

Bombay, for allowing them to use CDC-3600 versions of the *ORFLS* and *ORFFE* programs and for many discussions on computation.

References

- BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1962). *ORFLS, A Fortran Crystallographic Least-Squares Program*. USAEC Report ORNL-TM-305.
- BUSING, W. R., MARTIN, K. O. & LEVY, H. A. (1964). *ORFFE, A Fortran Crystallographic Function and Error Program*. USAEC Report ORNL-TM-306.
- CROMER, D. T. & WABER, J. T. (1965). *Acta Cryst.* **18**, 104.
- FRIDRICHSONS, J. & MATHIESON, A. McL. (1962). *Acta Cryst.* **15**, 119.
- FRUEH, A. J. (1953). *Acta Cryst.* **6**, 454.
- GABE, E. J. (1962). *Acta Cryst.* **15**, 759.
- International Tables for X-ray Crystallography* (1962). Vol. III. Birmingham: Kynoch Press.
- JEFFREY, G. A., ROSENSTEIN, R. D. & VLASSE, M. (1967). *Acta Cryst.* **22**, 725.
- KALYANI, V., MANOHAR, H. & MANI, N. V. (1967). *Acta Cryst.* **23**, 272.
- KALYANI, V. & VIJAYAN, M. (1969). *Acta Cryst.* **B25**, 252.
- KANNAN, K. K., VIJAYAN, M. & NAMBU DIRI, M. M. G. (1966). *A Unified Set of Crystallographic Programs*. Bangalore: Indian Institute of Science. Unpublished.
- KARLE, I. & KARLE, J. (1966). *Acta Cryst.* **20**, 555.
- KIM, S. H., JEFFREY, G. A., ROSENSTEIN, R. D. & CORFIELD, P. W. R. (1967). *Acta Cryst.* **22**, 733.
- KITAIGORODSKII, A. I. (1955). *Organic Chemical Crystallography*. New York: Consultants Bureau.
- MANOHAR, H., KALYANI, V., BHAT, M. V. & KAMATH, K. M. (1966). *Tetrahedron Letters*, **44**, 5413.
- PHILLIPS, D. C. (1954). *Acta Cryst.* **7**, 746.
- PRZYBYLSKA, M. & AHMED, F. R. (1958). *Acta Cryst.* **11**, 718.

Acta Cryst. (1969). **B25**, 1288

Effect of a Badly Misplaced Atom on the Residual

BY A. J. C. WILSON

School of Physics, Georgia Institute of Technology, Atlanta, Georgia 30332, U.S.A.
*and Department of Physics, University of Birmingham, P. O. Box 363, Birmingham 15, England**

(Received 24 June 1968 and in revised form 4 September 1968)

The effect of an entirely misplaced atom in $P1$ (or two equivalent atoms in $P\bar{1}$) is to give residuals of approximately the following values:

	$P1$	$P\bar{1}$
$R = \langle F_{\text{wrong}} - F_{\text{right}} \rangle / \langle F \rangle$	$16f/\pi^{5/2} \Sigma^{1/2}$	$8\sqrt{2}f/\pi^{3/2} \Sigma^{1/2}$
$R_1 = \langle I_{\text{wrong}} - I_{\text{right}} \rangle / \langle I \rangle$	$8f/\pi^{3/2} \Sigma^{1/2}$	$32\sqrt{2}f/\pi^{5/2} \Sigma^{1/2}$
$R_2 = \langle (I_{\text{wrong}} - I_{\text{right}})^2 \rangle / \langle I^2 \rangle$	$2f^2/\Sigma$	$16f^2/3\Sigma$

where f is the scattering factor of the misplaced atom and Σ is the sum of the squares of the scattering factors of all the atoms. These values seem too big to be overlooked unless the total number of atoms is very large.

