

matrices. Matrices provide an ideal mathematical notation for conceptualizing and analyzing symmetry. The mathematics and algorithms used to analyze symmetry become extremely simple since they are based on manipulating integers and simple rational numbers using elementary linear algebra. More importantly, the ability to abstract the symmetry from a problem and to represent it as a group of matrices leads to numerous practical and theoretical applications. In crystallography, for example, we have shown that the symmetry matrices may be used in practice to define the metric symmetry, the directions of the symmetry axes, the Laue symmetry, group-subgroup relationships, and conventional or standard cells. By providing the conceptual and practical framework required to perform experimental procedures in a logical and general manner, the matrix method should revolutionize the automation of diffractometers. Although evolving from research in lattices in crystallography, the matrix approach to symmetry is not limited to this discipline. Because of its fundamental nature, the matrix approach should provide the basis for further experimental and theoretical advances in symmetry and symmetry-related topics in crystallography as well as in chemistry, physics and mathematics.

#### Program availability

A Fortran program, *NBS\*LATTICE*, has been written to analyze lattice relationships and is available

*Acta Cryst.* (1987). **A43**, 384-393

## A Joint Probability Distribution of Invariants for all Space Groups

BY IAN R. CASTLEDEN

*Department of Physics, The University of Western Australia, Nedlands 6009, Australia*

(Received 15 May 1986; accepted 2 December 1986)

#### Abstract

A probability distribution is derived in the form of an infinite sum of cosine phase invariants between any number of structure factors and is applicable to any space group. No approximations are made. With the use of only the first-order terms, and the assumption that the reflection magnitudes are small and that all atoms are equal, the distributions reduce to the known invariant distribution forms. The effect of neighbourhood reflections is implicit in the form of the distribution.

#### 1. Introduction

Joint probability distributions (j.p.d.'s) have been derived by many authors, e.g. Hauptman & Karle

for distribution by the NBS Crystal Data Center. The present version of *NBS\*LATTICE* performs several functions including the determination of metric lattice symmetry, the identification of unknown materials using lattice-formula matching techniques, the calculation of the reduced cell of the lattice, and the calculation and reduction of specified derivative supercells and/or subcells.

The authors thank Professor H. Ammon of the Chemistry Department of the University of Maryland for use of the X-ray diffractometer which was purchased through NSF support, Grant CHE-84-02155.

#### References

- ANTON, H. (1984). *Elementary Linear Algebra*, 4th ed. New York: John Wiley.
- CALLIGARIS, M., NARDIN, G., RANDACCIO, L. & CHIARAMONTI, P. C. (1982). *Acta Cryst.* **B38**, 602-605.
- HIMES, V. L. & MIGHELL, A. D. (1982). *Acta Cryst.* **A38**, 748-749.
- HIMES, V. L., MIGHELL, A. D. & SIEDLE, A. R. (1981). *J. Am. Chem. Soc.* **103**, 211-212.
- International Tables for Crystallography* (1983). Vol. A, edited by T. HAHN. Dordrecht: Reidel.
- JONES, P. G., SHELDRIK, G. M. & HORN, A. S. (1981). *Acta Cryst.* **B37**, 906-910.
- MORRIS, M. C., MCMURDIE, H. F., EVANS, E. H., PARETZKIN, B., HUBBARD, C. R. & CARMEL, S. J. (1980). *Natl Bur. Stand. (US) Monogr.* **25**, Section 17.
- SANTORO, A., MIGHELL, A. D. & RODGERS, J. R. (1980). *Acta Cryst.* **A36**, 796-800.

(1953), Klug (1958), Naya, Nitta & Oda (1965), Tsoucaris (1970), Hauptman (1974, 1975), Fortier & Hauptman (1977), Heinerman, Krabbendam & Kroon (1979), Giacovazzo (1974, 1975, 1976) and more recently Shmueli & Weiss (1985). Each of these treatments is applicable to a particular set of phases or a particular space group. The distribution described in this paper is general and is intended to serve as the starting point for the derivation of a specific formula, rather than deriving the distributions for each case of interest. The expressions presented have the advantage of separating the notational complexity of the derivation of distributions into two distinct areas. The first is the calculation of integral coefficients. The second is the search for a set of integers that satisfy the phase-invariant relationships.

This makes the contributions by neighbourhood reflections explicit.

Although the j.p.d. expression derived is exact, it will probably find most use in deriving approximate space-group-dependent (sem)invariant relationships. This involves calculating lowest-order terms and then appealing to the central limit theorem. The derivation in § 3 resembles that used by Heinerman, Krabben-dam & Kroon (1977) and Shmueli & Weiss (1985). A similar treatment of the triplet j.p.d. in *P1* for the equal-atom case has recently been published by Peschar & Schenk (1986).

## 2. The probability distribution

This section describes the probability distribution to be derived in § 3. The joint probability density distribution between a set of  $M$  structure factors  $F(\mathbf{h}) = F_h \exp(i\theta_h)$ ,  $\mathbf{h} = \{h_1, h_2, \dots, h_M\}$ , with the assumption that all  $N$  atoms are independently uniformly distributed, is given by

$$\begin{aligned} P(F_{h_1}, F_{h_2}, \dots, F_{h_M}; \theta_{h_1}, \theta_{h_2}, \dots, \theta_{h_M}) & \left( \prod_{i=1}^M dF_{h_i} d\theta_{h_i} \right) \\ & = \left( \prod_{i=1}^M F_{h_i} dF_{h_i} d\theta_{h_i} / 2\pi \right) \\ & \times \sum_{\gamma_1} \sum_{\gamma_2} \dots \sum_{\gamma_{N/G}} \left\{ \prod_{i=1}^M G[\mathbf{n}, \mathbf{f}(h_i), F_{h_i}] \right\} \\ & \times \cos \left[ \sum_{i=1}^M \sum_{g=1}^G \sum_{\mu=1}^{N/G} n(\mu, g, h_i) (2\pi \mathbf{h}_i^T \cdot \mathbf{t}_g - \theta_{h_i}) \right] \end{aligned} \quad (1)$$

where  $\gamma_\mu$ ,  $\mu = 1, \dots, N/G$ , are the sets of integers  $n(\mu, g, h)$  that satisfy the constraint

$$\sum_{g=1}^G \sum_{i=1}^M n(\mu, g, h_i) \mathbf{h}_i^T \cdot \mathbf{R}_g = 0 \quad \{\mu = 1, 2, \dots, N/G\} \quad (2)$$

for each  $\mu$ . The origin of these integer constraints can be compared with the mathematical origin of the summation conditions given by Bricogne [1984; see equation (3.19)].  $\mathbf{R}_g$ ,  $\mathbf{t}_g$  are the symmetry rotation and translation matrices of the space group. If the  $F(\mathbf{h})$ 's are to be treated as normalized structure factors  $E(\mathbf{h})$ , then  $F_h$  must be divided by the normalizing factor

$$\left[ G \sum_{\mu=1}^{N/G} f_\mu^2(\mathbf{h}) \right]^{1/2} \quad (3)$$

$G[\mathbf{n}, \mathbf{f}, F]$  is an integral over  $N+1$  Bessel functions of the form

$$\begin{aligned} G[\mathbf{n}, \mathbf{f}(h), F_h] & = \int_0^\infty \left[ \prod_{g=1}^G J_{n(1,g,h)}[f_1(h)\rho] J_{n(2,g,h)}[f_2(h)\rho] \times \dots \right. \\ & \left. \times J_{n(N/G,g,h)}[f_{N/G}(h)\rho] \right] J_\Sigma(F_h\rho) \rho \, d\rho \end{aligned}$$

where

$$\Sigma = \sum_{g=1}^G \sum_{\mu=1}^{N/G} n(\mu, g, h). \quad (4)$$

$N$  is the total number of atoms in the unit cell (none in special positions),  $f_\mu(h)$  is the scattering factor of atom  $\mu$  divided by the normalizing factor and  $G$  is the number of symmetry elements. Hence the number of independent atoms is  $N/G$ .  $J_n(z)$  are Bessel functions of integer order  $n$  and  $n(\mu, g, h)$  is an integer function of  $\mu, h, g$  with a domain of  $\mu = 1, \dots, N/G$ ,  $g = 1, \dots, G$ , and  $h$  is a set of  $M$  reciprocal-space vectors (so there are  $N \times M$  different integer values of  $n$  not all independent). These integers are grouped into  $N/G$  sets, each being labelled by the argument  $\mu$ . Each set contains  $M \times G$  integers which must satisfy the relation (2) of which at least the trivial set exists:  $\{n(\mu, g, h_i) = 0, g = 1, \dots, G, i = 1, \dots, M\}$ . Sets of this type will be denoted by  $\{0\}$ . The subscript of the last Bessel function is a summation of the integers  $n(\mu, g, h)$  over all  $g$  and  $\mu$ , i.e. over each atom in the unit cell.

