DISCREPANCY INDICES FOR USE IN CRYSTAL STRUCTURE ANALYSIS

S. PARTHASARATHY

(*Centre of Advanced Study in Physics,*
University of Madras, Guindy Campus, Madras 600 025, *India*)

Received on March 6, 1975

ABSTRACT

The use of discrepancy indices (i.e., *R-indices*) *in crystal structure analysis as an index for the correctness of trial structures in the structure completion stage and as an index indicating the extent of refinement stage are well known. A number of R-indices have been suggested by various workers for these purposes. The results obtained from a systematic comparative study of the various normalised R-indices for following three crystallographic situations, namely,* (i) *the structure completion process,* (ii) *refinement of an incomplete model and* (iii) *refinement of a complete model, are summarised after introducing the necessary nomenclature, notation and relevant crystallographic literature. The results obtained from a comparative study of the normalised and unnormalised Booth type indices based on intensity carried out under different crystallographic situations for assessing the advantage of the normalisation procedure during the structure completion stage are also summarised. The necessity of taking into account the space group symmetry of the crystal in the evaluation of R-indices is pointed out. General theoretical expressions for the normalised Booth type index based on intensity which are useful to test the correctness of any type of model for crystals of the triclinic monoclinic and orthorhombic systems are given. These are valid even for crystals with any number and types of atoms in the unit cell. The expression which accounts for random errors in the intensity data is also given. Other uses of R-indices are indicated.*

1. INTRODUCTION

The method of crystal structure analysis may be broadly described as a process in which a model (or trial structure) is first obtained by making use of the available intensity data (say *via* the Patterson function or the direct methods) and then completed if it were incomplete by Fourier methods [1] and finally refined by the least-squares method to an extent

---

* Contribution No. 000.

that will be consistent with (i) the amount and accuracy of the intensity data and (ii) the deficiency of the model. At any state during this process, it is useful to know whether the ϕ model is reasonably correct. The only sure way[+] to ascertain this is to compare and see whether the calculated structure factor magnitudes of the various reflections (*i.e.*, $|F_P{}^c|$'s) agree reasonably with their corresponding observed values (*i.e.*, $|F_0|$'s). Though such a comparison could be useful for specific reflections (for example to detect the extinction effects), this tedious procedure is not generally followed. In practice one often computes the ratio $R(F)$ of the sum of the discrepancies $||F_0| - |F_P{}^c||$ in the various reflections to the sum of their $|F_0|$ values and uses it as an index to indicate the extent of the discrepancy between the proposed model and the true structure. That is

$$R(F) = \sum_{hkl} ||F_0| - |F_P{}^c|| \Big/ \sum_{hkl} |F_0|. \tag{1}$$

Though (1) is the commonly used discrepancy index (*i.e.*, R-index) a number of other R-indices based on the discrepancy between $|F_0|$ and $|F_P{}^c|$ have also been suggested in the literature for this purpose (*see* Section 4). We shall refer to such R-indices as *unnormalised* R-indices. Srinivasan and Ramachandran [2] have pointed out that when the model is not complete, the discrepancy between $|F_0|$ and $|F_P{}^c|/\sigma_1$ (where $\sigma_1{}^2$ is the ratio of the mean square structure amplitude of the model to that of the true structure) would be more relevant than that between $|F_0|$ and $|F_P{}^c|$ for the calculation of R-indices. They have therefore considered a modified form of the conventional R-index using the discrepancy $||F_0| - |F_P{}^c|/\sigma_1|$. This idea has later been extended to other types of R-indices by different workers [3] and [4]. Following Srinivasan and Ramachandran [2] we shall refer to such R-indices as the normalised R-indices. Thus a variety of the R-indices, normalised and unnormalised types, are available in the literature for use in crystal structure analysis. It becomes therefore essential now to compare the relative efficiency of the various available R-indices under different crystallographic situations. This article is mostly confined to such a study.

A number of independent factors contribute to the value of any R-index. These are: (i) Random and systematic errors in the observed intensity data and (ii) Different types of deficiency in the model. The latter arises from a number of factors such as (*a*) imperfections due to

---

[+] Stereochemical criteria could also be used as additional check for the validity of a model particularly when a major part of the structure is known random errors in the atomic positions; (*b*) *incompleteness* due to non-inclusion of *all* the atoms of the unit cell of the crystal in the model; (*c*) errors due to non-inclusion of the bonding electrons, errors in the thermal parameters, etc. A theoretical treatment of any R-index taking all these factors into account becomes complicated and hence in the literature R-indices have often been evaluated only for certain special situations. In our comparative study also we shall make the following two assumptions: (i) The intensity data will be assumed to be known with perfect accuracy (*i.e.*, $|F_0| = |F_N|$), (ii) we shall consider only two types of deficiencies of the model which make significant contributions to the value of any R-index, namely, the imperfection and the incompleteness mentioned above in (*a*) and (*b*) respectively. In spite of these simplifying assumptions the physical deductions that are obtained regarding the relative efficiency of the various R-indices could be expected to broadly hold good in practice.

In Section 2 the nomenclature and notation needed for discussions in the later sections are explained. Since various workers have used different notations for the various R-indices it becomes now essential to adopt a uniform notation for avoiding confusion and in Section 3 the notation employed for denoting the various R-indices is explained. Section 4 summarises the various cases for which R-indices have been evaluated in the literature. In section 5 the results obtained from a comparative study of the various normalised R-indices under different crystallographic situations are summarised. Since this study shows that of the various R-indices the Booth type index based on intensity (see Section 3 for the definition of this index) is the best when the model is *incomplete*, it becomes necessary to compare the relative efficiency of the normalised and unnormalised Booth type indices during the structure completion stage. The results obtained from such a study are summarised in Section 6. Section 7 deals with general theoretical expressions for the Booth type indices needed to test for the correctness of any type of incomplete model of complex structures containing any number and species of atoms in the unit cell. These results are applicable to crystals of the triclinic, monoclinic and orthorhombic systems. In Section 8 other possible uses of R-indices are pointed out.

2. NOMENCLATURE AND NOTATION

Consider a non-centrosymmetric crystal containing N atoms in the unit cell. Suppose at any given stage during crystal structure analysis the proposed model contains $P\,(> N)$ atoms so that $N - P\,(= Q$, say) will be the number of atoms yet to be located (*i.e.*, unknown atoms). Let the contributions to the structure factor of a reflection H (= hkl) from the N-, P- and Q-atoms of the true structure be denoted by F_N, F_P and F_Q respectively and let $F_P{}^c$ be the calculated structure factor of that reflection. It is obvious that F_N is nothing but the true structure factor of the reflection $\boldsymbol{H}$ and that in the absence of errors in the observed data $|F_N| = |F_0|$. Let $\triangle\boldsymbol{r}_{Pj}$, $j = 1$ to P, represent the random positional errors suggested by the atoms in the model. In the theoretical treatment it is assumed that (i) the $\triangle\boldsymbol{r}_{Pj}$'s are mutually independent random vectors obeying the same Gaussian distribution and (ii) the $\triangle\boldsymbol{r}_{Pj}$'s are also independent of the position vectors $\boldsymbol{r}_{Pj}$'s of the atoms in the structure. Physically these assumptions are equivalent to the following: (i) the $\triangle\boldsymbol{r}_{Pj}$'s are the elements of a random sample of size P from a Gaussian population of a given σ and (ii) the error in the position of atom j is independent of the position of any of the atoms in the structure and the errors in the positions of the other atoms. From Luzzati's [5] work on the conventional R-index in the presence of random positional errors, it becomes obvious that the quantity $D_P(H)$ defined by

$$D_P(H) = \langle \cos 2\pi \boldsymbol{H} \cdot \triangle\boldsymbol{r} \rangle_P \tag{2}$$

is a useful parameter. Luzzati [5] has also shown that $D_P(H)$ (hereafter written as D for convenience) can be represented by the following Gaussian function with spherical symmetry

$$D = \exp(-CH^2 \langle |\triangle\boldsymbol{r}| \rangle_P) \tag{3}$$

where $C = \pi^3$, 4π and $\pi^3/4$ for the 1-, 2- and 3-dimensional cases respectively and $\langle |\triangle\boldsymbol{r}| \rangle_P$ is the mean of the magnitudes of the errors in the position vectors of the atoms of the model.

Types of Models

From the point of view of positional errors [*i.e.*, deficiency (*a*) described in section 1], models could be conveniently classified into four types, namely, the related, unrelated, imperfectly related and the semi-related types. Thus a model in which all the atoms are in correct locations is said to be of the *related* type. A model in which all the atoms are at completely wrong positions is said to be of the *unrelated* type. A model in which the atoms suffer random parameter errors (whcih could be minimised during a least-squares refinement) is said to be of the *imperfectly related* type. A model is said to be of the *semi-related* type if P_r atoms in it are in correct locations while the remaining $P - P_r$ (= P_w say) atoms are in completely wrong locations. From the point of view of incompleteness [*i.e.*, deficiency (*b*) described in section 1], models could be classified into two types, namely, the complete and incomplete types. A model is said to be of the *complete* type if it accounts for all the N atoms in the unit cell of the crystal (*i.e.*, $P = N$) and it is said to be of the *incomplete* type if it contains only P out of N atoms (*i.e.*, $P < N$). When the deficiencies (*a*) and (*b*) co-exist, the models are to be classified by taking these two deficiencies into account. Thus we can talk of an imperfectly related incomplete model, unrelated complete model, etc. The various types of models that the thus possible and their characteristics are summarised in Table I by taking the non-centrosymmetric case as example. It is useful to note the following points: (i) Models met with during the initial stages of crystal structure analysis are of the imperfectly related incomplete type. (ii) The related complete model is nothing but the true structure. (iii) The rated and the unrelated cases can be thought of as two limiting situations of the imperfectly related case (*i.e.*, for the related case $\langle | \triangle r | \rangle$ is zero and for the unrelated case $\langle | \triangle r | \rangle_P$ is large in value). (iv) The related and the unrelated cases can also be interpreted as two limiting situations of the semi-related case (*i.e.*, for the related case $P_r = P$ and $P_w = 0$ while for the unrelated case $P_r = 0$ and $P_w = P$).

