle E_1E_2\rangle_{E_3} = \frac{z_3\left[E_3 - \frac{z_4}{8}\{4H_3(E_3) + H_5(E_3)\} - \frac{1}{2}\left(\frac{z_5}{z_3} - z_4\right)H_3(E_3)\right]}{1 - \frac{z_4}{8}H_4(E_3)}\,,$$

which, by (4·20) and (4·20*a*), becomes

$$\langle E_1E_2\rangle_{E_3} = z_3E_3 + \frac{\frac{1}{2}\left(z_4 - \frac{z_5}{z_3}\right)H_3(E_3)}{1 - \frac{z_4}{8}H_4(E_3)}\,. \tag{4·35}$$

To the order z_3, this result agrees with Cochran & Woolfson's equation (4·33) above, but we see that there are correction terms to the latter, particularly important for large E_3, when z_4 differs appreciably from z_5/z_3.

For the variance of a Sayre product E_1E_2 we find

$$\sigma^2_{E_3}(E_1E_2) = \frac{1 - \frac{z_4}{4}H_4(E_3) + z_3^2(E_3^2 - 2)}{1 - \frac{z_4}{4}H_4(E_3)} \tag{4·36}$$

$$\fallingdotseq 1+z_3^2(E_3^2-2)\,. \qquad (4{\cdot}36a)$$

Here again the variance is increased as compared with (4·22*a*).

While our results are more accurate than those of Cochran & Woolfson, it should be noted that theory given here takes no account of the physical basis of the result (4·32). The value of Cochran & Woolfson's approximate method of calculating probabilities is that it brings to the fore the connexion between the sign relations and the squared structure (see left-hand member of (4·32)). Thus, while it is not so powerful nor indeed so simple as the full statistical theory, it is nevertheless invaluable for general discussions of the phase problem.

5. The phase problem

5·1. *Order of a sign determining relation*

Provided that there is a linear relation between their indices, a sign relation or else a sign-determining relation* can be found between any set of structure factors. We shall denote such a relation by $R(E_{\mathbf{h}_1}, E_{\mathbf{h}_2}\ldots)$. The *a priori* probability of the truth of the relation can be found by the methods given earlier. It should also be possible *in every case* to find an inequality between the structure factors involved by the method of Karle & Hauptman (1950). As in the cases of the two relations studied in detail in this paper, the probability distribution will tail smoothly in the bounding contour set by the limiting case of the inequality.

Now it is clear that not all relations that can be derived are equally powerful or effective. The question arises whether there is a simple method of classification. The theory given above suggests a natural classification according to the lowest power of $1/N^{\frac{1}{2}}$ that enters into the probability distribution between the structure factors involved in the relation. More precisely, we define the *order of a relation* as the order $\nu = p+q+r\ldots$ of the lowest-order non-vanishing mixed moment $m_{pqr\ldots}$ for the structure factors involved. The leading term in the asymptotic expression for the probability will then be of order $1/N^{\nu/2-1}$.

The sign-determining formula (3·25) between $E_{2\mathbf{h}}$ and $E_{\mathbf{h}}$ (lowest moment m_{12}), and the S.C.Z. sign relation (4·1) between $E_{\mathbf{h}}$, $E_{\mathbf{h}'}$ and $E_{\mathbf{h}+\mathbf{h}'}$ (lowest moment m_{111}) will then both be of order 3. An example of a relation of higher order in space group $P\overline{1}$ is the much-discussed formula Σ_4 of Hauptman & Karle (1953, equation (4·06))

$$S(E_{2\mathbf{h}}) \approx S\left\{\sum_{\mathbf{h}'}(E_{\mathbf{h}'}^2-1)(E_{\mathbf{h}+\mathbf{h}'}^2-1)\right\}. \qquad (5{\cdot}1)$$

The formula is based on the probability

$$P_+(E_{2\mathbf{h}}) = \tfrac{1}{2}+\frac{1}{8N^{3/2}}(E_{\mathbf{h}'}^2-1)(E_{\mathbf{h}+\mathbf{h}'}^2-1)\,, \qquad (5{\cdot}2)$$

which arises from the non-vanishing moment m_{122}. The relation (5·1) is thus of order 5. Cochran & Woolfson (1955) have shown that, despite the large number of terms in it, (5·1) gives a less reliable indication for the sign of $E_{2\mathbf{h}}$ than does the single term $E_{\mathbf{h}}^2-1$ (our equation (3·25)). The reason is that (5·1) is of a much higher order. Furthermore, Cochran (1954) has derived the exact relation

$$E_{2\mathbf{h}} = N^{\frac{1}{2}}[2(E_{\mathbf{h}}^2-1)-\mathrm{N}\overline{(E_{\mathbf{h}'}^2-1)(E_{\mathbf{h}+\mathbf{h}'}^2-1)}^{\mathbf{h}'}], \qquad (5{\cdot}3)$$

which supersedes the statistical formula (5·1), and which we shall discuss presently.

The mathematical operation of determining signs in reciprocal space by means of sign relations has its physical counterpart in various operations in physical space. This was first pointed out by Cochran & Woolfson (1954) and Vand & Pepinsky (1954) soon after the appearance of Hauptman & Karle's monograph, and in fact was the basis of Cochran's derivation of (5·3). Thus $R(E_{2\mathbf{h}}, E_{\mathbf{h}})$ is related to the Patterson, $R(E_{\mathbf{h}}, E_{\mathbf{h}'}, E_{\mathbf{h}'+\mathbf{h}})$ to the squared structure, both of which are functions of the second power of the electron density. $R({}_{2\mathbf{h}}, E_{\mathbf{h}'}, E_{\mathbf{h}'+\mathbf{h}})$ is related to the squared Patterson. If r is the power of the electron density in the operation in real space corresponding to the sign relation, it thus appears that the order of the latter as defined above is equal to $r+1$. The correspondence between operations in reciprocal and real space has been recently emphasized by Bertaut (1955*c*).

We shall now discuss why higher-order relations become much less useful as the number of atoms in the cell increases. The power or efficiency of a relation for sign estimation is given by the deviation of the probability from $\frac{1}{2}$, i.e. its bias. This bias will be $\propto 1/N^{\nu/2-1}$, where ν is the order of the relation. All relations will thus be weakened as N increases, the more so, the greater ν. Alternatively, from the point of view of the last paragraph, the higher-order relations will tend to fail because the number of peaks in, say, the (sharpened) Patterson, or its convolution with the structure etc., will increase as the number of atoms in the unit cell increases. Thus the determination of the signs of $E_{2\mathbf{h}}$ in space group $P\overline{1}$ depends on spotting (equation (3·25)) or cancelling out (equation (5·3)) all peaks in the Patterson except the peaks at $\pm 2\mathbf{r}_j$ (Cochran, 1954). This will become increasingly difficult as the number of peaks increases, since there will be more and more 'chance' peaks lying close to the neighbourhood of a $\pm 2\mathbf{r}_j$ peak. Only if there is a sufficiency of data to resolve the peaks completely (Cochran, 1954) can we expect the method to succeed. Even then, there could always be a chance coincidence of one of the cross peaks $\mathbf{r}_i-\mathbf{r}_k$ with a $2\mathbf{r}_j$ peak, and the likelihood of this increases with N.

---

* It is also possible to find relations between magnitudes, but these are of no interest for the phase problem.

5·2. *The variance of a relation*

The weakness of high-order sign relations when N is large may perhaps be understood more clearly by considering the variance of one of the terms contributing to a sign relation. We shall again consider the example of (5·1).

The joint probability distribution of $E_{2\mathbf{h}}$, $E_{\mathbf{h}'}$ and $E_{\mathbf{h}'+\mathbf{h}}$ may be very easily found by the methods of § 2 and Appendix C. The first non-vanishing mixed moment is m_{122}. (Strictly speaking the lowest-order moments are m_{400}, m_{040}, etc., but these do not involve all three structure factors.) The next is m_{322}, and so on, but we shall confine ourselves to the leading terms in the probability theory. From Appendix E

$$\langle E_{2\mathbf{h}}\rangle = \frac{\beta_{122}I^1_{122}+\ldots}{I^0_{000}+\beta_{040}I^0_{040}+\beta_{004}I^0_{004}}$$

$$= \frac{\frac{1}{4N^{3/2}}\,(E^2_{\mathbf{h}'}-1)\,(E^2_{\mathbf{h}+\mathbf{h}'}-1)}{1-\frac{1}{8N}\,[H_4(E_{\mathbf{h}'})+H_4(E_{\mathbf{h}+\mathbf{h}'})]}\,, \qquad (5·4)$$

an equation which is the basis of (5·1). In a similar way the variance of $E_{2\mathbf{h}}$ for fixed values of $E_{\mathbf{h}'}$ and $E_{\mathbf{h}+\mathbf{h}'}$ may be shown to be unity, to the order $1/N^{\frac{1}{2}}$.

