

Fourier Representations of Pdf's Arising in Crystallography

George H. Weiss

National Institutes of Health, Bethesda, MD 20205

and

Uri Shmueli

Tel-Aviv University, Ramat-Aviv 69978, Tel-Aviv, Israel

Accepted: June 24, 1985

A survey is given of some recent calculations of univariate and multivariate probability density functions (pdf's) of structure factors used to interpret crystallographic data. We have found that in the presence of sufficient atomic heterogeneity the frequently used approximations derived from the central limit theorem in the form of Edgeworth or Gram-Charlier series can be quite unreliable, and in these cases the more exact, but lengthier, Fourier calculations must be made.

Key words: characteristic functions; direct methods of phase determination; Fourier series; intensity statistics.

Few scientific disciplines depend so heavily on techniques based on the central limit theorem and associated expansions in orthogonal polynomials as does crystallography. Ever since the pioneering work of Wilson [1,2],¹ and Karle and Hauptman [3-5], the central limit theorem has played a vital role in translating crystallographic scattering data into structural information and, indeed, it is built into many computer routines for this purpose. As we will show, when the central limit theorem is applied to data from unit cells with a considerable variation in the atomic weights of the constituent atoms it can lead to serious qualitative errors. That this

is true is well known to crystallographers who have made heavy use of Edgeworth and related expansions to correct zeroth order approximations based on the central limit theorem [5,6]. It is not generally appreciated, however, that serious errors can persist even with these correction terms, provided that atomic heterogeneity is sufficiently great. This suggests the value that may be attached to exact results when these are available and are readily computed. This paper reports on recent efforts we and several collaborators [7-11] have made in this direction.

Two general classes of probabilistic methods are used to deduce structural information from radiation intensity diffracted from crystals, the so-called intensity statistics and direct methods of phase determination. In order to make this exposition self-contained, we will sketch how such information can be derived from data

About the Authors: George H. Weiss is an applied mathematician and Uri Shmueli is a chemist.

¹ Figures in brackets indicate literature references.

on intensities in a particularly simple case, and refer the interested reader to two monographs that give detailed accounts of these subjects in more general cases [12,13]. The arrangement of atoms in a unit cell of a crystal is most often restricted by the space group to which the crystal belongs [14], and in the general case, only the arrangement within the asymmetric part of the cell needs to be determined. The intensity of the diffracted radiation can be represented in terms of structure factors $F(\mathbf{h})$, where the vector \mathbf{h} and its components (h, k, l), the orders of diffraction, specify the geometric relation between incident and scattered beams and their relative orientation to the basis vectors of the lattice of the diffracting crystal [14]. The structure factors are Fourier coefficients of the (periodic) density function of the scattering matter, and both their magnitude and phase are required in order to reconstruct the density—i.e., the actual atomic arrangement. Thus, $F(\mathbf{h})$ is in general a complex quantity, which we write as $F(\mathbf{h}) = A(\mathbf{h}) + iB(\mathbf{h})$. The function $F(\mathbf{h})$ can be expressed as a sum of contributions from individual atoms in the unit cell as

$$F(\mathbf{h}) = \sum_j f_j \exp(2\pi i \mathbf{h} \cdot \mathbf{r}_j) = \sum_j f_j \exp(i\theta_j) \quad (1)$$

where \mathbf{r}_j is the location of atom j , the f_j are so-called scattering or form factors which can be approximated, in the normalized-structure-factor representation (see below), by the atomic numbers of the corresponding atoms, and $\theta_j = 2\pi \mathbf{h} \cdot \mathbf{r}_j$. The space of \mathbf{h} is surveyed by varying the orientation of the crystal with respect to the incident beam.

Since $F(\mathbf{h})$ is a complex quantity, it can be represented as a vector in a plane which is the sum of n vectors, the j 'th being $f_j \exp(i\theta_j)$. The fundamental difficulty faced by crystallographers is that only the magnitude $|F|$ is measurable (although some recent work may change this situation [15]), and the phase of $F(\mathbf{h})$ must be inferred indirectly. To do so, one can establish a correspondence between the vector F and a random walk first studied by Pearson [16]. Using theorem of Weyl [17], one can show that if the components of \mathbf{r} are rationally independent, i.e., there exists no vector of integers \mathbf{m} such that $\mathbf{m} \cdot \mathbf{r} = \text{integer}$, then the set of angles, $\{\theta_j\}$, can be regarded as consisting of independent random variables, each of which is uniformly distributed over the interval $(0, 1)$ [17]. Thus the properties of the $F(\mathbf{h})$ can be determined using probabilistic methods, as was first pointed out by Wilson [1,2].

