treatment may lower the GoF significantly. Corrections for thermal diffuse scattering (Helmholdt & Vos, 1975) may also be important. To what extent these changes in data reduction will affect the charge-density parameters is uncertain but it may well be that the resolution of the bonding-density features could be further improved.

This work has been supported financially by the Australian Research Grants Committee, by the Australian Institute of Nuclear Science and Engineering, and by the Research Committee of the University of Western Australia. Two of us (PFP and JNV) gratefully acknowledge receipt of Commonwealth Postgraduate Awards.

References

BACON, J. & SANTRY, D. P. (1971). *J. Chem. Phys.* **55**, 3743–3751.

BECKER, P. J. (1975). *Acta Cryst.* A**31**, S227.

BECKER, P. J. & COPPENS, P. (1974). *Acta Cryst.* A**30**, 129–149.

CHANDLER, G. S. & SPACKMAN, M. A. (1975). Private communication.

CLEMENTI, E. (1965). *IBM J. Res. Dev.* **9**. Suppl. *Tables of Atomic Functions.*

COULSON, C. A. & THOMAS, M. W. (1971). *Acta Cryst.* B**27**, 1354–1359.

CROMER, D. T., LARSON, A. C. & STEWART, R. F. (1975*a*). *Acta Cryst.* A**31**, S224.

CROMER, D. T., LARSON, A. C. & STEWART, R. F. (1975*b*). *Application of Generalised Scattering Factors. How Atoms in Molecules Really Look.* Program Abstr., 33rd Pittsburgh Diffr. Conf.

DELBAERE, L. T. J. & JAMES, M. N. G. (1973). *Acta Cryst.* B**29**, 2905–2912.

EPSTEIN, J. & STEWART, R. F. (1975). *Electron Densities of Diatomic Hydrides from Generalised X-ray Scattering Factors.* Program Abstr. 33rd Pittsburgh Diffr. Conf.

HEHRE, W. J., STEWART, R. F. & POPLE, J. A. (1969). *J. Chem. Phys.* **51**, 2657–2664.

HELMHOLDT, R. B. & VOS, A. (1975). *Acta Cryst.* A**31**, S226.

HUGHES, E. W. (1941). *J. Am. Chem. Soc.* **63**, 1731–1752.

LARSON, A. C. & CROMER, D. T. (1974). *J. Chem. Phys.* **60**, 185–192.

PRICE, P. F., MASLEN, E. N. & DELANEY, W. T. (1978). *Acta Cryst.* A**34**, 194–203

PRICE, P. F., MASLEN, E. N. & MOORE, F. H. (1978). *Acta Cryst.* A**34**, 173–183

SANTRY, D. P. (1974). Private communication.

STEWART, R. F. (1972). *Electron Population Analysis with Generalised X-ray Scattering Factors.* Unpublished.

STEWART, R. F. (1973*a*). *Acta Cryst.* A**29**, 602–605.

STEWART, R. F. (1973*b*). *J. Chem. Phys.* **58**, 4430–4438.

STEWART, R. F. (1974). *Charge Densities from X-ray Diffraction Data in Critical Evaluation of Chemical and Physical Structural Information.* pp. 540–555. New York: National Academy of Sciences.

VARGHESE, J. N. (1974). Ph.D. Thesis, Univ. of Western Australia.

VARGHESE, J. N., O'CONNELL, A. M. & MASLEN, E. N. (1977). *Acta Cryst.* B**33**, 2102–2108.

ZACHARIASEN, W. (1967). *Acta Cryst.* **23**, 558–564.

Acta Cryst. (1978). A**34**, 216–223

Pairs in $P2_1$: Probability Distributions which Lead to Estimates of the Two-Phase Structure Seminvariants in the Vicinity of 0 or π

BY EDWARD A. GREEN AND HERBERT HAUPTMAN

Medical Foundation of Buffalo, Inc., 73 *High Street, Buffalo, New York* 14203, *USA*

(*Received* 16 *May* 1977; *accepted* 17 *August* 1977)

The first sequence of nested neighborhoods of the two-phase structure seminvariant $\varphi_{12} = \varphi_{h_1k_1} - \varphi_{h_2k_2}$ in the space group $P2_1$ is defined, and conditional probability distributions associated with the first four neighborhoods derived. In the favorable case that the variance of a distribution happens to be small, the distribution yields a particularly reliable value for φ_{12}. The most reliable estimates are obtained when $\varphi_{12} \simeq 0$ or π.

1. Introduction

In the space group $P2_1$, the linear combination of two phases

$$\varphi_{12} = \varphi_{h_1kl_1} - \varphi_{h_2kl_2} \quad (1.1)$$

is a structure seminvariant if and only if

$$(h_1 - h_2, 0, l_1 - l_2) \equiv 0 \pmod{\omega_s} \quad (1.2)$$

where ω_s, the seminvariant modulus in $P2_1$, is defined by

$$\omega_s = (2,0,2). \quad (1.3)$$

In short, φ_{12} is a structure seminvariant if and only if

$h_1 - h_2$ and $l_1 - l_2$ are even integers. The algebraic theory of these seminvariants was initiated several years ago (Hauptman, 1972), but the results obtained in the present series of papers go far beyond those obtainable by algebraic means. In this and the following two papers (Hauptman & Green, 1978; Green & Hauptman, 1978), the probabilistic theory of these two-phase structure seminvariants is initiated *via* the Principle of Nested Neighborhoods (Hauptman, 1975*a*,*b*). Unlike the theory of the four-phase structure invariant in $P1$ and $P\bar{1}$ (Hauptman, 1977*b*,*c*), which is based on a single sequence of neighborhoods (Hauptman, 1977*a*), here there exist two distinct sequences of neighborhoods associated with the two-phase structure seminvariant. Estimates of φ_{12} obtained from the first sequence of neighborhoods tend to be most reliable when the estimates are in the vicinity of 0 or π. Estimates obtained from the second sequence of neighborhoods are most useful when φ_{12} is in the vicinity of $\pm\pi/2$. Thus the two sequences of neighborhoods complement each other.

In this paper, probability distributions associated with the first four neighborhoods of the first sequence are obtained. Probability distributions associated with the first four neighborhoods of the second sequence are derived in the second paper. In the third paper, conditional probability distributions of a structure seminvariant, given, not only magnitudes $|E|$, but the values of one or more structure seminvariants as well, are used to estimate the values, *i.e.* both magnitudes and signs, of a family of structure seminvariants consistent with a specified enantiomorph.

