Acta Cryst. (1975). A**31**, 825

Real-Space Analogue of Reflexion Quartets

By J. Gassmann

Max-Planck-Institut für Biochemie, D-8033 Martinsried, Germany (BRD)

(*Received* 16 *May* 1975; *accepted* 16 *June* 1975)

The use of reflexion quartets in combination with reflexion triplets in direct methods is shown to be equivalent to the maximum-density concept introduced with the procedure of phase correction. For a specific form of phase correction it is shown that in addition to the triplet contributions it includes the quartet contributions to a phase indication. Probability formulae for the different cases of combined triplet–quartet contributions are derived. The possibilities for extension and improvement by using other forms of phase correction are given.

1. Introduction

'Relations of the second kind' (Schenk, 1973) and related procedures are known and used in programs for direct sign or phase determination (Stewart, 1970; Germain, Main & Woolfson, 1970). In a paper by Schenk (1973) the use of reflexion quartets in direct methods as an extension of the use of reflexion triplets was considered. In this paper for mainly empirical reasons the weighting or probability of the quartet condition,

$$\cos(\varphi_{h_1}+\varphi_{h_2}+\varphi_{h_3}+\varphi_{h_4}) \to 1\,, \tag{1}$$

where → means 'is probably' and $\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3+\mathbf{h}_4=0$, was combined with the weight of corresponding triple-product relations

$$\begin{aligned}\cos(\varphi_{h_1}+\varphi_{h_2}+\varphi_{h_3+h_4}) &\to 1\\ \cos(\varphi_{h_1}+\varphi_{h_3}+\varphi_{h_4+h_2}) &\to 1\\ \cos(\varphi_{h_1}+\varphi_{h_4}+\varphi_{h_2+h_3}) &\to 1\end{aligned} \tag{2}$$

to improve the reliability of the quartet relation. The result of the combined weighting scheme was called a 'strengthened quartet relationship'. Though used mainly with centrosymmetric structures where the relations concern signs instead of phases the usefulness for non-centrosymmetric structures was stated.

In a paper by Hauptman (1974) a selection rule and an estimation were given for those reflexion quartets which probably result in a negative value for the cosine invariant of (1). This was called a 'negative quartet relation'. A following paper by Schenk (1974) showed the application of these negative quartets to centrosymmetric structures.

In a series of papers (Hoppe & Gassmann, 1968; Hoppe, Gassmann & Zechmeister, 1970; Gassmann & Zechmeister, 1972) a procedure called 'phase correction' was introduced and shown to be useful not only for completion of partial structures but also to replace and substitute the normal extension and refinement procedures of direct methods. The theory of phase correction explains the differences and advantages compared with normal direct methods.

In numerous experiments and actual structure determinations this theory has been confirmed. For a given initial phase set, application of phase correction results in a wider range of convergence towards the correct phases compared with the normal phase-determination procedures (*e.g.* 'tangent refinement').

A major point in the development of phase correction was the concept of maximum electron density. In phase correction this maximum-density concept introduces phase values or phase changes which are markedly different from the normal phase-determination process. It will be shown below that this effect is due to the inclusion of quartet and higher-order relations if a specific form of phase correction is applied.

2. Notation

$\mathbf{E}_1=\mathbf{E}_{h_1}$ = normalized structure factor with index $\mathbf{h}_1$.

E_1 = normalized structure factor amplitude.

φ_1 = structure factor phase.

$\mathbf{E}_{2+3}=\mathbf{E}_{h_2+h_3}$ = normalized structure factor with index $\mathbf{h}_2+\mathbf{h}_3$.

$\mathbf{E}^c_h$ = structure factor calculated from modification process.

E^{obs}_h = observed structure factor amplitude.

$\langle\varphi\rangle, \langle\cos\varphi\rangle, \langle\mathbf{E}_{h'}\mathbf{E}_{h-h'}\rangle, \langle\mathbf{E}_h\rangle$ = average, expectation values.

$\varrho(\mathbf{r})$ = electron density.

$\varrho^*(\mathbf{r})$ = modified electron density.

$g(\varrho)$ = modification function.

N = number of atoms in unit cell (assuming equal atoms).

T, a, b, c = parameter and coefficients for cubic-term phase correction.

$$A=\frac{2}{\sqrt{N}}\,E_lE_jE_k\,.$$

$$B=\frac{2}{N}\,E_1E_2E_3E_4\,.$$

$$C=\frac{1}{N}\,E_1^2(E_2^2+E_3^2+E_4^2)\,.$$

$$D=\frac{1}{N}\,E_2^2E_3^2E_4^2/(E_2^2+E_3^2+E_4^2)=\frac{B^2}{4C}\,.$$

P = probability.

