RESEARCH PAPERS

Acta Cryst. (1994). A**50**, 9–17

Intensity Statistics. I

BY IAN R. CASTLEDEN* AND SUZANNE FORTIER

Department of Chemistry, Queen's University, Kingston, Ontario, Canada K7L 3N6

(*Received* 16 *June* 1992; *accepted* 18 *March* 1993)

Abstract

Two methods for evaluating the probability distribution of X-ray intensities are presented. The first is an exact series representation in powers of the atomic form factors. Overall thermal-displacement and scale parameters are shown to be dependent only on the ratio of intensity to σ_2. The second method is a maximum-entropy distribution based on the first two moments of the intensity distribution. It is shown to be a robust and accurate method of incorporating heavy-atom effects and results in a simple generalization of the Wilson distribution.

Introduction

The determination of scale and thermal-displacement parameters is usually based on a least-squares fit to the measured intensities under the restriction that $\langle I \rangle = \sigma_2$ (Wilson, 1942; Subramanian & Hall, 1982). This restriction results from the assumption that all possible atomic configurations within the asymmetric unit are, *a priori*, equally likely. Thus, the scattered X-ray intensities are, in the asymptotic limit, distributed according to the standard Wilson distributions: $\exp(-I/\sigma_2)/\sigma_2$ for acentric reflections and $\exp(-I/2\sigma_2)/(2\pi\sigma I)^{1/2}$ for centric reflections. This is a very cautious assumption, the advantage of which is that it will not yield estimates of structure parameters (scales, temperature factors *etc.*) that are not implied by the data. However, when the only property of the distribution used is its first moment, the actual working assumption is even weaker than the assumption that the structure is drawn from an ensemble where all atoms are independently and uniformly distributed in the unit cell. For example, there may exist intensity distributions $P(I)$ that have $\langle I \rangle = \sigma_2$ but that cannot possibly be generated by sampling any set of atomic configurations. Hence, it is worthwhile to consider estimation methods, such as the Bayesian approach of French & Oatley (1982), that use the full functional forms of the intensity probability distributions. It is important, however that these distributions be represented accurately or the advantages of 'unbiased' estimation will be lost. Indeed, Shmueli (1982) has shown that the actual probability distributions can deviate dramatically from the Wilson distribution. As a first step, therefore, a robust yet accurate method of evaluating probability distributions must be found.

The ultimate aim of this work is to use probability distributions within a Bayesian framework to estimate scale and atomic displacement parameters. Since this requires some form of iterative optimization, the derivatives of the distributions [with respect to the $f_\mu(\mathbf{h})$, for example] must be easily evaluated. The methods for deriving accurate probability distributions of structure-factor magnitudes described by Shmueli & Wilson (1982) and Shmueli & Weiss (1987) have been based on the cumulant expansion or on the Fourier–Bessel series. The cumulant expansions have coefficients that are difficult to calculate beyond the fifth order and hence can only be used to approximate the exact distribution under certain restrictive conditions. On the other hand, the Fourier–Bessel series (see Appendix III) are, in principle, exact but may require up to 40 terms to achieve convergence. Another drawback of Fourier–Bessel series is that their dependence on the atomic scattering factors is 'buried' in sums of products of the Bessel functions, making the evaluation of the derivatives difficult.

In this paper, two alternative expressions of intensity probability distributions are presented. These expressions have the advantage that inhomogeneity in cell content (*i.e.* variation in the relative scattering powers of the atoms) is expressed exclusively in terms of sums of powers of the scattering factors, $\sigma_n = \sum_{\mu=1}^{N} f_\mu^n(\mathbf{h})$, making the derivatives easy to evaluate. Comparisons with the true (Fourier–Bessel-series) distributions are made.

* Present address: Cambridge Crystallographic Data Centre, 12 Union Road, Cambridge CB2 1EZ, England.

A convergent-series representation

Exact joint probability distributions (j.p.d.'s) between structure factors have been derived by Peschar & Schenk (1986) and Castleden (1987). Since all phase terms in the exact j.p.d. between structure factors appear within cosine invariants [see, for example, Castleden, 1987, equation (1)], they are easily integrated, so the j.p.d. between structure-factor magnitudes is given, to a first approximation, simply by the product of the zeroth-order (phase-independent) terms

$$\mathscr{P}(F_{\mathbf{h}_1},F_{\mathbf{h}_2},\ldots,F_{\mathbf{h}_M}) = \prod_{i=1}^{M} P(F_{\mathbf{h}_i}). \qquad (1)$$

With the assumption that no atom occupies a special position and that no reflection forms an invariant with its symmetry-related reflections* (excluding those related by an inversion), *i.e.* $\sum_{g=1}^{G/L\varepsilon} n_g \mathbf{h}_g = \mathbf{0}$ and $\sum_{g=1}^{G/L\varepsilon} n_g = 0$ implies $n_g = 0\,\forall g$, then $P(F_{\mathbf{h}})$ is equal to

$$\int_0^\infty \prod_{\mu=1}^{N/G} (G/L\varepsilon_{\mathbf{h}}) J_0[L\varepsilon_{\mathbf{h}} f_\mu(\mathbf{h})\rho] J_0(F_{\mathbf{h}}\rho)\rho\,\mathrm{d}\rho \qquad (2a)$$

for $F_{\mathbf{h}}$ an acentric reflection and

$$(1/2\pi)\int_{-\infty}^{\infty} \prod_{\mu=1}^{N/G} (G/L\varepsilon_{\mathbf{h}}) J_0[L\varepsilon_{\mathbf{h}} f_\mu(\mathbf{h})x]\exp(iF_{\mathbf{h}}x)\,\mathrm{d}x \qquad (2b)$$

for $F_{\mathbf{h}}$ a centric reflection. The symbol G is the order of the space group, N is the number of atoms in the unit cell, $L = 1$ for an acentric reflection, $L = 2$ for a centric reflection and $\varepsilon_{\mathbf{h}}$ is the statistical weight equal to the cardinality of the reflection's isotropy subgroup (see Bricogne, 1991, Appendices *A*2 and *A*3 and references therein).

Although work has been carried out to ensure that the original j.p.d. series [the left-hand side of (1)] can be summed to convergence (Peschar & Schenk, 1986), little consideration has been given to the evaluation of the integrals that form the higher-order coefficients of the series. It has been shown, however (Castleden, 1988, ch. 7), that the results of the evaluation of the zeroth-order coefficients in (1) (described below) can also be used to calculate the higher-order terms.

