Acta Cryst. (1963). **16**, 662

Probability Distribution Connected with Structure Amplitudes of Two Related Crystals. I. Probability Distribution of the Difference

By G. N. Ramachandran, R. Srinivasan and V. Raghupathy Sarma

Department of Physics, University of Madras, Madras 25, India

(*Received* 20 *August* 1962)

The paper deals with the probability distribution of the differences in structure amplitudes of two crystals, containing N atoms and P atoms respectively ($P < N$). Two cases are considered: (a) When P forms a part of N and (b) when the two sets of atoms are entirely independent. Both centrosymmetric and non-centrosymmetric cases are discussed. Based on these, a method is proposed for testing the isomorphism of two crystals, purely from their intensity data. A new method of determining the relative scale factor for the data from two isomorphous crystals also follows from the results of this paper.

1. Introduction

This paper deals with the problem of the probability distribution of the structure amplitudes of two crystals. We shall be concerned essentially with deriving expressions for the distribution of the difference in structure amplitudes of two crystals both when they are related to and independent of each other. The results thus obtained lead to a criterion for testing, in any practical case, the isomorphism of two crystals.

It may be of interest to mention how the present problem came to be considered. It was suggested by Ramachandran & Srinivasan (1960) that the direct probability distribution function, $P(y)$, would form a good basis for a test for finding the presence or absence of a centre of symmetry in a crystal and is probably better than the cumulative function, $N(z)$, of Howells, Phillips & Rogers (1950), because of the essential dissimilarity in the nature of the curves for the two cases with this function. These results were tested out more fully later (Srinivasan, 1960), where the effect of the presence of a small number of heavy atoms in the structure on the distribution $P(y)$ was also considered. During these studies it was noticed that the presence of a few heavy atoms in the structure invariably tended partly to destroy the dissimilarity of the ideal $P(y)$ curves. It was suggested to the authors that it would probably be worthwhile working out the distribution function for the structure amplitude with the heavy atom contribution removed, that is, to find the distribution of $(|F_N|-|F_H|)$ where the subscripts N and H refer to the entire structure consisting of N atoms, and the heavy atoms alone, respectively.* In practice it should be possible to compute this since the heavy atom position can often be determined with ease at the beginning of a structure analysis. However, we can treat the problem in its more general form, namely, to consider the distribution of $(|F_N|-|F_P|)$ where P now refers to a group of P known atoms, which form a part of the whole structure. The treatment of this situation is fairly straightforward, but it turned out that the removal of the known part $|F_P|$ from the structure amplitude $|F_N|$ does not in any way enhance the dissimilarity of the two original distributions, but in fact affects them adversely. Hence, as a statistical test, this does not improve the situation. However, the study led to a number of other interesting results. Thus, the quantity $(|F_N|-|F_P|)$ can equally well be taken to represent the difference in the structure amplitudes of two ideally isomorphous crystals (see section 3). In fact, by viewing the problem in this manner, the authors were also led to consider another situation, namely one in which N and P are completely independent.

The various formulae are derived in the next section and section 3 contains a discussion of the results and their possible applications.

* This suggestion arose as a result of the discussions in a small symposium arranged at Madras in which Prof. A. J. C. Wilson took part. The authors are grateful to Prof. S. Ramaseshan for pointing out to them this interesting possibility.

2. Derivation of the formulae

Consider a structure containing N atoms, P of which are assumed to be known ($N=P+Q$). Let us also assume that the group P contains a sufficiently large number of atoms so that the distribution of $|F_P|$ (and therefore that of $|F_N|$ also) would follow the ideal centrosymmetric or the non-centrosymmetric one, as the case may be. We first observe that, in the equation

$$F_N=F_P+F_Q\,, \tag{1}$$

the quantities that are available are $|F_N|$ and F_P. Denoting for convenience $(|F_N|-|F_P|)$ by Δ, the distribution for Δ can be worked out using the following general theorem in probability. If $z=x+y$, where x, y and z are random variables, then