R_h, S_h = proportionality factors for structure factor convolutions.

I_0, I_1 = modified Bessel functions of the second kind.

$\rightarrow$ = 'is probably'.

3. Theoretical considerations

Most existing phase-determination procedures are applied in reciprocal space for several reasons. The main reasons are that

(*a*) probability estimates are given for the phases as functions of structure factor amplitudes and

(*b*) for a few initial phases the convolution process to get new phases is fast, compared with the equivalent operation in real space.

Nevertheless the investigation of direct methods and their interpretation is often much easier in real space, as will be demonstrated.

For ease of argument in our presentation we assume equal-atom structures. The change to different atom types can easily be done and involves mostly replacement of N (=atoms/unit cell) against more complicated coefficients.

We consider now different possibilities of modifying the electron density in real space. The simple form of squaring

$$\varrho^* = \varrho^2 \tag{3}$$

(Sayre, 1952) is the real-space equivalent of the convolution of structure factors in reciprocal space (Karle & Karle, 1966):

$$\mathbf{E}_h = S_h \sum_{h'} \mathbf{E}_{h'} \mathbf{E}_{h-h'} = \sqrt{N} \langle \mathbf{E}_{h'} \mathbf{E}_{h-h'} \rangle . \tag{4}$$

Deduced from statistical arguments the probability that a single contribution of the right side of (4) has a phase $\langle \varphi_h \rangle$ different from the actual phase φ_h of the structure factor $\mathbf{E}_h$ is given by (Cochran, 1955):

$$\Delta\varphi = \varphi - \langle \varphi_h \rangle ; \quad A = \frac{2}{\sqrt{N}} E_h E_{h'} E_{h-h'}$$

$$P(\Delta\varphi) = \exp \{A \cos (\Delta\varphi)\} / 2\pi I_0(A) . \tag{5}$$

The important point is the dependence on the phase difference

$$P(\Delta\varphi) \sim \exp \{A \cos (\Delta\varphi)\} . \tag{5a}$$

This relation can be extended (Cochran, 1955), equation (10)] if several phase indications, *i. e.* several contributions to the convolution, are present:

$$P(\Delta\varphi) = \exp \{A' \cos (\Delta\varphi)\} / 2\pi I_0(A') \tag{5b}$$

where

$$A' = \frac{2}{\sqrt{N}} |\mathbf{E}_h \sum_{h'}^{m} \mathbf{E}_{h'} \mathbf{E}_{h-h'}| . \tag{5c}$$

Generalizing these probability considerations one finds that the essential form arises from the expression:

$$P(\Delta\varphi) \sim \exp [-|\mathbf{E}_h - \langle \mathbf{E}_h^c \rangle|^2] . \tag{6}$$

The interesting, $\Delta\varphi$-dependent part of the probability is calculated as

$$P(\Delta\varphi) \sim \exp [\mathbf{E}_h \langle \mathbf{E}_{-h}^c \rangle + \mathbf{E}_{-h} \langle \mathbf{E}_h^c \rangle] = \exp [2|E_h^{\text{obs}} \langle E_h^c \rangle| \cos \Delta\varphi] , \tag{6a}$$

where $\Delta\varphi$ now means the phase difference between the actual and the calculated phase. The normalization of the probability gives the proportionality factor.

The squaring principle automatically leads one to higher powers of modification:

$$\varrho^* = \varrho^n ; \quad n > 2 . \tag{7}$$

The case $n = 3$, equivalent to

$$\varrho^* = \varrho^3 , \tag{7a}$$

shows the general argument involved. The corresponding double convolution in reciprocal space is:

$$\mathbf{E}_h = R_h \sum_{h'} \sum_{h''} \mathbf{E}_{h''} \mathbf{E}_{h'-h''} \mathbf{E}_{h-h'} = N \langle \langle \mathbf{E}_{h''} \mathbf{E}_{h'-h''} \rangle \mathbf{E}_{h-h'} \rangle . \tag{8}$$

For a single contribution

$$\langle \mathbf{E}_h^c \rangle = \frac{1}{N} \mathbf{E}_{h''} \mathbf{E}_{h'-h''} \mathbf{E}_{h-h'} ,$$

which leads to a probable-phase indication derived from (6*a*):

$$P(\varrho) \sim \exp [\mathbf{E}_h \langle \mathbf{E}_h^c \rangle + \mathbf{E}_{-h} \langle \mathbf{E}_h^c \rangle] = \exp \left[\frac{2}{N} E_h E_{h''} E_{h'-h''} E_{h-h'} \cos \delta \right]$$

$$P(\delta) = \exp (B \cos \delta) / 2\pi I_0(B) . \tag{9}$$

The distribution of $P(\delta)$ again is peaked at $\delta = 0$ like (5) for large values of $B = (2/N) E_h E_{h''} E_{h'-h''} E_{h-h'}$. However depending upon the value N this peak will not be so distinct, *i.e.* the probability is less, compared with the triple-product peak of (5).

