Acta Cryst. (1966). **21**, 273

Isomorphous Substitution and Formulas for Phase Determination

BY J. KARLE

U.S. Naval Research Laboratory, Washington, D.C. 20390, *U.S.A.*

(*Received* 4 *January* 1966)

Methods for combining information from a single isomorphous replacement with phase determining relations for noncentrosymmetric crystals are discussed. The methods mainly concern those circumstances when the related probabilities are large enough to permit the application of phase determining relations in a step-by-step fashion as in the symbolic addition procedure, obtaining new phases, at the start, from relatively few known phases. The case when the isomorphous replacement makes a real contribution to all the structure factors is discussed and appropriate phase determining formulas with their associated probabilities are derived.

For the case when the isomorphous replacement involves a noncentrosymmetric substituent, Coulter has previously suggested a procedure for combining information from the replacement and phase determining formulas for large molecules such as proteins. A procedure is presented here for smaller molecules involving a noncentrosymmetric substituent which is an alternative to that of Coulter. It is discussed in terms of a phase determining procedure which has already been applied to the structure determination of noncentrosymmetric crystals without the aid of an isomorphous replacement.

1. Introduction

The phases of the structure factors can be immediately obtained for centrosymmetric crystals, as is well known, from a single isomorphous substitution. For noncentrosymmetric crystals, at least two different isomorphous substitutions are required for a direct generalization of the method to the evaluation of the phases of the complex structure factors. This is known as the method of multiple isomorphous replacement (Bokhoven, Schoone & Bijvoet, 1951). Coulter (1965) has suggested a procedure for phase determination for noncentrosymmetric crystals which combines the use of a single isomorphous replacement possessing generally placed substitutions with a phase determining formula, the tangent formula (Karle & Hauptman, 1956; Karle, 1964). It is the purpose of this paper to discuss further the combination of a single isomorphous substitution with phase determining formulas, and to describe a procedure for phase determination for noncentrosymmetric crystals which is applicable when the isomorphous substitution has the special property of making only a real contribution to all the structure factors.

2. Analysis

Phase determining relations

We define the quasi-normalized structure factor (Karle & Hauptman, 1959),

$$\mathscr{E}_{\mathbf{k}}=\sigma_2^{-\frac{1}{2}}\sum_{j=1}^{N} Z_j \exp(2\pi i\mathbf{k}\cdot\mathbf{r}_j)\,, \tag{2.1}$$

and the quasi-normalized structure factor for the squared structure

$$\mathscr{E}'_{\mathbf{k}}=\sigma_4^{-\frac{1}{2}}\sum_{j=1}^{N} Z_j^2 \exp(2\pi i\mathbf{k}\cdot\mathbf{r}_j)\,, \tag{2.2}$$

where Z_j is the atomic number of the jth atom having coordinates represented by the vector $\mathbf{r}_j$ in a unit cell containing N atoms, and

$$\sigma_n=\sum_{j=1}^{N} Z_j^n\,. \tag{2.3}$$

Then,

$$\mathscr{E}_{\mathbf{k}}\mathscr{E}_{\mathbf{h}-\mathbf{k}}=\sigma_2^{-1}\sum_{j=1}^{N} Z_j^2 \exp(2\pi i\mathbf{h}\cdot\mathbf{r}_j) \tag{2.4}$$

$$+\sigma_2^{-1}\sum_{\substack{j\neq j'\\1}}^{N}\sum Z_jZ_{j'}\exp[2\pi i\{\mathbf{k}\cdot\mathbf{r}_j+(\mathbf{h}-\mathbf{k})\cdot\mathbf{r}_{j'}\}]$$

and if we average over all values of $\mathbf{k}$, the double sum term becomes zero and we obtain,

$$\langle\mathscr{E}_{\mathbf{k}}\mathscr{E}_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}=\sigma_2^{-1}\sum_{j=1}^{N} Z_j^2\exp(2\pi i\mathbf{h}\cdot\mathbf{r}_j)$$
$$=\sigma_4^{\frac{1}{2}}\sigma_2^{-1}\mathscr{E}'_{\mathbf{h}}\,. \tag{2.5}$$

