

RAMASECHAN, S. & RAMACHANDRAN, G. N. (1953). *Acta Cryst.* **6**, 364.
 STURKEY, L. S. (1948). *Phys. Rev.* **73**, 183.

WEI, P. (1935). *Z. Kristallogr.* **92**, 355.
 ZACHARIASEN, W. H. (1945). *Theory of X-ray Diffraction in Crystals*. New York: Wiley.

Acta Cryst. (1959). **12**, 794

The Treatment of Errors in the Isomorphous Replacement Method

BY D. M. BLOW AND F. H. C. CRICK

Medical Research Council Unit for Molecular Biology, Cavendish Laboratory, Cambridge, England

(Received 28 July 1958 and in revised form 5 November 1958)

This treatment is intended for complex structures where conventional refinement is impossible. A method is described for assessing the errors which arise in applying the isomorphous replacement method. Both errors due to non-isomorphism and observational errors are considered. Probability functions are derived which give, in the centrosymmetric case, the probability of a correct sign determination, and in the non-centrosymmetric case the relative probabilities of different phases. These probabilities may be used to calculate a 'best' Fourier, in which the errors in electron density are minimized, and also to estimate the r.m.s. error in this 'best' Fourier.

There are two steps in the application of the isomorphous replacement method. The first is the determination of the position of the outstanding features (usually a small number of heavy atoms) which differentiate a pair of isomorphous structures. The contribution f_c of this part of the structure to the structure factors may then be calculated. The second step is the use of these calculated contributions to determine the phases of the reflexions. This is done by comparing them with the observed intensity differences. In this way the structure may be determined.

This paper is concerned with the second step. There will be many reflexions for which f_c is very small. The determination of phase will be correspondingly poor. How should these reflexions be treated? In non-centrosymmetric structures, as is well known, unambiguous phase determinations are possible only if at least three isomorphous compounds are available. How should the results from the two pairs be combined? With simple structures where atoms are resolved, a trial structure can be obtained, and refinement made, for instance by the least-squares method. With a large protein, there is no immediate prospect of resolving the individual atoms and therefore no way of refinement from a trial structure. The accuracy of the final Fourier is dependent on the best choice of weights and phases during the second step of the calculation. There will be cases of intermediate complexity where the right trial structure will be found only if the second step is done accurately enough.

We will describe a method for treating this question as rigorously as possible. A structural study where the method has been put to practical use has been published elsewhere (Blow, 1958).

Estimation of error

The errors with which we are concerned are those which arise in the use of the isomorphous replacement method. It will be assumed that the 'true' structure would be the Fourier transform of accurately observed structure factors, given the proper phases. Errors which arise due to series termination and extinction are not considered.

Let F , F_H be the structure factors of two isomorphous compounds, the latter containing additional heavy atoms. We will define

$$f_H = F_H - F. \quad (1)$$

The basis of the isomorphous replacement method is to calculate an approximation to f_H , which we will call f_c , usually by assuming the differences are entirely due to heavy atoms whose coordinates have been determined. (1) then gives information about the phases.

Centrosymmetric case

If the structures are centrosymmetric, then F , F_H , f_H , f_c are all real. Either

$$|f_H| = |F_H - F| \quad (2a)$$

or

$$|f_H| = |F_H + F| \quad (2b)$$

the latter case arising only when the signs of F_H and F are different. If we exclude all cases where $F_H + F < f_c(000)$, the maximum possible value of f_c , we can approach certainty that (2a) applies. In these cases, a direct assessment of error may be made.

We will use the symbols F_H and F to represent the observed amplitudes of \mathbf{F}_H and \mathbf{F} , and define $\Delta F_{\text{obs.}} = F_H - F$. If there were no error, whenever (2a) applies $|\Delta F_{\text{obs.}}|$ and $|f_c|$ would always be equal. The quantity $E = |\Delta F_{\text{obs.}}| - |f_c|$ is a measure of the total error.

The distribution of $|E|$ has been examined for the $h0l$'s of horse haemoglobin (Cullis, Dintzis & Perutz, 1957). The distribution closely follows the Gaussian distribution of random errors; the mean value usually varies with $\sin \theta$ (Figs. 1 and 2).

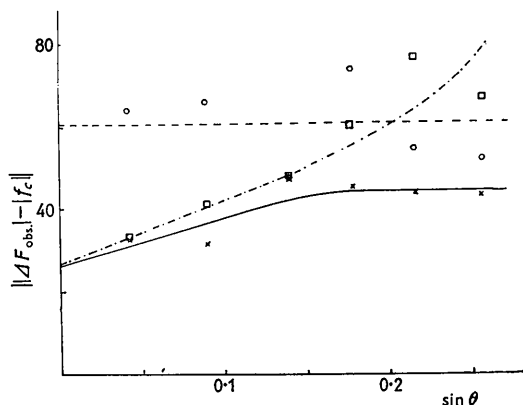


Fig. 1. The mean value of $|\Delta F_{\text{obs.}}| - |f_c|$ plotted as a function of $\sin \theta$ for various isomorphous derivatives of horse haemoglobin.

×	×		×		
×			×		mercury acetate:
haemoglobin compound					
○	○		○		dimercy acetic acid:
○			○		haemoglobin compound
□	□		□		p-chloro-mercuribenzoate:
□			□		haemoglobin compound

(Taken from the data of Cullis, Dintzis & Perutz (1957). We are grateful to Miss A. F. Cullis for preparing this figure.)

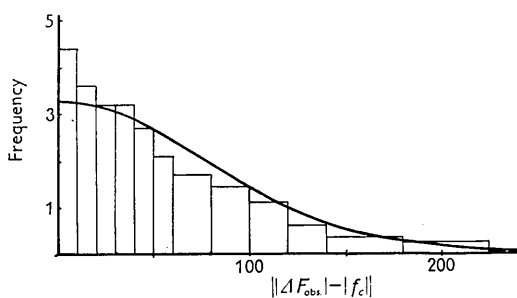


Fig. 2. Histogram showing the frequency distribution of $|\Delta F_{\text{obs.}}| - |f_c|$ for dimercy acetic acid haemoglobin. The full line is a Gaussian curve with the same mean square value. (Taken from the data of Cullis, Dintzis & Perutz (1957).)

