Acta Cryst. (1975). A31, 536

Discrepancy Indices for Use in Crystal Structure Analysis. VI. A Study on the Efficiency of the Indices ${}_4R_1(I)$ and ${}_4R(I)$

BY V. PARTHASARATHI AND S. PARTHASARATHY

Centre of Advanced Study in Physics, University of Madras, Guindy Campus, Madras*-600025, *India*

(*Received* 11 *January* 1975; *accepted* 14 *March* 1975)

Expressions for ${}_4R_1(I)$ and ${}_4R(I)$, which are based on the fourth powers of the discrepancies $(I_N-I_P^c/\sigma_1^2)$ and $(I_N-I_P^c)$ respectively, are worked out for the case of a crystal and model satisfying the requirements of the Wilson distributions. A comparative study of ${}_4R_1(I)$ and ${}_4R(I)$ is carried out to see whether the concept of a normalized index is useful. Since ${}_4R_1(I)$ is found to be better, it is also compared with ${}_BR_1(I)$. As ${}_4R_1(I)$ could be useful in the initial stages of the structure completion process, general expressions for this index for the related and unrelated types of models of triclinic, monoclinic and orthorhombic crystals containing any number and types of atoms at general positions in the asymmetric unit are obtained.

1. Introduction

A study in Part III of this series (Parthasarathy & Parthasarathi, 1975) on six different types of normalized R indices in three crystallographic situations showed that the Booth-type index ${}_BR_1(I)$ is preferable to the rest in the structure completion stage. This greater efficiency of ${}_BR_1(I)$ might be due to the fact that of all the indices, it is based on the highest power (namely, the square) of the discrepancies $(I_N-I_P^c/\sigma_1^2)$. It is now natural to ask whether it would be advantageous to use indices based on higher powers of $(I_N-I_P^c/\sigma_1^2)$. The index based on the cubes of $(I_N-I_P^c/\sigma_1^2)$ is not expected to be efficient since, even for the unrelated case, the value of this index would be small. Though this inefficiency could be overcome by taking the third power of $|I_N-I_P^c/\sigma_1^2|$, the handling of such a quantity involves theoretical difficulties. In this paper we shall therefore consider the normalized index ${}_4R_1(I)$ based on the fourth powers of $(I_N-I_P^c/\sigma_1^2)$ and the corresponding unnormalized index ${}_4R(I)$ based on those of $(I_N-I_P^c)$ [see (1) and (2) for a definition of these indices] and study their efficiency relative to the others studied in Part III. In this paper we shall consider only crystals in which all atoms in the asymmetric unit occur at general positions. In §2 we shall obtain the theoretical expressions for these indices for the case of an imperfectly related incomplete model when both the model and crystal satisfy the requirements of a Wilson distribution and use these to compare the normalized and unnormalized indices ${}_4R_1(I)$ and ${}_4R(I)$ during the structure completion stage. Such a study is useful to see whether the concept of the normalized index is useful. A comparative study of ${}_4R_1(I)$ and ${}_BR_1(I)$ is also carried out in §2. In §3 general expressions for ${}_4R_1(I)$ and ${}_4R(I)$ are derived for the related (*i.e.* R) and unrelated (*i.e.* UR) cases. These are valid for crystals and models containing any number and types of atoms at general positions in the asymmetric unit and belonging to the seven categories of space groups (for details see Part II, Parthasarathi & Parthasarathy, 1975*a*) of the triclinic, monoclinic and orthorhombic systems.

* Contribution No. 401.

