Acta Cryst. (1958). **11**, 579

Structure Factor Relations and the Phase Problem

By W. Cochran

Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England

(*Received* 3 *February* 1958)

A method for deriving equality relations between structure factors is described, and the conditions under which the results are valid are discussed. It is pointed out that most such relations require an amount of data much greater than the minimum necessary for structure determination. It is concluded that, for a crystal structure of a specified degree of complexity, no routine procedure has yet been devised which will enable the structure to be determined from the X-ray intensities alone. Conditions which are sufficient to exclude the possibility of a structure having a 'homometric mate' are given.

1. Introduction

In recent papers Hauptman & Karle (1957) and Karle & Hauptman (1957) have described a 'unified algebraic approach to the phase problem'. This has enabled them to derive new relations between the structure factors, for the particular case of a structure containing N equal atoms per unit cell. These relations, and others, can be derived by what might be called a 'Fourier transform approach to the phase problem'. This approach has the advantage of showing quite clearly the conditions under which the relations are valid, and here our conclusions are not identical with those of Karle & Hauptman. The question of the extent to which the relations can be used for the practical determination of crystal structures is more difficult to answer, but the approach given here throws some light on it. Some conclusions are also reached about the conditions under which a structure is uniquely determined by its structure amplitudes.

2. Results from Fourier transform theory

We consider a distribution $\varrho(\mathbf{r})$ consisting of N points of weights $W_1, \ldots, W_N$ at $\mathbf{r}_1 \ldots \mathbf{r}_N$, repeated by an infinite lattice. That is, the distribution in one unit cell is given by

$$\varrho(\mathbf{r}) = \sum_{j=1}^{N} W_j \delta(\mathbf{r}-\mathbf{r}_j)\,, \qquad (2\cdot1)$$

where $\delta(\mathbf{r})$ is a Dirac function. We also define

$$F(\mathbf{h}) = \sum_{j=1}^{N} W_j \exp\,[2\pi i \mathbf{r}_j.\mathbf{h}]\,, \qquad (2\cdot2)$$

and we shall refer to $\varrho(\mathbf{r})$ and $F(\mathbf{h})$ as Fourier transforms of one another, a relationship denoted by

$$\varrho(\mathbf{r}) \rightleftarrows F(\mathbf{h})\,.$$

(Strictly speaking, it is the continuous structure factor which is the transform of $\varrho(\mathbf{r})$.) We now state the following theorem:

If

$$\left.\begin{aligned} &\sum_j W_j'\delta(\mathbf{r}-\mathbf{r}_j) \rightleftarrows F_1(\mathbf{h}) \\ \text{and}\quad &\sum_j W_j''\delta(\mathbf{r}-\mathbf{r}_j) \rightleftarrows F_2(\mathbf{h}), \\ \text{then}\quad &\sum_j W_j'W_j''\delta(\mathbf{r}-\mathbf{r}_j) \rightleftarrows \langle F_1(\mathbf{k})F_2(\mathbf{h}-\mathbf{k})\rangle_{\mathbf{k}}\,, \end{aligned}\right\} \qquad (2\cdot3)$$

where $\langle\ \rangle_{\mathbf{k}}$ denotes an average over an infinite range. It should be noted that the points involved in the third relation are those which are *common* to the first two, since, if any W' or W'' is zero, a term disappears from the left-hand side of the third relation. The proof of (2·3) is given in a later section.

Two other useful results which are readily proved are that, with

$$\sum_j W_j\delta(\mathbf{r}-\mathbf{r}_j) \rightleftarrows F(\mathbf{h})\,,$$

then

$$\sum_{i,j} W_iW_j\delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)) \rightleftarrows |F(\mathbf{h})|^2, \qquad (2\cdot5)$$

or

$$\sum_{i\neq j} W_iW_j\delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)) \rightleftarrows |F(\mathbf{h})|^2-\sum_j W_j^2\,, \qquad (2\cdot5a)$$

and also

$$\sum_j (W_j \cos 2\pi\mathbf{H}.\mathbf{r}_j)\delta(\mathbf{r}-\mathbf{r}_j) \rightleftarrows \tfrac{1}{2}(F(\mathbf{h}+\mathbf{H}) + F(\mathbf{h}-\mathbf{H}))\,. \qquad (2\cdot6)$$

3. Derivation of some typical equalities

As an example of the application of the results given in the last section, we consider the derivation of certain equations given by Karle & Hauptman (1957). These equations are written in terms of a quantity $E(\mathbf{h})$, related to our $F(\mathbf{h})$ by

$E(\mathbf{h}) = F(\mathbf{h})/(\sum_j W_j^2)^{\frac{1}{2}}$, that is, $E(\mathbf{h}) = N^{-\frac{1}{2}}F(\mathbf{h})$ when all W's are unity. In our notation these equations become

$$F(\mathbf{h}) = \langle F(\mathbf{k})F(\mathbf{h}-\mathbf{k})\rangle_{\mathbf{k}}\,, \qquad (3\cdot1)$$

$$F(\mathbf{h}) = \langle (F^2(\mathbf{k})-N)F(\mathbf{h}-2\mathbf{k})\rangle_{\mathbf{k}}\,, \qquad (3\cdot2)$$

$$|F(\mathbf{h})|^2-N = \langle(|F(\mathbf{k})|^2-N)(|F(\mathbf{h}-\mathbf{k})|^2-N)\rangle_{\mathbf{k}}, \quad (3\cdot3)$$

$$|F(\mathbf{h}+\mathbf{H})F(-\mathbf{h}+\mathbf{H})F(-2\mathbf{H})|\cos(\varphi_{\mathbf{h}+\mathbf{H}}+\varphi_{-\mathbf{h}+\mathbf{H}}+\varphi_{-2\mathbf{H}}) = \tfrac{1}{2}\langle(|F(\mathbf{k}+\mathbf{H})|^2-N)(|F(\mathbf{k}-\mathbf{H})|^2-N) \times(|F(\mathbf{h}+\mathbf{k})|^2-N)\rangle_{\mathbf{k}}+(|F(\mathbf{h}+\mathbf{H})|^2+|F(\mathbf{h}-\mathbf{H})|^2 +|F(2\mathbf{H})|^2-2N), \quad (3\cdot4)$$

where

$$F(\mathbf{h}) = |F(\mathbf{h})|\exp[i\varphi_{\mathbf{h}}].$$

The first relation holds for both centrosymmetric and non-centrosymmetric structures, the second only for centrosymmetric and the third and fourth only for non-centrosymmetric structures.