By means of the joint probability distribution, a formula similar to (2.5) has been found in terms of normalized* structure factors (Karle & Hauptman, 1956),

$$E_{\mathbf{h}}\approx\sigma_2^{3/2}\sigma_3^{-1}\langle E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}\,. \tag{2.6}$$

An advantage of working with (2.6) rather than (2.5) is that probability formulas, expressing the probability

* For the case that none of the indices is zero, the quasi-normalized structure factors, $\mathscr{E}$, are the same as the normalized structure factors, E. When there are systematic absences due to space group extinctions among reflections for which at least one of the indices is zero, the quasi-normalized structure factors have to be rescaled to equal the normalized structure factors. If, for example, half of the reflections are extinctions within a particular set, such as the $h0l$ reflections in space group $P2_1/c$, then the appropriate quasi-normalized structure factors must be divided by $2^{\frac{1}{2}}$, *i.e.* $\mathscr{E}_{h0l}/2^{\frac{1}{2}}=E_{h0l}$. The rule is $|\mathscr{E}|^2(1-q)=|E|^2$, where q is the fraction of reflections in the set which are space group extinctions. We are considering here only primitive unit cells.

of the correctness of the phase indication of one or more of the contributors to the average, are readily defined in terms of the normalized structure factors. It is also seen that all the quantities in (2.6) are associated with the original structure. Actually, the normalized structure factors for the original and squared structures are equal when all the atoms are the same, and the larger ones do not differ significantly even when the atoms present are of quite unequal atomic number.

Equation (2.6) can be rewritten

$$E_{\mathbf{h}}=\alpha_{\mathbf{h}}+i\beta_{\mathbf{h}}\approx\sigma_2^{3/2}\sigma_3^{-1}\langle\alpha_{\mathbf{k}}\alpha_{\mathbf{h}-\mathbf{k}}-\beta_{\mathbf{k}}\beta_{\mathbf{h}-\mathbf{k}}+i(\beta_{\mathbf{k}}\alpha_{\mathbf{h}-\mathbf{k}}+\alpha_{\mathbf{k}}\beta_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}\,, \quad (2.7)$$

where the α's and β's are evidently the real and imaginary parts, respectively, of the normalized structure factors. Noting that an analysis similar to (2.1)–(2.5) for the real parts of the structure factors is equivalent, except for a factor of 2, to the derivation of (2.5) for centrosymmetric crystals, we may conclude from (2.7) that

$$\alpha_{\mathbf{h}}\approx 2\sigma_2^{3/2}\sigma_3^{-1}\langle\alpha_{\mathbf{k}}\alpha_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}=-2\sigma_2^{3/2}\sigma_3^{-1}\langle\beta_{\mathbf{k}}\beta_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}\,. \quad (2.8)$$

Also, since

$$\langle\beta_{\mathbf{k}}\alpha_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}=\langle\alpha_{\mathbf{k}}\beta_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}\,, \quad (2.9)$$

we have for the imaginary part,

$$\beta_{\mathbf{h}}\approx 2\sigma_2^{3/2}\sigma_3^{-1}\langle\beta_{\mathbf{k}}\alpha_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}=2\sigma_2^{3/2}\sigma_3^{-1}\langle\alpha_{\mathbf{k}}\beta_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}\,. \quad (2.10)$$

It is of interest to examine equations (2.8) and (2.10). Equation (2.8) indicates that the sign of the real part, $\alpha_{\mathbf{h}}$, can not only be obtained from an average over the product of real parts, $\alpha_{\mathbf{k}}\alpha_{\mathbf{h}-\mathbf{k}}$, as expected, but also from an average over the product of imaginary parts $\beta_{\mathbf{k}}\beta_{\mathbf{h}-\mathbf{k}}$. It is also seen from (2.10) that the imaginary part, $\beta_{\mathbf{h}}$, is obtained by taking averages over products of real and imaginary parts. A formula similar to equations (2.8) and (2.10) has been used quite generally and with success for sign determination for centrosymmetric crystals. A general discussion of the procedure to follow has been given by Karle & Karle (1966). Therefore, if the magnitudes of the real and imaginary parts of the structure factors were available, the phase problem for noncentrosymmetric crystals would be essentially as simple as that for centrosymmetric ones.