More relevant for our purpose are the mean and variance of a single product $(E^2_{\mathbf{h}'}-1)(E^2_{\mathbf{h}+\mathbf{h}'}-1) = H_2(E_{\mathbf{h}'})H_2(E_{\mathbf{h}+\mathbf{h}'})$ contributing to the sum in (5·1). These may be found by the method of Appendix E, § *b*(ii). After a somewhat lengthy calculation we find

$$\langle(E^2_{\mathbf{h}'}-1)(E^2_{\mathbf{h}'+\mathbf{h}}-1)\rangle_{E_{2\mathbf{h}}} = E_{2\mathbf{h}}/N^{3/2} \qquad (5·5)$$

and a variance of unity. Hence one term

$$N^{3/2}(E^2_{\mathbf{h}'}-1)(E^2_{\mathbf{h}'+\mathbf{h}}-1)$$

provides an estimate of $E_{\mathbf{h}}$ with a variance of N^3. This result agrees with that found by Bullough & Cruickshank (1955) by algebraic methods. As indicated by these authors, if there are p terms in the summation (5·1) the variance of this estimate will be reduced by a factor $1/p$. Now p will be approximately proportional to the square of the number of reflexions observed. On reasonable assumptions, this last number will be proportional to the volume of the unit cell and hence to the number of atoms in the unit cell, N. Thus $p \propto N^2$. Hence the variance of the estimate of $E_{2\mathbf{h}}$ through the sum (5·1) will be of order N. Now the E's are by definition normalized structure factors of variance unity, so this variance is in effect enormous when N is greater than a very small number. Thus we see that the reason from the point of view of statistical theory for the failure of Hauptman & Karle's Σ_4 is that, although it does in fact contain the correct bias, this is in practice not revealed, since the number of terms in the sum is not enough to outweight the variance of each term. The physical counterpart of this reasoning is given by Cochran & Woolfson (1954).

It should be noted that the arguments are based upon approximate formulae. From the results of previous sections we should expect that the inclusion of terms of higher order than $1/N^{3/2}$ in the expressions for the probability distribution and for the mean and variance will have the effect of making the situation slightly more favourable than as stated above. When the E's involved are large the variance of a single product $(E^2_{\mathbf{h}'}-1)(E^2_{\mathbf{h}+\mathbf{h}'}-1)$ will be less than unity and the efficiency of the estimate correspondingly increased. As Cochran & Woolfson's (1955) Table 3·1 shows, there will be some situations in which Σ_4 does give the correct sign, but an examination of their arguments will show that this happens essentially because $E_{\mathbf{h}}$ is large.

By a large value of E we mean one which is not too small compared with its maximum value, i.e. $N^{\frac{1}{2}}$ in the case of equal atoms. (The advantage of the U notation, of course, is just that the effective value of a structure factor is always kept before one.) Now the probability distribution for the normalized structure factor E is always approximately Gaussian with variance unity (neglecting terms in $1/N$ and higher):

$$(2\pi)^{-\frac{1}{2}}\exp\left(-\tfrac{1}{2}E^2\right)\,. \qquad (5·6)$$

Thus while the probability of finding a particular value of E is independent of N, this value will represent an increasingly smaller unitary structure factor as N increases.

As a measure of the probability of finding large values of E we may examine the distribution of the highest value (extreme value) or second highest value, and so on, of E. On the basis of previous assumptions the number of observations of E (i.e. the number of reflexions within the limiting sphere) will be proportional to N, say equal to θN. Then by Cramér (1946, p. 376) the expected (or mean) value of the largest E observed is

$$\langle E_{\text{extreme}}\rangle = (2\log\theta N)^{\frac{1}{2}},\quad \theta\sim 1\,, \qquad (5·7)$$

which decreases much more slowly than the square root or a power of N (cf. remark of Bullough & Cruickshank, 1955). Since any sign-determining formula is of order $1/N^{\frac{1}{2}}$, or higher, as N increases it will thus become increasingly rare to find E's large enough to use in, or even to start off, the process of sign determination. This is immediately obvious if we use unitary structure factors, which also have a Gaussian distribution, but with variance $1/N$. Then

$$\langle U_{\text{extreme}}\rangle = \{(2\log\theta N)/N\}^{\frac{1}{2}}\,. \qquad (5·8)$$

It is therefore misleading to state, as Hauptman & Karle (1954) did, that because (5·6) is always (approximately) obeyed, all structures are equally vulnerable to attack by statistical methods.

5·3. *Sign determination by means of a formula having strict validity*

One of the important effects of Cochran's (1954) introduction of the formula (5·3), having strict validity, into the discussion ranging round Hauptman & Karle's statistical formula (5·1) lies perhaps in the fact that it focused attention on the desirability of selecting the set of structure factors so that they were strictly related by a formula derived either by algebra or by physical considerations. It was thus seen that while the formulae relating $E_{2\mathbf{h}}$, $E_{\mathbf{h}'}$ and $E_{\mathbf{h}+\mathbf{h}'}$, (5·4), (5·5) and (5·1) etc., were certainly correct, they were nevertheless highly artificial in omitting $E_{\mathbf{h}}$, which would certainly be known in magnitude and could thus be taken into account in the process of sign determination of $E_{2\mathbf{h}}$. Cochran's formula (5·3) showed how to do this rationally.

From the point of view of probability theory we should thus consider the set of structure factors $E_{2\mathbf{h}}, E_{\mathbf{h}'}, E_{\mathbf{h}+\mathbf{h}'}, E_{\mathbf{h}}$. Besides the non-vanishing mixed moment m_{1220}, which led to (5·4) and (5·5), there are now lower-order moments m_{1002} and m_{0111} which must be taken into account in calculating the distribution $P(E_{2\mathbf{h}}, E_{\mathbf{h}'}, E_{\mathbf{h}+\mathbf{h}'}, E_{\mathbf{h}})$. The set of terms of order $1/N^{3/2}$ in the latter would then include, besides that in m_{1220} and those we have already considered in §§ 3 and 4, terms arising from the set Λ_3^3 (see § 2) like

$$\begin{aligned} m_{1002}m_{0111}^2 H_1(E_{2\mathbf{h}})H_2(E_{\mathbf{h}'})H_2(E_{\mathbf{h}+\mathbf{h}'})H_4(E_{\mathbf{h}}) \\ = m_{1002}m_{0111}^2 E_{2\mathbf{h}}(E_{\mathbf{h}'}^2-1)(E_{\mathbf{h}+\mathbf{h}'}^2-1)H_4(E_{\mathbf{h}}) \end{aligned} \qquad (5{\cdot}9)$$

and

$$m_{1002}^2 m_{0111}(E_{2\mathbf{h}}^2-1)E_{\mathbf{h}'}E_{\mathbf{h}+\mathbf{h}'}H_5(E_{\mathbf{h}})\,. \qquad (5{\cdot}10)$$

The term (5·9) is obviously one in which the intensities of $E_{\mathbf{h}}$, $E_{\mathbf{h}'}$ and $E_{\mathbf{h}+\mathbf{h}'}$ all play a part together in determining the sign of $E_{2\mathbf{h}}$, and is at least as important as (5·4). Indeed, by analogy with the case of § 4·2, where the probability theory gave Sayre's equation exactly, we would conjecture that if we calculated the conditional mean $\langle(E_{\mathbf{h}'}^2-1)(E_{\mathbf{h}'+\mathbf{h}}^2-1)\rangle$ as a function of $E_{\mathbf{h}}$ and $E_{2\mathbf{h}}$ from the full expression for $P(E_{2\mathbf{h}}, E_{\mathbf{h}'}, E_{\mathbf{h}'+\mathbf{h}}, E_{\mathbf{h}})$, we would arrive at the result

$$\frac{2}{N}(E_{\mathbf{h}}^2-1)-\frac{1}{N^{3/2}}E_{2\mathbf{h}}\,,$$

corresponding to Cochran's equation (5·3). The labour involved in such a proof would be very great and hardly seems worth while at this stage. We have already discussed the implications of (5·3) and have seen that the use of an accurate formula is still subject to limitations that depend on the number of atoms present in the cell.

Cochran (1954) also derived a number of third-order formulae having strict validity, which determine signs from intensities in higher space groups. We have already discussed in § 3·4 the range of application of one of these formulae. We shall now discuss a fourth-order formula.

5·4. *The fourth order formula in $P\bar{1}$*

5·41.—A formula having strict validity and useful for yielding information about structure factors not all of the form $E_{2\mathbf{L}}$ was derived by Bullough & Cruickshank (1955):

$$N\overline{(E_{\mathbf{h}}^2-1)E_{\mathbf{h}'}}^{2\mathbf{h}+\mathbf{h}'} = E_{2\mathbf{h}+\mathbf{h}'}\,. \qquad (5{\cdot}11)$$

The first non-vanishing moment in the probability distribution $P(E_{2\mathbf{h}+\mathbf{h}'}, E_{\mathbf{h}}, E_{\mathbf{h}'})$ is m_{121}, so the formula is of the fourth order. It corresponds physically to the convolution of the structure and the Patterson, i.e. to the superposition method.

Hauptman & Karle (1953, equation (3·31)) had already derived the probability of $E_{2\mathbf{h}+\mathbf{h}'}$ being positive, namely

$$P_+(E_{2\mathbf{h}+\mathbf{h}'}) = \tfrac{1}{2}+1/4N|E_{2\mathbf{h}+\mathbf{h}'}|(E_{\mathbf{h}}^2-1)E_{\mathbf{h}'}\,, \qquad (5{\cdot}12)$$

an expression which we now see is accurate to the order $1/N$. (The first correction term in this expression will be of order $1/N^2$.) They had then used this expression to find their sign relation Σ_3:

$$S(E_{2\mathbf{h}+\mathbf{h}'}) \approx S\left\{\sum_{2\mathbf{h}+\mathbf{h}'=\text{const.}}(E_{\mathbf{h}}^2-1)E_{\mathbf{h}'}\right\}. \qquad (5{\cdot}13)$$

Bullough & Cruickshank's work thus showed that the formula Σ_3 has a basis that is strictly and not merely statistically valid. However, while it is certainly desirable to try to use formulae or relations having strict validity, we must remember that the efficacy of such a formula is still determined by the order of the formula as defined above. The formula (5·13) expresses the desired structure factor as a sum of contributions involving the related structure factors. From the experience of § 5·2 it will be clear that the formula will be useful only if the number of contributions is such as to outweigh the variance of one of them. The variance of a single term, we recall, is related to the order of the formula.

