Acta Cryst. (1965). **19**, 1008

Probability Distribution Connected with Structure Amplitudes of Two Related Crystals. V. The Effect of Errors in the Atomic Coordinates on the Distribution of Observed and Calculated Structure Factors*

BY R. SRINIVASAN AND G. N. RAMACHANDRAN

Centre of Advanced Study in Physics, University of Madras, Madras 25, *India*

(*Received* 9 *February* 1965)

The problem of the probability distribution of the observed and calculated structure factors for the general case when the latter include not only a part of the structure (P out of a total of N atoms), but also when these P atoms have finite errors $\Delta\mathbf{r}_j$, is considered in this paper. This forms essentially an extension of the theory developed in the earlier parts where, however, only two limiting situations were considered, namely one in which the P atoms have no errors (related case) and the other in which the errors are very large (unrelated case). It is found that there is a formal identity in the mathematical results for the distributions arising out of errors in atomic coordinates and out of difference in number of atoms, which leads to elegant results for the general case. The results of the earlier parts, in particular those in terms of the normalized variables (parts II to IV), could be taken over completely to the present case with the only change that the parameter $\sigma_1^2(=\sum_j^P f_j^2/\sum_j^N f_j^2)$ of the earlier parts has to be replaced by $\sigma_A^2=\sigma_1^2D^2$ where $D=\langle\cos 2\pi\mathbf{H}\,.\,\Delta\mathbf{r}_j\rangle_P$, $\mathbf{H}$ being the reciprocal vector. The results can be used to estimate the errors in the positions of the known atoms. In particular, the normalized reliability index $R_1=\Sigma||F_N|-|F_P^c|/\sigma_1|/\Sigma|F_N|$ has been worked out as a function of σ_A both for centrosymmetric and non-centrosymmetric crystals, which enables one readily to estimate σ_A and hence D. For the centrosymmetric case R_1 is simply given by $R_1=\sqrt{2(1+\sigma_A)}+\sqrt{2(1-\sigma_A)}-2$.

1. Introduction

In the first three parts of this series (part I: Ramachandran, Srinivasan & Sarma, 1963; part II: Srinivasan, Sarma & Ramachandran, 1963*a*; part III: Srinivasan, Subramanian & Ramachandran, 1964) the probability distributions of the structure amplitudes $|F_N|$ and $|F_P|$ of two crystals containing N and P atoms and various variables connected with them, such as the difference, product and the quotient, were considered. The results were shown to be useful in testing the 'relatedness' or 'isomorphism' between a pair of crystals (see also Ramachandran & Srinivasan, 1963; Srinivasan, Sarma & Ramachandran, 1963*b*). It was pointed out in part IV (Srinivasan & Ramachandran, 1965) that the results obtained earlier were particularly applicable to the data on the 'observed' and 'calculated' structure factors during the early stages of a crystal structure analysis, when only a part of the structure is known. It was also shown in part IV that in such a case the proper types of variable to consider are the ones in which the $|F_N|$ and $|F_P|$ enter in their normalized form. In particular, the proper difference variable was shown to be δ, the normalized difference, defined by

$$\delta=y_N-y_P=(|F_N|/\sigma_N-|F_P|/\sigma_P)\,. \qquad (1)$$

(The notation of the earlier parts will be continued in this part also.)

The study of the distribution $P(\delta)$ led to the suggestion of a new reliability index R_1, termed the normalized reliability index defined by

$$R_1=\Sigma||F_N|-|F_P|/\sigma_1|/\Sigma|F_N|\,. \qquad (2)$$

R_1 is related to δ and in fact $R_1=\langle|\delta|\rangle/\langle y_N\rangle$ and could therefore be worked out theoretically from the available distribution $P(\delta)$. The value of R_1 for any structure could be used to assess the accuracy of the assumed positions of the P atoms by comparing them with the theoretical values available for the related (correct) and unrelated (incorrect) cases.

However, as will be clear from what was mentioned above, only two limiting situations were considered earlier, namely one in which there are no errors in the positions of the P atoms and the other in which the errors are so large that the P atoms could be taken to be in completely wrong positions. It is obvious that in actual practice, the situation would correspond to one in between, namely when there are finite errors in the coordinates of the P atoms. The value of R_1 for any such practical case would depend on the magnitudes of the errors $\Delta\mathbf{r}_j$ in the positions of the P atoms, (in addition to the actual proportion of the P-atoms in the structure). Thus, there should exist quantitative relations between the two. In this paper we shall be primarily concerned with examining this aspect. We shall thus work out the various distributions considered earlier for the general case when the errors in the coordinates of the P atoms are finite.

* Contribution 166 from the Centre of Advanced Study in Physics, University of Madras.

As is to be expected, the related and unrelated cases turn out to be limiting cases of the problem considered here.

It may be mentioned that Luzzati (1952) has considered the statistical distribution of the difference between observed and calculated structure factors in relation to the errors in the atomic coordinates. His treatment is, however, limited to the case when the calculated structure factors include all the atoms in the unit cell (*i.e.* $P=N$ in our notation), while the problem we are concerned with here is more general and is valid for the inclusion of any fraction P of the N atoms for calculating the structure factors. Luzzati's results should therefore be expected to come out as a limiting case of the present problem when $P \to N$, and in fact they do, as will be shown later.