A good measure of the random positional errors in the quantity $\langle | \triangle r | \rangle_P$ which is related to D as in (3). A good measure of incompleteness of the model is the quantity σ_1^2 defined by

$$\sigma_1^2 = \langle | F_P^c |^2 \rangle / \langle | F_N |^2 \rangle \tag{4}$$

which denotes the fractional contribution to the local mean intensity from the atoms in the model (relative to that of the true structure). When all the atoms in the unit cell are similar, σ_1^2 in (4) takes the simple form

$$\sigma_1^2 = P/N \tag{5}$$

which is the fractional number of input atoms in the model (relative to the true structure). When both deficiencies (*a*) and (*b*) coexist, Srinivasan and Ramachandran [6] have shown that the quantity of interest is σ_A defined by

$$\sigma_A = \sigma_1 D \tag{6}$$

Following Srinivasan and Ramachandran [6] we shall also define σ_B to be

$$\sigma_B = \sqrt{1 - \sigma_A^2} \tag{7}$$

It is convenient to define σ_2^2 by

$$\sigma_2^2 = \langle | F |^2 \rangle / \langle | F_N |^2 \rangle \tag{8}$$

TABLE I

Definition of different types of models relative to a given crystal structure

Crystal Structure: $(f_{Ni}, \mathbf{r}_{Ni}) = (f_{Pj}, \mathbf{r}_{Pj}) + (f_{Qk}, \mathbf{r}_{Qk})$, $i = 1$ to N, $j = 1$ to P, $k = 1$ to Q.

Model Structure: $(f_{Pj}, \mathbf{r}_{P_j}{}^c)$, $j = 1$ to P.

Mean coordinate error $\langle \lvert \Delta \mathbf{r} \rvert \rangle$	Number of atoms in the Model	
	P	N
Zero	Related in complete model	Related complete model (= true structure)
Finite but small	Imperfectly related incomplete model	Imperfectly related complete model
very large	Unrelated incomplete model	Unrelated complete model
Zero for some of the atoms and very large for the others	Semi-related incomplete model	Semi-related complete model

Note.—A group of P atoms of scattering power f_{P_j}'s at locations $\mathbf{r}_{P_j}$'s is denoted by $(f_{Pj}, \mathbf{r}_{Pj})$, $j = 1$ to P.

so that

$$\sigma_1{}^2 + \sigma_2{}^2 = 1 \tag{9}$$

The probable values of σ_1, D and σ_A for the various types of models are summarised in Table II.

In order to simplify our discussions, it is convenient to divide the process of crystal structure analysis into two stages, namely, the *structure completion stage* and the *conventional refinement stage*. The former is defined to be the stage during which only a part of the structure is known (*i.e.*, $\sigma_1{}^2 < 1$) while the latter is defined to be the stage during which all the atoms in the structure would be known but for the random parameter errors (we are not concerned here with the hydrogen atoms). For convenience we shall speak of the following two situations in the structure completion stage, namely, the *structure completion process* and the *refinement of an incomplete model*. The former refers to the process during which more and more unknown atoms are located by Fourier methods and added to an incomplete model. Thus σ_1^2 increases towards 1 during this process. In the latter situation, σ_1^2 remains fixed in value while $\langle |\Delta \boldsymbol{r}| \rangle_P$ decreases.

TABLE II

Probable values of σ_1, D and σ for the different types of models

No.	Type of Model	σ_1	D	σ_A
1.	Related incomplete Model	$0 < \sigma_1 < 1$	1	$0 < \sigma_A < 1$
2.	Imperfectly related incomplete Model	$0 < \sigma_1 < 1$	$0 < D < 1$	$0 < \sigma_A < 1$
3.	Unrelated incomplete Model	$0 < \sigma_1 < 1$	0	0
4.	Related complete Model	1	1	1
5.	Imperfectly related complete Model	1	$0 < D < 1$	$0 < \sigma_A < 1$
6.	Unrelated complete Model	1	0	0
7.	Semi-related incomplete Model	$0 < \sigma_1 < 1$	..	..
8.	Semi-related complete Model	1	..	..

Note.—$\sigma_A = \sigma_1 D$, $\sigma_2 = \sqrt{1 - \sigma_1^2}$, $\sigma_B = \sqrt{1 - \sigma_A^2}$. For the semi-related case D and σ_A are not defined.

It is obvious that the results in this section apply to the centrosymmetric case as well. However it is necessary to note that the positional errors for this case will be given by $\Delta \boldsymbol{r}_{Pj}$, $j = 1$ to $P/2$ where $P/2$ is the number of known atoms in the asymmetric unit.

In this paper we shall use the abbreviations C and NC to stand for the terms “centrosymmetric” and “non-centrosymmetric” respectively.

and the abbreviations *R*, *UR*, *IR* and *SR* for the terms "related", "unrelated", "imperfectly related" and "semi-related" respectively.

3. NOTATION FOR THE DESCREPANCY INDICES

The discrepancy of a model relative to the true structure could be conveniently measured in terms of the expectation values of the discrepancies between the quantities in any one of the following pairs: $|F_N|$, $|F_P^c|$; I_N, I_P^c (where $I_N = |F_N|^2$ and $I_P^c = |F_P^c|^2$); $|F_N|$, $|F_P^c|/\sigma_1$ and I_N, I_P^c/σ_1^2. Some simple types of discrepancies between the qunatities in these pairs are defined in Table III. Of these, those in rows (2), (3), (5) and (6) could be conveniently referred to as fractional discrepancies. A number of different types of *R*-indices could be defined using these or any of their powers (*e.g.*, squares). The *R*-indices that are of some interest

TABLE III

Some possible types of discrepancies between the structure factor magnitudes (or intensities) or the true structure and its model

No.	Discrepancy Based On	
	Structure Factor Magnitude	Intensity
1.	$\left\| \, \|F_N\| - \|F_P^c\| \, \right\|$	$\|I_N - I_P^c\|$
2.	$\left\| \dfrac{\|F_N\| - \|F_P^c\|}{\|F_N\| \text{ or } \|F_P^c\|} \right\|$	$\left\| \dfrac{I_N - I_P^c}{I_N \text{ or } I_P^c} \right\|$
3.	$\left\| \dfrac{\|F_N\| - \|F_P^c\|}{\frac{1}{2}(\|F_N\| + \|F_P^c\|)} \right\|$	$\left\| \dfrac{I_N - I_P^c}{\frac{1}{2}(I_N + I_P^c)} \right\|$
4.	$\left\| \, \|F_N\| - \|F_P^c\|/\sigma_1 \right\|$	$\|I_N - I_P^c/\sigma_1^2\|$
5.	$\left\| \dfrac{\|F_N\| - \|F_P^c\|/\sigma_1}{\|F_N\| \text{ or } \|F_P^c\|/\sigma_1} \right\|$	$\left\| \dfrac{I_N - I_P^c/\sigma_1^2}{I_N \text{ or } I_P^c/\sigma_1^2} \right\|$
6.	$\left\| \dfrac{\|F_N\| - \|F_P^c\|/\sigma_1}{\frac{1}{2}(\|F_N\| + \|F_P^c\|/\sigma_1)} \right\|$	$\left\| \dfrac{I_N - I_P^c/\sigma_1^2}{\frac{1}{2}(I_N + I_P^c/\sigma_1^2)} \right\|$

Note.—The (or in the denominator of (2) and (5) denotes that whichever is greater is to be used in the evaluation of the fractional discrepancy. Also $\sigma_1^2 = \langle |F_P^c|^2 \rangle / \langle |F_N|^2 \rangle$.

are defined in Table IV along with their symbols. The normalised R-indices are denoted by using '1' as post-subscript to the R-index symbol (*e.g.*, R_1). The fractional type index based on either of the discrepancies defined in rows (2) or (5) of Table III is denoted by using an asterisk mark as a post-superscript (*e.g.*, R^*). The fractional type index based on either of the discrepancies defined in rows (3) or (6) of Table III are denoted by using an 'f' as a post-superscript for the R-index symbol (*e.g.*, R^f). The indices based on the square of any of the discrepancies is called the Booth type index and are denoted by using the letter 'B' as a presubscript to the R-index symbol (*e.g.*, ${}_BR$). R-index could be defined with either the structure amplitude $|F|$ or the intensity I and this is indicated by using F or I in parantheses (*e.g.*, $R(F)$ or $R(I)$). The conventional R-index defined in (1) is thus to be denoted by $R(F)$.

The following points may be noted: (*i*) Though R-indices for models met with in crystal structure analysis are in general monotonically increasing function of $\sin\theta/\lambda$, these are usually computed by taking all the available reflections as a single group. Such an overall value for any R-index is denoted by $(R)_H$ in this article. (ii) In the case of any model of a given crystal, the overall value of any R-index could be calculated in two ways namely, (*a*) by using the definition in column (2) of Table IV which involves the use of the actual structure factor magnitudes or intensities as the case may be or (*b*) by using the results in the last column of Table IV which involves the use of the normalised structure factor magnitudes (*i.e.*, y's) or the normalised structure factor intensities (*i.e.*, z's) as the case may be. For any given fractional type R-index the two ways would yield the same result and hence need not be distinguished. However, for the others, overall values computed by the two ways† would in general differ [see point (iv) for an exception] and hence it becomes necessary to incorporate this aspect also in the notation for R-indices. R-indices computed using y-values will be denoted by using y in parantheses and those computed using z-values by using z in parantheses. Thus while $R_1(F) = \Sigma\,|\,|F_N| - |F_P{}^c|/\sigma_1\,|/\Sigma|F_N|$, $R_1(y) = \Sigma\,|y_N - y_P{}^c|/\Sigma\,y_N$. (iii) The value of any R-index computed by either of the two ways by using reflections in a sufficiently narrow region of $\sin\theta/\lambda$ will be nearly equal. (iv) Though for an equal atom structure, the value of any R-index for the related case would be independent of $\sin\theta/\lambda$, this is not so for crystals with dissimilar atoms. Hence for practical applications of the theoretical results this aspect must be taken into account in

---

† In the case of any non-fractional type R-index, the former method gives greater weight to reflections with smaller value of sin 4|7 and hence would yield smaller value for the R-index (*see* [13]).