These errors can be thought of as arising from two distinct causes. They arise partly from the difference ϵ between f_H and the estimate f_c . This is partly due to errors in placing and weighting the heavy atoms, wrongly estimated temperature factors and so on,

but also to the differences which arise from incomplete isomorphism, in some cases caused by slight shifts and rotations of the molecules (Crick & Magdoff, 1956), and perhaps by the introduction of lighter atoms into the structure whose parameters have not been determined.

A separate source of error is experimental inaccuracy in the determination of the structure amplitudes, F , F_H . We will call

$$\delta = |\Delta F_{\text{obs.}}| - |\mathbf{F}_H| - |\mathbf{F}|. \quad (3)$$

Its r.m.s. value $\langle \delta \rangle$ is easily checked by comparing values of F derived from different crystals of the same compound. (This method of estimation neglects some systematic errors, such as that due to absorption. But since isomorphous replacement depends on the comparison of intensities from similar crystals, these errors are unimportant in calculating the differences.)

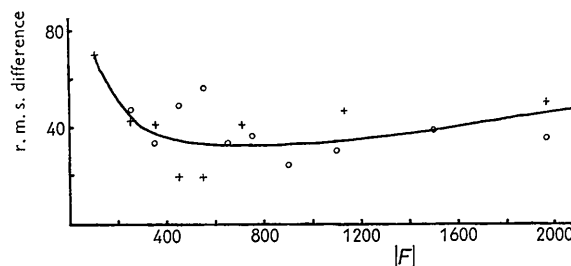


Fig. 3. R.m.s. difference between independent observations of F from two similar haemoglobin crystals. Each symbol represents the average of about 10 reflexions of similar F . The results come from two different experiments. The data were taken from precession photographs and the intensities measured photometrically (Blow, 1958).

The result of such a check with the $0kl$'s of horse haemoglobin is shown in Fig. 3. $\langle \delta \rangle$ was found to be a function of F . Since

$$\langle E \rangle^2 = \langle \delta \rangle^2 + \langle \epsilon \rangle^2 \quad (4)$$

the estimates of $\langle \delta \rangle$ and $\langle E \rangle$ permit $\langle \epsilon \rangle$ to be estimated.

Non-centrosymmetric case

In the non-centrosymmetric case \mathbf{F} , \mathbf{F}_H , f_H have arbitrary phase and there is no direct means of comparing observed and calculated differences, though $\langle \delta \rangle$ can be estimated in the same way as before. A reasonable estimate of $\langle \epsilon \rangle$ will have to be made. Any non-centrosymmetric space groups whose symmetry includes dyads or screw dyads possess centrosymmetric zones, and in these cases it is reasonable to use this zone to estimate $\langle \epsilon \rangle$ for the whole structure. f_H need no longer be real, and the estimate of it may be thought of as a circularly symmetrical 'cloud' of uncertainty centred on f_c . We shall now consider how the breadth of this cloud may be estimated.

Consider the unit cell split up into two similar units related by a dyad or a screw dyad. The heavy atom contributions of these two units are f_1 and f_2 . When projected down the dyad, the structure appears

centrosymmetric; thus in the zone corresponding to this projection the relation

$$\mathbf{f}_2 = \mathbf{f}_1^*$$

will hold. Suppose an error ε_1 has been made in estimating \mathbf{f}_1 . Then in this zone (see Fig. 4)

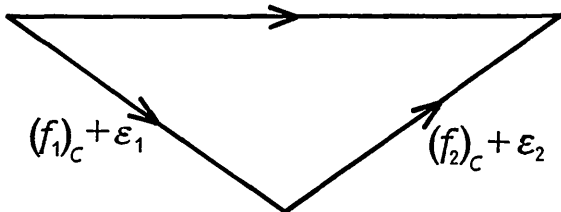


Fig. 4.

$$\mathbf{f}_H = (\mathbf{f}_1)_c + \varepsilon_1 + (\mathbf{f}_2)_c + \varepsilon_2 = \mathbf{f}_c + \varepsilon_1 + \varepsilon_1^* \quad (5)$$

The error in estimating \mathbf{f}_H is ε where

$$|\varepsilon| = |\varepsilon_1 + \varepsilon_1^*| = 2|\varepsilon_1| \cos \varphi' \quad (6a)$$

and φ' is the phase of \mathbf{f}_1 . We may assume φ' is random, so that

$$\langle \varepsilon \rangle^2 = 4\overline{\varepsilon_1^2} \overline{\cos^2 \varphi'} = 2\overline{\varepsilon_1^2}.$$

The r.m.s. value $\langle \varepsilon \rangle$ in this zone may be assessed in just the same way as for a centrosymmetric structure.

To apply this estimate to reflexions not in the centrosymmetric zone we have to assume that $|\varepsilon_1|$ and $|\varepsilon_2|$ have the same distribution for these other reflexions as they do for the centrosymmetric reflexions. The first equality in equation (5) applies, but ε_2 can no longer be replaced by ε_1^* . (6) becomes

$$|\varepsilon| = |\varepsilon_1 + \varepsilon_2| = (\varepsilon_1^2 + \varepsilon_2^2 - 2|\varepsilon_1||\varepsilon_2| \cos \varphi'')^{\frac{1}{2}} \quad (6b)$$

where φ'' is the angle between ε_1 and ε_2 and is assumed random, so we find $\langle \varepsilon \rangle^2 = 2\overline{\varepsilon_1^2}$ as before. The circularly symmetrical cloud of uncertainty around \mathbf{f}_c in the non-centrosymmetric case has the same breadth as the linear probability function representing the estimate of \mathbf{f}_H in the centro-symmetric case, namely $\langle \varepsilon \rangle$. In systems with only one axis of symmetry there are uncertainties about the heavy atom co-ordinates parallel to this axis, which do not affect the centrosymmetric projection (Harker, 1956; Perutz, 1956). In these cases a further allowance needs to be made for the corresponding uncertainty in the phase of \mathbf{f}_c .

