

peut être tenté de le faire, généralisant ainsi rapidement le cas tridimensionnel...) au volume total d'un empilement de n -sphères, la somme des volumes de ces dernières, provoque une erreur importante (plusieurs ordre de grandeur) même pour des valeurs encore faibles de n .

Je remercie vivement MM. J. Berthou, H. Lifchitz, A. Rimsky, A. Soulard et Melle E. Volpe pour leurs interventions, toujours appréciées, au cours de ce travail.

Acta Cryst. (1976). A32, 53

Effects of Misplaced Atoms on the Residual R_2 in Space Groups of Higher Symmetry

BY A. J. C. WILSON

Department of Physics, University of Birmingham, Birmingham B15 2TT, England

(Received 27 May 1975; accepted 11 June 1975)

In space groups of symmetry higher than $P1$ misplacement of a single atom implies the misplacement of all those in the same Wyckoff position, and hence the effect on the residual is larger than in $P1$. 'Exact' values of the residual R_2 are obtained, with explicit allowance for dispersion.

1. Introduction

Wilson (1950) considered the values obtained for what was then called the reliability index when none of the atoms was in its right position. Because of the different degrees of dispersion of the intensity distribution functions different values were obtained for space groups $P1$ and $P\bar{1}$. Later Wilson (1969) obtained expressions for the value of three residuals when only one atom and those related to it by symmetry were misplaced. Second approximations were obtained for the space groups $P1$ and $P\bar{1}$, and a first approximation for the space group $P2/m$. He speculated that space groups of higher symmetry would lead to results of the same type, but with larger numerical coefficients. Recently there has been renewed interest in the subject (*e.g.* Parthasarathy & Parthasarathi, 1972; Lenstra, 1974 and private communications; Wilson, 1974*b*; Parthasarathy, 1975). Lenstra, in particular, has approached the subject in a new way, from the coincidence of vectors in the Patterson map, instead of basing his calculations on the intensity distribution functions. The present paper extends the work systematically to space groups of higher symmetry; the approach is through the expressions for the structure factors in *International Tables for X-ray Crystallography* (1952).

In general, the speculations of Wilson (1969) about space groups of higher symmetry are confirmed. In particular, if there are several non-equivalent misplaced

Références

- BASU, J. P. & ODELL, P. L. (1974). *Pattern Recognition*, Vol. 6, p. 13–16. Oxford: Pergamon Press.
BOUCHER, M. (1917). *Introduction à la Géométrie à Quatre Dimensions d'après les Méthodes de la Géométrie Élémentaire*, p. 147. Paris: Librairie scientifique, A. Hermann et fils.
HUANG, K. (1963). *Statistical Mechanics*, p. 152. New York: John Wiley.
SAYRE, D. (1952). *Acta Cryst.* 5, 60–65.
SAYRE, D. (1971). *Acta Cryst.* A28, 210–212.

atoms R_2 is given by

$$(1)R_2 = 2\Sigma^{-1} \sum_i p_i f_i^2 \quad (1)$$

for non-centrosymmetric space groups, and by

$$(\bar{1})R_2 = 8(3\Sigma)^{-1} \sum_i p_i f_i^2 \quad (2)$$

for centrosymmetric space groups, where Σ is the sum of the squares of the moduli of the scattering factors of all the atoms, equal to the mean intensity of reflexion (Wilson, 1942), p_i and f_i are respectively the multiplicity and the modulus of the scattering factor of the i th misplaced atom, and the summation is over all misplaced atoms in the asymmetric unit. These expressions are in each case first approximations, valid when the scattering power of the misplaced atoms is small compared with the total scattering power of the unit cell; 'exact' expressions, involving also sums of fourth powers of the scattering factors and making allowance for dispersion, are given in (20) and §3.

A recent paper by Parthasarathi & Parthasarathy (1975) appears at first sight to be dealing with the same questions, but there are two important differences. First, they use a scaling factor that makes the average intensity for a partial structure equal to the average value for the full structure when some of the atoms are omitted because, for example, their positions are unknown. (The relation of this procedure to the scaling

of intensities is touched on in §4.) Secondly, they include the fourth-power terms for some atoms (the 'heavy' atoms in their application) and omit them for the rest (the 'light' atoms). This procedure is justifiable when there are many atoms in the second group and few in the first, but in effect they have set themselves a different problem, and the solution is correspondingly different.