Thus in effect, this paper is an attempt to synthesize the treatments available for the two different but limited aspects of the general problem of distribution of the observed and calculated structure factors, namely (*a*) when the latter do not necessarily include all the atoms, and (*b*) when there are finite errors in the coordinates of the atoms and all the atoms in the unit cell are included in the structure factor calculation.

The interesting result that comes out of this study is that there exists a formal identity of the problem of the statistical distibution arising out of the differences in scattering power and that due to errors in atomic coordinates. One is dependent on the parameter represented by $\sigma_1=\sigma_P/\sigma_N$ and the other by $D=\langle\cos 2\pi\mathbf{H}\,.\,\Delta\mathbf{r}_j\rangle$ and the expressions for probability distributions and other quantities in terms of σ_1 and D come out to be identical.

For reasons mentioned already, we shall treat the problem in this paper mainly from the point of view of the relation between observed and calculated structure factors. The results, however, can also be applied for testing the relatedness or isomorphism between different crystals, if necessary.

Non-centrosymmetric case

Consider a non-centrosymmetric crystal containing N atoms having position vectors $\mathbf{r}_j$ ($j=1$ to N) and denote the structure factors corresponding to $\mathbf{r}_j$ as F_N. When the observational errors are negligible we can assume that the observed structure amplitudes F_N would lead, after final refinement of the structure, to these position vectors $\mathbf{r}_j$. The structure factor corresponding to P out of the N atoms will be denoted by F_P, this being calculated from the 'true' coordinates $\mathbf{r}_j$ ($j=1$ to P) of the P out of the N atoms. If now we take a set of coordinates $\mathbf{r}_j^c$ ($j=1$ to P) for the P atoms, differing from $\mathbf{r}_j$ by $\Delta\mathbf{r}_j$ the structure factor corresponding to $\mathbf{r}_j^c$ will be denoted by F_P^c. The discrepancies between F_P^c's and F_N's will thus be, in general, due to two causes – first to the fact that not all atoms have been used in the calculation of the F_P^c's and secondly to the fact that the positions of the P-atoms used differ from their true values. The relations between the various structure factors are brought out on the Argand diagram in Fig. 1, where the symbol Q stands for the remaining atoms ($N=P+Q$). Here, $F_P=F_P^c+\Delta F_P$ and $F_N=F_P+F_Q$. The significance of V_P in Fig. 1 will be clear in §4.

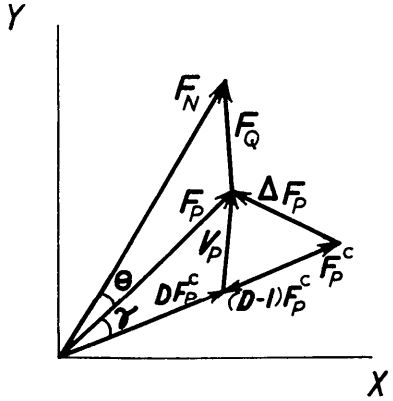

Fig. 1. Argand diagram showing the relation between the various vectors.

To start with, we observe that when the errors are all zero, $\Delta F_P=0$ and the discrepancy $F_N-F_P^c=\Delta$ (say) (see Fig. 1) is only due to F_Q. This corresponds to the 'related case' considered earlier. In the general case when $\Delta\mathbf{r}_j$ are finite and $P\neq N$, both F_Q and ΔF_P contribute to the difference Δ. When the $\Delta\mathbf{r}_j$'s become very large, the two structure factors F_N and F_P^c can be considered to be effectively independent of each other. This corresponds to the 'unrelated case' of the earlier parts.

Let us first consider the probability distribution $P(|F_P^c|;\,|F_P|)$ of $|F_P^c|$ for a given $|F_P|$. This can be obtained from expression (45), p. 806 of Luzzati* (1952) with appropriate changes in the notation to suit this paper. Thus we get

$$P(|F_P^c|;\,|F_P|)\,d|F_P^c| = \frac{2|F_P^c|\,d|F_P^c|}{\sigma_P^2(1-D^2)}\,\frac{1}{\pi}\int_0^{\pi} \times \exp\left\{-\frac{D^2|F_P|^2+|F_P^c|^2-2D|F_P|\,|F_P^c|\cos\gamma}{\sigma_P^2(1-D^2)}\right\}d\gamma \quad (3)$$

where γ is the angle between F_P^c and F_P and $D=\langle\cos 2\pi\mathbf{H}\,.\,\Delta\mathbf{r}_j\rangle_P$. The integral in equation (3) may be integrated in terms of the Bessel functions of imaginary argument $I_0(x)$, and thus we have

$$P(|F_P^c|;\,|F_P|)=\frac{2|F_P^c|}{\sigma_P^2(1-D^2)}\,I_0\left[\frac{2D|F_P|\,|F_P^c|}{\sigma_P^2(1-D^2)}\right] \times \exp\left\{-\frac{|F_P^c|^2+D^2|F_P|^2}{\sigma_P^2(1-D^2)}\right\}. \quad (4a)$$

