Inter-pair symmetry

Consider the summation over six non-zero indices, where none of (u_1, u_2), (u_3, u_4) and (u_5, u_6) is a pair of equal numbers. This case is the most time-consuming part of the calculation. As explained above, the summation can be restricted to positive indices only, where $u_1 > u_2$, $u_3 > u_4$ and $u_5 > u_6$. Suppose the sixfold summation converges when the upper limit of a single summation is some number, say M, and each of the indices u_1, u_3 and u_5 ranges from 1 to M. It is easily verified that when the three $|E|$ values are equal, the above index configuration leads to some redundant calculations the result of which is invariant under a permutation of the pairs of indices. Accordingly, in order to simplify the calculation it is convenient to define new indices as

$$\nu_k = \tfrac{1}{2}(u_{2k-1} - 1)u_{2k-1} + u_{2k}, \quad k = 1, 2, 3. \qquad (A.27)$$

It is then sufficient to let u_1 range from 1 to M, u_3 from 1 to u_1 and u_5 from 1 to u_3. We further introduce multiplicity factors depending on whether ν_1, ν_2, ν_3 are all different, $\nu_1 = \nu_2 \neq \nu_3$, $\nu_1 \neq \nu_2 = \nu_3$ or $\nu_1 = \nu_2 = \nu_3$. The E-dependent Bessel functions are stored outside the summation loop in arrays of the form

$$D_j(p, \nu_k) = J_p[\pi\alpha E_j(u_{2k-1}^2 + u_{2k}^2)^{1/2}] \qquad (A.28)$$

and the triple products such as those in $(A.4)$ can be computed as

$$T_{123} = D_1(p, \nu_1)D_2(p, \nu_2)D_3(p, \nu_3), \qquad (A.29)$$

where the subscripts on T pertain to the subscripts on the $|E|$ values. In the general case, *i.e.* for unequal $|E|$'s, we can still retain the restricted ranges of the summations by symmetrizing T_{123}. This is simply achieved by computing the expression

$$\tfrac{1}{6}(T_{123} + T_{312} + T_{231} + T_{132} + T_{321} + T_{213}). \qquad (A.30)$$

Since only some of the terms require such a symmetrization, the computing effort is thus again significantly reduced.

References

Castleden, I. R. (1987). *Acta Cryst.* **A43**, 384-393.
Cochran, W. (1955). *Acta Cryst.* **8**, 473-478.
Giacovazzo, C. (1974). *Acta Cryst.* **A30**, 626-630.
Giacovazzo, C. (1980). *Direct Methods in Crystallography.* London: Academic Press.
Gradshteyn, I. S. & Ryzhik, I. M. (1980). *Table of Integrals, Series, and Products.* New York: Academic Press.
Hauptman, H. (1971). *Z. Kristallogr.* **134**, 28-43.
Hauptman, H. & Karle, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal.* ACA Monogr. No. 3. Pittsburgh: Polycrystal Book Service.
Karle, J. (1972). *Acta Cryst.* **B28**, 3362-3369.
Karle, J. & Gilardi, R. D. (1973). *Acta Cryst.* **A29**, 401-407.
Naya, S., Nitta, I. & Oda, T. (1965). *Acta Cryst.* **19**, 734-747.
Peschar, R. & Schenk, H. (1986). *Acta Cryst.* **A42**, 309-317.
Shmueli, U., Rabinovich, S. & Weiss, G. H. (1989). *Acta Cryst.* **A45**, 367-371.
Shmueli, U. & Weiss, G. H. (1985). *Acta Cryst.* **A41**, 401-408.
Shmueli, U. & Weiss, G. H. (1986). *Acta Cryst.* **A42**, 240-246.
Shmueli, U., Weiss, G. H., Kiefer, J. E. & Wilson, A. J. C. (1984). *Acta Cryst.* **A40**, 651-660.