The effect of modification with higher powers of ϱ is evident. The reciprocal-space representation involves multiple convolutions. For a single contribution of such a representation the phase indication is again calculated after (6*a*). Raising the power from n to $n+1$ means essentially the inclusion of a factor $E/\sqrt{N}$ into the exponent of (6*a*). Assuming a reasonable structure size, it is obvious that the precision of such a single phase indication decreases.

For the process of phase correction a general modification function was introduced (Hoppe & Gassmann, 1968):

$$\varrho^* = g(\varrho) \,.\, \varrho . \tag{10}$$

As explained in detail in a later paper (Hoppe, Gassmann & Zechmeister, 1970) this function $g(\varrho)$ must not necessarily be analytical if applied in direct space. For reciprocal-space application however, it is best to develop $g(\varrho)$ in a power series to see the convolutions involved:

$$\varrho^* = \sum_n a_n \varrho^n . \tag{10a}$$

The process of phase correction is successful if at least three terms of (10*a*) are present:

$$\varrho^* = a\varrho + b\varrho^2 + c\varrho^3 . \tag{10b}$$

Simple calculations give the coefficients a, b, c as:

$$a = \frac{1-2T}{1-T}, \quad b = \frac{1+T}{1-T}, \quad c = -\frac{1}{1-T} ;$$

where $0 \le T < 1$ is a parameter for changing the electron density (Hoppe & Gassmann, 1968). For simplicity and comparison with published results in reciprocal space (Hauptman, 1974)† we take $T = \frac{1}{2}$ and we get:

$$\varrho^* = 3\varrho^2 - 2\varrho^3 \tag{10c}$$

or

$$\varrho^* \sim \varrho^2 - \tfrac{2}{3}\varrho^3 \tag{10d}$$

which may be compared with the former density modifications (Fig. 1).

The form of the reciprocal-space representation of (10*c*) is:

$$\mathbf{E}_h^c = 3\sqrt{N}\langle \mathbf{E}_{h'}\mathbf{E}_{h-h'}\rangle - 2N\langle\langle \mathbf{E}_{h''}\mathbf{E}_{h'-h''}\rangle \mathbf{E}_{h-h'}\rangle . \tag{11}$$

As can be seen from Fig. 1 the cubic term, *i.e.* the double convolution, is necessary to keep an upper limit of electron density and to confine approximate densities to sensible physical values (see below).

What is now the phase indication from single contributions of (11)? If one assumes three structure factors which form one triple product of the double convolution $\mathbf{E}_{h''}\mathbf{E}_{h'-h''}\mathbf{E}_{h-h'}$ in (11) to be known, one obtains from reasoning similar to that set out in the literature the expected value of $\mathbf{E}_h$ for fixed values of the structure-factors products (Woolfson, 1954):

$$\langle \mathbf{E}_h^c \rangle = -\frac{2}{N}\mathbf{E}_{h''}\mathbf{E}_{h'-h''}\mathbf{E}_{h-h'}$$
$$+ \frac{3}{\sqrt{N}}(\mathbf{E}_{h''}\mathbf{E}_{h-h''} + \mathbf{E}_{h'-h''}\mathbf{E}_{h-h'+h''} + \mathbf{E}_{h'}\mathbf{E}_{h-h'}) . \tag{12}$$

In this equation are also included the double products of the single convolution arising from the structure factors $\mathbf{E}_{h''}$, $\mathbf{E}_{h'-h''}$ and $\mathbf{E}_{h-h'}$. These double products contain the structure factors $\mathbf{E}_{h-h''}$, $\mathbf{E}_{h-h'+h''}$ and $\mathbf{E}_{h'}$, which are also assumed to be known.

For clarity, indices are now replaced by numbers, one summation of numbers means vector addition of indices:

$$\mathbf{h} \triangleq 1; \quad \mathbf{h}'' \triangleq 2; \quad \mathbf{h}' - \mathbf{h}'' \triangleq 3; \quad \mathbf{h} - \mathbf{h}' \triangleq 4;$$

Equation (12) then reads:

$$\langle \mathbf{E}_1^c \rangle = \frac{3}{\sqrt{N}}(\mathbf{E}_2\mathbf{E}_{3+4} + \mathbf{E}_3\mathbf{E}_{2+4} + \mathbf{E}_4\mathbf{E}_{2+3}) - \frac{2}{N}\mathbf{E}_2\mathbf{E}_3\mathbf{E}_4 . \tag{12a}$$

The phase probability for this single contribution is

$$\begin{aligned} P(\varepsilon) &\sim \exp\left[2E_1\langle E_1^c\rangle \cos\varepsilon\right] \\ &= \exp\left[\frac{6}{\sqrt{N}}\{E_1E_2E_{3+4}\cos\alpha + E_1E_3E_{4+2} \right. \\ &\quad \times \cos\beta + E_1E_4E_{2+3}\cos\gamma\} \\ &\quad \left. - \frac{4}{N}E_1E_2E_3E_4\cos\delta\right] . \end{aligned} \tag{13}$$

The angles are defined in Fig. 2.