It is interesting to compare this expression with the distribution obtained for the related case (expression (4), part I) which is given by

* The expression given by Luzzati is actually the conditional distribution of his z ($=|F_c|$) given t ($=|F_o|$), although the form in which he has given it does not express it explicitly. Note that, in our notation, his $z=|F_P^c|$, $t=|F_P|$ and $\varphi=\sigma_P^2$.

$$P(|F_N|;|F_P|)=\frac{2|F_N|}{\sigma_Q^2}I_0\left[\frac{2|F_N|\,|F_P|}{\sigma_Q^2}\right]\times\exp\left\{-\frac{|F_N|^2+|F_P|^2}{\sigma_Q^2}\right\}. \quad (4b)$$

The close similarity between the two is obvious. In fact the formal identity between the two cases is more clearly brought out if we consider the normalized variables. Thus, defining

$$y_N=|F_N|/\sigma_N,\ y_P=|F_P|/\sigma_P,\ \sigma_1^2=\sigma_P^2/\sigma_N^2,\ \sigma_2^2=\sigma_P^2/\sigma_N^2 \quad (5a)$$

so that $\sigma_1^2+\sigma_2^2=1$, $P(y_N;y_P)$ can be deduced from expression (4*b*) to be

$$P(y_N;y_P)=\frac{2y_N}{\sigma_2^2}I_0\left[\frac{2\sigma_1 y_N y_P}{\sigma_2^2}\right]\times\exp\left\{-\frac{y_N^2+\sigma_1^2 y_P^2}{\sigma_2^2}\right\}. \quad (6a)$$

Similarly, from expression (4*a*) we obtain

$$P(y_P^c;y_P)=\frac{2y_P^c}{1-D^2}I_0\left[\frac{2Dy_P^c y_P}{1-D^2}\right]\times\exp\left\{-\frac{y_P^{c2}+D^2y_P^2}{1-D^2}\right\}. \quad (6b)$$

Expressions (6*a*) and (6*b*) may be seen to be identical in form, if we consider that the parameter D in (6*b*) plays the same role as σ_1 in (6*a*) (Note $\sigma_2^2=1-\sigma_1^2$). We shall discuss the physical significance of this later but shall now make use of the above result to work out the expression for the general case, namely for the calculation of $P(|F_N|;|F_P^c|)$. For the reasons mentioned earlier, we shall treat the problem in terms of the normalized variables. Thus, we may write the required distribution $P(y_N;y_P^c)$ as

$$P(y_N;y_P^c)=\int_0^\infty P(y_N;y_P);P(y_P;y_P^c)dy_P \quad (7)$$

where $P(y_N;y_P)$ is given by (6*a*). $P(y_P;y_P^c)$ can be obtained from (6*b*) since

$$P(y_P;y_P^c)=\frac{P(y_P^c;y_P)\,P(y_P)}{P(y_P^c)}. \quad (8)$$

This gives

$$P(y_P;y_P^c)=\frac{2y_P}{1-D^2}I_0\left[\frac{2Dy_P^c y_P}{1-D^2}\right]\times\exp\left\{-\frac{y_P^2+D^2y_P^{c2}}{1-D^2}\right\}, \quad (9)$$

which, when compared with (6*b*), shows that y_P and y_P^c have simply interchanged their roles. It might be noted that this symmetrical form of the expressions is a consequence of the normalized form of the variables we have chosen. It may be verified that such a result does not hold good, for instance, when we consider the pair of distributions of the structure amplitudes themselves, namely $P(|F_N|;|F_P|)$ and $P(|F_P|;|F_N|)$. Substituting (6*a*) and (9) in (7), we obtain

$$P(y_N;y_P^c)=\int_0^\infty\frac{4y_N y_P^c}{\sigma_2^2(1-D^2)}I_0\left[\frac{2Dy_P^c y_P}{1-D^2}\right]I_0\left[\frac{2\sigma_1 y_N y_P}{\sigma_2^2}\right]\times\exp\left\{-\frac{D^2y_P^{c2}+y_P^2}{1-D^2}\right\}\exp\left\{-\frac{\sigma_1^2y_P^2+y_N^2}{\sigma_2^2}\right\}dy_P. \quad (10)$$

This integral can be evaluated (see Appendix I), and it reduces to

$$P(y_N;y_P^c)=\frac{2y_N}{\sigma_2^2+\sigma_1^2(1-D^2)}I_0\left[\frac{2\sigma_1 Dy_P^c y_N}{\sigma_2^2+\sigma_1^2(1-D^2)}\right]\times\exp\left\{-\frac{y_N^2+\sigma_1^2D^2y_P^{c2}}{\sigma_2^2+\sigma_1^2(1-D^2)}\right\}. \quad (11)$$

Let us now define two new symbols

$$\sigma_A^2=\sigma_1^2D^2,\ \sigma_B^2=\sigma_2^2+\sigma_1^2(1-D^2). \quad (12a)$$

