Acta Crystallographica Section A
Foundations of Crystallography
ISSN 0108-7673

Received 18 September 2000
Accepted 1 May 2001

The joint probability distributions of structure-factor doublets in displacive incommensurately modulated structures and their applicability to direct methods

René Peschar,[a] Randy Israël[b] and Paul T. Beurskens[b]*

[a]Laboratory for Crystallography, Institute of Molecular Chemistry, University of Amsterdam, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands, and [b]Crystallography Laboratory, Department of Inorganic Chemistry, University of Nijmegen, Toernooiveld 1, 6525 ED Nijmegen, The Netherlands. Correspondence e-mail: ptb@sci.kun.nl

In 1993, alternative normalized structure factors for incommensurately modulated structures were defined [Lam, Beurskens & van Smaalen (1993). *Acta Cryst.* A**49**, 709–721]. The probability distribution associated with the structure invariants $E(-\mathbf{H})E(\mathbf{H}')E(\mathbf{H}-\mathbf{H}')$ has approximately the same functional form as the Cochran distribution. It was shown, however, that triplet-phase relations are relatively less reliable when satellites are involved [de Gelder, Israël, Lam, Beurskens, van Smaalen, Fu & Fan (1996). *Acta Cryst.* A**52**, 947–954]. In the present paper, an alternative approach is presented: instead of studying the distribution of a three-phase invariant, the probability distribution of the phase sum of two first-order satellite reflections $(h,k,l,1$ and $h',k',l',-1)$ has been derived under the assumption that the phase of the associated main reflection $(h+h',k+k',l+l',0)$ can be calculated from the known main (or averaged) structure. Intensive tests with randomly generated artificial structures and one real structure show a significant improvement of direct-methods phase-sum statistics. Functional similarities with conventional direct methods, employing normalized structure factors and the Cochran distribution, are discussed.

1. Introduction

In the past, normalized structure factors (E) were defined for main reflections and satellite reflections of incommensurately modulated structures (Lam *et al.*, 1992, 1993, 1994). The conventional definition of normalized structure factors is

$$E(\mathbf{H}) = F(\mathbf{H})/g(\mathbf{H}), \tag{1}$$

where $g(\mathbf{H})$ is a real function that compensates for the fall-off of the scattering power of the atoms with increasing $\sin(\theta)$, including the effect of the overall temperature factor. By introduction of a modified expression for the $g(\mathbf{H})$ that includes the overall modulation effects, this definition was extended to modulated structures. Different expressions for $g(\mathbf{H})$ were required for main reflections and satellites. The statistical distributions of the magnitudes of the newly defined E values seem to obey similar distributions to those known for non-modulated crystals.

The $\sum_2$ relationship

$$\varphi(\mathbf{H}) \approx \varphi(\mathbf{H}') + \varphi(\mathbf{H}-\mathbf{H}') \tag{2}$$

($\approx$ means 'probably close to'), which is more reliable if all three corresponding $|E|$ values are large, was tested for a series of incommensurately modulated structures by de Gelder *et al.* (1996). This was performed by determining the probability density of the three-phase structure invariant, being defined as the phase sum

$$\Phi(\mathbf{H},\mathbf{H}') = \varphi(-\mathbf{H}) + \varphi(\mathbf{H}') + \varphi(\mathbf{H}-\mathbf{H}')$$

and given by the Cochran distribution[1] (Cochran, 1955):

$$P(\Phi) = [2\pi I_0(\kappa)]^{-1} \exp[\kappa\cos(\Phi)], \tag{3}$$

where I_n is a modified Bessel function of the first kind and

$$\kappa(\mathbf{H},\mathbf{H}') = 2C|E(-\mathbf{H})E(\mathbf{H}')E(\mathbf{H}-\mathbf{H}')| \tag{4}$$

with

$$C = \sum_{j=1}^{N} Z_j^3 \left[\sum_{j=1}^{N} Z_j^2\right]^{-3/2}, \tag{5}$$

[1] For centrosymmetric structures, the phases φ are restricted to 0 and π. The probability that the sign relationship $s(\mathbf{H}) \cong s(\mathbf{H}')s(\mathbf{H}-\mathbf{H}')$ is true is given by Cochran & Woolfson (1955) as $P^+(x) = \frac{1}{2} + \frac{1}{2}\tanh(x)$ with $x(\mathbf{H},\mathbf{H}') = C|E(-\mathbf{H})E(\mathbf{H}')E(\mathbf{H}-\mathbf{H}')| = \kappa(\mathbf{H},\mathbf{H}')/2$.

N is the number of atoms in the unit cell. For equal-atom structures, $C = N^{-1/2}$.

From these tests, it was concluded that the probability distribution associated with these structure invariants approximately has the same functional form as the Cochran distribution. However, when in triplet-phase relations (first-order) satellite reflections are involved, it appears that they are less reliable than when only main reflections are involved (de Gelder *et al.*, 1996). Since probability distributions have not been mathematically derived so far, it may well be possible that Cochran's description of the probability of triplet phase relations should not be applied as such when satellite reflections are involved. Therefore, an investigation to obtain a better theoretical foundation and practical approach was necessary.

This report deals with the derivation of the probability distribution of structure factors of two first-order satellite reflections; the validity and the actual applicability of this new distribution function were tested and compared to the Cochran distribution applying the E values as defined by Lam *et al.* (1993).

2. The derivation of the probability distribution of two satellite structure factors

In a three-dimensionally periodic crystallite or single crystal with reciprocal lattice ($\mathbf{a}^*$, $\mathbf{b}^*$, $\mathbf{c}^*$), the structure factor for the direction $\mathbf{K} = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*$ can be expressed as a function of the contents of a single unit cell since all unit cells are assumed to be identical,

$$F_{\mathbf{K}} = \sum_{j=1}^{N} f_j(\mathbf{K}) \exp(2\pi i \mathbf{K} \cdot \mathbf{r}_j). \tag{6}$$

In contrast, in modulated crystals, the atomic positions $\mathbf{r}_j$ are not constant but they can vary from cell to cell as a result of one (or more) modulation waves. This report discusses the influence of a single displacive modulation wave only. Such a modulation leads to an additional index (m) with respect to the modulation vector $\mathbf{q}$ in order to index the total diffraction pattern,

$$q = \alpha\mathbf{a}^* + \beta\mathbf{b}^* + \gamma\mathbf{c}^*. \tag{7}$$

As a result, the satellite diffraction vectors are expressed using four integer indices (h, k, l, m),

$$\mathbf{H} = \mathbf{K} + m\mathbf{q}. \tag{8}$$

In this report, only first-order satellites ($m = \pm 1$) will be considered.

A useful way to handle the unit-cell dependency of $\mathbf{r}_j$ is to introduce so-called 'basic' coordinates $\mathbf{r}_j(b)$, which are assumed to be constant from cell to cell, on top of which atomic modulation functions $\mathbf{u}_j$ are present. Together they determine the actual position $\mathbf{r}_{j\mathbf{n}}$ of atom number j in a unit cell referred to by a lattice vector $\mathbf{n}$,

$$\mathbf{r}_{j\mathbf{n}} = \mathbf{r}_j(b) + \mathbf{u}_j(0) + \mathbf{n} + \mathbf{u}_j\{\mathbf{q} \cdot [\mathbf{r}_j(b) + \mathbf{n}]\}. \tag{9}$$

The atomic modulation functions are assumed to be periodic functions [$\mathbf{u}_j(t) = \mathbf{u}_j(t+1)$] so they can be expanded in a Fourier series (*e.g.* Yamamoto, 1982; Perez-Mato *et al.*, 1986). The zero-order term of this series [$\mathbf{u}_j(0)$] is usually combined with $\mathbf{r}_j(b)$ into the average position $\mathbf{r}_j(0)$,

$$\mathbf{r}_j(0) = \mathbf{r}_j(b) + \mathbf{u}_j(0). \tag{10}$$

In practice, the determination of an incommensurately modulated structure involving main reflections ($m = 0$) only results in the average structural positions $\mathbf{r}_j(0)$. Owing to the modulation, the $\mathbf{r}_j(0)$ should not be looked upon strictly as being point positions but be interpreted as being the centers of the area(s) in the unit cell to which the possible positions $\mathbf{r}_{j\mathbf{n}}$ are limited.

Occasionally, one encounters in the literature $\mathbf{q} \cdot \mathbf{r}_j(0)$ as argument of $\mathbf{u}_j$ instead of $\mathbf{q} \cdot \mathbf{r}_j(b)$. The latter is more logical because otherwise $\mathbf{u}_j$ contains its own zero term. From the point of deriving joint probability distributions, it does not make much difference since either choice results in $\mathbf{r}_j$ changing continuously from cell to cell. Moreover, incommensurately modulated structures exist for which no basic structural positions can be defined (Janner & Janssen, 1980).

2.1. Structure-factor expressions in the presence of a single displacive modulation. No symmetry (except identity operation)

In this section, only the basic expressions of the structure factor will be discussed, the role of the superspace-group symmetry is left out.