$$P(z) = \int P_1(x)P_2[(z-x);\ x]dx\ , \qquad (2)$$

where $P_1(x)$ is the probability distribution function for the variable x and $P_2(y; x)$ is the conditional probability that y lies between y and $y+dy$ given the value x for the first variable. The limits of integration are determined by the appropriate domain in which the function $P_1(x)$ is defined. Equation (2) can also be written in its other equivalent form by interchanging the roles of x and y. When the variables x and y are independent, $P_2(y; x)$ equals $P_2(y)$ for all values of y, so that equation (2) reduces to the well known convolution integral. It is clear that, in our problem, the two variables $|F_N|$ and $|F_P|$ are not independent and equation (2) is to be applied in its original form. We shall refer to this case, in which $|F_N|$ and $|F_P|$ are not independent, as the case of 'related structure amplitudes', since later we shall be dealing with the case when they are independent, which will be called the 'unrelated case'.

2·1. *Case of related structure amplitudes*

(*a*) *Non-centrosymmetric case.*—Since the variable we are interested in is $\Delta=(|F_N|-|F_P|)$, we have from equation (2)

$$P(\Delta) = \int P_1(|F_P|)P_2[(\Delta+|F_P|);\ |F_P|]\,d|F_P|\ . \qquad (3)$$

where $P_2(|F_N|; |F_P|)$ is the conditional probability of having a value $|F_N|$ given $|F_P|$. When the structure is non-centrosymmetric, it can be shown that (see *e.g.* Sim, 1958)

$$P_2(|F_N|;\ |F_P|) = \frac{2|F_N|}{\sigma_Q^2}\exp\left\{-\frac{|F_N|^2+|F_P|^2}{\sigma_Q^2}\right\}I_0\left(\frac{2|F_N||F_P|}{\sigma_Q^2}\right). \qquad (4)$$

The function $P_1(|F_P|)$ is given by (Srinivasan, 1960)

$$P_1(|F_P|) = \frac{2|F_P|}{\sigma_P^2}\exp\left\{-\frac{|F_P|^2}{\sigma_P^2}\right\}. \qquad (5)$$

Here σ_P^2 and σ_Q^2 are the mean square values of the structure amplitudes $|F_P|$ and $|F_Q|$ and $I_0(x)$ is the Bessel function of imaginary argument. It is clear that Δ can have any value between $-\infty$ and $+\infty$, though both $P_1(|F_P|)$ and $P_2(|F_N|)$ exist only in the range 0 to ∞. Accordingly the lower limit of integration in (3) becomes 0 for $\Delta>0$ and $|\Delta|$ for $\Delta<0$. We therefore have

$$P(\Delta) = \frac{\Delta}{\sigma_P^2\sigma_Q^2}\exp\left\{-\frac{\Delta^2}{\sigma_Q^2}\right\}\int_{0\text{ or }|\Delta|}^{\infty}|F_P|(\Delta+|F_P|) \times\exp\left\{-\frac{|F_P|^2(\sigma_Q^2+2\sigma_P^2)+2\sigma_P^2\Delta|F_P|}{\sigma_P^2\sigma_Q^2}\right\} \times I_0\left[\frac{2|F_P|(\Delta+|F_P|)}{\sigma_Q^2}\right]d|F_P|\ . \qquad (6)$$

It is convenient to work out the results in terms of the normalized variable $w=\Delta/\sigma_N$, since the final expressions take simple forms in terms of this quantity. Similarly, define

$$\sigma_1^2=\sigma_P^2/\sigma_N^2;\ \sigma_2^2=\sigma_Q^2/\sigma_N^2,\ \text{ with }\ \sigma_1^2+\sigma_2^2=1\ . \qquad (7)$$

