Acta Cryst. (1955). 8, 473

Relations between the Phases of Structure Factors

BY W. COCHRAN

Crystallographic Laboratory, Cavendish Laboratory, Cambridge, England

(*Received* 12 *February* 1955)

Relations between the phase angles of complex structure factors (analogous to the sign relations which are known to apply for real structure factors) are derived, and their possible application to crystal-structure analysis is discussed.

1. Introduction

It is well known that for a centrosymmetric crystal the signs of structure factors $F(\mathbf{h})$, $F(\mathbf{h}')$, $F(\mathbf{h}+\mathbf{h}')$ and $F(\mathbf{h}-\mathbf{h}')$ tend to be related by the 'sign relation'

$$\begin{aligned} s(\mathbf{h}) &= s(\mathbf{h}')s(\mathbf{h}+\mathbf{h}') \\ &= s(\mathbf{h}')s(\mathbf{h}-\mathbf{h}') . \qquad (1) \end{aligned}$$

When the crystal is not centrosymmetric one might expect an analogous relation to hold between the phase angles:

$$\alpha(\mathbf{h}) = \alpha(\mathbf{h}')+\alpha(\mathbf{h}-\mathbf{h}') , \qquad (2)$$

and this turns out to be correct. Whereas $s(\mathbf{h})$ may be positive or negative only, $\alpha(\mathbf{h})$ has an infinite number of possible values in the range 0–2π. Equation (2) does, however, give the *most probable* value of $\alpha(\mathbf{h})$ when $\alpha(\mathbf{h}')$ and $\alpha(\mathbf{h}-\mathbf{h}')$ are known. In this paper we consider what distribution of values of $\alpha(\mathbf{h})$ about the most probable value is to be expected, how this distribution is altered when various values of $\alpha(\mathbf{h}')$ and $\alpha(\mathbf{h}-\mathbf{h}')$, for different $\mathbf{h}'$, are given, and a few related questions. Whether these results may be of use in the determination of crystal structures, particularly of proteins, is considered in a final section.

2. Notation

The notation is consistent with that used by Cochran & Woolfson (1955). A few amended definitions are given, and the less obvious ones repeated.

$$\begin{aligned} U(\mathbf{h}) &= \sum_{j=1}^{N} n_j \exp\,[2\pi i\mathbf{h}\,.\,\mathbf{r}_j] = |U(\mathbf{h})| \exp\,[i\alpha(\mathbf{h})] \\ &= A(\mathbf{h})+iB(\mathbf{h}) . \end{aligned}$$

$$U^*(\mathbf{h}) = A(\mathbf{h})-iB(\mathbf{h}).$$

$$V(\mathbf{h}) = \sum_{j=1}^{N} n_j^2 \exp\,[2\pi i\mathbf{h}\,.\,\mathbf{r}_j] = |V(\mathbf{h})| \exp\,[i\beta(\mathbf{h})] .$$

N = number of atoms per unit cell.

$$\varrho_s(\mathbf{r}) = \frac{1}{v}\sum_{h} U(\mathbf{h}) \exp\,[-2\pi i\mathbf{h}\,.\,\mathbf{r}] .$$

v = unit cell volume.

$$\varepsilon_m = \sum_{j=1}^{N} n_j^m .$$

$$E(\mathbf{h}) = U(\mathbf{h})/(\overline{|U(\mathbf{h})|^2})^{\frac{1}{2}} .$$

$\langle X\rangle$ = the expected value of any quantity X.

3. The relation $\alpha(\mathbf{h}) = \alpha(\mathbf{h}')+\alpha(\mathbf{h}-\mathbf{h}')$

For a centrosymmetric crystal it has been shown (Cochran & Woolfson, 1955) that $V(\mathbf{h}) = \overline{U(\mathbf{h}')U(\mathbf{h}+\mathbf{h}')}^{h'}$

The average may equally well be written

$$\overline{U(-\mathbf{h}')U(\mathbf{h}-\mathbf{h}')}^{h'},$$

or

$$V(\mathbf{h}) = \overline{U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')}^{h'} \qquad (3)$$

since $U(\mathbf{h}') = U(-\mathbf{h}')$ in this case.