Probabilities

At the beginning of a procedure for phase determination for structures of not too great complexity, single terms are used in (2.8) and (2.10), rather than averages over many contributors. The initial steps would involve the use of the real and imaginary parts of largest magnitude. In order to judge the reliability of each step in the procedure, it is important to have probability formulas expressing the probability that the sign of a real or imaginary part is positive in terms of the contributors to (2.8) and (2.10).

The appropriate probability formulas can be obtained in a fashion similar to that used to derive equation (3.36) in Monograph 3 (Hauptman & Karle, 1953). These formulas are then put into the convenient hyperbolic tangent form of Woolfson (1954). The following definitions of the complex structure factor, $F_{\mathbf{h}}$, and the real and imaginary parts of the normalized structure factor, $E_{\mathbf{h}}$, are employed:

$$F_{\mathbf{h}}=X_{\mathbf{h}}+iY_{\mathbf{h}}=\sum_{j=1}^{N/n} f_{j\mathbf{h}}[\xi(x_j,y_j,z_j,\mathbf{h})+i\eta(x_j,y_j,z_j,\mathbf{h})]\,, \quad (2.11)$$

$$E_{\mathbf{h}}=\frac{F_{\mathbf{h}}n^{\pm}}{(m_2^0+m_0^2)^{\pm}\sigma_2^{\pm}}=\alpha_{\mathbf{h}}+i\beta_{\mathbf{h}}\,, \quad (2.12)$$

so that

$$\alpha_{\mathbf{h}}=\frac{X_{\mathbf{h}}n^{\pm}}{(m_2^0+m_0^2)^{\pm}\sigma_2^{\pm}}\,, \quad (2.13)$$

and

$$\beta_{\mathbf{h}}=\frac{Y_{\mathbf{h}}n^{\pm}}{(m_2^0+m_0^2)^{\pm}\sigma_2^{\pm}}\,, \quad (2.14)$$

where $f_{j\mathbf{h}}$ is the scattering factor of the jth atom whose coordinates are x_j, y_j, z_j, n is the symmetry number of the space group and the m are moments defined in terms of the real and imaginary parts of $F_{\mathbf{h}}$, ξ and η resp., which can be found in *International Tables for X-ray Crystallography* (1952). A general mixed moment is defined by

$$m_{\gamma,\ldots,\gamma_r}^{\delta,\ldots\delta_r}=\int_0^1\int_0^1\int_0^1\xi^{\gamma_1}(x,y,z,\mathbf{h}_1)\cdots\xi^{\gamma_r}(x,y,z,\mathbf{h}_r)\times\eta^{\delta_1}(x,y,z,\mathbf{h}_1)\cdots\eta^{\delta_r}(x,y,z,\mathbf{h}_r)dxdydz\,. \quad (2.15)$$

The probability formulas, obtained in accordance with the definitions (2.11) to (2.15) with the use of Monograph 3, are then put into the hyperbolic tangent form by means of a procedure indicated in a previous review article (Karle, 1964, p. 73). We thus obtain for the probability, $P_+(\alpha_{\mathbf{h}})$, that the sign of $\alpha_{\mathbf{h}}$ is positive,

$$P_+(\alpha_{\mathbf{h}})\approx\tfrac{1}{2}+\tfrac{1}{2}\tanh\left[\frac{n^{\pm}}{m_{200}^{000}}(m_{200}^{000}+m_{000}^{200})^{\pm}\times(m_{020}^{000}+m_{000}^{020})^{\pm}(m_{002}^{000}+m_{000}^{002})^{\pm}\sigma_3\sigma_2^{-3/2}|\alpha_{\mathbf{h}}|\times\left(\frac{m_{111}^{000}}{m_{020}^{000}m_{002}^{000}}\sum_{\mathbf{k}}\alpha_{\mathbf{k}}\alpha_{\mathbf{h}-\mathbf{k}}+\frac{m_{100}^{011}}{m_{000}^{020}m_{000}^{002}}\sum_{\mathbf{k}}\beta_{\mathbf{k}}\beta_{\mathbf{h}-\mathbf{k}}\right)\right], \quad (2.16)$$