In the presence of a displacive modulation, the structure factor scattered in the direction $\mathbf{H}$ can be expressed as

$$F_{\mathbf{H}} = \sum_{j=1}^{N} \sum_{\mathbf{n}} f_j(\mathbf{H}) \exp(2\pi i \mathbf{H} \cdot \mathbf{r}_{j\mathbf{n}}), \tag{11}$$

in which now an explicit summation is present over all contributing unit cells, each cell being referred to by a different lattice vector $\mathbf{n}$ while the atomic positions $\mathbf{r}_{j\mathbf{n}}$ are defined as in (9). It is assumed that the atomic scattering factors are not modulated, so $f_{j\mathbf{n}}(\mathbf{H}) = f_j(\mathbf{H})$, but they are allowed to be complex valued as a result of anomalous scattering:

$$f_j(\mathbf{H}) = |f_j(\mathbf{H})| \exp[i\delta_j(\mathbf{H})]. \tag{12}$$

Taking (8), (9) and (10) into account, (11) becomes

$$F_{\mathbf{H}} = \sum_{j=1}^{N} \sum_{\mathbf{n}} f_j(\mathbf{H}) \exp\{2\pi i[\mathbf{K} \cdot \mathbf{r}_j(0) + m t_{j\mathbf{n}}]\} \exp[2\pi i \mathbf{H} \cdot \mathbf{u}_j(t_{j\mathbf{n}})] \tag{13}$$

with $t_{j\mathbf{n}} = \mathbf{q} \cdot [\mathbf{r}_j(b) + \mathbf{n}]$.

Because $\mathbf{q}$ is incommensurate with respect to ($\mathbf{a}^*$, $\mathbf{b}^*$, $\mathbf{c}^*$), $t_{j\mathbf{n}}$ can take any value in the interval [0, 1] when $\mathbf{n}$ ranges over all unit cells. In practice, the summation over the contributing $t_{j\mathbf{n}}$ (all unit cells) is usually carried out by replacing the summation by an integral involving a continuous variable τ_j [0, 1], which is often referred to as an 'internal' or 'fourth-dimensional' variable. Note that τ_j depends explicitly on j because it

is defined in relation to $\mathbf{r}_j(b)$. By means of the integration, *via* $\mathbf{u}_j$, a 'string' of atomic positions is defined that in fact corresponds with the atomic modulation function,

$$F_{\mathbf{H}} = \sum_{j=1}^{N} f_j(\mathbf{H}) \exp[2\pi i \mathbf{K} \cdot \mathbf{r}_j(0)] \times \int_0^1 \exp(2\pi i m \tau_j) \exp[2\pi i \mathbf{H} \cdot \mathbf{u}_j(\tau_j)] \, \mathrm{d}\tau_j. \quad (14)$$

After choosing a displacive modulation function, the integration of (14) can be carried out so an internal-coordinate-independent $F_{\mathbf{H}}$ is obtained. At this point, we notice that this approach, common practice in analysis of incommensurately modulated structures and therefore referred to here as 'conventional', affects not only the satellite structure-factor expressions but also (the derivation of) their joint probability distribution. Thus, before actually discussing the latter in more detail, we will briefly outline the general derivation procedure and argue that also an alternative to the 'conventional' approach is possible. Mathematical details of this alternative will be discussed elsewhere.

The derivation of a joint probability distribution of structure factors in normal direct methods involves the definition of a series expansion in moments, each moment consisting of a product of trigonometric parts of one or more structure factors. When this expansion in the moments is set up, the primitive random variables that are associated with the atomic coordinates are kept constant. Only after the moments expansion has been set up can the actual contribution of each moment to the probability distribution be calculated by allowing the primitive random variables to take on their possible values. In this way, the influence of the primitive random variables on the joint distribution of structure factors can be assessed. In the absence of any prior knowledge, the primitive random variables are assumed to be uniformly distributed in the unit cell. This comes down to integrating $\mathbf{r}_j$ on the interval [0, 1].

In the case of satellite structure factors, the $\mathbf{r}_j(0)$ together with τ_j determine a continuous sequence of atomic positions $\mathbf{r}_{j\mathbf{n}}$ of an atom (labeled j) in the subsequent unit cells (labeled $\mathbf{n}$). The prior functional distribution is described by means of a modulation function. The coordinate function of $\mathbf{r}_j(0)$ together with τ_j suggests a joint probability distribution should be set up with a structure-factor expression like (13) and the τ_j integration be carried out only after setting up the joint distribution of structure factors.

2.1.1. Sinusoidal displacive modulations. In descriptions of modulated crystals, it should be allowed for that with each atom j a different atomic modulation function $\mathbf{u}_j$ is associated. In literature, various expressions have been proposed for a displacive $\mathbf{u}_j$. In this report, we will consider only harmonic sinusoidal type modulations that in its general form can be noted (Petricek *et al.*, 1985, Petříček & Coppens, 1988; Paciorek & Kucharczyk, 1985; Steurer, 1987) as

$$\mathbf{u}_j(t_{j\mathbf{n}}) = \mathbf{U}_j \sin[2\pi(t_{j\mathbf{n}} - \Psi_{jk})], \quad k = 1, 2, 3, \quad (15)$$

in which six independent variables are present, the three components of $\mathbf{U}_j$ and three phase components Ψ_{jk} ($k = 1, 2, 3$). Expression (15) can be rearranged into an expression involving two (independent) vectors $\mathbf{V}_j$ and $\mathbf{W}_j$ (McConnell & Heine, 1984):

$$\mathbf{u}_j(t_{j\mathbf{n}}) = \mathbf{V}_j \sin(2\pi t_{j\mathbf{n}}) + \mathbf{W}_j \cos(2\pi t_{j\mathbf{n}}). \quad (16)$$

A simpler atomic modulation function, referred to as a rectilinear modulation (Paciorek & Kucharczyk, 1985) is arrived at under the assumption that the phase components Ψ_{jk} along the three axes are identical ($\Psi_j = \Psi_{jk}$ for $k = 1, 2, 3$),

$$\mathbf{u}_j(t_{j\mathbf{n}}) = \mathbf{U}_j \sin[2\pi(t_{j\mathbf{n}} - \Psi_j)] \quad \text{with } \Psi \in [0, 1]. \quad (17)$$

This model, which has been applied amongst others by Petricek *et al.* (1985), contains only four variables, $\mathbf{U}_j$ and Ψ_j, for each modulated atom.

2.2. Structure-factor expressions in the presence of a sinusoidal modulation

After inserting a modulation expression like (16) or (17) in (13) and changing over to a continuous variable τ_j, like from (13) to (14), the integral involving the continuous variable τ_j can be carried out if $\exp[2\pi i \mathbf{H} \cdot \mathbf{u}_j(\tau_j)]$ is expanded in a series of Bessel functions using

$$\exp[iz \sin(\alpha)] = \sum_{n=-\infty}^{\infty} J_{-n}(z) \exp(-in\alpha). \quad (18)$$

In the case of the rectilinear modulation (17), an attractive simplification of the structure factor is arrived at,

$$F_{\mathbf{H}} = \sum_{j=1}^{N} f_j(\mathbf{H}) \exp[2\pi i \mathbf{K} \cdot \mathbf{r}_j(0)] J_{-m}(2\pi \mathbf{H} \cdot \mathbf{U}_j) \exp(-i2\pi m \Psi_j) \quad (19)$$

[see *e.g.* equation (6) in Petricek *et al.* (1985)]. This expression has been shown to be very useful in the refinement of modulated structures.

The same procedure for the general sinusoidal modulation (16) gives the integral

$$\int_0^1 \sum_{n1=-\infty}^{\infty} \sum_{n2=-\infty}^{\infty} J_{-n1}(2\pi \mathbf{H} \cdot \mathbf{V}_j) J_{-n2}(2\pi \mathbf{H} \cdot \mathbf{W}_j) \times \exp[2\pi i \tau(m - n_1 - n_2) - i n_2 \pi/2] \, \mathrm{d}\tau. \quad (20)$$

After interchanging the integral and the summations, the integral can be carried out, leading to the condition $n_2 = m - n_1$. Using $J_{-n}(x) = J_n(x) \exp(i\pi n)$, the integral part of (20) becomes

$$\exp(-im\pi/2) \sum_{n1} J_{n1}(2\pi \mathbf{H} \cdot \mathbf{V}_j) J_{n1-m}(2\pi \mathbf{H} \cdot \mathbf{W}_j) \exp(i3\pi n_1/2). \quad (21)$$

After applying Graf's theorem (Abramowitz & Stegun, 1965),

$$J_n(\omega) \exp(in\chi) = \sum_{k=-\infty}^{\infty} J_{n+k}(u) J_k(v) \exp(ik\alpha) \quad (22)$$

with

$$\omega = [u^2 + v^2 - 2uv\cos(\alpha)]^{1/2}$$

and

$$\exp(i\chi) = \omega^{-1}[u - v\exp(-i\alpha)]$$

to (21), the structure-factor expression is obtained as

$$F_{\mathbf{H}} = \sum_{j=1}^{N} f_j(\mathbf{H})\exp[2\pi i\mathbf{K}\cdot\mathbf{r}_j(0)]J_{-m}(\omega_j)\exp[im(\chi_j - \pi/2)] \tag{23}$$

$$\omega_j = 2\pi|\mathbf{H}|[|\mathbf{V}_j|^2\cos^2(\psi') + |\mathbf{W}_j|^2\cos^2(\psi'')]^{1/2}$$

$$\mathbf{H}\cdot\mathbf{V}_j = |\mathbf{H}||\mathbf{V}_j|\cos(\psi'), \quad \mathbf{H}\cdot\mathbf{W}_j = |\mathbf{H}||\mathbf{W}_j|\cos(\psi'').$$

See Petricek *et al.* (1985) and Petříček & Coppens (1988) for an alternative approach.