The above result, in the form (3) *only*, can easily be shown to apply also when the crystal lacks a centre of symmetry. When the atoms are equal it reduces to

$$U(\mathbf{h}) = N\,\overline{U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')}^{h'}. \qquad (4)$$

Suppose now that $U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')$ is known for only one value of $\mathbf{h}'$. We write

$$U(\mathbf{h}) = \frac{1}{N}\sum_{j=1}^{N} \exp\,[2\pi i\mathbf{h}'\,.\,\mathbf{r}_j] \exp\,[2\pi i(\mathbf{h}-\mathbf{h}')\,.\,\mathbf{r}_j] .$$

(For the present the atoms are assumed to be equal.) From reasoning closely similar to that set out by Woolfson (1954), it then follows that for a fixed value of $U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')$, the expected value of $U(\mathbf{h})$ is

$$\langle U(\mathbf{h})\rangle = U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}') . \qquad (5)$$

This is the same result as was found by Woolfson, but it now relates complex numbers and requires both

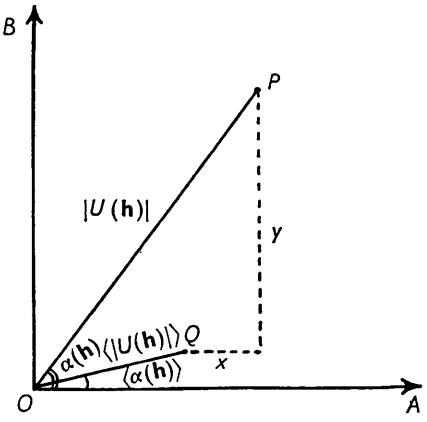

Fig. 1.

$$\langle|U(\mathbf{h})|\rangle = |U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')| \quad \text{and} \quad \langle\alpha(\mathbf{h})\rangle = \alpha(\mathbf{h}')+\alpha(\mathbf{h}-\mathbf{h}') .$$

Referring to Fig. 1, values of $U(\mathbf{h})$ are thus distributed about the point Q in the complex plane. With this point as origin, let

$$x = A(\mathbf{h})-\langle A(\mathbf{h})\rangle , \quad y = B(\mathbf{h})-\langle B(\mathbf{h})\rangle .$$

It then follows that

$$\overline{x^2}^h = \overline{y^2}^h \simeq \frac{1}{2N} ,$$

so that

$$P(x, y)dxdy = \frac{N}{\pi}\exp\left[-N(x^2+y^2)\right]dxdy$$

gives the probability that the point representing $U(\mathbf{h})$ occupies the element of area $dxdy$. When $\langle U(\mathbf{h})\rangle = 0$ this probability distribution reduces to that given by Wilson (1949). In any practical situation, the value of $|U(\mathbf{h})|$ will be known, and one is interested only in the distribution of $\alpha(\mathbf{h})$. From Fig. 1 it can be deduced that

$$P(\alpha(\mathbf{h})) \propto \exp\left[-N(x^2+y^2)\right] ,$$

and that

$$\begin{aligned} x^2+y^2 = \{&|U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')|^2 \\ &+|U(\mathbf{h})|^2-2|U(\mathbf{h})U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')| \\ &\times\cos\left(\alpha(\mathbf{h})-\langle\alpha(\mathbf{h})\rangle\right)\} . \end{aligned}$$

Since $\int_0^{2\pi} P(\alpha)d\alpha$ is necessarily unity, we then have

$$P(\alpha(\mathbf{h})) = \exp\left[-N\{\ldots\}\right]\Big/\int_0^{2\pi}\exp\left[-N\{\ldots\}\right]d\alpha(\mathbf{h}) ,$$

where the quantity in brackets $\{\ldots\}$ is as above. On cancelling common factors which are constant,

$$P(\alpha(\mathbf{h})) = \frac{\exp\left[2N|U(\mathbf{h})U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')|\cos\left(\alpha(\mathbf{h})-\langle\alpha(\mathbf{h})\rangle\right)\right]}{\int_0^{2\pi}\exp\left[2N|U(\mathbf{h})U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')|\cos\alpha\right]d\alpha} . \qquad (6)$$