2. Studies pertaining to crystals and models satisfying the requirements of Wilson distributions for the imperfectly related case

We define ${}_4R_1(I)$ and ${}_4R(I)$ to be

$$ {}_4R_1(I)=\sum(I_N-I_P^c/\sigma_1^2)^4/\sum I_N^4=\langle(z_N-z_P^c)^4\rangle/\langle z_N^4\rangle \quad (1)$$

$$ {}_4R(I)=\sum(I_N-I_P^c)^4/\sum I_N^4=\langle(z_N-\sigma_1^2 z_P^c)^4\rangle/\langle z_N^4\rangle . \quad (2)$$

We know that $\langle(z_N-z_P^c)^4\rangle=\langle z_d^4\rangle=144\sigma_B^4$ for the C case and $24\ \sigma_B^4$ for the NC case (see Parthasarathy & Srinivasan, 1967, hereafter PS, 1967). Hence it follows from (1) that

$$ {}_4R_1(I)=\tfrac{48}{35}\sigma_B^4 \text{ for } C \text{ and } \sigma_B^4 \text{ for } NC . \quad (3)$$

From the expansion of $(z_N-\sigma_1^2 z_P^c)^4$ by the binomial theorem it can be shown from (2) that

$$ {}_4R(I)=[\langle z_N^4\rangle-4\sigma_1^2\langle z_N^3 z_P^c\rangle+6\sigma_1^4\langle z_N^2(z_P^c)^2\rangle -4\sigma_1^6\langle z_N(z_P^c)^3\rangle+\sigma_1^8\langle(z_P^c)^4\rangle]/\langle z_N^4\rangle . \quad (4)$$

It is known that $\langle z_N^2(z_P^c)^2\rangle=\langle z_p^2\rangle=4(1+4\sigma_A^2+\sigma_A^4)$ for the NC case and $9(1+8\sigma_A^2+\frac{8}{3}\sigma_A^4)$ for the C case (PS, 1967). Using the joint p.d.f. of z_N and z_P^c available in PS (1967), we can show that $\langle z_N^3 z_P^c\rangle=\langle z_N(z_P^c)^3\rangle=6+18\sigma_A^2$ for the NC case and $15+90\sigma_A^2$ for the C case. Making use of these results in (4) and simplifying, we obtain for the C case

$$ {}_4R(I)=1-\tfrac{4}{7}\sigma_1^2+\tfrac{18}{35}\sigma_1^4-\tfrac{4}{7}\sigma_1^6+\sigma_1^8-\tfrac{24}{7}(1-\tfrac{6}{5}\sigma_1^2+\sigma_1^4)\sigma_1^2\sigma_A^2+\tfrac{48}{35}\sigma_1^4\sigma_A^4 \quad (5)$$

and for the NC case

$$ {}_4R(I)=1-\sigma_1^2+\sigma_1^4-\sigma_1^6+\sigma_1^8-(3-4\sigma_1^2+3\sigma_1^4)\sigma_1^2\sigma_A^2+\sigma_1^4\sigma_A^4 . \quad (6)$$

Comparison of the normalized index ${}_4R_1(I)$ with unnormalized index ${}_4R(I)$

Structure completion process: The relevant curves are those in Fig. 1(*a*). A study of this figure using the slope criterion of Part III shows that for $\sigma_1^2 \lesssim 0{\cdot}5$, the normalized is preferable to the unnormalized index. In the region $\sigma_1^2 > 0{\cdot}5$, though they appear to be equally good, the normalized index is to be preferred since for the *UR* case it is practically flat while the unnormalized index has a shallow minimum.

Refinement of an incomplete model: The relevant curves are those in Fig. 1(*b*) and Fig. 1(*c*) which show that the normalized is preferable to the unnormalized index.

Since the above study has shown that ${}_4R_1(I)$ is preferable to ${}_4R(I)$ and since among the indices studied in Part III ${}_BR_1(I)$ is the best (see also Parthasarathi & Parthasarathy, 1975*b*) during the structure completion stage, it would suffice if we compare the relative efficiency of ${}_4R_1(I)$ and ${}_BR_1(I)$.

