series. The procedure is a generalization of the one described for two terms in one dimension.*

Each term in the Fourier series determines an index $i$ by means of (14). For a fixed $\mu$ and each such $i$, the number $m_i$ is defined to be that integer which is nearest to

$$h_i x_\mu + k_i y_\mu + l_i z_\mu - \eta_i \,. \tag{35}$$

Consequently, the plane

$$h_i x + k_i y + l_i z = m_i + \eta_i \tag{36}$$

is that plane of the family (22) which passes nearest to the point $x_\mu, y_\mu, z_\mu$, i.e. the maximum of

$$C_i \left\{ \begin{array}{l} \cos 2\pi (h_i x + k_i y + l_i z) \\ \sin 2\pi (h_i x + k_i y + l_i z) \end{array} \right\}$$

(the cosine or sine is used according as $C_i = |C_{h_i k_i l_i}|$ or $C_i = |C'_{h_i k_i l_i}|$) closest to $x_\mu, y_\mu, z_\mu$ is in the plane (36). As in one dimension, the term

$$C_i \left\{ \begin{array}{l} \cos 2\pi (h_i x + k_i y + l_i z) \\ \sin 2\pi (h_i x + k_i y + l_i z) \end{array} \right\}$$

is replaced by its Taylor expansion in the region of the plane (36) for each $i$, and the results substituted into (12). The unique maximum of the resulting function of the second degree is readily found by the standard method of partial differentiation to be the solution, $x, y, z$, of

$$\left. \begin{array}{l} x\sum\limits_i C_i h_i^2 + y\sum\limits_i C_i h_i k_i + z\sum\limits_i C_i h_i l_i = \sum\limits_i C_i h_i (m_i + \eta_i) \,, \\ x\sum\limits_i C_i k_i h_i + y\sum\limits_i C_i k_i^2 + z\sum\limits_i C_i k_i l_i = \sum\limits_i C_i k_i (m_i + \eta_i) \,, \\ x\sum\limits_i C_i l_i h_i + y\sum\limits_i C_i l_i k_i + z\sum\limits_i C_i l_i^2 = \sum\limits_i C_i l_i (m_i + \eta_i) \,, \\ i = 1, 2, 3, \ldots \,. \end{array} \right\} \tag{37}$$

Using the solution $x, y, z$ of (37) instead of $x_\mu, y_\mu, z_\mu$, the refinement procedure may be repeated to yield a still better approximation to the location of the maximum, and the cycle may be repeated again and again. This iterative process ordinarily converges within ten cycles. The final values of the triples $x, y, z$, so obtained, as $\mu$ ranges through the values 1, 2, 3, ..., are the coordinates of the most prominent maxima of (12), arranged approximately in decreasing order.

The procedure described herein has been programmed for I.B.M. equipment by Mr Peter O'Hara of the Computation Laboratory of the National Bureau of Standards. His excellent cooperation is deeply appreciated.

### References

Dickson, L. E. (1936). *Introduction to the Theory of Numbers*. Chicago: University of Chicago Press.

Hauptman, H. & Karle, J. (1952). *Acta Cryst.* **5**, 48.

---

* The refinement procedure may be applied no matter how the approximate locations $x_\mu, y_\mu, z_\mu$ have been obtained.

---

*Acta Cryst.* (1953). **6**, 473

# Application of Statistical Methods to the Naphthalene Structure

By J. Karle and H. Hauptman

*U.S. Naval Research Laboratory, Washington 25, D.C., U.S.A.*

(*Received* 7 *November* 1952)

A method for locating the principal maxima of a Fourier series was applied to a function describing the probability distribution for interatomic vectors using the data of Abrahams, Robertson & White for naphthalene. A structure comparable to that found by Abrahams *et al.* was obtained. No attempt was, however, made to obtain the ultimate accuracy inherent in this method.

## Introduction

This paper concerns the application of the statistical methods developed in a previous paper (Hauptman & Karle, 1952) in order to determine the structure of the carbon frame in naphthalene. The principal maxima of formula (58) (Hauptman & Karle, 1952) expressing the probability distribution for interatomic vectors, have been located by a method described previously (Hauptman & Karle, 1953). The validity of the results therefore constitutes a test not only of the statistical method, but also of the method for locating the principal maxima of a Fourier Series. Since the X-ray scattering data of Abrahams, Robertson & White (1949) were used, a comparison with their results is significant.