2.3. The joint probability distribution of structure factors in the case of a single displacive sinusoidal modulation

2.3.1. Random and primitive random variables. The problem of phasing satellite structure factors of incommensurately modulated structures by means of direct methods differs considerably from the phase problem in *ab initio* direct methods. The mathematical model behind the latter is based upon N primitive random variables, each one being associated with one of the N atomic positions. Lack of knowledge concerning the atomic positions is translated into the model that the N primitive random variables are independent and uniformly distributed in the unit cell.

In incommensurately modulated compounds, the phase problem of the main reflections ($\mathbf{K}$) can be assumed to be solved by a standard *ab initio* technique so that the set of phased structure factors ($F_{\mathbf{K}}$) determines the average structural model [$\mathbf{r}_j(0)$].

Taking this structural model as a start, the problem is to phase the satellite structure factors correctly or, equivalently, to find the correct parameter values of the atomic modulation functions.[2] This has the following consequences:

(i) If the $F_{\mathbf{K}}$, and correspondingly $\mathbf{r}_j(0)$, are (assumed to be) completely known, it is not necessary to consider the $F_{\mathbf{K}}$ as random variables, *i.e.* as functions of primitive random variables for which a uniform distribution in the unit cell has to be assumed.

(ii) In contrast to $F_{\mathbf{K}}$, the phases of $F_{\mathbf{H}}$ are unknown while the $F_{\mathbf{H}}$ are functions of the atomic modulation functions. Because the latter determine the cell-dependent atomic positions $\mathbf{r}_{jn}$, it seems a reasonable assumption to take the atomic modulation function parameters as primitive random variables. In the case of a rectilinear modulation, this comes down to taking the $\mathbf{U}_j$ and the Ψ_j as the primitive random variables while, for the general sinusoidal modulation, the $\mathbf{V}_j$ and $\mathbf{W}_j$ fulfil this role.

[2] This problem is somewhat comparable to the case in which a superstructure needs to be recovered while a substructure is available (*e.g.* Cascarano *et al.*, 1987; Cascarano & Giacovazzo, 1988).

Like in normal direct methods, the primitive random variables are assumed to be independently distributed and, unless more precise information is available, also to be uniformly distributed within a certain range. Concerning the phase parameters Ψ_j, this comes down to a uniform distribution in the interval [0, 1]. The magnitudes of the vectorial parameters $\mathbf{U}_j$ in the rectilinear modulation and $\mathbf{V}_j$ and $\mathbf{W}_j$ for the general modulation are certainly not uniformly distributed over the complete unit cell. Two possible cases will be considered in some detail:

(i) The magnitudes of the vectors, *i.e.* $|\mathbf{U}|$ in the case of a rectilinear modulation or $|\mathbf{V}|$ and $|\mathbf{W}|$ in the case of a general sinusoidal modulation, are assumed to be known, *e.g.* from a fitting procedure like that devised by Lam *et al.* (1992), while the orientation of the vector(s) is assumed to be arbitrary.

This approach reduces the problem to the calculation of orientational averages, as will be discussed in the following sections.

(ii) In addition to the orientational average of the vector(s), their magnitudes are also not known. One option is to assume a uniform distribution on the interval $[0, |\mathbf{U}_{\max}|]$.

2.4. The joint probability distribution of two first-order satellite structure factors in the case of a single sinusoidal displacive modulation

The joint probability distribution of two structure factors of first-order satellites can be set up in the same way as described in Peschar & Schenk (1991). Let us denote the first-order satellite reflections by $\mathbf{H}_1 = \mathbf{K}_1 + \mathbf{q}$ and $\mathbf{H}_2 = \mathbf{K}_2 - \mathbf{q}$ and consider the phase sum

$$\Phi_2 = \varphi_1 + \varphi_2 \tag{24}$$

with φ_1 being the phase of satellite reflection $\mathbf{H}_1$ and φ_2 being the phase of satellite reflection $\mathbf{H}_2$. In general, (24) will not be a structure invariant. However, it is easily shown by calculating the moments of order N^0 that (19) does become a doublet phase sum invariant, as present between two isomorphously related structure factors, if the modulation parameters $\mathbf{U}_j$ and Ψ_j (or $\mathbf{V}_j$ and $\mathbf{W}_j$) are used as primitive random variables. In order to show this, we follow the definition of the moments as given in Peschar & Schenk (1991) and it is readily observed that the definition of the structure factor therein is analogous to expressions (19) and (23) provided the atomic scattering factors are now replaced by $f_j(\mathbf{H})\exp[2\pi i\mathbf{K}\cdot\mathbf{r}_j(0)]$ while the quantity $\exp(2\pi i\mathbf{H}\cdot\mathbf{r}_j)$, which is associated with the primitive random variables, is now replaced by $J_{-m}(2\pi\mathbf{H}\cdot\mathbf{U}_j)\exp(-i2\pi m\Psi_j)$ and $J_{-m}(\omega_j)\exp[-im(\chi_j - \pi/2)]$ as defined in (19) and (23), respectively. For simplicity, we will discuss now the former only.

2.4.1. Rectilinear modulation. In view of the above and referring to §§2–4 in Peschar & Schenk (1991), the moments expression for two structure factors $F(\mathbf{H}_1)$ and $F(\mathbf{H}_2)$ in the presence of a rectilinear modulation can be expressed as

$$m_{\alpha_1\alpha_2}^{\beta_1\beta_2} \propto \prod_{\nu=1}^{2} \{|f_\nu|^{\alpha_\nu+\beta_\nu} \exp[i\delta_{j\nu}(\alpha_\nu-\beta_\nu)] \exp[2\pi i(\alpha_\nu-\beta_\nu) \times \mathbf{K}_\nu \cdot \mathbf{r}_j(0)]\} \Big\langle \prod_{\nu=1}^{2} \{J_{-m_\nu}^{\alpha_\nu+\beta_\nu}[2\pi(\mathbf{H}_\nu \cdot \mathbf{U}_j)] \times \exp[2\pi i \Psi_j(\alpha_\nu-\beta_\nu)m_\nu]\} \Big\rangle. \quad (25)$$

The average in (25) can be carried out under the assumption of a prior distribution of the primitive random variables associated with the vector $\mathbf{U}_j$ and the phase parameter Ψ_j. Concerning the (primitive random variable associated with) Ψ_j, it is assumed that it is uniformly distributed on the interval [0, 1] and up to order N^0 the same conditions for non-zero moments are obtained as in Peschar & Schenk (1991),

$$(\alpha_1-\beta_1)-(\alpha_2-\beta_2)=0 \quad \text{with } \alpha_1+\alpha_2+\beta_1+\beta_2 \le 2.$$

This gives the following non-zero moments:

α_1	β_1	α_2	β_2	F
1	1	0	0	$F(\mathbf{H}_1)$
0	0	1	1	$F(\mathbf{H}_2)$
1	0	1	0	$F(\mathbf{H}_1)$ and $F(\mathbf{H}_2)$
0	1	0	1	$F(\mathbf{H}_1)$ and $F(\mathbf{H}_2)$.

(26)

The first two are moments of the individual structure factors only, the last two involve the doublet terms.

In the case of a single structure factor, $\alpha_\nu = \beta_\nu = 1$, for each moment an orientational average,

$$\langle J_{-1}^2[2\pi(\mathbf{H}_\nu \cdot \mathbf{U}_j)]\rangle, \quad (27)$$

has still to be carried out because of the presence of $\mathbf{U}_j$ in (25). The calculation of this type of average is discussed in detail in Appendix *A*. Expression (27) resembles closely expression (12) from Lam *et al.* (1992) although their calculation of the orientational average is slightly different from the one given here in Appendix *A*.

The subsequent operations in the derivation of the joint probability distribution of structure factors, the moments-cumulants transformation and summing the contributions of all *N* primitive random variables, can be carried out completely analogously to Peschar & Schenk (1991) and results in

$$z_{\nu\nu} = \sum_{j=1}^{N} |f_j(\mathbf{H}_\nu)|^2 \langle J_{-1}^2[2\pi(\mathbf{H}_\nu \cdot \mathbf{U}_j)]\rangle. \quad (28)$$

The moments in (26) involving two structure factors can be handled in a similar manner. The moments with $\alpha_1 = \alpha_2 = 1$ are:

$$m_{11}^{00} \propto \prod_{\nu=1}^{2} \{|f_{j\nu}| \exp(i\delta_{j\nu}) \exp[2\pi i \mathbf{K}_\nu \cdot \mathbf{r}_j(0)]\} \times \langle J_{-1}[2\pi(\mathbf{H}_1 \cdot \mathbf{U}_j)] J_1[2\pi(\mathbf{H}_2 \cdot \mathbf{U}_j)]\rangle. \quad (29)$$

The calculation of the orientational average present in (29) is explained in detail in Appendix *B*.