'The best Fourier'

Let us suppose we have some information about the value of a structure factor $\mathbf{F}(hkl)$. To include the non-centrosymmetric case, we will put

$$\mathbf{F}(hkl) = A(hkl) + iB(hkl) = F e^{i\varphi} \quad (8)$$

though in the centrosymmetric case it will be known that B is always zero. We can represent the informa-

tion we have about \mathbf{F} on an Argand diagram with A and B as coordinates. This can be done by plotting the probability that \mathbf{F} lies at any point on the diagram.

If, for instance, the amplitude $|\mathbf{F}|$ had been measured with perfect accuracy but there was no information about the phase, the diagram would in the non-centrosymmetric case consist of a circle radius $|\mathbf{F}|$, centred on the origin. In the centrosymmetric case it would consist of two equal points of probability density, $\mathbf{F} = +|\mathbf{F}|$ and $\mathbf{F} = -|\mathbf{F}|$. In practice, neither amplitude nor phase is known with perfect accuracy, and the probability will be a smoothly varying function, representing all available information about $\mathbf{F}(hkl)$. We will call this function $P_{hkl}(A, B)$, and it will be normalised so that

$$\iint_{-\infty}^{+\infty} P_{hkl}(A, B) dA dB = 1. \quad (9)$$

In order to portray this information in structural terms, it is necessary to make a Fourier transformation. The most general way to do this would be to work out the transforms for all possible values of A and B for each hkl . The probability of a transform being the correct one would be the product

$$\prod_{hkl} P_{hkl}(A, B). \quad (10)$$

In this way one can imagine obtaining a continuum of structures, each with an assigned probability. The structures whose probability fell below a certain value might be rejected, and the remainder examined to see what characteristics they had in common.

Such an idealised procedure could not be undertaken in practice. However, these considerations enable us to define two transforms which can be calculated more easily.

The more obvious is the 'most probable Fourier'. This may be calculated by choosing the value of $\mathbf{F}(hkl)$ for each reflexion corresponding to the highest value of $P_{hkl}(A, B)$. This corresponds to using (1) and neglecting all errors. This Fourier will clearly have the maximum value of $\prod_{hkl} P_{hkl}(A, B)$ and is the most likely to be correct.

If the form of the probability functions $P_{hkl}(A, B)$ were usually unimodal, with a single peak on the diagram, the most probable Fourier would not be a bad function to use, though it might tend to give too much weight to uncertain phases. Unfortunately, it turns out that in the isomorphous replacement method there is a strong tendency for $P_{hkl}(A, B)$ to be bimodal. When two peaks have nearly equal weight, there is a strong chance of making a large error, if the most probable value of \mathbf{F} is used. A compromise is clearly needed. It will now be shown that the centroid of the distribution provides just the required compromise.

For the time being we shall assume that only one reflexion $\mathbf{F}(hkl)$, and its conjugate $\mathbf{F}(\bar{h}\bar{k}\bar{l})$, are unknown; the others are known with perfect accuracy.

(The treatment can be extended directly to the case where all reflexions are unknown.) Suppose that we decide to use some value $F_o(hkl)$ for this reflexion, and that the true value is $F_T(hkl)$. In calculating a Fourier synthesis, we introduce an error

$$\Delta \varrho_{hkl} = \frac{1}{V} [(\mathbf{F}_o - \mathbf{F}_T) \exp \{-2\pi i(hx + ky + lz)\} + (\mathbf{F}_o^* - \mathbf{F}_T^*) \exp \{2\pi i(hx + ky + lz)\}]. \quad (11)$$

The mean square value of $\Delta \varrho$ over the whole unit cell is

$$\langle \Delta \varrho \rangle_{hkl}^2 = \frac{2}{V^2} (\mathbf{F}_o - \mathbf{F}_T)^2. \quad (12)$$

In practice, our knowledge of \mathbf{F}_T can only be expressed in the form of a probability distribution, and we must take a weighted mean over all possible values:

$$\langle \Delta \varrho_{hkl} \rangle^2 = \frac{2}{V^2} \int_{-\infty}^{\infty} P_{hkl}(A, B) [\mathbf{F}_o - (A + iB)]^2 dA dB, \quad (13)$$

which we may rewrite

$$\langle \Delta \varrho \rangle_{hkl}^2 = 2r^2/V^2, \quad \text{thus defining } r.$$

r is the r.m.s. length of the vector between \mathbf{F}_o and the probability function we use to describe the observations of \mathbf{F} . It is our estimate of the error introduced in deciding to use \mathbf{F}_o . The integral is exactly analogous to the expression for the radius of gyration r of a plane lamina with density $P_{hkl}(A, B)$ about the point \mathbf{F}_o . As in the parallel-axes theorem of mechanics, it is easy to show that this integral has its minimum value when

$$\mathbf{F}_o = \iint_{-\infty}^{\infty} P_{hkl}(A, B)(A + iB) dA dB; \quad (14)$$

that is, when \mathbf{F}_o is the centroid of the distribution.

We define the 'best Fourier' as *that Fourier transform which is expected to have the minimum mean square difference from the Fourier transform of the true \mathbf{F} 's when averaged over the whole unit cell. The best Fourier is the Fourier transform obtained by using the centroid of the probability distribution for \mathbf{F} .*

Application to centrosymmetric case

The probability distribution in the centrosymmetric case is a function of A alone, since B must be zero. In general, it will have two parts, centred on $+F$ and $-F$, which will be represented by Gaussian curves of breadth $\langle \delta' \rangle$ (Fig. 5). (It should be noted that $\langle \delta' \rangle$ is not quite the same as $\langle \delta \rangle$: for errors due to absorption etc., which are eliminated in comparing two similar crystals are important in estimating the true magnitude of $\langle \delta' \rangle$.)

The relative weight of these two parts of the distribution is our estimate of the probability of a correct sign determination. This estimate is made by compar-

ing the discrepancies between theory and experiment with the estimate of error $\langle E \rangle$, assuming

- (a) \mathbf{F} is positive;
- (b) \mathbf{F} is negative.

When $|\mathbf{F}|$ is small it will be important to remember that \mathbf{F}_H need not have the same sign as \mathbf{F} (Table 1).