## Treatment of data

The logarithm of equation (58) (Hauptman & Karle, 1952) may be expressed as the Fourier series

$$\sum_{hkl} C_{hkl} \cos 2\pi(hx+ky+lz) =$$

$$2\sum_{hkl}\left\{[A-\sqrt{(1-B^2)}]\left[\frac{1/B}{\sqrt{(1-B^2)}}-1/B\right]\cos 2\pi(hx+ky+lz)\right.$$

$$-[A-\tfrac{1}{2}\sqrt{(1-B^2)}]\left[\frac{2/B^2-1}{\sqrt{(1-B^2)}}-2/B^2\right]\cos 4\pi(hx+ky+lz)$$

$$+[A-\tfrac{1}{3}\sqrt{(1-B^2)}]\left[\frac{4/B^3-3/B}{\sqrt{(1-B^2)}}-(4/B^3-1/B)\right]$$

$$\times \cos 6\pi(hx+ky+lz)$$

$$-[A-\tfrac{1}{4}\sqrt{(1-B^2)}]\left[\frac{8/B^4-8/B^2+1}{\sqrt{(1-B^2)}}-(8/B^4-4/B^2)\right]$$

$$\times \cos 8\pi(hx+ky+lz)$$

$$+[A-\tfrac{1}{5}\sqrt{(1-B^2)}]\left[\frac{16/B^5-20/B^3+5/B}{\sqrt{(1-B^2)}}\right.$$

$$\left.\left.-(16/B^5-12/B^3+1/B)\right]\cos 10\pi(hx+ky+lz)-\ldots\right\}, \quad (1)$$

where $A = A(h, k, l)$ is a function of the observed intensity $|F_{hkl}|^2$ adjusted by a procedure to be described below, $B=B_{ij}(h, k, l)=2f_i(h, k, l)f_j(h, k, l)/\sigma_2$, $\sigma_2 = \sum_{\nu=1}^{N} f_\nu^2(h, k, l)$, $f_\nu(h, k, l)$ is the atomic scattering factor of the $\nu$th atom and $N$ is the total number of atoms per unit cell. The first, second, third, fourth or fifth terms of the right side of (1) will contribute to $C_{hkl}$ depending on which ones of 1, 2, 3, 4 or 5 divide the greatest common divisor of $h$, $k$, and $l$. For convenience, the functions of $B$ appearing in (1) were tabulated to four significant figures in the range $0{\cdot}000 \leq B \leq 0{\cdot}500$, $\Delta B = 0{\cdot}001$. In any given case, therefore, the coefficients of the cosine terms appearing in (1) are linear functions of $A$ with coefficients that have been tabulated.

The atomic scattering factors $f_\nu(s)$, where $s=\sin\theta/\lambda$, have been tabulated in the range $0{\cdot}000 \leq s \leq 1{\cdot}500$, $\Delta s = 0{\cdot}001$, by interpolation from existing tables. The observed intensities $I_0(h, k, l)$ were expressed as functions of $s$ and the entire interval of $s$ was divided into 9 sub-intervals, each corresponding to 100 intensities (except the last which contained only 32). Since $\langle I(s)/\varepsilon\sigma_2(s)\rangle = 1$* (Wilson, 1949), where $I$ is the value of the intensity adjusted for scale and vibrational motion, the expression

$$K_j = \frac{\sum\limits_{s_{j-1}\leq s<s_j} \varepsilon\sigma_2(s)}{\sum\limits_{s_{j-1}\leq s<s_j} I_0(s)}, \; j = 1, 2, \ldots, 9, \qquad (2)$$

is an average factor by which the observed intensities in each sub-interval could be multiplied in order to obtain the adjusted values. In practice, however, each of the 9 values of $K$ obtained from (2) is plotted at the mid-point of the corresponding interval and a smooth curve $K(s)$ is drawn among the points, giving a scale factor for each observed intensity. This procedure not only corrects for vibrational motion but puts the data on an absolute scale. The curve obtained for naphthalene appears in Fig. 1. It is noteworthy that this curve passes through unity at $s = 0$, showing that the procedure of Robertson for adjusting to an absolute scale is in very good agreement with the statistical method. It is now possible to define $A(h, k, l)$ of equation (1) by means of the adjusted intensities,