TABLE IV

Definition of the discrepancy indices of interest, their notation and their relation to the normalised variables

No.	Notation	Definition	Relation to the Normalised Variables
1.	$R\,(F)$	$\dfrac{\Sigma \,\|\,\|F_N\| - \|F_P^c\|\,\|}{\Sigma \,\|F_N\|}$	$\dfrac{\langle \,\|\,y_N - \sigma_1 y_P^c\,\|\,\rangle}{\langle y_N \rangle}$
2.	$R_1\,(F)$	$\dfrac{\Sigma \,\|\,\|F_N\| - \|F_P^c\|/\|\sigma_1\|}{\Sigma \,\|F_N\|}$	$\dfrac{\langle \,\|\,y_N - y_P^c\,\|\,\rangle}{\langle y_N \rangle}$
3.	$B^{R\,(F)}$	$\dfrac{\Sigma\,(\,\|F_N\| - \|F_P^c\|\,)^2}{\Sigma\,\|F_N\|^2}$	$\langle\,(\,y_N - \sigma_1 y_P^c\,)^2\,\rangle$
4.	$B^{R_1\,(F)}$	$\dfrac{\Sigma\,(\,\|F_N\| - \|F_P^c\|/\sigma_1\,)^2}{\Sigma\,\|F_N\|^2}$	$\langle\,(\,y_N - y_P^c\,)^2\,\rangle$
5.	$R^{*\,(F)}$	$\dfrac{1}{n}\sum \left\| \dfrac{\|F_N\| - \|F_P^c\|}{\|F_N\| \text{ or } \|F_P^c\|} \right\|$	$\left\langle \left\| \dfrac{y_N - \sigma_1 y_P^c}{y_N \text{ or } \sigma_1 y_P^c} \right\| \right\rangle$
6.	$R_1{}^*\,(F)$	$\dfrac{1}{n}\sum \left\| \dfrac{\|F_N\| - \|F_P^c\|/\sigma_1}{\|F_N\| \text{ or } \|F_P^c\|/\sigma_1} \right\|$	$\left\langle \left\| \dfrac{y_N - y_P^c}{y_N \text{ or } y_P^c} \right\| \right\rangle$
7.	$R^t{}_1\,(F)$	$\dfrac{1}{n}\sum \left\| \dfrac{\|F_N\| - \|F_P^c\|/\sigma_1}{\frac{1}{2}(\,\|F_N\| + \|F_P^c\|/\sigma_1\,)} \right\|$	$\left\langle \left\| \dfrac{y_N - y_P^c}{\frac{1}{2}(\,y_N + y_P^c)} \right\| \right\rangle$
8.	$R\,(I)$	$\dfrac{\Sigma\,\|\,I_N - I_P^c\,\|}{\Sigma\, I_N}$	$\langle\,\|\,z_N - \sigma_1{}^2 z_P^c\,\|\,\rangle$
9.	$R_1\,(I)$	$\dfrac{\Sigma\,\|\,I_N - I_P^c/\sigma_1{}^2\,\|}{\Sigma\, I_N}$	$\langle\,\|\,z_N - z_P^c\,\|\,\rangle$
10.	$B^{R\,(I)}$	$\dfrac{\Sigma\,(\,I_N - I_P^c)^2}{\Sigma\, I_N{}^2}$	$\dfrac{\langle\,(\,z_N - \sigma_1{}^2 z_P^c\,)^2\,\rangle}{\langle\, z_N{}^2\,\rangle}$
11.	$B^{R_1\,(I)}$	$\dfrac{\Sigma\,(I_N - I_P^c/\sigma_1{}^2\,)^2}{\Sigma\, I_N{}^2}$	$\dfrac{\langle\,(\,z_N - z_P^c\,)^2\,\rangle}{\langle\, z_N{}^2\,\rangle}$
12.	$R^*\,(I)$	$\dfrac{1}{n}\sum \left\| \dfrac{I_N - I_P^c}{I_N \text{ or } I_P^c} \right\|$	$\left\langle \left\| \dfrac{z_N - \sigma_1{}^2 z_P^c}{z_N \text{ or } \sigma_1{}^2 z_P^c} \right\| \right\rangle$
13.	$R_1{}^*\,(I)$	$\dfrac{1}{n}\sum \left\| \dfrac{I_N - I_P^c}{I_N \text{ or } I_P^c} \right\|$	$\left\langle \left\| \dfrac{z_N - z_P^c}{z_N \text{ or } z_P^c} \right\| \right\rangle$
14.	$R^t{}_1\,(I)$	$\dfrac{1}{n}\sum \left\| \dfrac{I_N - I_P^c/\sigma_1{}^2}{\frac{1}{2}(\,I_N + I_P^c/\sigma_1{}^2\,)} \right\|$	$\left\langle \left\| \dfrac{z_N - z_P^c}{\frac{1}{2}(\,z_N + z_P^c\,)} \right\| \right\rangle$

Note.—Here Σ is the summation over the n available reflections.

The notation in column (2) for the R-indices is due to Professor R. Srinivasan. In case of any given fractional type R-index, the expressions in columns 3 and 4 in the relevant row are equivalent independent of the value of $\sin\theta/\lambda$ However for an R-index of other types such an equivalence is valid only when the reflections used belong to a sufficiently narrow region of $\sin\theta/\lambda$ within which the atomic scattering factors can be treated as constants.

the evaluation of $\langle R \rangle_H$. Similar statement is true for the unrelated case also. (v) In the conventional refinement stage $\sigma_1^2 \simeq 1$ and hence the normalised *R*-index of any given type becomes identical to the corresponding type of unnormalised R-index.

4. THE CASES FOR WHICH R-INDICES HAVE BEEN EVALUATED IN THE LITERATURE

It is relevant here to summarise the cases for which *R*-indices have been theoretically evaluated in the literature. The types of crystals for which such evaluations have been carried out can be conveniently grouped into two categories, namely, (i) those which satisfy the requirements of the basic Wilson distributions and (ii) those which do not. Crystals with (a few) heavy atoms, with pseudosymmetry in the structure, with dissimilar atoms, etc., come under category (ii) while those with a sufficiently large number of similar atoms (at general positions) in the asymmetric unit come under category (i).

Studies pertaining to crystals containing a large number of similar atoms:

The conventional index $R(F)$ is the one that is commonly used in crystal structure analysis. Wilson [7] has worked out the values of $R(F)$ for the *C* and the *NC* cases when the model is of the unrelated complete type. Lizzati [5] has obtained its relationship to the parameter *D* for an imperfectly related complete model and used his results to study the variation of $R(F)$ as a function of sin θ corresponding to different specific values of $\langle |\triangle r| \rangle$. Srinivasan *et al* [8] have evaluated it as a function of σ_1^2 corresponding to two limiting situations (*i.e.*, the *R* and *UR* cases). The effect of random errors in thermal parameters on $\langle R(F) \rangle_H$ has been considered by Stanley [9] for a complete model. He also evaluated $\langle R(F) \rangle_H$ as a function of the standard deviation $\sigma(r)$ in the positions of the atoms for different values of the temperature factor *B* (*see* [10] to [12]). Srinivasan and Ramachandran [6] have evaluated $R_1(y)$ as a function of σA. Chandrasekharan and Srinivasan [13] have obtained curves of $\langle R_1(y) \rangle_H$ *vs* $\langle |\triangle r| \rangle$ for different values of σ_1^2 and these are applicable for 2-dimensional data. Theoretical expressions for $R_1(z)$ for the *C* and *NC* cases have been derived by Parthasarathy and Srinivasan 3 and these are valid for imperfectly related incomplete models. Booth [14] has worked out, using certain reasonable assumptions, the relationship of $\langle B^{R(F)} \rangle_H$ to the r.m.s. error in the coordinates for a complete model corresponding to an *NC* crystal. Parthasarathy and Parthasarathi [15] have obtained theoretical expressions for $B^{R(y)}$ and $B^{R(z)}$ for the *SR* case from which they have deduced the results for the *R* and *UR* cases.

They have also obtained expressions for the latter index for the C and NC cases which take into account Gaussian errors in the observed intensities and those in atomic positions. The effect of a badly misplaced atom in an otherwise correct complete model on the indices $R(F)$, $R(I)$ and $B^{R(I)}$ has been discussed by Wilson [16]. He has also considered the overall effect of different types of errors such as errors in the intensity data, random errors in positional and thermal parameters, etc. Lenstra [17] has considered the effect of a badly misplaced atom in an otherwise correct incomplete model on $B^{R(I)}$. Theoretical expressions for the index $\Sigma(I_N - I_N^c)^2/\Sigma I_N^{c2}$ (which can be taken to be a modified form of the index $B^{R(I)}$) for the C and NC cases have been derived by Luzzati [18] and these are valid for models of the imperfectly related complete type. Srinivasan and Srikrishnan [19] have evaluated the indices $R^*(F)$ and $R^*(I)$ for a model which is of the unrelated complete type. The curves of $\langle R_1^*(F)\rangle_H$ *vs* $\langle|\Delta r|\rangle$ and $\langle R_1^*(I)\rangle_H$ *vs* $\langle|\Delta r|\rangle$ for different specific values of σ_1^2 have been obtained by Srikrishnan and Srinivasan [4] and these are valid for the 2-dimensional case. They have also later obtained the corresponding results for the 3-dimensional case and the results are not accurate due to the inappropriate weighting used in the calculation [20]. Recently, Parthasarathy and Parthasarathi [21] (*see* also [22]) have made a comparative study of the various normalised R-indices under different crystallographic situations. In that paper they have also obtained explicit expressions for a number of normalised R-indices for which such results were not available in the literature. Their results for the overall values of the R-indices are for the 3-dimensional case and are applicable for indices calculated using the normalised variables (*i.e.*, y or z).

Studies pertaining to crystals with pseudosymmetry heavy atoms or dissimilar atoms

Owing to theoretical difficulties R-indices for such crystals have been evaluated only for the R and UR cases. Doughlas and Woolfson [23] have derived the maximum probable value of $R(F)$ for a hypercentric structure. The maximum probable values of $B^{R(F)}$ and $B^{R(I)}$ for structures with different types of pseudosymmetry of the Rogers-Wilson types (*see* [24] and [25]) have been worked out by Wilson [26] and Parthasarathi and Parthasarathy [27]. Theoretical expressions for the index $\Sigma(I_N - I_N^c)^2/\Sigma I_N^{c^2}$ and the index $B^{R(y)}$ for an NC crystal with a degree of centrosymmetry have been derived by Luzzati [18] and Swaminathan and Srinivasan [28] respectively. The values of $R(y)$, $R_1(y)$, $R_1^*(F)$ and $R_1^*(I)$ (for the R and UR cases) for triclinic crystals containing two heavy atoms in the unit cell when the heavy atom part constitutes the model are also available in the literature (*see* [29] and [30]). The expressions for $B^{R(y)}$ and $B^{R(z)}$ for triclinic crystals with two or many heavy atoms when these atoms are taken to constitute the model have been obtained for the R and UR cases by Parthasarathy and Parthasarathi [15]. Wilson [26] has obtained expressions for the largest likely values of the indices $B^{R(F)}$ and $B^{R(I)}$ and his results are valid even for cases where the number of atoms in the unit cell are small or differ in scattering power. He has also discussed the problem of scaling with reference to the index $B^{R(I)}$. Parthasarathi and Parthasarathy ([31] and [32]) have obtained theoretical expressions for $B^{R_1(z)}$ and $B^{R(z)}$ for triclinic, monoclinic and orthorhombic crystals containing one or two heavy atoms in the asymmetric unit when the heavy atom part constitutes the model. They have also carried out a comparative study of the normalised and unnormalised Booth type indices under different crystallographic situations [32].