The value of the definite integral in the denominator of this expression is $2\pi I_0(2N|U(\mathbf{h})U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')|)$, where I_0 is a modified Bessel function of the second kind (Watson, 1922, p. 181). Equation (6) is more conveniently written as

$$P(\alpha(\mathbf{h})) = \frac{\exp\left[-2X\sin^2\frac{1}{2}\left(\alpha(\mathbf{h})-\langle\alpha(\mathbf{h})\rangle\right)\right]}{2\pi e^{-X}I_0(X)} , \qquad (7)$$

$$\left.\begin{aligned} \text{where} \quad & X = 2N|U(\mathbf{h})U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')| \\ \text{and} \quad & \langle\alpha(\mathbf{h})\rangle = \alpha(\mathbf{h}')+\alpha(\mathbf{h}-\mathbf{h}') . \end{aligned}\right\} \qquad (8)$$

Equation (7) thus gives the required distribution of $\alpha(\mathbf{h})$ when $|U(\mathbf{h})|$, $U(\mathbf{h}')$ and $U(\mathbf{h}-\mathbf{h}')$ are known. For small values of $\alpha(\mathbf{h})-\langle\alpha(\mathbf{h})\rangle$ the distribution is

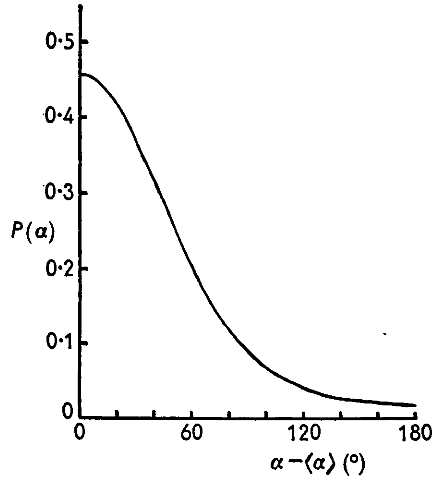

Fig. 2. Distribution of phase angles about the expected value, in a particular case.

Table 1. *Values of* $\dfrac{\exp\left[-2X\sin^2\frac{1}{2}(\alpha-\langle\alpha\rangle)\right]}{2\pi e^{-X}I_0(X)}$

X ↓	$(\alpha-\langle\alpha\rangle)$ (°) 0	20	40	60	80	100	120	140	160	180
0	0·159	0·159	0·159	0·159	0·159	0·159	0·159	0·159	0·159	0·159
1	0·342	0·322	0·271	0·207	0·149	0·105	0·076	0·059	0·049	0·046
2	0·516	0·457	0·323	0·190	0·099	0·049	0·025	0·015	0·011	0·009
3	0·655	0·547	0·325	0·146	0·055	0·019	0·007	0·003	0·002	0·002
4	0·769	0·604	0·302	0·104	0·028	0·007	0·002			
5	0·867	0·642	0·269	0·071	0·014	0·002				
6	0·965	0·673	0·237	0·048	0·007					
7	1·035	0·679	0·261	0·031	0·003					
8	1·110	0·686	0·171	0·020						
9	1·179	0·686	0·144	0·013						
10	1·245	0·682	0·120	0·008						
11	1·308	0·674	0·099	0·006						
12	1·367	0·664	0·083	0·003						
13	1·424	0·651	0·068							
14	1·479	0·637	0·056							
15	1·532	0·621	0·046							

almost Gaussian. Values of the function (7) are given in Table 1, and the form of the distribution for the particular case $N = 30$,

$$|U(\mathbf{h})| = |U(\mathbf{h}')| = |U(\mathbf{h}-\mathbf{h}')| = 0{\cdot}3\,,$$

is shown graphically in Fig. 2. In this instance it can be deduced by numerical integration that the probability that $\alpha(\mathbf{h})$ is within ±20° of $\alpha(\mathbf{h}')+\alpha(\mathbf{h}-\mathbf{h}')$ is 0·31, within ±40°, 0·57, etc.