Notation

The notation of this paper is similar to that of the author's previous work, but a few changes are introduced to avoid multiplicity of subscripts. In order to include dispersion bold-face type is used for complex scattering factors and signed or complex structure factors. The sum of the squares of the moduli of the scattering factors is denoted by Σ , as before, but the sum of the squares of the complex scattering factors is denoted by S . It should be noted that the modulus of S is equal to Σ only if all atoms have the same phase angle. To a first approximation

$$S \equiv |S| \doteq \Sigma [1 - 2\sigma^2(\delta)], \quad (3)$$

where $\sigma(\delta)$ is the standard deviation (square root of the variance) of the phase angle (δ) of the scattering factors. Also, the sum of the fourth powers of the moduli of the scattering factors is represented by T instead of by Σ_4 .

2. General calculation

Suppose that there are two structures of the same symmetry, in which there is a common part with structure factor G , and a non-common part with structure factors H_1 and H_2 respectively. The total structure factor of the first structure is then

$$F_1 = G + H_1, \quad (4)$$

and the corresponding intensity is

$$I_1 = F_1 F_1^* = G^2 + G H_1^* + G^* H_1 + H_1^2. \quad (5)$$

The first and fourth terms are always positive, but the second and third have randomly varying phases or signs for different reflexions. The mean value of I_1 , averaged over a range of general reflexions hkl , is thus

$$\langle I_1 \rangle = \langle G^2 \rangle + \langle H_1^2 \rangle + \sim 0 \quad (6)$$

$$= \Sigma_0 + \Sigma_1, \quad (7)$$

where Σ_0 and Σ_1 are the sums of the squares of the moduli of the scattering factors of the common and non-common parts of the structure. The value and mean value of I_2 , the intensity of reflexion for the second structure, are given by analogous expressions with 2 replacing 1. The difference in the intensities for the two structures is

$$I_1 - I_2 = H_1^2 - H_2^2 + (H_1 - H_2)G^* + (H_1^* - H_2^*)G, \quad (8)$$

so that

$$\begin{aligned} (I_1 - I_2)^2 &= H_1^4 - 2H_1^2 H_2^2 + H_2^4 \\ &+ (H_1^2 - 2H_1 H_2 + H_2^2) (G^*)^2 \\ &+ [(H_1^*)^2 - 2H_1^* H_2^* + (H_2^*)^2] G^2 \\ &+ 2(H_1^2 - H_1 H_2^* - H_1^* H_2 + H_2^2) G^2 \\ &+ \text{terms averaging to zero.} \end{aligned} \quad (9)$$

The mean value of terms like G^2 is the corresponding Σ . The mean value of terms like G^2 depends on whether or not the structure is centrosymmetric. If centrosymmetric, an argument similar to that of Wilson (1942) shows that the average is S , the sum of the squares of the complex scattering factors f^2 , whereas if it is non-centrosymmetric the average is zero, so that

$$\langle F^2 \rangle = \Sigma, \quad (10)$$

$$\langle F^2 \rangle = (k-1)S, \quad (11)$$

where k has the value 1 for the general reflexions of non-centrosymmetric space groups and the value 2 for centrosymmetric. (For zones and rows of reflexions the symmetry of the corresponding projection determines the value of k .) In (10) and (11) F represents F_1, F_2, G, H_1, H_2 as required, suitable subscripts being used to distinguish the corresponding values of Σ and S .

The mean value of the fourth powers of structure factors has been discussed by, for example, Foster & Hargreaves (1963) and Wilson (1951, 1976). In the present notation

$$\langle F^4 \rangle = 2\Sigma^2 + (k-1)S^2 - \sum_i [(1+k)p_i^2 - q_i] f_i^4, \quad (12)$$

where p_i is the multiplicity of the Wyckoff position occupied by the i th atom, q_i is an integer depending on the space group and the Wyckoff position (actually the mean value of the fourth power of the trigonometrical structure factor for the Wyckoff position in question), and the summation is over the atoms of the asymmetric unit. Foster & Hargreaves point out that the space groups of the triclinic, monoclinic and orthorhombic systems can be arranged in seven categories for which the values of p and q are the same (within trivial factors depending on the lattice; $Fdd2$ and $Fddd$ do not fit in with the scheme). The fundamental classification is presumably the point group, though $\bar{1}$, 2 and m all have $p=2$ and $q=6$ for the general position. Equality of the q 's holds also within each crystal class of the trigonal system, but there are many exceptions in the tetragonal, hexagonal (*sensu stricto*) and cubic systems. Empirically, one might generalize to say that if the space group contains no single symmetry element of multiplicity greater than three, the value of q depends only on the point group, whereas if there is a symmetry element of multiplicity of four or six (such as d in $Fdd2$ and $Fddd$ or 4_1 in $P4_1$) a special calculation is necessary. The values of q have been tabulated for the general positions of many space groups by Wilson (1975b); for those common to the two listings they are equivalent