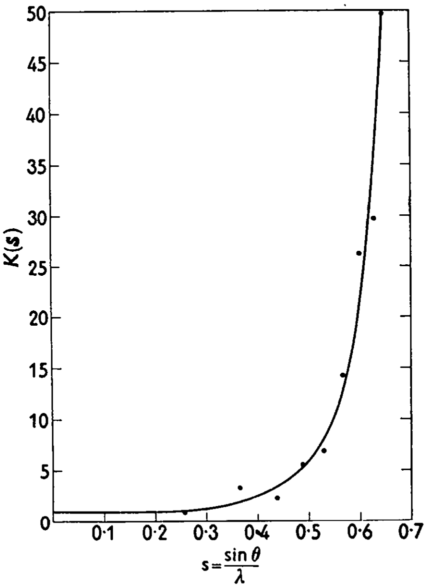

Fig. 1. The function $K(s)$.

$$A = A(h, k, l) = I(h, k, l)/\sigma_2(h, k, l)\,. \qquad (3)$$

Using these values of $A$ and the tabulated values of $B$, the coefficients of Fourier series (1) were computed. With these values of the coefficients $C_{hkl}$, the procedure for locating the principal maxima of (1) was carried out.

The reference mesh (Hauptman & Karle, 1953) was defined by the three families of planes

Table 1. *A row in the interatomic vector matrix corresponding to the two molecules in the unit cell*

The $m$ triples determine points on the reference mesh closest to maxima of (1)

| | Molecule I | | | Atom | Molecule II | | |
|---|---|---|---|---|---|---|---|
| Atom | $m_1$ | $m_2$ | $m_3$ | | $m_1$ | $m_2$ | $m_3$ |
| $A$ | 2 | 2 | 7 | $A_1$ | $\overline{11}$ | $\overline{9}$ | $\overline{7}$ |
| $B$ | 3 | 5 | 6 | $B_1$ | $\overline{11}$ | $\overline{5}$ | $\overline{4}$ |
| $C$ | 2 | 4 | 2 | $C_1$ | $\overline{9}$ | $\overline{5}$ | $\overline{1}$ |
| $D$ | 2 | 8 | 19 | $D_1$ | 9 | 16 | 2 |
| $E$ | 1 | 7 | 16 | $E_1$ | $\overline{9}$ | $\overline{4}$ | $\overline{15}$ |
| $A'$ | 0 | $\overline{18}$ | $\overline{6}$ | $A_1'$ | $\overline{7}$ | $\overline{9}$ | $\overline{13}$ |
| $B'$ | $\overline{1}$ | $\overline{1}$ | $\overline{3}$ | $B_1'$ | 12 | 10 | 5 |
| $C'$ | 0 | 0 | 0 | $C_1'$ | 12 | 10 | 2 |
| $D'$ | 0 | $\overline{4}$ | $\overline{18}$ | $D_1'$ | $\overline{7}$ | $\overline{13}$ | $\overline{1}$ |
| $E'$ | 0 | 18 | 6 | $E_1'$ | $\overline{8}$ | $\overline{12}$ | $\overline{4}$ |

---

* $\varepsilon = \varepsilon_{hkl} = 2$ if $k = 0$ or $h = l = 0$; otherwise $\varepsilon = 1$.

Table 2. *Approximate trimetric coordinates obtained from the triples of molecule* I *in Table* 1 *and two refinements based on two different origins in the molecule*

For purposes of comparison all coordinates were related to the same origin after the refinement

| | Approximate | | | Refined ($I_a$) | | | Refined ($I_b$) | | |
|---|---|---|---|---|---|---|---|---|---|
| | *x* | *y* | *z* | *x* | *y* | *z* | *x* | *y* | *z* |
| *A* | 0·1041 | 0·0828 | 0·3448 | 0·1182 | 0·1145 | 0·3449 | 0·1232 | 0·1136 | 0·3651 |
| *B* | 0·1618 | 0·2354 | 0·2919 | 0·1558 | 0·2670 | 0·2470 | 0·1665 | 0·2587 | 0·2655 |
| *C* | 0·1098 | 0·1953 | 0·0945 | 0·0930 | 0·1989 | 0·0778 | 0·0930 | 0·1989 | 0·0778 |
| *D* | 0·1176 | 0·3528 | −0·0559 | 0·1202 | 0·3456 | −0·0446 | 0·1263 | 0·3500 | −0·0264 |
| *E* | 0·0655 | 0·3102 | −0·2033 | 0·0685 | 0·2977 | −0·2067 | 0·0628 | 0·2842 | −0·2095 |
| *A′* | −0·0442 | 0·1149 | −0·2978 | −0·0325 | 0·0847 | −0·2908 | −0·0208 | 0·0883 | −0·2673 |
| *B′* | −0·0521 | −0·0426 | −0·1474 | −0·0734 | −0·0601 | −0·1873 | −0·0628 | −0·0681 | −0·1692 |
| *C′* | 0·0000 | 0·0000 | 0·0000 . | 0·0000 | 0·0000 | 0·0000 | 0·0000 | 0·0000 | 0·0000 |
| *D′* | −0·0078 | −0·1550 | 0·1004 | −0·0337 | −0·1508 | 0·1035 | −0·0276 | −0·1471 | 0·1221 |
| *E′* | 0·0442 | −0·1149 | 0·2978 | 0·0325 | −0·0847 | 0·2908 | 0·0328 | −0·0864 | 0·2938 |