To derive (3·1) we put $F_1(\mathbf{h}) = F_2(\mathbf{h}) = F(\mathbf{h})$ in relation (2·3), giving

$$\sum_j W_j^2\delta(\mathbf{r}-\mathbf{r}_j) \rightleftarrows \langle F(\mathbf{k})F(\mathbf{h}-\mathbf{k})\rangle_{\mathbf{k}}. \quad (3\cdot5)$$

Comparing this with the relation defining $F(\mathbf{h})$ itself,

$\sum_j W_j\delta(\mathbf{r}-\mathbf{r}_j) \rightleftarrows F(\mathbf{h})$, we see that if all $W_j = 1$,

relation (3·1) follows immediately. A necessary and sufficient condition for the truth of (3·1) is (aside from $W = 1$, and remembering that $\langle\ \rangle_{\mathbf{k}}$ is an average over an infinite range)

$$\mathbf{r}_i-\mathbf{r}_j \neq 0. \quad (3\cdot7)$$

(We shall always take this to include the condition $\mathbf{r}_i-\mathbf{r}_j \neq \mathbf{t}$, where $\mathbf{t}$ is any lattice translation.)

The basis of (3·2) may be understood by writing it as

$$F(\mathbf{h}) = \langle(F^2(\tfrac{1}{2}\mathbf{k})-N)F(\mathbf{h}-\mathbf{k})\rangle_{\mathbf{k}}. \quad (3\cdot2a)$$

Since the structure is centrosymmetric, $\mathbf{r}_i = -\mathbf{r}_{N-i}$, and the Fourier transform of $F^2(\mathbf{h})$ will include peaks at $2\mathbf{r}_i$; in fact

$$F^2(\mathbf{h})-N \rightleftarrows \sum_i W_i^2\delta(\mathbf{r}-2\mathbf{r}_i)+\sum_{i\neq j}2W_iW_j\delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)), \quad (3\cdot8)$$

If in the relation (2·3) we now take $F_1(\mathbf{h}) = F^2(\frac{1}{2}\mathbf{h})-N$ and $F_2(\mathbf{h}) = F(\mathbf{h})$, we see that (3·2$a$) depends on the fact that the transforms of $F_1(\mathbf{h})$ and $F_2(\mathbf{h})$ have common points at $\mathbf{r}_1, \ldots, \mathbf{r}_N$, and that

$$\langle(F^2(\tfrac{1}{2}\mathbf{k})-N)F(\mathbf{h}-\mathbf{k})\rangle_{\mathbf{k}} \rightleftarrows \sum_j W_j^3\delta(\mathbf{r}-\mathbf{r}_j). \quad (3\cdot9)$$

Taking all $W_j = 1$, relation (3·2a) follows. The conditions for the validity of this result are those we have already met, together with

$$2\mathbf{r}_i-(\mathbf{r}_n-\mathbf{r}_m) \neq 0. \quad (3\cdot10)$$

This last condition ensures that there are no accidental coincidences of points in the second term of (3·8) with points in the first term. If in (3·10) we take $m = N-n$, we see that it includes the condition $2(\mathbf{r}_i-r_m) \neq 0$, and if we now take $i = N-m$, we see that $4\mathbf{r}_i \neq 0$ is also included. Remembering that this is meant to include $4\mathbf{r}_i + \mathbf{t}$, we see that such special positions as $(\frac{1}{2}, 0, 0)$, $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, etc. must not be occupied. These conditions are all included in (3·10), or in the statement that all the points $2\mathbf{r}_i$ of the vector set are to be resolved from one another, and from points $\mathbf{r}_n-\mathbf{r}_m$.

The result (3·3) can be derived immediately by putting $F_1(\mathbf{h}) = F_2(\mathbf{h}) = |F(\mathbf{h})|^2-N$ in (2·3). The conditions for exact validity are found to be (3·7), plus the condition.

$$(\mathbf{r}_n-\mathbf{r}_m)-(\mathbf{r}_p-\mathbf{r}_q) \neq 0, \quad (3\cdot11)$$

or simply that all points of the vector set are to be resolved from one another (and of equal weight). The condition deduced by Karle & Hauptman (1957, footnote to p. 156) is more restrictive, being

$$A(\mathbf{r}_n-\mathbf{r}_m)+B(\mathbf{r}_p-\mathbf{r}_q) \neq 0, \quad (3\cdot12)$$

where A and B are any integers. As a check, we calculated values of $|F|^2-N$ for a one-dimensional structure with

$$r_1 = 0, \quad r_2 = \tfrac{1}{8}, \quad r_3 = \tfrac{5}{16}, \quad r_4 = \tfrac{3}{4}.$$

Condition (3·11) is satisfied, (3·12) is not, but it was verified by numerical calculation that (3·3) is valid.

The derivation of relation (3·4) is given in Appendix I, and it is shown that it is valid provided that

$$(\mathbf{r}_i-\mathbf{r}_j)-\{(\mathbf{r}_n-\mathbf{r}_m)-(\mathbf{r}_p-\mathbf{r}_q)\} \neq 0. \quad (3\cdot13)$$

In a discussion of the probability of the correctness of certain statistical relations between structure factors, Klug (1958) found it useful to define an 'order' for each relation. On this definition, (3·1), (3·2), (3·3) and (3·4) are respectively of order 3, 4, 5 and 7. We see from conditions (3·7), (3·10), (3·11) and (3·13) that the order of each equality is one more than the number of independent atomic positions involved in each condition.

