Acta Cryst. (1975). A**31**, 149

Some Statistical Results for Structures with Pseudosymmetry*

BY V. PARTHASARATHI AND S. PARTHASARATHY

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

(*Received* 31 *August* 1974; *accepted* 11 *September* 1974)

The values of the second, third and fourth-order moments of the normalized intensity z and the maximum probable values of two types of discrepancy index, namely R_B and R_2, are worked out for various types of pseudosymmetric structures considered by Rogers & Wilson [*Acta Cryst.* (1953). **6**, 439–449].

Introduction

Rogers & Wilson (1953, hereafter referred to as RW) have derived a number of statistical results for a variety of probable types of pseudosymmetric structures and these results, which are useful in the recognition of type of pseudosymmetry in crystal structures, include the evaluation of the cumulative function of the normalized intensity z, the test ratio ϱ and the variance of z. The types of pseudosymmetry† considered by these authors are (i) hypercentrosymmetry, (ii) hyperparallelism, (iii) parallel repetition of a motif at regular intervals along a straight line and (iv) many repetitions of a motif at random in a centrosymmetric or non-centrosymmetric space group. Their results show that in general such structures are characterized by a large percentage of very weak reflexions. The study by Parthasarathy (1966*a*) has shown that the higher-moment test is best suited for structures having a large percentage of unobserved reflexions. We shall therefore work out the values of the higher moments of z for the various types of pseudosymmetric structures considered in RW. It may incidentally be noted here that the second-order moment of z is also needed for evaluating the maximum probable value of R_2.

Rogers & Wilson (1953) have discussed qualitatively the value of the conventional R index to be expected for incorrect hypersymmetric structures but, owing to theoretical difficulties, no quantitative results are available for the various types of pseudosymmetric structure. However, Douglas & Woolfson (1954) were able to obtain the maximum probable value of R for a particular type of pseudosymmetry, namely the hypercentric case (Lipson & Woolfson, 1952). Since R_2 [see equation (12) below for the definition of this index] is easier to handle theoretically (Wilson, 1969), we shall work out the maximum probable values of R_2 for the various types of pseudosymmetric structures considered in RW. We shall also evaluate the Booth-type index R_B [Booth, 1945; Parthasarathy & Parthasarathi, 1972; see equation (11) below for the definition] for the various cases.

In this paper we shall follow the notation employed in RW. The probability density functions of z needed for the evaluation of the moments of z for the various types of pseudosymmetric structure are given in Table 1 and these are taken from RW. The values of the test ratio (ϱ) for the various cases are also given in Table 1, since these are required for the evaluation of R_B.

Evaluation of the moments of z

Hypercentrosymmetry and hyperparallelism

It has been shown in RW that the distributions of z for both these types are given by the same expression, namely equation (26) of RW and hence the moments for these two types will also be given by the same expression. From equation (33) of RW, the mth moment of z could be readily shown to be

$$\langle z^m \rangle_n = \pi^{-n/2} 2^{nm} [\Gamma(m+\tfrac{1}{2})]^n [\Gamma(m+1)]^{-n+1}. \quad (1)$$

Regular parallel repetition in a line

(*a*) *Non-centrosymmetric motif*: From equation (2) of Table 1 we obtain the mth moment of z to be

$$\langle z^m \rangle_n = \frac{2n}{\pi} \int_0^\infty \int_0^{\pi/2} z^m \frac{\sin^2 \psi}{\sin^2 n\psi} \exp\left\{-\frac{nz \sin^2 \psi}{\sin^2 n\psi}\right\} \mathrm{d}\psi \mathrm{d}z \quad (2)$$

$$= \frac{2}{\pi} \frac{\Gamma(m+1)}{n^m} \int_0^{\pi/2} \left(\frac{\sin^2 n\psi}{\sin^2 \psi}\right)^m \mathrm{d}\psi\,, \quad (3)$$

where we have used equation (3·478–1) on p. 342 of Gradshteyn & Ryzhik (1965; hereafter referred to as GR). The integral in (3) may be evaluated with the aid of the Fourier cosine expansion of $\sin^2 n\psi/\sin^2 \psi$ (see p. 171 of Whittaker & Watson, 1952). It can be shown that

$$\frac{1}{\pi} \int_0^{\pi/2} \left(\frac{\sin^2 n\psi}{\sin^2 \psi}\right)^2 \mathrm{d}\psi = (2n^3+n)/6 \quad \text{for } n=1,2,\ldots \quad (4)$$

* Contribution No.383 from the Centre of Advanced Study in Physics, University of Madras, Guindy Campus, Madras-600025, India.