The remaining steps in the derivation, the moments-cumulants transformation and the summation of the contributions of all *N* primitive random variables, are again standard and result in

$$z_{12} = \sum_{j=1}^{N} \left\{ \prod_{\nu=1}^{2} f_j(\mathbf{H}_\nu) \exp[2\pi i \mathbf{K}_\nu \cdot \mathbf{r}_j(0)] \right\} \times \langle J_{-1}[2\pi(\mathbf{H}_1 \cdot \mathbf{U}_j)] J_1[2\pi(\mathbf{H}_2 . \mathbf{U}_j)]\rangle. \quad (30)$$

A comparison of the current derivation and the derivation given in Peschar & Schenk (1991) leads to the conclusion that the only difference between the distributions resides in the exact definition of the functions z_{11}, z_{22} and z_{12}. This implies that the joint probability expression given in Peschar & Schenk (1991) for two isomorphously related structure factors {see Peschar & Schenk [1991, equation (60)] with $\mathbf{H}_2 = -\mathbf{H}_1$ and $s_{12} = -1$ in equation (19)} carries over provided the appropriate definitions of $z_{\nu\nu}$ and z_{12} are used. Thus, the joint probability distribution of the random variables ϕ_1 and ϕ_2, associated with the phases φ_1 and φ_2, respectively, and the random variables R_1 and R_2, associated with the structure-factor magnitudes $|F_1|$ and $|F_2|$, respectively, can be expressed as [for the benefit of this report we use a notation that is slightly different from that used in Peschar & Schenk (1991)]

$$P(\phi_1, \phi_2, R_1, R_2) \propto \exp[-G_1^2 - G_2^2 + 2G_1G_2|L_{12}| \times \cos(\phi_1 + \phi_2 - \Delta_{12})] \quad (31)$$

with

$$L_{12} = |L_{12}| \exp(i\Delta_{12}) = |d_{12}| \exp(i\Delta_{12})[1-|d_{12}|^2]^{-1/2}$$
$$d_{12} = z_{12}(z_{11}z_{22})^{-1/2} = |z_{12}| \exp(i\Delta_{12})(z_{11}z_{22})^{-1/2}$$
$$G_1 = R_1(z_{11})^{-1/2}[1-|d_{12}|^2]^{-1/2} \quad (32)$$
$$G_2 = R_2(z_{22})^{-1/2}[1-|d_{12}|^2]^{-1/2} \quad (33)$$

in which $z_{\nu\nu}$ and z_{12} are defined in (28) and (30), respectively. Eventually, working out (31) in more detail gives the new expression for the joint probability distribution of two first-order satellite structure factors in the case of a single sinusoidal displacive modulation

$$P(\phi_1, \phi_2, R_1, R_2) \propto \exp[(z_{11}z_{22} - |z_{12}|^2)^{-1}(-R_1^2 z_{22} - R_2^2 z_{11})] \times \exp[2R_1R_2|z_{12}|(z_{11}z_{22} - |z_{12}|^2)^{-1} \times \cos(\phi_1 + \phi_2 - \Delta_{12})]. \quad (34)$$

The above described expression for the probability distribution in the case of the presence of a rectilinear modulation is a simplification of the expression valid in the presence of a general sinusoidal modulation. The assumption is made that the phase components Ψ_{jk} along the three axes are identical (for $k = 1, 2, 3$).

The joint probability distribution in the presence of a general sinusoidal modulation can be obtained in a way similar to that described for a rectilinear modulation. Expression (31) again holds and only $z_{\nu\nu}$ and z_{12} need to be redefined because the orientational averages are different. In our investigation, we only studied the validity of (34), thus using rectilinear modulations for the randomly generated artificial structures and assuming a rectilinear modulation in our 'real structure' test case.

3. Comparison of the new distribution (34) and the Cochran distribution

Before discussing the test results of (34), it is useful to compare (34) with the standard direct-methods expression of Cochran [equations (3)–(5)].

3.1. Triplets *versus* doublets

The triplet phase relationship is the most used phase relation in direct methods. For incommensurately modulated compounds with a known (or assumed known) basic structure, the origin is fixed and any phase of a main reflection is a seminvariant under these conditions: the phases of the satellites, however, are not fixed by the main (averaged) structure. Now consider a doublet of two satellites $\mathbf{H}_1$ and $\mathbf{H}_2$ with

$$\mathbf{H}_1 = (h_1, k_1, l_1, 1) \quad \text{and} \quad \mathbf{H}_2 = (h_2, k_2, l_2, -1)$$

and with

$$\mathbf{K}_1 = (h_1, k_1, l_1, 0) \quad \text{and} \quad \mathbf{K}_2 = (h_2, k_2, l_2, 0)$$

so

$$\mathbf{K}_{12} = \mathbf{H}_1 + \mathbf{H}_2 = \mathbf{K}_1 + \mathbf{K}_2 = (h_1 + h_2, k_1 + k_2, l_1 + l_2, 0).$$

The main reflection $\mathbf{K}_{12}$ (defined this way, with phase φ_{12}) is the third reflection in the Cochran triplet relation:

$$\varphi_1 + \varphi_2 - \varphi_{12} \cong 0.$$

This is to be compared with the result inferred from (34):

$$\varphi_1 + \varphi_2 - \Delta_{12} \cong 0.$$

Once it has been ascertained that Δ_{12}, closely resembling the phase of the triplet-related main reflection $\mathbf{K}_{12}$, is indeed (almost) equal to φ_{12}, in principle almost all direct-methods procedures can be used, and in practice in a very powerful way because all phases φ_{12} are known at the outset.

3.2. Similarity of Δ_{12} and φ_{12}

The phase of z_{12} [in (30)] is determined by N atomic contributions of the form

$$f_j(\mathbf{H}_1) f_j(\mathbf{H}_2) \exp[2\pi \mathrm{i} \mathbf{K}_{12} \cdot \mathbf{r}_j(0)] \langle J_{-1}[2\pi(\mathbf{H}_1 \cdot \mathbf{U}_j)] J_1[2\pi(\mathbf{H}_2 \cdot \mathbf{U}_j)] \rangle. \tag{35}$$

This is almost identical to

$$f_j^2(\mathbf{K}_{12}) \exp[2\pi i \mathbf{K}_{12} \cdot \mathbf{r}_j(0)] \langle J_{-1}[2\pi(\mathbf{H}_1 \cdot \mathbf{U}_j)] J_1[2\pi(\mathbf{H}_2 \cdot \mathbf{U}_j)] \rangle. \tag{36}$$

Expression (36) can be compared with the atomic contributions to the structure factor of reflection $\mathbf{K}_{12}$,

$$f_j(\mathbf{K}_{12}) \exp[2\pi i \mathbf{K}_{12} \cdot \mathbf{r}_j(0)]. \tag{37}$$

The difference between (36) and (37) is the weighting power of atom j: for the structure factor it is f_j (including the atomic displacement factor) while for z_{12} [using (36)] it is

$$f_j^2 \langle J_{-1}[2\pi(\mathbf{H}_1 \cdot \mathbf{U}_j)] J_1[2\pi(\mathbf{H}_2 \cdot \mathbf{U}_j)] \rangle. \tag{38}$$

The last factor in (38), which describes the influence of the modulation on the structure factor, often shows in practice as a (possibly large) pseudo-temperature factor, but it is not expected to influence the calculated phase of Δ_{12} very much. Therefore, the differences between the phases Δ_{12} and φ_{12} are expected to be very small for equal-atom structures. For heavy-atom structures, these differences are (almost) entirely caused by the larger impact of f^2 in (36) relative to f in (37) for the heavy atoms. If relatively few heavy atoms are present in the structure, only a few terms in (37) have this enhanced weight, and the corresponding phase Δ_{12} should not differ much from φ_{12}. For strong reflections, which are of prime importance in practical direct methods, most atoms have similar constructive contributions to both Δ_{12} and φ_{12}, so again the difference will be small. Only for weak reflections in heavy-atom structures may significant phase differences be expected. This is subject to further tests, see §4.

3.3. Normalization

As commonly known, the derivation of the joint probability distribution of structure-factor magnitudes and phases automatically ends up in an expression in which normalized magnitudes ($|E|$'s) are easily recognized. Expressing such a distribution in terms of E's simplifies the equations and the computational procedures. The major properties of E's are: (i) the average of $|E^2|$ is equal to unity; (ii) the 'local' average is independent of the scattering angle; and (iii) the E's appear in a simple form in the probability equations (*e.g.* Cochran, 1955). In the present doublet distribution function (34), property (iii) may not be obvious. Nevertheless, the concept of normalization is hidden in the formulations. First of all, the 'normalized' average, re (i), is not of any relevance, except that the observed structure-factor amplitudes are on an approximately absolute scale. The normalization with respect to the $\sin(\theta)$ dependency, however, is important in classical direct methods: the phasing power of a participating reflection depends on its relative intensity (relative to the average intensity at a given scattering angle) and *not at all* on its scattering angle. This normalization aspect is apparent in (34).