Table 1. *Discrepancy between theory and experiment in various cases*

Sign of \mathbf{F}	Sign of \mathbf{F}_H	Discrepancy
+	+	$\mathbf{f}_c - (\mathbf{F}_H - \mathbf{F})$
+	-	$\mathbf{f}_c + (\mathbf{F}_H + \mathbf{F})$
-	+	$\mathbf{f}_c - (\mathbf{F}_H + \mathbf{F})$
-	-	$\mathbf{f}_c + (\mathbf{F}_H - \mathbf{F})$

The probability of the two cases are then as follows:

Probability \mathbf{F} is positive

$$P_+ = N \left[\exp \left\{ -\frac{(\mathbf{f}_c - [\mathbf{F}_H - \mathbf{F}])^2}{2\langle E \rangle^2} \right\} + \exp \left\{ -\frac{(\mathbf{f}_c - [\mathbf{F}_H + \mathbf{F}])^2}{2\langle E \rangle^2} \right\} \right] \quad (15a)$$

Probability \mathbf{F} is negative

$$P_- = N \left[\exp \left\{ -\frac{(\mathbf{f}_c + [\mathbf{F}_H - \mathbf{F}])^2}{2\langle E \rangle^2} \right\} + \exp \left\{ -\frac{(\mathbf{f}_c + [\mathbf{F}_H + \mathbf{F}])^2}{2\langle E \rangle^2} \right\} \right] \quad (15b)$$

where N is a normalising factor such that $P_+ + P_- = 1$. Rearranging these gives

$$P_+ = 2N \exp \left\{ -\frac{(F + \mathbf{f}_c)^2 + F_H^2}{2\langle E \rangle^2} \right\} \sinh \left\{ \frac{F_H(F + \mathbf{f}_c)}{\langle E \rangle^2} \right\},$$

$$P_- = 2N \exp \left\{ -\frac{(F - \mathbf{f}_c)^2 + F_H^2}{2\langle E \rangle^2} \right\} \sinh \left\{ \frac{F_H(F - \mathbf{f}_c)}{\langle E \rangle^2} \right\},$$

$$\frac{P_+}{P_-} = \exp \left\{ -\frac{2F\mathbf{f}_c}{\langle E \rangle^2} \right\} \frac{\sinh \{F_H(F + \mathbf{f}_c)/\langle E \rangle^2\}}{\sinh \{F_H(F - \mathbf{f}_c)/\langle E \rangle^2\}}. \quad (16)$$

The expressions may be simplified if \mathbf{f}_c and E are small compared to $(F + F_H)$. In this case the possibility of \mathbf{F} and \mathbf{F}_H having different sign may be neglected and the second terms in (15) disappear. Under these conditions

$$N^{-1} = 2 \exp \left\{ -\frac{(F - F_H)^2 + \mathbf{f}_c^2}{2\langle E \rangle^2} \right\} \cosh \left\{ \frac{\mathbf{f}_c(F - F_H)}{\langle E \rangle^2} \right\}.$$

If we write

$$t = \mathbf{f}_c(F_H - F)/\langle E \rangle^2, \quad (17)$$

P_+ , P_- may be written

$$\frac{P_+}{P_-} = \frac{e^t/2 \cosh t}{e^{-t}/2 \cosh t}. \quad (18)$$

Thus when $(F + F_H)$ is large compared to $\langle E \rangle$, $|\mathbf{f}_c|$,

the value of F which must be used to give the best Fourier is

$$F_0 = (P_+ - P_-)F = F \tanh t \quad (19)$$

and the mean square value of $\langle F_0 - F_T \rangle^2$, taken over the whole probability distribution is

$$\begin{aligned} r^2 &= (F_0 - F)^2 P_+ + (F_0 + F)^2 P_- + \langle \delta' \rangle^2 \\ &= F^2 \operatorname{sech}^2 t + \langle \delta' \rangle^2. \end{aligned} \quad (20)$$

It will help to discuss these results in terms of a concrete example. The data given below are typical of a weak reflexion in haemoglobin where the experimenter would certainly doubt the significance of his sign determination:

$$F = 200, F_H = 225, f_c = +75 \quad (\text{electrons/unit cell}).$$

Typical estimates of $\langle E \rangle$ and $\langle \delta' \rangle$ would be

$$\langle E \rangle = 50, \langle \delta' \rangle = 35 \quad (\text{electrons/unit cell}).$$

The discrepancies in the results according to the various possible sign combinations may be found from Table 1, and are set out in Table 2. Using equations

Table 2. Probabilities of various sign combinations in an actual case:

$$F = 200, F_H = 225, f_c = +75, \langle E \rangle = 50$$

Sign of F	Sign of F_H	Discrepancy (from Table 1)	Probability, assuming Gaussian error distribution
+	+	$50 = \langle E \rangle$	$Ne^{-\frac{1}{2}} = 0.606N$
+	-	$500 = 10\langle E \rangle$	$Ne^{-50} = 10^{-21}N$
-	+	$350 = 7\langle E \rangle$	$Ne^{-24.5} = 10^{-10}N$
-	-	$100 = 2\langle E \rangle$	$Ne^{-2} = 0.135N$

(15), after finding the normalising factor N , we obtain

$$P_+ = 0.818, P_- = 0.182.$$

Alternatively, using the approximate treatment we find $t = 0.750$, giving exactly the above result. Even when F is only three times $|f_c|$, the inaccuracy is quite negligible.

Equations (19) and (20) give

$$F_0 = +127, r^2 = 23,800 + 1,200 = 158^2.$$

The treatment thus shows that this reflexion should be put into the Fourier synthesis with a weight of about 0.6, and even then the radius of gyration is surprisingly large. It is instructive to compare it with the values corresponding to

- (a) giving the reflexion full weight with positive sign;
- (b) omitting the reflexion from the synthesis.