A more concise notation for the integers  $n(\mu, g, h)$  has been devised by Bricogne (1984) but the resulting compression in notation has not been thought necessary here. In this paper,  $\mathbf{h}$ ,  $\mathbf{k}$ ,  $\mathbf{l}$  and  $\mathbf{q}$  are reciprocal-lattice vectors,  $\mu$  is an integer that identifies the unique atoms and  $g$  is an integer that labels the space-group symmetry elements. The abbreviations AS and GR used below refer to Abramowitz & Stegun (1972) and Gradshteyn & Ryzhik (1980) respectively. References to GR indicate that a particular integral identity has been used.

## 3. The derivation of the probability distribution

This derivation differs from previous j.p.d. proofs [except for Shmueli & Weiss (1985) and Peschar & Schenk (1986)] in that the characteristic function is not converted into a truncated power series before the inverse Fourier transform is performed. The structure factors  $F(\mathbf{h})$  as a function of atomic position are given by

$$F(\mathbf{h}) = \sum_{g=1}^G \sum_{\mu=1}^{N/G} f_\mu(\mathbf{h}) \exp[2\pi i \mathbf{h}^T \cdot (\mathbf{R}_g \mathbf{r}_\mu + \mathbf{t}_g)], \quad (5)$$

where  $\mathbf{r}_\mu$  is the position of the  $\mu$ th atom and is assumed to be a random variable uniformly distributed over the unit cell. Since  $F(\mathbf{h}) = A_h + iB_h$ , the probability density distribution of  $A_h$  and  $B_h$  is given by

$$\begin{aligned} P(A_{h_1}, B_{h_1}, A_{h_2}, B_{h_2}, \dots, A_{h_M}, B_{h_M}) & = (2\pi)^{-2M} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \exp \left[ -i \sum_{i=1}^M (A_{h_i} x_{h_i} + B_{h_i} y_{h_i}) \right] \\ & \times Q(x_{h_1}, y_{h_1}, x_{h_2}, y_{h_2}, \dots, x_{h_M}, y_{h_M}) \\ & \times dx_{h_1} dy_{h_1} \dots dy_{h_M}, \end{aligned} \quad (6)$$

where  $Q$  is by definition the characteristic function of  $P$ . Since  $A_h$  and  $B_h$  are functions of the random variables  $r_\mu$  [from (5)], the characteristic function can be expressed as an expectation value over the probability distribution  $p(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N/G})$  of atom 1 being at  $\mathbf{r}_1$ , atom 2 being at  $\mathbf{r}_2$  etc:

$$Q(x_{h_1}, y_{h_1}, \dots, x_{h_M}, y_{h_M}) = \int_V \int_V \dots \int_V \exp \left[ i \sum_{i=1}^M (A_{h_i} x_{h_i} + B_{h_i} y_{h_i}) \right] \times p(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N/G}) d^3 \mathbf{r}_1 d^3 \mathbf{r}_2 \dots d^3 \mathbf{r}_{N/G} \quad (7)$$

where the integrals are over the unit cell. In polar coordinates,  $x_h = \rho_h \cos \xi_h$ ,  $y_h = \rho_h \sin \xi_h$ ,  $A_h = F_h \cos \theta_h$ ,  $B_h = F_h \sin \theta_h$ , and (6) becomes (Karle & Hauptman 1958)

$$P(F_{h_1}, F_{h_2}, \dots, F_{h_M}; \theta_{h_1}, \theta_{h_2}, \dots, \theta_{h_M}) \prod_{i=1}^M dA_{h_i} dB_{h_i} = \left[ \prod_{i=1}^M F_{h_i} dF_{h_i} d\theta_{h_i} / (2\pi)^2 \right] \times \int_0^{2\pi} \int_0^{2\pi} \dots \int_0^{2\pi} \exp \left[ -i \sum_{i=1}^M F_{h_i} \rho_{h_i} \cos(\xi_{h_i} - \theta_{h_i}) \right] \times Q(\rho_{h_1}, \rho_{h_2}, \dots, \rho_{h_M}; \xi_{h_1}, \xi_{h_2}, \dots, \xi_{h_M}) \times \rho_{h_1} \rho_{h_2} \dots \rho_{h_M} d\rho_{h_1} d\rho_{h_2} \dots d\rho_{h_M} \times d\xi_{h_1} d\xi_{h_2} \dots d\xi_{h_M}. \quad (8)$$

From (5) the argument in the exponential of (7) becomes

$$i \sum_{j=1}^M \rho_{h_j} \sum_{g=1}^G \sum_{\mu=1}^{N/G} f_\mu(h_j) \cos [2\pi \mathbf{h}_j^T \cdot (\mathbf{R}_g \mathbf{r}_\mu + \mathbf{t}_g) - \xi_{h_j}]. \quad (9)$$

The summations can be removed from the argument of the exponential and expressed as the product

$$\prod_{j=1}^M \prod_{g=1}^G \prod_{\mu=1}^{N/G} \exp \{ i \rho_{h_j} f_\mu(h_j) \times \cos [2\pi \mathbf{h}_j^T \cdot (\mathbf{R}_g \mathbf{r}_\mu + \mathbf{t}_g) - \xi_{h_j}] \}. \quad (10)$$

From the identity (AS 9.1.45)

$$\exp(iz \cos \beta) = \sum_{n=-\infty}^{\infty} i^n J_n(z) \cos(n\beta) \quad (11)$$

and the relationships  $J_{-n}(z) = (-1)^n J_n(z)$ ,  $i^{-n} = (-1)^n i^n$  it follows that

$$\exp(iz \cos \beta) = \sum_{n=-\infty}^{\infty} J_n(z) \exp[in(\beta + \pi/2)]. \quad (12)$$

Substitution of this in relation (10) yields

$$\prod_{j=1}^M \prod_{g=1}^G \prod_{\mu=1}^{N/G} \left( \sum_{n=-\infty}^{\infty} J_n[f_\mu(h_j) \rho] \times \exp \{ in [2\pi \mathbf{h}_j^T \cdot (\mathbf{R}_g \mathbf{r}_\mu + \mathbf{t}_g) - \xi_{h_j} + \pi/2] \} \right). \quad (13)$$

The products can be taken inside the summation giving an  $M \times N$ -tuple sum over the integers  $n(\mu, g, h)$  of the form

$$\sum_{n(\mu=1, g=1, h=1)=-\infty}^{\infty} \dots \sum \dots \sum_{n(\mu=N/G, g=G, h=h_M)=-\infty}^{\infty} \times \left\{ \prod_{\mu=1}^{N/G} \prod_{g=1}^G \prod_{i=1}^M J_{n(\mu, g, h_i)} [f_\mu(h_i) \rho_{h_i}] \right\} \times \exp \left\{ i \sum_{\mu=1}^{N/G} \sum_{g=1}^G \sum_{i=1}^M n(\mu, g, h_i) \times [2\pi \mathbf{h}_i^T \cdot (\mathbf{R}_g \mathbf{r}_\mu + \mathbf{t}_g) - \xi_{h_i} + \pi/2] \right\}. \quad (14)$$