Comparing the exponential form in (34) with a similar form in the Cochran distribution,

$$2C|E(\mathbf{K}_1)E(\mathbf{K}_2)E(-\mathbf{K}_{12})| \cos(\varphi_1 + \varphi_2 - \varphi_{12}), \tag{39}$$

suggests the definitions of normalized structure factors G_1 and G_2 and a quasi-normalized magnitude G_{12}, *via* (32) and (33) (upon substitution of $|d_{12}|$):

$$G_1 = F_1(z_{22})^{-1/2}[z_{11}z_{22} - |z_{12}|^2]^{-1/2} \tag{40}$$

$$G_2 = F_2(z_{11})^{-1/2}[z_{11}z_{22} - |z_{12}|^2]^{-1/2} \tag{41}$$

$$G_{12} = z_{12}[z_{11}z_{22}]^{-1/2} \tag{42}$$

and we define the structure-invariant triplet factor

$$K = 2|G_1G_2G_{12}| = 2|F_1F_2z_{12}|[z_{11}z_{22} - |z_{12}|^2]^{-1}, \tag{43}$$

which is present in (34) and plays the same role as κ in non-incommensurate direct methods.

K is expected to be 'normalized' with respect to the scattering angle. Both $|F_1|$ and $|F_2|$ depend on f while z_{11}, z_{22} and z_{12} depend on f^2 so the order dependency on f in $|F_1F_2z_{12}|$ and

Table 1
Characteristics of artificial non-centrosymmetric (A) and centrosymmetric (C) $C_{50-x-y}S_xFe_y$ structures with displacive rectilinear modulations.

The orientational averages were calculated according to Appendices *A* and *B*.

Compound code names	Composition (in the asymmetric unit)	Modulation amplitude†	Modulation direction
A1 to A4‡	50 C atoms	Random	Random
A1a and A3a	50 C atoms	Fixed	Fixed
A1b to A4b	50 C atoms	Fixed	Random
$A1S_6$ to $A4S_6$	44 C and 6 S atoms	Fixed	Random
$A1S_{16}$ to $A4S_{16}$	34 C and 16 S atoms	Fixed	Random
$A1Fe_6$ to $A4Fe_6$	44 C and 6 Fe atoms	Fixed	Random
$A1Fe_{16}$ to $A4Fe_{16}$	34 C and 16 Fe atoms	Fixed	Random
C1 to C4	25 C atoms	Random	Random
C1a and C3a	25 C atoms	Fixed	Fixed
C1b to C4b	25 C atoms	Fixed	Random
$C1S_6$ to $C4S_6$	22 C and 3 S atoms	Fixed	Random
$C1S_{16}$ to $C4S_{16}$	17 C and 8 S atoms	Fixed	Random
$C1Fe_6$ to $C4Fe_6$	22 C and 3 Fe atoms	Fixed	Random
$C1Fe_{16}$ to $C4Fe_{16}$	17 C and 8 Fe atoms	Fixed	Random

† The random amplitude varies between 0.0 and 0.4 Å, the fixed amplitude is 0.2 Å. ‡ A1 to A4 means A_i, $i = 1, 4$.

$[z_{11}z_{22} - |z_{12}|^2]^{-1}$ is the same. We note that G_{12} can be interpreted as a quasi-normalized structure factor of the triplet-related main reflection $\mathbf{K}_{12}$ so the normalization concepts for normal and modulated structures turn out to be largely similar. For more details about practical aspects of the normalization of modulated structures, see Lam *et al.* (1993).

4. Numerical test procedures

Numerical tests have been performed with structure factors from several randomly generated non-centrosymmetric and centrosymmetric structures, with and without heavy atoms, and with different one-dimensional rectilinear displacive modulations. All structures have 50 atoms in the unit cell; the equal-atom structures contain C atoms, the heavy-atom structures contain also S or Fe atoms. In all cases, main reflections and first-order satellites were generated with $\sin(\theta)/\lambda < 0.75$ Å^{-1}.

Table 1 shows an overview of the generated structures together with their code names, compositions and kinds of modulation. The structures used have the same atomic positions and code names as used by de Gelder *et al.* (1996), but now heavy atoms are introduced and different modulation functions are used (details given in Table 1). Structure factors were calculated in the same way as described by de Gelder *et al.* (1996)

In addition, we also used a data set that was obtained from an experimentally determined centrosymmetric structure, but instead of the measured intensities calculated structure factors were employed, thus avoiding complications due to an incomplete data set as published in the literature. The code name for this structure is PECO; for further details, we refer to §6.

We distinguish the following triplets:

mmm for a triplet formed by 3 main reflections;

ssm for a triplet (or doublet!) formed by 2 satellites and 1 main reflection;

sss for a triplet of 3 satellites (not used in this paper).

As a first test, we wished to see the influence of the different types of modulation as compared to the tests made by de Gelder *et al.* (1996). Therefore all our test structures were subjected to the same calculations as performed by de Gelder *et al.* (1996) with the same pseudo-normalized structure-factor definition of Lam *et al.* (1993). Fig. 1 shows the Cochran distribution plot [(*a*) for *mmm* triplets and (*b*) for *ssm* triplets) as obtained by de Gelder *et al.* (1996) for 'their' structure A1. Fig. 2 shows the results obtained for 'our' structure A1.

The overall statistics depicted in Fig. 1 do not differ significantly from those in Fig. 2 and similar results were obtained for other equal-atom test structures.

5. Test calculations I. Comparison of Δ_{12} and φ_{12}

Earlier it was shown that z_{12} is associated with the main reflection $\mathbf{K}_{12}$ that completes the triplet phase relationship $\mathbf{H}_1 + \mathbf{H}_2 = \mathbf{K}_{12}$. It was argued that the phase Δ_{12} occurring in (34) is expected to be (for most cases) approximately equal to the phase φ_{12} of the main reflection $\mathbf{K}_{12}$. This is an observation, which is of importance for the calculation of phase relationships for solving an unknown incommensurate structure with direct methods. In order to establish whether the doublet phase-sum estimate Δ_{12} equals the known phase φ_{12} of the main reflection $\mathbf{K}_{12}$, some statistics have been made.

Table 2 lists some representative individual phase differences $|\Delta_{12} - \varphi_{12}|$ at several K levels for various non-centrosymmetric structures. Table 3 gives the root-mean-square (r.m.s.) deviation, $\delta_{\rm rms}$, between φ_{12} and Δ_{12} for the two non-centrosymmetric structures. The r.m.s. deviation is defined as

$$\delta_{\rm rms} = \left\{ \left[\sum_1^n |\varphi_{12} - \Delta_{12}|^2 \right] \Big/ n \right\}^{1/2}$$

with n being the number of doublets participating in a certain interval of K.

Table 2
Phase differences $|\Delta_{12} - \varphi_{12}|$ (in °) *vs* intervals of the largest K values in the case of non-centrosymmetric structures.

Structure	Interval of K	Individual phase differences of single doublets
A1	5.0–7.0	0.02 0.81
A1a	9.0–11.0	0.00 0.00
A1b	6.0–10.0	1.18 0.04 1.18
A1S6	13.0–14.0	13.66
A1S16	8.0–9.0	1.71
A1Fe6	18.0–23.0	3.65 0.51
A1Fe16	7.0–8.0	5.66

Table 3
Results of the r.m.s. deviation, $\delta_{\rm rms}$, between Δ_{12} and φ_{12} for two non-centrosymmetric structures *vs* intervals of K.

	A1b		$A1Fe_{16}$	
K	n	$\delta_{\rm rms}$ (°)	n	$\delta_{\rm rms}$ (°)
5.0–8.0	11	1.0	34	3.9
4.0–5.0	58	1.2	169	4.7
3.0–4.0	262	2.2	1137	5.0
2.0–3.0	2389	2.2	6799	6.0
1.0–2.0	42349	3.4	80218	7.4
0.0–1.0	39906587	17.7	39863299	28.2

In Table 4, averages are given of the absolute difference, $|\Phi_D|$, of the doublet phase sum and its estimate according to (43):

$$\Phi_D = \varphi_1 + \varphi_2 - \Delta_{12} \tag{44}$$

and the absolute triplet phase sum (*ssm* case), $|\Phi_T|$,

$$\Phi_T = \varphi_1 + \varphi_2 + (-\varphi_{12}) \tag{45}$$

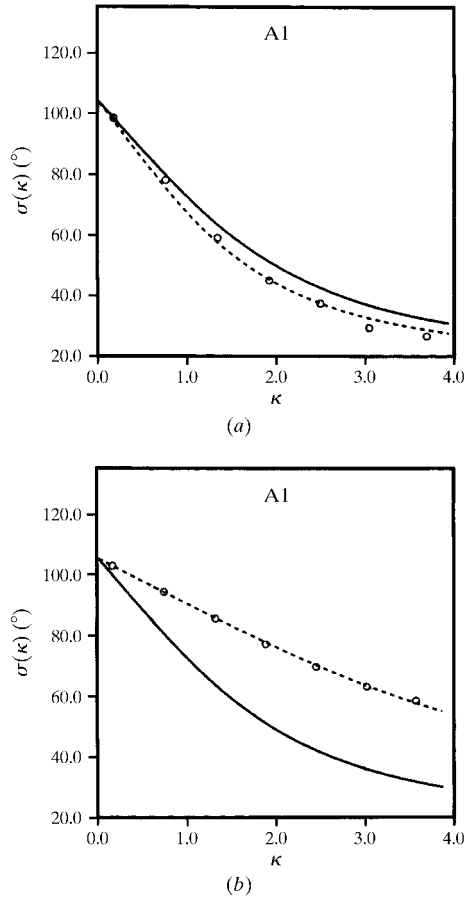

Figure 1
Standard deviation of the triplet phase sum, $\sigma(\kappa)$, as a function of κ for structure A1 as was defined in de Gelder *et al.* (1996). Circles represent experimental values. Solid curves represent the theoretical distribution for non-modulated structures according to Cochran. Dashed curves represent a fitted curve through experimental points. (*a*) Standard deviations involving *mmm* triplets. (*b*) Standard deviations involving *ssm* triplets.

at various K levels and for several non-centrosymmetric structures. For the centrosymmetric structures, the quantities P_D and P_T are given as the percentage of correct doublet signs and the percentage of correct triplet signs with triplets ranked according to decreasing K, respectively. For shortness, the derivation of the joint probability distribution of two first-order satellite structure factors in the case of a single sinusoidal displacive modulation in the centrosymmetric case has not been given. Its functional form is that of the Cochran & Woolfson (1955) expression (see footnote 1 in *Introduction*) with $x(\mathbf{H}, \mathbf{H}')$ being replaced by $|F_1F_2z_{12}/(z_{11}z_{22} - |z_{12}|^2)|$.