The exponent of (13) consists of three triple-product contributions and one quartet term. We now consider the different cases for the evaluation of (13).

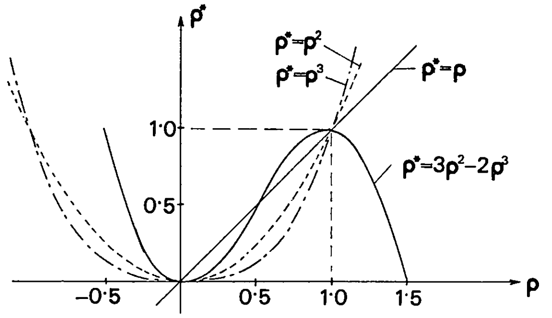

Fig. 1. Different types of density modifications which are applicable in reciprocal space because of the limited number of structure factor convolutions involved.

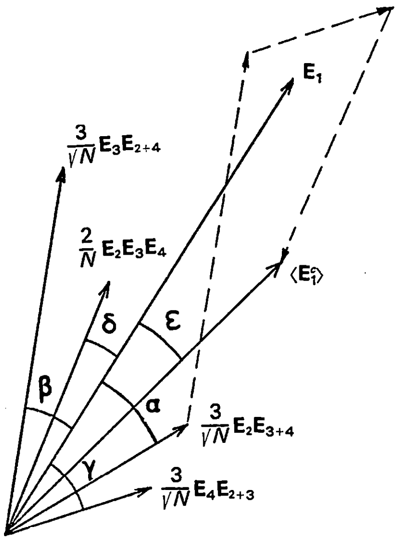

Fig. 2. Phase diagram representing the vector components to form the expected structure factor $\langle \mathbf{E}_1^c \rangle$.

---

† The derivation in this paper corresponds more closely to a value $T=0$: $\varrho^* = \varrho + \varrho^2 - \varrho^3$ which assumes a good estimate for E_h; see below.

(*a*) *Negative quartets*

If the triple products vanish because the structure factor amplitudes are zero:

$$E_{3+4}=E^{\text{obs}}_{h-h''}=0;\quad E_{2+3}=E^{\text{obs}}_{h'}=0;$$
$$E_{2+4}=E^{\text{obs}}_{h+h''-h'}=0;\qquad(14)$$

the only contribution to $\langle\mathbf{E}_1^c\rangle$ is the last term in (12*a*). Equation (13) then reads with a proper normalization denominator:

$$P(\varepsilon)=\exp\left[-2B\cos\delta\right]/2\pi I_0(2B).\qquad(13a)$$

Since $\mathbf{E}_2\mathbf{E}_3\mathbf{E}_4$ is the only contribution to $\langle\mathbf{E}_1^c\rangle$

$$\cos\varepsilon=\cos(\pi-\delta)=-\cos\delta\ \text{(see Fig. 2)}.$$

The so-called cosine invariant is then:

$$\langle\cos\varepsilon\rangle=\int_0^{2\pi}P(\varepsilon)\cos\varepsilon\,\mathrm{d}\varepsilon=-\frac{I_1(2B)}{I_0(2B)};\qquad(14)$$

In contrast to (9) the probability of (13*a*) is peaked at $\delta=\pi$, indicating a high probability that the product $\mathbf{E}_2\mathbf{E}_3\mathbf{E}_4$ has the opposite phase of $\mathbf{E}_1$. This is the case of a so-called 'negative quartet'.

(*b*) *Strengthened quartets*

If all (or at least one) of the triple products in (13) are large in comparison with the quartet term, the phase of $\langle\mathbf{E}_1^c\rangle$ tends to the phase of $\mathbf{E}_1$ and $P(\varepsilon)$ is peaked around $\varepsilon=0$. This is the case of a so-called 'strengthened quartet', because equations (5*b*) and (9) are peaked at $\Delta\varphi=0$, $\delta=0$ respectively, indicating that the phase of $\langle\mathbf{E}_1^c\rangle$ has a large probability of being near the phase of $\mathbf{E}_1$.