In their present form, the infinite integrals in (2) can only be numerically integrated for each value of $F_{\mathbf{h}}$. Appendix III shows how they can be re-expressed as a series of special functions permitting a term-by-term evaluation. Expressed as intensity probability distributions, the series are

$$P_{\mathrm{AC}}(I)\mathrm{d}I = (\mathrm{d}I/\varepsilon\sigma_2)\exp(-I/\varepsilon\sigma_2)\sum_{n=0}^{\infty} K_n^{\{1\}}\frac{\Gamma(n+1)}{\Gamma(1)} \times (1/\varepsilon\sigma_2)^n M(-n,1;I/\varepsilon\sigma_2) \qquad (3a)$$

for an acentric reflection and

$$P_C(I)\mathrm{d}I = \frac{\mathrm{d}I}{(2\pi\varepsilon\sigma_2 I)^{1/2}}\exp(-I/2\varepsilon\sigma_2)\sum_{n=0}^{\infty} K_n^{\{\bar{1}\}} \times \frac{\Gamma(n+1/2)}{\Gamma(1/2)}(2/\varepsilon\sigma_2)^n M(-n,1/2;I/2\varepsilon\sigma_2) \qquad (3b)$$

for a centric reflection. The $M(a,b;z)$ are confluent hypergeometric functions and are equal to the more familiar Hermite (for the centric case) and Laguerre (for the acentric case) polynomials due to the relations 13.6.17 and 13.6.9 of Abramowitz & Stegun (1972). These series (and any subseries) are automatically normalized for any values of $K_n\{n>0\}$ because of integral identities 7.621.4 of Gradshteyn & Ryzhik (1965) and 15.7.3 of Abramowitz & Stegun (1972). The first 14 coefficients $K_n^{\{1,\bar{1}\}}$ are given in Table 1. The method of evaluation differs from the standard moment expansion method (Shmueli, 1982) and is explained in Appendix III. The sum to ten terms neglects terms of order N^{-11} or smaller. The *general* seventh-order term, which has been derived by Castleden (1988), permits a series development that neglects terms of order N^{-14} or smaller. From the first seven terms, it is easy to deduce the form for all higher orders.

The benefit of this convergent-series representation is that the atomic contributions enter only in the form of sums of powers of their scattering factors. Moreover, σ_{2n}/σ_2^n is independent of variations in overall scale and thermal-displacement parameters and, therefore, so are the coefficients K_n/σ_2^n. Thus, the overall scale and thermal-displacement parameters appear only within the ratio I/σ_2, a fact not evident from the Fourier–Bessel-series expansion.

The series (3) are convergent and, unlike asymptotic distributions, may require large numbers of terms to represent the integrals accurately. The exact distributions will deviate strongly from the Wilson distributions when the cell contents are highly heterogeneous. As a worst case, Fig. 1 shows the series expansion for a structure with cell content of $C_{29}U$ in the space group $P\bar{1}$ (*cf.* Shumeli, Weiss, Keifer & Wilson, 1984, Fig. 1*b*). There is a slow convergence at the origin where the distribution even fails to remain positive. For such a structure, more than 20 terms seem to be required.

The simplicity hoped for in expressing all heterogeneity in terms of σ_n has been realized. However it is confounded by the increasing complexity of the coefficients K_n as n increases. Slow convergence and unwieldy terms reduce the distributions' acceptability for iterative crystallographic computations. For this reason, another method that is less accurate but much more robust was investigated.

* See Appendix I for a discussion.

Table 1. *Coefficients K_n for the series expansions* (3) *of the intensity probability distributions*

The σ'_n are equal to $(L\varepsilon)^{n-1}\sum_{\mu=1}^{N} f_\mu^n$. For noncentric reflections $L = 1$ whereas for centric reflections $L = 2$.

$$K_0 = 1$$

$$K_1 = 0$$

$$K_2 = \tfrac{-1}{4}\sigma'_4$$

$$K_3 = \tfrac{-1}{9}\sigma'_6$$

$$K_4 = \tfrac{1}{32}\left(\sigma'^2_4 - \tfrac{11}{6}\sigma'_8\right)$$

$$K_5 = \tfrac{1}{36}\left(\sigma'_6\sigma'_4 - \tfrac{57}{50}\sigma'_{10}\right)$$

$$K_6 = \tfrac{-1}{384}\left(\sigma'^3_4 - \tfrac{11}{2}\sigma'_4\sigma'_8 + \tfrac{626}{135}\sigma'_{12}\right)$$

$$K_7 = \tfrac{-1}{288}\left(\sigma'_6\sigma'^2_4 - \tfrac{57}{25}\sigma'_4\sigma'_{10} - \tfrac{11}{6}\sigma'_8\sigma'_6 + \tfrac{458}{147}\sigma'_{14}\right)$$

$$K_8 = \tfrac{1}{6144}\left(\sigma'^4_4 - 11\sigma'^2_4\sigma'_8 + \tfrac{121}{12}\sigma'^2_8 + \tfrac{2504}{135}\sigma'_4\sigma'_{12} - \tfrac{234461}{12600}\sigma'_{16}\right)$$

$$K_9 = \tfrac{1}{3456}\left(\sigma'^3_4\sigma'_6 - \tfrac{11}{2}\sigma'_4\sigma'_6\sigma'_8 - \tfrac{171}{50}\sigma'^2_4\sigma'_{10} + \tfrac{627}{100}\sigma'_8\sigma'_{10} + \tfrac{626}{135}\sigma'_6\sigma'_{12} + \tfrac{458}{49}\sigma'_4\sigma'_{14} - \tfrac{839003}{68040}\sigma'_{18}\right)$$

$$K_{10} = \tfrac{-1}{122880}\left(\tfrac{1535}{1536}\sigma'^5_4 - \tfrac{84425}{4608}\sigma'^3_4\sigma'^2_8 + \tfrac{928675}{18432}\sigma'_4\sigma'^2_8 + \tfrac{480455}{10368}\sigma'^2_4\sigma'_{12} - \tfrac{5285005}{62208}\sigma'_8\sigma'_{12} - \tfrac{71979527}{774144}\sigma'_4\sigma'_{16} + \tfrac{45046342399}{457228800}\sigma'_{20}\right)$$

$$K_{11} = \tfrac{-1}{55296}\left(\sigma'^4_4\sigma'_6 - 11\sigma'^2_4\sigma'_6\sigma'^2_8 + \tfrac{121}{12}\sigma'_6\sigma'^2_8 - \tfrac{114}{25}\sigma'^3_4\sigma'_{10} + \tfrac{627}{25}\sigma'_4\sigma'^2_8\sigma'_{10} + \tfrac{2504}{135}\sigma'_4\sigma'_6\sigma'_{12} - \tfrac{23788}{1125}\sigma'_{10}\sigma'_{12} + \tfrac{916}{49}\sigma'_4\sigma'_4\sigma'_{14} - \tfrac{5038}{147}\sigma'_8\sigma'_{14}\right.$$
$$\left. - \tfrac{234461}{12600}\sigma'_6\sigma'_{16} - \tfrac{839003}{17010}\sigma'_4\sigma'_{18} + \tfrac{187248779}{2858625}\sigma'_{22}\right)$$

$$K_{12} = \tfrac{1}{2949120}\left(\sigma'^6_4 - \tfrac{55}{2}\sigma'^4_4\sigma'^2_8 + \tfrac{605}{4}\sigma'^2_4\sigma'^2_8 - \tfrac{6655}{72}\sigma'^3_8 + \tfrac{2504}{27}\sigma'^3_4\sigma'_{12} - \tfrac{13772}{27}\sigma'_4\sigma'_8\sigma'_{12} + \tfrac{783752}{3645}\sigma'^2_{12} - \tfrac{234461}{840}\sigma'^2_4\sigma'_{16} + \tfrac{2579071}{5040}\sigma'_8\sigma'_{16}\right.$$
$$\left. + \tfrac{293461514}{496125}\sigma'_4\sigma'_{20} - \tfrac{462637929679}{70727580}\sigma'_{24}\right)$$