In this case it is easy to show that the variance of the product $H_2(E_{\mathbf{h}})H_1(E_{\mathbf{h}'}) = (E_{\mathbf{h}}^2-1)E_{\mathbf{h}'}$ is of the form

$$2+1/N\ (\text{function of } E_{2\mathbf{h}+\mathbf{h}'})+\dots\,.$$

Hence the variance of $N(E_{\mathbf{h}}^2-1)E_{\mathbf{h}'}$, which is an estimate of $E_{2\mathbf{h}+\mathbf{h}'}$ is

$$2N^2+N\ (\text{function of } E_{2\mathbf{h}+\mathbf{h}'})\,. \qquad (5{\cdot}14)$$

For small or moderate values of $E_{2\mathbf{h}+\mathbf{h}'}$, the dominant terms is $2N^2$, so that, to this approximation our result agrees with that of Bullough & Cruickshank. (Cf. § 3·3, for another instance where the variance as calculated correctly from the statistical theory contains correction terms to that found by the simple algebraic methods.)

On our previous assumptions, the number of terms in the summation (5·13) will be of the order N^2, so that the variance of the estimate of $E_{2\mathbf{h}+\mathbf{h}'}$ through (5·11) or (5·13) will generally be of the order 2. Only for large values of $E_{2\mathbf{h}+\mathbf{h}'}$ can we expect the variance to be much less than this, but these will occur only rarely, as discussed in § 5.2; our former remarks will apply and we can ignore large values in a general discussion. Now for an estimate to be effective for sign determination its variance must be small compared with unity (since $\langle E^2 \rangle = 1$), and we see that this is not the case here. Hence, although equation (5·11) is a formula having strict validity like Sayre's equation (4·18), and can ideally be used in the same way in the process of sign determination, it must in practice be much less effective. From our point of view the reason for this is that (5·11) is a higher-order formula.

Despite this, it should be emphasized that the variance of an estimate of $E_{2\mathbf{h}+\mathbf{h}'}$ through Σ_3 or (5·11) is (approximately) independent of N. We should thus not expect its efficiency to drop as N increases, as does that of the formula Σ_4 (equation (5·1)). In this connexion, it should be noted that Σ_3 played an important part in the successful solution of the structure of *p,p'*-dimethoxybenzophenone by statistical methods (Karle *et al.*, 1957). This is a fairly complex structure without any atoms heavier than the rest, and its solution would be creditable by any method. However, without further investigation into the distribution of the magnitudes of the structure factors in this particular case and an examination of the Patterson, it cannot be concluded that the sign determination will also be successful for all other comparable structures. Indeed, we should expect the reverse to be true; i.e. that only occasionally (statistically!) will sign determination by direct methods of this kind be successful for complex structures with all atoms equal. But only a good deal of experience can show just how useful Σ_3 will turn out to be in practice.

5·42.—Some further points need to be made concerning the order in which the formulae available for sign determination are to be used in the actual process of sign determination. In $P\bar{1}$, the only formula available for determining signs from intensities is either (3·25) with probability calculable from (3·23), or else the complete equation (5·3). The next step would be either (i) to use the S.C.Z. relations (4·1) to extend the signs available, or else (ii) to apply formula (5·11). Cochran & Woolfson (1955, § 4·5) have suggested by an example that higher probabilities for $P_+(E_{2\mathbf{h}+\mathbf{h}'})$ may be found by following the first path rather than the second. This does not seem to be necessarily always the case, but it does raise the question of which path is to be preferred, and we shall now consider what our theory has to say about this.

Although the set of structure factors $E_{2\mathbf{h}+\mathbf{h}'}$, $E_\mathbf{h}$ and $E_{\mathbf{h}'}$ are related by the equality (5·11), there is no reason why we should not at the same time consider $E_{2\mathbf{h}}$ as well. The formula relating the four structure factors will now only have statistical validity, but its power for sign determination will still be of the same order, if not greater than that of (5·11). The non-vanishing mixed moments of the set $E_{2\mathbf{h}+\mathbf{h}'}$, $E_\mathbf{h}$, $E_{\mathbf{h}'}$ and $E_{2\mathbf{h}}$ are m_{1011}, m_{0201}, m_{1210}, and the expression for $P(E_{2\mathbf{h}+\mathbf{h}'}, E_\mathbf{h}, E_{\mathbf{h}'}, E_{2\mathbf{h}})$ will contain among its terms of order $1/N$, not only the term

$$\frac{\lambda_{1210}}{1!\,2!\,1!\,0!}\,E_{2\mathbf{h}+\mathbf{h}'}(E_\mathbf{h}^2-1)E_{\mathbf{h}'} \tag{5·15}$$

corresponding to (5·12), but also the term

$$\frac{\lambda_{1011}}{1!\,0!\,1!\,1!}\,\frac{\lambda_{0201}}{0!\,2!\,0!\,1!}\,E_{2\mathbf{h}+\mathbf{h}'}(E_\mathbf{h}^2-1)E_{\mathbf{h}'}(E_{2\mathbf{h}}^2-1)\,, \tag{5·16}$$

which arises from the set $\frac{1}{2}\Lambda_3^2$ in equation (2·17). Putting in values for the mixed moments, we find to the order $1/N$

$$\begin{aligned}
P(E_{2\mathbf{h}+\mathbf{h}'}, E_\mathbf{h}, E_{\mathbf{h}'}, E_{2\mathbf{h}}) &= (2\pi)^{-2} \\
&\times \exp\left[-\tfrac{1}{2}(E_{2\mathbf{h}+\mathbf{h}'}^2+\text{cyc.})\right] \\
&\times\Big\{1+\frac{1}{2N^{\frac{1}{2}}}\,(E_\mathbf{h}^2-1)E_{2\mathbf{h}}+\frac{1}{N^{\frac{1}{2}}}E_{2\mathbf{h}+\mathbf{h}'}E_{\mathbf{h}'}E_{2\mathbf{h}} \\
&+\frac{1}{2N}E_{2\mathbf{h}+\mathbf{h}'}(E_\mathbf{h}^2-1)E_{\mathbf{h}'} \\
&+\frac{1}{2N}E_{2\mathbf{h}+\mathbf{h}'}(E_\mathbf{h}^2-1)E_{\mathbf{h}'}(E_{2\mathbf{h}}^2-1)-\frac{1}{8N}\,[H_4(E)+\text{cyc.}] \\
&+\frac{1}{2N}H_2(E_{2\mathbf{h}+\mathbf{h}'})H_2(E_{\mathbf{h}'})H_2(E_{2\mathbf{h}})+\frac{1}{8N}H_4(E_\mathbf{h})H_2(E_{2\mathbf{h}}) \\
&+\text{terms in } 1/N^{3/2}\ \text{etc.}\Big\}\,.
\end{aligned} \tag{5·17}$$

Let us consider the sign of $E_{\mathbf{h}'}$ as known. When the sign of $E_{2\mathbf{h}}$ is unknown, the second (non-constant) term in (5·17) corresponding to the S.C.Z. relation is of no use, and we have approximately

$$S(E_{2\mathbf{h}+\mathbf{h}'})\approx S\left\{\frac{1}{2N}\sum_{\mathbf{h},\mathbf{h}'}(E_\mathbf{h}^2-1)E_{\mathbf{h}'}+(E_\mathbf{h}^2-1)E_{\mathbf{h}'}(E_{2\mathbf{h}}^2-1)\right\} \tag{5·18}$$

$$= S\left\{\sum_{\mathbf{h},\mathbf{h}'}(E_\mathbf{h}^2-1)E_{\mathbf{h}'}E_{2\mathbf{h}}^2\right\}. \tag{5·19}$$

The two terms on the right-hand side of (5·18) correspond to the paths (i) and (ii) above, but we now see that there is really no question of preference of one above the other. Path (ii) based on the fourth-order formula, and path (i), which is equivalent to using two third-order formulae in succession, are of equal power (order $1/N$) and must be used together as indicated by (5·19). It is not correct to say, as did Cochran & Woolfson (1955, § 4·5), that the sign-determining formula (5·13) of path (ii) is not really independent of the lower-order formulae (3·25) and (4·1). The correct point to make is that (5·13) is not any more powerful than the latter two formulae used in succession, and indeed that the latter path makes use of the known magnitude of $E_{2\mathbf{h}}$. When $|E_{2\mathbf{h}}| > 1$, the sign of the contribution from paths (i) and (ii) separately will agree with that found from (5·19). When $|E_{2\mathbf{h}}| < 1$, i.e. $E_{2\mathbf{h}}$ is weaker than the average structure factor, the two paths (i) and (ii) would give discordant results, and hence particular care must be taken to use the full formula (5·19) in which the two paths are essentially given their correct weighting.