The result (7) applies when the atoms are equal. When they are not, we may assume by analogy with the corresponding result for the centrosymmetric crystal (Cochran & Woolfson, 1955) that

$$\langle U(\mathbf{h})\rangle = KU(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')\,,$$

where K is entirely real. The 'least squares' value of K is found to be

$$K = \overline{U(\mathbf{h})U^*(\mathbf{h}')U^*(\mathbf{h}-\mathbf{h}')}^{\,h,h'} \div \overline{|U(\mathbf{h}')|^2|U(\mathbf{h}-\mathbf{h}')|^2}^{\,h,h'} = \varepsilon_3/\varepsilon_2^2\,,$$

as for the centrosymmetric crystal. Remembering that now

$$\overline{x^2}^{\,h} = \overline{y^2}^{\,h} \simeq \tfrac{1}{2}\varepsilon_2\,,$$

one finds that (7) still applies, with

$$\left.\begin{aligned} X &= \frac{2\varepsilon_3}{\varepsilon_2^3}\,|U(\mathbf{h})U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')| \\ \text{and}\quad \langle\alpha(\mathbf{h})\rangle &= \alpha(\mathbf{h}')+\alpha(\mathbf{h}-\mathbf{h}')\,. \end{aligned}\right\} \quad (9)$$

4. The relation between $\alpha(\mathbf{h})$ and the phase of $\Sigma\, U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')$

How are a number of 'phase indications' from values of $U(\mathbf{h}_1)U(\mathbf{h}-\mathbf{h}_1)$, $U(\mathbf{h}_2)U(\mathbf{h}-\mathbf{h}_2)$, ... to be combined? From the result (3) it is clear that from a large number of such indications the phase angle $\beta(\mathbf{h})$ can be determined precisely, but only when the atoms are equal does this give $\alpha(\mathbf{h})$. We assume now that the atoms are unequal, and that values of $U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')$ are available for only some limited number m of different values of $\mathbf{h}'$. Considerations of symmetry suggest we take

$$\langle U(\mathbf{h})\rangle = k\sum_1^m U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')\,,$$

where k is entirely real. The least-squares value of k is found to be

$$k = \frac{\varepsilon_3}{\varepsilon_2^2+(m-1)\varepsilon_4}\,,$$

and writing

$$\varrho^2 = \overline{(|U(\mathbf{h})-k\sum_1^m U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')|)^2}^{\,h,h'}\,,$$

it is found that[†]

$$\varrho^2 = \frac{\varepsilon_2^3+(m-1)\varepsilon_2\varepsilon_4-m\varepsilon_3^2}{\varepsilon_2^2+(m-1)\varepsilon_4}\,.$$

These results are exactly as for the centrosymmetric crystal. Following the line of argument given in § 3, one then arrives at (7), with

$$\left.\begin{aligned} X &= \frac{2\varepsilon_3}{\varepsilon_2^3+(m-1)\varepsilon_2\varepsilon_4-m\varepsilon_3^2}\,|U(\mathbf{h})\sum_1^m U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')| \\ \text{and}\quad \alpha(\mathbf{h}) &= \text{phase of} \sum_1^m U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')\,. \end{aligned}\right\} \quad (10)$$

The following properties of this result are worth noting:

(i) When $m = 1$, (10) reduces to (9) when we remember that ε_3^2 ($\sim N^{-4}$) can be neglected compared with ε_2^3 ($\sim N^{-3}$).

(ii) When the atoms are equal,

$$X = 2N|U(\mathbf{h})\sum_1^m U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')|\,.$$

(iii) When the number of terms in the summation is very large

$$\sum_1^m U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}') \to mV(\mathbf{h})\,,$$

so that

$$X \to \frac{2\varepsilon_3}{\varepsilon_2\varepsilon_4-\varepsilon_3^2}\,|U(\mathbf{h})V(\mathbf{h})| \quad\text{and}\quad \langle\alpha(\mathbf{h})\rangle \to \beta(\mathbf{h})\,.$$

If, in addition, the atoms are equal, $P(\alpha(\mathbf{h})) = 0$ except when $\alpha(\mathbf{h}) = \langle\alpha(\mathbf{h})\rangle$, that is, the phase angle is precisely determined.

It would be useful to know the *a priori* distribution of the angle $\alpha(\mathbf{h})-\beta(\mathbf{h})$, that is, the expected distribution when only $|U(\mathbf{h})|$ is known. This would indicate for what range of values of $|U(\mathbf{h})|$ it was worth while to begin to calculate $V(\mathbf{h})$—if $\alpha(\mathbf{h})$ is going to be widely distributed about $\beta(\mathbf{h})$ it is not worth while to estimate the value of the latter. It is readily shown that, as for the centrosymmetric crystal,

$$\langle V(\mathbf{h})\rangle = \frac{\varepsilon_3}{\varepsilon_2}\,U(\mathbf{h})\,,$$

and that σ^2, defined as

$$\overline{\left|V(\mathbf{h})-\frac{\varepsilon_3}{\varepsilon_2}\,U(\mathbf{h})\right|^2}^{\,h}\,,$$

is equal to $\varepsilon_4-\varepsilon_3^2/\varepsilon_2$.