To get a measure of how the triple products strengthen the probability of (13) compared with the quartet probability of (9) one must replace $\mathbf{E}_{2+3}, \mathbf{E}_{2+4}, \mathbf{E}_{3+4}$ with their expected values from $\mathbf{E}_2, \mathbf{E}_3, \mathbf{E}_4$:

$$\mathbf{E}_4\mathbf{E}_{2+3}=(\mathbf{E}_2\mathbf{E}_3+\mathbf{E}_1\mathbf{E}_{-4})\,.\,\mathbf{E}_4\cdot\frac{1}{\sqrt{N}}$$
$$\mathbf{E}_2\mathbf{E}_{3+4}=(\mathbf{E}_3\mathbf{E}_4+\mathbf{E}_1\mathbf{E}_{-2})\,.\,\mathbf{E}_2\,.\frac{1}{\sqrt{N}}$$
$$\mathbf{E}_3\mathbf{E}_{4+2}=(\mathbf{E}_4\mathbf{E}_2+\mathbf{E}_1\mathbf{E}_{-3})\,.\,\mathbf{E}_3\,.\frac{1}{\sqrt{N}}\,.\qquad(15)$$

Summing these equations results in

$$\mathbf{E}_4\mathbf{E}_{2+3}+\mathbf{E}_2\mathbf{E}_{3+4}+\mathbf{E}_3\mathbf{E}_{4+2}$$
$$=[3\mathbf{E}_2\mathbf{E}_3\mathbf{E}_4+\mathbf{E}_1(E_2^2+E_3^2+E_4^2)]\,.\frac{1}{\sqrt{N}}\,.\qquad(15a)$$

Equation (12*a*) then reads:

$$\langle\mathbf{E}_1^c\rangle=[3\mathbf{E}_2\mathbf{E}_3\mathbf{E}_4+\mathbf{E}_1(E_2^2+E_3^2+E_4^2)]\,.\frac{3}{N}$$
$$-\frac{2}{N}\mathbf{E}_2\mathbf{E}_3\mathbf{E}_4=\frac{1}{N}[3\mathbf{E}_1(E_2^2+E_3^2+E_4^2)$$
$$+7\mathbf{E}_2\mathbf{E}_3\mathbf{E}_4]\qquad(16)$$

$$2E_1\langle E_1^c\rangle\cos\varepsilon=\frac{2}{N}[3E_1^2(E_2^2+E_3^2+E_4^2)$$
$$+7E_1E_2E_3E_4\cos\delta]\,.\qquad(16a)$$

The result for the probability is calculated by leaving out the constant term which changes only the normalization constant:

$$P(\varepsilon)\sim\exp\left[7\frac{2}{N}E_1E_2E_3E_4\cos\delta\right]=\exp\left[7B\cos\delta\right].\qquad(17)$$

For a specific value of ε there are two possible values δ_1 and δ_2 as can be seen from Fig. 3. The probability is then

$$P(\varepsilon)\sim\tfrac{1}{2}\{\exp(7B\cos\delta_1)+\exp(7B\cos\delta_2)\}\qquad(17a)$$

with a limiting value of

$$\sin\varepsilon_{\max}=-\cos\delta_0=\tfrac{7}{3}E_1E_2E_3E_4/E_1^2(E_2^2+E_3^2+E_4^2)\,.$$

The extreme values of the probability are therefore:

$$\varepsilon=0:\ P(0)\sim\cosh(7B)$$
$$\varepsilon=\varepsilon_{\max}:\ P(\varepsilon_{\max})$$
$$\sim\exp\left[-49BE_1E_2E_3E_4/3E_1^2(E_2^2+E_3^2+E_4^2)\right]$$
$$P(\varepsilon_{\max})\sim\exp\left(-\tfrac{98}{3}D\right),$$

where

$$D=\frac{1}{N}E_2^2E_3^2E_4^2/(E_2^2+E_3^2+E_4^2)\,.$$

The triple-product contributions alone without the quartet contribution would result in a circle with $\sin\varepsilon'_{\max}=9/7\sin\varepsilon_{\max}$ and a corresponding less distinct phase indication.

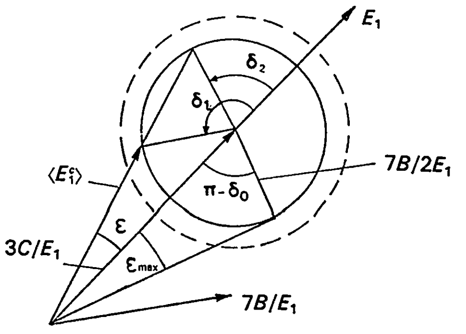

Fig. 3. Phase indication from combined triplet–quartet indications. The expected structure factor $\langle\mathbf{E}_1^c\rangle$ lies on a circle with radius $7B/2E_1$. For a specific value of ε two values of δ are possible. The maximum value for phase deviation is $\varepsilon_{\max}$. Increasing the component of $\langle\mathbf{E}_1^c\rangle$ in direction of $\mathbf{E}_1$ by a good estimate of $\mathbf{E}_1$ means further reduction of $\varepsilon_{\max}$. The dashed circle shows the phase indication from the triplet contributions alone.