4. The effect of averaging over a limited range

In practice the range of $\mathbf{k}$ over which averages can be taken is restricted. From Fig. 1 we see that the condition for both vectors $\mathbf{k}$ and $\mathbf{h}-\mathbf{k}$ to lie inside limiting spheres of radius S_0, is that $\mathbf{k}$ should lie inside the volume whose projection is shaded in the diagram. Now

$$\langle F_1(\mathbf{k})F_2(\mathbf{h}-\mathbf{k})\rangle'_{\mathbf{k}} = \langle(\sum_j W'_j\exp[2\pi i\mathbf{r}_j.\mathbf{k}]) \times(\sum_j W''_j\exp[2\pi i\mathbf{r}_j.(\mathbf{h}-\mathbf{k})])\rangle'_{\mathbf{k}}, \quad (4\cdot1)$$

where $\langle\ \rangle'_{\mathbf{k}}$ denotes an average over this *limited* range of $\mathbf{k}$. Transferring the origin from O by a distance $\frac{1}{2}\mathbf{h}$, and introducing a new coordinate in reciprocal space, $\mathbf{u} = \mathbf{k}-\frac{1}{2}\mathbf{h}$ (see Fig. 1), we find, since $\mathbf{k} = \frac{1}{2}\mathbf{h}+\mathbf{u}$ and $\mathbf{h}-\mathbf{k} = \frac{1}{2}\mathbf{h}-\mathbf{u}$, that

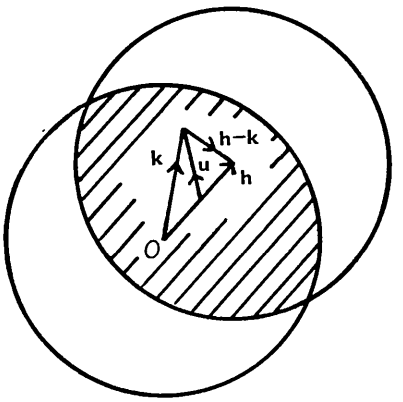

Fig. 1. The two spheres in the diagram each of radius S_0, have their centres separated by a distance **h**. The condition that both $|\mathbf{k}|$ and $|\mathbf{h}-\mathbf{k}|$ should be less than S_0 is that **k** should lie inside the volume common to the two spheres.

$$\langle F_1(\mathbf{k})F_2(\mathbf{h}-\mathbf{k})\rangle'_{\mathbf{k}} = \langle \sum_j W'_j W''_j \exp[2\pi i \mathbf{r}_j.\mathbf{h}]\rangle_{\mathbf{u}} + \langle \sum_{i\neq j} W'_i W''_j \exp[\pi i(\mathbf{r}_i+\mathbf{r}_j).\mathbf{h}] \exp[2\pi i(\mathbf{r}_i-\mathbf{r}_j).\mathbf{u}]\rangle'_{\mathbf{u}}, \quad (4\cdot2)$$

that is,

$$\langle F_1(\mathbf{k})F_2(\mathbf{h}-\mathbf{k})\rangle'_{\mathbf{k}} = \sum_j W'_j W''_j \exp[2\pi i \mathbf{r}_j.\mathbf{h}] + \sum_{i\neq j} W'_i W''_j \exp[\pi i(\mathbf{r}_i+\mathbf{r}_j).\mathbf{h}]\langle \exp[2\pi i(\mathbf{r}_i-\mathbf{r}_j).\mathbf{u}]\rangle'_{\mathbf{u}}. \quad (4\cdot3)$$

For an infinite range of **u**, the second term on the right vanishes and we have the result given earlier as (2·3). When the range of **u** is finite, we have an additional contribution from every point $\frac{1}{2}(\mathbf{r}_i+\mathbf{r}_j)$, the magnitude of the contribution depending on $(\mathbf{r}_i-\mathbf{r}_j)$, and decreasing as this separation increases. This contribution depends explicitly on

$$f(\mathbf{r}_i-\mathbf{r}_j) = \langle \exp[2\pi i(\mathbf{r}_i-\mathbf{r}_j).\mathbf{u}]\rangle'_{\mathbf{u}} = \langle C(\mathbf{u}) \exp[2\pi i(\mathbf{r}_i-\mathbf{r}_j).\mathbf{u}]\rangle_{\mathbf{u}}, \quad (4\cdot4)$$

where $C(\mathbf{u}) = 1$ inside the shaded region of Fig. 1, but is zero outside. Relation (4·4) will be seen to express the fact that $f(\mathbf{r}_i-\mathbf{r}_j)$ is the Fourier transform of the function $C(\mathbf{u})$. Now this function has a width in any direction of about S_0, so by the well known reciprocal relation between the widths of a function and its transform, $f(\mathbf{r}_i-\mathbf{r}_j)$ must fall nearly to zero in a distance of about S_0^{-1}. This can be confirmed for small values of **h**, for then $f(\mathbf{r}_i-\mathbf{r}_j)$ is given explicitly by

$$f(\mathbf{r}_i-\mathbf{r}_j) = \frac{3\,(\sin 2\pi|\mathbf{r}_i-\mathbf{r}_j|S_0 - 2\pi|\mathbf{r}_i-\mathbf{r}_j|S_0 \cos 2\pi|\mathbf{r}_i-\mathbf{r}_j|S_0)}{(2\pi|\mathbf{r}_i-\mathbf{r}_j|S_0)^3}. \quad (4\cdot5)$$

For greater values of **h**, $f(\mathbf{r}_i-\mathbf{r}_j)$ spreads out and is no longer spherically-symmetrical, but for $h < S_0$ it will not differ greatly from the result given as (4·5).