$$K_{13} = \tfrac{1}{1105920}\left(\sigma'^5_4\sigma'_6 - \tfrac{55}{3}\sigma'^3_4\sigma'_6\sigma'^2_8 + \tfrac{605}{12}\sigma'_4\sigma'_6\sigma'^2_8 - \tfrac{57}{10}\sigma'^4_4\sigma'_{10} + \tfrac{627}{10}\sigma'^2_4\sigma'^2_8\sigma'_{10} - \tfrac{2299}{40}\sigma'^2_8\sigma'_{10} + \tfrac{1252}{27}\sigma'^2_4\sigma'_6\sigma'_{12} - \tfrac{6886}{81}\sigma'_6\sigma'_8\sigma'_{12} - \tfrac{23788}{225}\sigma'_4\sigma'_{10}\sigma'_{12}\right.$$
$$+ \tfrac{4580}{147}\sigma'^3_4\sigma'_{14} - \tfrac{25190}{147}\sigma'_4\sigma'_8\sigma'_{14} + \tfrac{573416}{3969}\sigma'_{12}\sigma'_{14} - \tfrac{234461}{2520}\sigma'_4\sigma'_6\sigma'_{16} + \tfrac{4454759}{42000}\sigma'_{10}\sigma'_{16} - \tfrac{839003}{6804}\sigma'^2_4\sigma'_{18} + \tfrac{9229033}{40824}\sigma'_8\sigma'_{18} + \tfrac{146730757}{1488375}\sigma'_6\sigma'_{20}$$
$$\left. + \tfrac{187248779}{571725}\sigma'_4\sigma'_{22} - \tfrac{4807697606533}{110675556500}\sigma'_{26}\right)$$

$$K_{14} = \tfrac{-1}{82575360}\left(\sigma'^7_4 - \tfrac{77}{2}\sigma'^5_4\sigma'^2_8 + \tfrac{4235}{12}\sigma'^3_4\sigma'^2_8 - \tfrac{46585}{72}\sigma'_4\sigma'^3_8 + \tfrac{4382}{27}\sigma'^4_4\sigma'_{12} - \tfrac{48202}{27}\sigma'^2_4\sigma'_8\sigma'_{12} + \tfrac{265111}{162}\sigma'^2_8\sigma'_{12} + \tfrac{5486264}{3645}\sigma'_4\sigma'^2_{12} - \tfrac{234461}{360}\sigma'^3_4\sigma'_{16}\right.$$
$$\left. + \tfrac{2579071}{720}\sigma'_4\sigma'_8\sigma'_{16} - \tfrac{73386293}{24300}\sigma'_{12}\sigma'_{16} + \tfrac{146730757}{70875}\sigma'^2_4\sigma'_{20} - \tfrac{1614038327}{425250}\sigma'_8\sigma'_{20} - \tfrac{46263729679}{10103940}\sigma'_4\sigma'_{24} + \tfrac{153578765444633}{29499294825}\sigma'_{28}\right)$$

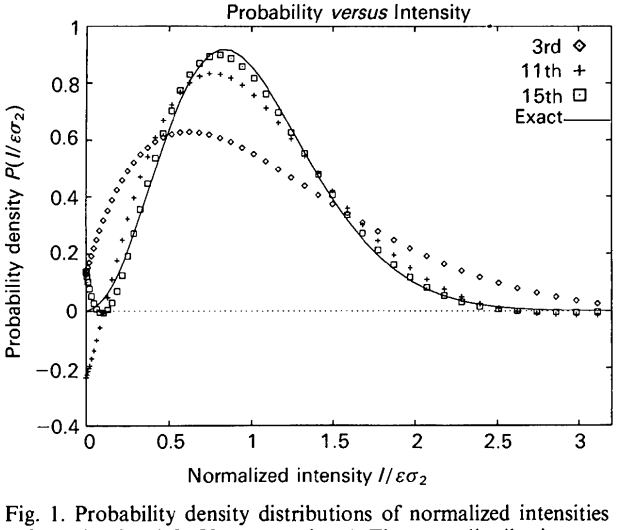

Fig. 1. Probability density distributions of normalized intensities for a simulated $C_{29}U$ structure in $P1$. The exact distribution was calculated using the Fourier–Bessel series method of Shmueli, Weiss, Keifer & Wilson (1984) (see Appendix III) and is shown in the diagram by the solid line. The distributions calculated using (2) with the series terminating at the K_3, K_{11} and K_{15} terms are also plotted.

Maximum-entropy representation

The exact distributions calculated by Shmueli & Weiss (1987) and Shmueli, Weiss, Keifer & Wilson (1984) and the series representation described above show that the mode of the intensity distribution can shift away from zero (see Fig. 1). The simplest way to mimic this shift while keeping the probability everywhere positive is to make the log of the probability quadratic in the intensities. This can be done *via* the principle of maximum entropy (PME) (Levine, 1980), assuming that the first two even moments are known. For a non-centric distribution, these are the averages (neglecting dispersion and assuming that all atoms are randomly distributed) $\langle I\rangle = \varepsilon\sigma_2$ and $\langle I^2\rangle = 2\varepsilon^2\sigma_2^2 - \varepsilon^3 Q_h^1\sigma_4$ (Wilson, 1978). The Q_h^1 are space-group-dependent numbers that have been tabulated for all general and most symmetry-specific reflections by Wilson (1978) and Shmueli & Kaldor (1981).* The problem of extending

* The notation used here is slightly different from that defined by these authors. This is to ensure that $Q_n = 1$ for space groups $P1$ and $P\bar{1}$ (see Appendix II).

the analysis to symmetry-specific reflections is considered in Appendix I. The maximum-entropy distribution for an acentric intensity under these constraints is

$$P(I)\mathrm{d}I = [\mathrm{d}I/Z(\alpha,\beta)]\exp(-\alpha I - \beta I^2), \quad (4a)$$

where

$$Z(\alpha,\beta) = (\pi/4\beta)^{1/2}\exp(+\alpha^2/4\beta)\{1 - \Phi[\alpha/(4\beta)^{1/2}]\} \quad (4b)$$

and where α, β satisfy $-\partial\ln Z/\partial\alpha = \varepsilon\sigma_2$ and $-\partial\ln Z/\partial\beta = 2\varepsilon^2\sigma_2^2 - \varepsilon^3 Q_h^1\sigma_4$. It would be convenient to express α and β directly in terms of the constraining intensity averages but, for general values of σ, this does not appear to be possible. However, with the definition $\gamma = \varepsilon Q_h^1\sigma_4/\sigma_2^2$, $\alpha \simeq (1-\gamma)/\varepsilon\sigma_2$ and $\beta \simeq \gamma(1-\gamma)^2/4(\varepsilon\sigma_2)^2$. These approximations are better the smaller γ becomes (see Table 2). For a structure consisting of equal randomly distributed atoms, γ is approximately a $1/N$ correction term. The maximum-entropy distribution is automatically greater than zero for all intensity values and is also normalized. This distribution bears some resemblance to that obtained by Kronenburg, Peschar & Schenk (1991) [equations (3.1) and (3.4)] for a single structure factor in $P1$. It should be noted, however, that in that case the maximum-entropy constraint used is the maximum of the distribution, while in our case it is the expectation value of the first two even moments.