Note that σ_A^2 and σ_B^2 have the property that

$$\sigma_A^2+\sigma_B^2=1. \quad (12b)$$

With these new symbols, equation (11) reduces to the simple form

$$P(y_N;y_P^c)=\frac{2y_N}{\sigma_B^2}I_0\left[\frac{2\sigma_A y_N y_P^c}{\sigma_B^2}\right]\times\exp\left\{-\frac{y_N^2+\sigma_A^2y_{Pt}^{c2}}{\sigma_B^2}\right\}, \quad (13)$$

which is exactly similar to expressions (6*a*) and (6*b*). The formal identity of this final expression for the general case with expressions (6*a*) and (6*b*) is very interesting. It will be noticed that when the errors $\Delta\mathbf{r}_j$ are all zero, D equals unity, so that $\sigma_A=\sigma_1$ and $\sigma_B=\sigma_2$, and expression (13) reduces to (6*a*) of the related case, as obviously it should. On the other hand when $P=N$, we see that $\sigma_1=1$ so that $\sigma_A=D$ and $\sigma_B=\sqrt{1-D^2}$ and expression (13) reduces to that in (6*b*) (with $P=N$). This is also to be expected physically, since the discrepancies in this case are entirely due to the errors in the coordinates. Lastly, when the errors are large, D tends to zero, so that $\sigma_A\to 0$, $\sigma_B\to 1$, and expression (13) tends to

$$P(y_N;y_P^c)\to 2y_N\exp(-y_N^2). \quad (14)$$

This corresponds to the unrelated case and is dependent of y_P^c.

The distribution of the actual structure amplitudes can be obtained by a simple transformation from (13). It is given here, since it may be useful. Thus

$$P(|F_N|;|F_P^c|)=\frac{2|F_N|}{\sigma_Q^2+\sigma_P^2(1-D^2)}I_0\left[\frac{2D|F_N|\,|F_P^c|}{\sigma_Q^2+\sigma_P^2(1-D^2)}\right]\times\exp\left\{-\frac{D^2|F_P^c|^2+|F_N|^2}{\sigma_Q^2+\sigma_P^2(1-D^2)}\right\}. \quad (15)$$

Centrosymmetric case

In view of the details given for the non-centrosymmetric case, only an outline of the derivation will be

given for the present case. From equation (11) of part IV, we have

$$P(y_N; y_P) = \sqrt{\frac{2}{\pi\sigma_2^2}} \times \exp\left\{-\frac{y_N^2+\sigma_1^2 y_P^2}{2\sigma_2^2}\right\} \cosh\left[\frac{\sigma_1 y_N y_P}{\sigma_2^2}\right]. \quad (16)$$

The distribution $P(y_P; y_P^c)$ may be shown to be given by (see Appendix II):

$$P(y_P; y_P^c) = \sqrt{\frac{2}{\pi}} \frac{1}{\sqrt{(1-D^2)}} \exp\left\{-\frac{y_P^2+D^2 y_P^{c2}}{2(1-D^2)}\right\} \times \cosh\left[\frac{y_P y_P^c D}{1-D^2}\right]. \quad (17)$$

Substituting (16) and (17) in (7), we have

$$P(y_N; y_P^c) = \frac{2}{\pi\sigma_2\sqrt{1-D^2}} \int_0^\infty \exp\left\{-\frac{y_N^2+\sigma_1^2 y_P^2}{2\sigma_2^2}\right\} \times \exp\left\{-\frac{y_P^2+D^2 y_P^{c2}}{2(1-D^2)}\right\} \cosh\left[\frac{y_N y_P \sigma_1}{\sigma_2^2}\right] \times \cosh\left[\frac{y_P y_P^c D}{1-D^2}\right] dy_P . \quad (18)$$

The integral can be evaluated (See Appendix III) and it reduces to

$$P(y_N; y_P^c) = \sqrt{\frac{2}{\pi}} \frac{1}{[\sigma_2^2+\sigma_1^2(1-D^2)]^{\frac{1}{2}}} \times \exp\left\{-\frac{y_N^2+\sigma_1^2 D^2 y_P^{c2}}{2[\sigma_2^2+\sigma_1^2(1-D^2)]}\right\} \quad (19)$$

$$\times \cosh\left[\frac{D\sigma_1 y_N y_P^c}{[\sigma_2^2+\sigma_1^2(1-D^2)]}\right] = \sqrt{\frac{2}{\pi\sigma_B^2}} \times \exp\left\{-\frac{y_N^2+\sigma_A^2 y_P^{c2}}{2\sigma_B^2}\right\} \cosh\left[\frac{\sigma_A y_N y_P^c}{\sigma_B^2}\right]. \quad (20)$$

The identity of form of expressions (16), (17) and (20) is noteworthy, and follows exactly the results derived for the non-centrosymmetric case.