1. Introduction

In connexion with the refinement of one of the structures of the apatite group, the question arose of the size of the residual to be expected if one atom were entirely misplaced, and the following calculations were therefore undertaken. Three disagreement indices or residuals are in common use, R based on the difference of the moduli of the structure factors, R_1 based on the difference of the squares of the moduli, and R_2 based on the square of the difference of the squares. The last, being effectively the variance of $|F|^2$, is considerably simpler to manipulate theoretically than are the other two, and the three residuals are, therefore, treated in reverse order below.

The calculations depend on averaging the difference between quantities calculated with the atom in the wrong place and similar quantities calculated with the atom in the right place. The vector difference between two atomic positions that can be regarded as a complete misplacement is, of course, a function of the Bragg angle, or more accurately of the scattering vector. If the 'right' position of the atom is \mathbf{r}_0 and its assumed position is \mathbf{r} , the difference in phase between their contributions to the reflexion with scattering vector \mathbf{S} is $2\pi\mathbf{S} \cdot (\mathbf{r}_0 - \mathbf{r})$. For the calculations to proceed smoothly it is necessary for this phase difference to change by at least 2π as \mathbf{S} takes on all orientations, so that an atom is entirely misplaced only if its distance from its correct position is larger than the reciprocal of the magnitude of the scattering vector. For the lowest-angle reflexions, therefore, the displacement

* Permanent address.

must be a large fraction of the cell size, but for higher-angle reflexions the necessary displacement becomes progressively smaller. This is in accordance with the common knowledge that accuracy in the determination of atomic parameters is obtained mainly from the high-angle reflexions.

In §§ 2–4 the calculation pictured is that of the residual arising if in an entirely correct structure the coordinates of one atom (and any related to it by crystallographic symmetry) are altered. In the practical determination of a crystal structure there would also be a residual arising from experimental error in the measurements of the structure factors and from defects in the model (such as doubts concerning the proper atomic scattering factors). In § 5 the joint residual arising from a misplaced atom in conjunction with experimental errors and certain other practical matters are discussed. The calculations refer specifically to structures having the space groups $P1$ and $P\bar{1}$, but similar methods could be used for any space group. In general one would expect the same dependence on f and Σ but larger numerical coefficients in space groups of higher symmetry.

2. Space group $P1$

(a) Residual R_2

We require to find the value of

$$R_2 = \langle (I_2 - I_1)^2 \rangle / \langle I^2 \rangle \quad (1)$$

$$= \langle (|F_2|^2 - |F_1|^2)^2 \rangle / \langle |F|^4 \rangle, \quad (2)$$

where the angle brackets indicate averages over all values of \mathbf{S} . Let the structure factor due to all atoms except the misplaced one be $F_0 \exp(i\psi)$, where F_0 is its magnitude and ψ its phase, the correct position of the misplaced atom be \mathbf{r}_0 , and its assumed position \mathbf{r} . Then the correct structure factor is

$$F_1 = F_0 \exp(i\psi) + f \exp(i\theta), \quad (3)$$

where

$$\theta = 2\pi \mathbf{S} \cdot \mathbf{r}_0, \quad (4)$$

but it is taken to be

$$F_2 = F_0 \exp(i\psi) + f \exp(i\varphi), \quad (5)$$

where

$$\varphi = 2\pi \mathbf{S} \cdot \mathbf{r}; \quad (6)$$

f is the scattering factor of the misplaced atom and \mathbf{S} is the scattering vector. From equations (3) and (5)

$$I_1 = F_1 F_1^* = F_0^2 + f^2 + 2fF_0 \cos(\psi - \theta), \quad (7)$$

$$I_2 = F_2 F_2^* = F_0^2 + f^2 + 2fF_0 \cos(\psi - \varphi), \quad (8)$$

so that

$$I_2 - I_1 = 2fF_0 \{ \cos(\psi - \varphi) - \cos(\psi - \theta) \} \quad (9)$$

$$= -4fF_0 \sin(\psi - \frac{1}{2}\theta - \frac{1}{2}\varphi) \sin(\frac{1}{2}\theta - \frac{1}{2}\varphi), \quad (10)$$

$$(I_2 - I_1)^2 = 16f^2 F_0^2 \sin^2(\psi - \frac{1}{2}\theta - \frac{1}{2}\varphi) \sin^2(\frac{1}{2}\theta - \frac{1}{2}\varphi). \quad (11)$$