We now apply this result to the case $F_1(\mathbf{h}) = F_2(\mathbf{h}) = F(\mathbf{h})$, taking all $W_j = 1$. We find, using the above results, that

$$\langle F(\mathbf{k})F(\mathbf{h}-\mathbf{k})\rangle'_{\mathbf{k}} = F(\mathbf{h}) + \sum_{i\neq j} f(\mathbf{r}_i-\mathbf{r}_j) \exp[\pi i(\mathbf{r}_i+\mathbf{r}_j).\mathbf{h}]. \quad (4\cdot6)$$

The condition for the second term to vanish is that all $f(\mathbf{r}_i-\mathbf{r}_j)$ should be negligible, which will be true if

$$|\mathbf{r}_i-\mathbf{r}_j| \gtrsim S_0^{-1}. \quad (4\cdot7)$$

This replaces the condition $(\mathbf{r}_i-\mathbf{r}_j) \neq 0$. It is readily shown that a corresponding change has to be made to other conditions such as (3·11), which becomes $|(\mathbf{r}_n-\mathbf{r}_m)-(\mathbf{r}_p-\mathbf{r}_q)| \gtrsim S_0^{-1}$, for example.

The interpretation of condition (4·7) in terms of functions in crystal, as opposed to reciprocal, space is evident. When the range of **h** is infinite, each atom in an electron-density map whose Fourier coefficients are F's as defined in § 2, is represented by a δ-function, and $(\mathbf{r}_i-\mathbf{r}_j) \neq 0$ requires the δ-functions not to coincide. When the range of **h** is restricted to a sphere of radius S_0, each atom is represented by a peak of width $\approx S_0^{-1}$, and (4·7) requires such peaks to be resolved from one another. We therefore come to appreciate that, for example, the *exact* validity of (3·3) depends on the *complete* resolution of all peaks in the Patterson function.

5. Some further equality relations

By successive applications of (3·1) it may be shown that

$$F(\mathbf{h}) = \langle F(\mathbf{k})F(\mathbf{l}-\mathbf{k})F(\mathbf{m}-\mathbf{l})\ldots F(\mathbf{h}-\mathbf{m})\rangle_{\mathbf{k},\mathbf{l},\ldots,\mathbf{m}}. \quad (5\cdot1)$$

There are similar results corresponding to (3·2) etc.

From the results (2·5a) and (I·9) it is readily shown that

$$|F(\mathbf{h})|^2 - N = 2(N-1)\langle (|F(\mathbf{k})|^2-N)(|F(\mathbf{h}-\mathbf{k})|^2-N)^2\rangle_{\mathbf{k}}. \quad (5\cdot2)$$

From (I·5), (I·6) and (I·10) we find that

$$\begin{aligned} &|F(\mathbf{h}+\mathbf{H})F(-\mathbf{h}+\mathbf{H})F(-2\mathbf{H}) \\ &\quad\times|\cos(\varphi_{\mathbf{h}+\mathbf{H}}+\varphi_{-\mathbf{h}+\mathbf{H}}+\varphi_{-2\mathbf{H}}) - |F(2\mathbf{H})|^2 \\ &= \tfrac{1}{4}\langle\{(|F(\mathbf{k}+\mathbf{H})|^2-N)(|F(\mathbf{k}-\mathbf{H})|^2-N) - (2N-8) \\ &\quad\times(|F(\mathbf{k}+\mathbf{H})|^2+|F(\mathbf{k}-\mathbf{H})|^2-2N)\} \\ &\quad\times\{|F(\mathbf{h}-\mathbf{k})|^2-N\}. \end{aligned} \quad (5\cdot3)$$

Equality relations can be found even when the atoms of the structure are unequal. For example, corresponding to (3·2),

$$\langle F(\mathbf{k})F(\mathbf{l}-\mathbf{k})F(\mathbf{h}-\mathbf{l})\rangle_{\mathbf{k},\mathbf{l}} = \langle (F^2(\mathbf{k})-\Sigma W_j^2)(F(\mathbf{h}-2\mathbf{k}))\rangle_{\mathbf{k}}. \quad (5\cdot4)$$

Suppose the atoms of a structure are equal, the

structure is centrosymmetric and the position of any *one* atom is known, say $\mathbf{r}_1$, then it can be shown that

$$F(\mathbf{h}) = \tfrac{1}{4}\langle (F^2(\mathbf{k})-N+4)(F^2(\mathbf{h}-\mathbf{k})-N+4) \times \cos 2\pi\mathbf{r}_1.(\mathbf{h}-2\mathbf{k})\rangle_{\mathbf{k}} . \quad (5{\cdot}5)$$

Since this result holds for any $\mathbf{r}_j$, it follows that

$$F(\mathbf{h}) = \frac{1}{4N}\langle (F^2(\mathbf{k})-N+4)(F^2(\mathbf{h}-\mathbf{k})-N+4) \times F(\mathbf{h}-2\mathbf{k})\rangle_{\mathbf{k}} . \quad (5{\cdot}6)$$

If the structure is non-centrosymmetric, then, provided that the origin is chosen to coincide with an atomic position,

$$F(\mathbf{h}) = \langle (|F(\mathbf{k})|^2-N+1)F(\mathbf{h}-\mathbf{k})\rangle_{\mathbf{k}} . \quad (5{\cdot}7)$$

If the origin is chosen midway between atoms 1 and 2, and the separation $\mathbf{r}_1-\mathbf{r}_2 = 2\mathbf{r}_1$ is known, it can be shown that

$$2(|F(\mathbf{h})|\cos\varphi_{\mathbf{h}}-\cos 4\pi\mathbf{r}_1.\mathbf{h}) = \langle (|F(\mathbf{k})|^2-N+1)(|F(\mathbf{h}-\mathbf{k})|^2-N+1)\cos 4\pi\mathbf{r}_1.\mathbf{k}\rangle_{\mathbf{k}} . \quad (5{\cdot}8)$$