As before, for convenience of reference, the distribution of the actual structure amplitudes obtained from (20) is given below:

$$P(|F_N|; |F_P^c|) = \sqrt{\frac{2}{\pi}} \frac{1}{[\sigma_Q^2+\sigma_P^2(1-D^2)]^{\frac{1}{2}}} \times \exp\left\{-\frac{|F_N|^2+\sigma_P^2 D^2|F_P^c|^2}{2[\sigma_Q^2+\sigma_P^2(1-D^2)]}\right\} \times \cosh\left[\frac{|F_N|\,|F_P^c| D}{\sigma_Q^2+\sigma_P^2(1-D^2)}\right] \quad (21)$$

3. The distribution of the difference, product and quotient of the quantities F_N and F_P^c

The formal identity of the expressions for the distributions in the present problem with those considered in earlier parts makes it possible to use many of the results obtained earlier, with an appropriate reinterpretation of the parameters concerned. However, we shall discuss mainly the results concerning the normalized variables (parts II to IV) in view of the simplicity of interpretation that ensues. Moreover, in order to avoid confusion in notation, the variables for the present case will all be denoted by a superscript c. This is used to denote that it corresponds to the general case in which the P (out of the N) atoms used in the structure factor calculation have *finite* errors in their coordinates. Thus, the normalized product, quotient and difference variables are denoted respectively by Z^c, v^c and δ^c. Their distributions are all obtained by simply replacing the parameter σ_1 in the corresponding expressions for the related case of the earlier parts (parts II to IV) by the parameter σ_A. For convenience, a summary of the formulae with this simple alteration is given in Table 1. It is not necessary to reproduce the various curves. They are the same as the figures given in the earlier parts, with the only difference that the ordinate is to be construed as the appropriate variable for the general case (namely Z^c, v^c and δ^c as the case may be in Parts II, III and IV respectively), while the abscissa should in all cases be treated as σ_A^2 instead of σ_1^2.

Needless to say, the other parameters and tests based on the normalized variables such as R_1, $\langle Z\rangle$ *etc.* could also be taken over for the present case, with the parameter σ_1 simply replaced by σ_A.

4. Discussion

Before entering into a discussion of the results and their applications, we might make a few remarks here regarding the equivalence result obtained in § 2. A comparison of expressions (4*a*), (4*b*) and (15) shows that instead of the parameter σ_Q^2 in (4*b*), $\sigma_Q^2+\sigma_P^2(1-D^2)$ occurs in (15), that is the sum of the terms σ_Q^2 and $\sigma_P^2(1-D^2)$ occurring individually in (4*b*) and (4*a*) respectively. This in fact is to be expected from the physical nature of the origin of the difference between F_N and F_P^c as shown in Fig. 1. It will be noticed that this difference $\Delta = F_N - F_P^c$ arises from two components (i) $\Delta F_P = F_P - F_P^c$ and (ii) F_Q. Taking ΔF_P alone, it is convenient (Luzzati, 1952) to decompose it into two vectors; one parallel to F_P equal to $(D-1)F_P$ and the other, say V_P, equal to $\Delta F_P - (D-1)F_P^c$ (Fig. 1). It can be shown (Luzzati, 1952) that the distribution of the magnitude of the component V_P has a Rayleigh distribution* given by

$$P(|V_P|)\,d|V_P| = \frac{2|V_P|}{\sigma_P^2(1-D^2)} \exp\left\{-\frac{|V_P|^2}{\sigma_P^2(1-D^2)}\right\} \times d|V_P| \quad (22)$$

* We use the term Rayleigh distribution to denote a distribution of the type $P(r)dr = (2r/\alpha^2)\exp(-r^2/\alpha^2)dr$ (*e.g.* see p. 181 of Parzen, 1960). The distribution of the magnitude $r = \sqrt{(x^2+y^2)}$ of the resultant vector **r** in a two-dimensional random walk problem has essentially this distribution, where the rectangular components x and y have each a Gaussian distribution with a characteristic $\sigma^2 = \alpha^2/2$.

and is thus independent of the orientation of the vector in the Argand plane and also of the value of F_P. This could be compared with the distribution of $|F_Q| = |F_N - F_P|$, which is well-known to be

$$P(|F_Q|)\,d|F_Q| = \frac{2|F_Q|}{\sigma_Q^2}\exp\left\{-\frac{|F_Q|^2}{\sigma_Q^2}\right\}d|F_Q| \qquad (23)$$

(see also part I). The identity of forms of expressions (22) and (23) is noteworthy. Also we have

$$F_P = DF_P^c + V_P,\ F_N = F_P + F_Q \qquad (24)$$

so that V_P and F_Q play identically parallel roles in the two distributions $P(F_P;\ DF_P^c)$ and $P(F_N;\ F_P)$ with the only difference that we have actually worked out $P(F_P;\ F_P^c)$ which can be obtained from $P(F_P;\ DF_P^c)$ in a very simple way since D is a constant.

We now have from (24*a*) and (24*b*), the result that $F_N = DF_P^c + V_P + F_Q$ where the distributions of V_P and F_Q are given by (22) and (23) respectively. Since the vectors V_P and F_Q arise from independent causes, their sum will have the same distribution as (22) and (23), with the only difference that the "σ^2" will be the sum of those occurring in (22) and (23), namely equal to $\sigma_Q^2 + \sigma_P^2(1-D^2)$. When these factors are taken into account, the identity of forms of expressions (4*a*), (4*b*) and (15), as well as the relations between the parameters occurring in them, become obvious.