$$\left.\begin{aligned} 20x-y &= m_1\,, \\ 20y+z &= m_2\,, \\ x+20z &= m_3\,. \end{aligned}\right\} \qquad (4)$$

This mesh is sufficiently fine to resolve the maxima but not so fine that the rough location of the maxima requires an undue amount of computing. It also leads to greatest common divisors (Hauptman & Karle, 1953, equation (29) ff.) which rarely exceed unity. Using the largest two hundred values of $C_{hkl}$ (Hauptman & Karle, 1953, equation (12) ff.), the one thousand largest maxima of the Fourier series were obtained, arranged roughly according to size. These one thousand points could have been further refined and reweighted according to the procedure already described. However, this would have been prohibitively expensive and time-consuming for the procedure as developed for I.B.M. equipment. Instead, it was assumed that the bonded C–C distance was between 1·1 and 1·7 Å and that the molecular structure was roughly the planar, double-hexagon shape. These assumptions, together with the assumption of crystal symmetry, were sufficient to lead to a *unique* choice of twenty points* defined by their *m* triples, (4), shown in Table 1. To the approximation involved, the orientations of the two molecules were the same as those reported by Abrahams *et al.* (1949). The coordinates corresponding to the ten *m* triples of Table 1 related to molecule I are listed in the first three columns of Table 2. The next three columns were obtained from the approximate coordinates by the refinement procedure described in the previous paper (Hauptman & Karle, 1953, equation (37)). By selecting a new origin in molecule I, giving another row in the interatomic vector matrix, and applying the refinement procedure to the coordinates referred to the new origin, the last three columns of Table 2 were obtained.

The discrepancies between the two refinements are a consequence of the limitations of the method and the inaccuracies of the data. Table 3 shows a comparison of the bonded distances obtained from the approximate and refined coordinates of Table 2. The average deviations show clearly the improvement resulting from the refinement. The results for molecule II are similar to those which have been presented for molecule I.

Table 3. *Bonded distances obtained from the approximate coordinates in Table* 2 *and the averages of the two sets of refined coordinates in Table* 2

The values of the lattice parameters which were used were $a = 8{\cdot}235$, $b = 6{\cdot}003$, $c = 8{\cdot}658$ Å and $\beta = 122°\ 55'$ (Abrahams *et al.*, 1949). All values in Ångström units.

| | Approximate | Refined |
|---|---|---|
| *AB* | 1·229 | 1·397 |
| *BC* | 1·538 | 1·380 |
| *CD* | 1·639 | 1·445 |
| *DE* | 1·133 | 1·343 |
| *EA′* | 1·434 | 1·398 |
| *A′B′* | 1·639 | 1·420 |
| *B′C′* | 1·133 | 1·380 |
| *C′D′* | 1·298 | 1·444 |
| *D′E′* | 1·538 | 1·397 |
| *E′A* | 1·265 | 1·352 |
| *CC′* | 1·434 | 1·380 |
| Av. | 1·389 | 1·394 |
| Av. dev. | 0·161 | 0·025 |

It would have been of interest to refine the coordinates in molecule I referred to each of the ten possible origins in molecule I, and similarly with molecule II. The average of these results would be expected to be sufficiently accurate to permit a more significant comparison with the results of Abrahams *et al.* (1949). This was not done, however, because of the expense and time involved, as well as the fact that recent developments in the application of statistical methods to the direct determination of phase hold promise of more efficient procedures.