This corresponds to finding the radius of gyration of the diagram (Fig. 5) about the point +200 and about the origin, respectively. The results are

$$(a) (400^2 \times 0.182) + 35^2 = 174^2$$

$$(b) 200^2 + 35^2 = 203^2.$$

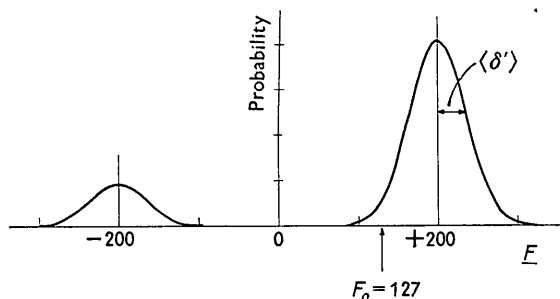


Fig. 5. Probability distribution for F in a centrosymmetric case. Calculated for

$$F = 200, F_H = 225, f_c = +75, \langle \delta' \rangle = 35, \langle E \rangle = 50.$$

We may take estimate (b), the mean square error due to omitting the reflexion, as a standard. In this case, when we are only about 80% sure of a correct sign determination, the mean square error due to giving the reflexion full weight is about 0.73 of the standard. By using the correct weighting function, it may be further reduced to 0.60 of the standard.

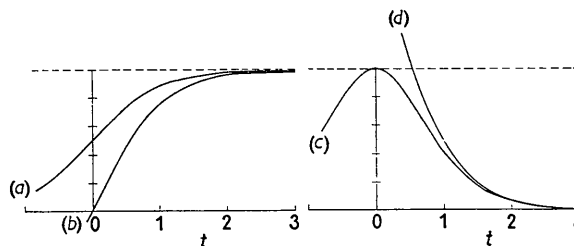


Fig. 6. Curves showing (a) the probability F is positive, $P_+ = \exp(t)/2 \cosh t$; (b) the best weight to apply to F , $F_0/F = \tanh t$; (c) the mean square error when this weight is used, $r^2/F^2 = \operatorname{sech}^2 t$; (d) the mean square error without weighting, $4P_-$.

Fig. 6 shows P_+ , the best weight F_0/F and the r.m.s. error r/F as a function of t . The curves have been calculated from equations (18)–(20), assuming that $\langle \delta' \rangle$ makes a negligible contribution to r . For comparison, the r.m.s. error due to including the term with full positive weight is also plotted (this is found to be $4P_-$). If an experimenter does not wish to use weighting factors, he should omit from his Fourier synthesis all terms where he is less than 75% sure of the signs, (i.e. $|P_+ - \frac{1}{2}| < \frac{1}{4}$) since to include them would make his synthesis worse.

Non-centrosymmetric case

The probability distribution in the non-centrosymmetric case has a much more complicated form and we cannot attempt to treat it so rigorously. Initially we shall assume that $|F|$ is known accurately, and that the observational error δ lies entirely in the observation of $|F_H|$. The extent to which this assumption is justified will be considered later.

Neglecting error, phases may be determined by forming a closed triangle from the known vector f_c and the known amplitudes F , F_H (Fig. 7(a)). (Since

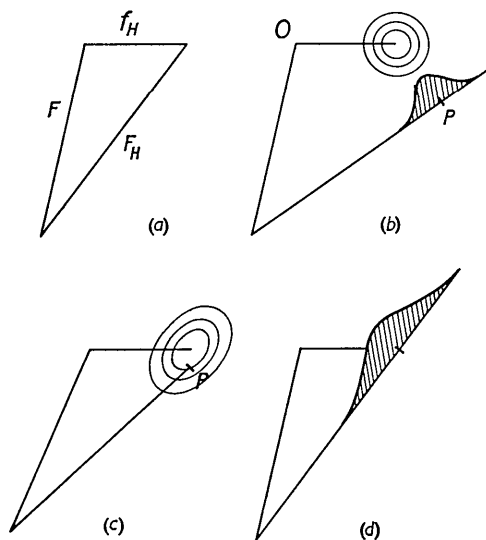


Fig. 7. Effect of errors on phase determination: non-centrosymmetric case. (See text).

this triangle may be drawn in two ways, there is still an ambiguity in the phase which can only be settled by using another isomorphous replacement in which the vector f_c has a different argument.) Then

$$F_H^2 = F^2 + f_c^2 + 2F|f_c| \cos \varphi. \quad (21)$$

To deal with errors we have to consider F_H , f_H as represented by probability functions. In drawing these on the diagram we shall consider the vertex O of the triangle as a fixed point (Fig. 7(b)). The two probability functions may now be combined into one, showing the probability that the point marked P falls at that point in the diagram. This distribution is the convolution of the two previous ones; if $|f_c|$ is much smaller than F and F_H , the result is close to a Gaussian of elliptical contour with major axis $\sqrt{\langle \varepsilon \rangle^2 + \langle \delta \rangle^2}$ and minor axis $\langle \varepsilon \rangle$ (Fig. 7(c)). Moving P along the minor axis of the ellipse has little effect on the phase of F relative to that of f_c , which is what we are trying to determine. So far as this phase is concerned, the probability function may be represented by its projection on its major axis, which is very nearly a Gaussian of breadth $\sqrt{\langle \varepsilon \rangle^2 + \langle \delta \rangle^2} = \langle E \rangle$ (Fig. 7(d)). This is clearly equivalent to regarding the whole error as residing in the determination of F_H .

In practice, errors will occur in the determination of both F and F_H . To show that an error in F causes a phase error of the same magnitude as an error in F_H , (21) may be differentiated with respect to these two quantities. So long as $|f_c|$ is small compared to F , F_H we find

$$\frac{d\varphi}{dF} \approx -\frac{d\varphi}{dF_H}.$$

Thus the assumption that all the error lies in the determination of F_H does not affect the magnitude of the phase error. All these approximations are most accurate when F and F_H are large, which is when estimation of error is most important. We can be satisfied with a much cruder estimate for the weaker terms.

