Acta Cryst. (1979). A**35**, 765–772

Statistical Properties of Normalized Difference-Structure Factors for Non-Centrosymmetric Structures

By Paul T. Beurskens, Peter A. J. Prick and H. M. Doesburg

Crystallography Laboratory, University of Nijmegen, Toernooiveld, 6525 ED Nijmegen, The Netherlands

and R. O. Gould

Chemistry Department, University of Edinburgh, West Mains Road, Edinburgh EH9 3JJ, Scotland

(*Received* 22 *November* 1978; *accepted* 2 *April* 1979).

Abstract

The probability density of phase angles for structures with one or more atoms on known positions [Sim (1959). *Acta Cryst.* **12**, 813–818] is expressed in terms of normalized difference-structure factors, and used to calculate standard deviations of phases and expectation values for amplitudes ($|E_r|$) of the normalized difference-structure factors. Numerical results are tabulated for various values of $|E_1|$ and $|E_2|$, *i.e.* the minimum and maximum $|E_r|$ value a given reflection can have. Applications to the *DIRDIF* procedures [van den Hark, Prick & Beurskens (1976). *Acta Cryst.* A**32**, 816–821] are described. New applications are the calculation of the *a priori* probability density function for $|E_r|$ values, and a statistical method for the detection of a centre of symmetry in the remaining part of the structure. For the statistical method, a centricity parameter, X_c, which is unity for centric and zero for acentric distributions, is defined; for the difference structure, X_C is calculated by an iterative procedure, extrapolating towards zero contribution of the known part of the structure. Numerical results for 13 test structures are given.

1. Introduction

In our *DIRDIF* procedures (van den Hark, Prick & Beurskens, 1976 and references therein), direct methods are applied to the solution of partially known structures, usually containing heavy atoms. Difference structure factors are calculated, making the usual assumption that the observed and calculated structure factors have the same phase. These are then 'normalized', and both amplitudes and phases are refined by a weighted tangent procedure.

Applications of the *DIRDIF* procedures prompted the present study. In particular, when the distribution of known atoms is centrosymmetric, calculation of the distribution of expectation values for normalized difference-structure factors may indicate whether or not the total structure is centrosymmetric. The statistical properties of difference structure factors in centrosymmetric structures have been described (Woolfson, 1956; Gould, van den Hark & Beurskens, 1975).

The discussion here is for space group $P1$, but is applicable to general reflections in other space groups if the symmetry enhancement factor ε is included.

2. Notation

The known part of the structure consists of N_p atoms whose positions in the unit cell are known. This is the *partial structure*. The remaining N_r atoms of unknown position, form the *rest* or *difference structure*. The *total structure*, then, has $N = N_p + N_r$ atoms per unit cell.

For any reflection **h**, F, F_p and F_r represent the structure factors for the total, partial and rest structures, respectively. Of these, F_p with phase φ_p may be calculated, and is usually denoted by $F_{\rm calc}$; $|F|$, usually denoted by $|F_{\rm obs}|$ is known experimentally; F_r is unknown. The relationship

$$F = F_p + F_r \qquad (1)$$

cannot be solved, as the phase of F is unknown.

Normalized structure factors may be approximated by:

$$E = \left(\sum_{j=1}^{N} Z_j^2\right)^{-1/2} \sum_{j=1}^{N} Z_j \exp(2\pi i \mathbf{h}.\mathbf{r}_j), \qquad (2)$$

where Z, the atomic number, is used in place of f, the scattering factor. Similar expressions may be written to define E_p and E_r, the sums being over N_p and N_r atoms, respectively.

For any reflection **h**, $|E|$ is known experimentally and E_p is calculated using the known atomic positions.

It is possible to write, analogous to (1)

$$E = pE_p + rE_r, \qquad (3)$$

0567-7394/79/050765-08$01.00

where

$$p^2 = \sum_{j=1}^{N_p} Z_j^2 \Big/ \sum_{j=1}^{N} Z_j^2, \tag{4}$$

and r is defined analogously to make $p^2 + r^2 = 1$. For equal atom structures, (4) reduces to $p^2 = N_p/N$ and $r^2 = N_r/N$. The quantities p^2 and r^2 are the fraction of the scattering power attributable to the partial and rest structures, respectively. In this way, the expectation value for the average of the squared normalized structure amplitudes will be 1·0 for the total, partial and rest structures simultaneously.

3. The probability density $P(\theta)$

For the situation given in Fig. 1, pE_p and $|E|$ are known and rE_r is considered to be a random variable. This implies that the rest structure is considered to be independent of the partial structure, or, knowing the position of some heavy atoms, or of a molecular fragment, does not affect the *randomness* of the position of the remaining atoms. This will generally be true, except for some low-order reflections in the case where a large molecular fragment is known. Thus we can apply the probability density formulae to the rest structure.