Acta Cryst. (1989). **A45**, 367-371

Exact Conditional Distribution of a Three-Phase Invariant in the Space Group *P*1. II. Calculations and Comparison with the Cochran Approximation

By Uri Shmueli and Savely Rabinovich

School of Chemistry, The Sackler Faculty of Sciences, Tel Aviv University, 69 978 *Tel Aviv, Israel*

and George H. Weiss

Division of Computer Research and Technology, National Institutes of Health, Bethesda, Maryland 20892, *USA*

(*Received* 7 *June* 1988; *accepted* 3 *January* 1989)

Abstract

The conditional probability density function of a three-phase invariant is computed from exact expressions derived and discussed in the first paper of this series [Shmueli, Rabinovich & Weiss (1989). *Acta Cryst.* **A45**, 361-367] and comparisons are presented of these computations with the approximate conditional density due to Cochran [*Acta Cryst.* (1955), **8**, 473-478]. Conditional variances computed from the exact and approximate expressions are also compared. The computations are carried out for the space group *P*1. This is the first numerical comparison of conditional phase-invariant statistics evaluated from exact and approximate expressions. The discrepancy between these two kinds of statistics appears to be negligible if the E values involved are small and the number of atoms in the cell is moderately

0108-7673/89/060367-05$03.00

large. Very significant discrepancies are observed when the number of the atoms in the unit cell is small, as well as in the case of the presence of an outstandingly heavy atom. For a given triplet of magnitudes of E values or a fixed atomic composition of the unit cell the exact conditional density functions tend to have sharper peaks than the approximate ones.

Introduction

The triple product of normalized structure factors

$$E(hkl)E(h'k'l')E(-h-h', -k-k', -l-l')$$

or briefly $E(\mathbf{h})E(\mathbf{k})E(-\mathbf{h}-\mathbf{k})$ is one of the most important relationships in the theory and practice of direct methods of phase determination. These methods rely rather heavily on probabilistic approaches. In particular, the conditional probability density function (c.p.d.f.) of the three-phase structure invariant

$$\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{-\mathbf{h}-\mathbf{k}}\equiv\Phi \qquad (1)$$

is of interest; the conditioning is here on the knowledge of the corresponding magnitudes $|E_{\mathbf{h}}|$, $|E_{\mathbf{k}}|$ and $|E_{-\mathbf{h}-\mathbf{k}}|$, hereafter denoted by E_1, E_2 and E_3, respectively.

The c.p.d.f. in general use is that due to Cochran (1955) which, in a slightly modified form (*e.g.* Giacovazzo, 1980), reads

$$p(\Phi|E_1, E_2, E_3)\simeq[2\pi I_0(\kappa)]^{-1}\exp(\kappa\cos\Phi), \qquad (2)$$

where

$$\kappa=2\sigma_3\sigma_2^{-3/2}E_1E_2E_3, \qquad \sigma_n=\sum_{j=1}^{N}Z_j^n$$

and $I_0(x)$ is a modified Bessel function of the first kind. Various approximations to (2) have been published over the years in terms of series expansions in orthogonal polynomials (*e.g.* Naya, Nitta & Oda, 1965; Hauptman, 1971; Giacovazzo, 1974), as well as exponentiated forms of these series (Karle, 1972; Karle & Gilardi, 1973; Peschar & Schenk, 1986). However, since no exact and computable results have appeared in the literature, it has so far been impossible to assess the accuracy of the Cochran approximation and its generalizations.

In paper I of this series (Shmueli, Rabinovich & Weiss, 1989; hereafter SRW) we gave an exact representation for the c.p.d.f. of Φ based on its expansion into a sixfold Fourier series. The coefficients of this series have been shown to be expressible in terms of characteristic functions, and we have pointed out a number of techniques that allow us to make accurate numerical calculations of $p(\Phi|E_1, E_2, E_3)$. In the present paper we examine the accuracy of the Cochran approximation, paying particular attention to three factors that are important in practice: (1) the effect of a small number of atoms in the unit cell; (2) dependence of $p(\Phi|E_1, E_2, E_3)$ on the individual values of the E's; and (3) the presence of an outstandingly heavy atom. We will see that, in agreement with our earlier studies comparing approximate and exact values of the c.p.d.f.'s required for the use of methods based on $\sum_1$ and $\sum_2$ (Shmueli & Weiss, 1985, 1986), approximations based on the central limit theorem produce probabilities that may be significantly lower than those found from our more accurate calculations.