Using the approximate result given above, we can now estimate the probability, from the observations, of a particular phase, φ , being the true phase. This is done by finding the error $x(\varphi)$ needed to be added to the observed F_H to close the triangle when F is given some phase φ . Using the cosine law (21) we have

$$F^2 + f_c^2 + 2|f_c|F \cos \varphi = F_H^2 + 2F_H x + x^2. \quad (22)$$

The probability $P(\varphi)$ of a given phase is

$$N \exp \{-x^2/2E^2\}, \quad (23)$$

where N is a normalising factor such that

$$\int_0^{2\pi} P(\varphi) d\varphi = 1.$$

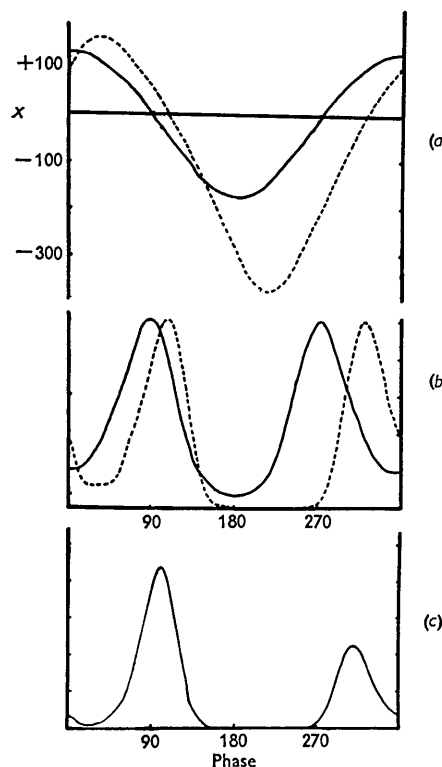


Fig. 8. Derivation of a phase probability curve in a typical case. (a) Error x involved in closing the phase triangle, as a function of phase. The full line and the dotted line represent results from different isomorphous replacements. (b) Relative probability of the phases, assuming Gaussian error distribution. (c) Product of the two curves shown in (b), showing the joint phase probability curve, resulting from taking all the information together. The ambiguity is partly resolved.

This is inconvenient to handle analytically, and a graphical method has been developed which can be used very rapidly. This consists of two rulers, hinged together, each of which carries a cursor. One cursor is pivoted at the centre of the coordinate system; the length of the attached ruler represents F . The length of the other represents F_H , and its cursor carries a scale graduated directly in units of $x^2/\langle E \rangle^2$. (A variety of scales is made for different values of $\langle E \rangle$). The vector f_c is represented by laying a marker on the baseboard. The probability of any phase φ may be estimated by setting F to the appropriate phase and reading off the value of $x^2/\langle E \rangle^2$ from the cursor when the triangle is closed. This is proportional to $\log(P(\varphi)/N)$.

Fig. 8 illustrates the way these values of $x^2/\langle E \rangle^2$ may be used to calculate the relative probabilities of various phases. As already mentioned, a single isomorphous replacement gives an ambiguous result for the phase. This ambiguity may be resolved by the use of a further member of the isomorphous series for which the vector f_c has a different phase. The probabilities may be multiplied together to give a joint probability curve. When this product has been taken over all pairs of compounds in the isomorphous series, the curve gives all available information about the phase of the reflexion.

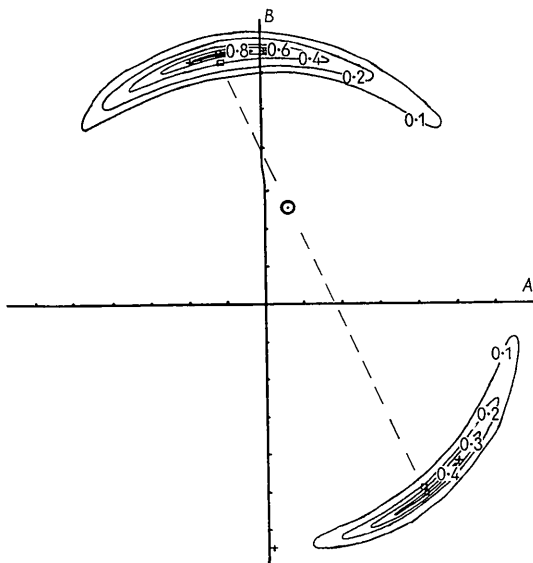


Fig. 9. The phase probability curve of Fig. 8(c), replotted on an Argand diagram. The centroid of the whole distribution, and of each of its parts, is indicated.

This phase probability curve may be used to plot a probability map on an Argand diagram as previously described (Fig. 9). (The breadth of the annular region is again $\langle \delta' \rangle$, as in the centrosymmetric case). In order to obtain the 'best Fourier', we need a way of finding the centroid of such a probability map. We also need a method of estimating the radius of gyra-

tion, r , since this will be used to give a measure of the accuracy of the 'best Fourier'.

In practice, these phase probability curves turn out to be of two types. One of these, the unimodal type, may be closely represented by a Gaussian probability distribution of phase; the other, the bimodal type, is close to the sum of two Gaussians which have the same breadth, given appropriate weights. It is shown in the appendix that the centroid of such a unimodal Gaussian distribution lies at a radius $F \exp(-\frac{1}{2}\sigma^2)$ from the origin, where σ radians is the breadth of the Gaussian. In the case where the probability curve is unimodal, this is the weight which must be given to the reflexion in the 'best Fourier'. In cases like that illustrated in Figs 7 and 8, where the curve is bimodal, the weighted mean of two such centroids must be used.

In the appendix the radius of gyration, r , has also been calculated. Allowing for the uncertainty $\langle \delta' \rangle$ in the magnitude of $|F|$, the result is

$$r^2 = F^2(1 - \exp(-\sigma^2)) + \langle \delta' \rangle^2. \quad (24)$$

The accuracy of the 'best Fourier'

Previous treatments of errors in crystallographic results have usually been concerned with the accuracy of atomic coordinates. Here we are chiefly interested in cases with no prospect of resolving the atomic positions, and since no refinement can be carried out the errors may be much larger than in the conventional treatments. However, we have found that an approach similar to that of Cruickshank (1949) may be used.

Using the weighting systems given above a 'best Fourier' may be calculated. The estimates of r^2 may now be used to form an idea of how accurate this Fourier is. This means, in principle, considering the probability of other Fouriers, where terms are given different values, and seeing to what extent they agree with the 'best Fourier'.