To calculate the characteristic function, this summation must be multiplied by the atomic position probability distribution  $p(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N/G})$  and integrated over the unit cell. The assumption of independent uniformly distributed atoms requires that  $p(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_{N/G}) = (1/V)^{N/G}$ . It follows that all terms in the above summation vanish when integrated unless the set of integers  $n(\mu, g, h)$  satisfies the relationship

$$\sum_{g=1}^G \sum_{i=1}^M n(\mu, g, h_i) \mathbf{h}_i^T \cdot \mathbf{R}_g = 0 \quad \{\mu = 1, 2, \dots, N/G\}. \quad (15)$$

The inverse Fourier transform of the characteristic function  $Q(\rho_h, \dots, \xi_h)$  can be performed by concentrating on each reflection  $h$  in turn. Collection of all terms containing  $\xi_h$  in (8) yields

$$\exp \left\{ i \sum_{g=1}^G \sum_{\mu=1}^{N/G} [n(\mu, g, h) (2\pi \mathbf{h} \cdot \mathbf{t}_g - \xi_h)] \right\} \times \exp [-i F_h \rho_h \cos(\xi_h - \theta_h)]. \quad (16)$$

Equivalent terms due to other reflections  $\mathbf{h}_i \neq \mathbf{h}$  multiply this expression but can be ignored for the purposes of integration. The relation (AS 9.1.21)

$$J_n(z) = (i^n / 2\pi) \exp(-in\xi) \times \int_{\delta}^{2\pi+\delta} \exp \{ i [n\beta - z \cos(\beta - \xi)] \} d\beta \quad (17)$$

can be used to integrate this function over  $2\pi$ . The integration with respect to  $\xi$  yields

$$2\pi J_\Sigma(F_h \rho) \exp \left[ -i \sum_{\mu, g} n(\mu, g, h) (\theta_h - \pi/2) \right], \quad (18)$$

where

$$\Sigma = \sum_{\mu, g} n(\mu, g, h)$$

and the final integration over  $\rho_h$  gives

$$(2\pi)^{-2M} \sum_{\gamma_1, \gamma_2, \dots, \gamma_{N/G}} \times \left[ \prod_{i=1}^M \left( \int_0^\infty \left\{ \prod_{g=1}^G \prod_{\mu=1}^{N/G} J_{n(\mu, g, h_i)} [f_\mu(h_i) \rho] \right\} J_{\Sigma n}(F_h \rho) \rho \, d\rho \right) \right. \\ \left. \times \exp \left( i \sum_{j=1}^M \left\{ \sum_{g=1}^G \sum_{\mu=1}^M [n(\mu, g, h_j) (2\pi \mathbf{h}_j \cdot \mathbf{t}_g - \theta_{h_i})] \right\} \right) \right] \quad (19)$$

where

$$\Sigma \mathbf{n} = \sum_{g=1}^G \sum_{\mu=1}^{N/G} n(\mu, g, h_i). \quad (20)$$

Provided (15) is true for the integer function  $n(\mu, g, h)$  then it is also true for  $-n(\mu, g, h)$  and since the integrals  $G(n, \mathbf{f}, F_h)$  are invariant under this inversion the summation reduces to (1).

#### 4. Discussion on the probability distribution

It is not obvious that this probability distribution, although exact, is usable in any form because of the multiple summations over invariants ( $N/G$  of them). A typical procedure for calculating a probability distribution from (1) would be:

(i) Remove all unwanted phases by integration. Since the expansion is a sum of cosine functions, all terms that contain the unwanted phases will vanish  $[\int_0^{2\pi} \cos(\xi - \alpha) \, d\xi = 0]$ .

(ii) Find the set of integers that satisfy

$$\sum_{g=1}^G \sum_{i=1}^M n(\mu, g, h_i) \mathbf{h}_i^T \cdot \mathbf{R}_g = 0 \\ \{\mu = 1, 2, \dots, N/G\}. \quad (21)$$

This is a straightforward but tedious operation that may be automated using a computer algebraic package.

(iii) Evaluate the integrals  $G(n, \mathbf{f}, F_h)$ .

##### 4.1. Evaluating the integrals $G(n, \mathbf{f}, F_h)$

The integrals  $G$  are the coefficients of the cosine terms. No linearization formulae for Bessel functions exist as in, for example, the Laguerre polynomials. The largest contributions will come from the integrals that contain the greatest number of zero-order Bessel functions since the integrand will then be highly peaked at the origin. Exact Fourier-Bessel series exist for these integrals (Barakat & Cole 1979; Watson, 1942) but here only simple series approximation will be needed. As expected, integrals of the form of equation (4) are identically zero if [GR 6.573(1)]

$$F_h > G \sum_{\mu=1}^{N/G} f_\mu(h). \quad (22)$$

In the more interesting case of physically obtainable reflection magnitudes integration of the functions  $G$  may be carried out if an approximation to the zero-order Bessel function is made [see, for example, Giacobazzo (1980)]. If one collects Bessel functions that are of zero order, the approximation is

$$\prod_{g=1}^G \prod_{\mu=1}^{N/G} J_0[f_\mu(h) \rho] \approx \exp \left[ -(G\rho^2/4) \sum_{\mu=1}^{N/G} f_\mu^2(h) \right]. \quad (23)$$

The integrals necessary for describing phase invariants and seminvariants with their first neighbourhoods are the following:

Case 0:

$$n(\mu, g, h) = 0 \quad \text{for all } \mu, g, h.$$

Case 1:

$$n(\mu_1, g_1, h) = \pm 1 \quad \text{for some } g_1, \mu_1,$$

all others zero.

Case 2:

$$n(\mu_1, g_1, h) = -n(\mu_2, g_2, h) = \pm 1$$

for some  $g_1, \mu_1, g_2, \mu_2$ , all others zero.

Although this is a very restricted set of integers, a later example will show that this is all that is needed to reproduce the well established triplet and quartet relationships. Note that in case 1 there are  $N/G$  different sets of integers corresponding to  $\mu$  ranging over the  $N/G$  unique atoms. For case 2 there are  $N/G$  choose 2\* different integer sets corresponding to the  $N/G$  choose 2 pairs of atoms  $\mu_1 \mu_2$ . Hence the procedure effectively searches for invariant relationships given by the integers and then sums the contributions over the atoms or pairs of atoms.

With the approximation for case 0 above, the integrals (4) become

$$\int_0^\infty \exp \left[ -G \sum_{\mu=1}^{N/G} f_\mu^2(h) \rho^2/4 \right] J_0(F_h \rho) \rho \, d\rho \quad (24)$$

and integrate to [GR 6.631(4)]

$$(2/\alpha_0) \exp(-F_h^2/\alpha_0) \quad \text{where } \alpha_0 = G \sum_{\mu=1}^{N/G} f_\mu^2(h). \quad (25)$$

The approximation in (23) is good for large  $N$ . A more detailed examination of (23) shows that the left-hand side can be expressed as the right-hand side multiplied by a power series in  $\rho^2$ . The coefficients of the power series  $a_n$  depend on the number of atoms as  $(N^{-1/2})^n$ ,  $n = 0, 1, 2, \dots$ . Hence (25) is correct to order  $N^{-1/2}$ .

\* The notation  $N$  choose  $n$  refers to the combinatorial product  $N!/[n!(N-n)!]$  equal to the number of ways to form unique sets of  $n$  elements from a group of  $N$  elements.