If the coordinate of a third atom relative to this origin, $\mathbf{r}_3$, is known it can be shown that

$$F(\mathbf{h}) = 2\langle (|F(\mathbf{k})|\cos\varphi_{\mathbf{k}}-\cos 4\pi\mathbf{r}_1.\mathbf{k}) \times (|F(\mathbf{h}-\mathbf{k})|^2-N+1)\exp[2\pi i\mathbf{r}_3.(\mathbf{h}-\mathbf{k})]\rangle_{\mathbf{k}} . \quad (5{\cdot}9)$$

Proofs of the results from (5·5) onwards will not be given; they have at most a curiosity value. Their physical basis will be pointed out in the next section.

6. The Patterson-function aspect of structure determination

It was shown by Wrinch (1939) that in general a set of points can be recovered from its vector set. We shall consider first in this section a structure (set of points) whose space group is $P1$; the atoms may be unequal and we need not know their weights or the value of N. We suppose, however, that all points of the vector set are resolved,

$$(\mathbf{r}_n-\mathbf{r}_m)-(\mathbf{r}_p-\mathbf{r}_q) \neq 0 . \quad (6{\cdot}1)$$

In the vector set we look for three points at $\mathbf{v}_1, \mathbf{v}_2$ and $\mathbf{v}_3$ such that $\mathbf{v}_3 = \mathbf{v}_2-\mathbf{v}_1$. The identification of three such points, with their weights, fixes the weights and relative coordinates of three points of the structure, or of its enantiomorph, for we can assume that

$$\mathbf{v}_1 = \mathbf{r}_i-\mathbf{r}_j, \quad \mathbf{v}_2 = \mathbf{r}_i-\mathbf{r}_k, \quad \mathbf{v}_3 = \mathbf{r}_j-\mathbf{r}_k .$$

An incorrect identification is impossible if (6·1), and the following condition, are always satisfied:

$$(\mathbf{r}_i-\mathbf{r}_j)-\{(\mathbf{r}_n-\mathbf{r}_m)-(\mathbf{r}_p-\mathbf{r}_q)\} \neq 0 . \quad (6{\cdot}2)$$

These three points, in their correct relative positions, are then moved over the vector set until a triple coincidence is found with three points of the vector set whose weights are in the same ratio as those of the first three of the structure. This serves to fix the weight and relative position of a fourth point of the structure. A false indication is not possible if conditions (6·1) and (6·2) are satisfied. The coincidence-seeking process is then repeated with these four points until a fifth has been found, and so on. This process is essentially the same as was described by Wrinch (1939) and Buerger (1950, 1951). It is mentioned here in order to show that the conditions which ensure success for the method, are the same as those given in § 3, and which ensure the success of phase-determining relations. The considerations of the present section show that these geometrical conditions are the same whether the points (atoms) are equal in weight or not. In particular, if conditions (6·1) and (6·2) are satisfied, a structure whose space group is $P1$ (whether composed of equal atoms or not) can have no homometric mate other than its enantiomorph. These conditions are contained in the one deduced by Karle & Hauptman (1957), but are less restrictive. While (6·1) and (6·2) provide sufficient conditions to exclude a homometric solution, they are clearly not necessary conditions. Neither will their occasional failure be a serious bar to the determination of a crystal structure by a routine procedure. This is clear from the point of view of Patterson-function methods when we remember that there are so many ways in which the first three points, $\mathbf{r}_i$, $\mathbf{r}_j$, $\mathbf{r}_k$, of the above method can be chosen. Even if they are at first incorrectly identified because of a failure of (6·2), the coincidence-seeking process is not likely to proceed much beyond this stage, and a fresh start would be made with three other points. The real limitation of course lies in the contrast between a vector set of points, in terms of which the discussion has been conducted, and the Patterson function which can be obtained in practice. This is simply another aspect of the point discussed in § 4.

When the space group is $P\bar{1}$, the structure can be recovered from its vector set by the 'single superposition method' of Clastre & Gay (1950*a*, *b*). For the process to be free from ambiguity, condition (6·1) is required, together with

$$2\mathbf{r}_i-\{(\mathbf{r}_n-\mathbf{r}_m)-(\mathbf{r}_p-\mathbf{r}_q)\} \neq 0 . \quad (6{\cdot}3)$$

This method of Clastre & Gay, expressed in terms of operations in reciprocal space, leads to (5·5). The same process, applied in a situation where the space group is $P\bar{1}$, leads to (5·8). Their 'double superposition method' for the determination, from its vector set, of a structure whose space group is $P1$, leads to (5·9).

A structure whose space group is $P\bar{1}$, and for which conditions (6·1) and (6·3) are satisfied, can have no homometric mate.

7. The amount of data required for direct structure analysis

For simplicity the discussion in this section will refer to a structure whose space group is $P\bar{1}$, and which is composed of equal atoms. The conclusions, however, are quite general in their application. Suppose the structure is to be determined using equality relations such as

$$F(2\mathbf{h}) = 2(F^2(\mathbf{h})-N)-\langle(F^2(\mathbf{k})-N)(F^2(\mathbf{h}-\mathbf{k})-N)\rangle_{\mathbf{k}} . \quad (7\cdot1)$$

to determine certain signs from measured intensities, followed by the use of

$$F(\mathbf{h}) = \langle(F^2(\mathbf{k})-N)F(\mathbf{h}-2\mathbf{k})\rangle_{\mathbf{k}} \quad (7\cdot2)$$

and

$$F(\mathbf{h}) = \langle F(\mathbf{k})F(\mathbf{h}-\mathbf{k})\rangle_{\mathbf{k}} . \quad (7\cdot3)$$

The relation (7·1) has been derived in terms of the Fourier transforms of both sides of the equation in an earlier paper (Cochran, 1954); it is readily shown to be exact provided that

$$|2\mathbf{r}_i-(\mathbf{r}_n-\mathbf{r}_m)| \gtrsim S_0^{-1} ,$$

that is, it requires certain peaks in the Patterson function (those between atoms related by the centre of symmetry) to be resolved from the others.