to the coefficient of $S(4)$ in Table 1 of Foster & Hargreaves. For the space groups $P1$ ($p=q=1$) and $P\bar{1}$ ($p=2, q=6$), (12) reduces to the familiar forms (Wilson, 1951)

$$(1)\langle I^2 \rangle = 2\Sigma^2 - T \quad (13)$$

and

$$(\bar{1})\langle I^2 \rangle = 2\Sigma^2 + S^2 - 3T, \quad (14)$$

$$= 3\Sigma^2 - 3T \quad (15)$$

in the absence of dispersion.

By the use of (10), (11) and (12) the average value of the square of the difference of the intensities from the two structures can be found. From (9), (10) and (11)

$$\begin{aligned} \langle (I_1 - I_2)^2 \rangle &= \langle H_1^4 \rangle - 2\Sigma_1\Sigma_2 + \langle H_2^4 \rangle \\ &+ (k-1)\mathbf{S}_0^*(\mathbf{S}_1 + \mathbf{S}_2) \\ &+ (k-1)\mathbf{S}_0(\mathbf{S}_1^* + \mathbf{S}_2^*) \\ &+ 2\Sigma_0(\Sigma_1 + \Sigma_2), \end{aligned} \quad (16)$$

where the subscript 0 refers to the atoms common to both structures, 1 to the atoms unique to the first structure, and 2 to the atoms unique to the second structure. Symbols without subscript will refer to the complete structure 1, so that, for example,

$$\Sigma = \Sigma_0 + \Sigma_1. \quad (17)$$

From (12) the residual becomes

$$R_2 = \langle (I_1 - I_2)^2 \rangle / \langle I_1^2 \rangle \quad (18)$$

$$\begin{aligned} &= \{2\Sigma_0(\Sigma_1 + \Sigma_2) + 2(\Sigma_1^2 - \Sigma_1\Sigma_2 + \Sigma_2^2) + (k-1)(S_1^2 \\ &+ S_2^2) + (k-1)[(\mathbf{S}_1 + \mathbf{S}_2)\mathbf{S}_0^* + (\mathbf{S}_1^* + \mathbf{S}_2^*)\mathbf{S}_0] \\ &- \sum_i^{(1+2)} [(1+k)p_i^2 - q_i]f_i^4\} \\ &\div \{2\Sigma^2 + (k-1)S^2 - \sum_i [(1+k)p_i^2 - q_i]f_i^4\} \end{aligned} \quad (19)$$

$$\begin{aligned} &= \{2\Sigma(\Sigma_1 + \Sigma_2) - 2\Sigma_2(2\Sigma_1 - \Sigma_2) + (k-1)(S_1^2 + S_2^2) \\ &+ (k-1)[(\mathbf{S}_1 + \mathbf{S}_2)\mathbf{S}^* + (\mathbf{S}_1^* + \mathbf{S}_2^*)\mathbf{S} - (2S_1^2 + S_1S_2^* \\ &+ \mathbf{S}_1^*\mathbf{S}_2)] - \sum_i^{(1+2)} [(1+k)p_i^2 - q_i]f_i^4\} \\ &\div \{2\Sigma^2 + (k-1)S^2 - \sum_i [(1+k)p_i^2 - q_i]f_i^4\} \end{aligned} \quad (20)$$

from (17). If only a few atoms are misplaced the important terms are those containing Σ , \mathbf{S} and S ; the residual is then

$$\begin{aligned} R_2 &= \{2\Sigma(\Sigma_1 + \Sigma_2) + (k-1)[(\mathbf{S}_1 + \mathbf{S}_2)\mathbf{S}^* + (\mathbf{S}_1^* + \mathbf{S}_2^*)\mathbf{S}]\} \\ &\div \{2\Sigma^2 + (k-1)S^2\}, \end{aligned} \quad (21)$$

which reduces to (1) and (2) when dispersion is neglected.

3. Special cases

The general expression for R_2 can be simplified in various ways, depending on the assumptions that can plausibly be made about the atomic positions and about the way in which the calculations are handled if, for example, I_1 relates to the observed intensities and I_2 to

the calculated. Those considered worth special mention are:

- (i) All the atoms occupy the same Wyckoff position.
- (ii) There is no appreciable dispersion.
- (iii) The structures 1 and 2 are different arrangements of the same atoms.
- (iv) The structures have no common part, so that Σ_0 and T_0 are zero.
- (v) No assumption is made about the positions of the atoms peculiar to structure 2, and Σ_2 and T_2 are put equal to zero.