2. The first sequence of neighborhoods of the two-phase structure seminvariant, $\varphi_{h_1kl_1} - \varphi_{h_2kl_2}$, in $P2_1$

2.1. *The first neighborhoods*

Assume that (1.2) holds. Construct the four-phase structure invariant

$$\varphi_{h_1kl_1} - \varphi_{h_2kl_2} - \varphi_{\frac{1}{2}(h_1-h_2),q,\frac{1}{2}(l_1-l_2)} - \varphi_{\frac{1}{2}(h_1-h_2),\bar{q},\frac{1}{2}(l_1-l_2)} \quad (2.1)$$

where q is an arbitrary non-zero integer. In view of (1.2) and (1.3), the components $\frac{1}{2}(h_1 - h_2)$ and $\frac{1}{2}(l_1 - l_2)$ are integers. The theory of the first neighborhood of the four-phase structure invariant (Hauptman, 1975*a*,*b*) shows that the linear combination of phases (2.1) is probably close to zero if the four magnitudes of the first neighborhood are large. However, in $P2_1$,

$$|E_{\frac{1}{2}(h_1-h_2),q,\frac{1}{2}(l_1-l_2)}| = |E_{\frac{1}{2}(h_1-h_2),\bar{q},\frac{1}{2}(l_1-l_2)}| \quad (2.2)$$

and

$$\varphi_{\frac{1}{2}(h_1-h_2),q,\frac{1}{2}(l_1-l_2)} + \varphi_{\frac{1}{2}(h_1-h_2),\bar{q},\frac{1}{2}(l_1-l_2)}| = \pi q. \quad (2.3)$$

Therefore, in view of (2.1), if the three magnitudes

$$|E_{h_1kl_1}|,\ |E_{h_2kl_2}|,\ |E_{\frac{1}{2}(h_1-h_2),q,\frac{1}{2}(l_1-l_2)}| \quad (2.4)$$

are large, the quartet (2.1) is probably close to zero, and

$$\varphi_{12} \simeq \pi q. \quad (2.5)$$

In other words, if the magnitudes (2.4) are large,

$$\varphi_{12} \simeq 0 \text{ or } \pi \quad (2.6)$$

according as q is even or odd, respectively. Accordingly, the first neighborhood of φ_{12} is defined to consist of the three magnitudes (2.4) which are shown in the first shell of Fig. 1. Since q is an arbitrary non-zero integer, there are many first neighborhoods.

2.2. *The second neighborhoods*

The second neighborhood of the two-phase seminvariant φ_{12} is defined to be the second neighborhood of the four-phase structure invariant (2.1) (Hauptman, 1975*a*,*b*). Thus, the second neighborhood consists of the three magnitudes (2.4) and the three additional magnitudes

$$|E_{\frac{1}{2}(h_1+h_2),q+k,\frac{1}{2}(l_1+l_2)}|,\quad |E_{\frac{1}{2}(h_1+h_2),q-k,\frac{1}{2}(l_1+l_2)}|,\quad |E_{h_1-h_2,0,l_1-l_2}|, \quad (2.7)$$

shown in the second shell of Fig. 1.

In view of the quartet theory, if the six magnitudes (2.4) and (2.7) are all large, the quartet (2.1) is probably close to zero, and

$$\varphi_{12} \simeq \pi q. \quad (2.8)$$

However, if the three magnitudes (2.4) are large and the three magnitudes (2.7) are small, then the quartet (2.1)

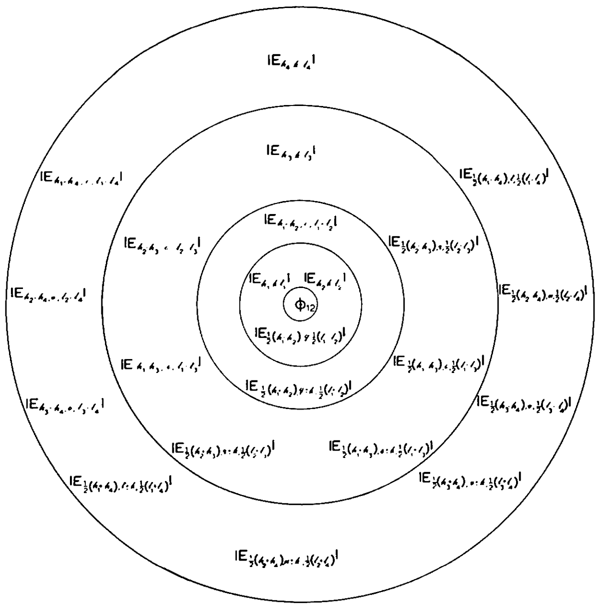

Fig. 1. The first sequence of nested neighborhoods of the two-phase structure seminvariant φ_{12} in $P2_1$; $h_\mu \equiv h_\nu$ (mod 2), $l_\mu \equiv l_\nu$ (mod 2) and q, r, s, t, u and v are arbitrary non-zero integers. The first neighborhood consists of the three magnitudes in the first shell, the second neighborhood of the six magnitudes in the first two shells, *etc.*

is probably close to π, and

$$\varphi_{12} \simeq \pi(q+1). \tag{2.9}$$

Since q is an arbitrary non-zero integer, there are many second neighborhoods.

2.3. *The third neighborhoods*

The third neighborhood of φ_{12} is again obtained by arguments similar to those used previously for the four-phase structure invariant (Hauptman, 1977*a*). If h_3kl_3 is a reciprocal lattice vector which satisfies

$$(h_2kl_2)-(h_3kl_3)\equiv 0 \pmod{\omega_s} \tag{2.10}$$

then

$$\varphi_{23}=\varphi_{h_2kl_2}-\varphi_{h_3kl_3} \tag{2.11}$$

is a two-phase structure seminvariant. The second neighborhood of φ_{23} consists of the six magnitudes

$$|E_{h_2kl_2}|,\ |E_{h_3kl_3}|,\ |E_{\frac{1}{2}(h_2-h_3),r,\frac{1}{2}(l_2-l_3)}|,\ |E_{h_2-h_3,0,l_2-l_3}|,\ |E_{\frac{1}{2}(h_2+h_3),r\pm k,\frac{1}{2}(l_2+l_3)}|, \tag{2.12}$$

where r is an arbitrary non-zero integer. Since φ_{12} and φ_{23} are both two-phase seminvariants,

$$\varphi_{31}=\varphi_{h_3kl_3}-\varphi_{h_1kl_1} \tag{2.13}$$

is also a structure seminvariant and has a second neighborhood consisting of the six magnitudes

$$|E_{h_1kl_1}|,\ |E_{h_3kl_3}|,\ |E_{\frac{1}{2}(h_3-h_1),s,\frac{1}{2}(l_3-l_1)}|,\ |E_{h_3-h_1,0,l_3-l_1}|,\ |E_{\frac{1}{2}(h_3+h_1),s\pm k,\frac{1}{2}(l_3+l_1)}|, \tag{2.14}$$

where s is an arbitrary non-zero integer. However, from (1.1), (2.11), and (2.13), the following identity holds:

$$\varphi_{12}+\varphi_{23}+\varphi_{31}\equiv 0. \tag{2.15}$$

Therefore, in the favourable case that the six-magnitude estimates yield values for φ_{12}, φ_{23}, and φ_{31} in accord with (2.15), φ_{12} will be well estimated in terms of the 18 magnitudes (2.4), (2.7), (2.12), and (2.14) of which only the following 15 are distinct:

$$|E_{h_1kl_1}|,\ |E_{h_2kl_2}|,\ |E_{h_3kl_3}|,\ |E_{\frac{1}{2}(h_1-h_2),q,\frac{1}{2}(l_1-l_2)}|,\ |E_{\frac{1}{2}(h_2-h_3),r,\frac{1}{2}(l_2-l_3)}|,\ |E_{\frac{1}{2}(h_3-h_1),s,\frac{1}{2}(l_3-l_1)}|,$$
$$|E_{\frac{1}{2}(h_1+h_2),q\pm k,\frac{1}{2}(l_1+l_2)}|,\ |E_{\frac{1}{2}(h_2+h_3),r\pm k,\frac{1}{2}(l_2+l_3)}|,\ |E_{\frac{1}{2}(h_3+h_1),s\pm k,\frac{1}{2}(l_3+l_1)}|,\ |E_{h_1-h_2,0,l_1-l_2}|,$$
$$|E_{h_2-h_3,0,l_2-l_3}|,\ |E_{h_3-h_1,0,l_3-l_1}|. \tag{2.16}$$

Thus, the third (15-magnitude) neighborhood of φ_{12} is obtained by adjoining to the second (six-magnitude) neighborhood, (2.4) and (2.7), the nine additional magnitudes shown in the third shell of Fig. 1,

$$|E_{h_3kl_3}|,\ |E_{\frac{1}{2}(h_2-h_3),r,\frac{1}{2}(l_2-l_3)}|,\ |E_{\frac{1}{2}(h_3-h_1),s,\frac{1}{2}(l_3-l_1)}|,\ |E_{h_2-h_3,0,l_2-l_3}|,\ |E_{h_3-h_1,0,l_3-l_1}|,\ |E_{\frac{1}{2}(h_2+h_3),r\pm k,\frac{1}{2}(l_2+l_3)}|,\ |E_{\frac{1}{2}(h_3+h_1),s\pm k,\frac{1}{2}(l_3+l_1)}|, \tag{2.17}$$

where r and s are arbitrary non-zero integers; hence there are many third neighborhoods.

One naturally anticipates that the conditional variance of the two-phase structure seminvariant φ_{12}, given the 15 magnitudes in its third neighborhood, will be small if the six-magnitude second neighborhoods of φ_{12}, φ_{23}, φ_{31} yield estimates for the latter in accord with the identity (2.15). Thus, those neighborhoods are most useful for which

$$|E_{\frac{1}{2}(h_1-h_2),q,\frac{1}{2}(l_1-l_2)}|,\ |E_{\frac{1}{2}(h_2-h_3),r,\frac{1}{2}(l_2-l_3)}|,\ |E_{\frac{1}{2}(h_3-h_1),s,\frac{1}{2}(l_3-l_1)}| \tag{2.18}$$

are large.

2.4. *The fourth neighborhoods*

Again, as in § 2.3, the fourth neighborhood of φ_{12} is obtained by the same method as that used for quartets (Hauptman, 1977*a*). If (h_4kl_4) is a reciprocal lattice vector satisfying

$$(h_1kl_1)-(h_4kl_4)\equiv 0 \pmod{\omega_s}, \tag{2.19}$$

then

$$\varphi_{14}=\varphi_{h_1kl_1}-\varphi_{h_4kl_4} \tag{2.20}$$

is a structure seminvariant. The second neighborhood of φ_{14} consists of the six magnitudes

$$|E_{h_1kl_1}|,\ |E_{h_4kl_4}|,\ |E_{\frac{1}{2}(h_1-h_4),t,\frac{1}{2}(l_1-l_4)}|,\ |E_{\frac{1}{2}(h_1+h_4),t+k,\frac{1}{2}(l_1+l_4)}|,\ |E_{\frac{1}{2}(h_1+h_4),t-k,\frac{1}{2}(l_1+l_4)}|,\ |E_{h_1-h_4,0,l_1-l_4}|, \tag{2.21}$$

where t is an arbitrary non-zero integer.

From (1.2), (1.3), (2.10) and (2.19), it is seen that h_3 and h_4 have the same parity, as do l_3 and l_4, so that

$$\varphi_{43}=\varphi_{h_4kl_4}-\varphi_{h_3kl_3} \tag{2.22}$$

is also a structure seminvariant. The second neighborhood of φ_{43} consists of the six magnitudes

$$|E_{h_3kl_3}|,\ |E_{h_4kl_4}|,\ |E_{\frac{1}{2}(h_4-h_3),v,\frac{1}{2}(l_4-l_3)}|,\ |E_{\frac{1}{2}(h_4+h_3),v+k,\frac{1}{2}(l_4+l_3)}|,\ |E_{\frac{1}{2}(h_4+h_3),v-k,\frac{1}{2}(l_4+l_3)}|,\ E|_{h_4-h_3,0,l_4-l_3}|, \tag{2.24}$$

where v is an arbitrary non-zero integer, and the three seminvariants φ_{31}, φ_{14}, φ_{43} satisfy the identity

$$\varphi_{31}+\varphi_{14}+\varphi_{43}\equiv 0. \tag{2.25}$$

In view of (1.2), (1.3) and (2.19) h_2 and h_4 have the same parity, as do l_2 and l_4, so that

$$\varphi_{42}=\varphi_{h_4kl_4}-\varphi_{h_2kl_2} \tag{2.26}$$

is also a structure seminvariant. The second neighborhood of φ_{42} consists of the six magnitudes

$$|E_{h_2kl_2}|,\ |E_{h_4kl_4}|,\ |E_{\frac{1}{2}(h_4-h_2),u,\frac{1}{2}(l_4-l_2)}|,\ |E_{\frac{1}{2}(h_4+h_2),u+k,\frac{1}{2}(l_4+l_2)}|,\ |E_{\frac{1}{2}(h_4+h_2),u-k,\frac{1}{2}(l_4+l_2)}|,\ |E_{h_4-h_2,0,l_4-l_2}|, \tag{2.27}$$

where u is an arbitrary non-zero integer. Since $\varphi_{ij}=-\varphi_{jl}$, in particular $\varphi_{43}=-\varphi_{34}$, the identity

$$\varphi_{23}+\varphi_{34}+\varphi_{42}\equiv 0 \tag{2.28}$$

holds. The 28-magnitude fourth neighborhood of φ_{12} is obtained by adjoining the 13 distinct magnitudes in

(2.21), (2.24) and (2.27), shown in the fourth shell of Fig. 1, to the 15 magnitudes of the third neighborhood, (2.16).