Some recent studies

Of the various R-indices, the Booth type index based on intensity is the easiest to manipulate theoretically [16] and hence theoretical expressions for this index for different cases of more general scope have been recently derived by different workers. Thus, Parthasarathy [33] has obtained general theoretical expressions for $B^{R_1(z)}$ and $B^{R(z)}$ applicable to any complex crystals of the triclinic, monoclinic and orthorhombic systems when any part of the structure constitutes the model. Wilson [34] has obtained expresions for $B^{R(I)}$ for the SR case and used his results to deduce results for other special cases. These are valid for crystals belonging to space groups of higher symmetry. He has given expressions which account for anomalous scattering. Lenstra [35] has discussed the effect of space group symmetry elements and pseudosymmetry in structure on $B^{R(I)}$ using an interesting approach based on the properties of the Patterson function.

Thus a variety of R-indices have been evaluated in the literature for different types of models and crystals and it becomes now essential to compare the relative efficiency of the different R-indices under different crystallographic situations. The results of such a study on normalised R-indices will be summarised in the next section.

5. COMPARATIVE STUDY OF THE VARIOUS NORMALISED R-INDICES

Evaluation of the overall indices $\langle R \rangle_H$

The theoretical expressions for the various normalised R-indices obtained by different workers [2], [3], [21] and [22] for the general case of an imperfectly related incomplete model when both the crystal and model satisfy the requirements of a Wilson distribution are summarised in Table V. It is seen from this table that each normalised R-index is a function of σA and hence depend on the quantities σ_1^2, $\langle |\triangle \mathbf{r}|\rangle$ and H $(= 2 \sin \theta/\lambda)$ implicitly. For a comparative study as well as for the practical application it would be useful to obtain the overall values of the various R-indices by taking all the reflections as a single group. If $H_{\max}$ denotes the maximum magnitude of the reciprocal lattice vector for the reflections included in the calculation of the R-index, then the overall value of the R-index denoted by $\langle R\rangle_H$ can be shown to be [21]

$$\langle R\rangle_H = \frac{3}{H^3_{\max}} \int_0^{H_{\max}} R\,(\sigma_1^2, \langle |\triangle r|\rangle, H)\, H^2\, dH. \tag{10}$$

Equation (10) is valid for the 3-dimensional case and that too for indices calculated using the normalised variables (y or z as the case may be). The value of $\langle R\rangle_H$ for any given values of σ_1^2 and $\langle |\triangle r|\rangle$ is to be obtained by substituting the expression for the relevant R-index from Table V in (10) and carrying out the resulting integration. For crystals containing similar atoms σ_1^2 $(= P/N)$ will be practically independent of $(\sin \theta/\lambda)$ and hence can be treated as a constant as far as the integration over H is concerned. In all cases the integral in (10) (index $B^{R_1(z)}$ is an exception and for the explicit expressions for $\langle B^{R_1(z)}\rangle_H$ for the C and NC cases *see* [21]) is to be evaluated numerically for given values of σ_1^2 and $\langle |\triangle \mathbf{r}|\rangle$. The variation of $\langle R\rangle_H$ as a function of σ_1^2 for a given $\langle |\triangle \mathbf{r}|\rangle$ or that of $\langle R\rangle_H$ as a function of $\langle |\triangle \mathbf{r}|\rangle$ for a given σ_1^2 could therefore be obtained. Parthasarathy and Parthasarathi [21] (*see* also [22]) have obtained such functional relationships in the form of suitable curves and used them for carrying out the comparative study of the various normalised R-indices under three different crystallographic situations. The situations considered are (i) structure completion process, (ii) refinement of an incomplete model and (iii) refinement of a complete model.

Structure completion process.—During this process, since the atoms in the model suffer random positional errors such that $\langle |\triangle \mathbf{r}|\rangle \simeq 0{\cdot}1$Å, the relevant curves for the present situation are those of R_H *vs* σ_1^2 for $\langle |\triangle \mathbf{r}|\rangle = 0{\cdot}1$Å. These are shown in Fig. (1 *a*) for the C case and Fig. (1 *b*) for the NC case. These figures show that $\langle B^{R_1(Z)}\rangle_H$ has a systematic and practically linear fall with increasing value of σ_1^2. Further it has the largest value for the slope for any given value of σ_1^2. Thus of the various

TABLE V

Theoretical expressions for the various normalised R-indices as a function of σ_A for the case of an imperfectly related incomplete model when the crystal and model satisfy the requirements of a Wilson distribution

Index	Centrosymmetric Case	Non-Centrosymmetric Case
$R_1(y)$	$\sqrt{2(1+\sigma_A)} + \sqrt{2(1-\sigma_A)} - 2$	$\frac{3}{2}\sigma_B^3 \int_0^1 \frac{{}_2F_1(-\frac{1}{4}, -\frac{3}{4}; \sigma_A^2 x^2)\,dx}{\sqrt{1+x}\,(1-\sigma_A^2 x^2)^2}$
$B^{R_1(y)}$	$2 - \frac{4}{\pi}[\sigma_B + \sigma_A \sin^{-1}(\sigma_A)]$	$2 - 2\left[E(\sigma_A) - \frac{\sigma B^2}{2} K(\sigma_A)\right]$
$R_1^*(y)$	$1 - \frac{2}{\pi}[\sigma_B \log_e(2\sigma_B) + \sigma_A \sin^{-1}(\sigma_A)$	$1 - 4\sigma_B^2 \int_0^1 \frac{x^2(1+x^2)\,dx}{[(1+x^2)^2 - 4\sigma_A^2 x^2]^{3/2}}$
$R_1^f(y)$	$\frac{2}{\pi}\left[\sqrt{\frac{1+\sigma_A}{1-\sigma_A}}\log_e\left(\frac{2}{1+\sigma_A}\right) + \sqrt{\frac{1-\sigma_A}{1+\sigma_A}}\log_e\left(\frac{2}{1-\sigma_A}\right)\right]$	$2 - \frac{2}{\sigma B}\log_e(1+\sigma_B)$
$R_1(z)$	$\frac{4}{\pi}\sigma_B$	σ_B
$B^{R_1(z)}$	$\frac{4}{3}\sigma_B^2$	σ_B^2
$R_1^*(z)$	$2\sigma_B^2 - \frac{4\sigma_B}{\pi}\left[1 + \frac{\sigma_A}{2}\log_e\left(\frac{1-\sigma_A}{1+\sigma_A}\right)\right]$	$2\sigma_B\left[(1 - \sigma_B \log_e\left(\frac{1+\sigma_B}{\sigma_B}\right)\right]$
	$\frac{1-{}_A}{1+{}_A}$	
$R_1^f(z)$	$\frac{2\sigma_B}{\pi\sigma_A}\log_e\left(\frac{1+\sigma_A}{1-\sigma_A}\right)$	$\frac{2\sigma_B}{1+\sigma_B}$

Note.—Here ${}_2F_1(a, b; c; x)$ is the hypergeometric function with two numerator and one denominator parameters. $K(k)$ and $E(k)$ are complete elliptic integrals of the first and second kind respectively. Though $R_1^*(F) \equiv R_1^*(y)$, we have used the latter symbol for uniformity. Similar statement holds good for other fractional type indices.

normalised R-indices, the index $B^{R_1(z)}$ seems to be the best for use during the structure completion process. Since $\langle B^{R_1(y)}\rangle_H$ also shows a practically linear fall with increasing value of σ_1^2, $B^{R_1(y)}$ may also be preferred next only to $B^{R_1(z)}$. It is relevant to not here that R_1 (y) which is the normalised form of the conventional R-index is inferior to the index $B^{R_1(y)}$. Further, the fractional type indices are seen to be least suited for this situation.

Refinement of an incomplete model.—A model which accounts for 50% of the atoms of the true structure is a typical example of an incomplete model. Hence Parthasarathy and Parthasarathi [21] obtained graphs of $\langle R\rangle_H$ *vs* $\langle|\triangle \boldsymbol{r}|\rangle$ for $\sigma_1^2 = 0 \cdot 5$ and these are shown in Fig. 2 *a* for the *C* case and Fig. 2 *b* for the *NC* case. It is seen from these figures that for this situation also the index $B^{R_1(z)}$ is the best. It is interesting to see that all the R-indices become practically insenstitive in the region $\langle|\triangle \boldsymbol{r}|\rangle \lessapprox 0 \cdot 05$ Å showing