For a centric reflection, the distribution for $F_{\mathbf{h}}$ was derived first and then converted to an intensity distribution. The expectation values $\langle A^2\rangle = \varepsilon\sigma_2$, $\langle A^4\rangle = 3\varepsilon^2\sigma_2^2 - 3\varepsilon^3 Q_h^1\sigma_4$ yield the distribution

$$P(I)\mathrm{d}I = [\mathrm{d}I/I^{1/2}Z(\alpha,\beta)]\exp(-\alpha I - \beta I^2), \quad (5a)$$

where

$$Z(\alpha,\beta) = (2\beta)^{-1/4}\Gamma(1/2) \times \exp(+\alpha^2/8\beta)D_{-1/2}[\alpha/(2\beta)^{1/2}]. \quad (5b)$$

Defining the centric equivalent of γ as $\gamma = \varepsilon Q_h^1\sigma_4/\sigma_2^2$, one obtains the approximations $\alpha \simeq (1 - 3\gamma/2)/2\varepsilon\sigma_2$ and $\beta \simeq 1/2\gamma(1-3\gamma/2)^2/(2\varepsilon\sigma_2)^2$ for γ small. Notice, again, how the overall thermal-displacement and scale parameters appear only in the ratio I/σ_2 because γ is independent of these values.

Attempts to include higher-order moments of the intensity distribution were hampered by the fact that the integrations required for the evaluation of Z could only be expressed as series of special functions. Numerical integration techniques were not tried.

The principle of maximum entropy does not assume an independent random atom distribution and is therefore applicable, at least theoretically, when this assumption does not hold. The simplest case is that of a known but unoriented molecular fragment for which the Debye distribution can be calculated. For example, the expectation values of an acentric reflection intensity, assuming that the interatomic distances are known, are given by

$$\langle I\rangle = \varepsilon\sigma_2 + G(s)$$

and (6)

$$\langle I^2\rangle = 2[\varepsilon\sigma_2 + G(s)]^2 - \varepsilon^3\sigma_4 - 2G(s)^2 + 4G'(s) + K(s),$$

where $s = |\mathbf{h}|$ and the G, G' and K functions are of the form

$$G(s) = 2\varepsilon^2 \sum_{\mu<\nu}^{N/\varepsilon} f_\mu f_\nu \mathrm{sinc}(2\pi s|\mathbf{r}_\mu - \mathbf{r}_\nu|)$$
$$G'(s) = 2\varepsilon^2 \sum_{\mu<\nu}^{N/\varepsilon} f_\mu^3 f_\nu \mathrm{sinc}(2\pi s|\mathbf{r}_\mu - \mathbf{r}_\nu|) \quad (7)$$
$$K(s) = 2\varepsilon^4 \sum_{\mu<\nu}^{N/\varepsilon}\sum_{\substack{\omega\neq\delta\\ \delta\neq\nu\\ \omega\neq\nu}}^{N/\varepsilon} f_\mu f_\nu f_\omega f_\delta \times \mathrm{sinc}(2\pi s|\mathbf{r}_\mu - \mathbf{r}_\nu + \mathbf{r}_\omega - \mathbf{r}_\delta|).$$

It can be seen that, for this case, $\gamma \simeq (\sigma_4 + 2G^2 - K)/(\sigma_2 + G)^2$ is not necessarily small because G and K are of order 1, making $\gamma \simeq 2(1 - K/G^2)$. Hence, significant deviation from the random atom distribution (*i.e.* the Gaussian type distribution) is not excluded *a*

Table 2. *Maximum-entropy values of* $\alpha' = \alpha\sigma_2$ *and* $\beta' = \beta\sigma_2^2$ *for various* γ *values applicable to expressions* (4) *and* (5)

The α' and β' numbers given are those that minimize the function $-\ln Z(\alpha',\beta') - \alpha' - \beta'(2-\gamma)$ for a noncentric reflection or the function $-\ln Z(\alpha',\beta') - \alpha' - \beta'(3-3\gamma)$ for a centric reflection. Thus, a given unit-cell content determines the γ value which, in turn, determines the applicable α' and β'. All numbers are accurate to six significant figures. Note that the ratio $\alpha'/\beta' \to -2$ as $\gamma \to 1$. For the acentric case, this implies that $\exp(-\alpha I - \beta I^2) \to A^{-1}\exp[-\beta'(I-\sigma_2)^2/\sigma_2^2]$ (A is a normalizing constant), which is a Gaussian distribution with mean σ_2 and standard deviation $\sigma_2/(2\beta')^{1/2}$.

	Acentric			Centric	
γ	$\alpha\sigma_2$	$\beta\sigma_2^2$	γ	$\alpha\sigma_2$	$\beta\sigma_2^2$
0.20	1.0	0.0	0/30	0.5	0.0
1/20	0.944848	0.0141415	1/30	0.475058	0.00455804
2/20	0.878609	0.0319450	2/30	0.437515	0.0111580
3/20	0.799761	0.0541187	3/30	0.395585	0.0193360
4/20	0.706317	0.0815787	4/30	0.344923	0.0298224
5/20	0.595751	0.115500	5/30	0.283873	0.0432254
6/20	0.464859	0.157394	6/30	0.210308	0.0603524
7/20	0.309547	0.209228	7/30	0.121492	0.0822842
8/20	0.124513	0.273590	8/30	0.0138571	0.110487
9/20	−0.0972171	0.353941	9/30	−0.117320	0.146981
10/20	−0.365031	0.455010	10/30	−0.278410	0.194603
11/20	−0.691903	0.583415	11/30	−0.478246	0.257434
12/20	−1.09638	0.748709	12/30	−0.729168	0.341436
13/20	−1.60618	0.965251	13/30	−1.04978	0.455821
14/20	−2.26507	1.25580	14/30	−1.46836	0.615113
15/20	−3.14781	1.65913	15/30	−2.03195	0.843983
16/20	−4.39674	2.24864	16/30	−2.82535	1.18762
17/20	−6.33177	3.18772	17/30	−4.02570	1.74065
18/20	−9.90171	4.95532	18/30	−6.10892	2.75371
19/20	−19.9983	9.99919	19/30	−11.3252	5.37509

priori even for large molecular structures. Debye averages have been shown to be unexpectedly *unhelpful* in estimating scale and temperature factor (Hall & Subramanian, 1982). However, (6) and (7) show that this may be because the Debye estimates and the Wilson distribution are incompatible. Work is in progress on the extent of this effect.

For arbitrary γ, the values of α and β must be evaluated by minimization of the function $-\ln Z(\alpha,\beta) - \alpha\langle I\rangle - \beta\langle I^2\rangle$. This function is guaranteed to have a global minimum because its Hessian is everywhere a Gram matrix. Table 2 shows the results of this optimization. A double-precision Fortran program was written employing a conjugate gradient-minimization routine. Convergence problems were encountered because the above functions are not defined for $\beta<0$, however, the final values were checked for accuracy against equivalent functions expressed in the arbitrary precision arithmetic of *Mathematica* (Wolfram, 1987). Other tests have shown that the variation of α and β is very smooth and so a simple interpolation scheme can be used for varying values of γ. Figs. 2 and 3 show the probability distributions parameterized by γ. For clarity, sections through the distribution are depicted in Figs. 2(*b*) and (*c*) and 3(*b*) and (*c*). As expected, in the limit of a single dominant scatterer, $\gamma\rightarrow 1$ and the probability distribution approaches a delta function around $I/\varepsilon\sigma_2 = 1$.

Fig. 4 shows a comparison between the maximum-entropy and exact distributions for the $C_{29}U$ structure used in Fig. 1. For only a small computational cost, the approximation follows the true contour quite well.