For a typical and important case in which probabilistic techniques allow one to derive structural information, consider how one can distinguish between centrosymmetric and noncentrosymmetric (space groups $P\bar{1}$ and $P1$, respectively) unit cells on the basis of

intensity statistics alone. A centrosymmetric unit cell is one in which, for every atom located at \mathbf{r}_j , there is an identical one at $-\mathbf{r}_j$. Consequently if we write $F = A + iB$ where

$$A = \sum_j f_j \cos \theta_j, \quad B = \sum_j f_j \sin \theta_j \quad (2)$$

it follows that $B \equiv 0$ by symmetry in the presence of centrosymmetry. When the unit cell is noncentrosymmetric B is not necessarily equal to 0. Hence the value of F can be represented as a one-dimensional random walk in $P\bar{1}$ and by a two-dimensional random walk in $P1$. In what follows we will use the physics notation that " $\langle \rangle$ " denotes the average of the variable contained in brackets. It will also prove convenient to work with the normalized structure factor $E = F / \langle |F|^2 \rangle^{1/2}$ which, since $\langle F \rangle = 0$, has the property that $\langle |E|^2 \rangle = 1$. Wilson's argument uses the central limit theorem to deduce the pdf of scattered intensities. In $P\bar{1}$, for which $B \equiv 0$, the form of the pdf of E that follows from the central limit theorem is

$$p(|E|) = (2/\pi)^{1/2} \exp(-E^2/2). \quad (3)$$

The corresponding pdf for the two dimensional case for unit cells without a crystallographic center of symmetry is

$$p(|E|) = 2|E| \exp(-|E|^2). \quad (4)$$

The qualitative difference between eqs (3) and (4) thus allows the experimental distinction to be drawn purely on a comparison of intensity data with the two forms for the pdf.

Notice, however, that the use of the central limit theorem presupposes the validity of certain assumptions, the major one of which is the presence of a large number of atoms in the unit cell and the second of which is that the f_j appearing in eq (1) should not exhibit too great a heterogeneity. The first of these assumptions holds for most crystalline materials of interest, but the second may be violated particularly when there are a small number of atoms that are considerably heavier than the majority of atoms comprising the molecule. When this is the case it is customary to replace, e.g., eq (3) by the Edgeworth series

$$p(|E|) = (2/\pi)^{1/2} \exp(-E^2/2) \{1 + \sum_n a_n H_n(|E|/\sqrt{2})\} \quad (5)$$

where the n 'th coefficient, a_n is expressible as a linear combination of the moments of A in eq (2) and $H_n(x)$ is the n 'th Hermite polynomial. These are readily calculated for the simpler space groups [18], and all space-group results are available for fourth, sixth, and eighth

moments[19,20]. Furthermore, the Edgeworth expansion may also not be too useful in the presence of extreme heterogeneity. This is illustrated in figure 1 in which the asymmetric unit of a cell in $P\bar{1}$ consists of 14 carbon atoms and one uranium atom, with a ratio of f 's approximately equal to 15 $\frac{1}{2}$. With 0 or 2 moments the Edgeworth series fails to reproduce the maximum and the 4 and 8 moment approximation locates the maximum quite far from its actual position. It is therefore desirable to have an exact easily computable representation for the pdf which is robust with respect to changes in atomic heterogeneity.

Just such a representation was first suggested by Barakat in a study of the freely jointed chain as a model for polymer configurations [21] and of laser speckle [22]. Let us write $g_j = f_j / (\sum f_j^2)^{1/2}$ so that