The normalized reliability index

In the light of the results obtained in the earlier sections of this paper, the interpretation and use of some of the tests developed in the earlier parts become particularly elegant. In particular, let us consider the normalized reliability index R_1. If we define R_1 as

$$R_1 = \Sigma\,\big||F_N| - |F_P^c|\big|/\sigma_1\big|/\Sigma\,|F_N|\,, \qquad (25)$$

all the discussions regarding this quantity given in part IV can be taken over completely, with the only difference that σ_1 of part IV has now to be replaced by σ_A. Thus, for instance, for a centrosymmetric crystal we have (analogous to equation (28) of part IV)

$$R_1 = \sqrt{2(1+\sigma_A)} + \sqrt{2(1-\sigma_A)} - 2\,, \qquad (26)$$

Table 1(*a*). *Expressions for the probability distribution functions*: non-centrosymmetric case*

Variable	Function	General	Unrelated
$Z^c = y_N\, y_P{}^c$	$P(Z^c)$	$4\,Z^c\,I_0\left(\frac{2Z^c\sigma_A}{\sigma_B{}^2}\right)K_0\left(\frac{2Z^c}{\sigma_B{}^2}\right)$	$4Z^c\,K_0\,(2Z^c)$
$v^c = y_N/y_p{}^c$	$P(v^c)$	$\frac{2v^c\sigma_B{}^2(1+v^{c\,2})}{[(1+v^{c2})^2 - 4\sigma_A{}^2 v^{c\,2}]^{3/2}}$	$\frac{2v^c}{(1+v^{c\,2})^2}$
$\delta^c = (y_N - y_p{}^c)$	$p(\delta^c)$†	$2\exp(-\delta^{c2}/\sigma_B{}^2)\int_{0\text{ or }\lvert\delta^c\rvert}^{\infty} x\exp(-x)\,I_0(\sigma_A\,x)\,dy_p{}^c$	$2\exp(-\delta^{c2}/\sigma_B{}^2)\int_{0\text{ or }\lvert\delta^c\rvert}^{\infty} 2y_p{}^c(y_p{}^c+\delta^c) \times\exp\{-2y_p{}^c(y_p{}^c+\delta^c)\}dy_p{}^c$

where $x = 2y_p{}^c(y_p{}^c + \delta^c)/\sigma_B{}^2$

* The expressions for the related case are not listed. They are obtained by putting $D=1$ (*i.e.* $\sigma_A = \sigma_1$) in the expressions for the general case. The unrelated case is obtained by putting $D=0$ ($\sigma_A = 0$) in the general case.

† The lower limit is 0 for $\delta^c > 0$ and $|\delta^c|$ for $\delta^c < 0$.

Table 1(*b*). *Expressions‡ for the probability distribution functions: centrosymmetric case*

Variable	Function	General	Unrelated
$Z^c = y_N y_p{}^c$	$P(z^c)$	$\frac{2}{\pi\sigma_B}\cosh\left(\frac{Z^c\sigma_A}{\sigma_B{}^2}\right)K_0\left(\frac{Z^c}{\sigma_B{}^2}\right)$	$\frac{2}{\pi}K_0(Z^c)$
$v^c = y_N\lvert y_p{}^c$	$P(v^c)$	$\frac{2}{\pi}\,\frac{\sigma_B(1+v^{c2})}{[(1+v^{c2})^2 - 4\sigma_A{}^2 v^{c\,2}]}$	$\frac{2}{\pi}\,\frac{1}{(1+v^{c2})}$
$\delta^c = (y_N - y_p{}^c)$	$P(\delta^c)$	$\frac{1}{2\sqrt{\pi}}\left[\frac{\exp\{-\delta^{c2}/4\,(1+\sigma_A)\}}{\sqrt{(1+\sigma_A)}}[1-\operatorname{erf}\delta^c/2\sqrt{(1-\sigma_A)}] + \frac{\exp\{-\delta^{c2}/4(1-\sigma_A)\}}{\sqrt{(1-\sigma_A)}}[1-\operatorname{erf}\delta^c/2\sqrt{(1+\sigma_A)}]\right]$	$\frac{1}{\sqrt{\pi}}\exp\left(-\frac{\delta^{c2}}{4}\right)(1-\operatorname{erf}\delta^{c2}/2)$

‡ See first footnote to Table 1(*a*)

which reduces to the equation (28) of part IV when the errors $\Delta\mathbf{r}_j$ are all zero, *i.e.* $D=1$, so that $\sigma_A=\sigma_1$. On the other hand when $P=N$, we have $\sigma_1=1$ and $\sigma_A=D$. In this case, R_1 reduces to the conventional reliability index R and we obtain

$$R = \sqrt{2(1+D)} + \sqrt{2(1-D)} - 2 . \qquad (27)$$

This expression is identical with that derived by Luzzati (1952). Unfortunately, for a non-centrosymmetric crystal, no such elegant expression can be obtained for R_1. However, the variation of R_1 with σ_A for the two cases is shown in Table 2.