Concluding remarks

Two different approaches to the accurate evaluation of intensity distributions have been described. Many other researchers have already presented work in this area. Unlike these previous efforts, however, the emphasis here has been primarily on representations that are easily machine computable while still retaining fidelity to the actual distributions. The convergent-series representations are quickly convergent except for highly heterogeneous cell contents. The maximum-entropy distribution offers a promising compromise between accuracy and robustness. In both the series and maximum-entropy representations, it is found that overall scale and thermal-displacement parameters appear only in the ratio I/σ_2. Thus, normalized structure factors $E = (I/\varepsilon\sigma_2)^{1/2}$ appear naturally from both the exact and approximate forms of the probability distribution. They are not an *ad hoc* definition resulting solely from the Taylor expansion of the distribution about the origin.

Financial assistance from the Natural Sciences and Engineering Research Council of Canada, in the form of an operating grant to SF, is gratefully acknowledged.

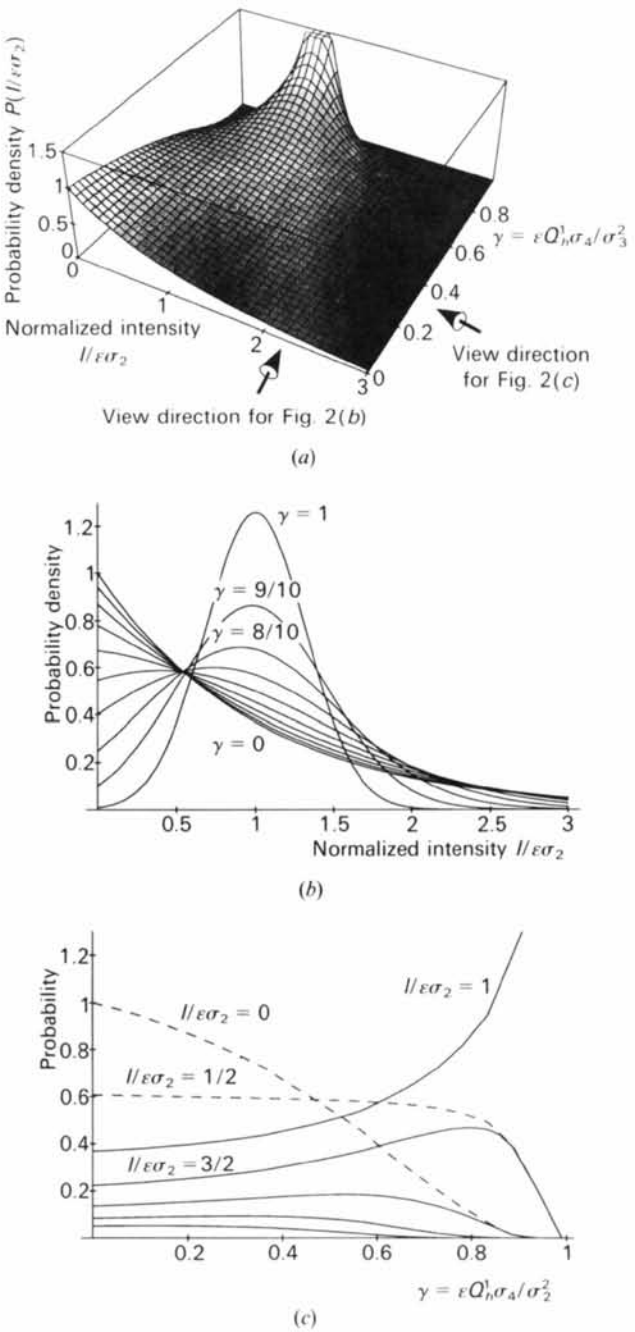

Fig. 2. (*a*) Acentric reflection probability distribution plotted with respect to the normalized intensity ($I/\varepsilon\sigma_2$) and $\gamma = \varepsilon Q_h^1\sigma_4/\sigma_2^2$. Plan views in the indicated directions are shown in (*b*) and (*c*). (*b*) Plots of the probability density *versus* $I/\varepsilon\sigma_2$ for values of γ equal to 0, 1/10, ..., 9/10. (*c*) Plots of the probability density *versus* γ for $I/\varepsilon\sigma_2$ equal to 0, 1/2, ..., 3. [Note that this is not a probability distribution $P(\gamma)$ for γ.]

APPENDIX I

The derivation of the maximum-entropy distributions (4) and (5) assumed that the moments $\langle I_{\mathbf{h}} \rangle \equiv \langle \mathbf{F}_{\mathbf{h}} \mathbf{F}_{-\mathbf{h}} \rangle$ and $\langle I_{\mathbf{h}}^2 \rangle \equiv \langle \mathbf{F}_{\mathbf{h}} \mathbf{F}_{-\mathbf{h}} \mathbf{F}_{\mathbf{h}} \mathbf{F}_{-\mathbf{h}} \rangle$ are known (dispersion is ignored so that Friedel pairs are related by complex conjugation). Other expectation values, for example, $\langle \mathbf{F}_{\mathbf{h}} \rangle$, $\langle \mathbf{F}_{\mathbf{h}} \mathbf{F}_{\mathbf{h}} \rangle$, $\langle \mathbf{F}_{\mathbf{h}} \mathbf{F}_{\pm_{\mathbf{h}} \mathbf{h}} \rangle$ and $\langle \mathbf{F}_{\mathbf{h}} \mathbf{F}_{\pm_{\mathbf{h}} \mathbf{h}} \mathbf{F}_{\pm_{\mathbf{h}} \mathbf{h}} \rangle$, have been ignored because under the assumption of a random uniform distribution of atoms they usually vanish and, thus, lead to Lagrange multipliers that also vanish. It is worth noting, though, that, owing to symmetry, products such as $\langle \mathbf{F}_{\mathbf{h}} \mathbf{F}_{\mathbf{h}} \mathbf{F}_{\mathbf{h}} \rangle$ may occasionally retain nonzero expectation values (Foster & Hargreaves, 1963). Let $\mathbf{h}_i$, $i = 1, \ldots, G/\varepsilon$ be the set of distinct acentric reflections that are symmetry related to $\mathbf{h}$, *i.e.* $\mathbf{h}_i = \mathbf{h}^T \cdot \mathbf{R}_i$, so that $\mathbf{h}_i \neq -\mathbf{h}_j \; \forall i, j$. The reflection $\mathbf{h}$ splits the space group into G/ε cosets all of the same cardinality ε (Wondratschek, 1983; Bricogne, 1991, Appendices *A*2 and *A*3). For each coset $\{(\mathbf{R}_i, \mathbf{t}_i) | i = 1, 2, \ldots, p\}$, the number $\mathbf{h} \cdot \mathbf{t}_i$ is the same $\forall i = 1, 2, \ldots, \varepsilon$. The third-order expectation value $\langle \mathbf{F}_{\mathbf{h}} \mathbf{F}_{\pm_{\mathbf{h}} \mathbf{F}_{\pm_{\mathbf{h}} \mathbf{h}}} \rangle$ will, in general, vanish unless $\pm \mathbf{h}_i + \mathbf{h}_j + \mathbf{h}_k = \mathbf{0}$ for some i, j, k; *i.e.* a reflection forms a triplet with its symmetry-related reflections. Because the reflections are related by a rotation they have the same magnitude. For three of them to form a triplet requires that they lie in the same plane and be related to one another by 60 or 120° rotations [*i.e.* form an equilateral triangle (see Shmueli & Kaldor, 1983, Table 1 and Appendix *A*)]. Such a condition can be satisfied, for example, in the space group $P3$, c unique. All $\mathbf{h} = hk0$ reflections form a triplet invariant with their symmetry-related reflections $-h-k,h,0$, $-k,-h,-k0$, which implies that $3\varphi_{\mathbf{h}}$ is a seminvariant. However, the origin used to fix the form of the Seitz matrices has not been made explicit so that in fact the expectation value $\langle \mathbf{F}_{\mathbf{h}} \mathbf{F}_{\mathbf{h}} \mathbf{F}_{\mathbf{h}} \rangle = 6\sigma_3 \langle \exp(6\pi i \mathbf{h} \cdot \mathbf{r}_0) \rangle$ is the expectation

