*Acta Cryst.* (1972). B28, 1601

# A Multiformula Procedure for a Single-Solution Approach to the Phase Problem. I. Generalized Product Relationships*

BY CHUN-CHE TSAI AND D. M. COLLINS

*Department of Chemistry, Cornell University, Ithaca, New York* 14850, *U.S.A.*

(*Received* 25 *June* 1971)

Any member of a set of complex structure factors can be approximately determined by various combinations of the others. In noncentrosymmetric structures, this property has been put to use mainly by applying phase-determining formulae related to the self-convolution, or squared structure. These are seen to comprise the simplest class of formulae obtainable from a generalized product relationship. Specific formulae that emphasize use of *a priori* phase information are developed for application. The underlying equivalence of centrosymmetric and noncentrosymmetric phase-determining formulae is demonstrated.

## 1. Introduction

It has been widely accepted that in crystal-structure problems of moderate complexity, the structure moduli obtained from X-ray diffraction experiments can be effectively employed in finding explicit restraints on the phases of structure factors. This has led to the development of algebraic relationships which make it possible to solve centrosymmetric phase problems by a routine arithmetic procedure. Hauptman & Karle (1953) have detailed such a procedure, and although their work has been subject to criticism and modification by many workers in the field it remains in use unchanged in essential concept, especially their $\sum_2$ formula or Sayre equation (Sayre, 1952). In spite of development of the basic triple-product relationship for noncentrosymmetric structures (Hughes, 1953; Cochran, 1955; Karle & Hauptman, 1956), the noncentrosymmetric phase problem remains beyond the kind of procedure so readily applied in the centrosymmetric case.

One purpose of this paper is to begin a unification between respective approaches to the solution of the centrosymmetric and noncentrosymmetric phase problems. Clearly, the principal information available for attacking any phase problem is the set of structure moduli. If by various criteria a structure is judged to be centrosymmetric, powerful relationships employing such quantities as $(|E_{\mathbf{k}}|^2-\overline{|E|^2})$ can be used to estimate some phases. An apparently trivial distinction can be made in observing that we can, equally well, write $(E_{\mathbf{k}}^2-\overline{E^2})$, since in the centrosymmetric case $2\varphi_{\mathbf{k}}=0$ for all $\mathbf{k}$. Nevertheless, the distinction is useful; we may now interpret the employment of $|E|^2$ in phase-determining formulae for the centrosymmetric case as use not only of structure moduli, but also of (perhaps trivial) phase information. These observations lead to the possibility that phase-determining formulae can be found that utilize *a priori* phase information, whether determined by permissible assignment or symmetry requirements, and are applicable to noncentrosymmetric structures, paralleling the existing centrosymmetric formulae.

For most structures, symmetry-imposed restraints on particular phases can be found and uniquely stated in the form $2\varphi=\alpha\pi$, where $\alpha$ is a rational number; indeed, it is just this information which has proved to be an invaluable aid in solving centrosymmetric phase problems. In the present paper, we find several phase-determining relationships which relate complex structure factors, and we emphasize those relationships that employ the information resident in $2\varphi$ to advantage. With the exception of (2.26), each equation in §2 (especially 2.9, 2.14, 2.19, 2.22, and 2.23) has its obvious analog among the accepted centrosymmetric formulae. Equation (2.26) is a new general product formula; the more general form (3.4) will probably be of little practical value other than in multiple-term equations such as (2.22). These relationships are summarized by equation (3.8), which retains its validity in the presence or absence of a center of symmetry.

## 2. Tangent formulae

In space group $P1$, assuming equal atoms, we take the normalized structure factor associated with reciprocal lattice vector $\mathbf{k}$, to be defined by

$$E_{\mathbf{k}}=|E_{\mathbf{k}}|\exp(i\varphi_{\mathbf{k}})=\frac{1}{N^{1/2}}\sum_{j=1}^{N}\exp(2\pi i\mathbf{k}\cdot\mathbf{r}_j) \qquad (2.1)$$

where the $N$ atoms in the unit cell lie at the ends of the position vectors $\mathbf{r}_j$. Then,

$$|E_{\mathbf{k}}|\cos\varphi_{\mathbf{k}}=\frac{1}{N^{1/2}}\sum_{j=1}^{N}\cos 2\pi\mathbf{k}\cdot\mathbf{r}_j\,, \qquad (2.2)$$

$$|E_{\mathbf{k}}|\sin\varphi_{\mathbf{k}}=\frac{1}{N^{1/2}}\sum_{j=1}^{N}\sin 2\pi\mathbf{k}\cdot\mathbf{r}_j\,. \qquad (2.3)$$

* Based on Report No. 1552 of the Materials Science Center, Cornell University. Part of this work was presented at the February 1971 meeting of the American Crystallographic Association in Columbia, South Carolina, U.S.A.