The r.m.s. error in electron density due to a single term in the synthesis is given by (12). r^2 is the estimate of the mean value of $(F_0 - F_T)^2$ for a particular term. The central limit theorem states, under very general conditions about the distribution of the $\Delta \rho$'s, that if there is a very large number of them of a similar order of magnitude, the total error

$$\Delta \rho_{\text{total}} = \sum_{hkl} \Delta \rho_{hkl} \quad (25)$$

will have a normal probability distribution (Cramér, 1937). The r.m.s. value

$$\langle \Delta \rho_{\text{total}} \rangle^2 = \frac{2}{V^2} \sum_{h=0}^{\infty} \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} r^2 \quad (26)$$

is the standard error of electron density.

In this way it is possible to allocate a level of significance to all the features in the Fourier. The accuracy of the 'best Fourier' will usually be the same

throughout the unit cell; except that in some space groups there will be special positions where errors are larger because $\exp \{2\pi i(hx + ky + lz)\}$ is unity for all permitted terms. Occasionally there may be special circumstances which introduce a small number of overpoweringly large error terms—in this case the above conditions are not satisfied, and a normal distribution of this kind cannot be assumed. A case of this sort is discussed in detail by Blow (1958). Another approach would be to synthesize the complete function $\langle \Delta \rho_{\text{total}} \rangle^2(x, y, z)$ from the values of $r^2(hkl)$.

Conclusion

We have described a method for minimising the effect of experimental errors on the results obtained by the isomorphous replacement method. A method for estimating the accuracy of the results has also been given. It may be asked whether it is worth while carrying out detailed calculations of this kind in practical cases.

Methods of the kind we have described here are laborious, and our experience with them is limited to two of the projections of haemoglobin—one centrosymmetric (Bragg & Perutz, 1954), and one non-centrosymmetric (Blow, 1958). The use of complicated weighting functions makes only small changes in the appearance of the final Fourier. More important effects arise because the cautious crystallographer prefers to omit from the Fourier synthesis terms whose phase is uncertain. Often a Fourier can be made worse by this omission, despite the uncertainty. It is in these cases that a reasonable estimate of error and a simple weighting function can be most useful.

What is more important in the case of proteins is to have a quantitative assessment of the accuracy of the final Fourier, at least in a few sample cases. The errors may appear so large that the validity of the results may be doubted; however, it turns out that with the large number of terms involved in a three-dimensional analysis, highly significant results may be obtained. (Blow, 1958; Kendrew *et al.*, 1958). In this way the results of X-ray studies of proteins may be interpreted on an objective basis.

APPENDIX

Calculation of weight when phase probability curve is a Gaussian

Fixed phase error

First consider the case where there is a known phase error, α , in the phase of a reflection. Its amplitude, F , is assumed to be known accurately. If we include the term in a Fourier synthesis with this wrong phase, what weight should it be given to minimise the errors? Let the chosen weight be w . Let the true phase be φ . The transform of this one reflexion (hkl) and its conjugate ($\bar{h}\bar{k}\bar{l}$) has the form

$$\begin{aligned} \rho_{hkl}(x, y, z) &= \frac{2}{V} F \cos(\varphi + 2\pi(hx + ky + lz)) \\ &= 2F \cos \psi / V, \end{aligned}$$

say, at a specified point in the transform.

Then the error introduced into the Fourier

$$\begin{aligned} \Delta \rho_{hkl} &= 2F [\cos \psi - w \cos(\psi - \alpha)] / V \\ &= 2F [\cos \psi (1 - w \cos \alpha) - \sin \psi (w \sin \alpha)] / V. \end{aligned} \quad (A1)$$

Let r^2 be the mean value of $2V^2 \Delta \rho_{hkl}^2$, taken over all values of ψ .

$$\begin{aligned} r^2 &= F^2 [(1 - w \cos \alpha)^2 + w^2 \sin^2 \alpha] \\ &= F^2 (1 - 2w \cos \alpha + w^2). \end{aligned} \quad (A2)$$

r^2 is a minimum when

$$\frac{dr^2}{dw} = 2F^2(w - \cos \alpha) = 0. \quad (A3)$$

This gives, for the best weighting factor,

$$w_0 = \cos \alpha. \quad (A4)$$

Gaussian distribution of error

We now consider the case where the phase is uncertain and there is a Gaussian probability distribution of phase errors

$$P(\alpha) = \frac{1}{(2\pi)^{\frac{1}{2}} \sigma} \exp \left\{ -\frac{\alpha^2}{2\sigma^2} \right\} \quad (A5)$$

Equation (A2) still holds for a given value of α .

The mean value of r^2 over the whole distribution of α is

$$\bar{r}^2(\sigma) = \frac{F^2}{(2\pi)^{\frac{1}{2}} \sigma} \int_{-\infty}^{\infty} (1 + w^2 - 2w \cos \alpha) \exp \left\{ -\frac{\alpha^2}{2\sigma^2} \right\} d\alpha. \quad (A6)$$

Using the result

$$\int_0^{\infty} e^{-a^2 x^2} \cos bx dx = \pi^{\frac{1}{2}} \exp \left\{ -b^2/4a^2 \right\} / 2a$$

this simplifies to

$$\bar{r}^2(\sigma) = F^2 [1 + w^2 - 2w \exp \left\{ -\frac{1}{2} \sigma^2 \right\}]; \quad (A7)$$

$$\frac{d\bar{r}^2}{dw} = 2F^2 (w - \exp \left\{ -\frac{1}{2} \sigma^2 \right\}); \quad (A8)$$

and the best weighting function is

$$w_0 = \exp \left\{ -\frac{1}{2} \sigma^2 \right\}, \quad (A9)$$

which approximates to (A4) with $\sigma = \alpha$, for small values of α .

Substituting this value back into (A7) gives

$$\bar{r}_0^2 = F^2 [1 - \exp \left\{ -\sigma^2 \right\}]. \quad (A10)$$

We are grateful to Dr Perutz and our colleagues and visitors at this Unit, both for access to their data prior to publication and for helpful discussion of the problems involved. One of us (D. M. B.) is indebted to the Medical Research Council for a scholarship.