We expect that, in the favorable case that the six-magnitude estimates of the structure seminvariants φ_{12}, φ_{13}, φ_{14}, φ_{23}, φ_{24}, φ_{34} conform to the identities (2.15), (2.25) and (2.28), then the variance of the conditional probability distribution of φ_{12}, given the 28 magnitudes in its fourth neighborhood, will be reduced and the corresponding estimate for φ_{12} better than those of the lower-order distributions.

3. Probabilistic background and notation

It is assumed that a crystal structure in $P2_1$ consisting of N atoms, not necessarily identical, in the unit cell is fixed, and that the three non-negative numbers R_1, R_2, $R_{1\bar{2}/10}$ are also specified. Suppose that the ordered pair $[(h_1kl_1), (h_2kl_2)]$ of reciprocal vectors is a random variable which is uniformly distributed over that subset of the two-fold Cartesian product $W \times W$ of reciprocal space W defined by (1.2), (1.3) and

$$|E_{h_1kl_1}| = R_1, \quad |E_{h_2kl_2}| = R_2, \quad |E_{\frac{1}{2}(h_1-h_2),q,\frac{1}{2}(l_1-l_2)}| = R_{1\bar{2}/10}. \tag{3.1}$$

The structure seminvariant, φ_{12}, is then a random variable whose conditional probability distribution, $P_{1|3}$, given the three magnitudes (3.1) in its first neighborhood, depends on the parameters R_1, R_2, $R_{1\bar{2}/10}$.

If in addition to (3.1), the three non-negative numbers $R_{12/11}$, $R_{12/1\bar{1}}$, $R_{1\bar{2}}$ are also specified, and it is assumed that the primitive random variable $[(h_1kl_1), (h_2kl_2)]$ is uniformly distributed over the subset of $W \times W$ defined by (1.2), (3.1) and

$$\begin{aligned} &|E_{\frac{1}{2}(h_1+h_2),q+k,\frac{1}{2}(l_1+l_2)}| = R_{12/11}, \\ &|E_{\frac{1}{2}(h_1+h_2),q-k,\frac{1}{2}(l_1+l_2)}| = R_{12/1\bar{1}}, \\ &|E_{h_1-h_2,0,l_1-l_2}| = R_{1\bar{2}}, \end{aligned} \tag{3.2}$$

then one is led to the conditional probability distribution, $P_{1|6}$, of the structure seminvariant φ_{12}, given the six magnitudes (3.1) and (3.2) in its second neighborhood.

One continues in this way first to specify the nine additional non-negative numbers R_3, $R_{2\bar{3}/20}$, $R_{23/21}$, $R_{23/2\bar{1}}$, R_{23}, $R_{3\bar{1}/30}$, $R_{31/31}$, $R_{31/3\bar{1}}$, $R_{3\bar{1}}$ and then to assume that the ordered triple $[(h_1kl_1), (h_2kl_2), (h_3kl_3)]$ is a random variable which is uniformly distributed over the subset of the threefold Cartesian product $W \times W \times W$ defined by (1.2), (1.3), (2.10), (3.1), (3.2) and by

$$\begin{aligned} &|E_{h_3kl_3}| = R_3, \ |E_{\frac{1}{2}(h_2-h_3),r,\frac{1}{2}(l_2-l_3)}| = R_{2\bar{3}/20}, \\ &|E_{\frac{1}{2}(h_2+h_3),r+k,\frac{1}{2}(l_2+l_3)}| = R_{23/21}, \\ &|E_{\frac{1}{2}(h_2+h_3),r-k,\frac{1}{2}(l_2+l_3)}| = R_{23/2\bar{1}}, \quad |E_{h_2-h_3,0,l_2-l_3}| = R_{2\bar{3}}, \\ &|E_{\frac{1}{2}(h_3-h_1),s,\frac{1}{2}(l_3-l_1)}| = R_{3\bar{1}/30}, \quad |E_{\frac{1}{2}(h_3+h_1),s+k,\frac{1}{2}(l_3+l_1)}| = R_{31/31}, \\ &|E_{\frac{1}{2}(h_3+h_1),s-k,\frac{1}{2}(l_3+l_1)}| = R_{31/3\bar{1}}, \quad |E_{h_3-h_1,0,l_3-l_1}| = R_{3\bar{1}}. \end{aligned} \tag{3.3}$$

Now one arrives at the conditional probability distribution, $P_{1|15}$, of φ_{12} given the 15 magnitudes (3.1)–(3.3) in its third neighborhood. In a similar way one is led to the conditional distribution of φ_{12}, $P_{1|28}$, given the 15 magnitudes (3.1)–(3.3) and the 13 additional magnitudes:

$$\begin{aligned} &|E_{h_4kl_4}| = R_4, \ |E_{\frac{1}{2}(h_1-h_4),t,\frac{1}{2}(l_1-l_4)}| = R_{1\bar{4}/40}, \\ &|E_{\frac{1}{2}(h_1+h_4),t+k,\frac{1}{2}(l_1+l_4)}| = R_{14/41}, \\ &|E_{\frac{1}{2}(h_1+h_4),t-k,\frac{1}{2}(l_1+l_4)}| = R_{14/4\bar{1}}, \\ &|E_{h_1-h_4,0,l_1-l_4}| = R_{1\bar{4}}, \quad |E_{\frac{1}{2}(h_4-h_2),u,\frac{1}{2}(l_4-l_2)}| = R_{4\bar{2}/50}, \\ &|E_{\frac{1}{2}(h_4+h_2),u+k,\frac{1}{2}(l_4+l_2)}| = R_{42/51}, |E_{\frac{1}{2}(h_4+h_2),u-k,\frac{1}{2}(l_4+l_2)}| = R_{42/5\bar{1}}, \\ &|E_{h_4-h_2,0,l_4-l_2}| = R_{4\bar{2}}, \ |E_{\frac{1}{2}(h_4-h_3),v,\frac{1}{2}(l_4-l_3)}| = R_{4\bar{3}/60}, \\ &|E_{\frac{1}{2}(h_4+h_3),v+k,\frac{1}{2}(l_4+l_3)}| = R_{43/61}, |E_{\frac{1}{2}(h_4+h_3),v-k,\frac{1}{2}(l_4+l_3)}| = R_{43/6\bar{1}}, \\ &|E_{h_4-h_3,0,l_4-l_3}| = R_{4\bar{3}}. \end{aligned} \tag{3.4}$$

