Acta Cryst. (1948). **1**, 138

The Fourier Method of Crystal-Structure Analysis

By W. Cochran

Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England

(*Received* 8 *March* 1948)

Alternative methods which could be used to derive accurate atomic co-ordinates from the data provided by an X-ray investigation of a crystal structure are discussed. It is shown that the customary Fourier series method is closely related to the least-squares method applied to this problem by Hughes. On the basis of this result a way in which the Fourier method can be modified to give less weight to certain coefficients has been discovered. The effect of terminating the Fourier series or introducing an artificial temperature factor is also investigated from the point of view of the relation between the Fourier and least-squares solutions. It is concluded that no significant increase in accuracy could be obtained by abandoning the Fourier method.

1. Introduction

Structure determinations by the methods of X-ray crystallography may be divided into two classes: those in which the object of the investigation is to determine the positions of the atoms with sufficient accuracy to give a general picture of the crystal structure, and those in which the object is to measure as accurately as possible the bond lengths and bond angles between the atoms. The method of Fourier synthesis from observed X-ray reflexion intensities has been widely used for both purposes, although the writer is not aware that any attempt has been made to show that it is the best way of utilizing the experimental data when the second of these objects is the one in view. Alternative methods have been suggested by Hughes (1941) and by Booth (1947*a*); their methods are similar in that the atomic co-ordinates are chosen so as to minimize some function of the observed and calculated structure amplitudes, F_o and F_c respectively. In the least-squares method as used by Hughes the function minimized was $\Sigma w(F_o-F_c)^2$, while Booth has suggested minimizing $\Sigma w(F_o^2-F_c^2)^2$, and has shown how this can be done systematically by the method of 'steepest descents'. Weighting factors w are introduced to allow for the fact that the F_o's are not all measurable with the same accuracy; an obvious weakness of the Fourier method is that all F_o's, including those few likely to be systematically in error because of extinction, etc., must be included as coefficients in the Fourier series and must therefore influence the result. The purpose of this note is to investigate the relationship between the Fourier and other methods and to point out ways in which the former can be improved. It is worth pointing out that we shall be concerned only with the problem of finding the best atomic co-ordinates—defined as those points at which the electron density reaches a maximum and around which the electron distribution is spherically symmetric in a small region. If it is desired to investigate the actual electron distribution, the Fourier method is the only one which can be used.

2. The relation between the Fourier and least squares methods

Consider a structure based on a space group which has a centre of symmetry, the structure consisting of m spherically symmetric atoms. To avoid repetition of equations put

$$x_1=x,\quad x_2=y,\quad x_3=z,$$
$$h_1=h,\quad h_2=k,\quad h_3=l,$$

so that x_j may denote x, y or z according as $j=1$, 2 or 3. Let the co-ordinates of the rth atom be x_{r1}, x_{r2}, x_{r3}. As a simplification it will be assumed that the three a-axes are orthogonal, although this is not an inherent limitation. Then

$$F(h_1h_2h_3)=\sum_r f_r\cos 2\pi\left\{\sum_j \frac{h_jx_{rj}}{a_j}\right\} \tag{2·1}$$

and
$$\rho(x_1x_2x_3)=\frac{1}{V}\sum_N F(h_1h_2h_3)\cos 2\pi\left\{\sum_j \frac{h_jx_j}{a_j}\right\}. \tag{2·2}$$

The symbols F, f and ρ have their usual meanings, $\sum\limits_N$ denotes a triple summation over h_1, h_2 and h_3 which includes all F's of finite magnitude.

Experimentally, one measures $F_o=F+\beta$, where β is an error of measurement, and derives therefore

$$\rho_o=\frac{1}{V}\sum_N F_o\cos 2\pi\left\{\sum_j \frac{h_jx_j}{a_j}\right\}. \tag{2·3}$$

The atomic co-ordinates derived experimentally therefore differ from the true co-ordinates. Let x'_{rj} be the experimentally determined co-ordinate corresponding to the true co-ordinate x_{rj}. Since the electron density is to be a maximum at $x_j=x'_{rj}$, it follows that

$$\left(\frac{\partial\rho_o}{\partial x_j}\right)_{x_j=x'_{rj}}=\frac{-2\pi}{a_jV}\sum_N h_jF_o\sin 2\pi\left\{\sum_j \frac{h_jx'_{rj}}{a_j}\right\}=0. \tag{2·4}$$

