AUGUST 15, 1938 PHYSICAL REVIEW VOLUME 54

Direct Determination of Crystal Structure from X-Ray Data

Melvin Avrami
University of Chicago, Chicago, Illinois

(Received May 23, 1938)

The empirical similarity in the behavior of scattering-factor curves for different atoms is used to replace the structure-factor formulae by an *approximating system.* This system is shown to lead to m'th degree algebraic equations, designated *fundamental equations,* whose coefficients are calculable in terms of crystal structure and atomic scattering factors. The roots of these equations give the coordinates of the unknown atomic positions. Thus if the values of a sufficient number ($2m$ in general; $m+1$ where there is a center of symmetry) of sequential structure-factors are known, the problem of direct structure determination is solved. If only the absolute magnitudes are known, one may derive reciprocal equations of the $m(m-1)$th degree whose coefficients are calculable in terms of the squared moduli of the structure factors. We can then write *fundamental equations* of degree $m(m-1)/2$ whose roots give the coordinates of the *interatomic* distances from which the unknown parameters themselves may readily be determined. If the number of unknown parameters in any projection is m, the minimum number of structure factor moduli required in this case is $m(m-1)+1$. The method developed in the paper is applied to the KH_2PO_4 crystal, and the results are found to compare very well with the values found by West on the basis of a trial-error and Fourier synthesis investigation.

In the problem of determining crystal structure from x-ray data, only the absolute values of the "crystal structure factors" F_{hkl}, associated with Laue indices hkl, are known.[1, 2] The F_{hkl} are related by theory to the distribution of scattering centers (atoms or ions) in the unit cell as follows:

$$F_{hkl}=\sum_{j=1}^{m} f_j(hkl)e^{-2\pi i(hx_j+ky_j+lz_j)}. \tag{1}$$

Here $f_j(hkl)$ is the scattering factor in the hkl direction of the j'th center, xyz are its coordinates in fractional parts of the unit axial ratios, and m is the number of scattering centers in the unit cell. The $f_j(hkl)$ may be taken as known from independent experimentation. From these and the absolute values of the F_{hkl} we seek to determine the x_jy_j and z_j. In the past crystallographers have resorted to methods of trial and error for solving (1). We shall here show how these repeated trials may be avoided and the structure obtained directly from the experimentally given quantities. This method can be seen to be connected with a quasi-direct method of Patterson,[3] but has, however, many advantages over the latter. A more detailed treatment will be given in a forthcoming paper in the *Zeitschrift für Kristallographie.*

I. The Approximating System

To carry through our treatment we make use of a simplification permitted by the experimental data. Examination of any table of atomic scattering factors[4] will show that the general variation at low temperatures of the $f_j(hkl)$ with hkl is similar for different atoms.[3] Thus to an approximation (which is better the lower the temperature and the larger the values of hkl) we may put $f_j(hkl)=f_jf(hkl)$ where f_j is a positive constant characteristic of atom j, and $f(hkl)$, the same for all atoms, may be taken as

$$\sum_{j=1}^{m} f_j(hkl)\Big/\sum_{j=1}^{m} f_j.$$

Thus (1) may be written

$$\frac{F_{hkl}}{f(hkl)}=\sum_{j=1}^{m} f_je^{-2\pi i(hx_j+ky_j+lz_j)} \tag{2}$$

and it is this *approximating system* which we shall solve for the $x_jy_jz_j$; these and the phases of the F's are the only unknowns. It can be shown that the quantities in (2) may be regarded as the formal Fourier coefficients of a step function whose highest order singularities (poles) coincide in position with the centers of scattering in the crystal.

It is possible to set up equations depending only on the (F/f)'s whose roots are the x_jy_j and

---

[1] W. L. Bragg and J. West, Zeits. f. Krist. **69**, 118 (1928).

[2] A. H. Compton and S. Allison, *X-Rays in Theory and Experiment.*

[3] W. L. Patterson, Zeits. f. Krist. **90**, 517 (1935).