Calculations and results

The results presented below were obtained from hypothetical atomic contents in a unit cell of the space group $P\bar{1}$. The computation was carried out using Fortran programs based on appropriately symmetrized versions of (10) in SRW, with a Cyber 180-990 computer and a NOS/VE operating system. The use of this system allowed us to take advantage of virtual memory and to work with large arrays that may be used to eliminate many repetitive calculations. The expressions found in paper I involve two kinds of summations: (i) those associated with the Fourier series and (ii) those associated with the inner expansion of exponentials in series of Bessel functions. Each of the results presented below is accompanied by a fraction $N1/N2$, where $N1$ is the number of terms in a single Fourier summation required for its approximate convergence, and $N2$ is the corresponding number of terms in a series of Bessel functions [*e.g.* (17) in SRW]. The results of most computations of the c.p.d.f. of Φ are presented in the 0-90° range, at 5° intervals (more detailed computations, not presented here, show that the c.p.d.f. in the 90-180° range is flat and featureless).

In addition to our calculation of the c.p.d.f.'s of Φ as function of the relevant parameters (*i.e.* atomic composition and $|E|$ values) we have also computed the conditional variance of Φ from the expression derived in Appendix *A*. These conditional variances are compared with those based on the Cochran approximation given in (2) (Karle & Karle, 1966) in Table 1.

An indication of the effect of the magnitudes of the $|E|$ values on the c.p.d.f. of a three-phase invariant is found in the comparison of exact and approximate results in Fig. 1. Fig. 1(*a*) compares the exact c.p.d.f. and the approximation (2) for $E_1=E_2=E_3=1\cdot50$, for an equal-atom structure containing 40 atoms in the unit cell of $P\bar{1}$. The comparison shown in Fig. 1(*b*) differs from that in Fig. 1(*a*) only in that the three $|E|$ values are each equal to $2\cdot50$. The agreement between the exact and approximate results shown in Fig. 1(*a*) is very good - although the approximate c.p.d.f. tends to underestimate slightly the exact result in the region (0-60°) of Φ. The discrepancy between exact and approximate c.p.d.f.'s is considerably

Table 1. *Conditional variance of the three-phase invariant* Φ

The table compares the square roots of the conditional variances of Φ, computed from the exact and approximate expressions. Each row of the table pertains to one of the figures in this paper. The first column contains the number of a figure, the second column contains the square root of $\langle\Phi^2|E_1, E_2, E_3\rangle$ computed from a symmetrized version of (A.3) and denoted by $\sigma_{ex}(\Phi)$ (in degrees), and the third column contains the square root of the corresponding conditional variance of Φ computed from the Cochran (1955) c.p.d.f. with the formula given by Karle & Karle (1966) and denoted by $\sigma_{Co}(\Phi)$ (in degrees).

Figure	$\sigma_{ex}(\Phi)$	$\sigma_{Co}(\Phi)$
1(*a*)	67·632	70·689
1(*b*)	21·987	27·504
2(*a*)	8·634	22·371
2(*b*)	13·820	25·046
3	8·600	18·415

greater in Fig. 1(*b*). This discrepancy is also evident in the comparison of corresponding conditional variances given in Table 1.

Far more drastic effects are caused by having a small number of atoms in the unit cell. The results displayed in Fig. 2 are for equal-atom structures for the specific normalized structure factors $E_1 = 2{\cdot}00$, $E_2 = 2{\cdot}25$ and $E_3 = 2{\cdot}50$; Fig. 2(*a*) compares the exact and approximate c.p.d.f.'s for ten atoms in the unit cell, and Fig. 2(*b*) for 15 atoms in the unit cell. The exact c.p.d.f. of Φ in Fig. 2(*a*) is much sharper than the approximate c.p.d.f. (2). This is of course consistent with the increased certainty of phase determination as one approaches the region of validity of inequalities derived from Hauptman–Karle determinants (*e.g.* Giacovazzo, 1980). The discrepancy in Fig. 2(*b*) is also significant; however, it is appreciably smaller than that in Fig. 2(*a*). A qualitatively similar result was obtained in comparing an exact c.p.d.f. for $\sum_1$ with an approximate c.p.d.f., in examining the effects of the paucity of atoms on these statistics (Shmueli & Weiss, 1985).