Table 2. *Normalized reliability index R_1 as a function of σ_A*

σ_A	Centrosymmetric case	Non-centrosymmetric case
0·0	0·828	0·586
0·1	0·826	0·583
0·2	0·815	0·576
0·3	0·798	0·563
0·4	0·770	0·544
0·5	0·733	0·516
0·6	0·685	0·479
0·7	0·620	0·434
0·8	0·532	0·369
0·9	0·399	0·273
1·0	0	0

Vice versa, if R_1 is evaluated from the data for a crystal, then σ_A can be estimated from Table 2. Since σ_1 is known, $\sigma_A/\sigma_1=D$ would provide an estimate of the errors in the atomic positions of the P-atoms used in the calculation of the structure factors. This will be useful in the early stages of a structure analysis, when a small number of atoms have been located, to check the accuracy of their locations. The tests indicated in the earlier parts provide only the values of the parameter concerned (such as R_1) for the related case (or correct positions) and the unrelated case (completely wrong positions). Using these one could previously make only a sort of an estimate of the errors. Now, on the other hand, it is possible to obtain a value of the parameter D, which is directly related to the errors $|\Delta\mathbf{r}_j|$.

In fact, from the definition of D, it is clear that it is a function of the reciprocal vector $\mathbf{H}$, and hence of the Bragg angle θ. Hence, it becomes possible to separate the effect due to errors in the atomic coordinates by studying the variation of σ_A $(=\sigma_1 D)$ with θ. In this, the variation of σ_1 with θ may not be as strongly dependent on θ as D, and in any case it is known. This aspect is being tested practically and will be reported elsewhere.

It may also be mentioned that the parameter $\langle Z\rangle$ and its analogue $\langle Z^c\rangle$ may also be used to estimate the value of σ_A. Its use in structure analysis for judging the accuracy of atomic parameters seems to be an interesting possibility. The calculation of $\langle Z^c\rangle$ from a set of experimental data does not require a knowledge of the scale factor, and the value of σ_A may be obtained by use of the curves shown in Fig. 3 of part II or Tables I and II of Srinivasan, Sarma and Ramachandran (1963*b*). However, the range of variation of $\langle Z^c\rangle$ is less than that of R_1, so that it is likely to be useful only in the very early stages of structure analysis.

APPENDIX I

Expression (10) can be written:

$$P(y_N;y_P^c) = \frac{4y_N}{\sigma_2^2(1-D^2)} \exp\left\{-\frac{D^2\sigma_2^2 y_P^{c2} + (1-D^2)y_N^2}{\sigma_2^2(1-D^2)}\right\} \times \int_0^\infty y_P \exp\left\{-y_P^2\left[\frac{\sigma_2^2+\sigma_1^2(1-D^2)}{\sigma_2^2(1-D^2)}\right]\right\} I_0\left[\frac{2Dy_P^c y_P}{1-D^2}\right] \times I_0\left[\frac{2\sigma_1 y_N y_P}{\sigma_2^2}\right] dy_P . \qquad (A1)$$

The integral in the above equation is of the form

$$\int_0^\infty x\, I_0(ax)\, I_0(bx) \exp(-p^2x^2)\, dx . \qquad (A2)$$

To evaluate this we use the result (Watson, 1944, p. 395)

$$\int_0^\infty \exp(-p^2t^2)\, J_\nu(at)\, J_\nu(bt)\, t\, dt = \frac{1}{2p^2}\exp\left\{-\frac{a^2+b^2}{4p^2}\right\} I_\nu\left(\frac{ab}{2p^2}\right). \qquad (A3)$$

With appropriate substitutions (*A*2) reduces to

$$\frac{1}{2p^2}\exp\left\{\frac{a^2+b^2}{4p^2}\right\} I_0\left(\frac{ab}{2p^2}\right). \qquad (A4)$$

Substitution of

$$y_P = x, \quad p^2 = \frac{\sigma_2^2+\sigma_1^2(1-D^2)}{\sigma_2^2(1-D^2)}, \quad a = \frac{2y_P^c D}{(1-D^2)}, \quad b = \frac{2y_N\sigma_1}{\sigma_2^2}$$

leads us to expression (11) given in the text.

APPENDIX II

For the centrosymmetric case, the law of distribution of $\Delta F_P = F_P^c - F_P$ is given expression (34) of Luzzati (1952), which in our notation becomes

$$P(\Delta F_P) = \frac{1}{\sqrt{2\pi\sigma_P^2(1-D^2)}} \exp\left\{-\frac{[\Delta F_P-(D-1)F_P]^2}{2\sigma_2^2(1-D^2)}\right\}. \qquad (A5)$$

This is obviously the conditional distribution of ΔF_P, given F_P. Since $\Delta F_P = F_P^c - F_P$ we see that the law of distribution of F_P^c should be

$$P(F_P^c;\, F_P) = \frac{1}{\sqrt{2\pi\sigma_P^2(1-D^2)}} \exp\left\{-\frac{(F_P^c - DF_P)^2}{2\sigma_2^2(1-D^2)}\right\}. \qquad (A6)$$

This consists of two possibilities, namely F_P and F_P^c having the same or opposite signs, so that separating out these two, we have

$$P(F_P^c;\, F_P) = \frac{1}{\sqrt{2\pi\sigma_2^2(1-D^2)}}\left[\exp\left\{\frac{(|F_P^c|-D|F_P|)^2}{2\sigma_P^2(1-D^2)}\right\} + \exp\left\{-\frac{(|F_P^c|+D|F_P|)^2}{2\sigma_P^2(1-D^2)}\right\}\right] \quad (A7)$$

$$= \sqrt{\frac{2}{\pi}}\frac{1}{[\sigma_P^2(1-D^2)]^{\frac{1}{2}}}\exp\left\{-\frac{|F_P^c|^2+D^2|F_P|^2}{2\sigma_P^2(1-D^2)}\right\} \times \cosh\left[\frac{D|F_P|\,|F_P^c|}{\sigma_P^2(1-D^2)}\right]. \quad (A8)$$

Since this involves only the magnitudes $|F_P^c|$ and $|F_P|$ it is seen that this is also the distribution, $P(F_P^c|;\, |F_P|)$.