Comparison of the index ${}_4R_1(I)$ and ${}_BR_1(I)$

Structure completion process: The relevant curves are those in Fig. 2(*a*). It is seen that ${}_4R_1(I)$ is preferable to ${}_BR_1(I)$, when σ_1^2 is not large (say $\sigma_1^2 \lesssim 0{\cdot}5$). It may also be noted that the distinction between the *R* and *UR* cases is more marked for ${}_4R_1(I)$ than for ${}_BR_1(I)$.

Refinement stage: The relevant curves are those in Fig. 2(*b*) and Fig. 2(*c*) which show that ${}_4R_1(I)$ could be preferred to ${}_BR_1(I)$ during the refinement of an incomplete model. However during the refinement of a complete model (*i.e.* $\sigma_1^2 = 1$) ${}_BR_1(I)$ is preferable.

An analogous comparative study of ${}_4R_1(I)$, ${}_4R(I)$ and ${}_BR_1(I)$ has also been carried out for the situation con-

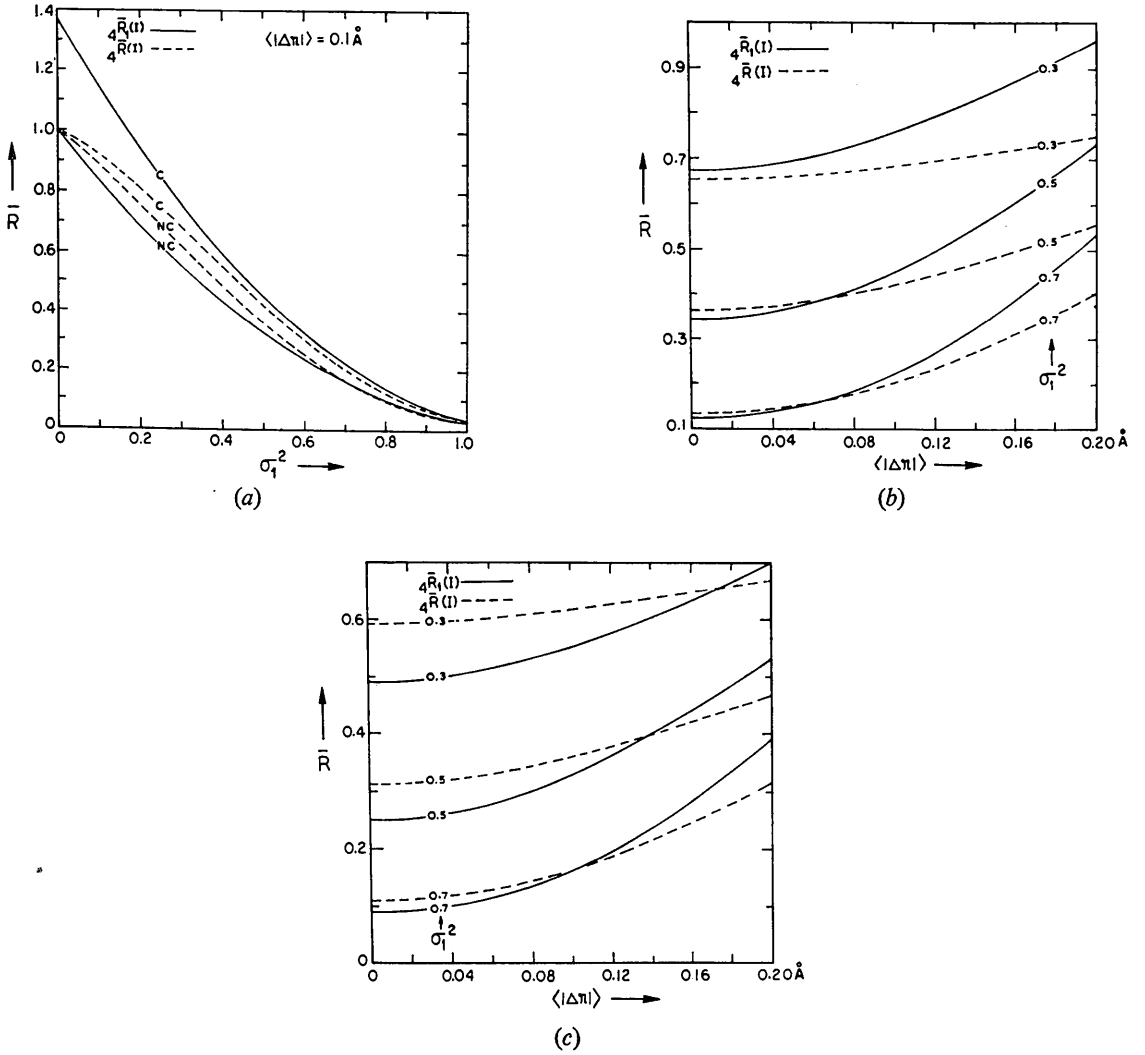

Fig. 1. Comparison of the relative efficiency of the indices ${}_4R_1(I)$ and ${}_4R(I)$ in the structure completion stage when both the crystal and the model satisfy the requirements of a Wilson distribution: (*a*) Variation of the overall values of these indices as a function of σ_1^2 when $\langle|\Delta\mathbf{r}|\rangle = 0{\cdot}1$ Å for the *C* and *NC* cases. (*b*) Variation of the overall values of these indices as a function of $\langle|\Delta\mathbf{r}|\rangle$ for different fixed values of σ_1^2, namely $\sigma_1^2 = 0{\cdot}3$, $0{\cdot}5$ and $0{\cdot}7$, for the *C* case. (*c*) Same as (*b*) for the *NC* case.