Fig. 3 illustrates the effect of the presence of an outstandingly heavy atom on the discrepancy between the exact and approximate c.p.d.f.'s. The effect of atomic heterogeneity on intensity (*e.g.* Shmueli,

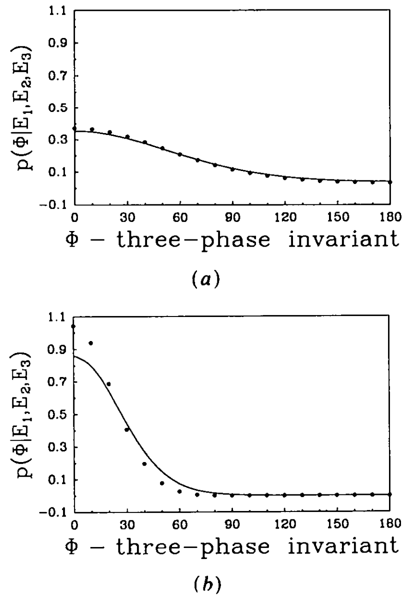

Fig. 1. Effect of the magnitude of E on the c.p.d.f. of Φ. Comparison of c.p.d.f.'s computed from the exact expression of SRW (black circles) and modified Cochran's (1955) approximation (2) (solid line) for 40 equal atoms in the unit cell of $P1$. The angle Φ is given in degrees. (*a*) $E_1 = E_2 = E_3 = 1{\cdot}50$, $N1 = 10$, $N2 = 9$. (*b*) $E_1 = E_2 = E_3 = 2{\cdot}50$, $N1 = 15$, $N2 = 9$.

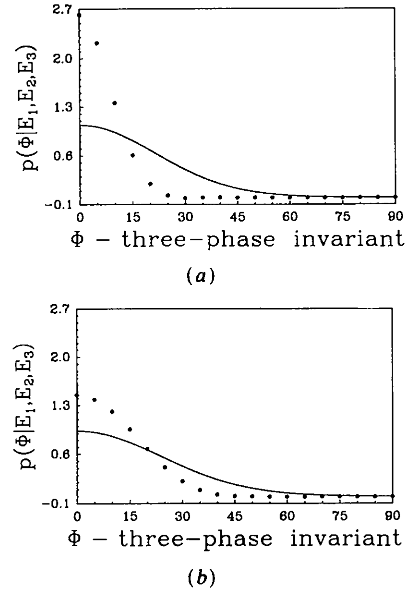

Fig. 2. Effect of the paucity of atoms on the c.p.d.f. of Φ. Comparison of c.p.d.f.'s computed from the exact expression of SRW (black circles) and modified Cochran's (1955) approximation (2) (solid line). The figure is based on the magnitudes: $E_1 = 2{\cdot}00$, $E_2 = 2{\cdot}25$, $E_3 = 2{\cdot}50$. The angle Φ is given in degrees. (*a*) Ten equal atoms in the unit cell of $P1$, $N1 = 17$, $N2 = 12$. (*b*) Fifteen equal atoms in the unit cell of $P1$, $N1 = 10$, $N2 = 9$.

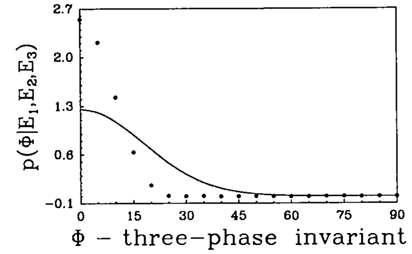

Fig. 3. Atomic heterogeneity and the c.p.d.f. of Φ. Comparison of c.p.d.f.'s computed from the exact expression of SRW (black circles) and modified Cochran's (1955) approximation (2) (solid line) for the hypothetical composition $C_{14}Br$ of the unit cell of $P1$. All three E's are assumed to equal 2·00, and the summation ranges are $N1 = 15$ and $N2 = 9$. The angle Φ is given in degrees.

Weiss, Kiefer & Wilson, 1984) and phase (*e.g.* Shmueli & Weiss, 1986) statistics was emphasized in our earlier studies, and we thought it of interest to examine this problem in the present work as well. The effect of the heavy atom is to sharpen the c.p.d.f. (and decrease the conditional variance) relative to those obtained from Cochran's c.p.d.f. (see Fig. 3 and Table 1). This is qualitatively similar to the effect caused by a small number of identical atoms in the unit cell. It follows that the use of Cochran-type c.p.d.f.'s in the presence of a small number of atoms in *P*1, or else a strongly heterogeneous atomic composition therein, significantly underestimates conditional probabilities for the three-phase invariant. This conclusion is consistent with those to which we were led in our previous studies of exact c.p.d.f.'s (Shmueli & Weiss, 1985, 1986). We point out that the heterogeneity examined here is relatively small and higher atomic numbers of the heavy atom will lead to a further sharpening of the c.p.d.f. However, few such calculations were carried out since high heterogeneities (just as very small numbers of atoms, *i.e.* less than 10) lead to difficulties in the convergence of the Fourier series for the c.p.d.f. - at least with the present version of our computer programs.