Now put

$$F_c(h_1h_2h_3)=\sum_r f_r\cos 2\pi\left\{\sum_j \frac{h_jx'_{rj}}{a_j}\right\}, \tag{2·5}$$

so that F_c denotes one of a set of structure amplitudes calculated from experimentally determined co-ordinates. Clearly since the Fourier series is assumed to be completely convergent a synthesis using the F_c's as coefficients must give back the atomic co-ordinates which were used in deriving the F_c's. Therefore

$$\left(\frac{\partial \rho_c}{\partial x_j}\right)_{x_j=x'_{rj}} = \frac{-2\pi}{a_j V}\sum_N h_j F_c \sin 2\pi\left\{\sum_j \frac{h_j x'_{rj}}{a_j}\right\}=0. \qquad (2\cdot 6)$$

Now from (2·5)

$$\frac{\partial F_c}{\partial x'_{rj}} = \frac{-2\pi}{a_j} h_j f_r \sin 2\pi\left\{\sum_j \frac{h_j x'_{rj}}{a_j}\right\}. \qquad (2\cdot 7)$$

Therefore, subtracting (2·6) from (2·4) and using (2·7),

$$\sum_N \frac{1}{f_r}(F_o - F_c)\frac{\partial F_c}{\partial x'_{rj}} = 0. \qquad (2\cdot 8)$$

Now consider the function

$$\phi_r = \sum_N \frac{1}{f_r}(F_o - F_c)^2.$$

The co-ordinates x'_{rj} which make ϕ_r a minimum are given by

$$\frac{\partial \phi_r}{\partial x'_{rj}} = 0, \quad \text{i.e.} \quad \sum_N \frac{1}{f_r}(F_o - F_c)\frac{\partial F_c}{\partial x'_{rj}} = 0.$$

But this is identical with (2·8), so that the atomic co-ordinates determined by Fourier synthesis are such as to render $\sum_N \frac{1}{f_r}(F_o - F_c)^2$ a minimum. The Fourier solution is thus a special case of the least-squares solution in which the weight given to each observation is inversely proportional to the magnitude of the corresponding atomic scattering factor. The fact that ϕ_r is a function of f_r means that a different function is minimized for each different kind of atom in the structure, but since the f-curves of all atoms are similar except in scale it is practically the same function that is minimized for all.

3. A modification of the Fourier method

The fact that the Fourier and least-squares methods are related and that in the latter observations can be weighted according to one's estimate of their probable error suggests that there may be some way of doing this when using the former method. That this is so can be shown as follows.

Let $(x'_{rj})_1$ be one of the atomic co-ordinates determined by the unmodified Fourier method. Each co-ordinate, by (2·8), satisfies

$$\sum_N \frac{1}{f_r}(F_o - F_{c1})\left(\frac{\partial F_c}{\partial x'_{rj}}\right)_1 = 0, \qquad (3\cdot 1)$$

suffix 1 denoting quantities calculated from co-ordinates $(x'_{rj})_1$. Consider a Fourier synthesis whose coefficients are $wF_o + (1-w)F_{c1}$, where w may be different for each F_o. This synthesis will give co-ordinates $(x'_{rj})_2$ which satisfy

$$\sum_N \frac{1}{f_r}[\{wF_o + (1-w)F_{c1}\} - F_{c2}]\left(\frac{\partial F_c}{\partial x'_{rj}}\right)_2 = 0,$$

or rearranging,

$$\sum_N \frac{1}{f_r}[w(F_o - F_{c1}) + (F_{c1} - F_{c2})]\left(\frac{\partial F_c}{\partial x'_{rj}}\right)_2 = 0. \qquad (3\cdot 2)$$

The synthesis can now be repeated, using as coefficients $wF_o + (1-w)F_{c2}$ and so on. When this has been done p times co-ordinates $(x'_{rj})_p$ are obtained which satisfy

$$\sum_N \frac{1}{f_r}[w(F_o - F_{c(p-1)}) + (F_{c(p-1)} - F_{cp})]\left(\frac{\partial F_c}{\partial x'_{rj}}\right)_p = 0.$$