The integrals for case 1 are of the form

$$\int_0^{\infty} \exp \left[ - \left( G \sum_{\mu=1}^{N/G} f_{\mu}^2(h) - f_i^2 \right) \rho^2 / 4 \right] \times J_{\pm 1}[f_i(h)\rho] J_{\pm 1}(F_h \rho) \rho \, d\rho \quad (26)$$

where the summation in the exponential argument is over all  $N/G$  scattering factors. This integrates to [GR 6.633(2)]

$$(2/\alpha_1) I_1(2f_i F_h / \alpha_1) \exp [-(f_i^2 + F_h^2) / \alpha_1], \quad (27)$$

where  $I_1$  is a first-order modified Bessel function and

$$\alpha_1 = G \left[ \sum_{\mu=1}^{N/G} f_{\mu}^2(h) \right] - f_i^2(h). \quad (28)$$

The integral for case 2 has the form

$$\int_0^{\infty} \exp (-\alpha_2 \rho^2 / 4) J_{+1}[f_i(h)\rho] J_{-1}[f_j(h)\rho] J_0(F_h \rho) \rho \, d\rho \quad (29)$$

or, more generally,

$$G_{nm} = \int_0^{\infty} \exp (-\alpha_2 \rho^2 / 4) \times J_n[f_i(h)\rho] J_m[f_j(h)\rho] J_{n-m}(F_h \rho) \rho \, d\rho, \quad (30)$$

where

$$\alpha_2 = G \left[ \sum_{\mu=1}^{N/G} f_{\mu}^2(h) \right] - [f_i^2(h) + f_j^2(h)] \quad (31)$$

and the sum extends over all  $N/G$  atomic scattering factors. If one expands the zero-order Bessel function as a power series and uses the integral relation GR 6.631(1), this reduces to

$$\begin{aligned} & (-\alpha_2)^m (2/\alpha_2) [f_i(h)/\alpha_2]^n [f_j(h)/\alpha_2]^m \\ & \times [F_h^{n-m}/(n-m)!] \exp(-F_h^2/\alpha_2) \\ & \times \sum_{k=1}^{\infty} \sum_{l=1}^{\infty} \{ (-1)^{k+l} [(f_i/f_j)^2 - 1]^k \\ & \times (n+m+2l)! \Gamma(n+l+1+k/2) \} \\ & \times [k! l! (n+1)! (m+n+1)! (m+1)!]^{-1} \\ & \times (f_2 \alpha_2^{-1/2})^{2l+k} \\ & \times {}_1F_1(-l-m-k/2, n-m+1; F_h^2/\alpha_2). \quad (32) \end{aligned}$$

${}_1F_1$  is a hypergeometric function. If  $\alpha$  is much larger than either  $f_i$  or  $f_j$  then only the  $l=k=0$  term is important and the summation for  $n=m=l$  reduces to

$$-2(f_i f_j / \alpha_2^2) \exp(-F_h^2/\alpha_2) (1 - F_h^2/\alpha_2). \quad (33)$$

Integrals with three or more non-zero-order Bessel functions can be calculated by replacing the Bessel functions with their series expansions. The relative contribution of such terms to the summation can be estimated if the structure factors  $F_h$  are taken to be

normalized  $E$  values. In this case  $f_i$  and  $f_j \sim N^{-1/2}$ . There are at most  $N^2$  integrals of the form of (29) and since the largest coefficients of (1) are of order  $N^{-1/2}$  (see example below) they give, in total, a relative contribution of order  $N^{[5-3(m+n)]/2}$ .

## 5. Application of the probability distribution

The  $N/G$  summation conditions (21) complicate the determination of the set of contributing terms. However, terms with the largest contribution to a particular invariant can easily be enumerated. These terms are equal to the argument in the exponential of the equivalent invariant distribution derived by the Gram-Charlier method since lowest-order terms in either series expansion must be equal (Heinerman *et al.*, 1977). To illustrate the effect of neighbourhood reflections, first consider the quartet distribution in  $P1$ . The effect of space-group symmetry will be illustrated by the triplet and seminvariant relations in  $P2_1$ .

### 5.1. Example 1: Quartet in $P1$

Only terms for which cosine arguments have phase values  $\theta_h, \theta_k$  etc. multiplied by  $\pm 1$  or 0 will be included. This is a subset of the set of integers satisfying  $n(\mu, g, h) = \{-1, 0, +1\}$ . All expressions containing  $n(\mu, g, h)$  will assume this restriction. In  $P1$  the only space-group element is the identity, hence  $G=1$  and  $\mathbf{R}=\mathbf{I}$ . This means that the  $N/G=N$  conditions on the integers  $n(\mu, g, h)$  reduce to

$$\sum_{i=1}^M n(\mu, 1, h_i) \mathbf{h}_i = \mathbf{0} \quad \{\mu = 1, 2, \dots, N/G\}. \quad (34)$$

Since there is only the single group element the argument  $g$  will be dropped.

Let  $\mathbf{h}, \mathbf{k}, \mathbf{l}, \mathbf{m}$  be a basis set of reflections such that  $\mathbf{h} + \mathbf{k} + \mathbf{l} + \mathbf{m} = \mathbf{0}$ . To derive a quartet relationship the terms containing phases  $\theta$  other than  $\theta_h, \theta_k, \theta_l, \theta_m$  must be integrated out. If this is done all such cosine terms will vanish. These terms are of the form

$$\cos \left\{ \sum_{i=1}^4 \left[ \sum_{\mu=1}^N n(\mu, h_i) \right] \theta_{h_i} \right\}. \quad (35)$$

Thus, the only terms which will survive are those for which

$$\sum_{\mu=1}^N n(\mu, h_i) = 0 \quad (36)$$

with  $\mathbf{h}_i$  not a basis reflection ( $h_i \neq \{h, k, l, m\}$ ).

As a first approximation let the integers  $n(\mu, q) = 0$  for  $\mu = 1, \dots, N$  and  $\mathbf{q}$  not a basis reflection  $\{\mathbf{h}, \mathbf{k}, \mathbf{l}, \mathbf{m}\}$ . This yields the phase relationships due solely to the basis reflections. The first neighbourhood appears when considering terms where  $n(\mu_1, q) = -n(\mu_2, q)$  for some  $\mu_1$  and  $\mu_2$  and all others are zero.

If this condition holds,  $\theta_q$  does not appear in the cosine argument and these terms will not vanish on integration. However,

$$n(\mu_1, h)h + n(\mu_1, k)k + n(\mu_1, l)l + n(\mu_1, m)m + n(\mu_1, q)q = 0$$

is true (similarly for  $\mu_2$ ) and hence  $\mathbf{q}$  must be a cross vector. In other words, the existence of cross vectors for the quartet allows the condition (36) to be true for cases other than  $\{n(\mu, q) = 0, \mu = 1, \dots, N\}$ .

Consider the set of integers  $n_i$  that satisfy

$$\sum_{i=1}^7 n_i h_i = 0 \tag{37}$$

for  $h_i = \{h, k, l, m = -h - k - l, h + k, h + l, k + l\}$ . Solutions for this equation are given in Table 1.

For each value of the integer  $\mu$ , one of the 14 rows can be selected although, in keeping with the above restrictions, the vast majority of the  $N/G$  sets  $\{n(\mu, q), q = h, k, \dots\}_\mu$  must contain only zero values. That is, line zero (set  $\{0\}$ ) is chosen from Table 1. In fact, at most two of the  $N$  sets can contain non-zero integers.

Three types of integer sets can be distinguished.

Set 0. The set  $n(\mu, q) = 0$  for all  $\mu, q$ : the zeroth-order term.

Table 1. The set of integers that satisfy equation (37) ( $n = +1$  or  $-1$ )

| Line | $h$  | $k$  | $l$  | $m = -h - k - l$ | $h + k$ | $h + l$ | $k + l$ |
|------|------|------|------|------------------|---------|---------|---------|
| 0    | 0    | 0    | 0    | 0                | 0       | 0       | 0       |
| 1    | $n$  | $n$  | 0    | 0                | 0       | $-n$    | 0       |
| 2    | $n$  | 0    | $n$  | 0                | 0       | 0       | $-n$    |
| 3    | 0    | $n$  | $n$  | 0                | 0       | 0       | $-n$    |
| 4    | $n$  | 0    | 0    | $n$              | 0       | 0       | $n$     |
| 5    | 0    | $n$  | 0    | $n$              | 0       | $n$     | 0       |
| 6    | 0    | 0    | $n$  | $n$              | $n$     | 0       | 0       |
| 7    | $-n$ | 0    | 0    | $n$              | $n$     | $n$     | 0       |
| 8    | 0    | $-n$ | 0    | $n$              | $n$     | 0       | $n$     |
| 9    | 0    | 0    | $-n$ | $n$              | 0       | $n$     | $n$     |
| 10   | 0    | $-n$ | $n$  | 0                | $n$     | $-n$    | 0       |
| 11   | $n$  | 0    | $-n$ | 0                | $-n$    | 0       | $n$     |
| 12   | $n$  | $-n$ | 0    | 0                | 0       | $-n$    | $n$     |
| 13   | $n$  | $n$  | $n$  | $n$              | 0       | 0       | 0       |

Set 1. The set where  $\{n(\mu, q)\}_\mu \neq \{0\}$  for one  $\mu$  (of which there are  $2N$ :  $N$  with  $n = +1$  and  $N$  with  $n = -1$ ).