In $P2_1$ the normalized structure factor E_{hkl} is defined by

$$\begin{aligned} E_{hkl} &= |E_{hkl}| \exp(i\varphi_{hkl}) \\ &= \frac{2}{(\varepsilon\sigma_2)^{1/2}} \sum_{j=1}^{N/2} f_j \cos 2\pi\left(\mathbf{h}\cdot\mathbf{r}_j + \frac{k}{4}\right) \\ &\quad \times \exp\left[2\pi i\left(ky_j - \frac{k}{4}\right)\right], \end{aligned} \tag{3.5}$$

where $\varepsilon = 2$ if $h = l = 0$ and 1 otherwise; finally, (x_j, y_j, z_j) is the position vector of the jth atom. The $\mathbf{h}$ and $\mathbf{r}_j$ are two-dimensional vectors defined by

$$\mathbf{h} = (h, l), \tag{3.6}$$

$$\mathbf{r}_j = (x_j, z_j), \tag{3.7}$$

and f_j is the zero-angle atomic scattering factor of the atom labeled j; the term σ_n is defined by

$$\sigma_n = \sum_{j=1}^{N} f_j^n. \tag{3.8}$$

In the case of X-ray diffraction, the f_j are equal to the atomic numbers Z_j. In the neutron diffraction case some of the f_j may be negative.

In the sequel conditional probability distributions of φ_{12}, given the magnitudes in each of its first four neighborhoods, are described. Only the briefest sketch of the derivations is given in Appendix I for the typical case of the second neighborhood; the heavy dependence on earlier work permits substantial abbreviation.

4. The conditional probability distribution of the two-phase structure seminvariant $\varphi_{12} = \varphi_{h_1kl_1} - \varphi_{h_2kl_2}$, given the three magnitudes in its first neighborhood

Suppose that the non-negative numbers R_1, R_2, $R_{12/10}$ defined in (3.1) are specified. Then, complete to terms of order $1/N$, $P_{1|3} = P_{1|3}(\Phi)$, the conditional probability distribution of φ_{12}, given the three magnitudes (3.1) of the first neighborhood, is found to be

$$P_{1|3} \simeq \frac{1}{K_{1|3}} \exp\left\{\frac{\sigma_4}{\sigma_2^2}\,[2(-1)^q R_1 R_2 (R^2_{1\bar{2}/10} - 1)\cos\Phi + R_1^2 R_2^2 \cos 2\Phi]\right\}, \tag{4.1}$$

which is correct through terms of order $1/N$. Although an analytical representation for the normalization factor $K_{1|3}$ may be found, it is not needed for the present purpose and, in any event, is more easily computed numerically in any given case.

5. The conditional probability distribution of φ_{12}, given the six magnitudes in its second neighborhood

Assume that the six non-negative numbers R_1, R_2 $R_{1\bar{2}/10}$, $R_{12/11}$, $R_{12/1\bar{1}}$, $R_{1\bar{2}}$ defined in (3.1), (3.2) are specified. Then, complete to terms of order $1/N$, the conditional probability distribution, $P_{1|6} = P_{1|6}(\Phi)$, of φ_{12}, given the six magnitudes of the second neighborhood, is found to be

$$\begin{aligned} P_{1|6} = \frac{1}{K_{1|6}} \exp\Bigg\{&\frac{-2(-1)^q R_1 R_2}{\sigma_2^3}\,[(3\sigma_3^2 - \sigma_2\sigma_4)R^2_{1\bar{2}/10} \\ &+ (\sigma_3^2 - \sigma_2\sigma_4)(R^2_{12/11} + R^2_{12/1\bar{1}}) - 3(\sigma_3^2 - \sigma_2\sigma_4)]\cos\Phi \\ &- \left(\frac{\sigma_3^2 - \sigma_2\sigma_4}{\sigma_2^3}\right) R_1^2 R_2^2 \cos 2\Phi\Bigg\} \\ &\times \cosh\left\{\frac{\sigma_3 R_{1\bar{2}}}{\sigma_2^{3/2}}\,[(-1)^q (R^2_{1\bar{2}/10} - 1) + 2R_1R_2\cos\Phi]\right\} \\ &\times I_0\left\{\frac{2\sigma_3}{\sigma_2^{3/2}} R_{1\bar{2}/10} R_{12/11} [R_1^2 + R_2^2 + 2(-1)^q R_1 R_2 \cos\Phi]^{1/2}\right\} \\ &\times I_0\left\{\frac{2\sigma_3}{\sigma_2^{3/2}} R_{1\bar{2}/10} R_{12/1\bar{1}} [R_1^2 + R_2^2 + 2(-1)^q R_1 R_2 \cos\Phi]^{1/2}\right\}. \end{aligned} \tag{5.1}$$

The numbers R_1, R_2, $R_{1\bar{2}/10}$, $R_{12/11}$, $R_{12/1\bar{1}}$, $R_{1\bar{2}}$ are parameters of the distribution. The normalization factor $K_{1|6}$ is best evaluated numerically, if desired, since its analytical representation is a complicated expression involving a multiple infinite series of products of Bessel functions. Details of the derivation of (5.1) are given in Appendix I.*

* Appendix I has been deposited with the British Library Lending Division as Supplementary Publication No. SUP 32949 (9 pp). Copies may be obtained through the Executive Secretary, International Union of Crystallography, 13 White Friars, Chester CH1 1NZ, England.

6. The conditional probability distribution of φ_{12}, given the 15 magnitudes in its third neighborhood

Assume that the 15 non-negative numbers R_1, R_2,..., $R_{3\bar{1}}$ defined by (3.1)–(3.3) are specified. Then, $P_{1|15} = P_{1|15}(\Phi)$, the conditional probability distribution of φ_{12}, given the 15 magnitudes (3.1)–(3.3) of the third neighborhood, is found to be

$$P_{1|15} \simeq \frac{1}{K_{1|15}}\, P_{1|6} \int_0^{2\pi} Q_1(\Phi_{23})\,\mathrm{d}\Phi_{23} \tag{6.1}$$