If the components of F_1 and F_2 are independently distributed in phase, the conditions for which are discussed in § 1, the average of $(I_2 - I_1)^2$ over a large number of values of \mathbf{S} can be obtained by averaging each of the factors in equation (11) separately, giving

$$\begin{aligned} \langle (I_2 - I_1)^2 \rangle &= 16f^2 \cdot \langle F_0^2 \rangle \cdot \frac{1}{2} \cdot \frac{1}{2} \\ &= 4f^2 (\Sigma - f^2), \end{aligned} \quad (12)$$

where Σ is the sum of the squares of the atomic scattering factors (Wilson, 1942). Since the mean-square value of I for a non-centrosymmetric structure is

$$\langle I^2 \rangle = 2\Sigma^2 - \Sigma_4, \quad (13)$$

where Σ_4 is the sum of the fourth powers of the atomic scattering factors (Wilson, 1951), the value of the residual is

$$R_2 = \frac{4f^2 (\Sigma - f^2)}{2\Sigma^2 - \Sigma_4}. \quad (14)$$

If the correction terms are neglected this reduces to

$$R_2 = 2f^2 / \Sigma, \quad (15)$$

and if the structure consists of n atoms of roughly equal scattering power $\Sigma \sim nf^2$, $\Sigma_4 \sim nf^4$, and

$$R_2 \sim 4(n-1)/n(2n-1) \sim 2n^{-1}. \quad (16)$$

(b) Residual R_1

The residual R_1 is given by

$$R_1 = \frac{\langle |I_2 - I_1| \rangle}{\langle I \rangle}. \quad (17)$$

From equation (10)

$$|I_2 - I_1| = 4f|F_0| |\sin(\psi - \frac{1}{2}\theta - \frac{1}{2}\varphi)| |\sin(\frac{1}{2}\theta - \frac{1}{2}\varphi)|, \quad (18)$$

and under the same assumptions of independence of the distributions

$$\langle |F_0| \rangle = \frac{1}{2} \{ \pi (\Sigma - f^2) \}^{1/2} \quad (19)$$

(Wilson, 1949) and

$$\langle |\sin x| \rangle = 2\pi^{-1}. \quad (20)$$

Since the mean value of I is Σ , the residual is

$$\begin{aligned} R_1 &= 4f \cdot \frac{1}{2} \{ \pi (\Sigma - f^2) \}^{1/2} \cdot 2\pi^{-1} \cdot 2\pi^{-1} \cdot \Sigma^{-1} \\ &= \frac{8f (\Sigma - f^2)^{1/2}}{\pi^{3/2} \Sigma} \end{aligned} \quad (21)$$

$$\sim 8f / \pi^{3/2} \Sigma^{1/2} \quad (22)$$

$$\sim \frac{8(n-1)^{1/2}}{\pi^{3/2} n} \quad (23)$$

$$\sim 8\pi^{-3/2} n^{-1/2}. \quad (24)$$

(c) *Residual R*

The values of R_2 and R_1 were obtained without making any new mathematical approximations, but it does not seem possible to obtain

$$R = \frac{\langle ||F_2| - |F_1|| \rangle}{\langle |F| \rangle} \quad (25)$$

in a simple fashion without making a series approximation. For all but a small fraction of the reflexions $F_0^2 + f^2$ is numerically greater than $2fF_0 \cos(\psi - \theta)$, so that $|F_1|$ can be obtained by taking the square root of equation (7) by means of the binomial theorem, giving

$$|F_1| = (F_0^2 + f^2)^{1/2} \left[1 + \frac{F_0 f}{F_0^2 + f^2} \cos(\psi - \theta) - \frac{F_0^2 f^2}{2(F_0^2 + f^2)^2} \cos^2(\psi - \theta) + \dots \right], \quad (26)$$

and similarly for $|F_2|$. The modulus of the difference is thus

$$\begin{aligned} ||F_2| - |F_1|| &= (F_0^2 + f^2)^{1/2} \\ &\times \left| \frac{F_0 f}{F_0^2 + f^2} [\cos(\psi - \varphi) - \cos(\psi - \theta)] \right. \\ &\left. - \frac{F_0^2 f^2}{2(F_0^2 + f^2)^2} [\cos^2(\psi - \varphi) - \cos^2(\psi - \theta)] + \dots \right|. \end{aligned} \quad (27)$$