The result (5·19) appears to be new and to supersede both Hauptman & Karle's formula (5·13) and Cochran & Woolfson's prescription when the sign of $E_{\mathbf{2h}}$ is not to be taken into consideration. With this proviso, it is formally correct. However, as soon as any signs of the $E_{\mathbf{2h}}$ become available, they must be taken into account in determining the sign of $E_{\mathbf{2h+h'}}$ by means of the S.C.Z. relation between $E_{\mathbf{2h+h'}}$, $E_{\mathbf{h'}}$ and $E_{\mathbf{2h}}$. From Bertaut's equation (3·22) and (5·17), the probability that $E_{\mathbf{2h+h'}}$ is positive or negative is

$$P_{\pm}(E_{\mathbf{2h+h'}}) = \tfrac{1}{2} \pm \tfrac{1}{2} \times \left[\frac{\frac{1}{2N}|E_{\mathbf{2h+h'}}|(E_{\mathbf{h}}^2-1)E_{\mathbf{h'}}E_{\mathbf{2h}}^2 + \frac{1}{N^{\frac{1}{2}}}|E_{\mathbf{2h+h'}}|E_{\mathbf{h'}}E_{\mathbf{2h}}}{1+\frac{1}{2N^{\frac{1}{2}}}E_{\mathbf{2h}}(E_{\mathbf{h}}^2-1)+\text{terms in }\frac{1}{N}} \right]. \tag{5·20}$$

When $E_{\mathbf{2h}}$ is ignored, we arrive at (5·19). When it is not, (5·19) must be replaced by

$$S(E_{\mathbf{2h+h'}}) \approx S\left\{\sum_{\mathbf{2h,h,h'}} [\text{terms in square brackets in (5·20)}]\right\}. \tag{5·21}$$

In practice, however, the S.C.Z. relation, being of order $1/N^{\frac{1}{2}}$, will dominate in the estimate of the sign $E_{\mathbf{2h+h'}}$. This is clearly revealed when the denominator of (5·20) is expanded by the binomial theorem, and terms higher than order $1/N$ are omitted:

$$\begin{aligned} P_{\pm}(E_{\mathbf{2h+h'}}) &= \tfrac{1}{2} \pm \tfrac{1}{2}\left[\frac{1}{2N}|E_{\mathbf{2h+h'}}|(E_{\mathbf{h}}^2-1)E_{\mathbf{h'}}E_{\mathbf{2h}}^2 \right. \\ &\quad \left. + \frac{1}{N^{\frac{1}{2}}}|E_{\mathbf{2h+h'}}|E_{\mathbf{h'}}E_{\mathbf{2h}}\left\{1-\frac{1}{2N^{\frac{1}{2}}}E_{\mathbf{2h}}(E_{\mathbf{h}}^2-1)\right\}\right] \\ &= \tfrac{1}{2} \pm \tfrac{1}{2}N^{-\frac{1}{2}}|E_{\mathbf{2h+h'}}|E_{\mathbf{h'}}E_{\mathbf{2h}}\,. \end{aligned}$$

This is the result we would get from the S.C.Z. relation alone. The contribution from the terms corresponding to (5·19) in (5·21) is thus negligible compared to that of the S.C.Z. terms.

5·43.—In higher space groups the non-vanishing moment m_{112} correlates (Bertaut, 1956*b*) structure factors of the same class, i.e. those undergoing the same change in sign when the origin is changed. Thus in space group $P2_1/a$, a relation will be obtained which enables one to compare the signs of two structure factors $E_{h'k'l'}$ and $E_{h''k'l''}$, where h' and h'' differ by an even integer, $2h$, say, and l' and l'' differ by $2l$ (see Hauptman & Karle, 1953, Table 11, Case 5). The statistical relation analogous to (5·12) correlates $E_{h'k'l'}$ and $E_{h''k'l''}$ through the magnitude $|E_{2h,k,2l}|$. Since k is a free index, the relation as used in practice contains more than one term. Likewise, the signs of $E_{h'k'l'}$ and $E_{h'k''l'}$ are correlated through the magnitude $|E_{hkl}|$ where $k = \frac{1}{2}(k'-k'')$, and h and l are free indices.

Bertaut (1956*a*, *b*) has made great use of the fourth-order relation in higher space groups, involving the moment m_{112}, to solve a number of structures. It seems that the solutions could equally well have been obtained by using first the analogue of (3·25) and then the S.C.Z. relations.

To facilitate the derivation of the relation in higher space groups, Bertaut (1955*c*) has developed a most ingenious 'structure-factor algebra', and has published (Bertaut & Dulac, 1955) tables of relations so obtained. The method is exactly analogous to calculation of the sets of structure factors for which the moment m_{112} does not vanish, but its advantage is that these sets are given easily and automatically. The structure-factor algebra is useful for other purposes and may also be used to generate higher-order relations, but Bertaut has not done this. In view of what we have already said concerning the weakness of relations of order higher than the fourth, it would not, however, seem worth while to do this.

5·5. *Formulae available for sign determination in $P\bar{1}$*

In their latest approach to the phase problem, Hauptman & Karle (1957) stress the importance of formulae having strict validity. The discussion in § 5·42 should serve to show that there is no particular advantage attached to such formulae. The desirable thing is to group all useful formulae of the same order, whether derived algebraically or else arising from successive application of lower-order formulae, and whether they are strictly or only statistically valid.

In $P\bar{1}$, the only third-order formulae available are (3·25) and (4·18). The fourth-order formulae available, apart from (5·19) based on the moment m_{112}, will be those based on the moments m_{1111} and m_{13}. The first involves four structure factors $E_{\mathbf{h}_i}$, connected by the condition

$$\mathbf{h}_1 \pm \mathbf{h}_2 \pm \mathbf{h}_3 \pm \mathbf{h}_4 = 0\,,$$

and it is easy to show (Simerská, 1956) that there is in fact an exact relation between them,

$$\overline{E_{\mathbf{h}_1}E_{\mathbf{h}_2}E_{\mathbf{h}_3}}^{\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3=\mathbf{h}_4} = \frac{1}{N}E_{\mathbf{h}_4}\,.$$

The use of this formula will require knowledge of three of the signs of the E's involved, and it will clearly not be as useful as the S.C.Z. relation. We shall not discuss it here (see Simerská (1956), who has tested the relation on a known structure). The moment m_{13} relates the structure factors $E_{\mathbf{3h}}$ and $E_{\mathbf{h}}$. The relation is one between signs since only odd powers of $E_{\mathbf{3h}}$ and $E_{\mathbf{h}}$ will be involved. The correlation is of order $1/N$ and is thus weaker than that between $E_{\mathbf{2h}}$ and $E_{\mathbf{h}}$ (order $1/N^{\frac{1}{2}}$). We shall not discuss it any further. (See Cochran & Woolfson, 1955, § 3·5.)

From our general remarks it will be clear that the fifth-order formulae will be too weak to afford reliable sign indications. Hauptman & Karle (1957) have listed a number of these and have expressed the hope that standards of significance may be set up for the reliability of means computed from the formulae, and suitable laws of rejection then established. The methods of the present paper may be used to write down probability distributions corresponding to all the formulae listed, and hence to estimate their reliability. We have not done this since we do not believe that any of the higher-order formulae will turn out to be useful in practice for anything but the very simplest structures.

Indeed it seems to us that the most useful approach to the phase problem is to concentrate on the low-(third-)order formulae, and to try to discover some systematic method of application. The first attempt in this direction was the method of Cochran & Douglas (1955), which sets out to find the signs of a limited set of structure factors so that the S.C.Z. relations are generally obeyed. More recent attempts in the same direction are those of Woolfson (1957) and Grant, Howells & Rogers (1957). It seems likely that this approach will eventually extend and return to Sayre's (1952) original method of finding a set of signs for all the structure factors that obey the complete Sayre equation (4·18*a*). The generalization of this kind of approach to finding a set of signs obeying the fourth-order formula (5·11) might also repay some study.

The work was stimulated by a colloquium given by Dr E. F. Bertaut in Cambridge in April 1956. It was done in the summer of 1956 at the Cavendish Laboratory, and I am grateful to Dr M. F. Perutz for his hospitality. I would also like to thank Dr W. Cochran for many helpful discussions during that time and for his general interest, and Dr H. E. Daniels for a general discussion of probability theory.

I wish to record the help of Mrs K. Cratchley and Mr A. Page with the computations and drawings.

APPENDIX A

Generating functions, moments and cumulants

We shall here briefly summarize the fundamental theorems of probability theory concerning generating functions used in the text. Full treatments are given in such standard works as Kendall (1943) and Cramér (1946).

Let $f(x)$ be a probability distribution function of the random variable x, for simplicity supposed to have the mean value zero, and where, as usual,

$$\int_{-\infty}^{\infty} f(x)dx = 1 \ .$$

The moments m_r are given by

$$m_r = \int_{-\infty}^{\infty} x^r f(x)dx \tag{A·1}$$

and it is convenient to introduce the moment generating function (m.g.f.)

$$\begin{aligned} M(u) &\equiv \mathscr{E}(\exp[ux]) \\ &= \int_{-\infty}^{\infty} \exp[ux]f(x)dx \\ &= 1+um_1+\frac{u^2}{2!}m_2+\frac{u^3}{3!}m_3+\ldots \ . \end{aligned} \tag{A·2}$$

Here the symbol $\mathscr{E}$ stands for the operation of finding the (mathematical) expectation or mean value of a function. Thus the mean value of a function $\psi(x)$ of the random variable x will be written

$$\mathscr{E}\psi(x) \equiv \int_{-\infty}^{\infty} \psi(x)f(x)dx \ , \tag{A·3}$$

The Fourier transform of $f(x)$,

$$C(u) \equiv \mathscr{E}(\exp[iux]) = \int_{-\infty}^{\infty} \exp[iux]f(x)dx \ , \tag{A·4}$$

is known in probability theory as the characteristic function (c.f.) and may also be used to generate the moments, since

$$C(u) = M(iu) = 1+ium_1+\frac{(iu)^2}{2!}m_2+\ldots \ . \tag{A·5}$$

The great utility of introducing the m.g.f. or c.f. lies in the ease with which the probability distribution of a sum of random variables may be calculated. If we have two random variables x_1 and x_2, with distributions $f(x_1)$ and $g(x_2)$, the distribution of their sum $X = x_1+x_2$ is given by

$$P(X) = \int_{-\infty}^{\infty} f(X-v)g(v)dv \ . \tag{A·6}$$

Then, by the convolution theorem, well known to crystallographers, the c.f. (Fourier transform) of $P(X)$, which we shall denote by $C_X(u)$, is simply the product of the c.f.'s of $f(x)$ and $g(x)$:

$$C_X(u) = C_{x_1}(u)C_{x_2}(u) \ . \tag{A·7}$$

Then $P(X)$ is obtained simply by taking the Fourier transform of (A·7):

$$P(X) = \frac{1}{2\pi}\int_{-\infty}^{\infty} C_X(u)\exp[-iuX]du \ . \tag{A·8}$$

In the case where the two random variables have the same probability distribution $f(x)$,

$$C_X(u) = [C_x(u)]^2 \ .$$

The last result is easily generalized to the case of the sum of t random variables:

$$X = x_1+x_2+\ldots+x_t \ . \tag{A·9}$$

Here

$$P(X) = \frac{1}{2\pi}\int C_X(u) \exp\,[-iuX]du\,, \qquad (\text{A·10})$$

where

$$C_X(u) = [C_x(u)]^t\,. \qquad (\text{A·11})$$

Note that there are exactly analogous results for the m.g.f. $M(u)$, and in practice it is more convenient to work with the latter to avoid having to use the complex i.