The probability density for the phase angle θ (see Fig. 1) has been derived (Sim, 1959) and the result is given here in terms of the previously defined quantities:

$$P(\theta) = \exp(Q \cos \theta)/2\pi I_0(Q), \tag{7}$$

where $I_0(Q)$ is a hyperbolic Bessel function of order zero and

$$Q = 2|E|\,|E_p|p/r^2. \tag{8}$$

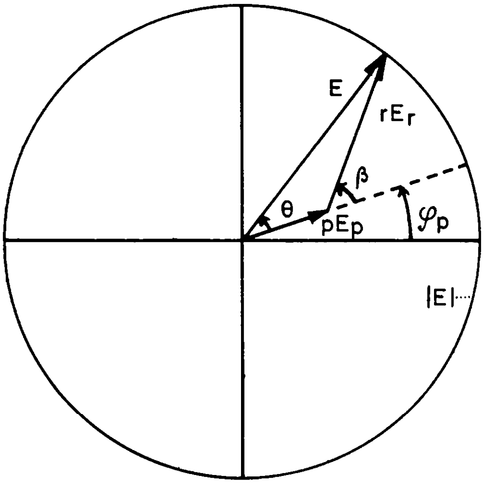

Fig. 1. Vector diagram in the complex plane of (3) for the case $p|E_p| < |E|$. The $|E|$ circle represents all possible vectors E.

(The derivation of these formulae in terms of E values is given in Appendix 1.)

It is useful to express the reflection data in terms of the minimum and maximum value that E_r can have for a given reflection.

Using (3), we have (see Fig. 1):

$$\begin{aligned} &\text{for } \theta = 0; \quad |E_1| = |(|E| - p|E_p|)|/r, \\ &\text{for } \theta = \pi; \quad |E_2| = (|E| + p|E_p|)/r. \end{aligned} \tag{9}$$

Using these definitions, we now write for (8):

$$Q = \tfrac{1}{2}(|E_2|^2 - |E_1|^2). \tag{10}$$

In the following section, (7) and (10) are used to calculate the expectation value for M for a reflection $\mathbf{h}$, given $|E_1|$ and $|E_2|$:

$$\langle M \rangle = \int_{-\pi}^{\pi} M P(\theta)\, d\theta, \tag{11}$$

where $M = \beta^2$, $\cos \beta$ or $|E_r|^2$. Note that (11) is valid for general reflections in all non-centrosymmetric space groups.

Corrections

Equation (11) and Table 1 replace our former incorrect results [van den Hark, Prick & Beurskens, 1976, equations (11) and (20), and Table 1].

4. Calculations of $\langle \beta^2 \rangle$, $\langle \cos \beta \rangle$ and $\langle |E_r|^2 \rangle$

We have used (11) for the numerical calculations of expectation values for β^2, $\cos \beta$ and $|E_r|^2$. Use was made of the relations:

$$r^2|E_r|^2 = p^2|E_p|^2 + |E|^2 - 2p|E_p|\,|E| \cos \theta; \tag{12}$$

$$\cos \beta = (|E| \cos \theta - p|E_p|)/r|E_r|. \tag{13}$$

These calculations were done for pairs of $|E_1|$, $|E_2|$ values. It is useful to express this in terms of $|E_1|$ and $|E_2|$, and the inverse of (9) is given for two cases. If the calculated structure amplitude is less than the observed structure amplitude we have:

$$p|E_p| < |E| \begin{cases} p|E_p| = \tfrac{1}{2}r(|E_2| - |E_1|), \\ |E| = \tfrac{1}{2}r(|E_2| + |E_1|), \end{cases} \tag{14}$$

Table 1. *Standard deviation (°) of the phase of E_1 for pairs of $|E_1|$ and $|E_2|$ values for reflections with $|F_{calc}|$ less than $|F_{obs}|$*

$\lvert E_2\rvert =$	0·3	0·6	0·8	1·0	1·2	1·4	1·7	2·0	2·5	3·0	4·0
$\lvert E_1\rvert$											
0·3	104	110	110	107	104	99	91	84	76	72	67
0·6		104	104	102	98	92	83	75	64	59	54
0·8			104	102	98	92	82	72	60	54	48
1·0				104	100	94	83	72	58	50	44
1·2					104	98	87	74	57	48	41
1·4						104	92	78	58	47	38
1·7							104	89	63	47	36
2·0								104	73	50	35

otherwise:

$$p|E_p| > |E| \begin{cases} p|E_p| = \tfrac{1}{2}r(|E_2| + |E_1|), \\ |E| = \tfrac{1}{2}r(|E_2| - |E_1|). \end{cases} \quad (15)$$

Substitution of (14) and (15) into (12) gives the same result for $|E_r|^2$:

$$2|E_r|^2 = |E_1|^2 + |E_2|^2 - (|E_2|^2 - |E_1|^2)\cos\theta, \quad (16)$$

but substitution into (13) gives

$$p|E_p| < |E|;$$

$$\cos\beta = [(|E_2| + |E_1|)\cos\theta - (|E_2| - |E_1|)]/2|E_r|,$$

$$p|E_p| > |E; \quad (17)$$

$$\cos\beta = [(|E_2| - |E_1|)\cos\theta - (|E_2| + |E_1|)]/2|E_r|. \quad (18)$$

The numerical integrations were carried out using steps of 1° in θ, and were repeated with steps of $\frac{1}{2}$°; the estimated errors of the results are in the order of the round-off errors of the data given in the tables.

The standard deviation of the phase of E_1 of a reflection **h**

If the value of E_1, with phase φ_p, is assumed to be the normalized difference-structure factor, then β (see Fig. 1) is the phase error ($\beta = \varphi_r - \varphi_p$, where φ_r is the correct phase of the difference structure factor). The standard deviation is the square root of the variance; the variance is calculated by (11) with $M = \beta^2$.