The second term is never more than a fraction of the first, so that the sign of the quantity within the modulus signs is the same as the sign of the first term, giving

$$\begin{aligned} \langle ||F_2| - |F_1|| \rangle &= \left\langle \frac{2F_0 f}{(F_0^2 + f^2)^{1/2}} \right\rangle \langle |\sin(\psi - \frac{1}{2}\varphi - \frac{1}{2}\theta)| \rangle \\ &\times \langle |\sin(\frac{1}{2}\theta - \frac{1}{2}\varphi)| \rangle - \left\langle \frac{F_0^2 f^2}{2(F_0^2 + f^2)^{3/2}} \right\rangle \\ &\times \langle \sin(2\psi - \varphi - \theta) \rangle \langle \sin(\theta - \varphi) \rangle + \dots \end{aligned} \quad (28)$$

The average value of $F_0 f / (F_0^2 + f^2)^{1/2}$ could be calculated, if necessary, from the distribution function of the structure factors for a non-centrosymmetric structure (Wilson, 1949). However, except for structure factors of the order of f , the ratio has nearly enough the value f . The mean value of the $|\sin x|$ terms is $2/\pi$ each, and the mean value of $|F|$ is given by equation (19), so that

$$R = \frac{16f}{\pi^{5/2} \Sigma^{1/2}} + \text{terms in } f^2 / \Sigma, \quad (29)$$

$$\sim 16f / \pi^{5/2} \Sigma^{1/2} \quad (30)$$

$$\sim 16\pi^{-5/2} n^{-1/2}. \quad (31)$$

Some justification for the neglect of reflexions with small F_0 is contained in § 3(b) below. Centrosym-

metric structures have a higher proportion of reflexions with small F_0 , and it is found that even then the effect of neglecting them is small.

3. Space group $P\bar{1}$ (a) *Residual R₂*

For a centrosymmetric structure there must be two misplaced atoms,* related by the centre of symmetry, and all the structure factors will be real. The phase angle ψ is thus either 0 or π , but it is now convenient to forget about it and simply regard F_0 as positive or negative. Then

$$\left. \begin{aligned} F_1 &= F_0 + 2f \cos \theta, \\ F_2 &= F_0 + 2f \cos \varphi, \\ I_1 &= F_0^2 + 4fF_0 \cos \theta + 4f^2 \cos^2 \theta, \\ I_2 &= F_0^2 + 4fF_0 \cos \varphi + 4f^2 \cos^2 \varphi, \end{aligned} \right\} \quad (32)$$

$$I_2 - I_1 = 8fF_0 \sin \frac{1}{2}(\theta + \varphi) \sin \frac{1}{2}(\theta - \varphi) + 4f^2 \sin(\theta + \varphi) \sin(\theta - \varphi). \quad (33)$$

Squaring and averaging in the same way as before gives

$$\begin{aligned} \langle (I_2 - I_1)^2 \rangle &= 64f^2 (\Sigma - 2f^2) \cdot \frac{1}{2} \cdot \frac{1}{2} + 0 + 16f^4 \cdot \frac{1}{2} \cdot \frac{1}{2} \\ &= 16f^2 \Sigma - 28f^4. \end{aligned} \quad (34)$$

