The distribution of interatomic distances, By J. ALBERTSSON* and P. M. SCHULTHEISS, *Becton Center, Yale University, New Haven, Connecticut* 06520, *U.S.A.*

The use of normal-probability-plot analysis to compare independently determined molecular geometries [De Camp, W. H. (1973). *Acta Cryst.* A**29**, 148–150] is justified by a calculation of the distribution of interatomic distances in a crystal structure under some simplifying conditions, which in most cases are at least approximately obeyed. The distribution is shown to be normal to a good approximation in spite of the non-linear transformation applied to the atomic coordinates in calculating the distances.

De Camp (1973) suggested that the use of probability-plot analysis (Abrahams & Keve, 1971) could be extended to the comparison of independently determined molecular geometries, each described by all of its independent bonded and non-bonded interatomic distances up to a certain limit. Examples of the use of the method can be found in De Camp's paper and in papers by Albertsson, Grenthe & Herbertsson (1973) and Oskarsson (1974). The technique requires that the differences between the compared distances are normally distributed. However, in calculating the interatomic distances one performs a non-linear transformation on the atomic coordinates. Even if the latter are normal

$$dF(x) = \begin{cases} \dfrac{x}{\sqrt{2\pi\lambda}} \{\exp[-(x-\sqrt{\lambda})^2/2] - \exp[-(x+\sqrt{\lambda})^2/2]\} \, dx & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (5)$$

variates, such a transformation causes the distribution of the distances, and hence of the difference of distances, to become non-normal.

Our first objective is therefore to find the distribution $dF(x)$ of an interatomic distance x in a crystal structure

$$dF(x) = \begin{cases} f(x)dx & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (1)$$

and to see how far from normal it is. The distribution $dH(y)$ of the difference y between two independent distances is then obtained as

$$dH(y) = \left\{ \int_0^\infty f_1(y+t)f_2(t)dt \right\} dy \quad (2)$$

where $f_1(x)$ and $f_2(x)$ are the frequency functions for the two distances (Kendall & Stuart, 1963). If the corresponding distributions $dF_1(x)$ and $dF_2(x)$ are normal, $dH(y)$ is also normal.

We assume that the atomic coordinates are obtained by least-squares refinement using a large sample of intensity data. The central-limit theorem then ensures that they are normal variates. To simplify our calculations we also make the following assumptions: (i) there is no correlation between coordinates of different atoms, (ii) the only correlation within each triple of coordinates is due to the possible inclination of the axes, and (iii) the variance of the position of each atom is isotropic. For a discussion of these assumptions see, *e.g.*, Templeton (1959). We are now able to make a linear coordinate transformation so that the squared interatomic distance z is the sum of the squares of three independent normal variates with non-zero mean and unit variance

$$z = x_1^2 + x_2^2 + x_3^2 . \quad (3)$$

* Present address: Inorganic Chemistry 1, Chemical Center, University of Lund, P.O.B. 740, S-220 07 Lund, Sweden.

The distribution $dG(z)$ of this statistic is the non-central χ^2 distribution with three degrees of freedom (Kendall & Stuart, 1961)

$$dG(z) = \begin{cases} \dfrac{1}{\sqrt{2\pi\lambda}} \exp[-(z+\lambda)/2] \sinh(\sqrt{z\lambda})dz & z \geq 0 \\ 0 & z < 0 . \end{cases} \quad (4)$$

The parameter λ is the sum of the squared means of x_1, x_2, and x_3. By differentiating $G(z)$ with respect to $x = \sqrt{z}$ we obtain