[4] James and Brindley, Phil. Mag. **12**, 81 (1931).

z_j, respectively. These cannot be used in their simple forms to determine the $x_j y_j$ and z_j in the general case because the phases of the (F/f)'s are unknown. Nevertheless, as we shall show later, it is possible to eliminate the unknown phases and get equations entirely similar in form, so that it is advantageous to work out these equations and their solutions.

Suppose that it is the x_j which we wish to determine. Then keeping k and l fixed and setting

$$F_{hkl}/f(hkl)=A_h, \quad e^{-2\pi i x_j}=\alpha_j, \quad f_j e^{-2\pi i(ky_j+lz_j)}=d_j$$

we may write from (2)

$$A_h=\sum_{j=1}^{m} d_j \alpha_j{}^h. \tag{2$'$}$$

An m'th degree equation of which the α's are the roots may be determined as follows. Into

$$a_0+a_1\alpha+a_2\alpha^2+\cdots+a_m\alpha^m=0 \tag{3}$$

we successively substitute the various α's and multiply by the associated d's. Adding and using (2′) we find

$$a_0A_p+a_1A_{p+1}+\cdots+a_mA_{p+m}=0, \tag{4}$$

which may be regarded as a recursion formula for the A's applicable for all p. Choosing a set of m+1 successive equations of the form (4) we find

$$a_r=(-1)^r \begin{vmatrix} A_p & \cdots & A_{p+r-1} & A_{p+r+1} & \cdots & A_{p+m} \\ A_{p+1} & \cdots & A_{p+r} & A_{p+r+2} & \cdots & A_{p+m+1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ A_{p+m-1} & \cdots & A_{p+m+r-2} & A_{p+m+r} & \cdots & A_{p+2m-1} \end{vmatrix} \tag{5}$$

or (3) may also be written

$$\begin{vmatrix} 1 & \alpha & \cdots & \alpha^m \\ A_p & A_{p+1} & \cdots & A_{p+m} \\ \cdot & \cdot & \cdot & \cdot \\ A_{p+m-1} & A_{p+m} & \cdots & A_{p+2m-1} \end{vmatrix}=0. \tag{3$'$}$$

While (3′) may be evaluated for different commencing indices p and different fixed values of k and l, it is actually independent of these except perhaps for a common proportionality factor. (3) or (3′), and its analogs for the determination of y_j and z_j, may be designated *fundamental equations*. Once these are solved, the identification of the roots, i.e. the correct association of each α_j with its d_j, may be carried out by solving the linear system obtained upon selecting any m equations from the system (2′).

In cases where the *fundamental equations* are too high in degree to be solved by elementary methods, general procedures for the approximation of roots must be resorted to. With the aid of the recursion formula (4) a valuable method of this kind may be established. This simple method, which is applicable to any system of the form (2′), with α's general, and therefore to the solution of any algebraic equation (since such may always be associated with a system (2′)), will be described elsewhere.

II. Application to the Case of Unknown Phase Factors

(3′) cannot be used in its simple form to determine the $x_j y_j$ and z_j because the phase of the (F/f)'s, and hence of the A's, are unknown. In certain cases, i.e., where the distribution has a center of symmetry, the (F/f)'s are all real and the ambiguity is merely one of sign[1] which can often be determined by other methods. In such cases the distribution of scattering centers may be obtained independently of our method by a Fourier synthesis. There are also important cases where (as in KH_2PO_4; see below) the phases are known of a sequence of F's sufficient in number to determine the unknown parameters. In the most general case, however, we are faced with an ignorance of the phases and the necessity of eliminating them. To do this we multiply (2) by its conjugate complex. We then get

$$|F_{hkl}|^2/f^2(hkl)=\sum_{j=1}^{m} f_j{}^2+\sum_{\substack{j,\,s=1\\ j\neq s}}^{m} f_j f_s \times e^{-2\pi i[h(x_j-x_s)+k(y_j-y_s)+l(z_j-z_s)]}.$$