On the usual assumptions, it then follows that

$$P(|V|,\,\beta)d|V|d\beta = \frac{1}{\pi\sigma^2}\exp\left[-\frac{1}{\sigma^2}\left|V-\frac{\varepsilon_3}{\varepsilon_2}\,U\right|^2\right]d|V|\,|V|d\beta\,.$$

[†] The symbol ϱ is not to be confused with ϱ_s defined in § 2, with which it has no connection.

Table 2. *Values of* $P(\Delta)$

	Δ (°)									
$R\lvert E\rvert$ ↓	0	20	40	60	80	100	120	140	160	180
0	0·159	0·159	0·159	0·159	0·159	0·159	0·159	0·159	0·159	0·159
1	0·579	0·487	0·305	0·159	0·080	0·044	0·027	0·019	0·015	0·015
2	1·128	0·664	0·166	0·029	0·005	0·002				
3	1·692	0·553	0·031	0·001						
4	2·256	0·326	0·002							
5	2·820	0·141								
6	3·384	0·048								
7	3·950	0·005								
8	4·514	0·002								
9	5·078									
10	5·642									

(The index $\mathbf{h}$ can be omitted here without leading to any confusion.) Writing $\Delta = \alpha-\beta$, the distribution of Δ when $|V|$ is unknown is thus

$$P(\Delta) = \frac{1}{\pi\sigma^2}\int_{V=0}^{\infty} \exp\left[-\frac{1}{\sigma^2}\left\{\left(\frac{\varepsilon_3|U|}{\varepsilon_2}\right)^2 + |V|^2 - \frac{2\varepsilon_3}{\varepsilon_2}|UV| \cos\Delta\right\}\right] V dV\,.$$

On evaluating the integral, it is found that

$$P(\Delta) = \exp[-R^2|E|^2]\left(\frac{1}{2\pi} + \frac{R|E| \cos\Delta}{2\sqrt{\pi}} \times \exp[R^2|E|^2 \cos^2\Delta]\{1+\operatorname{erf}(R|E| \cos\Delta)\}\right)\,, \quad (11)$$

where

$$R = \frac{\varepsilon_3}{\sigma\sqrt{\varepsilon_2}} = \frac{\varepsilon_3}{(\varepsilon_2\varepsilon_4-\varepsilon_3^2)^{\frac{1}{2}}} \quad \text{and} \quad |E|^2 = |U|^2/\varepsilon_2\,.$$

Thus the *a priori* distribution of α about β is sharpest for the larger structure factors ($|E|$ large), particularly for structures in which the atoms are nearly equal (R large). Some values of the distribution (11) are given in Table 2.

An idea of the order of magnitude of R can be obtained from the fact that for $CaB_3O_4(OH)_3.H_2O$ it is 1·92. It should be emphasized that $P(\Delta)$ is the probability estimated from $|U(\mathbf{h})|$ only. As soon as values of $U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')$ are known the probability should, at least in principle, be calculated from (7) with X and $\langle\alpha(\mathbf{h})\rangle$ as defined by (10).

5. Discussion

When a crystal is centrosymmetric, sign relations can in suitable cases be used to determine the structure completely. There is no hope of achieving this from relations between the phase angles when the crystal lacks a centre of symmetry. In certain circumstances these relations may be useful however. The method of isomorphous replacement, as used by Bokhoven, Schoone & Bijvoet (1951), for example, gives phase angles with an inherent twofold ambiguity. If now the ambiguity is removed for some limited number of the larger unitary structure factors (by finding an approximate structure), these phases can be used to indicate which phase is correct for at least some of the remaining structure factors. The success of this process would of course depend on the number of phase angles known without ambiguity, the number of atoms in the unit cell, etc. For example, suppose that $N = 30$ (equal atoms), $|U(\mathbf{h})| = 0{\cdot}3$ and $\alpha(\mathbf{h})$ is known from the isomorphous replacement results to be either +60° or −60°. If we also know that for a particular $\mathbf{h}'$, $|U(\mathbf{h}')U(\mathbf{h}-\mathbf{h}')| = 0{\cdot}09$ and $\alpha(\mathbf{h}')+\alpha(\mathbf{h}-\mathbf{h}') = 40°$, reference to Fig. 2 shows that from this *one* indication, $P(+60°)/P(-60°) = 0{\cdot}42/0{\cdot}069$, since $\langle\alpha(\mathbf{h})\rangle = 40°$. That is, $P(+60°) = 0{\cdot}86$. Dr F. H. Crick pointed this out to the writer, and suggested that a similar procedure might be used to determine additional phase angles for a protein crystal, once a number of phases had been unambiguously established.