Evidently, in (2.3) we may replace $\mathbf{k}$ by $\mathbf{h}-\mathbf{k}$, where $\mathbf{h}$ is fixed, and write

$$\langle(|E_{\mathbf{k}}|\cos\varphi_{\mathbf{k}})(|E_{\mathbf{h}-\mathbf{k}}|\sin\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}=$$

$$\left\langle\frac{1}{N}\left[\sum_{j,j'=1}^{N}\cos 2\pi\mathbf{k}\,.\,\mathbf{r}_j\sin 2\pi(\mathbf{h}-\mathbf{k})\,.\,\mathbf{r}_{j'}\right]\right\rangle_{\mathbf{k}}, \quad (2.4)$$

where $\langle\ldots\rangle_{\mathbf{k}}$ indicates an average over the vectors $\mathbf{k}$. If we write out the right-hand side of (2.4) in detail, and assume that for a reasonably large number of vectors $\mathbf{k}$, only the term $\cos^2 2\pi\mathbf{k}\,.\,\mathbf{r}_j\sin 2\pi\mathbf{h}\,.\,\mathbf{r}_j$ survives; then, since $\cos^2\alpha=\frac{1}{2}+\frac{1}{2}\cos 2\alpha$, it is seen that (2.4) is equivalent to

$$\langle|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|\cos\varphi_{\mathbf{k}}\sin\varphi_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}=\frac{1}{2N^{1/2}}|E_{\mathbf{h}}|\sin\varphi_{\mathbf{h}}, \quad (2.5)$$

provided the average over $\mathbf{k}$ consists of a sufficiently large number of terms.

In analogous fashion, we can show that

$$\langle|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|\sin\varphi_{\mathbf{k}}\cos\varphi_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}=\frac{1}{2N^{1/2}}|E_{\mathbf{h}}|\sin\varphi_{\mathbf{h}}, \quad (2.6)$$

$$\langle|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|\cos\varphi_{\mathbf{k}}\cos\varphi_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}=\frac{1}{2N^{1/2}}|E_{\mathbf{h}}|\cos\varphi_{\mathbf{h}}, \quad (2.7)$$

$$\langle|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|\sin\varphi_{\mathbf{k}}\sin\varphi_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}=-\frac{1}{2N^{1/2}}|E_{\mathbf{h}}|\cos\varphi_{\mathbf{h}}\,. \quad (2.8)$$

Combining (2.5) through (2.8) in a straightforward way, we have the standard tangent formula of Karle & Hauptman (1956), namely,

$$\tan\varphi_{\mathbf{h}}=T_{11}(\varphi_{\mathbf{h}})=\frac{\langle|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|\sin(\varphi_{\mathbf{k}}+\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}}{\langle|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|\cos(\varphi_{\mathbf{k}}+\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}}\,. \quad (2.9)$$

We have interpreted the employment of $(|E|^2-1)$ in centrosymmetric phase problems as a means of utilizing the phase information resident in $2\varphi$ $(=0)$, and it is natural to anticipate a similar use for $2\varphi$ in noncentrosymmetric problems. In order to find equations explicit in $2\varphi$ and of the type (2.9), we start with the square of (2.1). The real and imaginary parts of (2.1) and its square may be combined to give equations of the type (2.4); the procedure used in finding (2.5) then leads to:

$$\langle(|E_{\mathbf{k}}|^2\cos 2\varphi_{\mathbf{k}})(|E_{\mathbf{h}-2\mathbf{k}}|\sin\varphi_{\mathbf{h}-2\mathbf{k}})\rangle_{\mathbf{k}}=\frac{1}{2N}|E_{\mathbf{h}}|\sin\varphi_{\mathbf{h}}, \quad (2.10)$$

$$\langle(|E_{\mathbf{k}}|^2\sin 2\varphi_{\mathbf{k}})(|E_{\mathbf{h}-2\mathbf{k}}|\cos\varphi_{\mathbf{h}-2\mathbf{k}})\rangle_{\mathbf{k}}=\frac{1}{2N}|E_{\mathbf{h}}|\sin\varphi_{\mathbf{h}}, \quad (2.11)$$