If this approximation process converges,

$$F_{cp} \to F_{c(p-1)} \text{ as } p \to \infty,$$

and the final co-ordinates x'_{rj} (dropping the suffix) satisfy

$$\sum_N \frac{w}{f_r}(F_o - F_c)\frac{\partial F_c}{\partial x'_{rj}} = 0. \qquad (3\cdot 3)$$

However, this is exactly the condition for $\sum_N \frac{w}{f_r}(F_o - F_c)^2$ to be a minimum. An objection to the use of this procedure in practice is that the result is arrived at by successive approximations, each step apparently requiring the calculation of a Fourier synthesis and a set of structure amplitudes, and the process might converge slowly or not at all. This point is taken up in § 4.

4. The analogy of the weighted mean

The result arrived at in § 3 can be illustrated by a simple analogy. Given a set of numbers $x_1, x_2, \ldots, x_n$, which may be thought of as observations not all made with the same accuracy, the weighted mean is given by $M = \Sigma wx/\Sigma w$. This value of M makes $\Sigma w(x-M)^2$ a minimum. An alternative way of calculating M is to form first the mean M_1 from

$$\left.\begin{array}{l} nM_1 = \Sigma x \quad \text{(which makes } \Sigma(x-M_1)^2 \text{ a minimum),} \\ \text{then} \\ nM_2 = \Sigma\{wx + (1-w)M_1\} \quad \text{etc.,} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \text{and finally} \\ nM_p = \Sigma\{wx + (1-w)M_{p-1}\}. \end{array}\right\} \qquad (4\cdot 1)$$

This can be rearranged as

$$M_{p-1} = \frac{\Sigma wx}{\Sigma w} + \frac{n(M_{p-1} - M_p)}{\Sigma w}.$$

In this case it may readily be shown that as

$$p \to \infty, \quad M_p \to M_{p-1} \to M.$$

Of course, there would be no point in calculating a weighted mean in this indirect way, but if one wishes to give less weight to certain coefficients of a Fourier series then the method given in § 3, which is analogous to that given immediately above, is the only way in which this can be done. The analogy is useful because of the following result: the number p of approximations which have to be made in order to get a value M_p which is very close to M depends on the proportion of observations which have been given a weight less than unity, and the result of stopping the process at an intermediate stage, say (4·1), is to give a weight intermediate between 1 and w to an observation which is given a weight w in forming the true weighted mean. The proof of this statement is as follows.

Writing $\epsilon = 1 - \Sigma w/n$, (4·1) may be rearranged as

$$nM_2 = \Sigma(w+\epsilon)\,x,$$

or since

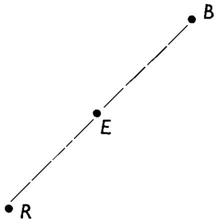
$$n = \Sigma(w+\epsilon), \quad M_2 = \{\Sigma(w+\epsilon)\,x\}/\Sigma(w+\epsilon).$$

Thus an observation which is given a weight w in forming M is given a relative weight $(w+\epsilon)/(1+\epsilon)$ in forming M_2, and since $1 > (w+\epsilon)/(1+\epsilon) > w$, and the amount by which $(w+\epsilon)/(1+\epsilon)$ differs from w depends on the value of ϵ (that is, on the proportion of observations which are given a reduced weight), the statement made above is proved. In view of the close similarity between the results of this section and those of § 3 one can predict that the effect of stopping the approximating process of § 3 at (3·2) will be to give a weight intermediate between 1 and w to those coefficients to which a weight w would be given if the process were continued to (3·3), and the smaller the proportion of F_o's that have been given reduced weight the closer is the approximation given by (3·2) to the final result which satisfies (3·3).