The distribution given by (7) can be expressed in terms of normalized structure factors by writing

$$X = (2\varepsilon_3/\varepsilon_2^{3/2})|E(\mathbf{h})E(\mathbf{h}')E(\mathbf{h}-\mathbf{h}')|\,.$$

Values of $|E|$ of the same order of magnitude are to be expected whatever the complexity of the crystal structure (Hauptman & Karle, 1953), so that the quantity $\varepsilon_3/\varepsilon_2^{3/2}$ is a measure of the efficiency of phase relations (or of sign relations). Since $\varepsilon_3/\varepsilon_2^{3/2}$ varies approximately as $N^{-\frac{1}{2}}$, it is clear that for $N \approx 5{\cdot}10^3$, as in a globular protein crystal, a very broad distribution of $\alpha(\mathbf{h})$ about $\langle\alpha(\mathbf{h})\rangle$ is inevitable. However, the theory given in previous sections is not entirely applicable since the data available from crystalline proteins do not extend to spacings sufficient to resolve individual atoms. One must take into account the effect of this limited resolution. On retracing the steps of the calculation, one finds that $\varepsilon_3/\varepsilon_2^{3/2}$ arose from the substitution

$$\frac{\overline{U(\mathbf{h})U^*(\mathbf{h}')U^*(\mathbf{h}-\mathbf{h}')}^{h,h'}}{\left(\overline{|U(\mathbf{h})|^2}\right)^{3/2}} = \varepsilon_3/\varepsilon_2^{3/2}\,. \quad (12)$$

This substitution is correct only when the range of indices is sufficient to produce well resolved maxima in a Fourier synthesis whose coefficients are the values of $U(\mathbf{h})$. Otherwise the left hand side of (12) should be used as it stands. Expressed in terms of normalized structure factors it is $\overline{E(\mathbf{h})E^*(\mathbf{h}')E^*(\mathbf{h}-\mathbf{h}')}^{h,h'}$. Whatever the limitations of the theory given in earlier sections, this quantity can in any case be taken to be a measure of the efficiency of phase relations since it measures the average degree of correlation between $E(\mathbf{h})$, $E(\mathbf{h}')$ and $E(\mathbf{h}-\mathbf{h}')$ (in such a way as to express the result as a real number). This quantity can also be expressed in terms of the electron distribution. Let

$$\varrho_s(r) = \frac{1}{v}\,\Sigma\, U(\mathbf{h}) \exp\,[-2\pi i\mathbf{h}.\mathbf{r}]\,,$$

where the sum is over n values of $\mathbf{h}$ contained in a limiting sphere of volume v_r in reciprocal space, so that $n = vv_r$, where v is the unit cell volume. Then

$$\int \varrho_s^2 dv = \frac{1}{v}\,\Sigma|U(\mathbf{h})|^2 = \frac{n}{v}\,\overline{|U(\mathbf{h})|^2}^h = v_r\overline{|U(\mathbf{h})|^2}^h$$

and

$$\int \varrho_s^3 dv = \frac{1}{v^2}\,\Sigma\Sigma\, U(\mathbf{h})U^*(\mathbf{h}')U^*(\mathbf{h}-\mathbf{h}') \propto v_r^2\overline{U(\mathbf{h})U^*(\mathbf{h}')U^*(\mathbf{h}-\mathbf{h}')}^{h,h'}\,.$$

The averages are now over a limited range of indices, viz., those contained in the limiting sphere of volume v_r. A 'proportional to' rather than an 'equals' sign has been used above since a factor, which is however independent of v_r, has been omitted.