where

$$\begin{aligned} Q_1(\Phi_{23}) = \exp\Bigg\{&-2\left(\frac{3\sigma_3^2 - \sigma_2\sigma_4}{\sigma_2^3}\right) \\ &\times [(-1)^r R_2 R_3 R^2_{2\bar{3}/20}\cos\Phi_{23} \\ &+ (-1)^s R_3 R_1 R^2_{3\bar{1}/30}\cos(\Phi_{23} + \Phi)] \\ &-2\left(\frac{\sigma_3^2 - \sigma_2\sigma_4}{\sigma_2^3}\right)[(1-1)^r R_2 R_3 (R^2_{23/21} + R^2_{23/2\bar{1}})\cos\Phi_{23} \\ &+ (-1)^s R_3 R_1 (R^2_{31/31} + R^2_{31/3\bar{1}})\cos(\Phi_{23} + \Phi)] \\ &+6\left(\frac{\sigma_3^2 - \sigma_2\sigma_4}{\sigma_2^3}\right)[(-1)^r R_2 R_3 \cos\Phi_{23} \\ &+ (-1)^s R_3 R_1 \cos(\Phi_{23} + \Phi)] \\ &-\left(\frac{\sigma_3^2 - \sigma_2\sigma_4}{\sigma_2^3}\right)[R_2^2 R_3^2 \cos 2\Phi_{23} \\ &+ R_3^2 R_1^2 \cos(2\Phi_{23} + 2\Phi)]\Bigg\} \\ &\times \cosh\left\{\frac{\sigma_3 R_{2\bar{3}}}{\sigma_2^{3/2}}\,[(-1)^r (R^2_{2\bar{3}/20} - 1) + 2R_2R_3\cos\Phi_{23}]\right\} \\ &\times \cosh\left\{\frac{\sigma_3 R_{3\bar{1}}}{\sigma_2^{3/2}}\,[(-1)^s (R^2_{3\bar{1}/30} - 1) + 2R_3R_1\cos(\Phi_{23} + \Phi]\right\} \\ &\times I_0\left\{\frac{2\sigma_3}{\sigma_2^{3/2}} R_{2\bar{3}/10} R_{23/21} [R_2^2 + R_3^2 + 2(-1)^r R_2 R_3 \times \cos\Phi_{23}]^{1/2}\right\} \\ &\times I_0\left\{\frac{2\sigma_3}{\sigma_2^{3/2}} R_{2\bar{3}/10} R_{23/2\bar{1}} [R_2^2 + R_3^2 + 2(-1)^r R_2 R_3 \times \cos\Phi_{23}]^{1/2}\right\} \\ &\times I_0\left\{\frac{2\sigma_3}{\sigma_2^{3/2}} R_{3\bar{1}/10} R_{31/31} [R_3^2 + R_1^2 + 2(-1)^s R_3 R_1 \times \cos(\Phi + \Phi_{23})]^{1/2}\right\} \\ &\times I_0\left\{\frac{2\sigma_3}{\sigma_2^{3/2}} R_{3\bar{1}/10} R_{31/3\bar{1}} [R_3^2 + R_1^2 + 2(-1)^s R_3 R_1 \times \cos(\Phi + \Phi_{23})]^{1/2}\right\}. \end{aligned} \tag{6.2}$$

A simple expression for (6.1) appears not to exist. Hence, one must either resort to an approximate analytical expression for the integral or evaluate the expression *via* numerical techniques. With either technique the integration leads to a function of order $1/N^2$, which contains the information associated with the 'trio relation', (2.15). Therefore, the conditional distribution, $P_{1|15}$, of φ_{12} given the 15 magnitudes (3.1)–(3.3) consists of all terms of order $1/N$ plus those terms of order $1/N^2$ which reflect the information in (2.15).

7. The conditional probability distribution of φ_{12}, given the 28 magnitudes in its fourth neighborhood

Denote by $P_{1|28} = P_{1|28}(\Phi)$ the conditional probability distribution of φ_{12}, given the 28 magnitudes in the fourth neighborhood defined by (3.1)–(3.4). Then

$$P_{1|28} = \frac{1}{K_{1|28}} P_{1|6} \int_0^{2\pi} Q_1(\Phi_{23}) \int_0^{2\pi} Q_2(\Phi_{23}, \Phi_{14})\, \mathrm{d}\Phi_{14}\, \mathrm{d}\Phi_{23} \tag{7.1}$$

where

$$\begin{aligned} Q_2(\Phi_{23}, \Phi_{14}) = \exp\Bigg\{ & -\frac{2}{\sigma_2^3}(3\sigma_3^2 - \sigma_2\sigma_4) \\ & \times [(-1)^t R_1 R_4 R^2_{1\bar{4}/40} \cos \Phi_{14} \\ & + (-1)^u R_4 R_2 R^2_{4\bar{2}/50} \cos(\Phi - \Phi_{14}) \\ & + (-1)^v R_4 R_3 R_{4\bar{3}/60} \cos(\Phi + \Phi_{23} - \Phi_{14})] \\ & - \frac{2}{\sigma_2^3}(\sigma_3^2 - \sigma_2\sigma_4)\, [(-1)^t R_1 R_4 (R^2_{14/41} + R^2_{14/4\bar{1}}) \cos \Phi_{14} \\ & + (-1)^u R_4 R_2 (R^2_{42/51} + R^2_{42/5\bar{1}}) \cos(\Phi - \Phi_{14}) \\ & + (-1)^v R_4 R_3 (R^2_{43/61} + R^2_{43/6\bar{1}}) \cos(\Phi + \Phi_{23} - \Phi_{14})] \\ & + \frac{6}{\sigma_2^3}(\sigma_3^2 - \sigma_2\sigma_4)\, [(-1)^t R_1 R_4 \cos \Phi_{14} \\ & + (-1)^u R_4 R_2 \cos(\Phi - \Phi_{14}) \\ & + (-1)^v R_4 R_3 \cos(\Phi + \Phi_{23} - \Phi_{14})] \\ & - \frac{(\sigma_3^2 - \sigma_2\sigma_4)}{\sigma_2^3} [R_1^2 R_4^2 \cos 2\Phi_{14} \\ & + R_4^2 R_2^2 \cos 2(\Phi - \Phi_{14}) + R_4^2 R_3^2 \cos 2(\Phi + \Phi_{23} - \Phi_{14})] \Bigg\} \\ & \times \cosh\left\{ \frac{\sigma_3 R_{1\bar{4}}}{\sigma_2^{3/2}} [(-1)^t (R^2_{1\bar{4}/40} - 1) + 2R_1 R_4 \cos \Phi_{14}] \right\} \\ & \times \cosh\left\{ \frac{\sigma_3 R_{4\bar{2}}}{\sigma_2^{3/2}} [(-1)^u (R^2_{4\bar{2}/50} - 1) + 2R_4 R_2 \cos(\Phi - \Phi_{14})] \right\} \\ & \times \cosh\left\{ \frac{\sigma_3 R_{4\bar{3}}}{\sigma_2^{3/2}} [(-1)^v (R^2_{4\bar{3}/60} - 1) + 2R_4 R_3 \cos(\Phi + \Phi_{23} - \Phi_{14})] \right\} \\ & \times I_0\left\{ \frac{2\sigma_3}{\sigma_2^{3/2}} R_{1\bar{4}/40} R_{14/41} [R_1^2 + R_4^2 + 2(-1)^t R_1 R_4 \cos \Phi_{14}]^{1/2} \right\} \\ & \times I_0\left\{ \frac{2\sigma_3}{} [R_{1\bar{4}/40} R_{14/4\bar{1}} [R_1^2 + R_4^2 + 2(-1)^t R_1 R_4 \cos \Phi_{14}]^{1/2} \right\} \\ & \times I_0\left\{ \frac{2\sigma_3}{\sigma_2^{3/2}} R_{4\bar{2}/50} R_{42/51} [R_4^2 + R_2^2 + 2(-1)^u R_4 R_2 \times \cos(\Phi - \Phi_{14})]^{1/2} \right\} \\ & \times I_0\left\{ \frac{2\sigma_3}{\sigma_2^{3/2}} R_{4\bar{2}/50} R_{42/5\bar{1}} [R_4^2 + R_2^2 + 2(-1)^u R_4 R_2 \times \cos(\Phi - \Phi_{14})]^{1/2} \right\} \\ & \times I_0\left\{ \frac{2\sigma_3}{\sigma_2^{3/2}} R_{4\bar{3}/60} R_{43/61} [R_4^2 + R_3^2 + 2(-1)^v R_4 R_3 \times \cos(\Phi + \Phi_{23} - \Phi_{14})]^{1/2} \right\} \\ & \times I_0\left\{ \frac{2\sigma_3}{\sigma_2^{3/2}} R_{4\bar{3}/60} R_{43/6\bar{1}} [R_4^2 + R_3^2 + 2(-1)^v R_4 R_3 \times \cos(\Phi + \Phi_{23} - \Phi_{14})]^{1/2} \right\}. \end{aligned} \tag{7.2}$$