5. The effect of terminating the Fourier series

In practice the Fourier series (2·2) does not contain as coefficients all F's which have a finite magnitude. If each F is thought of as being associated with a reciprocal lattice point, then only those F's which correspond to points contained in a sphere drawn about the origin in reciprocal space can be observed. The radius of this limiting sphere depends on the wave-length of X-rays used in the experiment. The omission of corresponding terms from the Fourier series causes a system of 'ripples' to run through the calculated electron density distribution, and these ripples displace the points of maximum electron density. Such 'finite summation' errors have been discussed by Booth (1946), who suggests the following technique for their elimination: 'Having calculated the structure amplitudes from the final atomic co-ordinates a synthesis is computed using the calculated values as coefficients. Any terms not included in the original synthesis with experimental coefficients are similarly omitted from this new synthesis. The co-ordinates derived will in general deviate from those used in calculating the F values, these deviations give the errors, with reversed signs, of the original co-ordinates.' In a later paper (Booth, 1947*b*) an attempt is made to justify the method by applying certain tests, and it is concluded that 'whilst it cannot be claimed that the foregoing treatment constitutes definite proof of the validity of the correction technique, it indicates that the errors derived from it are at least of the right order of magnitude'.

In § 3 it is shown that the correct way to allow for finite termination is to choose atomic co-ordinates which make $\sum_n \frac{1}{f_r}(F_o - F_c)^2$ a minimum, n being the number of structure amplitudes which can be observed. This gives zero weight to the $N-n$ which cannot be observed. That this condition is approximately fulfilled after correction of the atomic co-ordinates by Booth's method can be shown as follows. Let E be a point of maximum electron density given by a Fourier series containing n coefficients F_o. $(x'_{rj})_E$ is thus an experimentally determined atomic co-ordinate, uncorrected for finite termination. Let B be the corresponding point of maximum electron density given by a Fourier series containing n coefficients $(F_c)_E$. $(\ )_E$ denotes a quantity calculated from co-ordinates $(x'_{rj})_E$.



According to Booth R represents the position of the corresponding atomic centre corrected for finite summation, where R is defined by making REB a straight line with $RE = EB$. Since the ripples have a wavelength long compared with the displacements which they cause to the atomic co-ordinates, a Fourier synthesis containing n coefficients $(F_c)_R$ would have an electron density maximum at E. Now the synthesis using experimental coefficients F_o also has a maximum at E, and since the two electron density distributions have approximately the same shape in the neighbourhood of a maximum, their slopes $\partial\rho_o/\partial x_j$ and $\partial\rho_c/\partial x_j$ will be the same at the point R. That is, with

$$\rho_o = \frac{1}{V}\sum_n F_o \cos 2\pi\left\{\sum_j \frac{h_j x_j}{a_j}\right\} \tag{5·1}$$

and

$$\rho_c = \frac{1}{V}\sum_n (F_c)_R \cos 2\pi\left\{\sum_j \frac{h_j x_j}{a_j}\right\}, \tag{5·2}$$

$$\frac{\partial\rho_o}{\partial x_j} = \frac{\partial\rho_c}{\partial x_j} \tag{5·3}$$

at $x_j = (x'_{rj})_R$.

Since $$(F_c)_R = \sum_r f_r \cos 2\pi\left\{\sum_j \frac{h_j (x'_{rj})_R}{a_j}\right\}$$

(5·3) reduces to

$$\frac{1}{V}\sum_{n}\frac{1}{f_r}F_o\left(\frac{\partial F_c}{\partial x'_{rj}}\right)_R=\frac{1}{V}\sum_{n}\frac{1}{f_r}(F_c)_R\left(\frac{\partial F_c}{\partial x'_{rj}}\right)_R,$$

which is the condition that $\sum\limits_{n}\frac{1}{f_r}(F_o-(F_c)_R)^2$ should be a minimum. A more rigorous proof of the validity of Booth's method can be given in the case where the number of terms $N-n$ omitted from the Fourier series is small; since it merely confirms the above result, it was not thought to be worth reproducing here.