The distribution for w then takes the form:

$$P(w) = \frac{2}{\sigma_1^2}\exp-\frac{w^2}{\sigma_2^2}\int_0^\infty pe^{-p}I_0(p) \times\exp\left(-\frac{x^2}{\sigma_1^2}\right)dx;\ w>0 \qquad (8a)$$

$$P(w) = \frac{2}{\sigma_1^2}\exp-\frac{w^2}{\sigma_2^2}\int_{|w|}^\infty pe^{-p}I_0(p) \times\exp\left(-\frac{x^2}{\sigma_1^2}\right)dx;\ w<0 \qquad (8b)$$

where

$$x=|F_P|/\sigma_N \qquad (9a)$$

and

$$p=2x(w+x)/\sigma_2^2\ . \qquad (9b)$$

The above expression is easy to work out since tables of $e^{-p}I_0(p)$ are available. The function $P(w)$ has been evaluated by numerically integrating equation (8) for various values of σ_1^2. The results are given in Fig. 1.

(*b*) *Centrosymmetric case*

When the structure is centrosymmetric, we have (Srinivasan, 1960)

$$P_1(|F_P|) = \left(\frac{2}{\pi\sigma_P^2}\right)^{\frac{1}{2}}\exp\left\{-\frac{|F_P|^2}{2\sigma_P^2}\right\} \qquad (10)$$

and

$$P_2(|F_N|;\ |F_P|) = \frac{1}{\sqrt{(2\pi\sigma_Q^2)}} \times\left[\exp\left\{-\frac{(|F_N|-|F_P|)^2}{2\sigma_Q^2}\right\}+\exp\left\{-\frac{(|F_N|+|F_P|)^2}{2\sigma_Q^2}\right\}\right]. \qquad (11)$$

Substituting these in (3), we get

$$P(\Delta) = \int_{0\text{ or }|\Delta|}^{\infty}\frac{1}{\sigma_P\sigma_Q\pi}\exp\left(-\frac{|F_P|^2}{2\sigma_P^2}\right)\left[\exp\left\{-\frac{\Delta^2}{2\sigma_Q^2}\right\} +\exp\left\{-\frac{(\Delta+2|F_P|)^2}{2\sigma_Q^2}\right\}\right]d|F_P|\ . \qquad (12)$$

The integral (12) is expressible in terms of error functions. As before, substituting $w=\Delta/\sigma_N$ we get

$$P(w) = \frac{\exp(-w^2/2\sigma_2^2)}{\sqrt{(2\pi\sigma_2^2)}}+\frac{\exp(-w^2/2\sigma'^2)}{\sqrt{(2\pi\sigma'^2)}} \times\left[1-\operatorname{erf}\frac{\sqrt{(2)}\,\sigma_1 w}{\sigma'\sigma_2}\right],\ w>0 \qquad (13a)$$

$$P(w) = \frac{\exp(-w^2/2\sigma_2^2)}{\sqrt{(2\pi\sigma_2^2)}}\left[1-\operatorname{erf}\frac{|w|}{\sqrt{(2)}\sigma_1}\right] + \frac{\exp(-w^2/2\sigma'^2)}{\sqrt{(2\pi\sigma'^2)}}\left[1-\operatorname{erf}\frac{(\sigma_2^2+2\sigma_1^2)w}{\sqrt{(2)}\sigma_1\sigma_2\sigma'}\right],\ w<0 \quad (13b)$$

where

$$\sigma'^2 = \sigma_2^2+4\sigma_1^2 \text{ and } \operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}dt\,. \quad (14)$$

The nature of the function $P(w)$ for different values of σ_1^2 may be seen from Fig. 2.