Set 2. The set where  $\{n(\mu, q)\}_\mu \neq \{0\}$  and  $\{n(\eta, q)\}_\eta \neq \{0\}$  for some  $\mu, \eta$  such that  $\mu \neq \eta$ . [There are  $2(N \text{ choose } 2)$  terms of this type.]

For clarification in this example, one more simplification is made. All atoms are assumed equal so that  $f_\mu(h) = f$  for all atoms  $\mu$ . This results in the integrals  $G(n, \mathbf{f}, F_h)$  being invariant under  $\mu$ . Note that all  $2N$  integrals in set 0 above are equal. Let  $G_0(h), G_1(h), G_2(h)$  denote the integrals generated from the integer sets 0, 1, 2 respectively. Then under the restrictions made above the joint probability distribution for the quartet + cross vectors is

$$\begin{aligned}
 &P(F_h, F_k, F_l, F_{h+k+l}, F_{h+k}, F_{h+l}, F_{k+l}; \theta_h, \theta_k, \theta_l, \theta_{h+k+l}, \theta_{h+k}, \theta_{h+l}, \theta_{k+l}) \\
 &= \left[ \prod_q G_0(\mathbf{q}) \right] \tag{line 0} \\
 &+ 2N \left[ \prod_{q \neq h, k, h+k} G_0(\mathbf{q}) \right] G_1(h) G_1(k) G_1(h+k) \cos(\theta_h + \theta_k - \theta_{h+k}) \tag{line 1} \\
 &+ 2N \left[ \prod_{q \neq h, l, h+l} G_0(\mathbf{q}) \right] G_1(h) G_1(l) G_1(h+l) \cos(\theta_h + \theta_l - \theta_{h+l}) \tag{line 2} \\
 &+ 2N \left[ \prod_{q \neq k, l, k+l} G_0(\mathbf{q}) \right] G_1(k) G_1(l) G_1(k+l) \cos(\theta_k + \theta_l - \theta_{k+l}) \tag{line 3} \\
 &+ 2N \left[ \prod_{q \neq h, -h-k-l, k+l} G_0(\mathbf{q}) \right] G_1(h) G_1(-h-k-l) G_1(k+l) \cos(\theta_h + \theta_{-h-k-l} + \theta_{k+l}) \tag{line 4} \\
 &+ 2N \left[ \prod_{q \neq k, -h-k-l, h+l} G_0(\mathbf{q}) \right] G_1(k) G_1(-h-k-l) G_1(h+l) \cos(\theta_k + \theta_{-h-k-l} + \theta_{h+l}) \tag{line 5} \\
 &+ 2N \left[ \prod_{q \neq l, -h-k-l, h+k} G_0(\mathbf{q}) \right] G_1(l) G_1(-h-k-l) G_1(h+k) \cos(\theta_l + \theta_{-h-k-l} + \theta_{h+k}) \tag{line 6} \\
 &+ 2N \left[ \prod_{q \neq h, -h-k-l, h+k, h+l} G_0(\mathbf{q}) \right] G_1(h) G_1(-h-k-l) G_1(h+k) G_1(h+l) \cos(-\theta_h + \theta_{-h-k-l} + \theta_{h+k} + \theta_{h+l}) \tag{line 7}
 \end{aligned}$$

$$\begin{aligned}
& +2N \left[ \prod_{q \neq k, -h-k-l, h+k, k+l} G_0(\mathbf{q}) \right] G_1(k) G_1(-h-k-l) G_1(h+k) G_1(k+l) \cos(-\theta_k + \theta_{-h-k-l} + \theta_{h+k} + \theta_{k+l}) && \text{line 8} \\
& +2N \left[ \prod_{q \neq l, -h-k-l, h+l, k+l} G_0(\mathbf{q}) \right] G_1(l) G_1(-h-k-l) G_1(h+l) G_1(k+l) \cos(-\theta_l + \theta_{-h-k-l} + \theta_{h+l} + \theta_{k+l}) && \text{line 9} \\
& +2N \left[ \prod_{q \neq k, l, h+l, k+l} G_0(\mathbf{q}) \right] G_1(k) G_1(l) G_1(h+k) G_1(k+l) \cos(-\theta_k + \theta_l + \theta_{h+k} - \theta_{k+l}) && \text{line 10} \\
& +2N \left[ \prod_{q \neq h, l, h+k, k+l} G_0(\mathbf{q}) \right] G_1(h) G_1(l) G_1(h+k) G_1(k+l) \cos(+\theta_h - \theta_l - \theta_{h+k} + \theta_{k+l}) && \text{line 11} \\
& +2N \left[ \prod_{q \neq h, l, h+k, k+l} G_0(\mathbf{q}) \right] G_1(h) G_1(k) G_1(h+l) G_1(k+l) \cos(+\theta_h + \theta_k - \theta_{h+l} + \theta_{k+l}) && \text{line 12} \\
& +2N \left[ \prod_{q \neq h, k, l, -h-k-l} G_0(\mathbf{q}) \right] G_1(h) G_1(k) G_1(l) G_1(-h-k-l) \cos(+\theta_h + \theta_k + \theta_l + \theta_{-h-k-l}) && \text{line 13} \\
& +2 \binom{N}{2} \left[ \prod_{q \neq h, k, l, -h-k-l, h+k} G_0(\mathbf{q}) \right] G_1(h) G_1(k) G_1(l) G_1(-h-k-l) G_2(h+k) \cos(\theta_h + \theta_k + \theta_l + \theta_{-h-k-l}) && \text{lines 1+6} \\
& +2 \binom{N}{2} \left[ \prod_{q \neq h, k, l, -h-k-l, h+l} G_0(\mathbf{q}) \right] G_1(h) G_1(k) G_1(l) G_1(-h-k-l) G_2(h+l) \cos(\theta_h + \theta_k + \theta_l + \theta_{-h-k-l}) && \text{lines 2+5} \\
& +2 \binom{N}{2} \left[ \prod_{q \neq h, k, l, -h-k-l, k+l} G_0(\mathbf{q}) \right] G_1(h) G_1(k) G_1(l) G_1(-h-k-l) G_2(k+l) \cos(\theta_h + \theta_k + \theta_l + \theta_{-h-k-l}) && \text{lines 3+4} \quad (38)
\end{aligned}$$

plus terms involving two cross-vector phases. The expression

$$\binom{N}{2} \text{ means } N \text{ choose } 2 = N!/[2!(N-2)!].$$

The line numbers refer to Table 1 and indicate which set of integers were used to generate the term. The products appearing in each term are over those reflections  $\mathbf{q}$  not explicitly appearing as arguments in  $G_1$  or  $G_2$  functions for that term.