To the order of $1/\sqrt{\lambda}$ this distribution gives the mean $\sqrt{\lambda} + 1/\sqrt{\lambda}$ and unit variance for the variate x. The parameter λ is measured in units of the variance of each of the three normal variates in equation (3). For a reasonably well determined structure this unit is less than $3 . 10^{-3}$ Å^2 while the squared interatomic distances all are larger than 1 Å^2. Thus $\sqrt{\lambda} \gg 1$ and to a good approximation we can neglect the term $\exp[-(x+\sqrt{\lambda})^2/2]$ in equation (5) and take $\sqrt{\lambda}$ as the mean of x. After a rescaling to the usual unit of length we arrive at

$$dF(x) = \begin{cases} \dfrac{x}{d} \cdot \dfrac{1}{\sqrt{2\pi}\,\sigma} \exp[-(x-d)^2/2\sigma^2] \, dx & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (6)$$

as an approximate distribution of the interatomic distances. In equation (6) d is the mean and σ^2 the variance of x. Since $\sigma \ll d$, $\exp[-(x-d)^2/2\sigma^2]$ differs substantially from zero only in a small interval around $x = d$ and the resulting distribution of x is essentially normal. As a consequence the sought distribution of the difference between two independent distances is also very nearly normal. In fact, evaluation of the integral (2) using the complete frequency function of equation (5), assuming the same mean d and variance σ^2 for the two distances, results in the following expression, correct to the second order in σ/d*

* In terms of the parameter λ the exact expression is

$$dH(y) = \frac{1}{2\pi\lambda} \left\{ \frac{\sqrt{\lambda}}{2} \exp(-y^2/4) [\exp(-(|y|-2\sqrt{\lambda})^2/4) \right.$$
$$- \exp(-(|y|+2\sqrt{\lambda})^2/4]$$
$$+ \exp(-y^2/4) [\sqrt{\pi}(\tfrac{1}{2}+\lambda-y^2/4)(2-\mathrm{erf}(|y|/2-\sqrt{\lambda})$$
$$- \mathrm{erf}(|y|/2+\sqrt{\lambda})] - [\exp(-(|y|-2\sqrt{\lambda})^2/4)$$
$$\left. + \exp(-(|y|+2\sqrt{\lambda})^2/4]\sqrt{\pi}(\tfrac{1}{2}-y^2/4)(1-\mathrm{erf}(|y|/2)) \right\} dy$$

where

$$\mathrm{erf}(x) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^x \exp(-t^2)dt.$$

$$dH(y) = \frac{1}{2\sqrt{\pi}\,\sigma} [\exp(-y^2/4\sigma^2)] [1 + (2\sigma^2 - y^2)/4d^2] dy . \quad (7)$$

In this case the variance of the difference is $2\sigma^2$. As long as $\sigma \ll d$ the correction term is neglible and for all practical purposes the distribution $dH(y)$ can be treated as normal.

In calculating the distributions we have assumed that the atomic coordinates of the actual crystal structure are uncorrelated normal variates with small isotropic errors. In practice this condition is often rather well met and at least in these cases the distribution of interatomic distances is normal to a good approximation. The difference between two independently determined interatomic distances is then normally distributed to an even better approximation than the distances themselves [cf. equations (6) and (7)]. We therefore conclude that in most, if not all cases crystallographically independent molecular geometries can be compared using the powerful method of normal probability plot analysis, as suggested by De Camp (1973).

References

ABRAHAMS, S. C. & KEVE, E. T. (1971). *Acta Cryst.* A27, 157–165.
ALBERTSSON, J., GRENTHE, I. & HERBERTSSON, H. (1973). *Acta Cryst.* B29, 2839–2844.
DE CAMP, W. H. (1973). *Acta Cryst.* A29, 148–150.
KENDALL, M. G. & STUART, A. (1961). *The Advanced Theory of Statistics*, Vol. 2, pp. 227–229. New York: Hefner.
KENDALL, M. G. & STUART, A. (1963). *The Advanced Theory of Statistics*, Vol. 1, 2nd ed., p. 263. New York: Hefner.
OSKARSSON, Å. (1974). To be published.
TEMPLETON, D. H. (1959). *Acta Cryst.* 12, 771–773.

On symmetric variations of fault energies.