2·2. *The case of unrelated structure amplitudes*

Let us now consider the situation when N and P are completely independent of each other. Equation (3) then reduces to

$$P(\Delta) = \int P_1(|F_P|)P_2(\Delta+|F_P|)d|F_P|\,. \quad (15)$$

(*a*) *Non-centrosymmetric case.*—When the structure is non-centrosymmetric, $P_1(|F_P|)$ is given by equation (5) and

$$P_2(|F_N|) = \frac{2|F_N|}{\sigma_N^2}\exp\left\{-\frac{|F_N|^2}{\sigma_N^2}\right\}. \quad (16)$$

Substituting equations (5) and (16) in equation (15), we obtain

$$P(\Delta) = \frac{4}{\sigma_P^2\sigma_N^2}\int_0^\infty |F_P|(\Delta+|F_P|) \times\exp\left\{-\frac{|F_P|^2}{\sigma_P^2}\right\}\exp\left\{-\frac{(\Delta+|F_P|)^2}{\sigma_N^2}\right\}d|F_P|\,.$$

This is expressible in terms of incomplete gamma functions. As before, substituting $w=\Delta/\sigma_N$ we get

$$P(w) = \frac{2\sigma_1}{\Sigma^3}\exp\left(-\frac{w^2}{\Sigma^2}\right)\left\{\frac{\sqrt{\pi}}{2}\left[1-I\left(\sqrt{\tfrac{2}{3}}\frac{k^2w^2}{\Sigma^2};\tfrac{1}{2}\right)\right] + \frac{(1-\sigma_1^2)}{\Sigma\sigma_1}w\left[1-I\left(\frac{k^2w^2}{\Sigma^2};0\right)\right] - \frac{\sqrt{(\pi)}w^2}{\Sigma^2}\left[1-I\left(\sqrt{2}\cdot\frac{k^2w^2}{\Sigma^2};-\tfrac{1}{2}\right)\right]\right\} \quad (17)$$

with

$$k^2=\sigma_1^2 \text{ for } w>0,\quad k^2=1/\sigma_1^2 \text{ for } w<0\,, \quad (18a)$$

$$\Sigma^2=\sigma_1^2+1\,, \quad (18b)$$

and $I(x;p)$ is the incomplete gamma function (Pearson, 1922) defined by

$$I(x;p) = \int_0^{x\sqrt{(p+1)}} e^{-t}t^p dt/\Gamma(p+1)\,. \quad (19)$$

From a practical point of view, only the cases when σ_1^2 is nearly equal to unity will be of interest to us (see section 3) and therefore calculations were performed only in this region and the results are given in Fig. 3 for $\sigma_1^2=0{\cdot}8$ and $\sigma_1^2=1{\cdot}0$.

(*b*) *Centrosymmetric case.*—For this case, $P_1(|F_P|)$ is given by equation (10) and

$$P_2(|F_N|) = \left(\frac{2}{\pi\sigma_N^2}\right)^{\frac{1}{2}}\exp\left(-\frac{|F_N|^2}{2\sigma_N^2}\right). \quad (20)$$

$P(w)$ now takes the forms

$$P(w) = \left(\frac{2}{\pi\Sigma^2}\right)^{\frac{1}{2}}\left(\exp-\frac{w^2}{2\Sigma^2}\right)\times\left[1-\operatorname{erf}\frac{\sigma_1 w}{\sigma_2\Sigma\sqrt{2}}\right];\quad w>0 \quad (21a)$$

$$P(w) = \left(\frac{2}{\pi\Sigma^2}\right)^{\frac{1}{2}}\left(\exp-\frac{w^2}{2\Sigma^2}\right)\times\left[1-\operatorname{erf}\frac{\sigma_2 w}{\sigma_1\Sigma\sqrt{2}}\right];\quad w<0 \quad (21b)$$

where the notations are the same as above. The distribution functions for this non-centrosymmetric unrelated case are given also for $\sigma_1^2=0{\cdot}8$ and $\sigma_1^2=1$ in Fig. 4.