The probability of (17*a*) is however an overestimation, because the expectation values in (15) require the knowledge of $\mathbf{E}_1$.†

Assuming only the structure factors $\mathbf{E}_2$, $\mathbf{E}_3$ and $\mathbf{E}_4$ to be known, the phase of $\mathbf{E}_1$ calculated after (8) differs by δ from the exact phase φ_1. This changes (16*a*) to:

$$2E_1\langle E_1^c\rangle \cos\varepsilon = \frac{2}{N}\,[3E_1^2(E_2^2+E_3^2+E_4^2) + 7E_1E_2E_3E_4]\cos\delta, \tag{16b}$$

and (17) to:

$$P(\varepsilon)\sim\exp\,[6C+7B]\cos\delta\,. \tag{17b}$$

The 'strengthening' of the phase indication can be seen from (17*b*) in comparison with (9).

If the contribution to $\mathbf{E}_{2+3}$, $\mathbf{E}_{2+4}$ and $\mathbf{E}_{3+4}$ containing the value $\mathbf{E}_1$ is taken as totally unknown and left out, the result is as follows:

Assuming in (12*a*) that:

$$E_{2+3}\geq\frac{E_2E_3}{\sqrt{N}}$$

$$E_{3+4}\geq\frac{E_3E_4}{\sqrt{N}}$$

$$E_{4+2}\geq\frac{E_4E_2}{\sqrt{N}}, \tag{18}$$

the single-triple-product phase indications α, β and γ after (5) are better or equal to the value δ of the quartet phase indication. A lower limit $P^*(\varepsilon)$ for the probability of (13) is then given if one replaces $\cos\alpha$, $\cos\beta$ and $\cos\gamma$ by $\cos\delta$:

$$P(\varepsilon)\geq P^*(\varepsilon)\sim\exp\left[\frac{6}{\sqrt{N}}\{E_1E_2E_{3+4}+E_1E_3E_{2+4} + E_1E_4E_{2+3}\}-\frac{4}{N}E_1E_2E_3E_4\right]\cos\delta\,. \tag{18d}$$

For the case of the equality in at least one of the equations (18) one gets:

$$P^*(\varepsilon)\sim\exp\frac{2}{N}\,E_1E_2E_3E_4\cos\delta=\exp B\cos\delta\,. \tag{19}$$

which is equal to the quartet indication of (9). The probability $P(\varepsilon)$ from (13), *i.e.* the resulting phase indication ε is therefore always superior to the quartet indication δ if at least one of the inequalities in (18) is fulfilled.

(*c*) *Weakened triplets*

Instead of naming the case of the just-assumed phase distribution a 'strengthened quartet relation', it could be called a 'weakened triple-product relation' because the negative contribution of the quartet term $E_1E_2E_3E_4$ to $\langle\mathbf{E}_1^c\rangle$ reduces the probability for a correct phase indication compared with the phase indication calculated from the triple-product contributions alone.

To get a measure of how the quartet term reduces the probability calculated in (13) compared with the probability of the triple products alone, one takes the average of the quartet contribution:

$$\langle P(\delta)\rangle=\frac{1}{I_0(2B)}\,.\,\exp\{-2B\langle\cos\delta\rangle\} =\frac{1}{I_0(2B)}\exp\left\{-2B\frac{I_1(2B)}{I_0(2B)}\right\}. \tag{20}$$

Defining

$$P(\alpha,\beta,\gamma)\sim\exp\left[\frac{6}{\sqrt{N}}\,(E_1E_2E_{3+4}\cos\alpha + E_1E_3E_{4+2}\cos\beta+E_1E_4E_{2+3}\cos\gamma\right]$$

as the probability arising from the triple products, one gets for the overall probability:

$$P(\varepsilon)=P(\alpha,\beta,\gamma)\exp\left(-2B\frac{I_1(2B)}{I_0(2B)}\right)/I_0(2B)\,. \tag{21}$$

For small values of B this may be written as:

$$P(\varepsilon)=P(\alpha,\beta,\gamma)\,.\,\exp\,(-3B^2) \tag{21a}$$

where the exponential term shows the decrease in probability for the triple products involved.

4. Further improved phase indications

The probability expressions derived above are due to a very specific form of phase correction. The question arises if it is possible to improve further the phase indications by generalizing the modification functions.