If the data available are contained within a limiting sphere of radius S_0, the number M of structure amplitudes (magnitudes of structure factors) is given by

$$M = (4/3)\pi S_0^3 V, \quad (7\cdot4)$$

where V is the unit-cell volume. The width of a peak in the Patterson function (whose coefficients are values of $|F|^2$, as defined in this paper) is then of the order S_0^{-1}, and for maxima to be even approximately resolved from one another we require (since there are approximately N^2 separate peaks)

$$V/N^2 \approx S_0^{-3}, \quad (7\cdot5)$$

which, combined with (7·4), requires

$$M \approx (4/3)\pi N^2. \quad (7\cdot6)$$

Relations of higher order require an amount of data involving a higher power of N. This provides us with an estimate of the data required to determine a structure by a routine procedure, whether based on Patterson-function methods or sign-determining relations makes no difference. This conclusion, based on (7·6), has also been reached by Vaughan (private communication), to whom I owe the idea.

The conclusion appears at first to contradict the result of Hauptman & Karle (1950), who showed that the problem of determining $3N$ atomic coordinates is in general completely determinate when $M = 3(N-1)$ independent structure amplitudes are known. There is, however, no contradiction; relations of the lowest order (which is three) require the minimum amount of data. For example, relation (7·3) is valid as soon as

$$|\mathbf{r}_i-\mathbf{r}_j| \gtrsim S_0^{-1}, \quad (7\cdot7)$$

it merely requires the *atoms* to be resolved from one another. This is the case as soon as $V/N \approx S_0^{-3}$ (7·8), i.e., when

$$M \approx (4/3)\pi N . \quad (7\cdot9)$$

This is close to the Hauptman & Karle value. We therefore conclude that if a structure is to be determined from the minimum amount of data, only relations of lowest order (see § 3) can be used. These depend on relations between the electron density and powers of the electron density, for example (7·3) expresses a relation between ϱ and ϱ^2, (5·1) expresses a relation between ϱ and any power of ϱ. It has been shown by Woolfson (1958) that similar relations can be found when the atoms are unequal. Those equalities which give signs of structure factors directly from *intensities* inevitably involve relations between ϱ and the Patterson function P, or between ϱ and P^2, or between ϱ and the product ϱP etc. For their exact validity they therefore require more than the minimum amount of data. Unfortunately there is no way as yet in which an equation such as (7·3) can be solved by a routine procedure, although for fairly simple structures something approaching this can be done (Sayre, 1952; Cochran & Penfold, 1952).

The number of X-ray intensities measurable in practice increases roughly in proportion to N, since V is approximately proportional to N. For organic compounds, for example, one finds (neglecting hydrogens) that $V \simeq 12N$, where V is in Å^3. The range of measurement seldom extends as far as $S_0 = 2$ Å^{-1}. It may therefore be concluded that the limit of structure determination by Patterson-function methods, or by methods based on equality relations of order greater than three, is reached when

$$(4/3)\pi(2)^3\times 12N \approx (4/3)\pi N^2, \quad \text{or} \quad N \approx 100 .$$

The situation is of course completely altered if a few atoms are of much greater weight than the others, as the Patterson function may then contain recognizable features. The estimates we have made are if anything too generous, and the figure $N = 100$ may well be too great by a factor of four when the atoms are equal, or nearly so.

APPENDIX I

Proof of the relation

$$\begin{aligned}|F(\mathbf{h}+\mathbf{H})F(-\mathbf{h}+\mathbf{H})F(-2\mathbf{H})| \cos(\varphi_{\mathbf{h}+\mathbf{H}}+\varphi_{-\mathbf{h}+\mathbf{H}}+\varphi_{-2\mathbf{H}}) \\ = \tfrac{1}{2}\langle(|F(\mathbf{k}+\mathbf{H})|^2-N)(|F(\mathbf{k}-\mathbf{H})|^2-N) \\ \times(|F(\mathbf{h}+\mathbf{k})|^2-N)\rangle_{\mathbf{k}} \\ +(|F(\mathbf{h}+\mathbf{H})|^2+|F(\mathbf{h}-\mathbf{H})|^2+|F(2\mathbf{H})|^2-2N) . \quad (\text{I}\cdot1)\end{aligned}$$

We write P for the quantity occurring on the left-

hand side. We regard **H** as a constant, and note that since P is a structure invariant we can choose an origin so that $\varphi_{2\mathbf{H}} = 0$. $F(2\mathbf{H})$ is then real and positive, and we have

$$P = F(2\mathbf{H})|F(-\mathbf{h}+\mathbf{H})F(\mathbf{h}+\mathbf{H})|\cos(\varphi_{-\mathbf{h}+\mathbf{H}}+\varphi_{\mathbf{h}+\mathbf{H}}) . \quad (I\cdot2)$$

We then make use of the general result

$$\tfrac{1}{2}(F_1F_2+F_1^*F_2^*) = |F_1F_2|\cos(\varphi_1+\varphi_2)$$

to obtain

$$P = \tfrac{1}{2}F(2\mathbf{H})\{F^*(\mathbf{h}-\mathbf{H})F(\mathbf{h}+\mathbf{H})+F(\mathbf{h}-\mathbf{H})F^*(\mathbf{h}+\mathbf{H})\},$$

or

$$P = \tfrac{1}{2}F(2\mathbf{H})\{|F(\mathbf{h}+\mathbf{H})+F(\mathbf{h}-\mathbf{H})|^2 - (|F(\mathbf{h}+\mathbf{H})|^2+|F(\mathbf{h}-\mathbf{H})|^2)\} . \quad (I\cdot3)$$