5.1.1. *Comparisons with other quartet relations.* Denote by  $\Delta_i(h)$  the ratio  $G_i(h)/G_0(h)$ . Then if we divide by the zeroth-order term we can remove the products  $\prod_q G_0(q)$ . This essentially removes the Wilson term effect from the arguments. By integrating over the cross-vector phases, all terms involving these phases vanish (including those not explicitly written), to leave

$$\begin{aligned}
& \prod_q G_0(\mathbf{q}) \left[ 1 + 2N \Delta_1(h) \Delta_1(k) \Delta_1(l) \Delta_1(-h-k-l) \right. \\
& \quad \times \cos(\theta_h + \theta_k + \theta_l + \theta_{-h-k-l}) && \text{line 13}
\end{aligned}$$

$$\begin{aligned}
& +2 \binom{N}{2} \Delta_1(h) \Delta_1(k) \Delta_1(l) \Delta_1(-h-k-l) \Delta_2(h+k) \\
& \quad \times \cos(\theta_h + \theta_k + \theta_l + \theta_{-h-k-l}) && \text{lines 1+6} \\
& +2 \binom{N}{2} \Delta_1(h) \Delta_1(k) \Delta_1(l) \Delta_1(-h-k-l) \Delta_2(h+l) \\
& \quad \times \cos(\theta_h + \theta_k + \theta_l + \theta_{-h-k-l}) && \text{lines 2+5} \\
& +2 \binom{N}{2} \Delta_1(h) \Delta_1(k) \Delta_1(l) \Delta_1(-h-k-l) \Delta_2(k+l) \\
& \quad \times \cos(\theta_h + \theta_k + \theta_l + \theta_{-h-k-l}) && \text{lines 3+4} \quad (39)
\end{aligned}$$

The expressions for  $G_0(h)$ ,  $G_1(h)$ ,  $G_2(h)$  are given by functions (25), (27), (33):

$$G_0(h) = (1/2\alpha_0) \exp(-F_h^2/4\alpha_0) \quad (40)$$

$$G_1(h) = [\exp(-f^2/4\alpha_1)/2\alpha_1] \times I_1(fF_h/2\alpha_1) \exp(-F_h^2/4\alpha_1), \quad (41)$$

$$G_2(h) = (f^2/2\alpha_2) \exp(-F_h^2/4\alpha_2) [(F_h/4\alpha_2) - 1], \quad (42)$$

where

$$\alpha_0 = \frac{1}{4} \sum_{\mu=1}^N f_{\mu}^2 = \frac{1}{4} N f^2, \quad (43)$$

$$\alpha_1 = \frac{1}{4} \sum_{\mu=1, \mu \neq \mu_1}^N f_{\mu}^2 = \frac{1}{4} (N-1) f^2, \quad (44)$$

$$\alpha_2 = \frac{1}{4} \sum_{\mu=1, \mu \neq \mu_1, \mu \neq \mu_2}^N f_{\mu}^2 = \frac{1}{4} (N-2) f^2. \quad (45)$$

With the approximation that  $\alpha_0 \approx \alpha_1 \approx \alpha_2 \approx \alpha = Nf^2/4$ ,

$$\Delta_1(h) = \exp(-f^2/4\alpha) I_1(fF_h/2\alpha) \quad (46)$$

and

$$\Delta_2(h) = (f^2/2\alpha)[(F_h^2/4\alpha) - 1]. \quad (47)$$

If  $(N \text{ choose } 2) = N(N-1)/2 = N^2/2$ , the quartet relationship may be written

$$\begin{aligned} & \left[ \prod_{\kappa=h, k, \dots, k+l} (1/2\alpha) \exp(-F_{\kappa}^2/Nf^2) \right] \\ & \times [1 + 2N \exp(-4/N) I_1(2F_h/Nf) I_1(2F_k/Nf) \\ & \times I_1(2F_l/Nf) I_1(2F_{-h-k-l}/Nf) \\ & \times \{1 + \frac{1}{2}[(F_{h+k}^2 + F_{h+l}^2 + F_{k+l}^2)/Nf^2 - 3] \\ & \times \cos(\theta_h + \theta_k + \theta_l + \theta_{-h-k-l})\}]. \quad (48) \end{aligned}$$

If the sum of the squares of the cross-vector reflections is small, then the coefficient multiplying the cosine term is negative. This indicates a maximum of the probability density at  $\pi$ . This relationship is comparable to equation (5) of Heinerman (1977).

Equation (38) can be related to Hauptman (1975) [equation (2.5)] under the condition that  $F_h \ll Nf$ . In this event,  $\Delta_1(h) \approx \exp(-1/N) F_h/Nf \approx E_h/N^{1/2}$ ,  $\Delta_2(h) \approx -1/N$ . Substitution of these values into (39) and the taking of the exponential of the log and then expansion of the log function as a power series to first order yields Hauptman's result.

## 5.2. Example 2: The effect of symmetry

5.2.1. *The triplet in  $P2_1$ .* In this example the effect of  $P2_1$  symmetry on the probability distribution between three reflections  $\mathbf{h}$ ,  $\mathbf{k}$ ,  $\mathbf{l}$  is discussed. The symmetries of  $P2_1$  can be generated by the two Seitz matrices  $S_g = \mathbf{R}_g/t_g$ ,  $g = 1, 2$ :

$$\begin{array}{cc} \mathbf{R}_1 & \mathbf{t}_1 \\ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \end{array} \quad \begin{array}{cc} \mathbf{R}_2 & \mathbf{t}_2 \\ \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} & \begin{bmatrix} 0 \\ 1/2 \\ 0 \end{bmatrix} \end{array}.$$

The equation

$$\begin{aligned} & n_1 \mathbf{h}^T \mathbf{R}_1 + n_2 \mathbf{k}^T \mathbf{R}_1 + n_3 \mathbf{l}^T \mathbf{R}_1 \\ & + n_4 \mathbf{h}^T \mathbf{R}_2 + n_5 \mathbf{k}^T \mathbf{R}_2 + n_6 \mathbf{l}^T \mathbf{R}_2 = 0 \quad (49) \end{aligned}$$

needs to be solved in order to generate the correct set of integers. If no relationships can be specified between the six reflections  $\mathbf{h}^T \mathbf{R}_1, \mathbf{h}^T \mathbf{R}_2, \dots, \mathbf{l}^T \mathbf{R}_2$  then the only solution to (49) is  $n_1 = n_2 = n_3 = n_4 = n_5 = n_6 = 0$ . If, however,  $\mathbf{l} = -\mathbf{h} - \mathbf{k}$  and a triplet is formed then (49) can be simplified to

$$\begin{aligned} & (n_1 - n_3) \mathbf{h}^T \mathbf{R}_1 + (n_2 - n_3) \mathbf{k}^T \mathbf{R}_1 \\ & + (n_4 - n_6) \mathbf{h}^T \mathbf{R}_2 + (n_5 - n_6) \mathbf{k}^T \mathbf{R}_2 = 0. \quad (50) \end{aligned}$$

If  $\mathbf{h}$ ,  $\mathbf{k}$  are general reflections there are no other simplifications to (50) and the solution is  $n_1 = n_2 = n_3$  and  $n_4 = n_5 = n_6$ . For this combination of integers the argument of the cosine terms reduces to a sum over the phases of  $\mathbf{h}$ ,  $\mathbf{k}$ ,  $\mathbf{l}$  and the translation components  $2\pi \mathbf{h} \cdot \mathbf{t}_g$  vanish. The integrals  $G(n, f, F_h)$  are also the same as would be calculated from (4) assuming no symmetry and therefore  $G \times N/G = N$  unique atoms. In other words, the distribution is the same as for  $P1$ . If an extra condition such as  $\mathbf{h}^T \mathbf{R}_1 = \mathbf{h}^T \mathbf{R}_2$  [ $\mathbf{h} = (0, k_1, 0)$ ] is added then (50) reduces to

$$\begin{aligned} & (n_1 - n_3 + n_4 - n_6) \mathbf{h}^T \mathbf{R}_1 \\ & + (n_2 - n_3) \mathbf{k}^T \mathbf{R}_1 + (n_5 - n_6) \mathbf{k}^T \mathbf{R}_2 = 0. \quad (51) \end{aligned}$$

This yields the solutions  $n_2 = n_3$ ,  $n_5 = n_6$ ,  $n_1 + n_4 = n_3 + n_6$  and introduces an array of extra integers that are allowable and that alter the distribution. Table 2 shows the sets of allowable integers along with the coefficients that multiply the cosines. The cosine arguments reduce to

$$\begin{aligned} & - \sum_{\mu=1}^N n_6(\mu) (\theta_h + \theta_k + \theta_l + 2\pi \mathbf{h} \cdot \mathbf{t}_2) \\ & + \sum_{\mu=1}^N n_3(\mu) (\theta_h + \theta_k + \theta_l) + \sum_{\mu=1}^N n_4(\mu) 2\pi \mathbf{h} \cdot \mathbf{t}_2. \quad (52) \end{aligned}$$

The coefficients with the largest magnitudes (*i.e.* that vary as  $N^{-1/2}$ ) and also multiplying a non-constant cosine include lines 1, 2, 4, 8, 9, 11, 18, 24. Adding these terms together yields

$$4\Delta_1(h)\Delta_1(k)\Delta_1(l)[1 + (-1)^{k_1}] \cos \Phi. \quad (53)$$

For odd  $k_1$  the phase relationship vanishes, as expected for a systematically absent reflection. It should be noted the terms that contribute to the invariant relation (53) are completely specified by the integers which provide a solution to (51). Hence the main difficulty is in keeping track of the various terms.