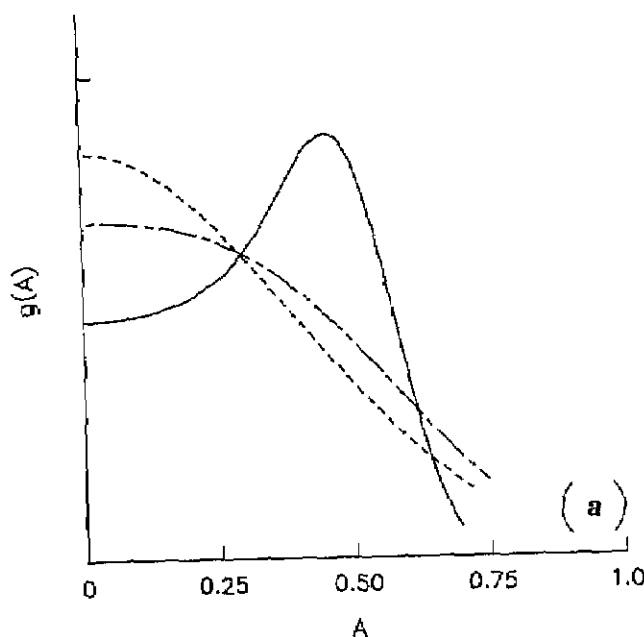
$$E = \sum g_j \exp(i\theta_j) = A + iB \quad (6)$$

and let us set

$$S = \sum g_j \quad (7)$$

so that $-S \leq A, B \leq S$. As an example we consider the case of a centrosymmetric unit cell for which $B=0$. The pdf of A , $g(A)$, has the property that it can differ from zero only in the interval $S^2 \geq A^2$. Within this interval we will expand $g(A)$ in a Fourier series:

$$g(A) = \frac{1}{2S} \left\{ 1 + 2 \sum_{n=1}^{\infty} \alpha_n \cos\left(\frac{\pi n A}{S}\right) \right\} \quad (8)$$



where

$$\begin{aligned} \alpha_n &= \int_{-S}^S g(A) \cos\left(\frac{\pi n A}{S}\right) \alpha A = \int_{-S}^S g(A) \cos\left(\frac{\pi n A}{S}\right) \alpha A \\ &= C\left(\frac{\pi n}{S}\right) \end{aligned} \quad (9)$$

where $C(\omega)$ is the characteristic function generated by $g(A)$. The Fourier series in eq (8) corresponds to a sampling theorem [23] for pdfs with a compact support. When the unit cell is noncentrosymmetric so that $B \neq 0$ in general, it is more convenient to expand the pdf of $|E| = (A^2 + B^2)^{1/2}$ in a Fourier-Bessel function series

$$p(|E|) = \frac{2|E|}{S^2} \sum_{i=0}^{\infty} D_i J_0(\gamma_i |E|/S) \quad (10)$$

where the γ_i are successive roots of $J_0(\gamma) = 0$ and the coefficients, D_i , are

$$D_i = C(\gamma_i/S) / J_1^2(\gamma_i) \quad (11)$$

again written in terms of the characteristic function.

Two questions that require an answer relate to the advantage of representations such as those in eqs (8) or (10) and the feasibility of numerical evaluation of the series. In the absence of atomic heterogeneity, or when there is a very large number of atoms in a unit cell, the central limit results are perfectly adequate for crys-

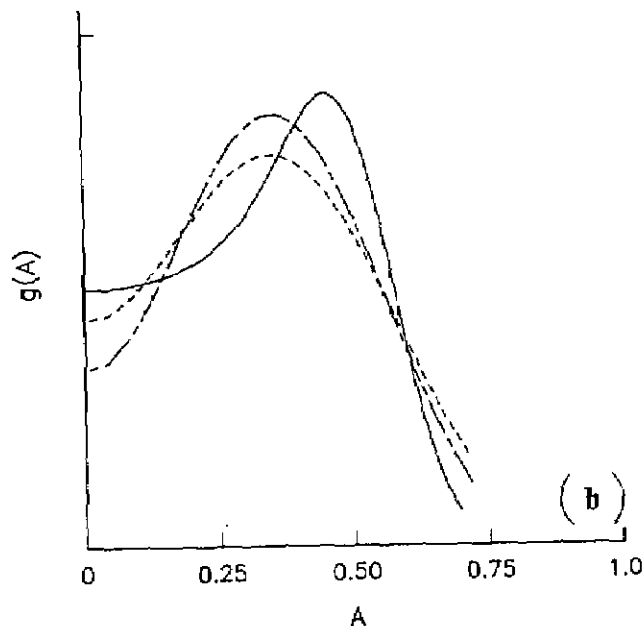


Figure 1a-(a) Approximations to the exact pdf $g(A)$ (denoted by the solid line) for a unit cell in space group $P\bar{1}$, consisting of 14 carbon atoms and a single uranium atom (atomic weight ratio 15 \pm 1) in the asymmetric unit. For convenience A_{\max} has been set equal to 1. Note that the pdf is symmetric around $A=0$. The approximations are a Gaussian (---) and the Gaussian corrected by two moments (---). (b) Approximations to the same pdf as in figure 1a by an Edgeworth series using 4 moments (---) and 8 moments (---).