and for the probability, $P_+(\beta_{\mathbf{h}})$, that the sign of $\beta_{\mathbf{h}}$ is positive,

$$P_+(\beta_{\mathbf{h}})\approx\tfrac{1}{2}-\tfrac{1}{2}\tanh\left[\frac{n^{\pm}m_{010}^{101}}{m_{000}^{200}m_{020}^{000}m_{000}^{002}}(m_{200}^{000}+m_{000}^{200})^{\pm}\times(m_{020}^{000}+m_{000}^{020})^{\pm}(m_{002}^{000}+m_{000}^{002})^{\pm}\sigma_3\sigma_2^{-3/2}|\beta_{\mathbf{h}}|\sum_{\mathbf{k}}\alpha_{\mathbf{k}}\beta_{\mathbf{h}-\mathbf{k}}\right]. \quad (2.17)$$

In general (2.16) and (2.17) reduce to

$$P_+(\alpha_{\mathbf{h}})\approx\tfrac{1}{2}+\tfrac{1}{2}\tanh 2\sigma_3\sigma_2^{-3/2}|\alpha_{\mathbf{h}}|(\sum_{\mathbf{k}}\alpha_{\mathbf{k}}\alpha_{\mathbf{h}-\mathbf{k}}-\sum_{\mathbf{k}}\beta_{\mathbf{k}}\beta_{\mathbf{h}-\mathbf{k}})\,, \quad (2.18)$$

and

$$P_+(\beta_{\mathbf{h}})\approx\tfrac{1}{2}+\tfrac{1}{2}\tanh 2\sigma_3\sigma_2^{-3/2}|\beta_{\mathbf{h}}|\sum_{\mathbf{k}}\alpha_{\mathbf{k}}\beta_{\mathbf{h}-\mathbf{k}}\,. \quad (2.19)$$

In special cases, calculation of the moments shows that the 2 in the argument of tanh must be replaced. For example, when α_h, α_k and α_{h-k} all correspond to centrosymmetric reflections of the same type, *e.g.* the *h*0*l* reflections in $P2$ or in $P222$, the 2 in the argument of tanh is replaced by 1. However if the α_h, α_k, α_{h-k} correspond to centrosymmetric reflections of the type *h*0*l*, 0*kl*, *hk*0, respectively, in space group $P222$, then the numerical coefficient is again 2. Additional variations in the coefficient occur when the reflections concerned come from subsets which include absences arising from the space group symmetry. In any given case the moments required for (2.16) and (2.17) can be readily computed from (2.15).

3. Real isomorphous replacement

The description of an isomorphous pair in terms of structure factors can be made as follows:

$$F_{R+X}+F_{Y-X}=F_{R+Y}\,, \tag{3.1}$$

where F_{R+X} is the structure factor for the structure consisting of invariant atoms R and replaceable atoms X, F_{R+Y} concerns the invariant atoms R and replaceable atoms Y and F_{Y-X} is the structure factor for the configuration of the difference between atoms Y and atoms X. The relationship of isomorphous replacement to phase determination in noncentrosymmetric crystals was clarified by Bijvoet (1952). We shall consider now the case when $Y-X$ refers to a centrosymmetric configuration, whereas $R+X$ and $R+Y$ are noncentrosymmetric. An appropriate diagram is given in Fig. 1 in which two circles of radius $|F_{R+X}|$ and $|F_{R+Y}|$ are drawn. The vector F_{Y-X} is assumed to point in the positive real direction, giving the two ambiguous closed vector polygons. Information concerning F_{Y-X} is generally obtained from a calculation of the Patterson function, or a modification thereof.