Introducing the abbreviations

$$G_{hkl}=|F_{hkl}|^2/f^2(hkl)-\sum_{j=1}^{m}f_j^2,$$

$$\begin{aligned} & & e^{-2\pi i(x_j-x_s)}&=\xi_{js}=\xi_w,\\ f_jf_s&=g_{js}=g_w, & e^{-2\pi i(y_j-y_s)}&=\eta_w,\\ m(m-1)&=n, & e^{-2\pi i(z_j-z_s)}&=\zeta_w,\end{aligned} \tag{6}$$

where $w=1, 2, \cdots n$ runs over the entire range for $j, s\ j\neq s$, we get

$$G_{hkl}=\sum_{w=1}^{n}g_w\xi_w{}^h\eta_w{}^k\zeta_w{}^l, \tag{7}$$

which is exactly the same as (2) in form, and may be treated in the same way to get *fundamental equations* whose roots are the ξ's, η's and ζ's. These roots occur in conjugate complexes, i.e. the *fundamental equations* are reciprocal and may be replaced by equations of half their degree. The roots give the coordinates of the *interatomic* distances rather than those of the atomic positions. The latter may easily be deduced from the former when the associated pair of atoms is identified by the procedure outlined above for (2).

Upon examination it may be seen that these roots correspond very closely to the positions of the peaks in Patterson's "sharpened up" series.[3] These positions, which Patterson's method can rarely locate because of the slow convergence of his series, are here determined in a simple algebraic way.

In case some of the positions of the scattering centers are known, their contributions may also be subtracted from $|F_{hkl}|^2/f^2(hkl)$ and an expression of the form (7) with n smaller than $m(m-1)$ will result.

When the experimental data are inaccurate, it will be advisable to get the coefficients of the *fundamental equations* by averaging the results from a number of the available sequences of F's (G's). Since both our assumptions and the experimental data are generally more valid for F's with higher indices, it will be wise to restrict ourselves to the use of these wherever possible. In the case of symmetry, $F_p=F_{-p}$, and the minimum number of sequential F's (including F_0) which must be known to determine an n'th degree *fundamental equation* is $n+1$.

Table I.

hkl	sin θ	F (corr.)	f_K+f_P	f_0	$\bar{G}$
060	0.248	40	10.8	2.3	−1.4
260	0.262	49	10.3	2.1	3.7
460	0.298	39	9.1	1.7	1.5
660	0.351	18	7.5	1.4	−8.6
860	0.413	20	6	1.05	−3.8
1060	0.482	24	4.4	0.75	8.5
1260	0.554	14	2.35	0.6	7.7

III. Application to KH_2PO_4

We shall now apply our method to the simple case of the KH_2PO_4 crystal which has been investigated very carefully by J. West.[5] The space group is $V_d{}^{12}$. The unit cell contains four molecules,

4 P atoms at $000;\ \frac{1}{2}0\frac{1}{4};\ \frac{1}{2}\frac{1}{2}\frac{1}{2};\ 0\frac{1}{2}\frac{3}{4}$
4 K atoms at $00\frac{1}{2};\ \frac{1}{2}0\frac{3}{4};\ \frac{1}{2}\frac{1}{2}0;\ 0\frac{1}{2}\frac{1}{4}$

and 16 O atoms in the general positions xyz; $y\bar{x}\bar{z}$; $\bar{x}\bar{y}z$; $\bar{y}x\bar{z}$; $(\frac{1}{2}-x), y, (\frac{1}{4}-z)$; $(\frac{1}{2}-y), \bar{x}, (z+\frac{1}{4})$; $(x+\frac{1}{2}), \bar{y}, (\frac{1}{4}-z)$; $(y+\frac{1}{2}), x, (z+\frac{1}{4})$ and these $+\frac{1}{2}\frac{1}{2}\frac{1}{2}$. It is the parameters xyz which we wish to determine.

With West we may neglect the effect of the eight H atoms. Then, considering first the hk_0 reflections, it is found that "Owing to the special positions occupied by the potassium and phosphorus atoms, the (hk_0) spectra reduce to two types . . . *viz.*

Type I. (h even) Contribution of atoms is potassium +phosphorus+oxygens.
[Type II.] (h odd) Contribution from oxygen atoms only.