sidered in Part II. For this situation also it has been found that* (i) ${}_4R_1(I)$ is better than ${}_4R(I)$ and (ii) ${}_4R_1(I)$ is preferable to ${}_BR_1(I)$ when the heavy-atom contribution is not large (*i.e.* $\sigma_1^2 \lesssim 0{\cdot}5$). Thus it would be useful to derive expressions for ${}_4R_1(I)$ for the general case when any part of a crystal containing any number and types of atoms in the asymmetric unit constitutes the model. Such expressions are given in §3 for the *R* and *UR* cases and are applicable to crystals belonging to the seven categories of space groups of the triclinic, monoclinic and orthorhombic systems. For the sake of completeness the expressions for ${}_4R(I)$ are also listed.

* Though the details of the results, tables and figures for this case are available, at the suggestion of the referees they are not given here.

3. General expressions for ${}_4R_1(I)$ and ${}_4R(I)$

Though the method of obtaining expressions for ${}_4R_1(I)$ and ${}_4R(I)$ is more tedious than for ${}_BR_1(I)$ and ${}_BR(I)$, it is similar in principle. It consists of the following steps: (i) Expansion of $(z_N - z_P^c)^4$ or $(z_N - \sigma_1^2 z_P^c)^4$ by the binomial theorem. (ii) Using the properties that z_N and z_P^c are independent for the *UR* case and that $z_P^c = z_P$ for the *R* case. (iii) Using the following properties of $\langle \varepsilon_{PQ}^n \rangle$ ($=\varepsilon_n$, say), namely, that for odd values of n, $\varepsilon_n = 0$ (for both the *C* and *NC* cases); for any n, $\varepsilon_{2n} = 1$ for the *C* case and $\varepsilon_2 = \frac{1}{2}$ and $\varepsilon_4 = \frac{3}{8}$ for the *NC* case. (iv) Substituting the expressions for the moments of z_P, z_Q and z_N (see Parthasarathy, 1973) at the relevant places. The results that were obtained are summarized below.