3. Discussion

Considering first Figs. 1 and 2, which give the probability distribution function $P(w)$ for the *difference* in structure amplitudes of two related crystals, it will

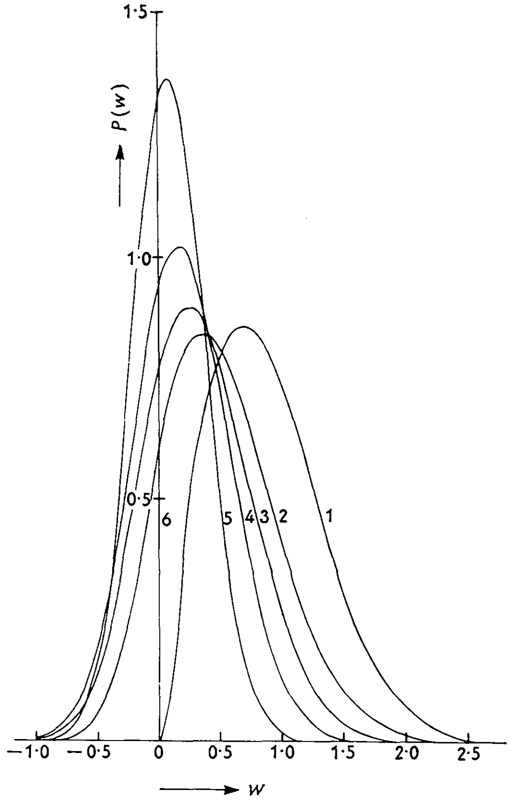

Fig. 1. Probability distribution function $P(w)$ for the related non-centrosymmetric case, corresponding to $\sigma_1^2=$ 0(1); 0·2(2); 0·4(3); 0·6(4); 0·8(5); 1·0(6).

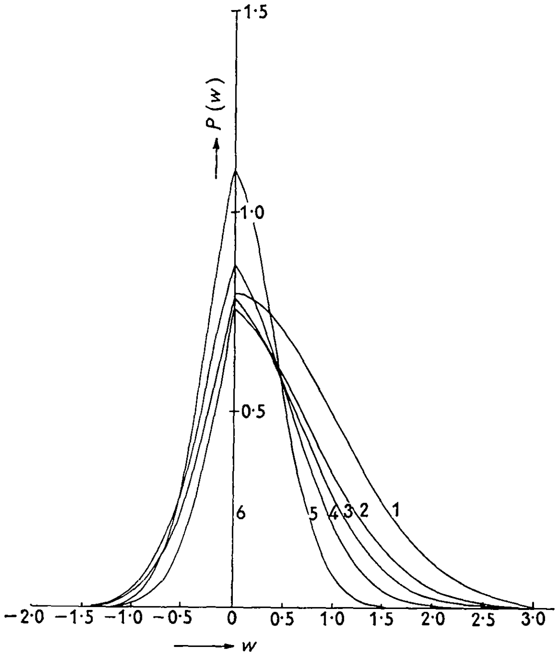

Fig. 2. Probability distribution function $P(w)$ for the related centrosymmetric case, corresponding to $\sigma_1^2=0(1)$; $0{\cdot}2(2)$; $0{\cdot}4(3)$; $0{\cdot}6(4)$; $0{\cdot}8(5)$; $1{\cdot}0(6)$.