The numerical calculations are performed for the case $p|E_p| < |E|$ using (17) for the calculation of β. The result is tabulated in Table 1. Note that the probability density function $P(|E|)$ has a maximum at $|E| \simeq 0·7$; for $|E_1| < 0·7$, the value of $|E_2|$ is often more probable than the value of E_1, and $|\theta|$ is expected to be greater than 90°. As a consequence, the table shows standard deviations exceeding the random-phases value (103·9°).

Note. For reflections with $p|E_p| > |E|$, the maximum absolute value of the phase β is restricted (see van den Hark, Prick & Beurskens, 1976); if $|E_1|$ is significantly greater than zero, the phase β will be very reliable, and this reflection will get unit weight in our procedures.

The weight of the phase of E_1

Sim (1960) introduced the expectation values for $\cos\theta$ as weighting factors for the calculation of a Fourier synthesis. Similarly, we may use expectation values for $\cos\beta$ as weights for a difference Fourier synthesis. These weights are calculated by (11) with $M = \cos\beta$. The results for $p|E_p| < |E|$ are given in Table 2. The negative values in Table 2 correspond with the larger standard deviations in Table 1; the corresponding reflections should probably be given zero weight.

Table 2. *Weights (expectation values for $\cos\beta$) for pairs of $|E_1|$ and $|E_2|$ values for reflections with $|F_{calc}|$ less than $|F_{obs}|$*

$\lvert E_2\rvert =$	0·3	0·6	0·8	1·0	1·2	1·4	1·7	2·0	2·5	3·0	4·0
$\lvert E_1\rvert$											
0·3	0·0	−0·11	−0·11	−0·08	−0·03	0·03	0·12	0·21	0·31	0·36	0·42
0·6		0·0	0·0	0·03	0·08	0·15	0·27	0·37	0·49	0·55	0·62
0·8			0·0	0·03	0·09	0·17	0·30	0·42	0·55	0·62	0·69
1·0				0·0	0·06	0·15	0·29	0·43	0·59	0·67	0·74
1·2					0·0	0·09	0·25	0·42	0·61	0·70	0·77
1·4						0·0	0·18	0·37	0·60	0·71	0·80
1·7							0·0	0·22	0·55	0·72	0·82
2·0								0·0	0·44	0·69	0·83

In the *DIRDIF* procedure, we use a different weighting scheme (W_1, see van den Hark, Prick & Beurskens, 1976), which is comparable with the weighting scheme based upon the results of the tangent refinement of the phases. For comparison we give the weights W_1 in Table 3. The formula used for the calculation of W_1 is not suitable for reflections with negative weights in Table 2; these reflections, however, are not used at all in the *DIRDIF* procedures.

Expectation value for $|E_r|^2$

The $|E_r|$ value for a reflection **h** has $|E_1|$ and $|E_2|$ as possible minimum and maximum values; the expectation value for $|E_r|^2$ may be calculated by (11) with $M = |E_r|^2$.

The analytical expression for the estimate

$$\langle |E_r|^2 \rangle = |E_1|^2 + Q\left[1 - \frac{I_1(Q)}{I_0(Q)}\right] \quad (19)$$

is obtained from (12) and (9), using

$$\langle \cos\theta \rangle = I_1(Q)/I_0(Q) \quad \text{(Sim, 1960).}$$

The result (19) was also obtained by Srinivasan (1968) and by Hull & Irwin (1978). The numerical results are given in Table 4.

The expectation values are used in *DIRDIF* to check the performance of the phase refinement.

If the expectation value for $|E_r|^2$ is calculated for all reflections, the average is expected to be nearly equal to one. In the *DIRDIF* procedure, the average value is calculated (denoted *a priori*). For those reflections of which the phases are to be refined, the expectation

Table 3. *DIRDIF weights W_1 (see text)*

$\lvert E_2\rvert =$	0·3	0·6	0·8	1·0	1·2	1·4	1·7	2·0	2·5	3·0	4·0
$\lvert E_1\rvert$											
0·3	0·0	0·04	0·04	0·02	0·00	0·03	0·24	0·58	0·93	1·00	1·00
0·6		0·0	0·00	0·00	0·04	0·13	0·40	0·70	0·96	1·00	1·00
0·8			0·0	0·01	0·04	0·13	0·40	0·71	0·96	1·00	1·00
1·0				0·0	0·02	0·09	0·35	0·67	0·95	1·00	1·00
1·2					0·0	0·03	0·25	0·60	0·93	1·00	1·00
1·4						0·0	0·12	0·47	0·91	0·99	1·00
1·7							0·0	0·20	0·82	0·98	1·00
2·0								0·0	0·59	0·96	1·00

Table 4. *Expectation values for* $|E_r|^2$

$\lvert E_2\rvert =$	0·3	0·6	0·8	1·0	1·2	1·4	1·7	2·0	2·5	3·0	4·0
$\lvert E_1\rvert$											
0·3	0·1	0·2	0·3	0·4	0·5	0·6	0·7	0·7	0·7	0·6	0·6
0·6		0·4	0·5	0·6	0·8	0·9	1·0	1·0	0·9	0·9	0·9
0·8			0·6	0·8	1·0	1·1	1·2	1·2	1·2	1·2	1·2
1·0				1·0	1·2	1·4	1·5	1·6	1·6	1·5	1·5
1·2					1·4	1·7	1·9	2·0	2·0	2·0	2·0
1·4						2·0	2·3	2·5	2·6	2·5	2·5
1·7							2·9	3·3	3·5	3·5	3·4
2·0								4·0	4·6	4·6	4·5

value is replaced by $|E_1|^2$. The average now is less than the *a priori* average. During the refinement of the phases, the $|E_1|^2$ values are replaced by the current (more probable) $|E_r|^2$ values, and the average will increase and approach the *a priori* average. A typical example is given in Appendix 2.