## Concluding remarks

The results of this paper indicate that the statistical methods and the procedure for locating the principal

---

* The 400 interatomic vectors based upon these twenty triples were computed and found among the 1000 largest maxima. All but 60 were found among the first 500 maxima. 130 occurred exactly; 158 occurred with one *m* value in error by unity; 96 occurred with two *m* values in error by unity; and 16 occurred with all three *m* values in error by unity. This distribution is to be expected on the basis of a simple statistical argument.

maxima of a Fourier series are sound. It is also clear that formula (1) is capable of resolving the maxima corresponding to interatomic vectors.

The statistical procedure in vector space constitutes a direct method for attacking structure problems. The validity of these statistical concepts, coupled with the attractiveness of working directly in coordinate space, indicate the direction of future developments, namely, the search for probability distributions for the phases of the structure factors as well as for the atomic coordinates, rather than the interatomic vectors.

The computations of this paper were performed by Mr Peter O'Hara of the Computation Laboratory of the National Bureau of Standards. His cooperation is gratefully acknowledged.

## References

ABRAHAMS, S. C., ROBERTSON, J. J. & WHITE, J. G. (1949). *Acta Cryst.* **2**, 233.

HAUPTMAN, H. & KARLE, J. (1952). *Acta Cryst.* **5**, 48.

HAUPTMAN, H. & KARLE, J. (1953). *Acta Cryst.* **6**, 469.

WILSON, A. J. C. (1949). *Acta Cryst.* **2**, 318.

---

*Acta Cryst.* (1953). **6**, 476

# A Three-dimensional Coordinate Model for Demonstration of Inorganic Crystal Structures

BY A. J. E. WELCH

*Department of Inorganic and Physical Chemistry, Imperial College, London S.W. 7, England*

(*Received* 15 *January* 1953 *and in revised form* 29 *January* 1953)

A transparent plastic framework is described in which coloured pins may be inserted to show the arrangements of atoms in simple unit cells.

Numerous types of three-dimensional model have been devised to show arrangements of atoms in crystal lattices, but few are quickly adaptable to the study of symmetry or coordination relationships in a wide variety of structures. Most X-ray crystallographers have at some time coveted a three-dimensional blackboard.

A model of the type shown in Fig. 1 has proved useful, in both research and teaching, for setting out atom positions in simple inorganic structures, particularly those belonging to the cubic and tetragonal systems. It comprises sheets of clear 'Perspex', $\frac{1}{8}$ in. in thickness, each having a coordinate 'net' of small holes drilled at suitable intervals. Successive sheets are held at regularly spaced intervals by nuts and washers on threaded brass rods; the rods are conveniently mounted in a heavy base-plate of thick 'Sindanyo' sheet. Atom centres are marked by brightly coloured map-marking pins dropped into the holes; such pins are available in a variety of sizes, shapes, and colours from large stationers. The pins are conveniently dropped into position by means of a long pair of tweezers. The particular model shown in Fig. 1 has a cubic array of holes, and was originally designed to depict a single unit cell of the spinel structure. The holes are drilled to a square pattern at 1 in. intervals, and the upper surfaces of successive sheets are 1 in. apart. Nine sheets, each containing nine rows of nine holes, provide all the coordinates required to set up the ideal spinel cell, in which all the atom coordinates are multiples of $\frac{1}{8}$. Many simpler cubic cells can be shown on the same framework. Tetragonal cells can be accommodated by adjusting the spacing of the sheets to the required $c/a$ ratio, the $c$ axis being set perpendicular to the plane of the sheets.

Another similar model has been constructed for hexagonal (including rhombohedral) structures, having the sheets drilled to provide sixfold symmetry; the sheets again contain nine rows of holes in each of two directions 120° apart, allowing atom coordinates to be set off in multiples of $\frac{1}{8}$ in the conventional hexagonal cell.

The models are particularly convenient in showing coordination relationships clearly. They are proving useful in research in the examination of ionic environments in ferromagnetic solids; possible modes of interaction between ions can be visualized much more clearly from such models than from diagrams or lists of coordinates. In teaching work numerous typical structures can be built up at little expense. Random and ordered replacement in solid solutions are readily contrasted on the lecture bench by interchanging a few coloured pins. By running wires or coloured threads through rows of holes to represent symmetry axes, the operation of the different types of axes (including screw axes) is effectively demonstrated; pins are in-