Hence

$$\overline{E(\mathbf{h})E^*(\mathbf{h}')E^*(\mathbf{h}-\mathbf{h}')}^{h,h'} \propto \frac{\dfrac{1}{v_r^2}\displaystyle\int \varrho_s^3 dv}{\left(\dfrac{1}{v_r}\right)^{3/2}\left\{\displaystyle\int \varrho_s^2 dv\right\}^{3/2}}\,.$$

However, the average on the left-hand side at present includes terms $E(0)|E(\mathbf{h}')|^2$ which are always positive and in whose value we are not interested. They can be excluded by imagining the constant term $U(0)$ to be omitted from the Fourier synthesis, that is, by measuring the electron density from its average value, $\bar{\varrho}_s$. Thus, finally,

$$\overline{E(\mathbf{h})E^*(\mathbf{h}')E^*(\mathbf{h}-\mathbf{h}')}^{h,h'} \propto \frac{1}{v_r^{1/2}}\,\frac{\displaystyle\int (\varrho_s-\bar{\varrho}_s)^3 dv}{\left\{\displaystyle\int (\varrho_s-\bar{\varrho}_s)^2 dv\right\}^{3/2}} \qquad (13)$$

is a measure of the efficiency of phase relations, or of sign relations as the case may be, between those terms which lie within a limiting sphere of volume v_r. Phase relations, or sign relations, therefore work best for structures with high 'positive contrast', that is, a few regions of high electron density and wide regions of small, or constant, density. This condition may well be satisfied in certain projections of a protein crystal. Further implications of relation (13) can best be appreciated by considering a few one-dimensional (and for simplicity, centrosymmetric) examples.

The distribution shown in Fig. 3(a) would give $\overline{E(\mathbf{h})E(\mathbf{h}')E(\mathbf{h}-\mathbf{h}')}^{h,h'}$ the limiting value $+0{\cdot}33$. However, the average taken only over a limited range of indices (-6 to $+6$) sufficient merely to produce the distribution shown in Fig. 3(b), is actually negative, since $\int(\varrho_s-\bar{\varrho}_s)^3dv$ is then negative. Sign relations, in the usual sense, simply do not exist between structure factors in this range.

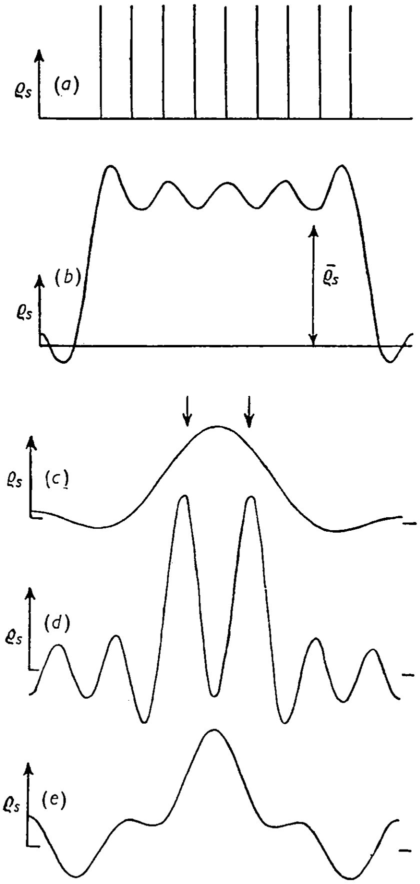

Fig. 3. (a) The limiting form of a one-dimensional distribution of nine equal atoms.

(b) The same distribution at low resolution. The 'hole' is now sharper than the 'peaks', giving negative contrast. Sign relations consequently break down.

(c) A distribution of two equal atoms, in positions marked by arrows, at low resolution.

(d) The same, at moderate resolution.

(e) An attempt to produce (d) with the structure factors which gave (c), extended in range by application of sign relations.