By A. G. CROCKER and P. D. BRISTOWE, *Department of Physics, University of Surrey, Guildford GU2 5XH, England*

The distribution of interatomic distances across unrelaxed stacking faults is shown to be independent of the sense of the fault vector if the shear plane is perpendicular to an n-fold axis or if the shear direction is parallel to an n-fold axis. Hence the unrelaxed energies of such faults, calculated using pair potentials, are symmetric with respect to the sense of the shear.

In a recent computer simulation study of generalized stacking faults in body-centred cubic crystals (Bristowe, Crocker & Norgett, 1974) it was found that the unrelaxed energy of the faults was independent of the sense of the shear direction. The faults were all on (112) planes and the displacements were in $\pm[11\bar{1}]$ directions. In addition the interatomic pair potential used to evaluate their energy terminated between second and third nearest neighbours. The aim of the work was to determine the structure and energy of the faults which arise on allowing the atoms to relax, subject to the pair potential, from their initial sheared positions to equilibrium configurations. However, the symmetric variation of the unrelaxed energy has also proved to be of interest. This energy is the sum of the pair interactions between atoms which, for a given potential $\varphi(r)$, are defined by the interatomic spacings r. Thus a symmetric variation of fault energy indicates that positive and negative rigid-body shears produce identical distributions of distorted bond lengths. This is illustrated for the special case of the b.c.c. fault in Fig. 1. It suggests the following general problem which will be the subject matter of the present note: Given a crystal which undergoes a rigid-body shear on a plane $h_l = (h_1 h_2 h_3)$ in a direction $u^i = [u^1 u^2 u^3]$ under what conditions are the distributions of distorted bond lengths for positive and negative shears identical?

Consider first a bond defined by the lattice vector $p^i = [p^1 p^2 p^3]$ in the undistorted crystal. If the shear plane h_l cuts this bond the vector becomes $p^i \pm au^i$ in the faulted crystal. Here a is a parameter defining the magnitude of the displacement and the two signs correspond to the two senses

of the shear. The length of the bond is thus changed from $l = (p^i p^j c_{ij})^{1/2}$ to

$$l^\pm = [(p^i p^j + a^2 u^i u^j \pm 2ap^i u^j) c_{ij}]^{1/2}$$

where c_{ij} is the direct metric tensor. Clearly $l^+ \neq l^-$ and thus if the collected bond lengths for opposite shears are to be identical another vector q^i must exist which gives rise to distorted bond lengths $L^\pm$ satisfying $L^\pm = l^\mp$ for all values of a. For this to occur p^i and q^i must be crystallographically equivalent variants r_1^i and r_2^i, of a particular vector form r^i,
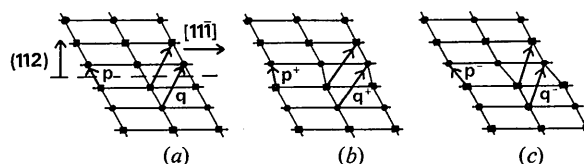

Fig. 1. Faults on a (112) plane in a b.c.c. crystal projected on to a (1$\bar{1}$0) plane, atoms represented by circles and squares lying in adjacent (1$\bar{1}$0) planes. The nearest-neighbour bonds $\mathbf{p} = \frac{1}{2}[1\bar{1}1]$ and $\mathbf{q} = \frac{1}{2}[111]$ shown in the perfect crystal (*a*) are decreased in length and increased in length respectively by the shear in the [11$\bar{1}$] direction shown in (*b*). These changes in length of $\mathbf{p}$ and $\mathbf{q}$ in (*b*) are exactly equal to the changes in length of $\mathbf{q}$ and $\mathbf{p}$ respectively arising from the equal and opposite shear shown in (*c*). Note that the fault plane intersects two bonds parallel to $\mathbf{q}$ but only one parallel to $\mathbf{p}$ for each atom in the interface. However in this projection the vector $\mathbf{p}$ may also be considered to represent the nearest-neighbour bond $\frac{1}{2}[\bar{1}11]$ so that equal numbers of bonds are in fact extended and contracted.