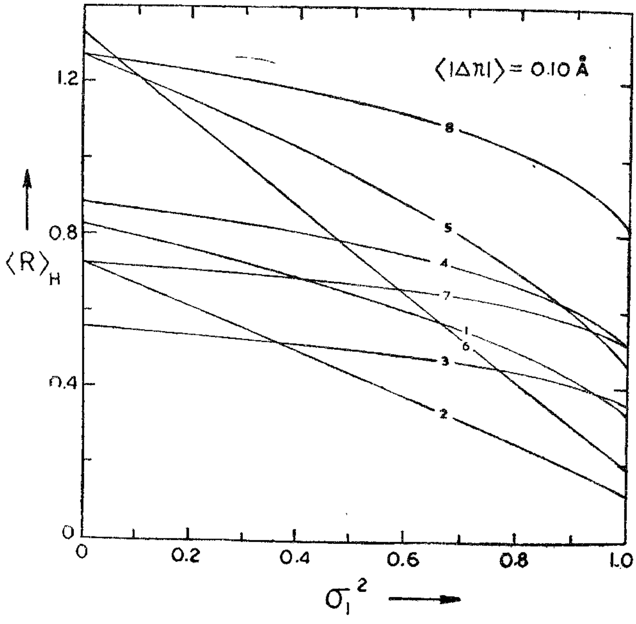

FIG. 1 *a*

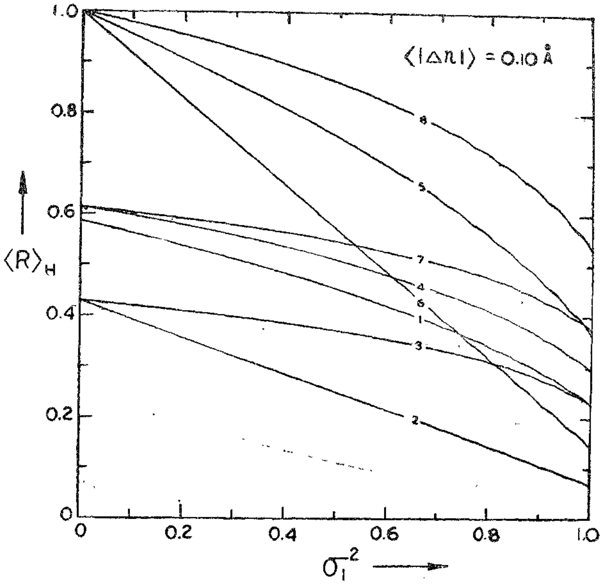

FIG. 1 *b*

FIG. 1. Representation of the overall values of the various normalised *R*-indices as a function of σ_1^2 when $\langle |\Delta \boldsymbol{r}| \rangle = 0 \cdot 1$ Å. The curves in (*a*) are for the *C* case and those in (*b*) are for the *NC* case. The numbers 1 to 8 against the various curves are used to denote the *R*-indices such that

$1 \rightarrow R_1(y)$, $2 \rightarrow B^{R_1(y)}$, $3 \rightarrow R_1^*(y)$, $4 \rightarrow R_1'(y)$
$5 \rightarrow R_1(z)$, $6 \rightarrow B^{R_1(z)}$, $7 \rightarrow R_1^*(z)$ and $8 \rightarrow R_1'(z)$.

thereby that refinement of an incomplete model to a high order could not be judged by the use of *R*-indices. In this connection it is relevant to note that the study on the probability distribution of the phase angle error has also shown that the refinement of an incomplete model would not be rapid [36].

Refirement of a complete model.—During this stage $\sigma_1^2 \simeq 1 \cdot 0$ and $\langle |\Delta \boldsymbol{r}| \rangle$ tends to zero. Thus the curves of $\langle R \rangle_H$ as a function of $\langle |\Delta \boldsymbol{r}| \rangle$ for $\sigma_1^2 = 1 \cdot 0$ are the relevant ones for this situation and these are shown in Fig. 3 *a* for the *C* case and Fig. 3 *b* for the *NC* case. From a study of these figures and also from the calculation of the actual slopes of the curves at different points of the $\langle|\Delta \boldsymbol{r}|\rangle$-axis the following conclusions have been obtained [22]: (*a*) In the final refinement stage (say $\langle|\Delta \boldsymbol{r}|\rangle \lesssim 0\cdot05$ Å) the fractional type indices based on intensity (particularly $R_1{}^f(z)$) are preferable to the others. (*b*) When $\langle|\Delta \boldsymbol{r}|\rangle$ lies in the interval $0\cdot05 \lesssim \langle|\Delta \boldsymbol{r}|\rangle \lesssim 0\cdot13$ Å (this could be referred to as the initial stages of refinement), the index $R_1(z)$ seems to be the best for the *C* case. However, for the *NC* case,

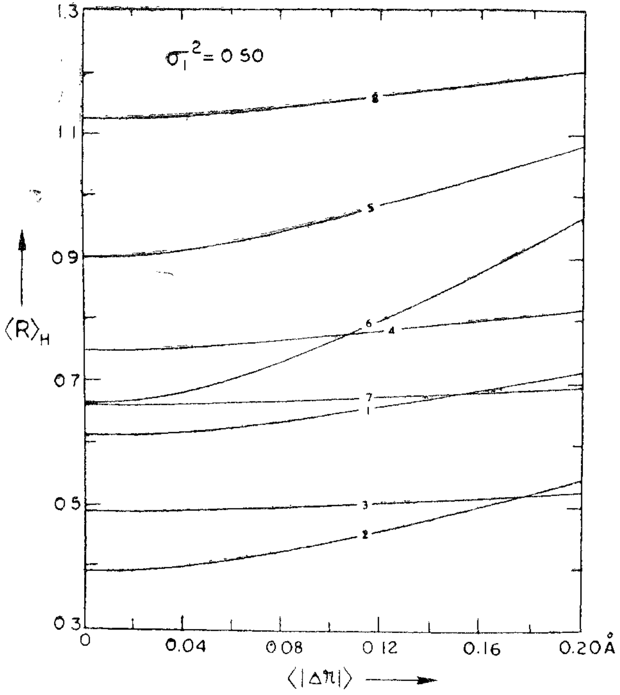

FIG. 2 *a*

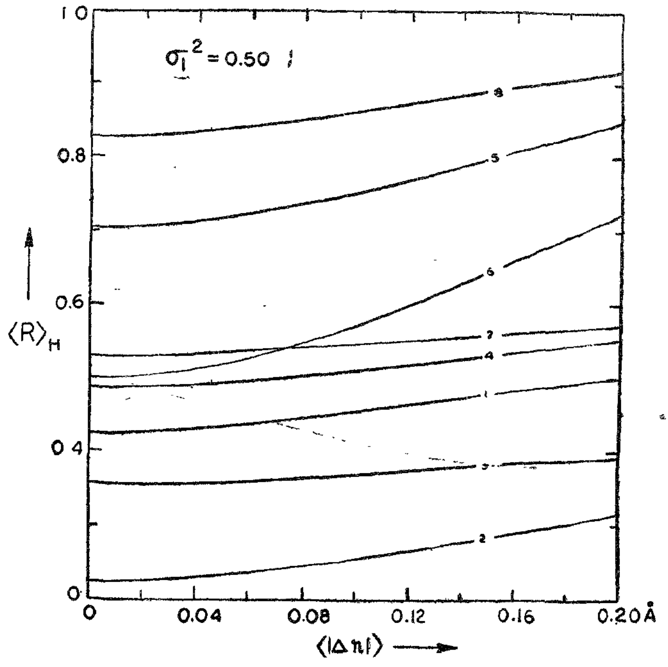

FIG. 2 *b*

FIG. 2. Representation of the overall values of the various normalised *R*-indices as a function of $\langle |\Delta r| \rangle$ for an imperfectly related incomplete model with $\sigma_1^2 = 0 \cdot 5$. The curves in (*a*) are for the *C* case and those in (*b*) are for the *NC* case. The numbers 1 to 8 against the various curves denote the various *R*-indices and for the details see the caption to Fig. 1.

a choice between $R_1 (z)$ and $R_1^f (z)$ seems to be difficult to make from purely theoretical considerations. Practical application in actual crystal structure refinements is needed to arrive at a unique decision.

From the above discussions it becomes clear that of the various normalised *R*-indices, the Booth type index $B^{R_1 (z)}$ is the best for use during the structure completion stage. During this stage since $\sigma_1^2 < 1$, the normalised and unnormalised indices of a given type differ. Thus it is useful to make a comparative study of the relative efficiency of the normalised and unnormalised Booth type indices $B^{R_1 (z)}$ and $B^{R (z)}$ during this stage in order to see whether the normalisation concept is a useful one. Such a comparative study has been carried by Parthasarathi and Parthasarathy [32] corresponding to the following cases: (i) the case when the crystal and the model satisfy the requirements of a Wilson distribution and (ii) the cases of triclinic, monoclinic and orthorhombic crystals containing p (= 1 or 2) heavy atoms in the asymmetric unit, the heavy atom part being the model. The results obtained for these cases will be summarised in section 6.

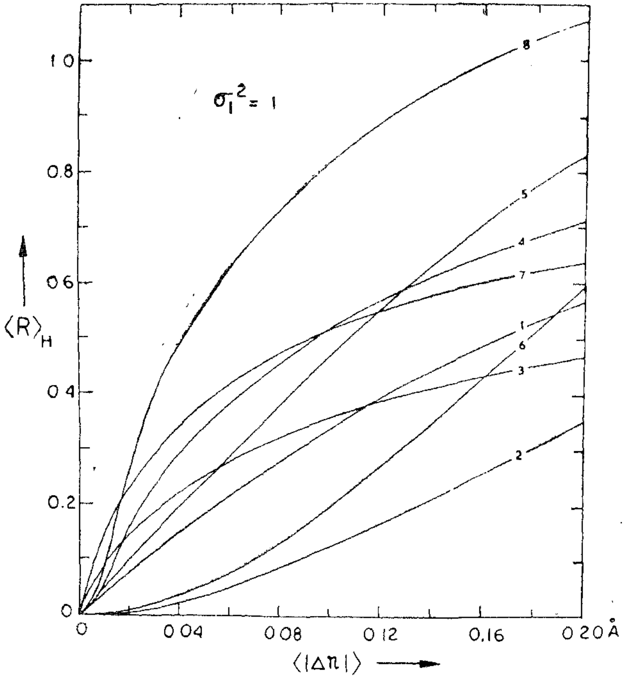

FIG. 3 *a*

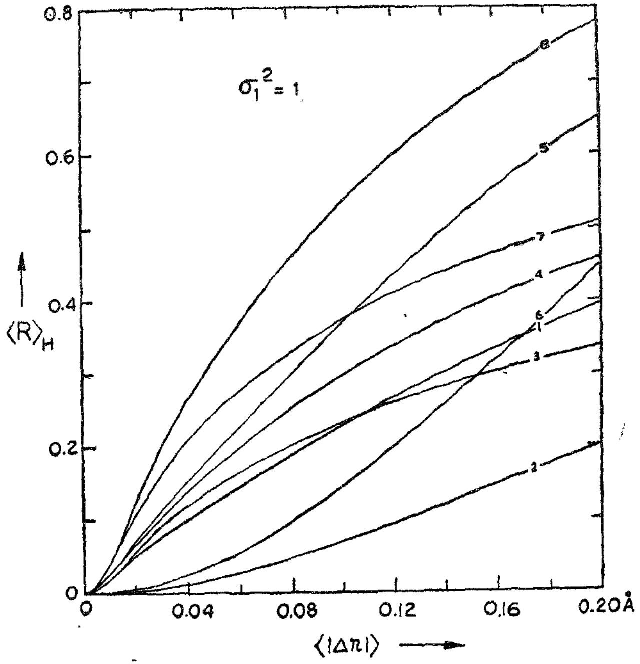

FIG. 3 *b*

FIG. 3. Representation of the overall values of the various normalised *R*-indices as a function of $\langle | \Delta r | \rangle$ for an imperfectly related complete model (*i.e.*, $\sigma_1^2 = 1$). The curves in (*a*) are for the *C* case and those in (*b*) are for the *NC* case. The numbers 1 to 8 against the various curves correspond to the different *R*-indices (for details see caption to Fig. 1).