5. Distribution function of $|E_r|$

If the expectation value for $|E_r|$ is calculated for all reflections, then one can plot the number of reflections present in intervals of $|E_r|$; this plot should give an approximation of the probability density function $P(|E_r|)$ for the rest structure. The experimental result should more or less match the probability density function for either centro-symmetric or non-centrosymmetric structures. The expectation values, however, tend to cluster about the mean; the exceptional cases, which are critical to the distribution, will almost never be near expectation values. The expectation value therefore is replaced by a weighted distribution of possible $|E_r|$ values.

The procedure used in our present programs is:

(*a*) The rest structure is centrosymmetric;

$|E_r|$ must be equal to either $|E_1|$ or $|E_2|$. For all reflections, calculate the probabilities P_1 and P_2 (Gould, van den Hark & Beurskens, 1975) and sum P_1 and P_2 in the appropriate ranges of $|E_r|$.

(*b*) The rest structure is non-centrosymmetric;

$|E_r|$ may now lie anywhere between $|E_1|$ and $|E_2|$. For all reflections calculate relative weights (using 7) for a number (say 18) of possible $|E_r|$ values in the range $|E_1|$, $|E_2|$, and sum the (18) weights of the possible $|E_r|$ values in the appropriate ranges of $|E_r|$.

Thus, knowing whether or not the structure is centrosymmetric, the experimental distribution of $|E_r|$ is easily obtained. An example is given in Appendix 2.

Experiments on a number of test structures have shown that when the known part of the structure consists of heavy atoms, reflections with relatively large $|E_p|$ or $\sin\theta/\lambda$ are likely to have unreliable $|E_1|$ and $|E_2|$ values; therefore the best results for the distribution function are obtained by imposing $|E_2| - |E_1|$ and $\sin\theta/\lambda$ limits (*e.g.* 2·0 and 0·5, respectively).

6. A statistical test for the presence of a centre of symmetry in the rest structure

A brief, practical method for testing for the presence of a centre is to compare the occurrence of $|E|$ values with theoretical populations of ranges of $|E|$. For this purpose we define a 'centricity fraction', X_C, as the least-squares solution to the set of condition equations:

$$P_{\text{exp}}(|E|) = X_C P_C(|E|) + (1 - X_C) P_A(|E|), \quad (20)$$

where P_{exp} is the experimental distribution and P_C and P_A are the theoretical centric and acentric distributions, respectively. One result (20) is obtained for each of the ranges of $|E|$. Minimizing the squares of the deviations in (20) gives:

$$X_C = \sum (P_{\text{exp}} - P_A)(P_C - P_A) / \sum (P_C - P_A)^2, \quad (21)$$

where the summation is over all ranges of $|E|$. $X_C = 1$ for an ideal centric distribution and $X_C = 0$ for an ideal acentric distribution. Experimental values for X_C, calculated for structures with centric and acentric distributions, come close to these theoretical values.

The application of this test to difference structure factors is slightly more complicated as the true $|E_r|$ values are not known. In special cases, where the known atoms are at special or pseudo-special positions, one can apply the test using a large number of reflections with zero contribution of the known atoms ($E_p = 0$, $|E_r| = |E_1| = |E_2|$).

In the general case, the result will depend on the presumed presence or absence of a centre; this dependency, however, is reduced if the application is limited to reflections with rather small $|E_p|$ values.

Define $X_C'(E_m)$ as the X_C value, obtained by using only reflections with $|E_p|$ less than a certain maximum value E_m. The calculations then are repeated for, say, $E_m = 1·0, 0·8, 0·6, 0·4, \cdots$. On reducing the maximum allowed contribution of the known atoms for any reflection, one reduces the range of possible $|E_r|$ values for that reflection, and thus the dependency on the presumed distribution is reduced.

Experiments clearly show that extrapolation of the results $X_C'(E_m)$ for $E_m \to 0$ leads to *one* solution:

$$X_C = \lim_{E_m \to 0} X_C'(E_m), \quad (22)$$

independent of the presumed presence or absence of a centre. This being established, we have now accepted a fixed number of possible $|E_r|$ (in the range $|E_1|$, $|E_2|$) in our calculations of $X_C'(E_m)$. For centrosymmetric rest structures the result of $X_C'(E_m)$ will usually *increase*, and for non-centrosymmetric rest structures, the result will usually *decrease*, on lowering E_m. In all cases $X_C'(E_m)$, plotted as a function of E_m, is linear except at low E_m values (where the results are unreliable because of the low number of reflections used in the calculations).

The extrapolation (equation 22), therefore, is done by calculating the least-squares line through the experimental points, $X'_C(E_m)$, using the number of contributing reflections as a weighting factor.