Since the mean-square value of I for a centrosymmetric structure is $3\Sigma^2 - 3\Sigma_4$ (Wilson, 1951) the residual is

$$R_2 = \frac{16f^2 \Sigma - 28f^4}{3\Sigma^2 - 3\Sigma_4} \sim \frac{16f^2}{3\Sigma}, \quad (35)$$

or approximately

$$R_2 = \frac{4(4n-7)}{3n(n-1)} \quad (36)$$

$$\sim 16/3n. \quad (37)$$

(b) *Residual R₁*

Equation (33) can be written

$$I_2 - I_1 = 8f \sin \frac{1}{2}(\theta + \varphi) \sin \frac{1}{2}(\theta - \varphi) \times [F_0 + 2f \cos \frac{1}{2}(\theta + \varphi) \cos \frac{1}{2}(\theta - \varphi)], \quad (38)$$

so that

$$|I_2 - I_1| = 8f |\sin \frac{1}{2}(\theta + \varphi)| |\sin \frac{1}{2}(\theta - \varphi)| |F_0 + k| \quad (39)$$

where

$$k = 2f \cos \frac{1}{2}(\theta + \varphi) \cos \frac{1}{2}(\theta - \varphi). \quad (40)$$

The quantity within the final modulus signs in equation (39) is actually positive if

$$F_0 > -k \quad (41)$$

and negative if

$$F_0 < -k. \quad (42)$$

* It is, of course, possible that a single atom might be misplaced on the wrong centre of symmetry, but the possibility hardly deserves a detailed calculation.

Averaging over all possible values of F_0 thus gives

$$\langle |F_0 + k| \rangle = - \int_{-\infty}^{-k} (F_0 + k) P(F_0) dF_0 + \int_{-k}^{\infty} (F_0 + k) P(F_0) dF_0 \quad (43)$$

where $P(F_0)dF_0$ is the probability distribution of F_0 . For a centrosymmetric structure this is (Wilson, 1949)

$$P(F_0)dF_0 = (2\pi \Sigma_0)^{-1/2} \exp(-F_0^2/2 \Sigma_0) dF_0 \quad (44)$$

where

$$\Sigma_0 = \Sigma - 2f^2. \quad (45)$$

The reduction of equation (43) follows slightly different routes, depending on whether k is actually positive or negative, but in either case it becomes

$$\langle |F_0 + k| \rangle = (2\pi \Sigma_0)^{-1/2} \left[2 \int_0^{\infty} F_0 \exp(-F_0^2/2 \Sigma_0) dF_0 + 2 \int_0^{|k|} (|k| - F_0) \exp(-F_0^2/2 \Sigma_0) dF_0 \right] \quad (46)$$

after a little rearrangement. The first integral is just the mean value of $|F_0|$, $(2 \Sigma_0/\pi)^{1/2}$. In the second integral $|k|$ is small compared with $\Sigma_0^{1/2}$, so that the exponential may be expanded in a power series, giving

$$\langle |F_0 + k| \rangle = (2 \Sigma_0/\pi)^{1/2} \left[1 + \frac{k^2}{2 \Sigma_0} - \frac{k^4}{24 \Sigma_0^2} + \dots \right]. \quad (47)$$

Equation (39) now has to be averaged over the angles θ and φ . Inserting the average over F_0 from equation (47) and writing k in full gives

$$\langle |I_2 - I_1| \rangle = 8f(2 \Sigma_0/\pi)^{1/2} \left\langle \left| \sin \frac{1}{2}(\theta + \varphi) \right| \left| \sin \frac{1}{2}(\theta - \varphi) \right| + \frac{2f^2}{\Sigma_0} \left| \cos^2 \frac{1}{2}(\theta + \varphi) \sin \frac{1}{2}(\theta + \varphi) \right| \left| \cos^2 \frac{1}{2}(\theta - \varphi) \right| \times \left| \sin \frac{1}{2}(\theta - \varphi) \right| + \dots \right\rangle \quad (48)$$

$$= \frac{32\sqrt{2}f \Sigma_0^{1/2}}{\pi^{5/2}} \left[1 + \frac{2f^2}{9 \Sigma_0} + \dots \right]. \quad (49)$$