The double integral in (7.1) is also evaluated by standard numerical integration techniques, and leads to terms of order $1/N^2$ which contain the information associated with the identities (2.15), (2.25), and (2.28).

8. The applications

The figures which accompany this section show $P_{1|3}$, $P_{1|6}$ and $P_{1|15}$ as functions of Φ in the interval $-180° \le \Phi \le +180°$. They illustrate the properties of these probability distributions for a structure containing $N = 100$ identical atoms in the unit cell. The values given for the various magnitudes are mostly selected to exemplify ideal behavior of these distributions (*i.e.* to minimize their variance), and thus to illustrate the most reliable estimates of φ_{12} which are possible in a structure of this size. They confirm the plausible reasoning of §2.

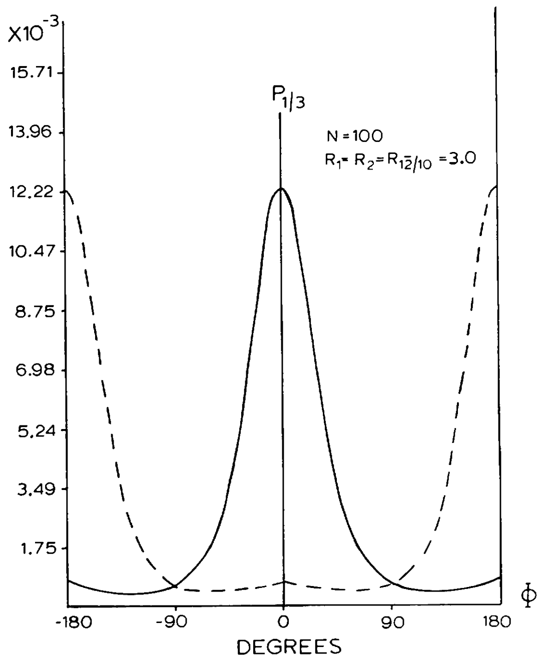

Fig. 2. The distribution $P_{1|3}$ for the values of the parameters shown. (*a*) q is even (——); (*b*) q is odd (– – –).

8.1. *The first neighborhood*

Fig. 2 shows that reliable estimates of φ_{12} occur if the three magnitudes (3.1) are large. If q is even, the most probable value of φ_{12} is zero. If q is odd the most probable value of φ_{12} is 180°.

8.2. *The second neighborhood*

Figs. 3 and 4 illustrate two kinds of distributions based on the six magnitudes of the second neighborhood. If the three magnitudes (3.1) are large and the three additional magnitudes (3.2) are also large, φ_{12} is likely to be near 0 if q is even or likely to be near $\pm 180°$ if q is odd. This is shown in Fig. 3. If the three magnitudes (3.1) are large and the three magnitudes (3.2) are small, then the most probable value of φ_{12} is 0 if q is odd and 180° if q is even. This result is shown in Fig. 4.

8.3. *The third neighborhood*

Fig. 5 shows the conditional probability distribution of φ_{12}, given that all 28 magnitudes of the third neighborhood are large. If q and $r + s$ are both even, then the most probable value of φ_{12} is 0. If, on the other hand, q and $r + s$ are both odd, then the most probable value of φ_{12} is 180°. In either of these two cases, the estimate for φ_{12} is extremely reliable. If, however, $q + r + s$ is odd, then the most probable value of φ_{12} is $\pm\alpha$ where $0 < \alpha < 180°$. Thus, if q is even and $r + s$ is odd, then $\varphi_{12} \simeq \pm 60°$; if q is odd and $r + s$ is even, then $\varphi_{12} \simeq \pm 120°$. In the latter cases the reliability of

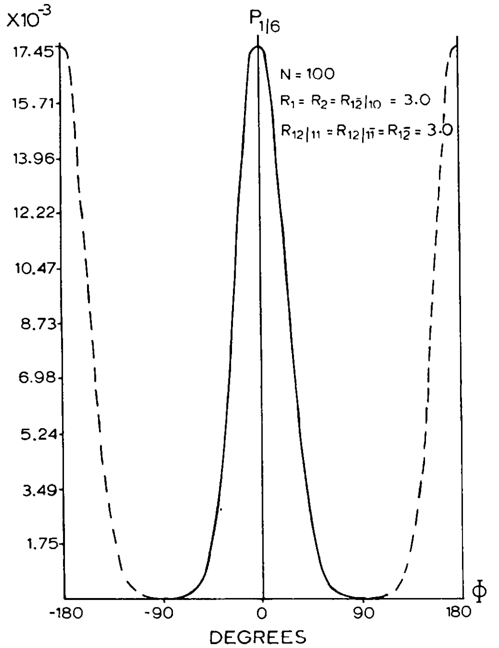

Fig. 3. The distribution $P_{1|6}$ for the values of the parameters shown. (*a*) q is even (——); (*b*) q is odd (– – –).

the estimate will in general be too low to be useful for very complex structures, as shown, for example, by the relatively large variance of Fig. 5(*c*). Fig. 6 shows the result obtained when the three 'cross-terms'