When X is a linear combination, rather than a sum, of random variables

$$X = a_1x_1+a_2x_2+\ldots+a_tx_t \qquad (\text{A·12})$$

then (A·7) is easily generalized to

$$C_X(u) = C_{x_1}(a_1u)\ldots C_{x_t}(a_tu)\,, \qquad (\text{A·13})$$

or, when all the a_i are equal to a, and all the x's have the same distribution,

$$C_X(u) = [C_x(au)]^t\,. \qquad (\text{A·14})$$

Now while the moments of a distribution are a set of parameters which are useful for measuring its properties, and for specifying it, they are not, however, the only set of parameters for this purpose, or even the best set. Another series of parameters, the so-called *cumulants*, have properties which are more useful from the theoretical standpoint.

Formally the cumulants k_r are defined by the expansion of the cumulant generating function (c.g.f.) $K(u)$

$$\begin{aligned} K(u) &= \log M(u) \\ &= k_1u+k_2\frac{u^2}{2!}+k_3\frac{u^3}{3!}+\ldots\,, \end{aligned} \qquad (\text{A·15})$$

or alternatively,

$$\begin{aligned} K(iu) &= \log C(u) \\ &= k_1(iu)+k_2\frac{(iu)^2}{2!}+\ldots\,. \end{aligned} \qquad (\text{A·16})$$

Note that it is sometimes more convenient to write the same equations as

$$M(u) = \exp\,[K(u)]$$

or

$$C(u) = \exp\,[K(iu)]\,. \qquad (\text{A·17})$$

The relations between cumulants and moments may be found by equating coefficients of powers of u in the expansions

$$k_1u+k_2\frac{u^2}{2!}+k_3\frac{u^3}{3!}+\ldots = \log\left(1+m_1u+m_2\frac{u^2}{2!}+\ldots\right)\,, \qquad (\text{A·18})$$

and the results are widely tabulated.

In terms of moments about the mean m_r, the cumulants k_r are

$$\begin{aligned} k_1 &= m_1 = \text{mean}\,, \\ k_2 &= m_2 = \text{variance}\,, \\ k_3 &= m_3\,, \\ k_4 &= m_4-3m_2^2\,, \\ k_5 &= m_5-10m_2m_3\,. \end{aligned} \qquad (\text{A·19})$$

The coefficients of skewness and excess, which are measures of the departure of a distribution from the normal, are $k_3/k_2^{3/2}$ and k_4/k_2^2 respectively.

The important property of cumulants for our purpose is that, in taking the sum of a set of random variables, the cumulants of the individual variables are additive. This introduces a great simplification in the mathematical manipulations. For it follows from (A·7) that

$$\begin{aligned} K_X(u) &= \log C_X(u) \\ &= \log C_{x_1}(u)+\log C_{x_2}(u) \\ &= K_{x_1}(u)+K_{x_2}(u)\,, \end{aligned} \qquad (\text{A·20})$$

and hence the coefficients in the expansion of the right-hand side of this equation are also additive.

We can obtain the characteristic function for the distribution of the sum X of t identical random variables (A·9) by using (A·17),

$$C_X(u) = \exp\,[K_X(iu)] = \exp\,[tK_x(iu)]\,, \qquad (\text{A·21})$$

where $K_x(u)$ is the c.g.f. of a single random variable. When the random variables are not identical, we shall have instead

$$C_X(u) = \exp\left[\sum_{i=1}^{t} K_{x_i}(u)\right]\,. \qquad (\text{A·22})$$

APPENDIX B

Generating functions and cumulants of multivariate distributions

We shall illustrate the definitions required for the joint probability distribution of a number of random variables by the bivariate case. The extension to more than two random variables is obvious. If $f(x, y)$ is the joint probability distribution of the random variables x and y, the mixed moments are defined by

$$m_{rs} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^ry^sf(x, y)\,dx\,dy\,. \qquad (\text{B·1})$$

Note that the first suffix in m_{rs} always refers to the first random variable considered, x, and the second to y. This conventional notation will be followed for mixed cumulants as well.

In the various generating functions, it is now necessary to introduce a carrying variable u_i for each random variable. Thus, for instance, the characteristic function is now

$$\begin{aligned} C(u_1, u_2) &= \mathscr{E} \exp\,[i(u_1x+u_2y)] \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\,[i(u_1x+u_2y)]f(x, y)\,dx\,dy \end{aligned} \qquad (\text{B·2})$$

$$= 1+i(m_{10}u_1+m_{01}u_2)$$
$$+\frac{i^2}{2!}(m_{20}u_1^2+2m_{11}u_1u_2+m_{02}u_2^2)$$
$$+\frac{i^3}{3!}(m_{30}u_1^3+\dots)+\dots . \qquad \text{(B·3)}$$

The c.g.f. will be related to the m.g.f. by the equations:

$$K(u_1, u_2) = \log M(u_1, u_2) \qquad \text{(B·4)}$$

or

$$\exp[K(u_1, u_2)] = M(u_1, u_2)\ , \qquad \text{(B·5)}$$

and the mixed cumulants to the mixed moments by

$$\frac{k_{10}}{1!\,0!}u_1+\frac{k_{01}}{0!\,1!}u_2+\frac{k_{20}}{2!\,0!}u_1^2+\frac{k_{11}}{1!\,1!}u_1u_2+\frac{k_{02}}{0!\,2!}u_2^2$$
$$+\dots+\frac{k_{rs}}{r!\,s!}u_1^ru_2^s+\dots$$
$$= \log\left(1+\frac{m_{10}}{1!\,0!}u_1+\frac{m_{01}}{0!\,1!}u_2+\frac{m_{20}}{2!\,0!}u_1^2+\dots\right). \qquad \text{(B·6)}$$

Instead of expanding the right-hand side of (B·6), the relation between cumulants and moments can most easily be derived by a process due to Kendall (1943), which considerably simplifies the labour. This method has been used by Cook (1951) to derive relations for the bivariate case, and we reproduce some of her results that we shall need here. When the origin is at the mean ($k_{10} = k_{01} = 0$):

Order

$$\begin{array}{ll}
2 & k_{20} = m_{20} \\
 & k_{11} = m_{11} \\
3 & k_{30} = m_{30} \\
 & k_{21} = m_{21} \\
4 & k_{40} = m_{40}-3m_{20}^2 \\
 & k_{31} = m_{31}-3m_{20}m_{11} \\
 & k_{22} = m_{22}-m_{20}m_{02}-2m_{11}^2 \\
5 & k_{50} = m_{50}-10m_{30}m_{20} \\
 & k_{41} = m_{41}-4m_{30}m_{11}-6m_{21}m_{20} \\
 & k_{32} = m_{32}-m_{30}m_{02}-6m_{21}m_{11}-3m_{12}m_{20}\ . \qquad \text{(B·7)}
\end{array}$$

The relations are symmetrical in the suffixes, and are the generalizations of relations (A·19), from which they are in fact derived. In the application to statistical problems in crystallography, many of the moments on the right-hand side are zero (e.g. for a random distribution of atoms $m_{30} = 0$, and for space group $P\bar{1}$ $m_{11} = 0$).

A search of the statistical literature has not revealed any tables of relations for the trivariate case, but it is easy to derive those we need from (B·7) by using Kendall's (1943) method. The results not already contained in (B·7) (one can add as many zero suffixes as required to any of the relations) are:

Order

$$\begin{array}{ll}
3 & k_{111} = m_{111} \\
4 & k_{211} = m_{211}-m_{200}m_{011}-2m_{110}m_{101} \\
5 & k_{311} = m_{311}-3m_{210}m_{101}-m_{300}m_{011} \\
 & \qquad -3m_{111}m_{200}-3m_{201}m_{110} \\
 & k_{221} = m_{221}-m_{200}m_{021}-4m_{111}m_{110} \\
 & \qquad -2m_{120}m_{101}-2m_{210}m_{011}-m_{020}m_{201}\ . \qquad \text{(B·8)}
\end{array}$$

These relations can be checked by seeing whether they reduce to (B·7) or (A·19) on amalgamating suffixes. We have written out the relations in full for, although many of the moments will vanish when the atoms are assumed to roam uniformly through the unit cell, they will not necessarily do so for other distributions that one might use (e.g. when restrictions are placed on the atomic positions, or the atomic positions are assumed to be correlated in some way).