tallographic applications. However, when there are fewer than about 40 atoms in the unit cell combined with one or two outstandingly heavy atoms, one has to convolve the Gaussian with the appropriate pdf for the heavy atoms [12]. In principle, using the series of eqs (8) and (10) finesses this difficulty, provided that the convergence properties are not overwhelming. In practice, in the case of intensity statistics we have found no problem in evaluating the Fourier or Fourier-Bessel function series, requiring no more than about 40 terms for the most extreme amounts of heterogeneity, and many fewer terms in the absence of heterogeneity. The evaluation of the analogous series for direct methods can present much tougher numerical problems, as we will see. Finally, a problem not so far discussed is the ease with which expressions for the characteristic function can be calculated. We have found that it is not too difficult to evaluate the characteristic function for all but a handful of space groups whose structure factor is found in the International Tables [24]. As an example, let us write the structure factor for $P\bar{1}$ as

$$A = 2 \sum_{j=1}^{n/2} g_j \cos \theta_j \quad (12)$$

where the g_j are known and the θ_j are uniformly distributed in $(0, 2\pi)$. The characteristic function is

$$\begin{aligned} C(\omega) &= \langle \exp(2i\omega \sum_{j=1}^{n/2} \cos \theta_j) \rangle \\ &= \prod_{j=1}^{n/2} C_j(\omega) \end{aligned} \quad (13)$$

where

$$\begin{aligned} C_j(\omega) &= \langle \exp(2i\omega g_j \cos \theta) \rangle \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(2i\omega g_j \cos \theta) d\theta = J_0(2\omega g_j). \end{aligned} \quad (14)$$

Other examples merely test one's ability to evaluate integrals. For example, we have recently examined the Fourier representation of the pdf of the intensity for a unit cell in $P\bar{1}$ in which there is an auxiliary or non-crystallographic center of symmetry located at \mathbf{d} , so that a single atom located at \mathbf{r}_j generates one at $-\mathbf{r}_j$ and $-\mathbf{r}_j \pm 2\mathbf{d}$ [9,10]. In this case one can show that

$$A(\mathbf{h}) = 4 \sum_{j=1}^{n/4} g_j \cos(2\pi \mathbf{h} \cdot \mathbf{d}) \cos[2\pi \mathbf{h} \cdot (\mathbf{r}_j - \mathbf{d})] \quad (15)$$

and the corresponding characteristic function is

$$C(\omega) = \frac{2}{\pi} \int_0^{\pi/2} \left(\prod_{j=1}^{n/4} J_0(4\omega g_j \cos \theta) \right) d\theta. \quad (16)$$

Again the numerical problems associated with this representation were not severe and allowed us to generalize the theory first presented by Rogers and Wilson for the equal-atom case [9,25]. It is possible though algebraically messy to generate the orthogonal polynomials corresponding to the Rogers-Wilson pdf, but the Fourier series representation is relatively straightforward. One can also analyze partially bicentric structures using the same techniques [10].

Our present development of Fourier representations of crystallographic pdf's has led us into the examination of direct methods in which one is interested in the joint pdf of several, usually correlated, structure factors. One of the simplest examples of these is the so-called Σ_1 relationship [13,26], in which one uses the joint pdf of $E(\mathbf{h})$ and $E(2\mathbf{h})$ to determine the probability that the phase of $E(2\mathbf{h})$ is positive given a knowledge of $|E(\mathbf{h})|$ and $|E(2\mathbf{h})|$. For simplicity we consider structures in $P\bar{1}$ letting $E(\mathbf{h}) = E$ and $E(2\mathbf{h}) = G$. Then

$$\begin{aligned} p(E, G) &= \frac{1}{4S^2} \sum_{r=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} C\left(\frac{\pi r}{S}, \frac{\pi s}{S}\right) \\ &\quad \cos\left(\frac{\pi r E}{S}\right) \cos\left(\frac{\pi s G}{S}\right) \end{aligned} \quad (17)$$

where

$$C(\omega_1, \omega_2) = \langle \exp(i\omega_1 E + i\omega_2 G) \rangle = \prod_{j=1}^{n/2} C_j(\omega_1, \omega_2) \quad (18)$$

