

will be about $\frac{1}{2} + 6/\sqrt{N}$. (The fact that this probability may exceed unity is noted by Karle & Hauptman.) The point we wish to make is that as N increases, even the highest values of $P_+(F_{2h})$ tend to $\frac{1}{2}$. A similar conclusion can be reached from the 'Patterson function' point of view which we outlined above. It is easily shown that the situation for the determination of F_{2h} as negative is even less favourable.

We conclude that relation (1) is of value only for very simple crystal structures.

At first sight equation (4.06) looks more promising, since Σ_4 may contain a large number of terms. It might appear that although the indication of sign given by a single term of Σ_4 was weak, a large number of such indications would be statistically significant. We shall therefore consider the case where Σ_4 contains an infinite number of terms.

Then, in our notation,

$$\begin{aligned} s(\Sigma_4) &= s\left\{ \sum_{\mathbf{h}'} (F_{\mathbf{h}'}^2 - N)(F_{\mathbf{h}+\mathbf{h}'}^2 - N) \right\} \\ &= s\left\{ \sum_{\mathbf{h}'} (F_{\mathbf{h}'}^2 F_{\mathbf{h}+\mathbf{h}'}^2 - N^2) \right\}, \end{aligned}$$

since $\overline{F_{\mathbf{h}}^2} = N$.

Then, by (4.06),

$$s(F_{2h}) \approx s\left\{ \sum_{\mathbf{h}'} (F_{\mathbf{h}'}^2 F_{\mathbf{h}+\mathbf{h}'}^2 - N^2) \right\}. \quad (2)$$

It may be shown that $\sum_{\mathbf{h}'} F_{\mathbf{h}'}^2 F_{\mathbf{h}+\mathbf{h}'}^2$ is the \mathbf{h} th Fourier coefficient of P^2 , the squared Patterson function, while ΣN^2 is the contribution to this coefficient of the origin peak of P^2 . We conclude that (2) is then true to the extent that all peaks in P^2 may be ignored except that at the origin and those at $\pm 2\mathbf{r}_j$. In this way the correlation between $s(F_{2h})$ and Σ_4 is seen to be even weaker than that between the former and Σ_1 . The effect of squaring the Patterson function is to decrease, to one quarter, the relative weight of the peaks at $2\mathbf{r}_j$, compared with the peaks at $(\mathbf{r}_i - \mathbf{r}_j)$. This situation cannot be altered by increasing the number of terms in Σ_4 .

Acta Cryst. (1954). **7**, 451

The statistical approach of Hauptman & Karle to the phase problem. By V. VAND and R. PEPIN-SKY, *X-ray and Crystal Analysis Laboratory, Department of Physics, The Pennsylvania State University, State College, Pa., U.S.A.*

(Received 4 March 1954)

The statistical approach to X-ray structure analysis has been treated by Hauptman & Karle (1953), who derived a number of formulae for the probabilities that certain structure factors are positive. However, a peculiar situation arises from their analysis. Although their mathematical approach is valid, and the formulae for the determining quantities Σ_n are correct within the limitations of their derivation, the conclusions, that a practical solution of the phase problem has been attained for all centrosymmetrical structures—provided a sufficient number of structure factors is available—cannot be maintained.

We have constructed more accurate distribution functions using the *Tables* of Vand (1953), and have also formed a simpler mathematical derivation of the statistical formulae. The true distribution functions yield not

Equation (4.05) has not been considered in any detail but it is clear that it, too, embodies the inherent faults of (4.03) and (4.06).

It will also be seen that, even if a number of signs can be determined for the set of indices for which h , k and l are all even, the extension to the other sets of terms will be very tenuous. Only (4.04) remains for this purpose, if we do not take the discredited (4.05) into account, and we have seen that this is equivalent to applying the normal sign relationship $s(\mathbf{h}) = s(\mathbf{h}')s(\mathbf{h}+\mathbf{h}')$. But the only additional information will be the three arbitrary signs which define the choice of origin. The determination of the next sign for a coefficient for which h , k and l are not all even will depend on one sign relationship! There would then possibly be two sign relationships for the next sign determination, and so on. It is our experience that attempts to build a knowledge of signs on too narrow a base inevitably fail unless the crystal structure problem is a very simple one.

We thus conclude that the solution of the phase problem for the centrosymmetrical crystal given by Hauptman & Karle offers nothing new for solving complex structures. While our criticisms have been directed to the procedure suggested for space group $P\bar{1}$, the simple 'Patterson function' interpretation also shows the weakness of that suggested for the only other space group considered by Hauptman & Karle in detail, namely $P2_1/a$.

The example which they have worked out, naphthalene, is far too obliging in many respects. The high symmetry gives extremely large unitary structure factors and there is no doubt that this structure can be solved by a number of previously published direct methods.

References

- COCHRAN, W. (1953). *Acta Cryst.* **6**, 810.
 HARKER, D. & KASPER, J. S. (1948). *Acta Cryst.* **1**, 70.
 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.

only probabilities, but also the inequalities of Harker & Kasper. The Hauptman & Karle formulae are confirmed as approximations when the number of atoms is large and the magnitudes of structure factors involved are small.