† For a description of these one may refer to the original paper (RW).

$$\frac{1}{\pi}\int_0^{\pi/2}\left(\frac{\sin^2 n\psi}{\sin^2\psi}\right)^3 d\psi=\begin{cases}0{\cdot}5 & \text{for } n=1\\ 10 & \text{for } n=2\\ 70{\cdot}5 & \text{for } n=3\\ 290 & \text{for } n=4\\ 875{\cdot}5 & \text{for } n=5\\ 2166 & \text{for } n=6\end{cases} \quad (5)$$

$$\frac{1}{\pi}\int_0^{\pi/2}\left(\frac{\sin^2 n\psi}{\sin^2\psi}\right)^4 d\psi=\begin{cases}0{\cdot}5 & \text{for } n=1\\ 35 & \text{for } n=2\\ 553{\cdot}5 & \text{for } n=3\\ 4046 & \text{for } n=4\\ 19082{\cdot}5 & \text{for } n=5\\ 67977 & \text{for } n=6\,.\end{cases} \quad (6)$$

By substitution of (4), (5) and (6) in (3), the quantity $\langle z^m\rangle_n$ could be evaluated for any given $n(\le 6)$ and $m(\le 4)$.

(*b*) *Centrosymmetric motif*: From equation (3) of Table 1 we obtain the mth moment of z to be

$$\langle z^m\rangle_n=\frac{1}{\pi}\sqrt{\frac{2n}{\pi}}\int_0^\infty\int_0^{\pi/2} z^{m-1/2}\exp\left\{-\frac{nz\sin^2\psi}{2\sin^2 n\psi}\right\}\times\left|\frac{\sin\psi}{\sin n\psi}\right| d\psi dz$$
$$=\frac{2^{m+1}}{\pi^{3/2}}\,\frac{\Gamma(m+\frac{1}{2})}{n^m}\int_0^{\pi/2}\left(\frac{\sin^2 n\psi}{\sin^2\psi}\right)^m d\psi\,, \quad (7)$$

where we have used equation (3·478–1) on p. 342 of GR. Substituting (4)–(6) in (7), the second, third and fourth moments of z for any $n(\le 6)$ could be evaluated.

Many parallel repetitions at random

(*a*) *Non-centrosymmetric motif in a non-centrosymmetric arrangement*: From equation (4) of Table 1 we obtain

$$\langle z^m\rangle=\int_0^\infty 2z^m K_0(2\sqrt{z})dz=[m!]^2, \quad (8)$$

where we have used equation (16) on p. 684 of GR.

(*b*) *Centrosymmetric motif in a non-centrosymmetric arrangement*. From equation (5) of Table 1 we obtain

$$\langle z^m\rangle=\frac{1}{\sqrt{2}}\int_0^\infty z^{m-1/2}\exp\{-\sqrt{2z}\}dz=2^{-m}(2m)!\,, \quad (9)$$

where we have used equation (3·351–3) on p. 310 of GR.

(*c*) *Centrosymmetric motif in a centrosymmetric arrangement*: From equation (6) of Table 1 we obtain

$$\langle z^m\rangle=\frac{1}{\pi}\int_0^\infty z^{m-1/2}K_0(\sqrt{z})dz=[\pi^{-1/2}2^m\Gamma(m+\tfrac{1}{2})]^2, \quad (10)$$

Table 1. *Wilson's ratio ϱ and probability density function $P(z)$ for structures with different types of pseudosymmetry considered by Rogers & Wilson* (1953)

The data of this table are collected from RW. Equations (1) and (2) are the same as equations (26) and (39) of RW. Equations (3)–(6) could be derived respectively from the expressions for the function $P(|F|)$ obtained in equations (48), (60), (71) and (78) of RW. The abbreviations *NC* and *C* stand for 'non-centrosymmetric' and 'centrosymmetric' respectively.

Hypercentrosymmetry and hyperparallelism

$$\varrho_n=2^{3n-2}\pi^{-2n+1}$$

$$P_n(z)=2^{n/2-1}\pi^{-n+1/2}\int_0^{\pi/2}\dots\int_0^{\pi/2}z^{-1/2}\exp\{-z\sec^2\psi_2\dots\sec^2\psi_n/2^n\}\sec\psi_2\dots\sec\psi_n d\psi_2\dots d\psi_n \quad (1)$$

Regular parallel repetition in line: *NC* motif*

$$P_n(z)=\frac{2n}{\pi}\int_0^{\pi/2}\frac{\sin^2\psi}{\sin^2 n\psi}\exp\left\{\frac{-zn\sin^2\psi}{\sin^2 n\psi}\right\}d\psi \quad (2)$$

Regular parallel repetition in line: *C* motif†

$$P_n(z)=\frac{1}{\pi}\sqrt{\frac{2n}{\pi}}\int_0^{\pi/2}z^{-1/2}\exp\left\{-\frac{zn\sin^2\psi}{2\sin^2 n\psi}\right\}\left|\frac{\sin\psi}{\sin n\psi}\right|d\psi \quad (3)$$