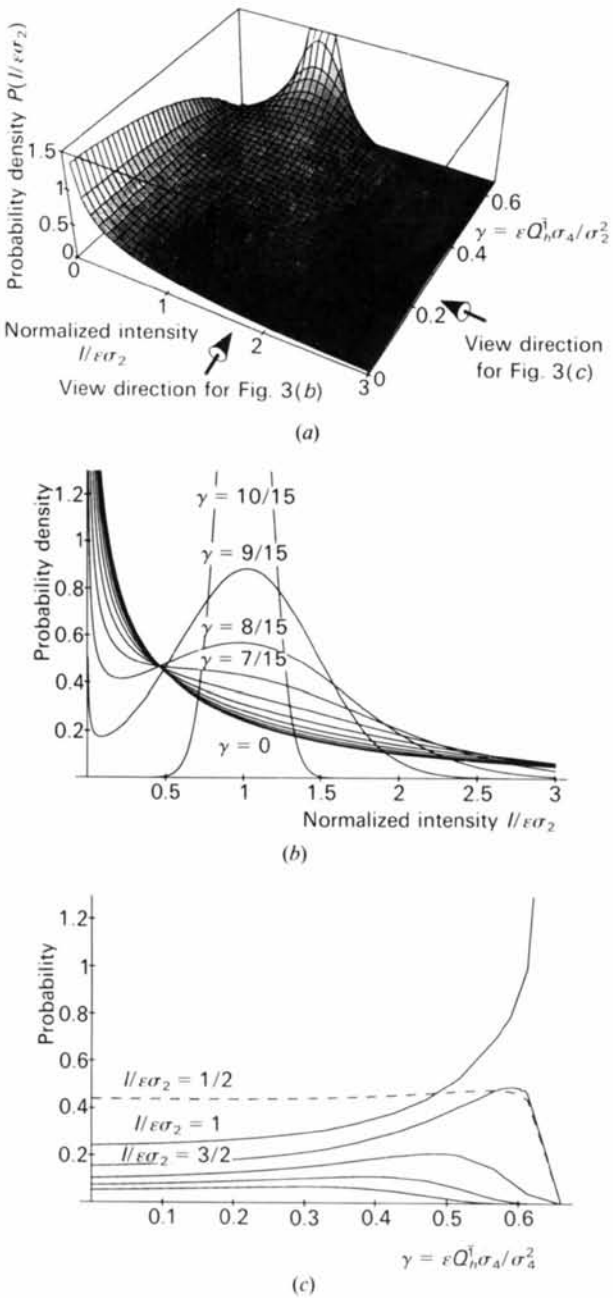

Fig. 3. (*a*) Centric reflection probability distribution plotted with respect to the normalized intensity ($I/\varepsilon\sigma_2$) and $\gamma = \varepsilon Q_h^{-1}\sigma_4/\sigma_2^2$. Plan views in the indicated directions are shown in (*b*) and (*c*). (*b*) Plots of the probability density *versus* $I/\varepsilon\sigma_2$ for values of γ equal to 0, 1/15, ..., 2/3. (*c*) Plots of the probability density *versus* γ for I/σ_2 equal to 0, 1/2, ..., 3. [Note that this is not a probability distribution $P(\gamma)$ for γ.]

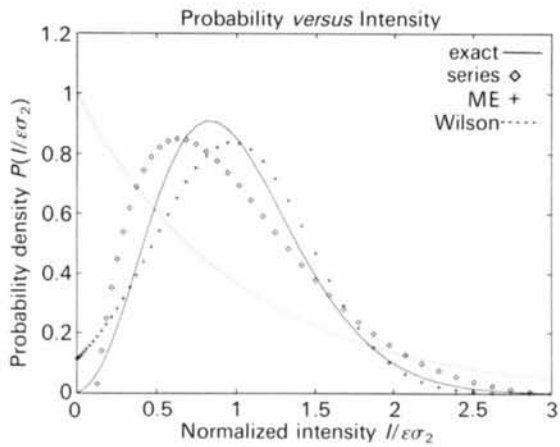

Fig. 4. Probability density distributions of normalized intensities for the $C_{29}U$ structure in $P1$. The exact distribution, as calculated by the Fourier–Bessel series (see Appendix III), is shown by the solid line. The maximum-entropy distribution is shown by the crosses. The Wilson distribution is represented by the dashed line. The series distribution is represented by the diamond and has been truncated at the 15th term.

value being evaluated, where, by convention, $\mathbf{r}_0 = \mathbf{0}$ is on the threefold axis. Two different cases can arise depending on whether the origin $\mathbf{r}_0$ is taken to be a random variable or not. In this work, the origin is taken as being uniformly distributed and therefore all odd phase-dependent terms must vanish. Although this averaging has no effect in space groups such as $P1$ and $P\bar{1}$, a careful space group by space group analysis needs to be undertaken for a fixed origin.

APPENDIX II

The definitions for the moments in Wilson (1978) vary slightly from those used here. Labelling the definitions of Wilson by the superscript W, one has the equivalences: $p^W = \varepsilon G$ and $Q^W = \varepsilon^3 GQ$ [*cf.* Wilson, 1978, equation (6), p. 988] where G is the order of the space group. For example, the Wilson fourth-order moment q^W is given here by

$$q^W = \left\langle \varepsilon^4 \sum_{i,j,k,l}^{G/\varepsilon} \exp\left[2\pi i(\mathbf{h}_i - \mathbf{h}_j + \mathbf{h}_k - \mathbf{h}_l)\cdot\mathbf{r} + 2\pi i\mathbf{h}\cdot(\mathbf{t}_i - \mathbf{t}_j + \mathbf{t}_k - \mathbf{t}_l)\right]\right\rangle. \quad (A1)$$

The expectation value is zero unless the coefficient of $\mathbf{r}$ is zero. There are three special cases for which the summation can be reduced: (1) $i = j$ and $k = l$, (2) $i = l$ and $j = k$ and (3) $i = j = k = l$. Conditions (1) and (2) give $(G/\varepsilon)^2$ and condition (3) sums to (G/ε), however, the first two include the third condition as a subcondition and so $2(G/\varepsilon)$ must be subtracted. Thus, $Q^W = 2(p^W)^2 - q^W$, which gives $2(G\varepsilon)^2 - p^4[2(G/\varepsilon)^2 + (G/\varepsilon) - 2(G/\varepsilon) + q]$ or, rearranging, $Q^W = G\varepsilon^3[1 - q/(G/\varepsilon)]$. If the condition $\mathbf{h}_i - \mathbf{h}_j + \mathbf{h}_k - \mathbf{h}_l \neq \mathbf{0}\ \forall i, j, k, l, i \neq j$ or $k \neq l$, holds, the fourth-order moment can be calculated as $Q = 1$ with $q = 0$, otherwise

$$q = 2 \sum_{i<j,\, k\neq j,\, l\neq i,\, l\neq k}^{G/L\varepsilon} \delta(\mathbf{h}_i - \mathbf{h}_j + \mathbf{h}_k - \mathbf{h}_l) \times \cos[2\pi(\mathbf{t}_i - \mathbf{t}_j + \mathbf{t}_k - \mathbf{t}_l)].$$

The centric case is similar, except that the sum is over the $G/2\varepsilon$ reflections not related by an inversion. Numerical values for q^W are given in Table 1 of Shmueli & Kaldor (1981).