Numerical results for 13 test structures are given in Appendix 3. It is shown that in most cases $X_C > 0{\cdot}6$ for centric, and $X_C < 0{\cdot}4$ for acentric distributions. In the case where X_C is close to 0·5, the rest structure may indeed have an 'intermediate' (neither centric nor acentric) distribution. It is, however, more probable that the subtraction of the known part is not done accurately; if that is the case, then the errors in $|E_r|$ are largest for high-order reflections, and improved results for X_C are obtained by imposing a limit on $\sin\theta/\lambda$ for contributing reflections.

The authors are grateful to Dr R. de Graaf (Math. Stat. Adv.) and to Drs J. C. Smit (Math. Inst.) for mathematical advice. Part of this work was supported by FOMRE with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO).

We would like to express our thanks to the referee for helpful criticism.

APPENDIX 1

Derivation of the distribution function $P(\theta)$

We will use the following well-known expression for the normalized probability density for normalized structure factors of non-centrosymmetric structures (Wilson, 1949; Ramachandran & Srinivasan, 1959):

$$P(|E|) = 2|E| \exp(-|E|^2). \tag{23}$$

$P(|E|)\,\mathrm{d}|E|$ is the probability that a normalized structure factor has an amplitude within the range $|E|$, $|E| + \mathrm{d}|E|$. E is a complex number, or a vector in the complex plane; define:

$$\begin{aligned} &E = X + iY, \\ &R = |E|; \quad R^2 = X^2 + Y^2, \\ &X = R\cos\theta; \quad Y = R\sin\theta. \end{aligned} \tag{24}$$

Equation (23) was originally derived (Wilson, 1948) from a Gaussian distribution density for the real and imaginary components of the structure factors. In our notation:

$$\begin{aligned} P(X) &= \pi^{-1/2} \exp(-X^2), \\ P(Y) &= \pi^{-1/2} \exp(-Y^2), \end{aligned} \tag{25}$$

where the variance of the distribution is $\frac{1}{2}$ to achieve the normalization of $P(|E|)$. X and Y are considered to be independent random variables. $P(X,Y)\,\mathrm{d}X\,\mathrm{d}Y$ is the probability that the components of E are in the ranges X, $X + \mathrm{d}X$ and Y, $Y + \mathrm{d}Y$.

The vector density is given by the joint probability density:

$$P(X,Y) = P(X)\,P(Y) = \frac{1}{\pi}\exp[-(X^2 + Y^2)]. \tag{26}$$

Transformation to polar coordinates, using (24) and $\mathrm{d}X\,\mathrm{d}Y = R\,\mathrm{d}R\,\mathrm{d}\theta$, gives

$$P(\theta, R) = \frac{1}{\pi} R \exp(-R^2). \tag{27}$$

Integration over all angles gives the distribution of the amplitudes:

$$P(R) = 2R\exp(-R^2),$$

which is identical to (23).

For the present problem we consider the *a priori* probability density for the normalized difference-structure factor $E_r = X_r + iY_r$.

Equation (26) reads:

$$P(X_r, Y_r) = \frac{1}{\pi}\exp[-(X_r^2 + Y_r^2)].$$

This gives for the vector

$$rE_r = rX_r + irY_r = X^1 + iY^1,$$

$$P(X^1, Y^1) = \frac{1}{\pi r^2}\exp[-(X^{1^2} + Y^{1^2})]/r^2. \tag{28}$$

Let us consider the calculated vector pE_p as a fixed number. For simplicity, we may assume that the calculated phase φ_p is zero (this may be achieved for any given reflection by an origin shift of the crystal axes); the corresponding situation is given in Fig. 2: pE_p now is a real, positive number.

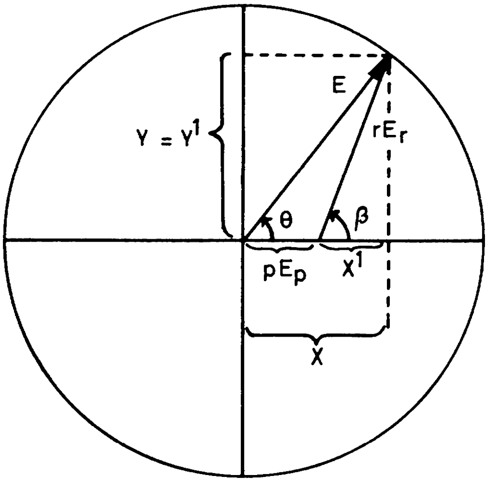

Fig. 2. Vector diagram for $\varphi_p = 0$.

The probability density $P(X, Y)$ of the vector $E = pE_p + rE_r = X + iY$, given the value of pE_p, now is obtained by substituting $X^1 = X - p|E_p|$ and $Y^1 = Y$ in the right-hand side of (28). Upon transformation to polar coordinates, as above, one obtains:

$$P(R,\theta) = \frac{R}{\pi r^2} \exp[-(R^2 + |pE_p|^2 - 2Rp|E_p| \cos \theta)]/r^2. \quad (29)$$

Let us now assume that the amplitude R is known from experiment. All possible vectors E will form a circle with radius R. The normalized probability density $P(\theta)$, given the value of R, then is obtained by substituting the observed R value in (27) and renormalizing the result:

$$P(\theta) = P(R,\theta) / \int_{-\pi}^{\pi} P(R,\theta)\, d\theta. \quad (30)$$

This immediately gives (7), which is identical to the result obtained by Sim (1959).