APPENDIX C

Case of unequal atoms

The point of departure for deriving the joint probability distribution of a set of structure factors when the atoms in the cell are not all equal, but have the same probability distribution, is equation (2·3) of the text.

$$\mathscr{M}(u_1, \dots, u_m)$$
$$= M(\varphi_1u_1, \dots, \varphi_1u_m)\dots M(\varphi_tu_1, \dots, \varphi_tu_m)\ . \qquad \text{(C·1)}$$

Introducing the multivariate cumulant generating functions, as in (2·7),

$$\begin{array}{l} \mathscr{K} = \log \mathscr{M}\ , \\ K = \log M\ , \end{array} \qquad \text{(C·2)}$$

where $K(u_1, \dots, u_m)$ is given by equation (2·8); it then follows from (C·1) that

$$\mathscr{K}(u_1, \dots, u_m) = \sum_{j=1}^{t} K(\varphi_ju_1, \dots, \varphi_ju_m)\ . \qquad \text{(C·3)}$$

Here

$$K(\varphi_ju_1, \dots, \varphi_ju_m)$$
$$= \frac{k_{200\dots}\varphi_j^2u_1^2}{2!\,0!\,0!\dots}+\frac{k_{020\dots}\varphi_j^2u_2^2}{0!\,2!\,0!\dots}+\frac{k_{300\dots}\varphi_j^3u_1^3}{3!\,0!\,0!\dots}+\dots$$
$$= \varphi_j^2\ (\text{terms in } p+q+r\dots = 2)+\dots$$
$$+\varphi_j^\nu\ (\text{terms in } p+q+r\dots = \nu)+\dots .$$

Hence

$$\mathscr{K}(u_1, \dots, u_m)$$
$$= \left(\sum_{j=1}^{t}\varphi_j^2\right)(\text{terms in } p+q+r\dots = 2)+\dots$$
$$+\left(\sum_{j=1}^{t}\varphi_j^\nu\right)(\text{terms in } p+q+r\dots = \nu)+\dots . \qquad \text{(C·4)}$$

Since the atoms are unequal it is not possible to rewrite this expression directly in terms of standardized cumulants, as in equation (2·12). The second-order term does, however, simplify to

$$\tfrac{1}{2}u_1^2+\tfrac{1}{2}u_2^2+\ldots,$$

since

$$k_{200\ldots} = m_{200\ldots} = \langle\xi^2\rangle = n$$

and

$$n\sum_{j=1}^{t}\varphi_j^2 = 1\,.$$

To deal with the higher terms we introduce the notation

$$\zeta_\nu = \sum_{j=1}^{t}\varphi_j^\nu\,. \tag{C·4}$$

The moment-generating function is then

$$\begin{aligned}\mathscr{M}(u_1,\ldots,u_m) &= \exp\left[\mathscr{K}(u_1,\ldots,u_m)\right]\\ &= \exp\left[\tfrac{1}{2}(u_1^2+u_2^2+\ldots)+\sum_{\nu=3}^{\infty}\zeta_\nu L_\nu\right],\end{aligned} \tag{C·6}$$

where (cf. (2·15))

$$L_\nu = \mathop{\mathsf{S}}_{p+q+\ldots=\nu}\frac{k_{pqr\ldots}}{p!\,q!\,r!\ldots}u_1^p u_2^q u_3^r\ldots, \tag{C·5}$$

the sum S being taken over all partitions of ν. Note that the k's in this expression are unstandardized cumulants and that

$$L_\nu = n^{\nu/2}\Lambda_\nu\,. \tag{C·6}$$

Proceeding now exactly as in § 2, we obtain

$$\begin{aligned}&\exp\left[-\tfrac{1}{2}(u_1^2+u_2^2+\ldots)\right]\mathscr{M}(u_1,\ldots,u_m)\\ &=1\\ &\quad+\zeta_3L_3\\ &\quad+\zeta_4L_4+\tfrac{1}{2}\zeta_3^2L_3^2\\ &\quad+\zeta_5L_5+\zeta_3\zeta_4L_3L_4+\tfrac{1}{6}\zeta_3^3L_3^3\\ &\quad+\ldots,\end{aligned} \tag{C·7}$$

where we have kept the same ordering of terms as in (2·16).

The inversion of (C·7) follows just as described in the text in § 2. We obtain

$$\begin{aligned}P(E_1,\ldots,E_m) &= \frac{1}{(2\pi)^{m/2}}\exp\left[-\tfrac{1}{2}(E_1^2+E_2^2+\ldots+E_m^2)\right]\\ &\times\left\{1+\zeta_3\left[\frac{k_{300\ldots}}{3!\,0!\,0!\ldots}H_3(E_1)+\ldots\right.\right.\\ &\qquad+\frac{k_{210\ldots}}{2!\,1!\,0!\ldots}H_2(E_1)H_1(E_2)+\ldots\\ &\qquad\left.+\frac{k_{111}}{1!\,1!\,1!}H_1(E_1)H_1(E_2)H_1(E_3)+\ldots\right].\\ &\quad+\zeta_4\left[\frac{k_{400\ldots}}{4!\,0!\,0!\ldots}H_4(E_1)+\ldots\right.\\ &\qquad\left.+\frac{k_{310\ldots}}{3!\,1!\,0!\ldots}H_3(E_1)H_1(E_2)+\ldots\right]\\ &\quad\left.+\text{ etc.}\right\}.\end{aligned} \tag{C·8}$$

APPENDIX D

Calculation of mixed moments and cumulants

Some of the required results are already tabulated in Hauptman & Karle's monograph (1953). A good general method is that due to Bertaut (1955*a*, *b*), in which the trigonometric structure factor is expressed as the linear sum over the equivalent points of the space group. For $P\bar{1}$, one writes

$$\xi(\mathbf{h}) = 2\cos 2\pi\mathbf{h}.\mathbf{r} = S(\mathbf{h})+S^{-1}(\mathbf{h})\,, \tag{D·1}$$

where

$$S(\mathbf{h}) = \exp\,(2\pi i\mathbf{h}.\mathbf{r})\,. \tag{D·2}$$

(i) Then to calculate, say, m_{14} for the pair of structure factors $E_{2\mathbf{h}}\equiv E_{\mathbf{h}_1}$ and $E_{\mathbf{h}}\equiv E_{\mathbf{h}_2}$ (§ 3) one writes

$$\begin{aligned}\xi_{\mathbf{h}_1}\xi_{\mathbf{h}_2}^4 &= (S_{\mathbf{h}_1}+S_{\mathbf{h}_1}^{-1})(S_{\mathbf{h}_2}+S_{\mathbf{h}_2}^{-1})^4\\ &= (S_{\mathbf{h}_1}+S_{\mathbf{h}_1}^{-1})(S_{\mathbf{h}_2}^4+4S_{\mathbf{h}_2}^2+6+4S_{\mathbf{h}_2}^{-2}+S_{\mathbf{h}_2}^{-4})\\ &= 4S_{\mathbf{h}_1}S_{\mathbf{h}_2}^2+4S_{\mathbf{h}_1}^{-1}S_{\mathbf{h}_2}^{-2}+\ldots\\ &= 4+4+\ldots,\end{aligned} \tag{D·3}$$

since

$$S_{\mathbf{h}_1}S_{\mathbf{h}_2}^2 = S_{\mathbf{h}_1}^{-1}S_{\mathbf{h}_2}^{-2} = 1$$

because

$$\mathbf{h}_1+2\mathbf{h}_2 = 0\,. \tag{D·4}$$

Then

$$m_{14} = \langle\xi_{\mathbf{h}_1}\xi_{\mathbf{h}_2}^4\rangle = 4+4 = 8\,, \tag{D·5}$$

since the averages of the other terms in (D·3) are zero when the atoms are uniformly distributed throughout the cell.

In a similar way we find that the only non-vanishing moments are

$$\begin{aligned}m_{02} = m_{20} = 2,\ m_{12} = 2,\ m_{40} = m_{04} = 6,\\ m_{22} = 4,\ m_{32} = 6,\ m_{14} = 8\,.\end{aligned} \tag{D·6}$$

From these we obtain the cumulants k_{pq} by (B·7):

$$\begin{aligned}k_{02} = k_{20} = 2,\ k_{12} = 2,\ k_{40} = k_{04} = -6,\\ k_{22} = 0,\ k_{32} = -6,\ k_{14} = -16\,.\end{aligned} \tag{D·7}$$

Finally the standardized cumulants λ_{pq} are obtained by dividing the latter by $(\sqrt{k_{02}})^{p+q} = 2^{(p+q)/2}$:

$$\begin{aligned}\lambda_{02} = \lambda_{20} = 1,\ \lambda_{12} = 1/2^{\frac{1}{2}},\ \lambda_{04} = \lambda_{40} = -3/2,\\ \lambda_{22} = 0,\ \lambda_{32} = -3/2^{3/2},\ \lambda_{14} = -8/2^{3/2}\,.\end{aligned} \tag{D·8}$$

We have written the λ_{pq} in terms of powers of $1/2^{\frac{1}{2}}$ so that the latter may be easily incorporated into the powers of $1/t^{\frac{1}{2}}$ that occur in the expression (3·3) to give terms in $1/N^{\frac{1}{2}}$.