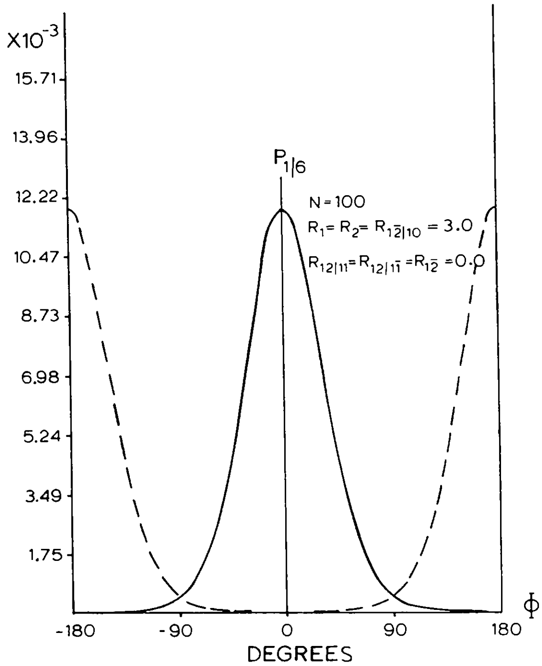

Fig. 4. The distribution of $P_{1|6}$ for the values of the parameters shown. (*a*) q is odd (——); (*b*) q is even (– – –).

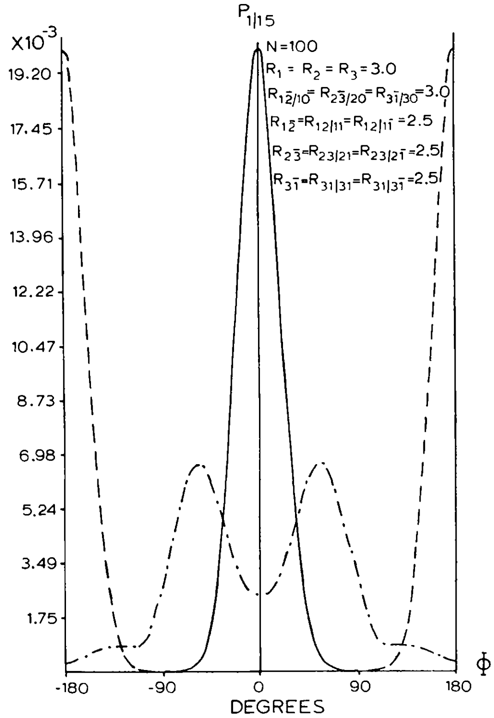

Fig. 5. The distribution of $P_{1|15}$ for the values of the parameters shown. (*a*) q and $r + s$ are both even (——); (*b*) q and $r + s$ are both odd (– – –); (*c*) q is even and $r + s$ is odd (– · – · –).

($R_{1\bar{2}}, R_{12/11}, R_{12/1\bar{1}}$, *etc.*) of the second neighborhoods of each of φ_{12} and φ_{31} are small. Reliable estimates of φ_{12} are obtained if q and $r + s$ are both odd or if q and $r + s$ are both even. In the former case $\varphi_{12} \simeq 0°$, and in the latter $\varphi_{12} \simeq 180°$. Other parity combinations of q and $r + s$ lead to estimates of φ_{12} having values between 0 and 180°, but with generally reduced reliability (*e.g.* Fig. 6*c*).

From the favorable cases considered in Figs. 2–6 it is seen that as the size of the neighborhood increases, one may obtain a more reliable estimate of φ_{12}; the larger the neighborhood, the greater is the potential for obtaining a distribution with a very small variance.

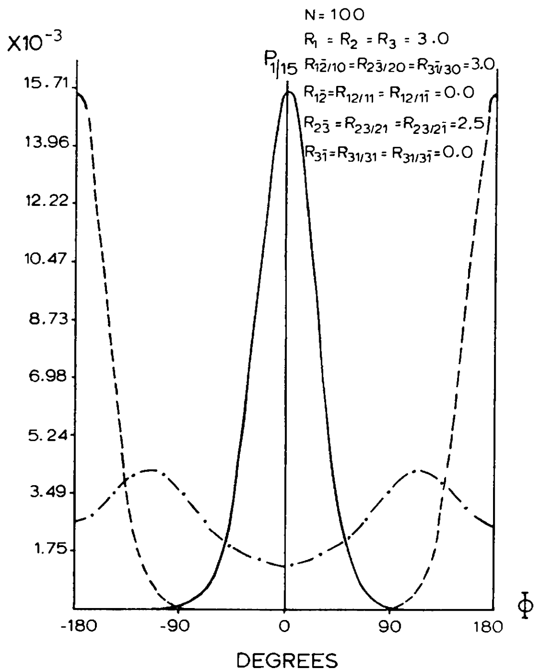

Fig. 6. The distribution $P_{1|15}$ for the values of the parameters shown. (*a*) q and $r + s$ are both odd (——); (*b*) q and $r + s$ are both even (– – –); (*c*) q is even and $r + s$ is odd (–·–·–).

In the applications one naturally selects those seminvariants and those of $q, r, s, \ldots$ which lead to distributions having the smallest possible variance, *i.e.* the favorable cases.

9. Concluding remarks

The first sequence of nested neighborhoods of the two-phase structure seminvariant φ_{12} in $P2_1$ has been found. The conditional probability distributions of φ_{12}, given, in the first instance, the three magnitudes of the first neighborhood; in the second instance, the six magnitudes of the second neighborhood; thirdly, the 15 magnitudes of the third neighborhood; and finally, the 28 magnitudes of the fourth neighborhood, have been derived. The distributions yield estimates for φ_{12} which may lie anywhere in the interval $(-\pi,\pi)$ but which are most reliable in the case that $\varphi_{12} \simeq 0$ or π. As anticipated, when more magnitudes are used more reliable estimates are obtainable, but in practice the gain in using distributions of order higher than $P_{1|6}$ may only be marginal, as Figs. 3–6 suggest.

This research was supported in party by Grant No. CHE76-17582 from the National Science Foundation, and DHEW Grants No. HL-15378 from the National Heart and Lung Institute and RR-05716-05 from the Division of Research Resources.

References

GREEN, E. A. & HAUPTMAN, H. (1978). *Acta Cryst.* A**34**, 230–241.
HAUPTMAN, H. (1972). *Acta Cryst.* B**28**, 2337–2340.
HAUPTMAN, H. (1975*a*). *Acta Cryst.* A**31**, 671–679.
HAUPTMAN, H. (1975*b*). *Acta Cryst.* A**31**, 680–687.
HAUPTMAN, H. (1977*a*). *Acta Cryst.* A**33**, 553–555.
HAUPTMAN, H. (1977*b*). *Acta Cryst.* A**33**, 556–564.
HAUPTMAN, H. (1977*c*). *Acta Cryst.* A**33**, 565–568.
HAUPTMAN, H. & GREEN, E. A. (1978). *Acta Cryst.* A**34**, 224–229.