APPENDIX 2
An example

The example compound $C_{15}H_{16}N_2O_2S$ crystallizes in space group $P2_12_12_1$ with $Z = 4$ (Noordik, Beurskens, Ottenheijm, Herscheid & Tijhuis, 1978). The crystal structure could not 'routinely' be solved by direct methods. The position of the S atom was found from a Patterson map as: $x \simeq 0{\cdot}0$, $y = 0{\cdot}09$, $z = 0{\cdot}14$. This position, however, is a pseudo-special position as the trial model (S atoms) includes a mirror plane at $x = 0$ (and a centre of symmetry at $\frac{1}{4},\frac{1}{4},0$). A conventional difference Fourier would have led to a superposition of the structure with its enantiomorph. This mirror symmetry was destroyed by shifting the atom about 0·15 Å ($x_{\text{new}} = 0{\cdot}02$, after refinement: $x = 0{\cdot}0165$).

A structure factor calculation gave an R value of 0·56 for 1684 reflections (including all 'zero-observed' reflections).

The refinement of the scale together with the temperature factor of the S atom (B_p) and the overall temperature factor for the remaining atoms (B_r) gave:

$$B_p = 3{\cdot}5\ \text{Å}^2, \quad B_r = 3{\cdot}4\ \text{Å}^2.$$

With these results, E_1 and E_2 values were calculated for all 1684 reflections. The distribution of these reflections over several ranges in $|E_1|$ and $|E_2|$ are given in Table 5.

The numbers in this table add up to 1216, *i.e.* the number of reflections having $|F_{\text{calc}}|$ less than $|F_{\text{obs}}|$. The remaining 468 reflections are relatively weak reflections: only nine of these have $|E_1| > 1{\cdot}0$.

Table 5. *Number of reflections in intervals of* $|E_1|$ *and* $|E_2|$ *for reflections with* $|F_{calc}|$ *less than* $|F_{obs}|$

(Example, see Appendix 2.)

$E_2 = 0{\cdot}0\text{–}0{\cdot}3\text{–}0{\cdot}6\text{–}0{\cdot}8\text{–}1{\cdot}0\text{–}1{\cdot}2\text{–}1{\cdot}4\text{–}1{\cdot}7\text{–}2{\cdot}0\text{–}2{\cdot}5\text{–}3{\cdot}0\text{–}4{\cdot}0\text{–}\infty$

E_1												
0·0	9	31	33	37	33	40	55	44	39	23	2	0
0·3	0	4	22	31	30	29	54	57	41	29	5	0
0·6	0	0	1	7	15	13	33	30	37	18	12	0
0·8	0	0	0	5	14	15	17	23	47	13	12	0
1·0	0	0	0	0	4	4	19	17	15	26	14	0
1·2	0	0	0	0	0	1	7	15	19	13	12	0
1·4	0	0	0	0	0	0	1	9	15	12	19	3
1·7	0	0	0	0	0	0	0	0	6	3	4	4
2·0	0	0	0	0	0	0	0	0	2	6	5	1
∞												

The distribution of $|E_r|$ is given in Fig. 3. The 'experimental' curve is obtained by calculating structure factors, using the refined parameters for all atoms but sulphur and normalizing the data in the usual way. The *a priori* 'expected' curve, as obtained by the procedure described in §5, is smoother, but otherwise close to the 'experimental' curve.

The results for the centricity test for the rest structure are given in Appendix 3; the extrapolation towards $E_p = 0$ clearly shows the rest structure to be non-centrosymmetric. Of course, there is no problem concerning the space group in this example. Neverthe-

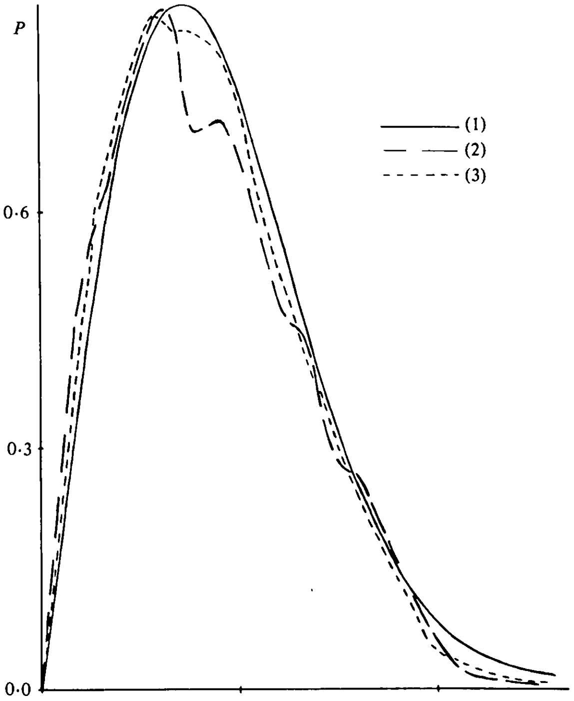

Fig. 3. Probability density $P(|E_r|)$. Curve (1) Theoretical curve, acentric. Curve (2) Experimental curve using calculated E_r's (Appendix 2). Curve (3) Results obtained by the present procedure (example, see Appendix 2).

less, we will use these statistical methods as a check before the application of direct methods to the difference structure factors (*DIRDIF*): if the statistical results are not as satisfactory as in this example, one should carefully check the 'input structure' or perhaps adjust the program parameters to be used.