6. The effect of an artificial temperature factor

A method designed to ensure the convergence of the Fourier series and therefore the elimination of finite summation errors has been used by Brill, Grimm, Hermann & Peters (1939) and consists in multiplying each F by a converging factor α, generally taken as having the form $\alpha=\exp[(-B\sin^2\theta)/\lambda^2]$. This produces the same effect on the electron distribution as the real temperature movement of the atoms, so that α is usually called the artificial temperature factor. f_r, F_o and F_c are replaced by αf_r, αF_o and αF_c respectively, so that the co-ordinates obtained from a synthesis whose coefficients are the αF_o's are such as to render $\sum\limits_{N}\frac{\alpha}{f_r}(F_o-F_c)^2$ a minimum. The effect of an artificial temperature factor is therefore to give relatively less weight to structure amplitudes of high order, as might have been anticipated. An interesting case occurs when $\alpha=f_r$, the co-ordinates given by the Fourier method now render $\sum\limits_{N}(F_o-F_c)^2$ a minimum, which is the least-squares solution with all observations given unit weight. One objection to the use of an artificial temperature factor in practice is that, owing to the spreading out of the electron distribution, neighbouring atoms begin to overlap, and the point of maximum electron density of one atom may be moved by the 'tail' of another. A method of correcting for this effect has been given by Booth (1946).

7. Discussion

We are now in a position to attempt an answer to the question whether the Fourier method is the best way of utilizing the experimental data. The theory of errors predicts that in the adjustment of indirect observations involving several unknowns the best values of the unknowns are those which make the sum of the squares of the residuals a minimum, provided that the observations have all been made with the same accuracy. In our case the observations are the values of the F_o's, the unknowns are the co-ordinates of the atoms, and the residuals the values of F_o-F_c. If the observations have not all been made with the same accuracy, the best atomic co-ordinates are those which make $\Sigma w(F_o-F_c)^2$ a minimum, where the weight w of each observation is inversely proportional to the square of its probable error. Since the methods of Hughes (1941) and of Booth (1947*a*) enable all observations to be weighted in this way and are not influenced by 'series termination', it would appear that they must yield a better result than the Fourier method. This conclusion involves two assumptions: (i) the atoms are spherically symmetric, and (ii) their atomic scattering factors are known. Unless these conditions are satisfied F_c cannot be calculated accurately. Since electrons are involved in the formation of chemical bonds, assumption (i) is in general not correct. However, only the outer electrons are affected and these contribute only to reflexions of low Bragg angle. This has been verified by the experimental work of James, Brindley & Wood (1929) on aluminium, and by calculation (Ewald & Hönl, 1936) in the case of carbon. (In this connexion see also Ehrenberg & Schäfer (1932).) For values of $(\sin\theta)/\lambda$ greater than about 0·3 the assumption that electrons are distributed around the nucleus with spherical symmetry appears to be an excellent approximation to the truth. The second objection may be more serious, because, although atomic scattering factors can be calculated accurately, such calculations refer to atoms at rest, and the effective scattering factor of an atom is always decreased by thermal movement. Perhaps the most serious objection to the least-squares method, however, is a practical one—the amount of work involved in setting up the observational equations, reducing these to a set of normal equations and solving the latter for $3m$ unknowns.

The Fourier method possesses the following advantages: (*a*) within limits no assumption regarding the electron distribution in atoms need be made; (*b*) the *relative* values of the atomic scattering factors need be known only with sufficient accuracy to enable one to calculate correctly which sign is to be associated with each $|F_o|$; and (*c*) the computations involved lend themselves more readily to mechanization. The disadvantages of the unmodified Fourier method are as follows: (*a*) no account is taken of the fact that some observations are much less accurate than others; (*b*) errors are introduced by the termination of the series; and (*c*) the treatment of § 2 shows that excessive weight is given to F_o's of high order.

Results arrived at in previous sections show that these disadvantages can be at least partially eliminated.

(1) In § 3 it has been shown that observations believed to be less accurate than the others can be given less weight by replacing F_o by $wF_o+(1-w)F_c$ in a subsequent Fourier synthesis. Observations affected by extinction can be given zero weight by replacing F_o by F_c

(2) In § 5 it is shown that the method of correcting for finite termination suggested by Booth can be justified theoretically.

(3) In § 6 it is shown that by introducing an appropriate artificial temperature factor unit weight can be given to all observations. This has the advantage of simultaneously reducing finite summation errors.

It may therefore be concluded that there are good reasons, theoretical as well as practical, why the Fourier method should be used in accurate crystal-structure determinations.

This work has been done as part of a programme of investigation of the structures of organic molecules by X-ray methods. I am grateful to the Department of Scientific and Industrial Research for financial aid, and to Prof. Sir Lawrence Bragg and Dr W. H. Taylor for the facilities they have provided and the interest they have shown.