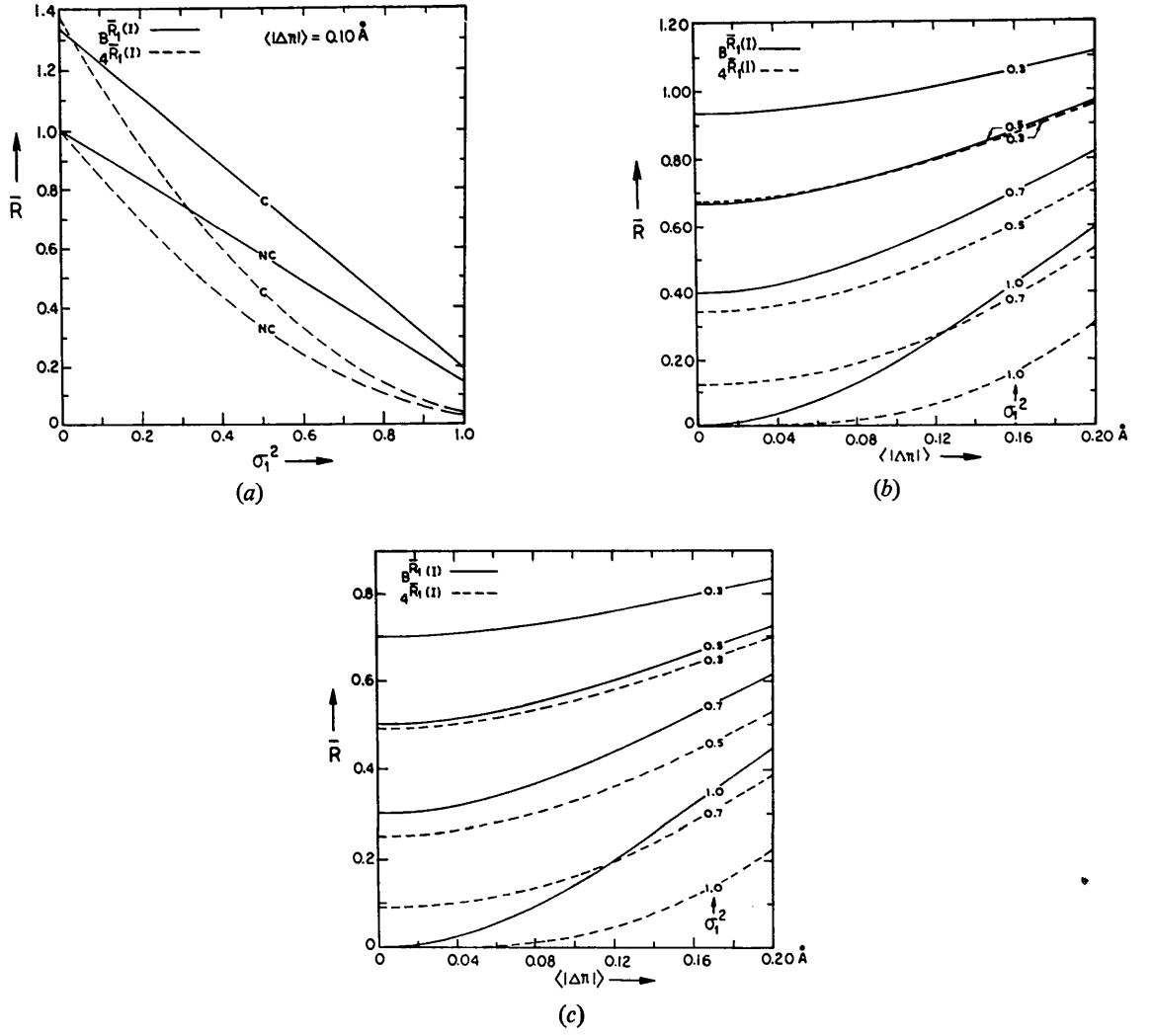

Fig. 2. Comparison of the relative efficiency of the indices ${}_4R_1(I)$ and ${}_BR_1(I)$ under different crystallographic situations when both the crystal and the model satisfy the requirements of a Wilson distribution: (*a*) Variation of the overall values of these indices as a function of σ_1^2 when $\langle |\Delta \mathbf{r}| \rangle = 0{\cdot}1$ Å for the *C* and *NC* cases. (*b*) Variation of the overall values of these indices as a function of $\langle |\Delta \mathbf{r}| \rangle$ for different fixed values of σ_1^2, namely, $\sigma_1^2 = 0{\cdot}3$, 0·5 and 0·7 and 1·0 for the *C* case. (*c*) Same as (*b*) for the *NC* case.