The solution of the structure by the *DIRDIF* procedure was continued as follows. There were 110 reflections having $W_1 > 0{\cdot}16$ and $|E_1| > 1{\cdot}30$ (the present program default values), and these reflections were used as input to the tangent refinement procedure; a total of 330 reflections with $|E_1| > 0{\cdot}90$ were refined in four cycles.

The *a priori* average of expectation values for $|E_r|^2$ was found to be:

0·973
using $|E_1|^2$ for the 330 reflections to be refined: 0·872
using the most probable $|E_r|^2$ value;
after cycle 1: 0·897
after cycle 2: 0·920
after cycle 3: 0·948
after cycle 4: 0·965

The weighted difference-structure factors were used as input to a Fourier program, resulting in a map of which the top 20 peaks corresponded to the 19 non-hydrogen atoms of the molecule (including the S atom!) and one false position.

The known part of the structure often will be found in the Fourier map; in the present example the fourth highest peak corresponds to the input S atom position.

APPENDIX 3

Numerical results for the centricity fraction X_C

The centricity fraction X_C, as defined in §6 and calculated according to the procedures described before, has been determined for 13 test structures. In all cases an extrapolation to zero contribution of the known part of the structure has been applied.

In the present procedure, special reflections, which have or may have a centric distribution, are skipped. Instead of the true, but unknown $|E_r|$ value, we accept three possible values with different weights, namely: $|E_r| = |E_1|$ with $w = \frac{1}{4}$; $|E_r| = \frac{1}{2}(|E_1| + |E_2|)$ with $w = \frac{1}{2}$; $|E_r| = |E_2|$ with $w = \frac{1}{4}$. These weights are summed in the appropriate ranges of $|E_r|$. After completing this summation, the resulting distribution is normalized. Two thresholds have been built in. First a sin θ/λ limit is used to reduce the influence of possible errors in the parameters of the known atoms; all calculations are repeated for the threshold values 0·50, 0·40, 0·32 and 0·25 Å^{-1}. The second criterion is the magnitude of the calculated structure factor of the reflection. For programming convenience we replace the threshold on $|E_p|$ by a threshold on $p|E_p|/r$, which is given by (14) and (15) as $\frac{1}{2}(|E_2| \pm |E_1|)$. The threshold values $p|E_m|/r$ are: 8·0, 2·0, 1·8, 1·6, 1·4, 1·2, 1·0, 0·8, 0·6, 0·4. For each of these values X'_C is calculated. Nine X'_C values enter in a weighted least-squares procedure (the threshold 8·0 is not used). The weights assigned to the X'_C values are proportional to the number of reflections contributing to an X'_C value. We have observed that at

Table 6. *Calculated* X_C *values of* 13 *test structures for various* sin θ/λ *limits*

Values between parentheses should be neglected because the total number of reflections contributing to the calculation of X_C is too small (<50).

Entry	Known atoms	Unknown atoms	Sin $\theta/\lambda < \infty$	0·50	0·40	0·32	0·25	Theoretical X_C*
(1)	Sb	$C_9N_{21}N_2S_4$	0·321	0·041	−0·587	−0·559	−0·425	0
(2)	Zn	$C_6H_{12}O_6N_2$	−0·097	−0·311	−0·012	0·126	(0·728)	0
(3)	Cl	$C_9H_{14}N_3O_4$	−0·066	0·043	−0·096	−0·062	0·179	0
(4)	S	$C_{15}H_{16}N_2O_2$	0·497	0·221	0·084	0·042	0·081	0
(5)	2Cl	$C_{27}H_{45}NO$	0·671	0·734	0·602	0·513	0·204	0
(6)	Au + 2Br	$C_{16}H_{23}NS_2$	0·885	1·119	1·509	1·905	2·206	1
(7)	4I	$(AsC_4H_9NS_2)_2$	1·184	1·505	1·614	1·931	1·963	1
(8)	Cl	$C_5H_{13}N_2O_4$	1·989	2·179	2·037	1·943	2·083	1
(9)	Ni + 2S	$C_{15}H_7N_4$	0·785	0·574	0·644	0·750	(−0·019)	1
(9)	Ni	$C_{15}H_7N_4S_2$	1·031	1·109	1·387	1·637	2·035	1
(10)	2Br	$C_{11}H_{12}O_3$	0·535	0·606	0·584	0·778	0·954	1
(11)	Au + 2Br	$C_{17}H_{25}NS_2$	0·970	1·087	1·094	1·271	1·361	1
(12)	Cu + I	$C_{12}H_{20}N_2S_3$	0·353	0·195	0·133	0·098	0·518	1
(13)	S	$C_{15}H_{16}N_2O_2$	0·923	1·003	1·107	1·093	1·001	1

References: (1) Cras & Willemse (1978). (2) van der Helm, Nicholas & Fisher (1970). (3) Subramanian & Hunt (1970). (4) Noordik, Beurskens, Ottenheijm, Herscheid & Tijhuis (1978). (5) Mootz & Berking (1970). (6) Wijnhoven, Bosman & Willemse (1979). (7) Beurskens, Beurskens, Noordik, Willemse & Cras (1979). (8) Prick & Beurskens (1979). (9) Manoharan & Noordik (1979). (10) Noordik & Groen (1978). (11) Bosman, Wijnhoven & Willemse (1979). (12) van de Leemput, Willemse, Cras & Groen (1979). (13) Noordik, Herscheid, Tijhuis & Ottenheijm (1977).