Making the usual transformation in (*A*8), namely writing $y_N = |F_N|/\sigma_N$ and $y_P^c = |F_P|^c/\sigma_P$, we obtain expression (17) given in the text.

APPENDIX III

The desired integral (18) can be written:

$$\frac{2}{\pi\sigma_2(1-D^2)}\exp\left\{-\frac{(1-D^2)\,y_N^2+\sigma_2^2D^2\,y_P^{c2}}{2\sigma_2^2(1-D^2)}\right\} \times \int_0^\infty \exp\left\{-\frac{[(1-D^2)\,\sigma_1^2+\sigma_2^2]\,y_P^2}{2\sigma_2^2(1-D^2)}\right\}\cosh\left[\frac{y_N y_P\,\sigma_1}{\sigma_2^2}\right] \times \cosh\left[\frac{y_P^c y_P D}{(1-D^2)}\right] dy_P\,. \quad (A9)$$

The integral in (*A*9) above is of the form

$$\int_0^\infty \exp(-p^2x^2)\cosh(ax)\cosh(bx)\,dx\,, \quad (A10)$$

where

$$p^2 = \frac{(1-D^2)\,\sigma_1^2+\sigma_2^2}{2\sigma_2^2(1-D^2)},\quad a = \sigma_1 y_N/\sigma_2^2,\quad b = y_P^c D/(1-D^2)\,.$$

Since we have the relation

$$I_{-1/2}(z) = \sqrt{\frac{2}{\pi z}}\cosh z\,, \quad (A11)$$

(*A*10) can be written in the form

$$\frac{\pi}{2}\sqrt{ab}\int_0^\infty \exp(-p^2x^2)\,I_{-1/2}(ax)\,I_{-1/2}(bx)\,x\,dx\,. \quad (A12)$$

The integral in (*A*12) is similar to (*A*2) considered in Appendix I, with the only difference that the order of the Bessel function is $-\frac{1}{2}$ in the former, while it is zero in the latter. Thus, using the same result (*A*3) of Appendix I, (*A*12) reduces to

$$\frac{\pi}{2}\sqrt{ab}\,\frac{1}{2p^2}\exp\left\{\frac{a^2+b^2}{4p^2}\right\} I_{-1/2}\left(\frac{ab}{2p^2}\right). \quad (A13)$$

Again making use of (*A*1) in (*A*13) and substituting for a, b, p^2, *etc.*, we obtain the expression (19) given in the text.

References

LUZZATI, V. (1952). *Acta Cryst.* **5**, 802.

PARZEN, E. (1960). *Modern Probability Theory and its Applications*. New York: John Wiley.

RAMACHANDRAN, G. N. & SRINIVASAN, R. (1963). *Nature, Lond.* **20**, 1090.

RAMACHANDRAN, G. N., SRINIVASAN, R. & SARMA, V. R. (1963). *Acta Cryst.* **16**, 662.

SRINIVASAN, R. & RAMACHANDRAN, G. N. (1965). *Acta Cryst.* **19**, 1003.

SRINIVASAN, R., SARMA, V. R. & RAMACHANDRAN, G. N. (1963*a*). *Acta Cryst.* **16**, 1151.

SRINIVASAN, R., SARMA, V. R. & RAMACHANDRAN, G. N. (1963*b*). In *Crystallography and Crystal Perfection*, Ed. G. N. Ramachandran, p. 83. London: Academic Press.

SRINIVASAN, R., SUBRAMANIAN, E. & RAMACHANDRAN, G. N. (1964). *Acta Cryst.* **17**, 1010.

WATSON, G. N. (1944). *A Treatise on the Theory of Bessel Functions*. Cambridge Univ. Press.

Acta Cryst. (1965). **19**, 1014

The Absorption Correction in Crystal Structure Analysis

BY J. DE MEULENAER AND H. TOMPA

Union Carbide European Research Associates, Brussels 18, *Belgium*

(*Received* 29 *June* 1964 *and in revised form 15 November* 1964)

An exact method for computing the absorption correction for any polyhedral crystal is described. First an analytical formula is derived for the contribution to the diffracted intensity from a tetrahedron in which the path length of the rays is a linear function of the coordinates of the diffracting element, and it is then shown how the crystal is to be divided into such tetrahedra. A computer program for the IBM 1620 machine to compute the absorption correction is described.

Introduction

In crystal-structure analysis the basic observed quantities are the intensities of the *hkl* reflexions. A number of corrections (Lorentz and polarization factors, absorption correction) have to be applied to the observed intensities before they can be used as the basis of a structure determination. The absorption correction