APPENDIX III

The derivation of the series (3) from the integrals (2) is based on a re-expression of the Bessel functions which, in the limit $F_h \to 0$, have a substantial contribution only near the origin. At this point, they are approximately exponentially decreasing, *i.e.* $J_o(x) \to \exp\{-x^2/4\}$, an approximation that permits the analytical integration of (2) and yields the Wilson distributions. To ensure that, in the limit of small F_h, the exact distributions converge to the Wilson distributions, the Bessel functions in the integrals (2) are first expressed as a product of the exponential term and a series F of 'corrections'

$$J_o(L\varepsilon fx) = \exp[-(L\varepsilon fx)^2/4]F(L\varepsilon fx) \quad \text{with} \quad (A2)$$

$$F(L\varepsilon fx) = \sum_{k=0}^{\infty} G_k(L\varepsilon fx/2)^{2k}/k!.$$

The series coefficients, G_k, can be determined by moving the exponential term to the left-hand side and equating powers of x. This yields the rational values $G_k = M(-k,1;1)$, where M is a confluent hypergeometric function that varies as $k^{-1/4}$ for $k \to \infty$. Equation (*A*2) is exact in that, order for order, the left- and right-hand sides are the same. The product over the Bessel functions occurring in (2) can be written as

$$\prod_{\mu=1}^{N/L\varepsilon} J_o(L\varepsilon f_\mu x) = \exp(-L\varepsilon\sigma_2 x^2/4) \sum_{n=0}^{\infty} K_n(x/2)^{2n}, \quad (A3)$$

where $L = 1$ for a noncentric reflection and $L = 2$ for a centric one. The coefficient K_n is, in turn, a sum

$$K_n = \sum_{\{\Gamma\}=n} K_\Gamma. \quad (A4)$$

The set $\{\Gamma\} = n$ is the set of unique (unordered) sets of positive integers $\gamma_1,\ldots,\gamma_M$ such that $\gamma_1 + \ldots + \gamma_M = n$ [of which there are $p(n)$ (Abramowitz & Stegun, 1972, ch. 24)] and

$$K_\Gamma = G_o^{\max(N-M,0)} \prod_{i=1}^{M} \frac{G_{\gamma_i}}{\gamma_i!} \frac{N_M(\gamma_1,\ldots)}{M!} \times \sum_{i_1=1}^{N/L\varepsilon} \sum_{\substack{i_2=1 \\ i_2\neq i_1}}^{N/L\varepsilon} \cdots \sum_{\substack{i_M=1 \\ i_M\neq i_1 \\ i_M\neq i_2 \\ \vdots \\ i_M\neq i_{M-1}}}^{N/L\varepsilon} L\varepsilon f_{i_1}^{2\gamma_1} L\varepsilon f_{i_2}^{2\gamma_2} \ldots L\varepsilon f_{i_M}^{2\gamma_M}, \quad (A5)$$

where N_M is the number of distinct ways to order the integers $\gamma_1,\ldots,\gamma_M$.

The K_n are independent of x and so, through substitution of the series (*A*3) into (2), the integration can be performed term by term using the identity 6.631.1 of Gradshteyn & Ryzhik (1965). The multiple conditional summations appearing in (*A*5) can be replaced by products of single unconditional summations. For example, the general third-order sum-

Table 3. *Mathematica program for evaluating the sums of products of scattering factors appearing in* (*A*5)

The command simp[e3[a,b,c]] converts the left-hand side of (*A*6) to the right-hand side. One can see that e3[a,b,c] accurately represents the multiple sum (*A*6) and that each conversion left-hand→side right-hand side in this program preserves arithmetic truth. For example: sum[−(e_),l_] := −sum[e_,l_] (line 4) merely takes the negative sign outside the summation. Higher-order terms can be easily expressed in the same manner as e3. All the terms in Table 1 are based on the results of this reduction.

```
(**     define expansions and transformations for sum          **)
sum[e_  ,{i_,l_}] := sum[e,{i}]-Sum[e/. i->l[[j]],{j,Length[l]}];
sum[el_+e2_, l_ ] := sum[el,l] + sum[e2,l];
sum[-(e_)  , l_ ] := -sum[e,l];
sum[c_ e_  ,{i_}] := c sum[e,{i}]            /; FreeQ[c,i];
sum[e_     ,{i_}] := sig[ss[e]]              /; prodQ[e,i];
(** check that the expression is a product f[i,a]f[i,b]f[i,c]..**)
prodQ[f[i_,a_],      i_]  = True;
prodQ[f[i_,a_] expr_,i_] := prodQ[expr,i];
prodQ[expr_,         i_] := False;
ss[f[_,a_]] = a; ss[f[_,a_] expr_] :=a+ss[expr];
(*********              S I M P L I F Y            **********)
simp[f[i_,a_] e_]  := f[i,a]simp[e];
simp[e_]           := simp /@ e                 /;  Head[e]=!=sum;
simp[e_]           := e                         /;  FreeQ[e,sum];
simp[c_ e_]        := c simp[e]                 /;  NumberQ[c];
simp[sum[e_,{i_}]] := simp[sum[simp[e],{i}]]    /; !FreeQ[e,sum];
simp[sum[e_,{i_}]] := simp[sum[Expand[e],{i}]] /;  FreeQ[e,sum];
e3[a_,b_,c_] := sum[sum[sum[f[i3,c],{i3,{i2,il}}]
       f[i2,b],{i2,{   il}}]
       f[il,a],{il        }];
simp[e3[a,b,c]]
```

mation can be rearranged to give

$$\sum_{i=1}^{N}\sum_{\substack{j=1\\ j\neq i}}^{N}\sum_{\substack{k=1\\ k\neq i\\ k\neq j}}^{N} f_i^{\alpha} f_j^{\beta} f_k^{\gamma}$$
$$= \sigma_\alpha\sigma_\beta\sigma_\gamma - \sigma_\beta\sigma_{\alpha+\gamma} - \sigma_\alpha\sigma_{\beta+\gamma} - \sigma_\gamma\sigma_{\alpha+\beta} + 2\sigma_{\alpha+\beta+\gamma}. \qquad (A6)$$