(*a*) *General cubic-term phase correction*

The general form of cubic-term phase correction is written in reciprocal space as:

$$\langle\mathbf{E}_1^c\rangle=a\mathbf{E}_1^{\rm est}+\frac{|b|}{\sqrt{N}}\,(\mathbf{E}_2\mathbf{E}_{3+4}+\mathbf{E}_3\mathbf{E}_{2+4}+\mathbf{E}_4\mathbf{E}_{2+3}) -\frac{|c|}{N}\,\mathbf{E}_2\mathbf{E}_3\mathbf{E}_4 \tag{22}$$

with

$$a\leq1;\quad b\geq1;\quad c\leq-1;$$

$\mathbf{E}_1^{\rm est}$ is an initial value for $\mathbf{E}_1$, not deduced from statistics. A calculation equivalent to the derivation of

---

† Furthermore because of the probabilistic nature of (12) and (15) there cannot be a boundary on the possible values of $\langle E_1^c\rangle$, implying that the consequences deduced [(17*a*) and Fig. 3)] are only an algebraic result. This has been pointed out by the referee.

(17*b*) leads to the following result:

$$2E_1\langle E_1^c\rangle \cos\varepsilon = 2aE_1^2\cos w + [2b\,.\,C+(3b+c)B]\cos\delta \quad (23)$$

where

$$C=\frac{1}{N}E_1^2(E_2^2+E_3^2+E_4^2)\,,\quad 3b+c=\frac{2+3T}{1-T}$$

and w is the phase difference of the estimate $\mathbf{E}_1^{\text{est}}$.

The important point is that the contribution connected with the second term in (23) gets smaller if a lower critical density T is assumed.

With a good estimate $\mathbf{E}_1^{\text{est}}$, corresponding to the first term in (23) the maximum probability will be moved away from the origin, mainly along the direction of $\mathbf{E}_1$. This means that the phase indication gets more distinct. Comparing (17*b*) and (23) gives an improvement of phase indication for:

$$2aE_1^2\cos w+[(3b+c)B + 2bC]\cos\delta \geq [7B+6C]\cos\delta\,; \quad (24)$$

with results in

$$\cos w \geq \frac{2}{N}\{\tfrac{5}{2}E_1E_2E_3E_4/E_1^2+E_2^2+E_3^2+E_4^2\}\,. \quad (24a)$$

This result means that for an approximate phase indication better than $\cos w\sim 1/N$, application of the general cubic-term phase correction is superior to the special cubic-term phase correction of (10*c*). At the same time it indicates for what values of approximate start phases one may expect convergence to the correct result through application of phase correction.

(*b*) *General real-space phase correction*

As can be seen from Fig. 1 even with a general cubic-term phase correction it is not possible to modify the electron density ϱ correctly over the whole range of ϱ. For negative regions and for positive regions $\varrho>1$ it should be $\varrho^*=0$ (see Fig. 4). Such a behaviour could be approximated by including higher coefficients according to (10*a*). This would result in alternating values for the coefficients of successive powers of ϱ. The reciprocal-space treatment corresponding to (11) would result in multiple convolutions. Assuming a modification function

$$\varrho^*=a\varrho+b\varrho^n+c\varrho^{n+1}\,, \quad (10e)$$

simple calculations give the coefficients as:

$$a=\frac{Q-(n+1)\,.\,T^{n-1}}{Q-n\,.\,T^{n-1}}$$

$$b=\frac{Q}{Q-n\,.\,T^{n-1}}$$

$$c=-\frac{Q-T^{n-1}}{Q-nT^{n-1}}$$

with

$$Q=\sum_{k=0}^{n-1}T^k\,.$$

Requiring $Q=(n+1)T^{n-1}$ we get the general form of (10*c*):

$$\varrho^*=(n+1)\varrho^n-n\varrho^{n+1};\quad n\geq 2\,. \quad (10f)$$

The assumption of a good estimate for $\mathbf{E}_h$ implies values for the coefficients a, b, c, by approaching $T=0$, which results in:

$$\varrho^*=\varrho+\varrho^n-\varrho^{n+1}\,. \quad (10g)$$

Between these extreme modifications are the intermediate forms with $0<T<\frac{1}{2}$.

The characteristic analytical behaviour of (10*f*) is shown in Fig. 4. Depending on the even or odd value of n, negative density regions are modified differently. The critial value T for increasing or decreasing the density can be shown to change with the power n as:

$$\varrho^*(T)=T\approx 1-\frac{2}{n(n+1)}\,;$$

The reciprocal-space analogue of (10*f*) contains n-tuple convolutions which might be difficult to calculate and to program in an actual phase determination procedure.

It is however much simpler to apply such modifications in real space. The electron density ϱ itself is modified to the corrected density ϱ^* and then transformed to get structure factor values with a backward Fourier transformation. The corrected density ϱ^* automatically contains all powers of ϱ pertaining to the chosen modification form. A simple real-space modification, the so called 'linear form' (see Fig. 4) has been shown to be successful.