in which since

$$G = 2 \sum_{j=1}^{n/2} g_j \cos(2\theta_j)$$

$$\begin{aligned} C_j(\omega_1, \omega_2) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp[2ig_j(\omega_1 \cos \theta + \omega_2 \cos 2\theta)] d\theta \\ &= R_j + iI_j \end{aligned} \quad (19)$$

where R_j and I_j can be expanded in terms of Bessel functions as

$$\begin{aligned} R_j(\omega_1, \omega_2) &= J_0(2g_j\omega_1)J_0(2g_j\omega_2) + 2 \sum_{m=1}^{\infty} \\ &\quad (-1)^m J_{4m}(2g_j\omega_1)J_{2m}(2g_j\omega_2) \end{aligned} \quad (20)$$

$$I_j(\omega_1, \omega_2) = 2 \sum_{m=0}^{\infty} (-1)^{m+1} J_{4m+2}(2g_j\omega_1)J_{2m+1}(2g_j\omega_2).$$

From eq (18) it follows that $C(\omega_1, \omega_2) = R(\omega_1, \omega_2) + iI(\omega_1, \omega_2)$ where R and I can be computed from the R_j and I_j . The probability that G is positive given E can now be written exactly as

$$p_+(2\mathbf{h}|\mathbf{h}) = \frac{1}{2}(1 + \frac{\Omega}{\Gamma}) \quad (21)$$

where

$$\Omega = \frac{1}{S^2} \sum_{s=1}^{\infty} \sum_{t=1}^{\infty} I\left(\frac{\pi s}{S}, \frac{\pi t}{S}\right) \cos\left(\frac{\pi s E}{S}\right) \sin\left(\frac{\pi t |G|}{S}\right)$$

$$\Gamma = \frac{1}{4S^2} \left\{ 1 + 2 \sum_{s=1}^{\infty} R\left(\frac{\pi s}{S}, 0\right) \left[\cos\left(\frac{\pi s E}{S}\right) + \cos\left(\frac{\pi t G}{S}\right) \right] + \right. \quad (22)$$

$$\left. + 4 \sum_{s=1}^{\infty} \sum_{t=1}^{\infty} R\left(\frac{\pi s}{S}, \frac{\pi t}{S}\right) \cos\left(\frac{\pi s E}{S}\right) \cos\left(\frac{\pi t G}{S}\right) \right\}.$$

The exact eq (21) should be compared to the much simpler approximation furnished by the use of the central limit theorem [26],

$$p_+(2\mathbf{h}|\mathbf{h}) \sim \frac{1}{2} \left[1 + \tanh\left(\frac{\sigma_2}{2\sigma_2^{3/2}} |G| (E^2 - 1) \right) \right] \quad (23)$$

where

$$\sigma_m = 2 \sum_{j=1}^{n/2} g_j^m. \quad (24)$$

Although eq (23) and generalizations of it are much used in the crystallographic literature, there has been no real test of its accuracy in the presence of heterogeneity, since until now there has been no attempt to calculate the exact pdf. A comparison of the result of evaluating eq (21) with that obtained from eq (23) for an assumed composition $C_{30}Kr_2$ in the half unit of a $P\bar{1}$ structure is shown in figure 2 for $G = 1.75$ [11]. A substantial difference between the two predictions is immediately evident. Further evidence of the inaccuracy of eq (23) in the presence of atomic heterogeneity is provided in figure 3 where we examine the effects of the variation in atomic weights for a unit cell in which the half unit is $C_{30}X_2$, where X varies. In the absence of heterogeneity eq (23) provides perfectly satisfactory results, but its utility decreases considerably with an increase in the atomic weight of the X atom.