(ii) In § 4 we shall require the mixed moments and cumulants for the evaluation of $P(E_{\mathbf{h}}, E_{\mathbf{h}'}, E_{\mathbf{h}+\mathbf{h}'})$ for $P\bar{1}$. The moments

$$m_{pqr} = \langle\xi(\mathbf{h}_1)^p\xi(\mathbf{h}_2)^q\xi(\mathbf{h}_3)^r\rangle\,, \tag{D·9}$$

subject to the condition

$$\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3 = 0\,, \tag{D·10}$$

are found as in the example (D·3)–(D·5) above. The cumulants k_{pqr} are then found by using (B·8) above, and are finally standardized by dividing by $2^{(p+q+r)/2}$. The results (which are, of course, symmetrical in $\mathbf{h}_1, \mathbf{h}_2$ and $\mathbf{h}_3$) for the non-vanishing moments are:

pqr	200	111	400	220	113
m_{pqr}	2	2	6	4	6
k_{pqr}	2	2	-6	0	-6
λ_{pqr}	1	$1/2^{\frac{1}{2}}$	$-3/2$	0	$-3/2^{3/2}$

(D·11)

APPENDIX E

Calculation of the characteristics of a conditional probability distribution

(a) Bivariate distribution

The method we shall use is due to Wicksell (1917), in which the moments of the distribution (3·13) are evaluated in the form (3·17). We require to calculate

$$I_{ij}^{s} = \int_{-\infty}^{\infty} x^{s}\,\frac{\partial^{i+j}\varphi(x,y)}{\partial x^{i}\,\partial x^{j}}\,dx\,, \tag{E·1}$$

where

$$\varphi(x,y) = \varphi(x)\varphi(y) = \frac{1}{2\pi}\exp\left(-\tfrac{1}{2}x^{2}-\tfrac{1}{2}y^{2}\right). \tag{E·2}$$

It is easy to show by integration by parts of (E·1) that

$$I_{ij}^{s} = \begin{cases} (-1)^{i}\dfrac{s!}{(s-i)!}\,\dfrac{d^{j}I_{00}^{(s-i)}}{dy^{j}}\,, & \text{when } i \le s\,,\\ 0\,, & \text{when } i > s\,. \end{cases} \tag{E·3}$$

Furthermore,

$$\begin{aligned} I_{00}^{s} &= \varphi(y)\int_{-\infty}^{\infty} x^{s}\varphi(x)\,dx\\ &= \begin{cases} 1.3.5\ldots(2m-1)\varphi(y)\,, & \text{when } s = 2m\,,\\ 0\,, & \text{when } s = 2m+1\,. \end{cases} \end{aligned} \tag{E·4}$$

From (E·3) and (E·4) we obtain, by using (1·16) and (1·17),

$$\begin{cases} I_{00}^{0} = \varphi(y)\,,\\ I_{0j}^{0} = (-1)^{j}H_{j}(y)\varphi(y)\,, \end{cases} \qquad \begin{cases} I_{0j}^{3} = 0\,,\\ I_{1j}^{3} = 3(-1)^{j+1}H_{j}(y)\varphi(y)\,,\\ I_{2j}^{3} = 0\,,\\ I_{3j}^{3} = 6(-1)^{j+1}H_{j}(y)\varphi(y)\,. \end{cases}$$

$$\begin{cases} I_{0j}^{1} = 0\,,\\ I_{1j}^{1} = (-1)^{j+1}H_{j}(y)\varphi(y)\,, \end{cases}$$

$$\begin{cases} I_{0j}^{2} = (-1)^{j}H_{j}(y)\varphi(y)\,,\\ I_{1j}^{2} = 0\,,\\ I_{2j}^{2} = 2(-1)^{j}H_{j}(y)\varphi(y)\,, \end{cases} \tag{E·5}$$

Hence every term in the numerator and denominator of (3·17) contains the factor $\varphi(y)$, which thus cancels out completely, leaving simple expressions involving products of β_{ij} and Hermite polynomials H_j. As an example we shall write down the expression for the mean $\langle E_{2\mathbf{h}}\rangle_{E_{\mathbf{h}}}$. From (3·17),

$$\langle x\rangle_{y} = \frac{I_{00}^{1}+\beta_{03}I_{03}^{1}+\beta_{12}I_{12}^{1}+\beta_{04}I_{04}^{1}+\beta_{13}I_{13}^{1}+\ldots}{I_{00}^{0}+\beta_{03}I_{03}^{0}+\beta_{04}I_{04}^{0}+\ldots} \tag{E·6}$$

where we identify $E_{2\mathbf{h}}$ ($\equiv E_1$) with x and $E_{\mathbf{h}}$ ($\equiv E_2$) with y. Substituting from (E·5) into (E·6), and including only those β_{ij} that do not vanish in the expression (3·4) for $P(E_1, E_2)$, we have

$$\langle E_1\rangle_{E_2} = \frac{-\beta_{12}H_2(E_2)-\beta_{14}H_4(E_2)-\beta_{16}H_6(E_2)}{1+\beta_{04}H_4(E_2)}\,.$$

The values of β_{ij} are obtained by comparing (3·4) with (3·14), and we finally obtain

$$\langle E_1\rangle_{E_2} = \frac{\dfrac{1}{2N^{\frac{1}{2}}}H_2(E_2)-\dfrac{1}{N^{3/2}}\left[\tfrac{1}{3}H_4(E_2)+\tfrac{1}{16}H_6(E_2)\right]}{1-\dfrac{1}{8N}H_4(E_2)}\,. \tag{E·7}$$

(b) Trivariate distribution

When three random variables x, y, z are involved, as in § 4 where they are $E_{\mathbf{h}}$, $E_{\mathbf{h}'}$ and $E_{\mathbf{h}+\mathbf{h}'}$ respectively, the probability distribution is of the form (see (4·5))

$$\begin{aligned} &F(x,y,z)\\ &= \varphi(x,y,z)+\sum_{i+j+k\ge 3}\beta_{ijk}\,\frac{\partial^{i+j+k}}{\partial x^{i}\partial y^{j}\partial z^{k}}\,\varphi(x,y,z)\\ &= \varphi(x,y,z)\left[1+\sum_{i+j+k\ge 3}(-1)^{i+j+k}\beta_{ijk}H_i(x)H_j(y)H_k(z)\right], \end{aligned} \tag{E·8}$$

where

$$\begin{aligned} \varphi(x,y,z) &= \varphi(x)\varphi(y)\varphi(z)\\ &= (2\pi)^{-3/2}\exp\left[-\tfrac{1}{2}(x^{2}+y^{2}+z^{2})\right]. \end{aligned} \tag{E·9}$$

(i) The s^{th} moment of x as a function of y and z will then be given by

$$\mu_s'(x)_{y,z} = \frac{I_{000}^{s}+\sum\limits_{i+j+k\ge 3}\beta_{ijk}I_{ijk}^{s}}{I_{000}^{0}+\sum\limits_{i+j+k\ge 3}\beta_{ijk}I_{ijk}^{0}}\,, \tag{E·10}$$

where

$$I_{ijk}^{s} \equiv \int_{-\infty}^{\infty} x^{s}\,\frac{\partial^{i+j+k}\varphi(x,y,z)}{\partial x^{i}\partial y^{j}\partial z^{k}}\,dx\,. \tag{E·11}$$

Comparing (E·1) and (E·11), and using equation (F·2), it is obvious that

$$I_{ijk}^{s} = (-1)^{k}H_k(z)I_{ij}^{s}\,, \tag{E·12}$$

where the I_{ij}^{s} are given in (E·5).

The β_{ijk} are obtained by comparing (E·8) with the particular probability distribution to be considered, namely (4·5):

$$\left.\begin{array}{l}\beta_{111}=-1/N^{\frac{1}{2}},\ \beta_{400}=-1/8N,\ \beta_{222}=1/2N\ ,\\ \beta_{113}=1/2N^{3/2},\ \beta_{115}=1/8N^{3/2},\ \beta_{333}=-1/6N^{3/2}.\end{array}\right\}\qquad(\mathrm{E}{\cdot}13)$$

The calculation of the mean $\langle x\rangle_{y,z}$ is then straightforward. To obtain the variance, the second moment $\mu_2'(x)_{y,z}$ is first calculated from (E·10), and then equation (3·8) is used. The final results are given in equations (4·12) and (4·13).

(ii) Further quantities required for the case of the trivariate distribution (see § 4·4) are the moments of the product xy as a function of z. We shall denote the mean and second moment by $\langle xy\rangle_z$ and $\langle x^2y^2\rangle_z$ respectively.

Quite generally, the product moment $\langle x^sy^t\rangle_z$ of the trivariate distribution (E·8) is given by

$$\langle x^sy^t\rangle_z=\frac{I^{st}_{000}+\sum\limits_{i+j+k\geq 3}\beta_{ijk}I^{st}_{ijk}}{I^{00}_{000}+\sum\limits_{i+j+k\geq 3}\beta_{ijk}I^{00}_{ijk}}\ ,\qquad(\mathrm{E}{\cdot}14)$$

where

$$I^{st}_{ijk}=\int_{-\infty}^{\infty}x^sy^t\,\frac{\partial^{i+j+k}\varphi(x,\,y,\,z)}{\partial x^i\,\partial y^j\,\partial z^k}\,dx\,dy\ .\qquad(\mathrm{E}{\cdot}15)$$

It is easy to show just as for the bivariate case that

$$I^{st}_{ijk}=\begin{cases}(-1)^{i+j}\dfrac{s!\,t!}{(s-i)!\,(t-j)!}\dfrac{d^k}{dz^k}I^{s-i,\,t-j}_{000}, & \\ \qquad\qquad\text{when}\quad i\leq s,\ j\leq t\ , & \\ 0,\qquad\qquad\text{when}\quad i>s\quad\text{or}\quad j>t\ , & \end{cases}\qquad(\mathrm{E}{\cdot}16)$$

and that

$$I^{st}_{000}=\begin{cases}1.3.5\ldots(2m-1).\ 1.3.5\ldots(2n-1)\varphi(z)\ , & \\ \qquad\qquad\text{when}\quad s=2m,\ t=2n\ , & \\ 0,\qquad\qquad\text{when}\quad s=2m+1\quad\text{or}\quad t=2n+1\ . & \end{cases}\qquad(\mathrm{E}{\cdot}17)$$