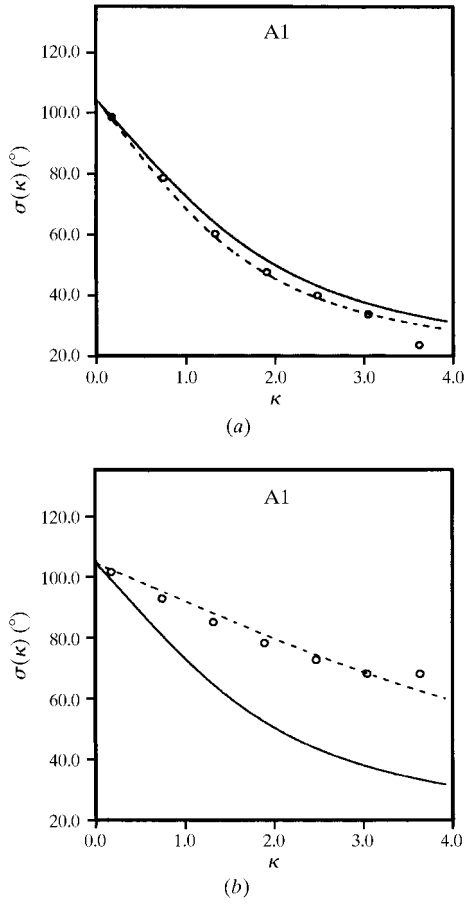

Figure 2
Standard deviation of the triplet phase sum, $\sigma(\kappa)$, as a function of κ for structure A1 as is defined in this study and presented in Table 1. See Fig. 1 for explanations.

Table 4
Average phase differences (in °), $\langle|\Phi_D|\rangle$ and $\langle|\Phi_T|\rangle$ *vs* intervals of K in the case of non-centrosymmetric structures.

Percentages (/100) of correct sign relations P_T (triplets) and correct sign relations P_D (doublets) *vs* K. Note that for statistical significance only intervals are reported with 20 or more participating doublets.

	A1		A1a		A1b		$A1S_6$		$A1S_{16}$		$A1Fe_6$		$A1Fe_{16}$	
K	$\langle\|\Phi_D\|\rangle$	$\langle\|\Phi_T\|\rangle$	$\langle\|\Phi_D\|\rangle$	$\langle\|\Phi_T\|\rangle$	$\langle\|\Phi_D\|\rangle$	$\langle\|\Phi_T\|\rangle$	$\langle\|\Phi_D\|\rangle$	$\langle\|\Phi_T\|\rangle$	$\langle\|\Phi_D\|\rangle$	$\langle\|\Phi_T\|\rangle$	$\langle\|\Phi_D\|\rangle$	$\langle\|\Phi_T\|\rangle$	$\langle\|\Phi_D\|\rangle$	$\langle\|\Phi_T\|\rangle$
12.0–24.0	–	–	–	–	–	–	–	–	–	–	14.3	16.5	–	–
10.0–12.0	–	–	–	–	–	–	15.1	17.8	–	–	17.0	17.1	–	–
8.0–10.0	–	–	–	–	–	–	20.5	20.7	–	–	19.6	20.6	–	–
7.0–8.0	–	–	–	–	–	–	24.6	26.1	–	–	20.0	21.8	–	–
6.0–7.0	–	–	40.4	40.4	–	–	23.4	25.2	25.1	24.5	22.3	22.7	–	–
5.0–6.0	–	–	36.8	36.8	–	–	26.5	28.3	31.1	30.7	23.8	24.1	–	–
4.0–5.0	28.8	28.3	40.7	40.7	37.0	37.1	29.6	30.7	32.9	32.9	26.7	27.1	24.8	25.5
3.0–4.0	39.4	39.0	46.6	46.4	43.0	42.7	32.8	34.0	37.5	37.0	31.3	31.5	23.7	23.6
2.0–3.0	46.7	46.4	55.6	55.6	49.7	49.4	41.4	42.6	45.8	45.2	38.8	38.9	30.5	30.4
1.0–2.0	61.1	60.8	63.4	63.4	62.3	61.9	53.6	54.6	58.3	57.9	51.0	51.0	44.2	43.9
0.0–1.0	87.2	86.3	86.1	86.1	87.1	86.5	84.3	84.0	86.5	85.7	83.5	82.7	85.3	84.3

	C1		C1a		C1b		$C1S_6$		$C1S_{16}$		$C1Fe_6$		$C1Fe_{16}$	
K	P_D	P_T	P_D	P_T	P_D	P_T	P_D	P_T	P_D	P_T	P_D	P_T	P_D	P_T
8.0–10.0	–	–	–	–	–	–	–	–	–	–	1.000	1.000	–	–
7.0–8.0	–	–	–	–	–	–	1.000	1.000	–	–	1.000	1.000	–	–
6.0–7.0	–	–	–	–	–	–	1.000	1.000	1.000	1.000	0.994	0.994	0.984	0.984
5.0–6.0	–	–	0.974	0.974	–	–	0.976	0.958	0.957	0.978	0.985	0.970	0.994	0.994
4.0–5.0	–	–	0.947	0.947	1.000	1.000	0.948	0.928	0.966	0.966	0.974	0.967	0.966	0.972
3.0–4.0	–	–	0.850	0.850	0.958	0.958	0.934	0.931	0.935	0.940	0.945	0.941	0.942	0.943
2.0–3.0	0.854	0.858	0.812	0.812	0.855	0.855	0.864	0.852	0.883	0.883	0.886	0.884	0.903	0.901
1.0–2.0	0.766	0.771	0.730	0.730	0.762	0.763	0.772	0.768	0.784	0.785	0.790	0.795	0.809	0.808
0.0–1.0	0.510	0.514	0.515	0.515	0.509	0.512	0.523	0.527	0.520	0.523	0.527	0.535	0.521	0.525

Please note that the number of contributions to each interval (not listed) varies per structure and ranges from almost 40000000 for the interval with the lowest K values to just a few per interval for the larger values of K.[3]

Conclusions from Tables 2 to 4

The phase statistics in Tables 2–4 show that φ_{12} and Δ_{12} do not differ much, especially for the equal-atom cases and at the larger K values. The average $|\Delta_{12} - \varphi_{12}|$ for the 15 doublets in Table 2 is only 2.1° while Table 4 shows that also the average Φ_D *vs* Φ_T values and P_D *versus* P_T do not differ much.

As expected, the differences are larger for S- and Fe-containing structures and at the lowest reliability levels (see Tables 3 and 4). This leads to the conclusion that, keeping the above restrictions in mind, the G_{12} acts as a 'third' structure factor that from a practical point of view changes the doublet into a triplet.

6. Test calculations II. The applicability of the doublet distribution

We have shown that (34), derived for non-centrosymmetric structures, can be quite easily calculated, resulting in an expression that has the same Von Mises form as the Cochran distribution. The question arises which distribution is best for the performance of direct methods.

In order to test this, the standard deviation, $\sigma(\Phi_D)$, of the doublet phase differences, Φ_D (44), as a function of K (43) was compared with the standard deviation, $\sigma(\Phi_T)$, of the triplet phase sum, Φ_T (45), as a function of κ (4) [in the latter case using E values defined according to Lam *et al.* (1993)]. For the tests, intervals (decreasing in K or κ) were chosen to contain approximately equal numbers of doublets and triplets. The calculations have been performed for a series of structures but for shortness only those of A1 are listed in Table 5.[4]

The behavior of the tested structures is different but the general trend with respect to the comparison of the cumulative standard deviations, $\sigma(\Phi_D)$ and σ (Φ_T), is the same. In order to show this, in Table 6 an excerpt is given of the cumulative standard deviations of Φ_D and Φ_T of the approximately 650 and 2700 strongest doublets or triplets, respectively, for all A1 and A3 related test structures.

From the data given in Tables 5 and 6, a tentative conclusion is possible concerning the quality of (43) compared to the procedure developed by Lam *et al.* (1993). At larger K or κ values, the (cumulative) standard deviations are systematically lower in the case of (43). The structures A1a and A3a, having fixed modulation amplitude and fixed modulation direction, show the above-mentioned tendency most convincingly. Using the procedure of Lam *et al.*, the standard deviations are large, even at higher values of κ (4), while, with (43), a large improvement is accomplished. On the other hand, the struc-

[3] The number of contributions to the intervals of K given in Table 4 are typical values.