We are presently examining the analogous properties of the Σ_2 relationship, in which one determines probabilistic relations between phases from properties of the joint pdf of $E(\mathbf{h})$, $E(\mathbf{k})$ and $E(-\mathbf{h}-\mathbf{k})$, which requires the evaluation of higher order Fourier series by the same basic techniques. While this investigation is very similar both in spirit and results to those for Σ_1 discussed in the last paragraphs, it appears to be much more difficult to evaluate the series for the pdf of the three-phase

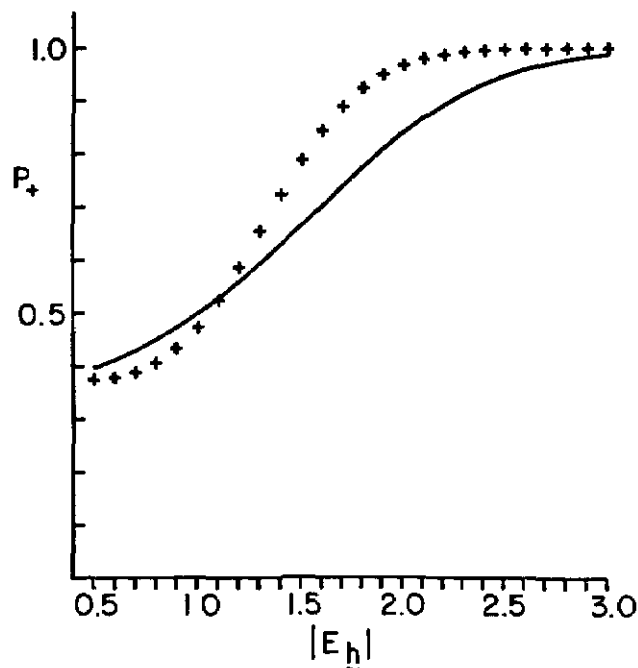


Figure 2—A graph of the exact expression for $p_+(2\mathbf{h}|\mathbf{h})$, the probability that the phase of $E(2\mathbf{h})$ is positive, as a function of E (+ + +) compared to the approximation provided by eq (23) (the solid curve) for a molecule with the assumed composition $C_{30}Kr_2$ in the asymmetric unit. The magnitude, $|E(2\mathbf{h})|$, was chosen equal to 1.75 for this example.

invariant, Φ , defined in terms of the phases of the triplet of structure factors $E(\mathbf{h})$, $E(\mathbf{k})$, $E(-\mathbf{h}-\mathbf{k})$, by

$$\Phi = \phi(\mathbf{h}) + \phi(\mathbf{k}) + \phi(-\mathbf{h}-\mathbf{k}). \quad (25)$$

To convey some notion of the difficulties we point out that the characteristic function to be evaluated is

$$C_j(\mathbf{w}) = \langle \exp \{ i g_j (\omega_1 A_1 + \omega_2 B_1 + \omega_3 A_2 + \omega_4 B_2 + \omega_5 A_3 + \omega_6 B_3) \} \rangle \quad (26)$$

where $E(\mathbf{h}) = A_1 + iB_1$, $E(\mathbf{k}) = A_2 + iB_2$, $E(-\mathbf{h}-\mathbf{k}) = A_3 + iB_3$. A detailed evaluation of $C_j(\mathbf{w})$ results in the expression

$$R_j(\mathbf{w}) = \prod_{k=1}^6 J_0(f_j \omega_k) + 2 \sum_{m=1}^{\infty} (-1)^m \prod_{k=1}^6 J_{2m}(f_j \omega_k)$$

$$I_j(\mathbf{w}) = 2 \sum_{m=0}^{\infty} (-1)^{m+1} \prod_{k=1}^6 J_{2m+1}(f_j \omega_k)$$

$$C(\mathbf{w}) = R(\mathbf{w}) + iI(\mathbf{w}) = \prod_j C_j(\mathbf{w}) = \prod_j (R_j + iI_j). \quad (27)$$

The resulting expression for the pdf of Φ is in terms of sevenfold Fourier series, each coefficient of which is an