6. COMPARATIVE STUDY ON THE RELATIVE EFFICIENCY OF THE INDICES $B^{R_1(z)}$ AND $B^{R(z)}$

The case when the crystal and model satisfy the requirements of a Wilson distribution

Structure Completion process.—The variation of $B^{R_1(z)}$ and $B^{R(z)}$ for the *R* and *UR* cases are shown in Fig. 4 *a* for the *C* case and Fig. 4

for the *NC* case. It is seen that the distruction between the *R* and *UR* cases is more marked for the normalised index than for the unnormalised index. Hence $B^{R_1(Z)}$ is preferable to $B^{R(Z)}$ for this situation.

Refinement of an incomplete model.—The curves of $\langle B^{R_1(Z)}\rangle_H$ and $\langle B^{R(Z)}\rangle_H$ as a function of $\langle|\Delta \mathbf{r}|\rangle$ for three different values of $\sigma_1{}^2$, namely, $\sigma_1{}^2 = 0\cdot3$, $0\cdot5$ and $0\cdot7$ are shown in Fig. 5 *a* for the *C* case and Fig. 5 *b* for the *NC* case. These figures clearly show that the normalised index is preferable to the unnormalised index for this situation as well.

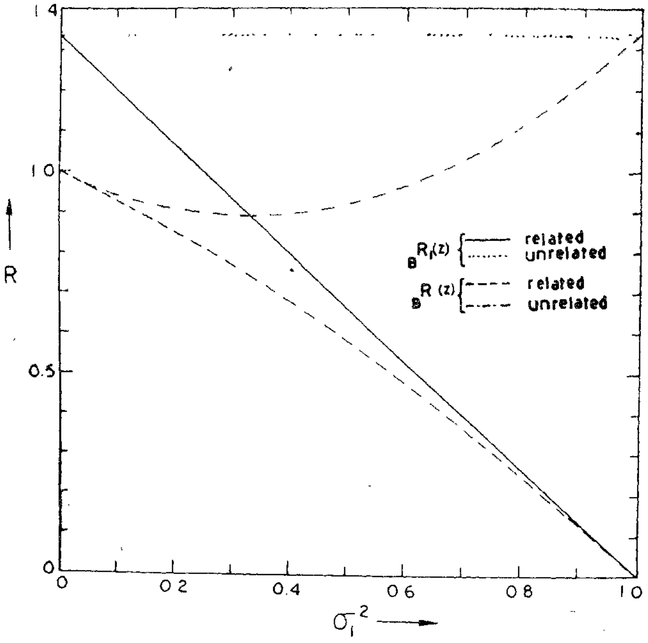

FIG. 4 *a*

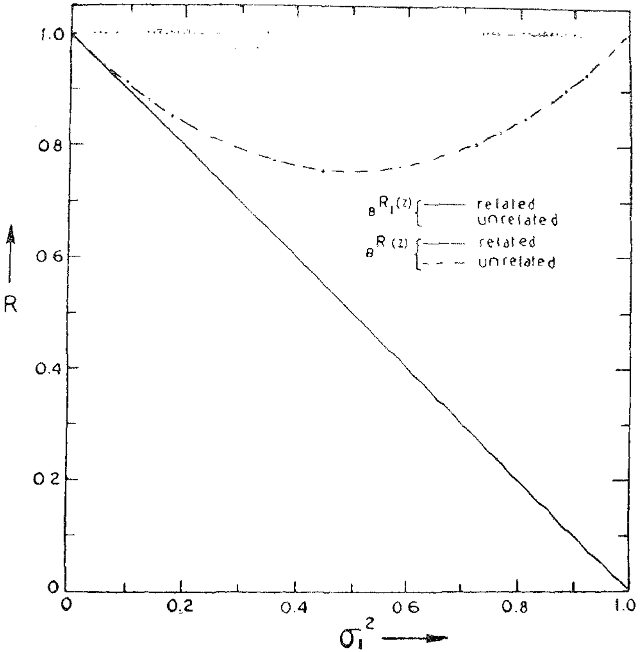

Fig. 4 *b*

Fig. 4. Graphical representation of the unnormalised and normalised Booth type indices based on normalised intensity as a function of σ_1^2 for the related and unrelated cases. The curves in (*a*) are for the *C* case and those in (*b*) are for the *NC* case. Note that the curves for the normalised and unnormalised indices corresponding to the related case have coincided.

The cases of triclinic, monoclinic and orthorhombic crystals containing p (= 1 or 2) heavy atoms in the asymmetric unit, the heavy-atom part being the model

Parthasarathi and Parthasarathy ([31] and [32]) have obtained the theoretical expressions for the indices $B^{R_1(z)}$ and $B^{R(z)}$ for the present situation under the following assumptions: (*a*) The structure factor F_Q due to the unknown atoms follow the acentric or centric Wilson distribution according as the crystal is *NC* or *C*. (*b*) All atoms in the asymmetric unit occur at general positions. (*c*) All the heavy atoms in the asymmetric unit are of the same type. The theoretical expressions have been derived for the *R* and *UR* cases and these are found to be functions of the fractional heavy-atom contribution σ_1^2 (note that $\sigma_2^2 = 1 - \sigma_1^2$). Further it has been found that the theoretical values of these indices for the *R* and *UR* cases for crystals belonging to all but two of the 74 space groups of the triclinic, monoclinic and orthorhombic systems (*Fdd*2 and *Fddd* are the exceptions) could be obtained from only 7 categories of expressions. The theoretical values of $B^{R_1(z)}$ and $B^{R(z)}$ as a function of σ_1^2 for the *R* and *UR* cases corresponding to the 7 categories of space groups are given in Table VI *a*

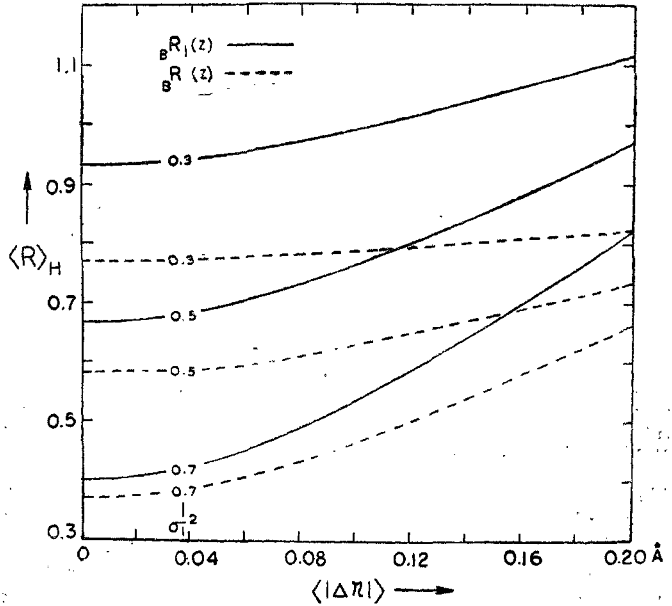

FIG. 5 *a*

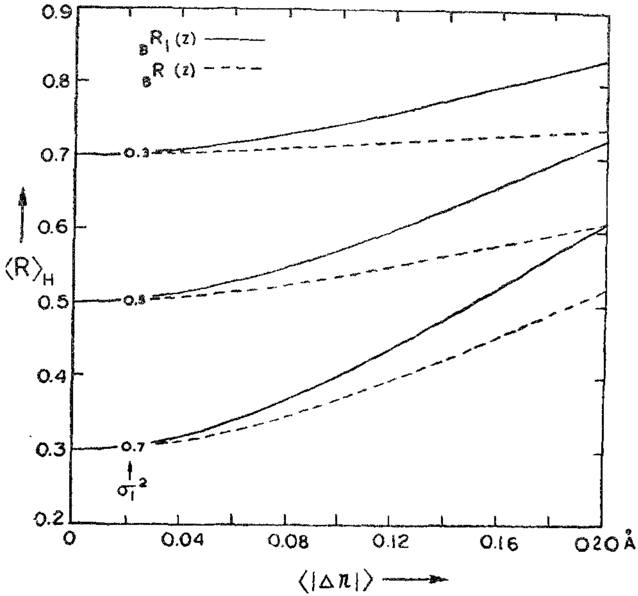

FIG. 5*b*

FIG. 5. Representation of the overall values of the normalised and unnormalised Booth type indices based on normalised intensity as a function of $\langle|\Delta r|\rangle$ for an imperfectly related incomplete model corresponding to three different values of σ_1^2, namely, 0·3, 0·5 and 0·7. The curves in (*a*) are for the *C* case and those in (*b*) are for the *NC* case.

for the case of crystals with one heavy atom in the asymmetric unit and in Table VI *b* for crystals with two heavy atoms in the asymmetric unit. In this table the numbers 1 to 7 in the first column correspond to the numbers in column 1 of Table I of Foster and Hargreaves [37] and these numbers are referred to as space group category number for convenience. From a knowledge of the space group of the crystal, the category number of the space group could be readily determined (for details *see* Foster and Hargreaves [37]). The space groups correspodning to categories 1, 3, 5

Values (in %) of the indices $B^{R(z)}$ and $B^{R_1(z)}$ for the related (R) and unrelated (UR) cases for crystals containing p (= 1 or 2) heavy atoms in the asymmetric unit for various categories of space groups belonging to the triclinic, monoclinic and arthorhombic systems*

(A) Crystals with one heavy atom in the asymmetric unit (*i.e.*, $p = 1$)