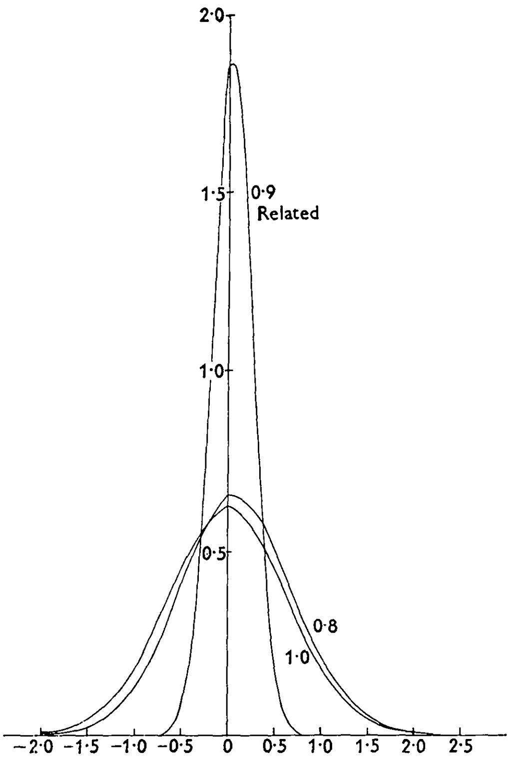

Fig. 3. Probability distribution function $P(w)$ for the unrelated non-centrosymmetric case. Only the curves for $\sigma_1^2=1{\cdot}0$ and $0{\cdot}8$ are shown. The corresponding curve for the related case, $\sigma_1^2=0{\cdot}9$ is also given for comparison.

be noticed that this function corresponds to $P(y)$ (*i.e.* the distribution function for the normalized structure amplitude for a single crystal with N atoms) when $\sigma_1^2=0$. Obviously, $P(y)=0$ for negative values of y. As σ_1^2 increases, the function $P(w)$ develops more and more on the negative side, and in the limit, when $\sigma_1^2=1$, it becomes a delta function, which is completely symmetric about $w=0$. This general behaviour is similar both for centrosymmetric and non-centrosymmetric crystals. However, there is a discontinuity in the derivative of the function $P(w)$ at the origin in the centrosymmetric case, although the function itself is continuous (Fig. 2).

A comparison of Figs. 1 and 2 shows that, for a particular value of σ_1^2, the curve is sharper for the non-centrosymmetric than for the centrosymmetric case. However, in both cases, the curves tend to be more similar to each other with increasing σ_1^2, and for this reason, it would not be profitable to use the function $P(w)$ as a statistical test for detecting a centre of symmetry in place of the function $P(y)$ itself.

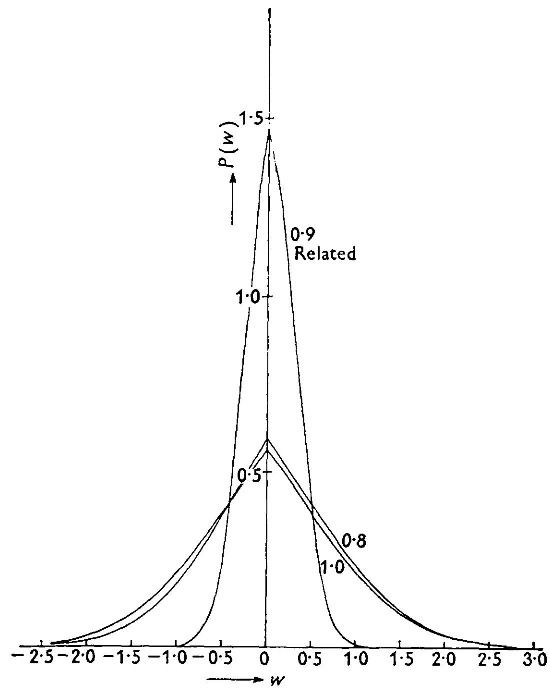

Fig. 4. Probability distribution function $P(w)$ for the unrelated centrosymmetric case. Only the curves for $\sigma_1^2=1{\cdot}0$ and $0{\cdot}8$ are shown. The corresponding curve for the related case $\sigma_1^2=0{\cdot}9$ is also given for comparison.

The curves for $\sigma_1^2=1$ in Figs. 3 and 4 for two unrelated crystals are particularly interesting. They correspond to the case of two crystal structures, both having the same number and types of atoms, but whose coordinates in the two structures are quite different from one another. It is possible to work out the R-values (reliability indices) for these cases from the curves for $P(w)$. It is readily seen that

$$R = \Sigma||F_N^{(1)}|-|F_N^{(2)}||/\Sigma|F_N^{(1)}| = \langle|\Delta|\rangle/\langle|F_N|\rangle \quad (22a)$$