Table 1. *Values of the constants k_j needed to evaluate the R indices ${}_4R_1(I)$ and ${}_4R(I)$*

Space-group category number*	k_1	k_2	k_3	k_4	k_5	k_6	k_7
1	1·000	9·000	4·00	72·00	64	33·0000	18·0000
2	1·500	22·500	10·00	315·00	280	144·0000	78·7500
3	0·500	4·500	1·00	36·00	16	4·1250	4·5000
4	0·750	11·250	2·50	157·50	70	18·0469	19·6875
5	0·250	2·250	0·25	18·00	4	2·7656	3·3750
6	−0·250	−2·250	−2·00	−18·00	−32	−8·0156	1·1250
7	−0·375	−5·625	−5·00	−78·75	−140	−35·0684	4·9219

* Categories 1, 3, 5 and 6 correspond to the *NC* case and 2, 4 and 7 to the *C* case.

Related case:

$$
\begin{aligned}
{}_4R_1(I) = [&16\varepsilon_4(W_2^2 - k_1W_2D_1 + k_1^2D_5)\sigma_1^4\sigma_2^4 + 24\varepsilon_2\{2W_3 \\
&- 2W_2^2 - (k_2 - 2k_1W_2)D_1 + k_3D_2 \\
&- 2k_1^2D_5\}\sigma_1^2\sigma_2^6 + \{2W_4 - 8W_3 + 6W_2^2 \\
&- (k_4 - 4k_2 + 6k_1W_2)D_1 + (k_5 - 4k_3)D_2 - k_6D_3 \\
&+ k_7D_4 + 6k_1^2D_5\}\sigma_2^8]/\langle z_N^4\rangle \qquad (7)
\end{aligned}
$$

$$
\begin{aligned}
{}_4R(I) = [&16\varepsilon_4(W_2^2 - k_1W_2D_1 + k_1^2D_5)\sigma_1^4\sigma_2^4 + 24\varepsilon_2\{W_3 \\
&- k_2C_q(4) + k_3C_q(6)\}\sigma_1^2\sigma_2^6 + \{W_4 - k_4C_q(4) \\
&+ k_5C_q(6) - k_6C_q(8) + k_7[C_q(4)]^2\}\sigma_2^8]/\langle z_N^4\rangle . \qquad (8)
\end{aligned}
$$

Unrelated case:

$$
\begin{aligned}
{}_4R_1(I) = [&2(W_4 - 4W_3 + 3W_2^2) - (k_4 - 4k_2 + 6k_1W_2) \\
&\times \{C_n(4) + C_p(4)\} + (k_5 - 4k_3)\{C_n(6) + C_p(6)\} \\
&- k_6\{C_n(8) + C_p(8)\} + k_7\{C_n(4)\}^2 + k_7\{C_p(4)\}^2 \\
&+ 6k_1^2C_n(4)C_p(4)]/\langle z_N^4\rangle \qquad (9)
\end{aligned}
$$

$$
\begin{aligned}
{}_4R(I) = [&(W_4 - 4W_3\sigma_1^2 + 6W_2^2\sigma_1^4 - 4W_3\sigma_1^6 + W_4\sigma_1^8) \\
&- (k_4 - 4k_2\sigma_1^2 + 6k_1W_2\sigma_1^4)C_n(4) + (k_5 - 4k_3\sigma_1^2) \\
&\times C_n(6) - k_6C_n(8) + k_7\{C_n(4)\}^2 - (k_4\sigma_1^4 - 4k_2\sigma_1^2 \\
&+ 6k_1W_2) + \sigma_1^4C_p(4) + (k_5\sigma_1^2 - 4k_3)\sigma_1^6C_p(6) \\
&- k_6\sigma_1^8C_p(8) + k_7\sigma_1^8\{C_p(4)\}^2 \\
&+ 6k_1^2\sigma_1^4C_n(4)C_p(4)]/\langle z_N^4\rangle . \qquad (10)
\end{aligned}
$$