The residual is thus

$$R_1 = \frac{32\sqrt{2}f (\Sigma - 2f^2)^{1/2}}{\pi^{5/2} \Sigma} \left[1 + \frac{2f^2}{9(\Sigma - 2f^2)} + \dots \right] \quad (50)$$

or approximately

$$R_1 = \frac{32\sqrt{2}f}{\pi^{5/2} \Sigma^{1/2}} \quad (51)$$

$$\sim \frac{32\sqrt{2}}{\pi^{5/2} n^{1/2}}. \quad (52)$$

It should be noted that if the effect of reflexions with $F_0 \sim f$ had been neglected, as was done in paragraph 1(c), only the first term of the series (50) would have been obtained. The correction arising from taking into account these weak reflexions is ordinarily negligible, amounting to 1% for a structure with about twenty atoms.

(c) Residual R

As in paragraph 1(c), for most reflexions $|F_0| \gg f$, and $|F_1|$ and $|F_2|$ can be expressed as a series by taking the square roots of equations (32) for I_1 and I_2 . Expansion by the binomial theorem gives

$$|F_1| = |F_0| + 2fF_0 \cos \theta / |F_0| + 0 + \dots, \quad (53)$$

$$|F_2| = |F_0| + 2fF_0 \cos \varphi / |F_0| + 0 + \dots, \quad (54)$$

the terms in f^2 cancelling. Then

$$\begin{aligned} \langle ||F_2| - |F_1|| \rangle &= \langle 2f| \cos \varphi - \cos \theta | + \dots \rangle \\ &= 4f \langle | \sin \frac{1}{2}(\theta + \varphi) | | \sin \frac{1}{2}(\theta - \varphi) | + \dots \rangle \\ &= \frac{16f}{\pi^2} + \dots \end{aligned} \quad (55)$$

The mean value of $|F|$ for a centrosymmetric structure is $(2 \Sigma/\pi)^{1/2}$ (Wilson, 1949), so that

$$R = \frac{8\sqrt{2}f}{\pi^{3/2} \Sigma} + \dots \quad (57)$$

$$\sim 8\sqrt{2}\pi^{-3/2} n^{-1/2}. \quad (58)$$

4. Other space groups

There is no difficulty in principle in carrying out similar calculations for space groups of higher symmetry, but the complexity increases with the multiplicity of the general position. The numerical values of the coefficients in R and R_1 for the centrosymmetric space group $P\bar{1}$ are about twice as great as for the non-centrosymmetric $P1$, so that for other space groups one might hazard the guess that

$$R \sim R_1 \sim pf/\Sigma^{1/2}, \quad (59)$$

where p is the multiplicity of the general position. For R_2 the increase in going from $P1$ to $P\bar{1}$ is a factor of $8/3$, so one might guess that

$$R_2 \sim 2\frac{1}{2} pf^2/\Sigma. \quad (60)$$

For the space group $P2/m$ a preliminary calculation gives

$$R_2 \sim 32f^2/3\Sigma, \quad (61)$$

whereas equation (60) would give the numerical coefficient 10.

5. Experimental errors and some other practical considerations

In the practical determination of a crystal structure the quantities corresponding to F_1 and I_1 will not be

calculated from the correct atomic positions, but will be obtained from the observed intensities, and will thus be subject to experimental error. The model of the crystal structure will be imperfect; the temperature factors may be anharmonic; the atomic scattering factors may be aspherical or vary incorrectly with the length of the scattering vector; there may be vacant atomic positions. One can thus write for the three residuals

$$R_2'' = \frac{\langle (I_2 - I_1 + \delta)^2 \rangle}{\langle I^2 \rangle}, \quad (62)$$

$$R_1'' = \frac{\langle |I_2 - I_1 + \delta| \rangle}{\langle I \rangle}, \quad (63)$$

$$R'' = \frac{\langle ||F_2| - |F_1| + \varepsilon| \rangle}{\langle |F| \rangle}, \quad (64)$$

where δ and ε are the corrections required to make the calculated values of I_1 and $|F_1|$ agree with those observed. Provided that there is no correlation between the corrections δ and the sign of the differences $I_2 - I_1$ the mean value of $\delta(I_2 - I_1)$ is zero, so that