σ_1^2 ↓		Space group category number †											
		2		3		4		5		6		7	
		$B^{R(z)}$	$B^{R(z)}$	$B^{R(z)}$	$B^{R_1(z)}$	$B^{R(z)}$	$B^{R_1(z)}$	$B^{R(z)}$	$B^{R_1(z)}$	$B^{R(z)}$	$B^{R_1(z)}$	$B^{R)z_1}$	$B^{R_1(z)}$
0·00	*R*	100·0	83·3	100·0	75·0	100·0	108·3	100·0	87·5	100·0	112·5	100·0	145·8
	UR	100·0	83·3	100·0	75·0	100·0	108·3	100·0	87·5	100·0	112·5	100·0	145·8
0·10	*R*	93·5	79·9	90·2	69·9	93·2	100·0	90·1	80·0	89·9	100·0	92·9	130·0
	UR	93·8	83·2	90·7	74·9	94·1	108·4	90·9	87·5	91·1	112·5	94·5	145·8
0·20	*R*	87·1	76·2	80·8	64·6	86·2	91·6	80·4	72·4	79·6	87·6	84·9	114·1
	UR	88·4	83·0	82·8	74·7	89·6	108·4	83·4	87·4	84·6	112·4	91·2	145·6
0·30	*R*	80·6	72·1	71·6	59·1	78·8	82·9	70·8	64·6	69·2	75·3	76·1	98·4
	UR	83·8	82·5	76·2	74·4	86·4	108·5	77·6	87·4	80·3	112·4	90·2	145·3
0·40	*R*	73·9	67·4	62·5	53·1	70·8	74·0	61·2	56·6	58·8	63·2	66·7	82·8
	UR	79·7	81·9	70·8	74·0	84·7	108·7	73·5	87·2	78·4	112·3	91·5	144·9
0·50	*R*	66·7	61·9	53·3	46·7	62·2	64·4	51·6	48·4	48·5	51·5	56·6	67·7
	UR	76·2	81·0	66·7	73·3	84·4	108·9	71·0	87·1	78·8	112·1	94·9	144·4
0·60	*R*	58·5	55·3	44·0	39·6	52·7	54·2	41·9	39·8	38·3	40·2	45·9	53·0
	UR	73·2	79·7	63·7	72·5	85·7	109·2	70·2	86·9	81·3	112·0	100·5	143·9
0·70	*R*	49·0	47·0	34·2	31·6	42·2	43·0	32·0	30·8	28·3	29·3	34·9	38·8
	UR	70·6	77·9	62·1	71·5	88·7	109·5	71·1	86·7	86·0	111·8	108·0	143·2
0·80	*R*	37·3	36·3	23·8	22·6	30·2	30·6	21·7	21·2	18·5	19·0	23·5	25·2
	UR	68·6	75·5	61·9	70·2	93·7	109·9	73·9	86·4	92·6	111·6	117·3	142·4
0·90	*R*	21·8	21·6	12·5	12·2	16·3	16·4	11·1	11·0	9·1	9·2	11·8	12·2
	UR	67·2	72·0	63·3	68·7	100·9	110·4	78·7	86·1	101·0	111·4	128·3	141·6
1·00	*R*	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0
	UR	66·7	66·7	66·7	66·7	111·1	111·1	85·7	85·7	111·1	111·1	140·7	140·7

σ_1^2 ↓		Space group category number													
		1		2		3		4		5		6		7	
		$B^{R(z)}$	$B^{R1(z)}$	$B^{R(z)}$	$B^{R1(z)}$	$B^{R(z)}$	$B^{R1(z)}$	$B^{R(z)}$	$B^{R_1(z)}$	$B^{R(z)}$	$B^{R_1(z)}$	$B^{R(z)}$	$B^{R_1(z)}$	$B^{R(z)}$	$B^{R_1(z)}$
0·00	*R*	100·0	75·0	100·0	108·3	100·0	87·5	100·0	120·8	100·0	93·8	100·0	106·3	100·0	139·6
	UR	100·0	75·0	100·0	108·3	100·0	87·5	100·0	120·8	100·0	93·8	100·0	106·3	100·0	139·6
0·10	*R*	90·2	69·9	93·2	100·0	90·1	80·0	93·1	110·0	90·1	85·0	89·9	95·0	92·9	125·0
	UR	90·7	74·9	94·1	108·4	90·9	87·5	94·2	120·9	90·9	93·7	91·1	106·2	94·4	139·6
0·20	*R*	80·8	64·6	86·2	91·6	80·4	72·4	85·8	90·2	80·2	76·2	79·8	83·8	85·1	110·4
	UR	82·8	74·7	89·6	108·4	83·4	87·4	90·1	120·9	83·7	93·7	84·3	106·2	90·9	139·5
0·30	*R*	71·6	59·1	78·8	82·9	70·8	64·6	77·9	88·2	70·4	67·3	69·6	72·7	76·6	95·9
	UR	76·2	74·4	86·4	108·5	77·6	87·4	87·7	121·1	78·3	93·7	79·7	106·2	89·6	139·4
0·40	*R*	62·5	53·1	70·8	74·0	61·2	56·6	69·4	77·0	60·6	58·3	59·4	61·6	67·3	81·4
	UR	70·8	74·0	84·7	108·7	73·5	87·2	87·1	121·3	74·7	93·7	77·2	106·2	90·4	139·2
0·50	*R*	53·3	46·7	62·2	64·4	51·6	48·4	60·2	65·6	50·8	49·2	49·2	50·8	57·4	67·2
	UR	66·7	73·3	84·4	108·9	71·0	87·1	88·2	121·5	73·0	93·7	76·9	106·2	93·3	139·0
0·60	*R*	44·0	39·6	52·7	54·2	41·9	39·8	50·3	53·8	40·9	39·9	39·1	40·1	46·9	53·1
	UR	63·7	72·5	85·7	109·2	70·2	86·9	91·1	121·8	73·1	93·6	78·7	106·1	98·3	138·7
0·70	*R*	34·2	31·6	42·2	43·0	32·0	30·8	39·4	41·4	30·9	30·4	29·1	29·7	35·9	39·4
	UR	62·1	71·5	88·7	109·5	71·1	86·7	96·0	122·2	75·2	93·6	82·6	106·1	105·2	138·4
0·80	*R*	23·8	22·6	30·2	30·6	21·7	21·2	27·5	28·4	20·8	20·6	19·2	19·5	24·4	25·9
	UR	61·9	70·2	93·7	109·9	73·9	86·4	102·9	122·6	79·2	93·5	88·5	106·0	114·1	138·1
0·90	*R*	12·5	12·2	16·3	16·4	11·1	11·0	14·5	14·7	10·5	10·5	9·5	9·6	12·4	12·8
	UR	63·3	68·7	100·9	110·4	78·7	86·1	112·1	123·2	85·2	93·4	96·3	105·9	124·8	137·7
1·00	*R*	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0	0·0
	UR	66·7	66·7	111·1	111·1	85·7	85·7	123·8	123·8	93·3	93·3	105·9	105·9	137·3	137·3

---

* The (p) heavy atoms in the asymmetric unit (assumed to be in general positions) constitute the atoms of the model structure.

† Since the case $p = 1$ in category 1 is trivial (*see* Parthasarathi and Parthasarathy [31]) the values of the *R*-indices are not given here for this case.

TABLE VII

General theoretical expressions for $B^{R_1}(z)$ *for the related and unrelated types of models for complex crystals of the triclinic, monoclinic and orthorhombic systems when all atoms in the asymmetric unit are at general positions*

Space group category number	Related Case	Unrelated Case
1	$(2t_1 + t_6)/t_9$	$(2 - t_5)/t_9$
2	$(8t_1 + 3t_6)/(3t_9)$	$(8 - 3t_5)/(3t_9)$
3	$(4t_1 + t_6)/t_{10}$	$(4 - t_5)/t_{10}$
4	$(16t_1 + 3t_6)/(3t_{10})$	$(16 - 3t_5)/(3t_{10})$
5	$(8t_1 + t_8)/t_{11}$	$(8 - t_5)/t_{11}$
6	$(8t_1 - t_6)/t_{12}$	$(8 + t_5)/t_{12}$
7	$(32t_1 - 3t_6)/(3t_{12})$	$(32 + 3t_5)/(3t_{12})$

Note.—

$t_1 = 1 - \sigma_1^2,$ $t_2 = 1 - \sigma_1^2 + \sigma_1^4,$

$t_3 = 3 - 2\sigma_1^2 - \sigma_1^4,$ $t_4 = 3 - 2\sigma_1^2 + 3\sigma_1^4,$

$t_5 = C_p + C_n,$ $t_6 = (2\sigma_1^2 - 1)\, C_p - C_n,$

$t_7 = \sigma_1^4\, C_p - C_n,$ $t_8 = \sigma_1^4\, C_p + C_n,$

$t_9 = 2 - C_n,$ $t_{10} = 4 - C_n,$

$t_{11} = 8 - C_n,$ $t_{12} = 8 + C_n.$

and 6 are *NC* and those corresponding to the categories 2, 4 and 7 are *C*. A study of Table VI shows that the normalised index is preferable to the unnormalised index in testing for the correctness of the heavy atom positions. This has been found to be so in the case of a number of actual crystals containing one heavy atom per asymmetric unit [32]. This greater efficiency of the normalised *R*-index might be due to the fact that it could be interpreted to correspond to the *R*-value of a point atom structure in the case of crystals with similar atoms and to a structure with atoms in which the electron density is more concentrated (*i.e.*, an approximation towards the point atom situation) in other cases [32]. It is also interesting to see that for a given value of σ_1^2 the R-indices become more efficient for space groups of higher category number. Thus it is clear that for obtaining better results, it is necessary to take into account the space group symmetry of the crystal in the evaluation of the R-indices.