Next consider a distribution which at a moderate degree of resolution is as in Fig. 3(d), and at a low degree of resolution as in Fig. 3(c). While the value of $\overline{E(\mathbf{h})E(\mathbf{h}')E(\mathbf{h}-\mathbf{h}')}^{h,h'}$ is positive for both the limited and the extended range of structure factors, it is clear from considerations set out in Sayre's (1952) paper that, given the signs of all structure factors corresponding to Fig. 3(*c*), one cannot deduce from them enough further correct signs to produce Fig. 3(*d*). Only a sharpening of the original maximum results, and not its resolution into individual maxima, as is shown in Fig. 3(*e*). Such considerations must be kept in mind in any use of sign relations or phase relations when the data are not sufficient to resolve individual atoms.

I am grateful to my colleagues Dr F. H. Crick and Dr M. M. Woolfson for the helpful advice they have given in the course of discussions on this topic.

References

Bokhoven, C., Schoone, J. C. & Bijvoet, J. M. (1951). *Acta Cryst.* **4**, 275.
Cochran, W. & Woolfson, M. M. (1955). *Acta Cryst.* **8**, 1.
Hauptman, H. & Karle, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal.* A.C.A. Monograph No. 3. Wilmington: The Letter Shop.
Sayre, D. (1952). *Acta Cryst.* **5**, 60.
Watson, G. N. (1922). *A Treatise on the Theory of Bessel Functions.* Cambridge: University Press.
Wilson, A. J. C. (1949). *Acta Cryst.* **2**, 318.
Woolfson, M. M. (1954). *Acta Cryst.* **7**, 61.

Acta Cryst. (1955). **8**, 478

New Calculations of Atomic Scattering Factors

By J. Berghuis, IJbertha M. Haanappel and M. Potters
Mathematical Centre, Amsterdam, The Netherlands

and B. O. Loopstra*, Caroline H. MacGillavry and A. L. Veenendaal
Laboratory for General and Inorganic Chemistry, University of Amsterdam, The Netherlands

(*Received* 17 *February* 1955)

Scattering factors for twenty-three atoms have been calculated from Hartree and Hartree–Fock radial wave functions. The results are compared with James & Brindley's values and with those recently obtained by other authors.

1

Since the calculations of scattering factors by James & Brindley (1931) many new data on electron distributions in atoms, computed by the self-consistent field method, have become available. The work of James & Brindley (quoted as JB in the following) was some years ago extended by Viervoll & Ögrim (1949), who included electron distribution-data on Na^+, K^+, and Cu^+ and extended the $\sin\theta/\lambda$ range. Viervoll & Ögrim were primarily interested in the f values at higher $\sin\theta/\lambda$; they therefore applied only wave functions calculated without exchange, as the corrections for exchange would presumably have only small effects on the scattering factors in this region.

Recently, new values for scattering factors have been calculated by: (*a*) McWeeny (1951) (McW), for all atoms lighter than Na, from wave functions given by Duncanson and Coulson; (*b*) Hoerni & Ibers (1954) (HI) for C, N and O from newer available self-consistent field data with exchange; (*c*) Henry (1954), for Au^+ and Hg^{++}, from self-consistent field data without exchange.

It is evident from comparison of the results of HI and JB that the taking into account of exchange gives effects which make the corresponding effects on calculated structure factors outside the error limit for structure determination by accurate modern methods. Also, the interval in $\sin\theta/\lambda$ chosen by JB is rather too large for suitable interpolation in the low $\sin\theta/\lambda$ region.

In view of the preparation of Volume 3 of the new edition of the *International Tables for X-ray Crystallography*, it was decided to recalculate atomic scattering factors from all the newest available self-consistent field data. This decision had, in fact, been taken before we were aware of the work being done by HI. We therefore included C, N and O in our computations, which has the advantage of giving the f values for these atoms at the same intervals as for the other atoms, and, moreover, affords the possibility of an extra check.

In this paper we give our results on all atoms, up to Rb^+, for which the required electron distributions are available†. We omitted H and He, however, for which

---

* Now at J.E.N.E.R., Kjeller per Lillestrøm, Norway.

† Besides the literature mentioned at the foot of Table 1, we found references to calculations on Sc and Ti (Hancock, 1934), and on Ni (Gray & Manning, 1941). We were unable to obtain the former; the latter is a short communication without numerical data.

Manning & Goldberg (1938) gave data on Fe. These, however, are on a logarithmic scale of r, whereas the data for the other atoms were on a linear r scale. It is therefore impossible to handle the Fe atom in the same mechanical way as the others. We are now computing the scattering factor of Fe separately and we intend to report on it in due time.