[4] Results for all other test structures have been deposited and are available from the IUCr electronic archives (Reference: JS0108). Services for accessing these data are described at the back of the journal.

Table 5
Comparison of the standard deviations of Φ_D and Φ_T as calculated using the new distribution (34) (on the left) and the Cochran distribution (4) (on the right) for the artificial structure A1.

Middle of interval K	$\sigma(\Phi_D)$ (°) per interval	$\sigma(\Phi_D)$ (°) cumulative	Cumulative no. of doublets	Middle of interval K	$\sigma(\Phi_T)$ (°) per interval	$\sigma(\Phi_T)$ (°) cumulative	Cumulative no. of triplets
6.37	35.2	35.2	11	5.19	42.7	42.7	11
4.62	45.5	42.3	32	4.12	60.1	54.9	33
4.05	40.1	41.1	74	3.87	63.1	59.6	75
3.59	47.3	44.5	158	3.58	67.9	64.3	166
3.19	51.6	48.3	331	3.29	61.2	62.7	338
2.79	56.4	52.6	670	3.04	70.5	66.8	680
2.43	60.6	56.7	1335	2.78	65.2	66.0	1388
2.11	63.0	60.0	2720	2.53	71.4	68.7	2740
1.83	69.0	64.6	5424	2.31	75.6	72.2	5479
1.56	73.8	69.5	11220	2.08	76.6	74.5	11157
1.32	76.2	73.0	22694	1.86	78.6	76.6	22594
1.10	79.1	76.1	45069	1.66	80.8	78.7	44586
0.91	82.6	79.5	91691	1.45	83.0	80.9	90424
0.73	85.8	82.8	188378	1.27	85.3	83.2	183139
0.57	88.8	85.9	386547	1.09	87.8	85.5	368002
0.43	91.5	88.8	793164	0.92	90.0	87.8	735907
0.31	93.9	91.4	1621082	0.75	92.4	90.1	1465945
0.21	96.1	93.8	3290097	0.60	94.5	92.3	2912006
0.12	98.5	96.5	7413146	0.45	96.8	94.6	5777485
0.06	100.7	98.7	15508954	0.30	99.0	96.9	12014996
0.02	102.1	100.8	39951656	0.12	102.2	100.7	39951656

Table 6
Cumulative standard deviations (in °) of Φ_D and Φ_T of approximately 650 and 2700 doublets and triplets with the largest K and κ values, as ranked according to decreasing K and κ, respectively.

No. of doublets	650		2700			650		2700	
Test structure	$\sigma(\Phi_D)$	$\sigma(\Phi_T)$	$\sigma(\Phi_D)$	$\sigma(\Phi_T)$	Test structure	$\sigma(\Phi_D)$	$\sigma(\Phi_T)$	$\sigma(\Phi_D)$	$\sigma(\Phi_T)$
A1	52.6	66.8	60.0	68.7	A3	62.0	66.5	65.2	70.6
A1a	52.4	84.7	57.4	68.5	A3a	50.6	84.3	58.3	69.7
A1b	50.5	55.3	59.3	61.9	A3b	60.1	56.5	65.1	62.4
$A1S_6$	32.9	33.5	38.2	42.0	$A3S_6$	31.9	42.6	35.4	43.1
$A1S_{16}$	41.8	44.9	46.3	51.0	$A3S_{16}$	42.9	42.6	46.6	49.6
$A1Fe_6$	24.8	28.0	30.1	32.7	$A3Fe_6$	23.7	41.3	27.8	39.3
$A1Fe_{16}$	33.9	49.3	38.0	48.9	$A3Fe_{16}$	40.1	39.2	43.1	43.8

tures A1b and A3b show no improvement using (43) (see supplementary material).

Since the above-mentioned structures are artificial, it would be useful to perform also a test for experimental data. Recently, the centrosymmetric incommensurately displacive modulated structure of PECO [short for (perylene)-Co(maleonitriledithiolate)$_2$(CH_2Cl_2)$_{0.5}$] was determined by direct methods (Lam *et al.*, 1995). The same structure was also used in the investigation of the applicability of the Cochran distribution for incommensurately modulated structures using the E values defined by Lam *et al.* (1993) (de Gelder *et al.*, 1996) and provides a first assessment of both approaches. In Table 7, the percentage (/100) of correct sign relations is listed when ranked according to K and $\kappa/2$ (defined as x, see footnote 1 in *Introduction*) [using $|E|$ values of Lam *et al.* (1992, 1993, 1994)], respectively. The data of PECO show a similar trend to that observed in the artificial structures.

Although the reliability of the triplets does not differ too much from one method to another, at higher K (or x) values the percentage of correctly predicted sign relations is slightly larger using (34).

7. Conclusions and discussion

The new doublet probability distribution for incommensurately modulated structures can be used for the determination of the phases of the satellite reflections for those cases where the atomic positions of the basic structure are known.

The type of modulation assumed at the outset of the present derivation of the doublet probability distribution leads to the same statistical behavior as other types of modulations used in former work on statistical properties of incommensurate structures [Lam *et al.* (1992, 1993, 1994, 1995); de Gelder *et al.* (1996)].

Under the assumption of a known basic structure, it was shown that the phase argument of the resulting probabilistic expression of two first-order satellite reflections $\mathbf{H}_1 = \mathbf{K}_1 + \mathbf{q}$ and $\mathbf{H}_2 = \mathbf{K}_2 - \mathbf{q}$ with phases φ_1 and φ_2 can be expressed as

Table 7
Results of the comparison between the Cochran distribution and the new distribution (34).

The percentage (/100) of correct sign relations P_D (doublets) as a function of K and correct sign relations P_T (triplets) as a function of x (= $\kappa/2$) for PECO.

Middle of interval K	P_D† per interval	P_D† cumulative	Cumulative no. of doublets	Middle of interval x	P_T† per interval	P_T† cumulative	Cumulative no. of triplets
11.91	1.000	1.000	11	29.85	1.000	1.000	11
9.51	1.000	1.000	32	20.23	1.000	1.000	32
8.39	1.000	1.000	74	15.98	1.000	1.000	73
7.29	1.000	1.000	155	13.53	0.976	0.987	155
6.20	1.000	1.000	316	11.62	0.994	0.991	318
5.21	1.000	1.000	641	10.07	0.994	0.992	642
4.32	0.997	0.998	1306	8.53	0.989	0.991	1296
3.49	0.992	0.995	2619	6.95	0.993	0.992	2506
2.75	0.977	0.986	5256	5.41	0.981	0.986	5222
2.08	0.961	0.973	10555	4.03	0.974	0.980	10430
1.50	0.934	0.953	21261	2.84	0.958	0.969	20980
1.02	0.875	0.914	42868	1.90	0.920	0.944	41977
0.64	0.795	0.854	86054	1.19	0.847	0.895	85876
0.37	0.688	0.768	179219	0.69	0.742	0.817	175348
0.20	0.588	0.673	378822	0.39	0.635	0.721	373150
0.10	0.526	0.594	822368	0.21	0.566	0.638	796899
0.04	0.503	0.539	2060021	0.10	0.535	0.581	1797331
0.01	0.496	0.506	8677498	0.03	0.512	0.526	8677498

† P_D and P_T were calculated by determining the ratio between the amount of correct sign relationships and the total amount of sign relationships.

$(\varphi_1 + \varphi_2 - \Delta_{12})$. Although the main reflection $\mathbf{K}_1 + \mathbf{K}_2$ was not considered a random variable, the doublet phase-sum estimate Δ_{12} is in most cases closely related to φ_{12}, the phase of the main reflection $\mathbf{K}_1 + \mathbf{K}_2$.

It is very important for the success of direct methods that as much phase information as possible is used in the very beginning of the procedures, *i.e.* we need to use many reliable doublet or triplet phase sums. In this respect, our test results show that the new doublet distribution is expected to work well, and that doublets with the largest K values can be used with confidence for the phase determination of the satellite reflections.

In most cases, the prior incommensurate information taken into account, a small sinusoidal displacive modulation with a magnitude smaller than 0.4 Å (or 0.2 Å in the case of a fixed magnitude), did not lead to a large improvement of the phase sum statistics. Nevertheless, a large improvement was gained in the case where both the modulation magnitude and the modulation direction were fixed (structures A1a and A3a), showing the usefulness of the new methodology.

The conclusion from foregoing results (*e.g.* de Gelder *et al.*, 1996) that triplets involving satellites are less reliable than triplets consisting of three main reflections still holds in the sense that at high K levels the number of triplets with two satellite reflections is relatively small. Nevertheless, with the input of all phases calculated from the known basic structure, the doublet formula for satellite phases should be sufficiently powerful to guarantee the straightforward solution of the phase problem.

It is suggested for future developments of direct methods for incommensurately modulated structures to reprogram the computerized procedures by Lam *et al.* (1993, 1994, 1995) with implementation of the newly defined quasi-normalized structure factors [equations (40), (41), (42)].