Here k_j, $j=1$ to 7, are constants which are defined in Table 1 and W_i is the ith-order moment of the normalized intensity for the Wilson distributions. That is, $W_i=2$, 6 and 24 for $i=2$, 3 and 4 for the *NC* case and $W_i=3$, 15 and 105 for $i=2$, 3 and 4 for the *C* case. $C_i(2m)$, $i=p$, q or n, are defined to be

$$C_i(2m) = S_i(2m)/[S_i(2)]^m . \qquad (11)$$

Here $S_i(m)$ is the sum of the mth powers of the scattering factors of the atoms of the ith group ($i=p$, q or n) in the asymmetric unit. The D_i's are defined to be

$$
\begin{aligned}
&D_1 = C_q(4) + C_p(4),\ D_2 = C_q(6) + C_p(6), \\
&D_3 = C_q(8) + C_p(8),\ D_4 = [C_q(4)]^2 + [C_p(4)]^2, \\
&D_5 = C_q(4)\ .\ C_p(4)\ . \qquad (12)
\end{aligned}
$$

$\langle z_N^4\rangle$ is to be obtained from

$$
\begin{aligned}
\langle z_N^4\rangle = {} & W_4 - k_4C_n(4) + k_5C_n(6) - k_6C_n(8) \\
& + k_7[C_n(4)]^2 . \qquad (13)
\end{aligned}
$$

4. Concluding remarks

The above study shows that ${}_4R_1(I)$ is preferable to ${}_BR_1(I)$ only during the initial stages of structure analysis (*i.e.* when $\sigma_1^2 \lesssim 0{\cdot}5$). If this observation could be extrapolated we could state that the index* ${}_6R_1(I)$ based on the sixth powers of $(I_N - I_P^c/\sigma_1^2)$ and those based on still higher powers would not be as useful as the Booth-type index ${}_BR_1(I)$ based on the squares of $(I_N - I_P^c/\sigma_1^2)$. That this is so can be seen for the Wilson case, since for this case corresponding to an *NC* crystal ${}_6R_1(I) = \sigma_B^6$ [a result that could be obtained from (A-17) of PS, 1967]. Thus ${}_6R_1(I)$ would fall to very low values even for small values of σ_1^2 and thereafter remain practically insensitive. Thus it appears that it might not be advantageous to make use of R indices based on higher powers than the fourth.

One of the authors (V.P.) thanks the Council of Scientific and Industrial Research, New Delhi, India for financial assistance. The authors thank the referees for their valuable suggestions.

* The index ${}_6R_1(I)$ is to be defined as ${}_6R_1(I) = \sum(I_N - I_P^c/\sigma_1^2)^6/\sum I_N^6 = \langle(z_N - z_P^c)^6\rangle/\langle z_N^6\rangle$.

References

PARTHASARATHI, V. & PARTHASARATHY, S. (1975*a*). *Acta Cryst.* A**31**, 38–41.

PARTHASARATHI, V. & PARTHASARATHY, S. (1975*b*). *Acta Cryst.* A**31**, 529–535.

PARTHASARATHY, S. (1973). *Acta Cryst.* A**29**, 725–726.

PARTHASARATHY, S. & PARTHASARATHI, V. (1975). *Acta Cryst.* A**31**, 178–185.

PARTHASARATHY, S. & SRINIVASAN, R. (1967). *Indian J. Pure Appl. Phys.* **5**, 502–510.