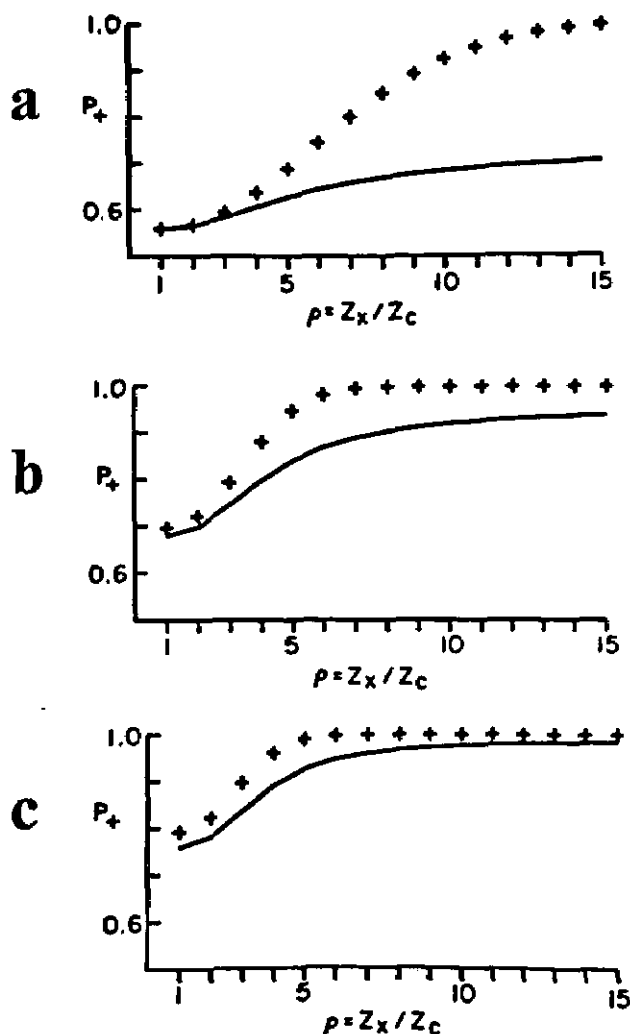


Figure 3—This figure shows the effects of heterogeneity on $p + (2h | h)$, as a function of the ratio of atomic numbers, $\rho = Z_x/Z_c$, for a molecule with the composition $C_{30}X_2$. The $++$'s are the exact results and the solid lines are the approximation of eq (23). The values chosen are (a) $|E(h)| = |E(2h)| = 1.5$, (b) $|E(h)| = |E(2h)| = 2.0$, (c) $|E(h)| = |E(2h)| = 2.25$. Note that the approximation is always on the conservative side. It is not known whether this is always true.

infinite series of the form shown in the last equation. Whether the resulting calculations can be made in a reasonable amount of time remains to be seen, but the difficulties to be overcome are exemplified by this problem.

A final word is in order about the philosophy behind the series of projects that we have undertaken. It would hardly be sensible to want to eliminate methods based on the central limit theorem that have served crystallographers so well in the past. However, it is useful to establish the limitations of these methods by having more exact representations available. Indeed we have explored such limitations in the case of tests based on the

Σ_1 relationship as indicated earlier and are presently considering more complicated crystallographic techniques. Furthermore, as the processing of crystallographic data becomes more and more automated it becomes increasingly attractive to have exact, rather than approximate formulae in the computer. We hope, in the coming years, to explore the feasibility of doing this for a variety of techniques, as well as contributing to the development of further ones based on the availability of exact representations.

References

- [1] Wilson, A. J. C. Determination of absolute from relative x-ray intensity data, *Nature* **150**, 151-152 (1942).
- [2] Wilson, A. J. C. The probability distribution of x-ray intensities, *Acta Crystallogr.* **2**, 318-321 (1949).
- [3] Hauptman, H., and J. Karle, Crystal structure determination by means of a statistical distribution of interatomic vectors, *Acta Crystallogr.* **5**, 48-54 (1952).
- [4] Hauptman, H., and J. Karle, The probability distribution of the magnitude of a structure factor, I. The centrosymmetric crystal, *Acta Crystallogr.* **6**, 131-135 (1953); II. The noncentro-symmetric crystal, *ibid* **6**, 136-141.
- [5] Hauptman, H., and J. Karle, *Solution of the Phase Problem I. The Centrosymmetric Crystal*, ACA Monograph No. 3 (Poly-crystal Book Service, New York) 1953.
- [6] Klug, A., Joint probability distributions of structure factors and the phase problem, *Acta Crystallogr.* **11**, 515-543 (1958).
- [7] Weiss, G. H., and J. E. Kiefer, The Pearson random walk with unequal step sizes, *J. Phys.* **A16**, 489-495 (1983).
- [8] Shmueli, U.; G. H. Weiss, J. E. Kiefer, and A. J. C. Wilson, Exact random walk models in crystallographic statistics. I. Space groups $P\bar{1}$ and $P1$. *Acta Crystallogr.* **A40**, 651-660 (1984).
- [9] Shmueli, U.; G. H. Weiss and J. E. Kiefer, Exact random walk models in crystallographic statistics, II. The biccetric distribution for the space group $P\bar{1}$, *Acta Crystallogr.* **A41**, 55-59 (1985).
- [10] Shmueli, U., and G. H. Weiss, Centric, biccetric, and partially biccetric intensity statistics in *Structure and Statistics in Crystallography* (Adenine Press, Quilderland New York) to appear.
- [11] Shmueli, U., and G. H. Weiss, Exact joint probability distribution for centrosymmetric structure factors. Derivation and application to the Σ_1 relationship in the space group $P\bar{1}$, *Acta Crystallogr.* (to appear).
- [12] Srinivisan, R. and Parthasarathy, S., *Some Statistical Applications in x-ray Crystallography*, Pergamon Press, Oxford (1976).
- [13] Giacovazzo, C., *Direct Methods in Crystallography*, Academic Press, New York (1980).
- [14] Woolfson, M. M., *An Introduction to x-ray Crystallography*, Cambridge University Press, Cambridge (1970).
- [15] Post B.; J. Nicolosi and J. Ladell, Experimental procedures for the determination of invariant phases of centrosymmetric crystals, *Acta Crystallogr.* **A40**, 684-688 (1984).
- [16] Pearson, K. The problem of random walk, *Nature* **72**, 294 (1905).
- [17] Hardy, G. H., and E. M. Wright, *Introduction to the Theory of Numbers*, Oxford University Press, Oxford, 5th ed. (1979).