Many parallel repetitions at random

(*a*) *NC* motif in a *NC* arrangement

$$\varrho=0{\cdot}617\,,\quad P(z)=2K_0(2\sqrt{z}) \quad (4)$$

(*b*) *C* motif in a *NC* arrangement

$$\varrho=0{\cdot}500\,,\quad P(z)=\frac{1}{\sqrt{2z}}\exp\{-\sqrt{2z}\} \quad (5)$$

(*c*) *C* motif in a *C* arrangement

$$\varrho=0{\cdot}405\,,\quad P(z)=\frac{1}{\pi\sqrt{z}}K_0(\sqrt{z}) \quad (6)$$

* The values of ϱ_n for this case are 0·785, 0·637, 0·540, 0·473, 0·424, 0·385 for $n=1,2,\dots 6$ respectively.
† The values of ϱ_n for this case are 0·637, 0·516, 0·438, 0·383, 0·343, 0·312 for $n=1,2,\dots 6$ respectively.

where we have used equation (16) on p. 684 of GR.

Discrepancy indices R_B and R_2

By definition, for a *complete* model [see equations (1) and (2) of Parthasarathy & Parthasarathi, (1972)]

$$R_B = \sum_{hkl}(|F_N|-|F_N^C|)^2/\sum_{hkl}|F_N|^2 \quad (11)$$

$$R_2 = \sum_{hkl}(I_N - I_N^C)^2/\sum_{hkl} I_N^2 \,. \quad (12)$$

In terms of the normalized variables $y_N = |F_N|/\langle|F_N|^2\rangle^{1/2}$, $y_N^C = |F_N^C|/\langle|F_N^C|^2\rangle^{1/2}$, $z_N = y_N^2$ and $z_N^C = y_N^{C2}$ equations (11) and (12) become

$$R_B = \langle(y_N - y_N^C)^2\rangle/\langle y_N^2\rangle \quad (13)$$

$$R_2 = \langle(z_N - z_N^C)^2\rangle/\langle z_N^2\rangle \,. \quad (14)$$

Since we are considering only an *unrelated* model here, the variables z_N and z_N^C are mutually independent and so are the variables y_N and y_N^C. Hence we have

$$\langle y_N y_N^C\rangle = \langle y_N\rangle\langle y_N^C\rangle \quad \text{and} \quad \langle z_N z_N^C\rangle = \langle z_N\rangle\langle z_N^C\rangle \,. \quad (15)$$

Since both the trial model and the true structure contain the same number and types of atoms with the same type of pseudosymmetry it follows that $\langle y_N\rangle = \langle y_N^C\rangle$ and $\langle z_N^2\rangle = \langle z_N^{C2}\rangle$. Further $\langle z_N\rangle = \langle z_N^C\rangle = 1$. We can therefore rewrite (15) as

$$\langle y_N y_N^C\rangle = \langle y_N\rangle^2 = \varrho \quad \text{and} \quad \langle z_N z_N^C\rangle = 1 \quad (16)$$

where ϱ is the test ratio of Wilson, [1949; see also equation (7*a*) of Parthasarathy, 1966*b*]. Making use of (16) we can show from (13) and (14) that

$$R_B = 2[1 - \langle y_N\rangle^2] = 2[1-\varrho] \quad (17)$$

$$R_2 = 2[\langle z_N^2\rangle - 1]/\langle z_N^2\rangle \,. \quad (18)$$

The values of ϱ for structures with different types of pseudosymmetry considered in RW are given in Table 1. Making use of this in (17) the values of R_B for the various cases can be obtained. The *m*th order moments of z for the various cases are obtained in equations (1), (3), (7), (8), (9) and (10) above and from these the values of $\langle z^2\rangle$ for any given case can be deduced. With the use of the values of $\langle z^2\rangle$ thus obtained in (18), R_2 can be evaluated. The values of R_B and R_2 thus evaluated are summarized in Table 2.

Table 2. *Values of the higher moments of the normalized intensity z and the maximum probable values of the discrepancy indices R_B and R_2 for the pseudosymmetric structures of Table* 1

Hypercentro symmetry and hyperparallelism

n	0	1	2	3	4	5	6
$\langle z^2\rangle_n$	2	3	4·5	6·75	10·12	15·18	22·78
$\langle z^3\rangle_n$	6	15	37·5	93·75	234·4	585·9	1465
$\langle z^4\rangle_n$	24	105	459·4	2010	8793	48468	268298
R_B	0·429	0·727	0·968	1·163	1·322	1·450	1·554
R_2	1·000	1·333	1·556	1·704	1·802	1·868	1·912