From (E·16) and (E·17) we find, *omitting the common factor* $\varphi(z)$,

$$\begin{array}{ll}I^{00}_{00k}=(-1)^kH_k(z)\ , & \left\{\begin{array}{l}I^{33}_{00k}=0\ ,\\ I^{33}_{11k}=9(-1)^kH_k(z)\ ,\\ I^{33}_{22k}=0\ ,\\ I^{33}_{33k}=36(-1)^kH_k(z)\ ,\\ I^{33}_{13k}=18(-1)^kH_k(z)\ .\end{array}\right.\\ \left\{\begin{array}{l}I^{11}_{00k}=0\ ,\\ I^{11}_{11k}=(-1)^kH_k(z)\ ,\end{array}\right. & \\ \left\{\begin{array}{l}I^{22}_{00k}=(-1)^kH_k(z)\ ,\\ I^{22}_{11k}=0\ ,\\ I^{22}_{22k}=4(-1)^kH_k(z)\ ,\end{array}\right. & \end{array}\qquad(\mathrm{E}{\cdot}18)$$

APPENDIX F

The Hermite polynomials

The Hermite polynomials $H_\nu(x)$ are defined by the equation

$$H_\nu(x)=(-1)^\nu\exp\left[\tfrac{1}{2}x^2\right]\frac{d^\nu}{dx^\nu}\exp\left[-\tfrac{1}{2}x^2\right]\qquad(\mathrm{F}{\cdot}1)$$

or

$$\varphi^{(\nu)}(x)=(-1)^\nu H_\nu(x)\varphi(x)\ ,\qquad(\mathrm{F}{\cdot}2)$$

where

$$\varphi(x)=(2\pi)^{-\frac{1}{2}}\exp\left(-x^2/2\right)\ .\qquad(\mathrm{F}{\cdot}3)$$

The lowest-order polynomials are

$$\begin{array}{ll}H_0(x)=1\ , & H_4(x)=x^4-6x^2+3\ ,\\ H_1(x)=x\ , & H_5(x)=x^5-10x^3+15x\ ,\\ H_2(x)=x^2-1\ , & H_6(x)=x^6-15x^4+45x^2-15.\\ H_3(x)=x^3-3x\ , & \end{array}$$

We take the following table of values from Wicksell (1917):

x	$H_2(x)$	$H_3(x)$	$H_4(x)$	$H_5(x)$	$H_6(x)$
0·0	−1·000	0·000	+3·000	0·000	−15·000
0·5	−0·750	−1·375	+1·563	+6·281	−4·672
1·0	−0·000	−2·000	−2·000	+6·000	+16·000
1·5	+1·250	−1·125	−5·438	−3·656	+21·703
2·0	+3·000	+2·000	−5·000	−18·000	−11·000
2·5	+5·250	+8·125	+4·563	−21·094	−75·547
3·0	+8·000	+18·000	+30·000	+18·000	−96·000
3·5	+11·250	+32·375	+79·563	+148·969	+123·578

References

Bertaut, E. F. (1955*a*). *Acta Cryst.* **8**, 537.

Bertaut, E. F. (1955*b*). *Acta Cryst.* **8**, 544.

Bertaut, E. F. (1955*c*). *Acta Cryst.* **8**, 823.

Bertaut, E. F. (1956*a*). Two lectures delivered at the Cavendish Laboratory, Cambridge.

Bertaut, E. F. (1956*b*). *Bull. Soc. franç. Minér. Crist.* **79**, 392.

Bertaut, E. F. (1956*c*). *Acta Cryst.* **9**, 769.

Bertaut, E. F. & Blum, P. (1956). *Acta Cryst.* **9**, 121.

Bertaut, E. F., Blum, P. & Magnano, G. (1955). *C. R. Acad. Sci., Paris*, **241**, 757.

Bertaut, E. F. & Dulac, J. (1955). *Tables de linearisation des Produits et Puissances des Facteurs de Structure.* Grenoble: Laboratoire d'Electrostatique et de Physique du Metal, Institut Fourier.

Bouman, J. (1956). *Acta Cryst.* **9**, 777.

Bullough, R. K. & Cruickshank, D. W. J. (1955). *Acta Cryst.* **8**, 29.

Cochran, W. (1952). *Acta Cryst.* **5**, 65.

Cochran, W. (1953). *Acta Cryst.* **6**, 810.

Cochran, W. (1954). *Acta Cryst.* **7**, 581.

Cochran, W. & Douglas, A. S. (1955). *Proc. Roy. Soc.* A, **227**, 486.

Cochran, W. & Woolfson, M. M. (1954). *Acta Cryst.* **7**, 450.

Cochran, W. & Woolfson, M. M. (1955). *Acta Cryst.* **8**, 1.

Cook, M. B. (1951). *Biometrika*, **38**, 179.

Cramér, H. (1946). *Mathematical Methods of Statistics.* Princeton: University Press.

Grant, D. F., Howells, R. G. & Rogers, D. (1957). *Acta Cryst.* **10**, 489.

Hauptman, H. & Karle, J. (1953). *The Solution of the Phase Problem: I. The Centrosymmetric Crystal.* A.C.A. Monograph No. 3. New York: Polycrystal Book Service.

Hauptman, H. & Karle, J. (1954). *Acta Cryst.* **7**, 452.

HAUPTMAN, H. & KARLE, J. (1957). *Acta Cryst.* **10**, 267.
HUGHES, E. W. (1953). *Acta Cryst.* **6**, 871.
KARLE, I., HAUPTMAN, H., KARLE, J. & WING, A. B. (1957). *Acta Cryst.* **10**, 481.
KARLE, J. & HAUPTMAN, H. (1950). *Acta Cryst.* **3**, 181.
KENDALL, M. G. (1943). *The Advanced Theory of Statistics*, vol. 1. London: Griffin.
KITAIGORODSKI, A. I. (1954). *Travaux de l'Institut de Cristallographie* (Communications au IIIe. Congres International de Cristallographie). Moscou: Academie des Sciences de l'U.R.S.S.
SAYRE, D. (1952). *Acta Cryst.* **5**, 60.
SAYRE, D. (1953). *Acta Cryst.* **6**, 430.
SIMERSKÁ, M. (1956). *Czech. J. Phys.* **6**, 1.
SLACK, M. (1946). *J. Inst. Elec. Eng.* **93** (Part III), 76.
VAND, V. & PEPINSKY, R. (1953). *The Statistical Approach to X-ray Structure Analysis*. State College: X-ray and Crystal Analysis Laboratory of the Pennsylvania State University.
VAND, V. & PEPINSKY, R. (1954). *Acta Cryst.* **7**, 451.
WICKSELL, S. D. (1917). *Kungl. Sv. Vet. Akad. Handl.* **58**, Nr. 3.
WOOLFSON, M. M. (1954). *Acta Cryst.* **7**, 61.
WOOLFSON, M. M. (1957). *Acta Cryst.* **10**, 116.

Acta Cryst. (1958). **11**, 543

Matrix-Theoretical Derivation of Inequalities

BY INAO TAGUCHI AND SHIGEO NAYA
Faculty of Science, Osaka University, Osaka, Japan

(*Received* 10 *January* 1958)

Using the matrix theory, a general type of inequalities for $P1$ and $P\bar{1}$ is derived. The results obtained for $P1$ are the same as those derived by Karle & Hauptman.

1. Introduction

Since Harker & Kasper (1948) derived inequalities among the structure factors, the same problem has been examined by many authors. Based upon the fact that the electron densities are non-negative, Karle & Hauptman (1950) derived some general limitations to be imposed upon the relations among the structure factors. Using their method, Bouman (1956) derived a complete set of fundamental inequalities for the structure factors possessing a centre of symmetry. The present paper will show that the same relations can be derived in a more compact form by means of the matrix representation of the Fourier transformation between the structure factors and electron densities.

2. Representation in matrix form

For the sake of simplicity we shall treat the one-dimensional case, though the method can be extended to the cases of more than one dimension without encountering any great difficulty. The electron density $\varrho(x)$ in the unit of the length L is represented by

$$\varrho(x) = \frac{1}{L}\sum_{h=-\infty}^{+\infty} F_h \exp\,[2\pi i h x], \quad h = 0, \pm 1, \pm 2, \ldots; \quad -\tfrac{1}{2} \le x \le \tfrac{1}{2}. \quad (1)$$

Let us divide our unit cell into $(2N+1)$ equal ranges, $\Delta \equiv L/(2N+1)$. Then the following formula approaches to equation (1) asymptotically with increase of the number of division:

$$\varrho\left(\frac{r}{2N+1}\right) = \frac{1}{(2N+1)\Delta}\sum_{h=-N}^{+N} F_h \exp\left[2\pi i h \frac{r}{2N+1}\right], \quad h, r = -N, \ldots, 0, \ldots, N. \quad (2)$$

Defining

$$\varrho_r \equiv \varrho\left(\frac{r}{2N+1}\right)\Delta,$$

(2) is replaced by

$$\varrho_r = \frac{1}{2N+1}\sum_{h=-N}^{+N} F_h \exp\left[2\pi i \frac{hr}{2N+1}\right], \quad h, r = -N, \ldots, 0, \ldots, N. \quad (3)$$

The Fourier series (3) can now be represented in terms of a $(2N+1)$-dimensional matrix product as follows:

$$\boldsymbol{\varrho} = \boldsymbol{UFU}^{-1}, \quad (4)$$

where

$$\boldsymbol{\varrho} = \begin{bmatrix} \varrho_{-N} & & & & & & \\ & \ddots & & & & & \\ & & \varrho_{-r} & & 0 & & \\ & & & \ddots & & & \\ & & & & \varrho_0 & & \\ & & & & & \ddots & \\ & & 0 & & & \varrho_r & \\ & & & & & & \ddots \\ & & & & & & & \varrho_N \end{bmatrix}, \quad (5)$$