It is, of course, possible for two or more of these possibilities to occur together; there is, for example, nothing incompatible in any combination of the first four.

3.1 All atoms in the same Wyckoff position

If all atoms are in the general position, or in the same special position with different parameters, p and q have the same value for all atoms, and can be taken outside the summation signs. Equation (20) then reduces to

$$\begin{aligned} R_2 &= \{2\Sigma(\Sigma_1 + \Sigma_2) - 2\Sigma_2(2\Sigma_1 - \Sigma_2) + (k-1)[S_1^2 + S_2^2] \\ &+ (k-1)[(\mathbf{S}_1 + \mathbf{S}_2)\mathbf{S}^* + (\mathbf{S}_1^* + \mathbf{S}_2^*)\mathbf{S} \\ &- (2S_1^2 + S_1S_2^* + \mathbf{S}_1^*\mathbf{S}_2)] - [(1+k)p - q/p](T_1 + T_2)\} \\ &\div \{2\Sigma^2 + (k-1)S^2 - [(1+k)p - q/p]T\}. \end{aligned} \quad (22)$$

The situation of having all atoms in the general position is very common among organic structures, but much less common among inorganic and metallic structures.

3.2 Dispersion negligible

If there is no appreciable dispersion it is unnecessary to distinguish between \mathbf{S} , \mathbf{S}^* , and Σ , so that (20) becomes

$$\begin{aligned} R_2 &= \{2k\Sigma(\Sigma_1 + \Sigma_2) - (k-1)\Sigma_1^2 - 2(k+1)\Sigma_1\Sigma_2 \\ &+ (k+1)\Sigma_2^2 - \sum_i^{(1+2)} [(1+k)p_i^2 - q_i]f_i^4\} \\ &\div \{(1+k)\Sigma^2 - \sum_i [(1+k)p_i^2 - q_i]f_i^4\}, \end{aligned} \quad (23)$$

or, if all the atoms have the same Wyckoff position,

$$\begin{aligned} R_2 &= \{2k\Sigma(\Sigma_1 + \Sigma_2) - (k-1)\Sigma_1^2 - 2(k+1)\Sigma_1\Sigma_2 \\ &+ (k+1)\Sigma_2^2 - [(1+k)p - q/p](T_1 + T_2)\} \\ &\div \{(1+k)\Sigma^2 - [(1+k)p - q/p]T\}. \end{aligned} \quad (24)$$

3.3 Different arrangements of the same atoms

If structures 1 and 2 consist of the same numbers and kinds of atoms in different arrangements the values of Σ_1 and Σ_2 ; \mathbf{S}_1 and \mathbf{S}_2 ; and T_1 and T_2 are equal, so that (20) becomes

$$\begin{aligned} R_2 &= \{4\Sigma\Sigma_1 - 2\Sigma_1^2 + 2(k-1)(\mathbf{S}_1\mathbf{S}^* + \mathbf{S}_1^*\mathbf{S} - S_1^2) \\ &- \sum_i^{(1+2)} [(1+k)p_i^2 - q_i]f_i^4\} \\ &\div \{2\Sigma^2 + (k-1)S^2 - \sum_i [(1+k)p_i^2 - q_i]f_i^4\}. \end{aligned} \quad (25)$$

If, in addition, the atoms all occupy the same Wyckoff position,

$$R_2 = \{4\Sigma\Sigma_1 - 2\Sigma_1^2 + 2(k-1)(S_1S^* + S_1^*S - S_1^2) - 2[(1+k)p - q/p]T_1\} \div \{2\Sigma^2 + (k-1)S^2 - [(1+k)p - q/p]T\}, \quad (26)$$

with a further obvious simplification if dispersion can be neglected.

3.4 No common part

If the structures have no common part Σ_0 , S_0 and S_0 are zero and Σ , S and S are equal to Σ_1 , S_1 and S_1 . The expression for R_2 is then given by (19) with the terms in Σ_0 , S_0 and S_0 omitted and subscripts 1 added in the denominator. If, in addition, the structures are different arrangements of the same atoms, the expression reduces to

$$R_2 = \{2\Sigma_1^2 + 2(k-1)S_1^2 - 2 \sum_i [(1+k)p_i^2 - q_i]f_i^4\} \div \{2\Sigma_1^2 + (k-1)S_1^2 - \sum_i [(1+k)p_i^2 - q_i]f_i^4\}. \quad (27)$$

If, further, all atoms are in the same Wyckoff position,

$$R_2 = \{2\Sigma_1^2 + 2(k-1)S_1^2 - 2[(1+k)p - q/p]T_1\} \div \{2\Sigma_1^2 + (k-1)S_1^2 - [(1+k)p - q/p]T_1\}. \quad (28)$$

Wilson (1974, 1975a) has given the values of R_2 for this case for the space groups $P1$ and $P\bar{1}$ (for which p , q and k are 1, 1, 1 and 2, 6, 2 respectively), not taking dispersion into account. Equation (28) correctly reduces to his equations (12) and (14) after the appropriate substitutions.