$$R_2'' = \frac{\langle (I_2 - I_1)^2 + 2\delta(I_2 - I_1) + \delta^2 \rangle}{\langle I^2 \rangle} \quad (65)$$

$$= R_2 + R_2', \quad (66)$$

where

$$R_2' = \frac{\langle \delta^2 \rangle}{\langle I^2 \rangle} \quad (67)$$

is the residual resulting from experimental errors and imperfections in the model. The values of R_2 are thus simply additive – a result that could have been predicted from the relation of R_2 to the variance of the intensities. The combination is not so simple, however, for R_1 and R . These can be written

$$R_1'' = \langle |X + x| \rangle \quad (68)$$

and

$$R'' = \langle |Y + y| \rangle \quad (69)$$

where

$$X = (I_2 - I_1) / \langle I \rangle, \quad (70)$$

$$x = \delta / \langle I \rangle, \quad (71)$$

$$Y = (|F_2| - |F_1|) / \langle |F| \rangle, \quad (72)$$

$$y = \varepsilon / \langle |F| \rangle, \quad (73)$$

and the averaging takes place over both X and x or both Y and y . Equations (68) and (69) are of the same form, and it suffices to discuss either one of them. Suppose that the fractional errors x have a probability distribution $p(x)dx$. Equation (68) becomes

$$R_1'' = \left\langle - \int_{-\infty}^{-X} (X+x)p(x)dx + \int_{-X}^{\infty} (X+x)p(x)dx \right\rangle, \quad (74)$$

since $|X+x|$ is equal to $(X+x)$ when x is greater than $-X$ and is equal to $-(X+x)$ otherwise. Equation (74) may be rearranged to give

$$R_1'' = \left\langle X \int_{-\infty}^{\infty} \text{sgn}(x)p(x)dx + \int_{-\infty}^{\infty} |x|p(x)dx + 2 \int_{-X}^0 (X+x)p(x)dx \right\rangle \quad (75)$$

$$= \langle X \rangle \langle \text{sgn}(x) \rangle + \langle |x| \rangle + 2 \left\langle \int_{-X}^0 (X+x)p(x)dx \right\rangle, \quad (76)$$

where $\text{sgn}(x)$ indicates the sign of x . If the distribution of either X or x is symmetrical about zero the first term vanishes. Both distributions would be expected to be nearly symmetrical, so that the first term can be neglected. The second term is just the function corresponding to R_1 for the errors, say R_1' . The third term cannot be evaluated unless the distributions of X and x are known or assumed. If the variance of x is large compared with the variance of X one can get an approximation by replacing $p(x)$ by $p(0)$, which gives

$$R_1'' = R_1' + p(0) \langle X^2 \rangle \quad (77)$$

$$= R_1' + p(0) \langle (I_2 - I_1)^2 \rangle / \langle I \rangle^2 \quad (78)$$

$$= R_1' + kp(0)R_2, \quad (79)$$

where

$$k = \langle I^2 \rangle / \langle I \rangle^2 \quad (80)$$

has approximately the value 2 for a non-centrosymmetric structure and approximately the value 3 for a centrosymmetric structure (Wilson, 1951). If the distribution $p(x)$ has a single maximum at $x=0$ equations (77) and (79) set an upper limit for R_1'' . Since equation (68) is symmetrical in X and x , equations (76), (77) and (79) can equally well be written as

$$R_1'' = \langle x \rangle \langle \text{sgn}(X) \rangle + \langle |X| \rangle + 2 \left\langle \int_{-X}^0 (X+x)P(X)dX \right\rangle \quad (81)$$

$$= R_1 + P(0) \langle x^2 \rangle \quad (82)$$

$$= R_1 + kP(0)R_2', \quad (83)$$

where $P(X)dX$ is the probability distribution of X and R_2' is defined by equation (67). Again, for $P(X)$ having a single maximum at the origin, equations (82) and (83) actually set an upper limit for R_1'' . If the variance of X is less than the variance of x the limit set by equations (82) and (83) may be lower than, and therefore preferable to, the upper limit set by equations (77) and (79). Lower limits for R_1'' , under the same conditions, can be obtained by replacing $p(0)$ by $p(\bar{X})$ and $P(0)$ by $P(\bar{x})$, but the probability distributions may fall off too rapidly for these lower limits to be appreciably greater than R_1' or R_1 . Exactly analogous equations can be written for the residual R in terms of the