References

- BLOW, D. M. (1958). *Proc. Roy. Soc. A*, **247**, 302.
BRAGG, W. L. & PERUTZ, M. F. (1954). *Proc. Roy. Soc. A*, **225**, 315.

- CRAMER, H. (1937). *Random Variables and Probability Distributions*. Cambridge: Cambridge University Press.
CRICK, F. H. C. & MAGDOFF, B. (1956). *Acta Cryst.* **9**, 901.
CRUICKSHANK, D. W. J. (1959). *Acta Cryst.* **2**, 65.
CULLIS, A. F., DINTZIS, H. M. & PERUTZ, M. F. (1957). *Conference on Haemoglobin*, p. 50. (National Academy of Sciences, Washington, 1958).
HARKER, D. (1956). *Acta Cryst.* **9**, 1.
KENDREW, J. C., BODO, G., DINTZIS, H. M., PARRISH, R. G., WYCKOFF, H. & PHILLIPS, D. C. (1958). *Nature, Lond.* **181**, 662.
PERUTZ, M. F. (1956). *Acta Cryst.* **9**, 867.

Acta Cryst. (1959). **12**, 802

A Unified Program for Phase Determination, Type $3P_1$

BY H. HAUPTMAN AND J. KARLE

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

(Received 5 March 1959)

The unified program for phase determination, valid for all the space groups and both the equal and unequal atom cases, is continued here. The present paper is concerned with the centrosymmetric space groups comprising type $3P_1$. A detailed procedure for phase determination is described for this type.

1. Introduction

This is the fourth in a series of papers concerned with a program for phase determination initiated by us (Karle & Hauptman, 1959, hereafter referred to as *1P*). The application of the new probability methods, based on the Miller indices as random variables, is made to the space groups of type $3P_1$, (Hauptman & Karle, 1953). This type consists of the eleven primitive centrosymmetric space groups in the hexagonal system. We present here a detailed procedure for phase determination which utilizes the same general formula and, at the same time, makes use of relationships among the structure factors characteristic of each space group.

2. Notation

The same notation as appears in *1P* (1959) is employed here.

3. Phase determining formulas

3.1. Basic formulas

$$B_{2,0}: \mathcal{E}'_{\mathbf{h}}{}^2 = 1 + \frac{4\pi\sigma_2^2}{2^{(p+q+2)/2} pq \Gamma\left(\frac{p+1}{2}\right) \Gamma\left(\frac{q+1}{2}\right) \sigma_4} \times \langle \lambda_{pk} \lambda_{q(\mathbf{h}+\mathbf{k})} \rangle_{\mathbf{k}} + R_{2,0}. \quad (3.1.1)$$

$$B_{3,0}: \mathcal{E}'_{\mathbf{h}_1} \mathcal{E}'_{\mathbf{h}_2} \mathcal{E}'_{\mathbf{h}_1+\mathbf{h}_2} = \frac{(2\pi)^{3/2} \sigma_2^3}{2^{(p+q+r+3)/2} pqr \Gamma\left(\frac{p+1}{2}\right) \Gamma\left(\frac{q+1}{2}\right) \Gamma\left(\frac{r+1}{2}\right) \sigma_4^{3/2}} \times \langle \lambda_{pk} \lambda_{q(\mathbf{h}_1+\mathbf{k})} \lambda_{r(\mathbf{h}_1+\mathbf{h}_2+\mathbf{k})} \rangle_{\mathbf{k}} - 2 \frac{\sigma_6}{\sigma_4^{3/2}} + \frac{\sigma_8^{1/2}}{\sigma_4} (\mathcal{E}'_{\mathbf{h}_1} \mathcal{E}'_{\mathbf{h}_1}{}'' + \mathcal{E}'_{\mathbf{h}_2} \mathcal{E}'_{\mathbf{h}_2}{}'' + \mathcal{E}'_{\mathbf{h}_1+\mathbf{h}_2} \mathcal{E}'_{\mathbf{h}_1+\mathbf{h}_2}{}'') + R_{3,0}. \quad (3.1.2)$$

3.2. Integrated formulas

$$I_{2,0}: \mathcal{E}'_{\mathbf{h}}{}^2 = 1 + \frac{2\sigma_2^2}{C_1^2(t) \sigma_4} \langle A_{t\mathbf{k}} A_{t(\mathbf{h}+\mathbf{k})} \rangle_{\mathbf{k}} + R'_{2,0}. \quad (3.2.1)$$

$$I_{3,0}: \mathcal{E}'_{\mathbf{h}_1} \mathcal{E}'_{\mathbf{h}_2} \mathcal{E}'_{\mathbf{h}_1+\mathbf{h}_2} = \frac{\sigma_2^3}{C_1^3(t) \sigma_4^{3/2}} \langle A_{t\mathbf{k}} A_{t(\mathbf{h}_1+\mathbf{k})} A_{t(\mathbf{h}_1+\mathbf{h}_2+\mathbf{k})} \rangle_{\mathbf{k}} - 2 \frac{\sigma_6}{\sigma_4^{3/2}} + \frac{\sigma_8^{1/2}}{\sigma_4} \times (\mathcal{E}'_{\mathbf{h}_1} \mathcal{E}'_{\mathbf{h}_1}{}'' + \mathcal{E}'_{\mathbf{h}_2} \mathcal{E}'_{\mathbf{h}_2}{}'' + \mathcal{E}'_{\mathbf{h}_1+\mathbf{h}_2} \mathcal{E}'_{\mathbf{h}_1+\mathbf{h}_2}{}'') + R'_{3,0}. \quad (3.2.2)$$

In these formulas, p, q, r and t are restricted to be positive. Ordinarily they are given values in the range 2-4.

The remainder terms are given in the appendix § 6 and in *1P* (1959). Equation (3.1.1) or (3.2.1) serves to determine the magnitudes of the structure factors $|\mathcal{E}'_{\mathbf{h}}|$ corresponding to the squared structure. By means of equation (3.1.2) or (3.2.2), the phases of these structure factors $\varphi'_{\mathbf{h}}$ may be determined. In the next