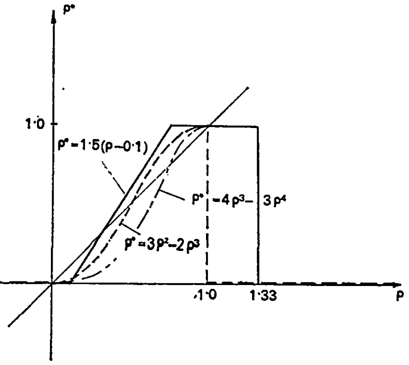

Fig. 4. Characteristic differences between modification functions $\varrho^*=a\varrho+b\varrho^n-|c|\varrho^{n+1}$ for the assumed values $n=2$ and $n=3$. The modification with $n=3$ needs better start phases since $T=0{\cdot}77$ and treats regions $\varrho>1{\cdot}43$ and $\varrho<-0{\cdot}43$ incorrectly. The 'linear-form' modification is derived from the power-series modification. For the range $1\leq\varrho\leq 8/6$ the value of ϱ^* has been fixed to unity to allow for the density of oxygen atoms.

The inclusion of higher powers of ϱ certainly increases the range of convergence to the correct phase. It is a safe assumption that the condition for the convergence angle w for all reasonable real-space modifications is relaxed to:

$$\cos w \geq 0 . \qquad (24b)$$

Some of the consequences arising from the modification described above will now be discussed.

5. Application for phase determination

The cyclic process of phase correction has been described (Gassmann & Zechmeister, 1972; Gassmann, 1976). The inclusion of very high powers of ϱ due to a nonanalytical form of the modification and the elimination of physically nonsensible regions (nonlocal weighting) leads to spurious, weak phase indications for some reflexions. To exclude these phases from the next cycle of phase correction, one must set an acceptance criterion for a newly determined phase. From (6*a*) one recognizes that the essential factor is $E^{\text{obs}}\langle E^c\rangle$, which must lie above a certain limit to have a reasonable probability for a correct phase indication. Only if

$$E^{\text{obs}}\langle E^c\rangle \geq \text{limit} \qquad (24)$$

is a phase accepted as input to the next phase-correction cycle.

The application of real-space phase correction has been simplified by the use of the fast Fourier transform (FFT) (see *e.g.* Barrett & Zwick, 1970). The use of the FFT-algorithm may however lead to difficulties since an intrinsic assumption of the algorithm is the nonexistence of frequencies (= structure factors) higher than $M/2$, where M is the array size of the FFT-transform. A modified electron density ϱ^* may contain very high resolution structure factors. These values are then intermixed with low resolution structure factors, if the array size M is not large enough (Gassmann & Hofmann, 1975).

Conclusions

The concept of maximum electron density introduced with phase correction leads to improved phase indications. The probability formulae derived show the possibilities of phase correction in structure determination.

The author is grateful to Professor W. Hoppe for his continuous support of this work and thanks Dr P. Colman for stimulating discussions.

References

Barrett, A. N. & Zwick, M. (1971). *Acta Cryst.* A**27**, 6–11.
Cochran, W. (1955). *Acta Cryst.* **8**, 473–478.
Gassmann, J. (1976). *Crystallographic Computing.* II, Edited by F. R. Ahmed. Copenhagen: Munksgaard. To be published.
Gassmann, J. & Hofmann, H. (1975). In preparation.
Gassmann, J. & Zechmeister, K. (1972). *Acta Cryst.* A**28**, 270–280.
Germain, G., Main, P. & Woolfson, M. M. (1970). *Acta Cryst.* B**26**, 274–285.
Hauptman, H. (1974). *Acta Cryst.* A**30**, 472–476.
Hoppe, W. & Gassmann, J. (1968). *Acta Cryst.* B**24**, 97–107.
Hoppe, W., Gassmann, J. & Zechmeister, K. (1970). *Crystallographic Computing*, pp. 26–36. Edited by F. R. Ahmed. Copenhagen: Munksgaard.
Karle, J. & Karle, I. L. (1966). *Acta Cryst.* **21**, 849–859.
Sayre, D. (1952). *Acta Cryst.* **5**, 60–65.
Schenk, H. (1973). *Acta Cryst.* A**29**, 77–82.
Schenk, H. (1974). *Acta Cryst.* A**30**, 477–481.
Stewart, J. M. (1970). *Crystallographic Computing*, pp. 71–74. Edited by F. R. Ahmed. Copenhagen: Munksgaard.
Woolfson, M. M. (1954). *Acta Cryst.* **7**, 61–64.