The spectra of type II are all weak compared with those of type I which receive a full contribution from potassium and phosphorus. We conclude that the signs of the phase factors (and therefore of the structure amplitudes $F(hk_0)$) for type I spectra are governed by the potassium and phosphorus atoms. It follows that . . . we shall know the signs of . . . $F(hk_0)$, for all planes with h even though not with h odd."

The knowledge of the signs of a sufficient sequence of F's permits us, in this case, to determine a *fundamental equation* for the atomic coordinates themselves. The contribution to $F(hk_0)$, with h even, from **K** and **P** is $4[f_k(hk_0)+f_P(hk_0)]$. Subtracting this and dividing by

---

[5] J. West, Zeits. f. Krist. **74**, 306 (1930).

$f_0(hk_0)$ we get a set of numbers, $\bar{G}_{hk_0}$, which, from (8), can be seen to be equal to

$$\bar{G}_{hk_0}=16[\cos 2\pi ky(e^{2\pi ihx}+e^{-2\pi ihx}) \\ +\cos 2\pi kx(e^{2\pi ihy}+e^{-2\pi ihy})]. \quad (9)$$

Thus, keeping k constant, and varying h over even values, we see that we have a system, with a center of symmetry, in which the roots are $e^{4\pi ix}$, $e^{4\pi iy}$ and their conjugate complexes. The *fundamental equation* for the determination of these roots is then a reciprocal equation of the fourth degree. To determine its coefficients we here use only one of the available sequences of F's since this yields adequate accuracy. The values, as calculated from F_{h60} and f curves given by West, are given in Table I. These values yield

$$\begin{vmatrix} 1 & \alpha & \alpha^2 & \alpha^3 & \alpha^4 \\ 3.7 & -1.4 & 3.7 & 1.5 & -8.6 \\ -1.4 & 3.7 & 1.5 & -8.6 & -3.8 \\ 3.7 & 1.5 & -8.6 & -3.8 & 8.5 \\ 1.5 & -8.6 & -3.8 & 8.5 & 7.7 \end{vmatrix} =$$

$$2130\alpha^4-608\alpha^3+3178\alpha^2+211\alpha+2020=0. \quad (10)$$

Since this *fundamental equation* should be reciprocal and the symmetric coefficients equal, i.e., $a_0=a_4$, $a_1=a_3$, we may average these. Doing this, and dividing through by the first (last) coefficient, we have

$$\alpha^4-0.42\alpha^3+1.5\alpha^2-0.42\alpha+1=0. \quad (10')$$

This reciprocal equation may be solved by elementary methods and the roots are found to be

$$e^{\pm 4\pi ix}=0.475\pm 0.88i, \quad e^{+4\pi iy}=-0.265\pm 0.965i,$$

or in degrees $x\approx 31°$, $y\approx 53°$,

which agrees well with West's values of 29° and 52°. Once x and y are known, the third parameter is readily found in the manner given by West.

The remarks of West concerning the trial-error and double Fourier series methods apply with equal force to our procedure. "The extent to which these methods are applicable without modification to complex structures must depend on the structures e.g. whether some of the parameters of the heavier atoms are fixed by symmetry considerations. In any case, however, it should not often be necessary to use a laborious trial and error method throughout to determine completely the structure. By the aid of certain reasonable assumptions and a qualitative examination of the experimental data a complex structure may often be partially solved. This partial solution may well be sufficient to determine the signs of a considerable portion of the F values for the reflections from planes in some important zone."

In conclusion we may make one general remark. Our method bears the same relation to the ordinary trial and error method as the systematic solution of a polynomial equation (e.g. the *fundamental equation*) bears to a cut and try attempt to determine its roots from linear combinations of homogeneous powers of these roots (e.g. 2′).

The author wishes to acknowledge his indebtedness for helpful advice and suggestion to Dr. W. H. Zachariasen and Dr. W. Bartky of the University of Chicago.