- [18] Foster, F., and A. Hargreaves, The use of moments of x-ray intensity in space group determination I. Derivation of theoretical moments, *Acta Crystallogr.* **16**, 1124–1133 (1963); II. Practical application, *ibid* **16**, 1133–1138 (1963).
- [19] Shmueli, U., and U. Kaldor, Calculation of even moments of the trigonometric structure factor. Methods and results, *Acta Crystallogr.* **A37**, 76–80 (1981).
- [20] Shmueli, U., and U. Kaldor, Moments of the trigonometric structure factor, *Acta Crystallogr.* **A39**, 619–621 (1983).
- [21] Barakat, R., Isotropic random flights, *J. Phys.* **A6**, 796–804 (1973).
- [22] Barakat, R., First-order statistics of combined random sinusoidal waves with applications to laser speckle patterns, *Optica Acta* **21**, 921 (1974).
- [23] Hamming, R., *Numerical Methods for Scientists and Engineers*, (McGraw-Hill, New York, 2d ed. (1973).
- [24] *International Tables for x-ray Crystallography* ed. N.F.M. Henry, K. Lonsdale, Kynoch Press, Birmingham, 2d ed. (1965).
- [25] Rogers, D., and A. J. C. Wilson, The probability distribution of x-ray intensities. V. A note on some hypersymmetric distributions, *Acta Crystallogr.* **6**, 439–449 (1953).
- [26] Cochran, W., and M. M. Woolfson, The theory of sign relations between structure factors. *Acta Crystallogr.* **8**, 1–12 (1955).

DISCUSSION

of the Weiss-Shmueli paper,
Fourier Representations of Pdf's
Arising in Crystallography

E. Prince

Institute for Materials Science and Engineering
National Bureau of Standards

This interesting paper by Drs. Weiss and Shmueli represents a substantially exact solution of a problem that has concerned crystallographers for more than 35 years, the analysis in terms of atomic structure of x-ray diffraction data. (Similar information can be obtained from the diffraction of electrons and neutrons, but, for reasons that are both experimental and theoretical, this information is mainly used to supplement that obtained from x-ray diffraction, which remains the basic tool of the structural crystallographer.) The observed intensity in x-ray diffraction is given by

$$I = SL |F(\mathbf{h})|^2,$$

where S is a scale factor, L is a geometrical factor, and $F(\mathbf{h})$, commonly called the structure factor, is the Fourier transform of the electron density in a crystal. It may be written in the form

$$F(\mathbf{h}) = \int \rho(\mathbf{r}) \exp(2\pi i \mathbf{h} \cdot \mathbf{r}) d\mathbf{r}.$$

The density function, $\rho(\mathbf{r})$, in a crystal is periodic in three dimensions, so that it can be represented as a convolution of a function consisting of δ functions located at the nodes of a space lattice and a density defined in a small region known as a unit cell. Because of the periodicity the Fourier transform has appreciable values only at the nodes of a lattice in transform space, called by crystallographers the reciprocal lattice. Because it is a physical quantity, $\rho(\mathbf{r})$ is non-negative, and, furthermore, because a crystal is composed of atoms, it can be