Regular parallel repetition in line: *NC* Motif

n	1	2	3	4	5	6
$\langle z^2\rangle_n$	2	3	4·22	5·50	6·80	8·11
$\langle z^3\rangle_n$	6	15	31·33	54·38	84·05	120·3
$\langle z^4\rangle_n$	24	105	328·0	758·6	1466	2518
R_B	0·429	0·727	0·920	1·054	1·152	1·230
R_2	1·000	1·333	1·526	1·636	1·706	1·753

Regular parallel repetition in line: *C* motif

n	1	2	3	4	5	6
$\langle z^2\rangle_n$	3	4·5	6·33	8·25	10·20	12·17
$\langle z^3\rangle_n$	15	37·5	78·33	135·9	210·1	300·8
$\langle z^4\rangle_n$	105	459·4	1435	3319	6412	11015
R_B	0·727	0·968	1·124	1·234	1·314	1·376
R_2	1·333	1·556	1·684	1·758	1·804	1·836

Many parallel repetitions at random

	NC motif in a *NC* arrangement	*C* motif in a *NC* arrangement	*C* motif in a *C* arrangement
$\langle z^2\rangle$	4	6	9
$\langle z^3\rangle$	36	90	225
$\langle z^4\rangle$	576	2520	11025
R_B	0·766	1·000	1·190
R_2	1·500	1·667	1·778

Discussion of the results

The values of the second, third and fourth-order moments of z for the various cases are summarized in Table 2. It is seen that the values of moments for the various pseudosymmetric structures considered here are very much greater than those for even the centric distributions. Structures for which the moments are found to have very high values compared to those of the centric distribution must therefore be expected to possess some type of pseudosymmetry in the atomic distribution. It may be pointed out here that, besides statistical criteria, chemical information and inspection of the Patterson map could also yield valuable information in demonstrating the existence and the extent of pseudosymmetry (see RW).

The maximum probable values of R_B and R_2 for the various cases are given in Table 2. From the work of Parthasarathy & Parthasarathi (1972), it can be seen that the maximum probable value of R_B for a completely 'wrong' structure is 0·429 for the acentric distribution and 0·727 for the centric distribution while the values of R_2 for the corresponding cases are 1·0 and 1·333 respectively. Thus structures with an acentric distribution having values* of $R_B \simeq 0{\cdot}25$ and $R_2 \simeq 0{\cdot}6$ might be expected to be essentially correct while the corresponding values for structures with a centric distribution could be taken as $R_B \simeq 0{\cdot}44$ and $R_2 \simeq 0{\cdot}8$. It may be seen from Table 2 that for structures with pseudosymmetric distributions, the maximum probable values of these indices are in general larger than those expected for the centric and acentric distributions, as the case may be. Thus, for structures with pseudosymmetry, trial structures with values of R_B and R_2 slightly greater than the respective values given above for the centric and acentric distributions, as the case may be, might be significant.† For example, a trial structure with a hypercentric intensity distribution having a value of R_B as high as, say, 0·58 and a value of R_2 as high as, say, 0·94 might be expected to refine, though the final values of these indices may not drop to very low values. This trend is similar to that expected for R from qualitative arguments (see RW).

One of the authors (V.P.) thanks the Council of Scientific and Industrial Research, New Delhi, India for financial support.

* These values are nearly 0·6 times the corresponding maximum probable values. See p.586 of Buerger (1960) for a similar criterion on the value of R for structures with acentric and centric distributions.

† The relevant value that could be used as a criterion for the essentially correct nature of the model structure with any given type of pseudosymmetry considered in RW could be taken to be about 0·6 times the corresponding maximum probable value obtained in Table 2. This criterion, though arbitrary, would suffice in practice.

References

Booth, A. D. (1945). *Phil. Mag.* **36**, 609–615.
Buerger, M. J. (1960). *Crystal Structure Analysis.* New York: John Wiley.
Douglas, A. S. & Woolfson, M. M. (1954). *Acta Cryst.* **7**, 517.
Gradshteyn, I. S. & Ryzhik, I. M. (1965). *Tables of Integrals, Series and Products.* New York & London: Academic Press.
Lipson, H. & Woolfson, M. M. (1952). *Acta Cryst.* **5**, 680–682.
Parthasarathy, S. (1966*a*). *Acta Cryst.* **21**, 642–647.
Parthasarathy, S. (1966*b*). *Z. Kristallogr.* **123**, 27–50.
Parthasarathy, S. & Parthasarathi, V. (1972). *Acta Cryst.* A**28**, 426–432.
Rogers, D. & Wilson, A. J. C. (1953). *Acta Cryst.* **6**, 439–449.
Whittaker, E. T. & Watson, G. N. (1952). *A Course of Modern Analysis.* Cambridge Univ. Press.
Wilson, A. J. C. (1949). *Acta Cryst.* **2**, 318–321.
Wilson, A. J. C. (1969). *Acta Cryst.* B**25**, 1288–1293.