Using general results given earlier, and taking the atoms to be equal from the outset, we find that

$$|F(\mathbf{h}+\mathbf{H})+F(\mathbf{h}-\mathbf{H})|^2-2(N+F(2\mathbf{H})) \rightleftarrows 4\sum_{i\neq j}(\cos 2\pi\mathbf{H}.\mathbf{r}_i\cos 2\pi\mathbf{H}.\mathbf{r}_j)\delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)), \quad (I\cdot4)$$

and that

$$|F(\mathbf{h}+\mathbf{H})|^2+|F(\mathbf{h}-\mathbf{H})|^2-2N \rightleftarrows 2\sum_{i\neq j}(\cos 2\pi\mathbf{H}.(\mathbf{r}_i-\mathbf{r}_j)\delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)) . \quad (I\cdot5)$$

Multiplying both sides of (I·4) and (I·5) by $\frac{1}{2}F(2\mathbf{H})$, subtracting one from the other, and making use of (I·3), we obtain

$$P-F^2(2\mathbf{H}) \rightleftarrows \sum_{i\neq j}(F(2\mathbf{H})\cos 2\pi\mathbf{H}.(\mathbf{r}_i+\mathbf{r}_j)) \delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)) , \quad (I\cdot6)$$

where P is given by (I·3) or (I·2).

Next we consider the function

$$\begin{aligned}&2(|F(\mathbf{h}+\mathbf{H})|^2-N)(|F(\mathbf{h}-\mathbf{H})|^2-N)\\ &\quad=\{(|F(\mathbf{h}+\mathbf{H})|^2-N)+(|F(\mathbf{h}-\mathbf{H})|^2-N)\}^2\\ &\qquad-\{(|F(\mathbf{h}+\mathbf{H})|^2-N)^2+(|F(\mathbf{h}-\mathbf{H})|^2-N)^2\}\\ &\quad=Q-R, \text{ say} . \end{aligned} \quad (I\cdot7)$$

Both terms Q and R are obviously related to $(|F(\mathbf{h})|^2-N)^2$; in fact the transforms of all three consist of δ-functions at the same points

$$\mathbf{r} = (\mathbf{r}_n-\mathbf{r}_m)-(\mathbf{r}_p-\mathbf{r}_q) , \quad (I\cdot8)$$

which are the points obtained by forming the vector set of points $\mathbf{r}_i$, $i = 1, \ldots, N$, subtracting the peak at the origin, and then forming in turn the vector set of this vector set. Certain points of this vector–vector set coincide with points $\mathbf{r}_i-\mathbf{r}_j$ of the vector set, for example in (I·8) take $m = q$, $n = i$ and $p = j$. As we shall see later, it is with these points only that we are concerned.

It is found that

$$(|F(\mathbf{h})|^2-N)^2 \rightleftarrows \sum_{i\neq j} 2(N-2)\delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)) + \text{other terms, not at points of the vector set.} \quad (I\cdot9)$$

Similarly,

$$Q \rightleftarrows \sum_{i\neq j}(4F(2\mathbf{H})\cos 2\pi\mathbf{H}.(\mathbf{r}_i+\mathbf{r}_j) +(4N-16)\cos 2\pi\mathbf{H}.(\mathbf{r}_i-\mathbf{r}_j))\delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)) + \text{other terms} , \quad (I\cdot10)$$

and

$$R \rightleftarrows \sum_{i\neq j} 4(N-2)\cos 2\pi\mathbf{H}.(\mathbf{r}_i-\mathbf{r}_j)\delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)) + \text{other terms} , \quad (I\cdot11)$$

so that, subtracting R from Q and multiplying by $\frac{1}{4}$,

$$\tfrac{1}{2}(|F(\mathbf{h}+\mathbf{H})|^2-N)(|F(\mathbf{h}-\mathbf{H})|^2-N) \rightleftarrows \sum_{i\neq j}(F(2\mathbf{H}) \times\cos 2\pi\mathbf{H}.(\mathbf{r}_i+\mathbf{r}_j)-2\cos 2\pi\mathbf{H}.(\mathbf{r}_i-\mathbf{r}_j)) \times\delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)) + \text{terms not at } (\mathbf{r}_i-\mathbf{r}_j) . \quad (I\cdot12)$$

Now we also have

$$|F(\mathbf{h})|^2-N \rightleftarrows \sum_{i\neq j}\delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)), \quad \text{with } \textit{no} \text{ other terms} . \quad (I\cdot13)$$

Therefore, from our central theorem (2·3)

$$\tfrac{1}{2}\langle(|F(\mathbf{k}+\mathbf{H})|^2-N)(|F(\mathbf{k}-\mathbf{H})|^2-N)(|F(\mathbf{k}-\mathbf{h})|^2-N)\rangle_{\mathbf{k}} \rightleftarrows \sum_{i\neq j}(F(2\mathbf{H})\cos 2\pi\mathbf{H}.(\mathbf{r}_i+\mathbf{r}_j)-2\cos 2\pi\mathbf{H}.(\mathbf{r}_i-\mathbf{r}_j)) \times\delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)) . \quad (I\cdot14)$$

The conditions for this result to be rigorously true are, first that (I·13) should be correct, which, as before requires

$$\mathbf{r}_n-\mathbf{r}_m \neq 0 \qquad (n, m \text{ different}) ,$$
$$(\mathbf{r}_n-\mathbf{r}_m)-(\mathbf{r}_p-\mathbf{r}_q) \neq 0 , \qquad (n, m, p, q \text{ all different}),$$

and that no point of the vector–vector set should *accidentally* coincide with a point of the vector set, i.e.