$$= \langle|w|\rangle\sigma_N/\langle|F_N|\rangle = \langle|w|\rangle/\langle|y|\rangle \quad (22b)$$

where we have used $F_N^{(1)}$ and $F_N^{(2)}$ to represent the structure amplitudes of the two crystals (corresponding to F_N and F_P, used earlier). The quantities $\langle|w|\rangle$ and $\langle|y|\rangle$ can be readily obtained from the distribution functions $P(w)$ and $P(y)$, and this was done both for the centrosymmetric and non-centrosymmetric cases. These gave the values

$$R_{\text{cs}}^{\text{unrel}} = 0{\cdot}828,\ R_{\text{ncs}}^{\text{unrel}} = 0{\cdot}586\ ,$$

which agree perfectly with the R-values given by Wilson (1950), for a proposed structure which is completely wrong. On the other hand, we have been able to obtain the actual distribution function for the differences in structure amplitude in these cases.

Figs. 3 and 4 also contain for comparison the distribution function for the *related* case, for $\sigma_1^2 = 0{\cdot}9$. It will be seen that the curve is much sharper in the related case, compared with the unrelated case. The distribution function $P(w)$, therefore, provides a good criterion for testing the isomorphism of two structures. This is possible because, when P forms a part of N, it corresponds to the case of two isomorphous crystals with P and N atoms respectively.

So also, we can calculate the probable fraction of reflections (say P_+) for which $|F_N| > |F_P|$ by finding the integral

$$P_+ = \int_0^\infty P(w)dw\,. \quad (23)$$

The variation of P_+ with σ_1^2 is shown both for centrosymmetric as well as non-centrosymmetric crystals in Fig. 5. Obviously, in both cases, $P_+ = 0{\cdot}5$ for $\sigma_1^2 = 1$ and $P_+ = 1$ for $\sigma_1^2 = 0$. The value of P_+ for any particular value of σ_1^2 can be used for finding the relative scale factor of two isomorphous crystal, for only when the relative scaling is correct would the fraction of reflections for which $|F_N|$ is greater than $|F_P|$ be equal to the theoretical value. The applications of the results of this paper to isomorphous crystals will be considered in detail in another paper, along with examples.

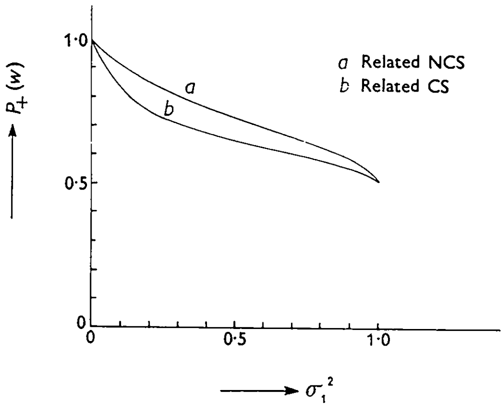

Fig. 5. Variation of P_+, the probable fraction of reflections for which $|F_N| > |F_P|$ with σ_1^2, for the related case. (*a*) Non-centrosymmetric, (*b*) centrosymmetric.

Finally it might be mentioned that, in addition to the case of isomorphous crystals, the results obtained above might also prove useful in studies where structural similarity is involved, as for instance in problems connected with phase transitions and order–disorder structures. We may sometimes meet with a family of compounds with a gradation in their similarities, *e.g.*, the feldspar compounds. Such cases will also be considered later.

References

Howells, E. R., Phillips, D. C. & Rogers, D. (1950). *Acta Cryst.* **3**, 210.

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

Pearson, K. (1922). *Tables of Incomplete Gamma Functions.*

Ramachandran, G. N. & Srinivasan, R. (1959). *Acta Cryst.* **12**, 410.

Sim, G. A. (1958). *Acta Cryst.* **11**, 123.

Srinivasan, R. (1960). *Acta Cryst.* **13**, 388.

Wilson, A. J. C. (1950). *Acta Cryst.* **3**, 397.