7. GENERAL THEORETICAL EXPRESSIONS FOR $B^{R_1(z)}$

Since the normalised Booth type index based on intensity has been found to be the best one for use during the structure completion stage, it would be quite useful to obtain general theoretical expressions for this index that could be used to test the correctness of any type of incomplete model of complex structures containing any number and species of atoms in the unit cell. Such expressions for the R and UR cases have been derived recently by Pathasarathy [33] for crystals belonging to the triclinic, monoclinic and orthorhombic systems (*Fdd*2 and *Fddd* are exceptions). The relevant expressions for $B^{R_1(z)}$ applicable to crystals containing all atoms in general positions are given in Table VII for the 7 categories of triclinic, monoclinic and orthorhombic space groups. It is seen that these expressions depend on three quantities, namely, σ_1^2, C_n and C_p, which are defined below.

$$\sigma_1^2 = S_p(2)/S_n(2) \tag{11}$$

$$C_j = S_j(4)/[S_j(2)]^2, \quad j = n \quad \text{or} \quad p \tag{12}$$

where

$$S_j(m) = \sum_{k=1}^{j} f_{jk}{}^m, \qquad j = n \quad \text{or} \quad p \tag{13}$$

Obviously $S_j(m)$ is the sum of the mth powers of the atomic scattering factors of the atoms in the group j ($j = n$ or p) of the asymmetric unit. Also n denotes the total number of atoms in the asymmetric unit while p denotes the number of known atoms of the asymmetric unit. Thus the theoretical values of $B^{R_1(z)}$ for the R and UR cases could be calculated from the relevant expressions of Table VII from a knowledge of the contents of the asymmetric unit of the crystal and the model. Methods of obtaining the theoretical values of this index for crystals and models with atoms in both general and special positions are also discussed by Pathasarathy [33] and for details the original paper may be referred. Theoretical expression for this index when the intensity data suffer random errors of the Gaussian type with parameters $(0, \sigma_\epsilon)$ has been shown to be [33]

$$B^{R_1(z_0)} = [B^{R_1(z)} + \sigma_\epsilon^2/\langle I_N^2\rangle]/[1 + \sigma_\epsilon^2/\langle I_N^2\rangle] \qquad (14)$$

In (14), $B^{R_1(z)}$ represents the index obtained with accurate intensity data (*i.e.*, the $|F_0| = |F_N|$ case) and $B^{R_1(z_0)}$ denotes the index calculated using normalised intensity data derived from the observed intensities.

8. OTHER APPLICATIONS OF *R*-INDICES

In crystal structure analysis, though *R*-indices are mostly used for indicating the correctness of trial structures in the structure completion stage and the extent of refinement during the refinement stage, other uses also been discussed by different workers. Thus Bhuiya and Stanley [38] (*see* also [10]) have suggested a minimum residual method for refining crystal structures and this method has been found to be successful even for cases where the conventional least squares method of refinement fails. Hamilton [39] has developed significance test based on the *R*-factor ratio and these tests allow one to decide whether the addition of parameters or the imposition of fixed relationships between parameters results in a significant improvement or a significant worsening of the agreement between the observed and calculated structure factor magnitudes. These tests are also useful for determining the absolute configuration of a non-centrosymmetric structure (*see* also Ibers and Hamilton [40]) without the necessity of measuring intensities for (hkl) and $(\bar{h}\bar{k}\bar{l})$ pairs. For details the original papers may be referred.

9. CONCLUSIONS

The various types of *R*-indices behave differently under different crystallographic situations. The index based on intensity is in general preferable to the corresponding index based on structure factor magnitude. During the structure completion stage the normalised index $B^{R_1(z)}$ is the best. During the initial refinement stage either $R_1(z)$ or $R_1^f(z)$ could be used in preference to the others. The fractional type indices based on intensity, particularly $R_1^f(z)$, are to be preferred during the final stages of refinement. The normalised index is more powerful than the unnormalised index during the structure completion stage.

ACKNOWLEDGEMENTS

The author wishes to express his sincere thanks to Professor G. N. Ramachandran for his constant interest and encouragement. He is thankful to Mr. V. Parthasarathi for this help in the preparation of the manuscript.

REFERENCES

[1] Ramachandran, G. N. and Srinivasan, R. *Fourier Methods in Crystallography*, John Wiley and Sons, New York, 1970.

[2] Srinivasan, R. and Ramachandran, G. N. Prabability distribution connected with structure amplitudes of two related crystals. IV. The distribution of the normalised difference. *Acta Crystallographica*, 1965, **19**, 1003–1007.

[3] Parthasarathy, S and Srinivasan, R. Probability distributions connected with intensities of two related crystals *Indian Journal of Pure and Applied Physics*, 1967, **5**, 502–510.

[4] Srikrishnan, T. and Srinivasan, R. Some practical aspects of the use of the new types of discrepancy factors in crystal structure refinement. *Zeitschrift fur Kristallographie*, 1968, **127**, 427–441.

[5] Luzzati, V. .. Traitement statistique des erreurs dans la determination des structures cristallines. *Acta Crystallographica*, 1952, **5**, 802–810.

[6] Srinivasan, R. and Ramachandran, G. N. 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. *Acta Crystallographica*, 1965, **19**, 1008–1014.

[7] Wilson, A. J. C. .. Largest likely values for the reliability index. *Acta Crystallographica*, 1950, **3**, 397–398.

[8] Srinivasan, R., Raghupathy Sarma, V. and Ramachandran, G. N. *Statistical Tests for Isomorphism* in *Crystallography and Crystal Perfection* (Editor: G. N. Ramachandran), Academic Press, London, 1963, pp. 85–98.

[9] Stanley, E. .. A statistical treatment of the effect of thermal parameter errors on the value of the residual. *Acta Crystallographica*, 1965, **18**, 579–580.

[10] Stanley, E. .. Rate of refinement of coordinates in the minimum residual method. *Acta Crystallographica*, 1964, **17**, 1028–1032.

[11] Stanley, E. .. The rate of refinement of coordinates for the minimum residual method in three dimensions. *Acta Crystallographica*, 1965, **18**, 1086–1087.

[12] Stanley, E. .. Standard deviation of coordinates as a function of the residual and temperature coefficient. *Acta Crystallographica*, 1965, **19**, 1055–1056.

[13] Chandrasekharan, R. and Srinivasan, R. On the use of the normalised discrepancy index and other statistical parameters in crystal structure refinement. *Zeitschrift fur Kristallographie*, 1969, **129**, 435–450.

[14] Booth, A. D. .. An expression for following the process of refinement in x-ray structure analysis using Fourier series. *Philosophical Magazine*, 1945, **36**, 609–615.

[15] Parthasarathy, S. and Parthasarathi, V. Discrepancy factors for use in crystal structure analysis. *Acta Crystallographica*, 1972, **A28**, 426–432.

[16] Ison, A. J. C. .. Effect of a badly misplaced atom on the residual. *Acta Crystallographica*, 1969, **B25**, 1288–1293.

[17] Lenstra, A. T. H. .. The residual and its use for partial structure evaluation. A critical note on the discriminator function. *Acta Crystallographica*, 1974, **A30**, 362–369.

[18] Luzzati, V. .. Traitement statistique des erreurs dans la cas d'une structure pseudo-centrosymmetrique. *Acta Crystallographica*, 1953, **6**, 550–552.

[19] Srinivasan, R. and Srikrishnan, T. New criteria for use in crystal structure refinement. *Zeitschrift fur Kristallographie*, 1966, **126**, 382–387.

[20] Srikrishnan, T. and Srinivasan, R. Studies on the use of the new types of discrepancy factors in an actual structure refinement. *Indian Journal of Pure and Applied Physics*, 1971, **9**, 212–214.

[21] Parthasarathy, S. and Parthasarathi, V. Discrepancy factors for use in crystal structure analysis. III. A theoretical comparison of the normalised indices. *Acta Crystallographica*, 1975, **A31**,

[22] Parthasarathi, V. .. Some Applications of probability methods in X-ray crystallography. *Thesis* (unpublished), Centre of advanced study in Physics, University of Madras, 1975.

[23] Douglas, A. S. and Woolfson, M. M. The maximum probable value of the reliability index for a hypercentric structure. *Acta Crystallographica*, 1954, **7**. 517.

[24] Rogers, D. and Wilson, A. J. C. The probability distribution of X-ray intensities. V. A note on hypersymmetric distributions. *Acta Crystallographica*, 1953, **6**, 439–449.

[25] Wilson, A. J. C. .. The probability distribution of x-ray intensities. VII. Some sesquisymmetric distributions. *Acta Crystallographica* 1956, **9**, 143–144.

[26] Wilson A. J. C. .. Largest likely values of residuals. *Acta Crystallographica*, 1974, **A30**, 836–838.

[27] Parthasarathi, V. and Parthasarathy, S. Some statistical results for structures with pseudosymmetry. *Acta Crystallographica*, 1975, **A31**.

[28] Swaminathan, P. and Srinivasan, R. Probability distribution connected with structure amplitudes of two related crystals. VII. The case of an approximately centrosymmetric structure. *Acta Crystallographica*, 1975 (In Press).

[29] Srikrishnan, T. and Parthasarathy, S. Discrepancy factors for crystals with two heavy atoms. *Indian Journal of Pure and Applied Physics*, 1972, **10**, 672–675.

[30] Parthasarathy, S., Parthasarathi, V. and Srikrishnan, T. Conventional reliability index for crystals containing two heavy atoms. *Indian Journal of Pure and Applied Physics*, 1973, **11**, 248–249.

[31] Parthasarathi, V. and Parthasarathy, S. Discrepancy factors for the use in crystal structure analysis. II. Normalised index R_2 for crystals with a few heavy atoms in the asymmetric unit belonging to the triclinic, monoclinic and orthorhombic systems. *Acta Crystallographica*, 1974, **A30**.

[32] Parthasarathi, V. and Parthasarathy, S. Discrepancy indices for use in crystal structure analysis. V. A comparative study of the normalised and unnormalised Booth-type indices on the structure completion stage. *Acta Crystallographica* (sent for publication).

[33] Parthasarathy, S. .. Discrepancy indices for use in crystal structure analysis IV. Theoretical expressions for the indices $B^{R_1(I)}$ and $B^{R(I)}$ for any incomplete model of complex structures. *Acta Crystallographica*, 1975, **A31**.

[34] Wilson, A. J. C. .. Personal communication.

[35] Lenstra, A. T. H. .. The influence of space group symmetry, pseudosymmetry and atomic scattering power on the path of the residual. *Acta Crystallographica*, 1975 (sent for publication).

[36] Parthasarathy, S. and Parthasarathi, V. The effect of coordinate errors on the phase angle distribution. *Acta Crystallographica*, 1974, **A30**, 43–46.

[37] Foster, F. and Hargreaves, A. The use of moments of X-ray intensity in space-group determination. II. Practical application. *Acta Crystallographica*, 1963, **16**, 1133–1139.

[38] Bhuiya, A. K. and Stanley, E. The refinement of atomic parameters by direct calculation. of the minimum residual. *Acta Crystallographica*, 1963, **16**, 981–984.

[39] Hamilton, W. C. .. Significance tests on the crystallographic R-factor. *Acta. Crystallographica*, 1965, **18**, 502–510.

[40] Ibers, J. A. and Hamilton, W. C. Dispersion connections and crystal structure refinements *Acta Crystallographica*, 1964, **17**, 781–782.