* Ideal X_C value for centrosymmetric structure: 1; ideal X_C value for non-centrosymmetric structure: 0.

low $|E_m|$ values, the results for $X'_C(E_m)$ become unreliable.

The calculated X_C values for 13 test structures applying four sin θ/λ thresholds, are collected in Table 6. This table clearly shows the tendency for centrosymmetric structures to have an X_C value near 1·0 and for non-centrosymmetric structures to have a value approaching 0. Upon decreasing the dependency on the 'input' structure, a better approximation for these ideal X_C values is obtained, especially for structures where at first sight a strong preference for 'heavy atom' centricity is clearly seen. Upon decreasing the sin θ/λ limit, the results sometimes become less reliable because of the small number of contributing reflections; on the other hand, this sometimes will improve the results (see Table 6, entry 5), as errors in thermal parameters become less important.

One example (Table 6, entry 12) does not give the expected X_C value. To explain this, we have calculated the distribution for the $|E|$ values of the rest structure, using the phases calculated after the refinement of the structure; it was found that this distribution resembles more the acentric than the centric curve. From this one example it is concluded that the results for X_C are not definite proof for the presence or absence of a center of symmetry.

References

Beurskens, P. T., Beurskens, G., Noordik, J. H., Willemse, J. & Cras, J. A. (1979). To be published.
Bosman, W. P., Wijnhoven, F. & Willemse, J. (1979). To be published.
Cras, J. A. & Willemse, J. (1978). *Recl Trav. Chim. Pays-Bas*, **97**, 28–29.
Gould, R. O., van den Hark, Th. E.M. & Beurskens, P. T. (1975). *Acta Cryst.* A**31**, 813–817.
Hark, Th. E. M. van den, Prick, P. A. J. & Beurskens, P. T. (1976). *Acta Cryst.* A**32**, 816–821.
Helm, D. van der, Nicholas, A. F. & Fisher, C. G. (1970). *Acta Cryst.* B**26**, 1172–1178.
Hull, S. E. & Irwin, M. J. (1978). *Acta Cryst.* A**34**, 863–870.
Leemput, P. J. H. A. M. van de, Willemse, J., Cras, J. A. & Groen, L. (1979). To be published.
Manoharan, P. T. & Noordik, J. H. (1979). To be published.
Mootz, D. & Berking, B. (1970). *Acta Cryst.* B**26**, 1362–1372.
Noordik, J. H., Beurskens, P. T., Ottenheijm, H. C. J., Herscheid, J. D. M. & Tijhuis, M. W. (1978). *Cryst. Struct. Commun.* pp. 669–677.
Noordik, J. H. & Groen, L. (1978). *Cryst. Struct. Commun.* **7**, 293–297.
Noordik, J. H., Herscheid, J. D. M., Tijhuis, M. W. & Ottenheijm, H. C. J. (1977). *Recl Trav. Chim. Pays-Bas*, **97**, 91–95.
Prick, P. A. J. & Beurskens, P. T. (1979). To be published.
Ramanchandran, G. N. & Srinivasan, R. (1959). *Acta Cryst.* **12**, 410–411.
Sim, G. A. (1959). *Acta Cryst.* **12**, 813–818.
Sim, G. A. (1960). *Acta Cryst.* **13**, 511–512.
Srinivasan, R. (1968). *Z. Kristallogr.* **126**, 175–181.
Subramanian, E. & Hunt, J. O. (1970). *Acta Cryst.* B**26**, 303–311.
Wijnhoven, F., Bosman, W. P. & Willemse, J. (1979). To be published.
Wilson, A. J. C. (1949). *Acta Cryst.* **2**, 318–321.
Woolfson, M. M. (1956). *Acta Cryst.* **9**, 804–810.

Acta Cryst. (1979). A**35**, 772–775

A Proposed Rigorous Definition of Coordination Number

By M. O'Keeffe

Department of Chemistry, Arizona State University, Tempe, Arizona 85281, *USA*

(*Received* 11 *December* 1978; *accepted* 21 *March* 1979)

Abstract

It is proposed that a weighted coordination number, Z^*, be used as a measure of the number of neighbours of an atom in a crystal. Coordinating atoms contribute faces to the Voronoi polyhedron around a central atom and their contributions are weighted in proportion to the solid angle subtended by that face at the centre. The advantages of this definition over other proposals are pointed out.

0567-7394/79/050772-04$01.00

The concept of coordination number (CN) of an atom or ion in a crystal is very widely used and has proved very fruitful in crystal chemistry. The CN of atoms in crystals has become accepted as a basic parameter describing a structure, and many atomic properties such as atomic (or ionic) radius are considered to depend upon it directly. It is nevertheless true that it is hard to find an unambiguous definition of CN that is not in conflict with intuition in many instances. In this paper a rigorous yet logical and useful definition is proposed.