5.2.2. *Seminvariants.* The concepts of neighbourhoods and representations used by Hauptman (1975) and Giacovazzo (1980) are encompassed in this

Table 2. The set of integers that solve equation (51) along with the corresponding arguments;  $\Phi = \theta_h + \theta_k + \theta_l$  and  $\Delta_{nm} = G_{nm}/G_0$ ;  $G_{nm}$  refers to the integral (30)

| Line | $n_1$ | $n_2$ | $n_3$ | $n_4$ | $n_5$ | $n_6$ | Coefficient  | Cosine argument    |
|------|-------|-------|-------|-------|-------|-------|--|--------------------|
| 0    | 0     | 0     | 0     | 0     | 0     | 0     | $\Delta_0(h)\Delta_0(k)\Delta_0(l)$                            | 0                  |
| 1    | -1    | 0     | 0     | 0     | -1    | -1    | $\Delta_1(h)\Delta_1(k)\Delta_1(l)$                            | $\Phi$             |
| 2    | 1     | 0     | 0     | 0     | 1     | 1     | $\Delta_1(h)\Delta_1(k)\Delta_1(l)$                            | $-\Phi$            |
| 3    | 1     | 0     | 0     | -1    | 0     | 0     | $\Delta_{1\bar{1}}(h)\Delta_1(k)\Delta_0(l)$                   | $-\pi k_1$         |
| 4    | 0     | 0     | 0     | -1    | -1    | -1    | $\Delta_1(h)\Delta_1(k)\Delta_{1\bar{1}}(l)$                   | $\Phi - \pi k_1$   |
| 5    | 2     | 0     | 0     | -1    | 1     | 1     | $\Delta_{2\bar{1}}(h)\Delta_1(k)\Delta_1(l)$                   | $-\Phi - \pi k_1$  |
| 6    | -1    | 0     | 0     | 1     | 0     | 0     | $\Delta_{1\bar{1}}(h)\Delta_0(k)\Delta_0(l)$                   | $\pi k_1$          |
| 7    | -2    | 0     | 0     | 1     | -1    | -1    | $\Delta_{2\bar{1}}(h)\Delta_1(k)\Delta_1(l)$                   | $\Phi + \pi k_1$   |
| 8    | 0     | 0     | 0     | 1     | 1     | 1     | $\Delta_1(h)\Delta_1(k)\Delta_1(l)$                            | $-\Phi + \pi k_1$  |
| 9    | -1    | -1    | -1    | 0     | 0     | 0     | $\Delta_1(h)\Delta_1(k)\Delta_1(l)$                            | $\Phi$             |
| 10   | -2    | -1    | -1    | 0     | 1     | 1     | $\Delta_2(h)\Delta_{1\bar{1}}(k)\Delta_{1\bar{1}}(l)$          | $2\Phi$            |
| 11   | 0     | -1    | -1    | 0     | 1     | 1     | $\Delta_0(h)\Delta_{1\bar{1}}(k)\Delta_{1\bar{1}}(l)$          | 0                  |
| 12   | 0     | -1    | -1    | -1    | 0     | 0     | $\Delta_1(h)\Delta_1(k)\Delta_1(l)$                            | $\Phi - \pi k_1$   |
| 13   | -1    | -1    | -1    | -1    | -1    | -1    | $\Delta_{1\bar{1}}(h)\Delta_{1\bar{1}}(k)\Delta_{1\bar{1}}(l)$ | $2\Phi - \pi k_1$  |
| 14   | 1     | -1    | -1    | -1    | 1     | 1     | $\Delta_{1\bar{1}}(h)\Delta_{1\bar{1}}(k)\Delta_{1\bar{1}}(l)$ | $-\pi k_1$         |
| 15   | -2    | -1    | -1    | 1     | 0     | 0     | $\Delta_{2\bar{1}}(h)\Delta_1(k)\Delta_1(l)$                   | $\Phi + \pi k_1$   |
| 16   | -3    | -1    | -1    | 1     | -1    | -1    | $\Delta_{3\bar{1}}(h)\Delta_{1\bar{1}}(k)\Delta_{1\bar{1}}(l)$ | $2\Phi + \pi k_1$  |
| 17   | -1    | -1    | -1    | 1     | 1     | 1     | $\Delta_{1\bar{1}}(h)\Delta_{1\bar{1}}(k)\Delta_{1\bar{1}}(l)$ | $\pi k_1$          |
| 18   | 1     | 1     | 1     | 0     | 0     | 0     | $\Delta_1(h)\Delta_1(k)\Delta_1(l)$                            | $-\Phi$            |
| 19   | 0     | 1     | 1     | 0     | -1    | -1    | $\Delta_0(h)\Delta_{1\bar{1}}(k)\Delta_{1\bar{1}}(l)$          | 0                  |
| 20   | 2     | 1     | 1     | 0     | 1     | 1     | $\Delta_2(h)\Delta_{1\bar{1}}(k)\Delta_{1\bar{1}}(l)$          | $-2\Phi$           |
| 21   | 2     | 1     | 1     | -1    | 0     | 0     | $\Delta_{2\bar{1}}(h)\Delta_1(k)\Delta_1(l)$                   | $-\Phi - \pi k_1$  |
| 22   | 1     | 1     | 1     | -1    | -1    | -1    | $\Delta_{1\bar{1}}(h)\Delta_{1\bar{1}}(k)\Delta_{1\bar{1}}(l)$ | $-\pi k_1$         |
| 23   | 3     | 1     | 1     | -1    | 1     | 1     | $\Delta_{3\bar{1}}(h)\Delta_{1\bar{1}}(k)\Delta_{1\bar{1}}(l)$ | $-2\Phi - \pi k_1$ |
| 24   | 0     | 1     | 1     | 1     | 0     | 0     | $\Delta_1(h)\Delta_1(k)\Delta_1(l)$                            | $-\Phi + \pi k_1$  |
| 25   | -1    | 1     | 1     | 1     | -1    | -1    | $\Delta_{1\bar{1}}(h)\Delta_{1\bar{1}}(k)\Delta_{1\bar{1}}(l)$ | $-\pi k_1$         |
| 26   | 1     | 1     | 1     | 1     | 1     | 1     | $\Delta_{1\bar{1}}(h)\Delta_{1\bar{1}}(k)\Delta_{1\bar{1}}(l)$ | $-2\Phi + \pi k_1$ |

formulation of the probability distribution. This stems from the fact that if the summation integers  $n(h, g = i) = -n(h, g = j)$  then the phase  $\Phi_h$  does not appear in the cosine argument. For example, in the two-phase seminvariant given by

$$\Phi_{\mathbf{h}_1} - \Phi_{\mathbf{h}_2} \quad (54)$$

where  $\mathbf{h}_1 = (h_1, k, l_1)$  and  $\mathbf{h}_2 = (h_2, k, l_2)$ , the following invariant can be constructed (Giacovazzo, 1979):

$$\mathbf{h}_1 - \mathbf{h}_2 + \mathbf{h}_1^T \mathbf{R}_2 - \mathbf{h}_2^T \mathbf{R}_2 = \mathbf{0}. \quad (55)$$