It is apparent that knowledge of the real replacement F_{Y-X} gives both the magnitudes and signs of the real parts of F_{R+X} and F_{R+Y}, and the magnitudes of the imaginary parts. With the assignment of the sign of a particular imaginary part to specify the enantiomorph, it should be readily possible to employ the information available from the isomorphous replacement with the phase determining formulas (2.8) and (2.10) to obtain the signs for the remaining imaginary parts. In this way it is possible to perform a complete phase determination, even when the isomorphous replacement is real, given sufficient experimental data. Sufficiency is determined by the possibility of obtaining adequately high probabilities, as measured by (2.18) and (2.19), to proceed with the determination. The situation here can be expected to be more favorable than in the ordinary application of the symbolic addition procedure for centrosymmetric crystals, since in applying (2.10) the sign of one of the factors in the product, that of α, is initially known. The detailed theoretical and practical aspects of the application of the symbolic addition procedure have recently been described (Karle & Karle, 1966).

4. Complex isomorphous replacement

In the case of a complex isomorphous replacement, equation (3.1) is still appropriate. However, here $Y-X$, as well as $R+X$ and $R+Y$, refers to a noncentrosymmetric configuration. An appropriate diagram is given in Fig. 2 in which two circles of radius $|F_{R+X}|$ and $|F_{R+Y}|$ are drawn. Two ambiguous closed vector polygons are obtained symmetric about the direction of F_{Y-X}, Clearly the problem of determining the phase of F_{R+X} or F_{R+Y} is reduced by the single isomorphous substitution from the consideration of all values between $-\pi$ and $+\pi$ to the choice of either of two fixed alternatives. The choice of one of the two alternatives

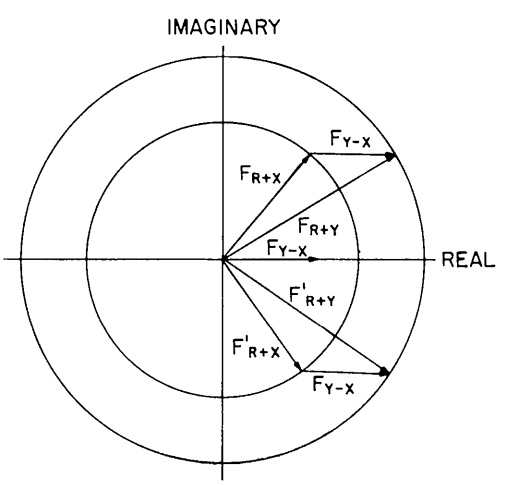

Fig. 1. Construction for isomorphous replacement showing the relationship among the vectors F_{R+X}, F_{R+Y} and F_{Y-X} where the latter is real.

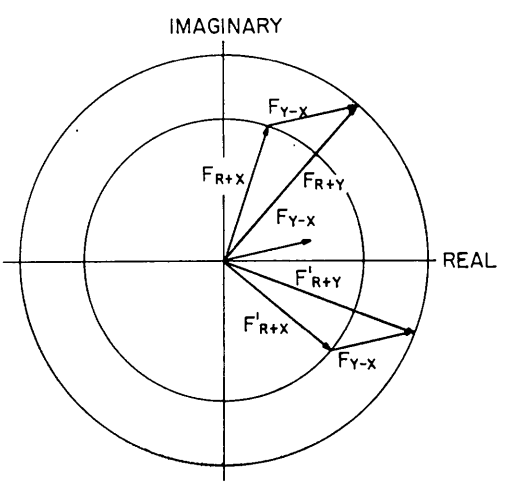

Fig. 2. Construction for isomorphous replacement showing the relationship among the vectors F_{R+X}, F_{R+Y} and F_{Y-X} where the latter is complex.

for an appropriate structure factor can determine, through phase determining formulas, the choice of the proper alternative for the remaining structure factors of interest.

When the probabilities or variances do not permit the application of phase determining relations in a step-by-step fashion by giving definitive choices between the alternative values for the phases, it is possible to proceed according to the suggestion of Coulter (1965). In his procedure a large amount of preliminary phase information can be assembled, since the phases of pure real and pure imaginary structure factors are determined by a single isomorphous substitution, and in many instances the ambiguous alternatives for the complex structure factors are not widely separated, thus permitting the averages of their alternative phase values, given by the phase of the substitution, to be used as a starting point. This preliminary phase information can then be refined and extended by use of formula (4.2) below.