The right-hand side is a function only of the sum of powers of the scattering factors

$$\sigma_\alpha = \sum_{i=1}^{N} f_i^{\alpha},$$

which are simple to evaluate. The routine re-expression of the multiple summations such as (*A*6) that occur in (*A*5) is discussed in detail in Appendix I of Foster & Hargreaves (1963). Shmueli & Wilson (1983) have proposed an algorithm to implement this reduction using *REDUCE* (Hearn, 1973). We have opted for the more expressive functional language of *Mathematica* (Wolfram, 1987) as shown in Table 3. The general *n*th-order form can be inferred from the form of the first seven terms* to be the following: for the group of n integers $\{\alpha_1,\alpha_2,\alpha_3,\alpha_4,\ldots,\alpha_n\}$ determine the set of all different partitions of these integers, *i.e.* all possible sets of sets so that each integer appears in one and only one element (a set) per set, for example, $\{\{\alpha_1, \alpha_2, \alpha_n\}, \{\alpha_3, \alpha_4\},\ldots\}$. A set has no implied order to its elements so that two partitions are different if there is no objective mapping between their elements giving pairwise equality (set equality is defined in the standard manner). There is a total of

$$\sum_{m=1}^{n} S2_n^{(m)} \qquad (A7)$$

of these sets, where $S2$ is a Stirling number of the second kind (Abramowitz & Stegun, 1972, 24.1.4). For each such set of sets, calculate the product

$$-1^n \prod_{\mu=1}^{m} [-1^{n_\mu}(n_\mu - 1)!\sigma_{\gamma_\mu}], \qquad (A8)$$

where n_μ is the number of elements in the set that forms the μth element and γ_μ is its sum. Summing up these values yields the appropriate *n*th-order term. Although this prescription is trivial to code in *Mathematica* (Wolfram, 1987) and is not memory intensive, the number of terms to be considered explodes exponentially.

Another method is to use the band-limited nature of the probability distribution [*i.e.* $P(F_h) \equiv 0$ for

* The *Mathematica* program was written with simplicity in mind and runs out of memory (32 Mbytes!) when applied to the general eighth-order term.

$F_{\mathbf{h}} > \sigma_1$] and the 'orthogonality' of the Bessel functions to produce a Fourier–Bessel series for the noncentric reflections. Thus, the probability can be expressed as

$$P(I_{\mathbf{h}}) = \sum_{m=1}^{\infty} a_m J_0(j_m I_{\mathbf{h}}^{1/2}/\sigma_1), \qquad (A9)$$

where j_m is the mth zero of J_0. The values of a_m can be determined using the orthogonality condition (Watson, 1942; see also Gradshteyn & Ryzhik, 1965, 6.521)

$$(1/2)\sigma_1^2[J_1(j_m)]^2 \int_0^{\sigma_1} J_0(Fx)J_0(Fj_m/\sigma_1)F\,dF = \delta(j_m - x\sigma_1), \qquad (A10)$$

which gives

$$a_m = \sigma_1^{-2}[J_1(j_m)]^{-2} \prod_{\mu=1}^{N/\varepsilon} J_0(\varepsilon f_\mu j_m/\sigma_1) = \sigma_1^{-2}[J_1(j_m)]^{-2}\exp - (\varepsilon\sigma_2/4\sigma_1^2)j_m^2 \times \sum_{n=0}^{\infty} K_n^1(\varepsilon j_m/2\sigma_1)^{2n}. \qquad (A11)$$

This series expansion is used to calculate the exact distributions for comparison in Figs. 1 and 4. At least 40 terms were used in any expansion.

These series have importance when the central-limit-theorem approximation (*i.e.* the Wilson term) is inapplicable. For example, if the cell is considered to be made up of a small number of fragments with structure factors $|\mathbf{F}_{\mathbf{h}}^{\mu}| = f_\mu(\mathbf{h}, \theta_\mu)$, then the expansions (3) still hold for random fragments but the higher-order terms have more significance. Thus, a joint probability distribution of fragment orientations θ_μ can be calculated.

References

ABRAMOWITZ, M. & STEGUN, I. (1972). *Handbook of Mathematical Functions*. New York: Dover.

BRICOGNE, G. (1991). *Acta Cryst.* A**47**, 803–829.

CASTLEDEN, I. R. (1987). *Acta Cryst.* A**43**, 384–393.

CASTLEDEN, I. R. (1988). PhD Thesis, Univ. of Western Australia, Perth, Australia.

FOSTER, F. & HARGREAVES, A. (1963). *Acta Cryst.* **16**, 1124–1133.

FRENCH, S. & OATLEY, S. (1982). *Crystallographic Statistics: Progress and Problems*, edited by S. RAMASESHAN, M. F. RICHARDSON & A. J. C. WILSON, pp. 19–51. Bangalore: Indian Academy of Sciences.

GRADSHTEYN. I. & RYZHIK, I. (1965). *Tables of Integrals Series and Products*. London: Academic Press.

HALL, S. R. & SUBRAMANIAN, V. (1982). *Acta Cryst.* A**38**, 590–598.

HEARN, A. C. (1973). *REDUCE User's Manual.* Univ. of Utah, Salt Lake City, Utah, USA.

KRONENBURG, M. J., PESCHAR, R. & SCHENK, H. (1991). *Acta Cryst.* A**47**, 469–480.

LEVINE, R. D. (1980). *J. Phys. A*, **13**, 91–108.

PESCHAR, R. & SCHENK, H. (1986). *Acta Cryst.* A**42**, 309–317.

SHMUELI, U. (1982). *Crystallographic Statistics: Progress and Problems*, edited by S. RAMASESHAN, M. F. RICHARDSON & A. J. C. WILSON, pp. 53–82. Bangalore: Indian Academy of Sciences.

SHMUELI, U. & KALDOR, U. (1981). *Acta Cryst.* A**37**, 76–80.

SHMUELI, U. & KALDOR, U. (1983). *Acta Cryst.* A**39**, 619–621.

SHMUELI, U. & WEISS, G. H. (1987). *Acta Cryst.* A**43**, 93–98.

SHMUELI, U., WEISS, G. H., KEIFER, J. E. & WILSON, A. J. C. (1984). *Acta Cryst.* A**40**, 651–660.

SHMUELI, U. & WILSON, A. J. C. (1982). *Crystallographic Statistics: Progress and Problems*, edited by S. RAMASESHAN, M. F. RICHARDSON & A. J. C. WILSON, pp. 83–97. Bangalore: Indian Academy of Sciences.

SHMUELI, U. & WILSON, A. J. C. (1983). *Acta Cryst.* A**39**, 225–233.

SUBRAMANIAN, V. & HALL, S. R. (1982). *Acta Cryst.* A**38**, 577–590.

WATSON, G. N. (1942). *A Treatise on the Theory of Bessel Functions*, 2nd ed. Cambridge Univ. Press.

WILSON, A. J. C. (1942). *Nature (London)*, **150**, 151–152.

WILSON, A. J. C. (1978). *Acta Cryst.* A**34**, 986–994.

WOLFRAM, S. (1987). *Mathematica, a System for Doing Mathematics by Computer*. New York: Addison-Wesley.

WONDRATSCHEK, H. (1983). *International Tables for Crystallography*, Vol. A, edited by T. HAHN, pp. 711–731. Dordrecht: Kluwer Academic Publishers.

Acta Cryst. (1994). A**50**, 17–22

Mathematical Structure of the Coherent Wave Field in the Statistical Theory of Dynamical Diffraction

BY NORIO KATO

Department of Physics, Faculty of Science and Technology, Meijo University, Tenpaku-ku, Nagoya, Japan

(*Received* 12 *April* 1993; *accepted* 25 *May* 1993)

Abstract

The coherent wave field, which is the ensemble average of the solution of the wave equation, is studied. The approach is similar to that used in the previous theory on extinction [Kato (1976). *Acta Cryst.* A**32**, 453–457, 458–466]. Here it is extended to deal with general cases where the single average and the