$$(\mathbf{r}_i-\mathbf{r}_j)-\{(\mathbf{r}_n-\mathbf{r}_m)-(\mathbf{r}_p-\mathbf{r}_q)\} \neq 0 \qquad (i, j, n, m, p, q, \text{ all different}) . \quad (I\cdot15)$$

From (I·14), and making use of (I·5), we have

$$\tfrac{1}{2}\langle(|F(\mathbf{k}+\mathbf{H})|^2-\mathrm{N})(|F(\mathbf{k}-\mathbf{H})|^2-N)(|F(\mathbf{h}+\mathbf{k})|^2-N)\rangle_{\mathbf{k}} +(|F(\mathbf{h}+\mathbf{H})|^2+|F(\mathbf{h}-\mathbf{H})|^2-2N \rightleftarrows \sum_{i\neq j}(F(2\mathbf{H}) \times\cos 2\pi\mathbf{H}.(\mathbf{r}_i+\mathbf{r}_j))\delta(\mathbf{r}-(\mathbf{r}_i-\mathbf{r}_j)) . \quad (I\cdot16)$$

But from (I·6), the right-hand side of (I·16) is also the transform of $P-F^2(2\mathbf{H})$. Hence, on rearranging terms, we have proved (I·1). The conditions for the truth of (I·1) are therefore those given as (I·15).

References

BUERGER, M. J. (1950). *Acta Cryst.* **3**, 87.
BUERGER, M. J. (1951). *Acta Cryst.* **4**, 531.

CLASTRE, J. & GAY, R. (1950*a*). *C. R. Acad. Sci., Paris*, **230**, 1976.
CLASTRE, J. & GAY, R. (1950*b*). *J. Phys. Radium*, **11**, 75.
COCHRAN, W. (1954). *Acta Cryst.* **7**, 581.
COCHRAN, W. & PENFOLD, B. R. (1952). *Acta Cryst.* **5**, 644.
HAUPTMAN, H. & KARLE, J. (1950). *Acta Cryst.* **3**, 478.
HAUPTMAN, H. & KARLE, J. (1957). *Acta Cryst.* **10**, 267.
KARLE, J. & HAUPTMAN, H. (1957). *Acta Cryst.* **10**, 515.
KLUG, A. (1958). *Acta Cryst.* **11**, 515.
SAYRE, D. (1952). *Acta Cryst.* **5**, 60.
WOOLFSON, M. M. (1958). *Acta Cryst.* **11**, 277.
WRINCH, D. (1939). *Phil. Mag.* **27**, 98.

Acta Cryst. (1958). **11**, 585

Extinction Effects in Neutron Scattering from Single Magnetic Crystals*

BY WALTER C. HAMILTON

Chemistry Department, Brookhaven National Laboratory, Upton, Long Island, New York, U.S.A.

(*Received* 2 *December* 1957)

Results arising from the treatment of magnetic scattering of neutrons, taking due account of the possibility of polarization of the neutron beam, indicate that the usual expressions for primary and secondary extinction must be modified in certain cases. Extinction, particularly primary extinction, will be generally more severe for reflections which have both nuclear and magnetic contributions than for either pure nuclear or pure magnetic reflections. Formulas and curves are presented for primary and secondary extinction corrections which are applicable to both magnetized and unmagnetized ferromagnetic or antiferromagnetic crystals. Some of the results obtained may be conveniently used to determine relative amounts of primary and secondary extinction, and consequently both mosaic block size and angular distribution. Many of the calculations are of interest in predicting the effects of extinction on experiments designed either to produce or use polarized neutron beams.

The precession of the neutron polarization axis about the magnetic axis can affect the reflected intensity if extinction is severe. This effect is discussed briefly and is shown to be serious only for magnetized crystals.

The appendix discusses the necessary changes in the scattering formulas if all the spins in the unit cell do not lie along a unique magnetic axis.

Introduction

In a recent paper (Hamilton, 1957) the author has discussed secondary extinction corrections for crystals of arbitrary geometrical cross-section. In the example which was chosen to illustrate some of the points discussed in that paper—a synthetic single crystal of magnetite which showed particularly severe extinction—it was noted that several of the reflections which had large magnetic contributions did not give as good a fit to the extinction curves as did the pure nuclear reflections. An empirical extinction curve was found to give a satisfactory fit to the observed intensities of the nuclear reflections, but the intensities of many of the mixed reflections were considerably lower than this curve would predict. This lowered intensity could not be accounted for by any reasonable changes in the parameters describing the magnetic structure, nor in the form factor and saturation curves. Preliminary considerations indicated that this behavior could be explained by a combination of polarization and extinction effects, and the present paper is a detailed elaboration of that point of view.

Following Halpern & Johnson (1939), we may write the wave function for the incident neutron beam as

$$\psi_0 = (2\pi M_0/hk)^{\frac{1}{2}} \exp\,[i\mathbf{k}.\mathbf{r}]\chi_s\,, \qquad (1)$$

where $\mathbf{k}$ is the wave vector $2\pi\mathbf{P}/h$ with $\mathbf{P}$ the neutron momentum, $\mathbf{r}$ is a position vector, M_0 is the neutron mass, and χ_s is the neutron spin function. The scattered wave from a single oriented magnetic ion may then be represented by

$$\psi_H = (2\pi M_0/hk)^{\frac{1}{2}} r^{-1} \exp\,[ikr](b+p\mathbf{q}.\mathbf{s})\chi_s\,, \qquad (2)$$

provided that there is no change in the spin state of the scattering ion. The nuclear and magnetic scattering amplitudes are given by b and p respectively with p defined as

$$p = (e^2\gamma_n S/mc^2)f\,. \qquad (3)$$

Here γ_n is the neutron magnetic moment in nuclear magnetons, S is the spin of the scattering ion, m is the mass of the electron, f is a form factor, and c and e

* Research performed under the auspices of the U.S. Atomic Energy Commission.