References

Brill, R., Grimm, H., Hermann, C. & Peters, C. (1939). *Ann. Phys., Lpz.*, **34**, 393.

Booth, A. D. (1946). *Proc. Roy. Soc.* A, **188**, 77.

Booth, A. D. (1947*a*). *Nature, Lond.*, **160**, 196.

Booth, A. D. (1947*b*). *Proc. Roy. Soc.* A, **190**, 490.

Ewald, P. P. & Hönl, H. (1936). *Ann. Phys., Lpz.*, **25**, 281.

Ehrenberg, W. & Schäfer, K. (1932). *Phys. Z.* **33**, 97.

Hughes, E. W. (1941). *J. Amer. Chem. Soc.* **63**, 1737.

James, R. W., Brindley, G. W. & Wood, R. G. (1929). *Proc. Roy. Soc.* A, **125**, 401.

Acta Cryst. (1948). **1**, 142

Vibration Amplitudes of Atoms in Cubic Crystals

By Kathleen Lonsdale

University College, London, W.C. 1, *England*

(*Received* 20 *February* 1948)

Root-mean square amplitudes of atomic vibration are given for twenty elements and twenty-four compounds crystallizing in the cubic system. The spatial distribution of atomic displacements due to thermal vibration is calculated from the elastic constants of nineteen cubic crystals, and attention is drawn to the relationships existing between the amplitudes of pure longitudinal and transverse waves travelling along the cube edges, face diagonals and cube diagonals, and the structures of these crystals.

From the data given, the intensity of diffuse scattering power could be plotted in reciprocal space, for comparison with experimental data from monochromatic Laue photographs.

It is possible to calculate a root-mean square amplitude of thermal vibration of atoms in cubic crystals if the Debye characteristic temperature is known. But the atoms are not just simple-harmonic oscillators, although their movements are resolvable into a series of harmonic vibrations. The object of the present paper is in § 1 to give the root-mean square amplitudes and in § 2 to investigate the spatial distribution of amplitudes in these component vibrations, and to consider what relation, if any, exists between the amplitudes of waves travelling in certain principal directions and the crystal structure.

1. Root-mean square amplitude of vibration

The Debye-Waller formula, $I_T = Ie^{-2M}$, which was confirmed by early experimenters, applies to cubic crystals composed of one kind of atom only, at temperatures not too near to the melting-point. It expresses the reduction of Bragg scattering of X-rays by crystals with increasing temperature, in terms of the mean square displacement of an atom from the average position which it occupies in the crystal. In this expression M is given by

$$M = \frac{8\pi^2 \sin^2\theta}{3\lambda^2}\,\overline{u^2}. \tag{1}$$

M may, however, also be expressed in terms of Θ, the Debye characteristic temperature of specific heat theory

$$M = \frac{6h^2}{mk\Theta}\,\frac{\sin^2\theta}{\lambda^2}\left\{\frac{\phi(x)}{x} + \frac{1}{4}\right\}, \tag{2}$$

where h, k are the Planck and Boltzmann constants, m the mass of the atom in grams, θ the Bragg angle, λ the wave-length of X-rays, $x = \Theta/T$, where T is the absolute temperature, and $\phi(x)$ is the Debye function of

$$\frac{1}{x}\int_0^x \frac{\xi\,d\xi}{e^\xi - 1}$$

(which is tabulated, for instance, in the *Internationale Tabellen zur Bestimmung von Kristallstrukturen* (1935), **2**, p. 574). The term $\frac{1}{4}$ allows for the existence of zero-point energy.

Equating (1) and (2), we find that

$$\overline{u^2} = \frac{9h^2}{4\pi^2 km\Theta}\left\{\frac{\phi(x)}{x} + \frac{1}{4}\right\} = \frac{4{\cdot}364 \times 10^{-14}}{A\Theta}\left\{\frac{\phi(x)}{x} + \frac{1}{4}\right\}, \tag{3}$$

where A is the atomic weight in terms of ${}^{16}_{8}\mathrm{O}$, the values of universal constants used being those given by Birge (1941).

Table 1 gives the root-mean square amplitudes at 293° K. calculated from equation (3) for a number of