When the probability measures are high enough, the phase determination can proceed in a step-by-step fashion by employing appropriate phase determining formulas for the noncentrosymmetric space groups. An advantage in proceeding in this way is that it would be initially possible to evaluate the phases associated with the largest E-magnitudes, even though the ambiguous values for the associated phases were widely separated. The phase determining formulas are

$$\varphi_{\mathbf{h}} \approx \langle \varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}} \rangle_{\mathbf{k}_r} , \tag{4.1}$$

and

$$\tan \varphi_{\mathbf{h}} \approx \frac{\sum_{\mathbf{k}} |E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}| \sin (\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}})}{\sum_{\mathbf{k}} |E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}| \cos (\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}})} , \tag{4.2}$$

where the $E_{\mathbf{k}}$ represent normalized structure factors whose phases are $\varphi_{\mathbf{k}}$ and the symbol $\mathbf{k}_r$ implies that the average extends only over those vectors, $\mathbf{k}$, associated with large $|E_{\mathbf{k}}|$ values. The application of these formulas in connection with the symbolic addition procedure has been discussed elsewhere (Karle & Karle, 1964, 1966). It is suggested that the symbolic addition procedure be followed in direct combination with the information from the isomorphous substitution. Clearly the addition of this latter information reduces the problem from determining the value of a phase from the continuous range of $-\pi$ to $+\pi$ to that of choosing between only two possible values permitted by the interpretation of the substitution. Measures of the variance (Karle & Karle, 1966, equation (3.33) and Fig. 2) may be employed in order to evaluate the reliability of the contributors in (4.1). In this way a set of phases may be generated in a stepwise fashion.

References

BIJVOET, J. M. (1952). Conference on Computing Methods and the Phase Problem in X-ray Crystal Analysis: Penn. State College. p. 84.

BOKHOVEN, C., SCHOONE, J. C. & BIJVOET, J. M. (1951). *Acta Cryst.* **4**, 275.

COULTER, C. L. (1965). *J. Mol. Biol.* **12**, 292.

HAUPTMAN, H. & KARLE, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal.* A.C.A. Monograph No. 3. Pittsburgh Polycrystal Book Service.

International Tables for X-ray Crystallography (1952). Vol. I. Birmingham: Kynoch Press.

KARLE, I. L. & KARLE, J. (1964). *Acta Cryst.* **17**, 835.

KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635.

KARLE, J. & HAUPTMAN, H. (1959). *Acta Cryst.* **12**, 404.

KARLE, J. (1964). *Advances in Structure Research by Diffraction Methods,* Vol. I. Edited by R. Brill. New York: Interscience Publishers.

KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* To be published.

WOOLFSON, M. M. (1954). *Acta Cryst.* **7**, 61.

Acta Cryst. (1966). **21**, 276

Some Growth Features on (111) Faces of Natural Diamonds

BY FEDERICO BEDARIDA* AND HIROSHI KOMATSU

Physics Department, Royal Holloway College, Englefield Green, Surrey, England

(*Received* 1 *November* 1965)

Some particular features on the (111) faces of a diamond are reported. They are vicinal faces built up by extremely thin growth layers. A growth mechanism is suggested.

Introduction

The crystal studied was picked from among some hundreds of natural diamonds from South Africa. It is a tabular twinned octahedron modified to a rounded hexaoctahedron on the sides, and the crystal is so flattened that only two octahedral faces, *viz.* (111) and ($\bar{1}\bar{1}\bar{1}$), appear as well as curved (*hkl*) type faces.

A small portion of the crystal had been chipped off (Fig. 1). This diamond weighs 0·58 carat, is transparent and as far as can be seen under the microscope has no

* Present address: Istituto di Mineralogia, Università di Genova, Via Leon Battista Alberti 4, Genova, Italy.