3.5 Atoms peculiar to second structure omitted

It may be thought advisable to leave the atoms peculiar to one structure out of consideration if, for example, I_1 relates to the observed intensities, and at the current stage of refinement the positions of some atoms are unknown. The calculated intensities I_2 will then relate to only the common part of the two structures, and Σ_2 , S_2 , S_2 and T_2 will be zero. Equation (20) becomes

$$R_2 = \{2\Sigma\Sigma_1 + (k-1)(S_1S^* + S_1^*S - S_1^2) - \sum_i (1+k)p_i^2 - q_i]f_i^4\} \div \{2\Sigma^2 + (k-1)S^2 - \sum_i [(1+k)p_i^2 - q_i]f_i^4\}. \quad (29)$$

This is approximately half the value of R_2 given by (25), in which the atoms peculiar to the second structure are put in the wrong positions. From the point of view of reducing the residual it is better to omit than to misplace – a conclusion in accordance with common sense.

4. Problems of scaling

In the preceding discussions it has been assumed that G , H_1 and H_2 and the corresponding intensities are all on the same 'absolute' scale. In practice this will not always be the case. If I_1 is the observed intensity and I_2 the calculated, the scaling factor implicit in I_2 should be adjusted at each stage of refinement so that the average values are equal – this has been discussed in more detail elsewhere (Wilson, 1974a; Lomer & Wilson, 1975). In the type of problem discussed here, however, for which some atoms may not be included in the calculation of I_2 , the best procedure will probably be to adjust the scaling so that

$$\langle I_1 \rangle = \langle G^2 \rangle + \Sigma_2; \quad (27)$$

one thus takes into account the actual average of the calculated intensity for the atoms whose positions are supposed known, and the average to be expected on statistical grounds for the atoms whose positions are unknown. The procedure of Parthasarathi & Parthasarathy (1975), following Srinivasan & Ramachandran (1965), goes some way towards this, but is not identical with it. Refining the scaling factor by least squares would seem to be unjustifiable, since the value obtained is known to be biased by 'defects in the model', which missing or misplaced atoms certainly are. It would be easy, but tedious, to elaborate on the consequences of doing so. Some extreme examples are given by Wilson (1974b).

References

- FOSTER, F. & HARGREAVES, A. (1963). *Acta Cryst.* **16**, 1133–1139.
- International Tables for X-ray Crystallography* (1952). Vol. I. Birmingham: Kynoch Press.
- LENSTRA, A. T. H. (1974). *Acta Cryst.* **A30**, 363–369.
- LOMER, T. R. & WILSON, A. J. C. (1975). *Acta Cryst.* **B31**, 646–647.
- PARTHASARATHI, V. & PARTHASARATHY, S. (1975). *Acta Cryst.* **A31**, 38–41.
- PARTHASARATHY, S. (1975). *Acta Cryst.* **A31**, 375–380.
- PARTHASARATHY, S. & PARTHASARATHI, V. (1972). *Acta Cryst.* **A28**, 426–432.
- SRINIVASAN, R. & RAMACHANDRAN, G. N. (1965). *Acta Cryst.* **19**, 1003–1007.
- WILSON, A. J. C. (1942). *Nature, Lond.* **150**, 151–152.
- WILSON, A. J. C. (1950). *Acta Cryst.* **3**, 397–398.
- WILSON, A. J. C. (1951). *Research*, **4**, 141–142.
- WILSON, A. J. C. (1969). *Acta Cryst.* **B25**, 1288–1293.
- WILSON, A. J. C. (1974a). Paper given the IUCr Conference, Madrid, 25 April 1974. Published in *Anomalous Scattering* (1975), Edited by S. RAMASESHAN & S. C. ABRAHAMS, pp. 325–332. Copenhagen: Munksgaard.
- WILSON, A. J. C. (1974b). *Acta Cryst.* **A30**, 836–838.
- WILSON, A. J. C. (1976). In preparation.