However, when the formulae are applied to a structure of space group $P\bar{1}$, the following difficulties arise: Only the equations (4.03) and (4.06), as numbered by Hauptman & Karle, can be used in the initial step, since these alone contain squares of structure factors only. For N equal atoms these become

$$\begin{aligned} \Sigma_1 &= \frac{1}{4\sqrt{N}} (E_\mu^2 - 1) && \text{for } h = 2h, \\ \Sigma_4 &= \sum \frac{1}{8N\sqrt{N}} (E_\mu^2 - 1)(E_\nu^2 - 1) && \text{for } h = 2h_\mu \pm 2h_\nu, \end{aligned}$$

where E is the structure factor on a scale $\langle E^2 \rangle = 1$.

Turning first to Σ_1 : it is immediately apparent that this is equivalent to the coefficient of a sharpened Patterson with the peak at the origin removed. It is well known that a Patterson of a structure having a center of symmetry contains a set of 'rotational' peaks corresponding to the vectors *across* the center of symmetry. These peaks correspond to an image of the structure on twice the scale. A detailed calculation (Vand & Pepinsky, 1953, 1954) indicates that the Hauptman-Karle statistical formulae are simply making use of the above property of the Patterson function in order to obtain the bias for the probability of sign of a structure factor. It unfortunately follows that when the statistical formulae are used *jointly*, nothing else than a Patterson is obtained, *whatever the number of structure factors used*. Therefore the solution converges to the highest peaks of the Patterson, and the statistical method becomes successful only when the highest peaks of the Patterson resemble the image of the structure on twice the scale. This seldom occurs in practice, however, since the wanted *rotational* peaks have half the weight of the *non-rotational* peaks. It follows that the only structures which can be solved in space group $P\bar{1}$ are those with two atoms per cell. We have tried to apply the method to a four-atom structure, and have obtained—as theoretically predicted—an incorrect solution resembling the Patterson rather than the structure itself.

Turning now to Σ_4 : we see that this is equivalent to a convolution in reciprocal space, and therefore to a product in Patterson space. This product turns out to be nothing else than a square of Patterson density. This sharpens the Patterson peaks, but brings out the unwanted peaks much higher than the wanted rotational peaks. It therefore can be concluded that if Σ_1 leads jointly to an incorrect structure, then Σ_4 certainly leads jointly to an incorrect structure. The claim that the formulae represent a solution for the space group $P\bar{1}$ is therefore fallacious.

The above considerations need to be modified when other elements of symmetry are present. We have shown (Vand & Pepinsky, 1953) that the statistical formulae then become equivalent to Patterson-Harker sections. If these sections resemble the image of the structure on twice the scale, the statistical formulae become successful; but the Harker sections then become so, too. This is the case with naphthalene, used as an example of the power of their method by Hauptman & Karle. In this respect, their method does not represent anything of greater power

than that of previously known methods. On the contrary, the intelligent use of Harker maps permits one to distinguish and allow for non-Harker peaks, whereas the statistical method conceals their influence in the mathematics involved.

This note is not meant to belittle entirely the value of the formulae derived by Hauptman & Karle. In fact, their approach has led us to derivations of maps which are superior to Harker maps for the purposes of structure determination. For example, one of the functions for space group $P2_1/c$ is a map having

$$U_{2h,0,2l} \sum_k (-1)^{k+l} (E_{hkl}^2 - 1)$$

as Fourier coefficients. This map shows non-Harker peaks considerably suppressed. Details of the derivations and properties of these functions have been made available elsewhere, along with a full account of all the above matters (Vand & Pepinsky, 1953).

Our conclusions are in full agreement with those obtained independently by Cochran & Woolfson (1954), which appear elsewhere in this issue of *Acta Crystallographica*. We are grateful to these authors for permitting us to see their communication before its publication.

We wish to acknowledge support of this analysis by the National Institute of Health, Institute for Arthritis and Metabolic Diseases, under Grant No. A-228. Computations of illustrative functions on X-RAC were supported by Contract No. N6onr-26916, T. O. 16, with the Office of Naval Research.

References

- COCHRAN, W. & WOOLFSON, M. M. (1954). *Acta Cryst.* **7**, 450.
 HAUPTMAN, H. & KARLE, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal*. American Crystallographic Association Monograph No. 3. Wilmington: The Letter Shop.
 VAND, V. (1953). *Tables for Direct Determination of Crystal Structures*. Glasgow: Chemistry Department of the University.
 VAND, V. & PEPINSKY, R. (1953). *The Statistical Approach to X-Ray Structure Analysis*. State College: X-ray and Crystal Analysis Laboratory of the Pennsylvania State University.
 VAND, V. & PEPINSKY, R. (1954). *Bull. Amer. Phys. Soc.* **29**, No. 3, 37.

Acta Cryst. (1954). **7**, 452

A note on the solution of the phase problem. By H. HAUPTMAN and J. KARLE, *U.S. Naval Research Laboratory, Washington 20, D.C., U.S.A.*

(Received 30 March 1954)

A routine solution of the phase problem, valid for all 92 centrosymmetric space groups, is described in our Monograph (Hauptman & Karle, 1953). The evidence for its validity we felt was overwhelming. The mathematical development is exact and rigorous, and is not challenged in the above notes (Cochran & Woolfson, 1954; Vand & Pepinsky, 1954). The probability theory yields the correct

structure invariants and seminvariants. The final formulas are capable of independent justification by means of a simple geometric interpretation. Finally, the application of these formulas to the naphthalene structure was successful. We consider this application non-trivial because no advantage was taken of the special properties of this simple structure, and only a small fraction of the