probability distributions $q(y)dy$ and $Q(Y)dY$, giving

$$R'' = R' + q(0) \frac{\langle (|F_2| - |F_1|)^2 \rangle}{\langle |F|^2 \rangle} \quad (84)$$

and

$$R'' = R + Q(0) \frac{\langle \varepsilon^2 \rangle}{\langle |\bar{F}|^2 \rangle}. \quad (85)$$

It is perhaps worth noting that the argument between equations (68) and (79) is mathematically just a more general version of that between equations (43) and (50), though the meanings attached to the symbols are different, and closely resembles that used in a different residual calculation by Wilson (1950). Equation (50) may be obtained from equation (79) by substituting the appropriate distribution function from equation (44). Some insight may be gained into the nature of the approximation made in equation (77) and its equivalents by employing Gaussian approximations for $p(x)$ and $P(X)$. The sum of two variables with Gaussian distributions has a Gaussian distribution with variance equal to the sum of the variances of the variables, and R_1 for a Gaussian distribution is $(2/\pi)^{1/2}$ times the square root of the variance, so that

$$R'_1 = [R_1^2 + (R'_1)^2]^{1/2}. \quad (86)$$

The first two terms of the binomial expansion of this in powers of the ratio R_1/R'_1 or R'_1/R_1 then lead to equation (77) or (82).

It has been tacitly assumed that the components of the misplacement of the atom are not rationally related to the edges of the unit cell. Dr T. R. Lomer has pointed out to me that a rational misplacement of, say, half a cell edge is quite a probable error. The atom will then contribute correctly to reflexions with the corresponding index even, but with the wrong phase to reflexions with the index odd. This affects the numerical value of the coefficients of some of the residuals to some extent (for example, the mean value of $|\sin x|$ becomes $1/2$ instead of $2/\pi$), but the order of magnitude is not affected.

As the calculations have turned out, the residuals resulting from a misplaced atom are large and unlikely

to be overlooked. When the measurements of intensity have been made with good accuracy and an advanced refinement program is used there will be other and better pointers to a misplaced atom; its temperature factors may be expected to be anomalously large; if occupancy weights are refined that for the misplaced atom will continuously decrease; an electron-density map will probably show two partial atoms, one in the right place and one in the wrong place; a difference Fourier will have large anomalies. Nevertheless, it seems worthwhile to have the calculations on record, if only to avoid their being unnecessarily repeated. In so far as they have any direct use it will be in the study of substances whose nature precludes the attainment of great accuracy in intensity measurement.

It has been assumed in the calculations that the misplacement of one atom does not affect the coordinates of those correctly placed. The question thus arises whether, by small adjustments of the parameters of the correctly placed atoms, it is possible to refine the residual to values appreciably lower than those calculated. A full discussion of this problem would be difficult, but a preliminary consideration of the minimization of R_2 suggests that there would be no systematic changes in the coordinates of the correctly placed atoms. The shifts would be proportional to the mean over all reflexions of $F_0(\partial F_0/\partial x)$, where x is the parameter in question. For the positional parameters the expected value of this average is zero, but in any particular case it may have some finite value, which may be appreciable if the number of reflexions averaged is small. Misplacement of one atom, therefore, will increase the variance of the coordinates of the remaining atoms.

As originally submitted, this paper consisted essentially of §§ 2 to 4. The author is indebted to the referees for suggestions which have led to the expansion of § 1 and the addition of § 5.

References

- WILSON, A. J. C. (1942). *Nature, Lond.* **150**, 151.
 WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318.
 WILSON, A. J. C. (1950). *Acta Cryst.* **3**, 397.
 WILSON, A. J. C. (1951). *Research*, **4**, 141.