For the  $N/G = N/2$  equal-atom case the lowest-order contribution for this invariant relationship is

$$(1/N) E_{\mathbf{h}_1}^2 E_{\mathbf{h}_2}^2 \cos [2(\Phi_{\mathbf{h}_1} - \Phi_{\mathbf{h}_2})]. \quad (56)$$

An equally large contribution can be obtained from considering the first representation of this invariant (Giacovazzo, 1980). This is the reflection(s)  $\mathbf{l}$  such that

$$\mathbf{h}_1 - \mathbf{h}_2 + \mathbf{l}^T \mathbf{R}_s - \mathbf{l}^T \mathbf{R}_p = \mathbf{0} \quad (57)$$

is true for some rotation matrices  $\mathbf{R}_s$  and  $\mathbf{R}_p$ . Again the phase of  $\mathbf{l}$  does not appear in the cosine argument and we have

$$(2/N) E_{\mathbf{h}_1} E_{\mathbf{h}_2} (E_{\mathbf{l}}^2 - 1) \cos [\Phi_{\mathbf{h}_1} - \Phi_{\mathbf{h}_2} + 2\pi \mathbf{l} \cdot (\mathbf{t}_s - \mathbf{t}_p)]. \quad (58)$$

These two relationships can be compared with the seminvariant formula derived by Hauptman & Green [(1978), equation (4.1)].

## 6. Concluding remarks

A general expression for the joint probability distribution between phases assuming a random atom distribution has been derived. The form of the integrals  $G(n, \mathbf{f}, F_h)$  is complex but adequate approximations can be made if the number of atoms is large. Identification of the integers  $n(\mu, g, h)$  that satisfy the constraints should be possible through a computer algebraic approach. The methods described reduce the complications involved in expanding the characteristic function as a power series (Naya *et al.*, 1965) if only an asymptotic distribution is required. The effect of neighbourhood reflections is made explicit.

Finally, it should be noted that joint probability distributions are usually expressed as an asymptotic series in powers of  $N^{-1/2}$ . These series are not convergent for a given  $N$  and the truncation error is of the order of the last term neglected. No analysis of the convergence properties of (1) has been conducted although the proof given in § 3 shows that the series must converge at least conditionally.

The author thanks Professor Syd Hall for many useful discussions. This work was supported by a Commonwealth Postgraduate Research Award.

## References

- ABRAMOWITZ, M. & STEGUN, I. (1972). *Handbook of Mathematical Functions*. New York: Dover.
- BARAKAT, R. & COLE, J. E. III (1979). *J. Sound Vib.* **62**(3), 365-377.
- BRICOGNE, G. (1984). *Acta Cryst.* **A40**, 410-445.

- FORTIER, S. & HAUPTMAN, H. (1977). *Acta Cryst.* **A33**, 694–696.  
 GIACOVAZZO, C. (1974). *Acta Cryst.* **A30**, 522–524.  
 GIACOVAZZO, C. (1975). *Acta Cryst.* **A31**, 252–259.  
 GIACOVAZZO, C. (1976). *Acta Cryst.* **A32**, 91–99.  
 GIACOVAZZO, C. (1979). *Acta Cryst.* **A35**, 296–305.  
 GIACOVAZZO, C. (1980). *Direct Methods in Crystallography*. London: Academic Press.  
 GRADSHTEYN, I. & RYZHIK, I. (1980). *Table of Integrals, Series, and Products*. New York: Academic Press.  
 HAUPTMAN, H. (1974). *Acta Cryst.* **A30**, 822–829.  
 HAUPTMAN, H. (1975). *Acta Cryst.* **A31**, 671–679.  
 HAUPTMAN, H. & GREEN, E. A. (1978). *Acta Cryst.* **A34**, 216–223.  
 HAUPTMAN, H. & KARLE, J. (1953). *Acta Cryst.* **6**, 136–141.  
 HEINERMAN, J. J. L. (1977). *Acta Cryst.* **A33**, 100–106.  
 HEINERMAN, J. J. L., KRABBENDAM, H. & KROON, J. (1977). *Acta Cryst.* **A33**, 873–878.  
 HEINERMAN, J. J. L., KRABBENDAM, H. & KROON, J. (1979). *Acta Cryst.* **A35**, 101–103.  
 KARLE, J. & HAUPTMAN, H. (1958). *Acta Cryst.* **11**, 264–269.  
 KLUG, A. (1958). *Acta Cryst.* **11**, 515–543.  
 NAYA, S., NITTA, I. & ODA, T. (1965). *Acta Cryst.* **19**, 734–747.  
 PESCHAR, R. & SCHENK, H. (1986). *Acta Cryst.* **A42**, 309–317.  
 SHMUELI, U. & WEISS, G. H. (1985). *Acta Cryst.* **A41**, 401–408.  
 TSOUCARIS, G. (1970). *Acta Cryst.* **A26**, 492–499.  
 WATSON, G. N. (1942). *A Treatise on the Theory of Bessel Functions*, 2nd ed. Cambridge Univ. Press.

*Acta Cryst.* (1987). **A43**, 393–400

## Diffraction Patterns from Tilings with Fivefold Symmetry

BY E. PRINCE

*Institute for Materials Science and Engineering, National Bureau of Standards, Gaithersburg, Maryland 20899, USA*

(Received 21 March 1986; accepted 1 December 1986)

### Abstract

A procedure involving projection from six-dimensional to three-dimensional space to describe objects that give sharp diffraction with fivefold symmetry can be reduced to the easier problem of projection from two dimensions to one dimension. This result is used to derive an explicit formula for the quasilattice contribution to the diffracted intensity for an arbitrary size and shape of the selection region. The predictions of this formula are compared with the electron diffraction patterns obtained from rapidly solidified aluminium–manganese alloys, and it is concluded that the edges of the rhombic faces of the three-dimensional objects from which models for these alloy structures may be constructed is larger than that used in previous analyses by a factor of  $\tau^3$ , where  $\tau$  is the golden mean. It is shown that the quasilattice density is proportional to the volume of the selection region in the complementary three-dimensional space into which a lattice point in six-dimensional space must project in order for the point to be included in the direct space; this results in important constraints on the possible structures of these alloys.

### Introduction

The recent discovery by Shechtman, Blech, Gratias & Cahn (1984) of electron diffraction patterns with icosahedral symmetry that have sharp spots has stimulated a new study of the conditions under which such diffraction patterns may exist. Sharp diffraction has traditionally been associated with periodic lat-

tices, and it has long been known that fivefold rotation symmetry is incompatible with an infinite translation lattice. It has been shown, however (Levine & Steinhardt, 1984), that sharp diffraction does occur from non-periodic patterns under some conditions. Moreover, as was first shown by Penrose (1974; also Mackay, 1976; Gardner, 1977), patterns with fivefold symmetry can be produced that have long-range order although they lack periodicity. The two-dimensional Penrose tilings have been generalized to three dimensions by Mackay (1981, 1982). [In fact, Mackay's results were anticipated by Baer (1970), who used the principles to construct a number of interesting architectural structures.] Kramer & Neri (1984), Duneau & Katz (1985) and Kalugin, Kitayev & Levitov (1985) have shown that the three-dimensional tilings can be understood in terms of projection from six dimensions, in which fivefold rotation is compatible with periodicity, into three dimensions.

Regardless of whether the alloy of aluminium and manganese that was studied by Shechtman *et al.* (1984) is actually an example of a three-dimensional object that has an inherent icosahedral diffraction pattern, the mathematical development of the diffraction properties of these tilings stands by itself, independent of its applications to any experimental observations. As in periodic crystals, the intensity of a diffraction spot for the 'quasicrystal' (Mackay, 1981) with fivefold symmetry is proportional to the product of two factors, one due to the projected lattice and one due to the arrangement of atoms associated with each point. In contrast to crystals, however, the spots