Concluding remarks

The three-phase invariant has been studied by a number of investigators because of its great importance in applications to direct methods of phase determination. However, all these studies are based on the central limit theorem and its extensions. Without an exact theory together with numerical evaluation of the results of the theory it has been impossible to assess the accuracy of Cochran's (1955) approximation, or any of its generalizations.

In the present paper we give the first accurate numerical evaluation of exact expressions for the c.p.d.f. of a three-phase invariant. Because of the complexities of the calculations the techniques that we have used are neither ripe nor essential for routine practical applications. Even with the many simplifications of the algorithms resulting from inherent symmetries in our expressions (Shmueli, Rabinovich & Weiss, 1989), computing times are still of the order of 10 min per invariant thereby rendering the computations too expensive for routine use.

The main merit of the present results is that they allow one to test the accuracy of various approximations to the exact c.p.d.f., and assess the circumstances under which currently available approximate methods deviate considerably from the conditional densities. The variances displayed in Table 1 show that the exact c.p.d.f.'s are often much sharper than the approximate ones, thus emphasizing the conservative character of the latter. This is concurrent with our earlier results on problems of sign determination (Shmueli & Weiss, 1985, 1986). The discrepancy appears to be negligible when the number of atoms in the unit cell is large and the degree of atomic heterogeneity is small. Our results indicate that when either of these conditions is violated this discrepancy can be most significant.

This research was supported in part by grant No. 84-00076 from the United States-Israel Binational Science Foundation (BSF), Jerusalem, Israel. All the computations related to this paper were carried out at the Tel Aviv University Computation Center on a Cyber 180-990 computer.

APPENDIX *A*

Derivation of the exact conditional variance of Φ

Our starting point for this derivation is equation (10) of SRW, with the various quantities defined by equations (11), (17) and (9) of that paper. If we collect the terms depending on Δ and $-\Delta$, we can write the c.p.d.f. of Φ as

$$p(\Phi|E_1, E_2, E_3) = K'\sum_{\mathbf{u}} C_{\mathbf{u}} Z'_{\mathbf{u}}, \qquad (A.1)$$

where $\mathbf{u}^T = (u_1, u_2, u_3, u_4, u_5, u_6)$ is the vector of summation indices, $C_{\mathbf{u}}(\Delta) = C_{\mathbf{u}}(-\Delta)$ are the Fourier coefficients defined by (17) in SRW, $Z'_{\mathbf{u}}$ is given by

$$\begin{aligned} Z'_{\mathbf{u}} = \mathbf{G}_0 &+ 2\sum_{p=1}^{\infty} (-1)^p \mathbf{G}_{2p} \cos(2p\Delta)\cos(2p\Phi) \\ &- 2i\sum_{p=1}^{\infty} (-1)^p \mathbf{G}_{2p-1} \cos[(2p-1)\Delta] \\ &\times \cos[(2p-1)\Phi], \end{aligned} \qquad (A.2)$$

and the quantities $\mathbf{G}_m$ are triple products of Bessel functions defined by (A.4) in SRW; the normalization constant K' is given by (11) in the latter reference.

It can be easily shown that the conditional mean, $\langle\Phi|E_1, E_2, E_3\rangle$ vanishes. Hence, straightforward integration shows that the conditional variance of Φ is given by

$$\begin{aligned} \langle\Phi^2|E_1, E_2, E_3\rangle = \pi^2/3 + 8\pi K'\mathcal{R}e\sum_{\mathbf{u}} C_{\mathbf{u}} \sum_{p=1}^{\infty} (-1)^p \\ \times (K_{2p}\mathbf{G}_{2p} + iK_{2p-1}\mathbf{G}_{2p-1}), \end{aligned} \qquad (A.3)$$

where

$$K_m = \cos(m\Delta)/m^2. \qquad (A.4)$$

The conditional variance of the Cochran (1955) c.p.d.f. has been given by Karle & Karle (1966). Note that the leading $\pi^2/3$ term appears in both the exact and approximate formulations of the variance.