APPENDIX *A*
The orientational average $\langle J_{-1}^2[2\pi(\mathbf{H}_\nu \cdot \mathbf{U}_j)]\rangle$

With respect to the orientational average of a function $A_{\rm orient}$ involving a triangle between the vectors **U** and **H** that can be oriented arbitrarily in space, two cases can be distinguished:

(i) Both $|\mathbf{H}|$ and $|\mathbf{U}|$ are known and can be kept fixed during the averaging process (Debye, 1915).

(ii) The $|\mathbf{H}|$ is known and fixed but the $|\mathbf{U}|$ is distributed (*e.g.* uniformly) in a certain interval $[0, \mathbf{U}_{\rm max}]$.

Case I: $|\mathbf{U}|$ is fixed.

Abbreviating $p = 2\pi|\mathbf{H}| \cdot |\mathbf{U}|$ and defining L to be a normalization factor, the orientational average can be expressed as

$$A_{\rm orient} = L^{-1} \int_0^{2\pi} \int_0^{\pi} J_{-1}^2[p\cos(\psi)]\sin(\psi)\,{\rm d}\psi\,{\rm d}\theta \qquad (46)$$

$$L = \int_0^{2\pi} \int_0^{\pi} \sin(\psi)\,{\rm d}\psi\,{\rm d}\theta = 4\pi. \qquad (47)$$

As a result, $A_{\rm orient}$ becomes

$$A_{\rm orient} = -\tfrac{1}{2}\int_0^{\pi} J_{-1}^2[p\cos(\psi)]\,{\rm dcos}(\psi) = \int_0^{1} J_{-1}^2(px)\,{\rm d}x. \qquad (48)$$

Expression (48) can be evaluated numerically.

Case II: $|\mathbf{U}|$ is uniformly distributed on $[0, U_m]$

The definition of the orientational integrals leads to a normalization factor L as follows:

$$L = \int_0^{U_m} \int_0^{2\pi} \int_0^{\pi} r^2 \sin(\theta)\,{\rm d}\theta\,{\rm d}\varphi\,{\rm d}r = \tfrac{4}{3}\pi U_m^3. \qquad (49)$$

Consequently, the normalized expression of $A_{\rm orient}$ becomes

$$A_{\rm orient} = 4\pi L^{-1} \int_0^{U_m} r^2 \int_0^1 J_{-1}^2(2\pi|\mathbf{H}|rx)\,{\rm d}x\,{\rm d}r. \qquad (50)$$

Again expression (50) can be evaluated numerically.

APPENDIX *B*
The orientational average $\langle J_{-1}[2\pi(\mathbf{H}_1\cdot\mathbf{U}_j)]J_1[2\pi(\mathbf{H}_2\cdot\mathbf{U}_j)]\rangle$

Orthogonal direct-cell spherical polar coordinates R, θ_r and ϕ_r with $0 \le R < \infty$, $0 \le \theta_r \le \pi$ and $0 \le \phi_r \le 2\pi$ can be related to the direct-cell coordinates X, Y and Z (all in Å) *via* the relations

$$\begin{aligned} X &= R\sin(\theta_r)\cos(\phi_r) \quad &\text{with } X = x\mathbf{a} \\ Y &= R\sin(\theta_r)\sin(\phi_r) \quad &\text{with } Y = y\mathbf{b} \\ Z &= R\cos(\theta_r) \quad &\text{with } Z = z\mathbf{c}. \end{aligned} \qquad (51)$$

Accordingly, a volume element ${\rm d}X\,{\rm d}Y\,{\rm d}Z = {\rm d}V$ is changed into ${\rm d}V = R^2\sin(\theta_r)\,{\rm d}r\,{\rm d}\theta_r\,{\rm d}\phi_r$ and a vector $\mathbf{R}$ in the direct cell is now written as

$$\mathbf{R} = R\sin(\theta_r)\cos(\phi_r)\mathbf{e}_1 + R\sin(\theta_r)\sin(\phi_r)\mathbf{e}_2 + R\cos(\phi_r)\mathbf{e}_3 \qquad (52)$$

with $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ unit length vectors along $\mathbf{a}$, $\mathbf{b}$ and $\mathbf{c}$, respectively. If the cell is not orthogonal, an orthogonalization should be carried out first.

In a similar way, spherical coordinates $Q = |\mathbf{H}|$, θ_q, ϕ_q, with $0 \le Q < \infty$, $0 \le \theta_q \le \pi$ and $0 \le \phi_q \le 2\pi$, in the orthogonal reciprocal system can be expressed in terms of the non-spherical coordinates H, K and L:

$$\begin{aligned} H &= Q\sin(\theta_q)\cos(\phi_q) \quad &\text{with } H = h\mathbf{a}^* \\ K &= Q\sin(\theta_q)\sin(\phi_q) \quad &\text{with } K = k\mathbf{b}^* \\ L &= Q\cos(\theta_q) \quad &\text{with } L = l\mathbf{c}^*. \end{aligned} \qquad (53)$$

Accordingly, a reciprocal vector $\mathbf{H}$ can be expressed as

$$\mathbf{H} = Q\sin(\theta_q)\cos(\phi_q)\mathbf{e}_1^* + Q\sin(\theta_q)\sin(\phi_q)\mathbf{e}_2^* + Q\cos(\phi_q)\mathbf{e}_3^* \qquad (54)$$

with $\mathbf{e}_1^*$, $\mathbf{e}_2^*$ and $\mathbf{e}_3^*$ unit length vectors along $\mathbf{a}^*$, $\mathbf{b}^*$ and $\mathbf{c}^*$, respectively.

The scalar product of $\mathbf{R}$ and $\mathbf{H}$ now becomes

$$\mathbf{H}\cdot\mathbf{R} = QR[\cos(\theta_q)\cos(\theta_r) + \sin(\theta_q)\sin(\theta_r)\cos(\phi_q - \phi_r)]. \qquad (55)$$

Both $\mathbf{H}_1\cdot\mathbf{U}_j$ and $\mathbf{H}_2\cdot\mathbf{U}_j$ can be expressed like (55). Let us introduce the abbreviations $Q_i = |\mathbf{H}_i|$ for $i = 1, 2$ and $R_j = |\mathbf{U}_j|$ so that

$$\mathbf{H}_i\cdot\mathbf{U}_j = Q_iR_j[\cos(\theta_{qi})\cos(\theta_r) + \sin(\theta_{qi})\sin(\theta_r)\cos(\phi_{qi} - \phi_r)], \quad i = 1, 2.$$

As a result, the orientational average reduces to a twofold integration that can be carried out numerically:

$$A_{\rm orient} = (4\pi)^{-1} \int_0^{\pi}\int_0^{2\pi} J_{-1}(2\pi\mathbf{H}_1\cdot\mathbf{U}_j)J_1(2\pi\mathbf{H}_2\cdot\mathbf{U}_j)\sin(\theta_r)\,{\rm d}\phi_r\,{\rm d}\theta_r.$$

References

Abramowitz, M. & Stegun, I. A. (1965). *Handbook of Mathematical Functions.* New York: Dover.
Cascarano, G. & Giacovazzo, C. (1988). *Acta Cryst.* A**44**, 176–183.
Cascarano, G., Giacovazzo, C. & Luic, M. (1987). *Acta Cryst.* A**43**, 14–22.
Cochran, W. (1955). *Acta Cryst.* **8**, 473–478.
Cochran, W. & Woolfson, M. M. (1955). *Acta Cryst.* **8**, 1–12.
Debye, P. (1915). *Ann. Phys. (Leipzing)*, **46**, 809–823.
Gelder, R. de, Israël, R., Lam, E. J. W., Beurskens, P. T., Smaalen, S. van, Fu, Z. Q. & Fan, H. F. (1996). *Acta Cryst.* A**52**, 947–954.
Janner, A. & Janssen, T. (1980). *Acta Cryst.* A**36**, 408–415.
Lam, E. J. W., Beurskens, P. T. & Smaalen, S. van (1992). *Solid State Commun.* **82**, 345–349.
Lam, E. J. W., Beurskens, P. T. & Smaalen, S. van (1993). *Acta Cryst.* A**49**, 709–721.
Lam, E. J. W., Beurskens, P. T. & Smaalen, S. van (1994). *Acta Cryst.* A**50**, 690–703.
Lam, E. J. W., Beurskens, P. T., Smits, J. M. M., Smaalen, S. van, Boer, J. L. de & Fan, H. F. (1995). *Acta Cryst.* B**51**, 779–789.
McConnell, J. D. C. & Heine, V. (1984). *Acta Cryst.* A**40**, 473–482.
Paciorek, W. A. & Kucharczyk, D. (1985). *Acta Cryst.* A**41**, 462–466.
Perez-Mato, J. M., Madariaga, G. & Tello, M. J. (1986). *J. Phys. C*, **19**, 2613–2622.
Peschar, R. & Schenk, H. (1991). *Acta Cryst.* A**47**, 428–440.
Petříček, V. & Coppens, P. (1988). *Acta Cryst.* A**44**, 235–249.
Petricek, V., Coppens, P. & Becker, P. (1985). *Acta Cryst.* A**41**, 478–483.
Steurer, W. (1987). *Acta Cryst.* A**43**, 36–42.
Yamamoto, A. (1982). *Acta Cryst.* A**38**, 87–92.