References

Cochran, W. (1955). *Acta Cryst.* **8**, 473-478.
Giacovazzo, C. (1974). *Acta Cryst.* **A30**, 626-630.
Giacovazzo, C. (1980). *Direct Methods in Crystallography.* London: Academic Press.
Hauptman, H. (1971). *Z. Kristallogr.* **134**, 28-43.
Karle, J. (1972). *Acta Cryst.* **B28**, 3362-3369.
Karle, J. & Gilardi, R. D. (1973). *Acta Cryst.* **A29**, 401-407.
Karle, J. & Karle, I. L. (1966). *Acta Cryst.* **21**, 849-859.
Naya, S., Nitta, I. & Oda, T. (1965). *Acta Cryst.* **19**, 734-747.
Peschar, R. & Schenk, H. (1986). *Acta Cryst.* **A42**, 309-317.
Shmueli, U., Rabinovich, S. & Weiss, G. H. (1989). *Acta Cryst.* **A45**, 361-367.
Shmueli, U. & Weiss, G. H. (1985). *Acta Cryst.* **A41**, 401-408.
Shmueli, U. & Weiss, G. H. (1986). *Acta Cryst.* **A42**, 240-246.
Shmueli, U., Weiss, G. H., Kiefer, J. E. & Wilson, A. J. C. (1984). *Acta Cryst.* **A40**, 651-660.

Acta Cryst. (1989). **A45**, 371-380

Comparison Between Gaussian and Exponential Charge Distributions in Ewald Surface Potentials and Fields: NaCl, Aragonite, Phlogopite

By C. S. Strom and P. Hartman

Institute of Earth Sciences, State University of Utrecht, PO Box 80021, 3508 TA Utrecht, The Netherlands

(*Received 29 September 1987; accepted 3 January 1989*)

Abstract

A comparative study has been undertaken of the employment of Gaussian and exponential charge distribution functions in calculating Coulomb potentials, energies and fields at arbitrary points due to lattice slices, using the Ewald method. The APL program *SURFPOT* has been developed for this purpose, for a general crystal structure with user-defined slice orientations and slice boundary configurations. Results are presented for the (111) face of NaCl, (001) and (112) faces of aragonite ($CaCO_3$) and (001) face of phlogopite ($KM_3T_4O_{10}(OH)_2$, M = divalent cation, $T_4 = Si_3Al$]. The convergence behaviour of the potential sums is consistently and considerably better when the Gaussian form is used.

1. Introduction

Since the predominant mechanism of crystal growth is growth by slices, the availability of generally applicable methods for calculating potentials, electric fields, energies and interaction energies of lamina-shaped lattices is essential to the study of crystal growth. The slices under consideration are parallel to (*hkl*), have infinite extent in two dimensions and thickness d_{hkl}, or some fraction thereof, in the third. They are charge-neutral and usually nonpolar. The points at which potentials and fields are required are arbitrary and may be outside the slice or within it or coinciding with ion sites. Formulating the problem for the face (001) instead of (*hkl*) does not constitute any real limitation. A new unit cell can always be found, such that the required face (*hkl*) in terms of the original cell parameters becomes transformed to (001) in terms of the new.

The most widespread method for calculating Coulomb potentials in ionogenic structures is the method of Ewald (1921). It is essentially based on the introduction of a continuous spherically symmetric charge distribution function σ dependent on an adjustable parameter η, in addition to the distance from a given ion (Tosi, 1964). Thus the charge density, and hence also the potential, are split into two contributions; one formulated in the direct lattice, the other in the reciprocal. Each potential sum converges independently but depends on η, whereas the total potential is independent of η. Thus η functions as the radius of a convergence sphere and is adjusted so as to optimize the convergence of both sums. As will be shown in § 2, neutrality of the unit cell guarantees convergence of the potential expression for an individual slice.

The Ewald method assumes its simplest form in the case of an infinite lattice but has been extended to laminas and to semi-infinite lattices, which are infinite only on one side of a plane, the other side being empty (Hartman, 1958; Heyes & van Swol, 1981). The lamina case is most useful because the laminas may be stacked at will. Should an infinite or semi-infinite lattice be of interest, then a limited number of slices need to be stacked to obtain the most satisfactory approximation, *i.e.* convergence in the third dimension, provided the chosen unit cell possesses no component of dipole moment perpendicular to (*hkl*).

The traditionally chosen functional dependence of σ on the distance from a given ion is Gaussian, but several other possibilities, including the exponential, have been proposed and applied, mostly to rather simple structures. Such an application to a cubic