$$\langle(|E_{\mathbf{k}}|^2\cos 2\varphi_{\mathbf{k}})(|E_{\mathbf{h}-2\mathbf{k}}|\cos\varphi_{\mathbf{h}-2\mathbf{k}})\rangle_{\mathbf{k}}=\frac{1}{2N}|E_{\mathbf{h}}|\cos\varphi_{\mathbf{h}}, \quad (2.12)$$

$$\langle(|E_{\mathbf{k}}|^2\sin 2\varphi_{\mathbf{k}})(|E_{\mathbf{h}-2\mathbf{k}}|\sin\varphi_{\mathbf{h}-2\mathbf{k}})\rangle_{\mathbf{k}}=-\frac{1}{2N}|E_{\mathbf{h}}|\cos\varphi_{\mathbf{h}}\,. \quad (2.13)$$

Combining (2.10) through (2.13) in a straightforward way, we have

$$\tan\varphi_{\mathbf{h}}=T_{21}(\varphi_{\mathbf{h}})=\frac{\langle|E_{\mathbf{k}}|^2|E_{\mathbf{h}-2\mathbf{k}}|\sin(2\varphi_{\mathbf{k}}+\varphi_{\mathbf{h}-2\mathbf{k}})\rangle_{\mathbf{k}}}{\langle|E_{\mathbf{k}}|^2|E_{\mathbf{h}-2\mathbf{k}}|\cos(2\varphi_{\mathbf{k}}+\varphi_{\mathbf{h}-2\mathbf{k}})\rangle_{\mathbf{k}}}\,. \quad (2.14)$$

Proceeding as before we can also show that

$$\langle(|E_{\mathbf{k}}|^2\cos 2\varphi_{\mathbf{k}})(|E_{\mathbf{h}-\mathbf{k}}|^2\sin 2\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}=\frac{1}{N}|E_{\mathbf{h}}|^2\sin 2\varphi_{\mathbf{h}}-\frac{1}{2N^{3/2}}|E_{2\mathbf{h}}|\sin\varphi_{2\mathbf{h}}, \quad (2.15)$$

$$\langle(|E_{\mathbf{k}}|^2\sin 2\varphi_{\mathbf{k}})(|E_{\mathbf{h}-\mathbf{k}}|^2\cos 2\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}=\frac{1}{N}|E_{\mathbf{h}}|^2\sin 2\varphi_{\mathbf{h}}-\frac{1}{2N^{3/2}}|E_{2\mathbf{h}}|\sin\varphi_{2\mathbf{h}}, \quad (2.16)$$

$$\langle(|E_{\mathbf{k}}|^2\cos 2\varphi_{\mathbf{k}})(|E_{\mathbf{h}-\mathbf{k}}|^2\cos 2\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}=\frac{1}{N}|E_{\mathbf{h}}|^2\cos 2\varphi_{\mathbf{h}}-\frac{1}{2N^{3/2}}|E_{2\mathbf{h}}|\cos\varphi_{2\mathbf{h}}, \quad (2.17)$$

$$\langle(|E_{\mathbf{k}}|^2\sin 2\varphi_{\mathbf{k}})(|E_{\mathbf{h}-\mathbf{k}}|^2\sin 2\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}=-\frac{1}{N}|E_{\mathbf{h}}|^2\cos 2\varphi_{\mathbf{h}}+\frac{1}{2N^{3/2}}|E_{2\mathbf{h}}|\cos\varphi_{2\mathbf{h}}\,. \quad (2.18)$$

If we follow Hauptman (1970) in ignoring terms of order $1/N^{3/2}$, the combination of (2.15) through (2.18) leads to

$$\tan 2\varphi_{\mathbf{h}}=T_{22}(2\varphi_{\mathbf{h}})=\frac{\langle|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|^2\sin(2\varphi_{\mathbf{k}}+2\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}}{\langle|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|^2\cos(2\varphi_{\mathbf{k}}+2\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}}\,. \quad (2.19)$$

If, on the other hand, we retain terms of order $1/N^{3/2}$, the combination of (2.15) through (2.18) leads to

$$\tan\varphi_{2\mathbf{h}}=T_{2,22}(\varphi_{2\mathbf{h}})=\frac{2|E_{\mathbf{h}}|^2\sin 2\varphi_{\mathbf{h}}-N\langle|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|^2\sin(2\varphi_{\mathbf{k}}+2\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}}{2|E_{\mathbf{h}}|^2\cos 2\varphi_{\mathbf{h}}-N\langle|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|^2\cos(2\varphi_{\mathbf{k}}+2\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}}}, \quad (2.20)$$

which was to be expected from the work of Cochran (1954). For if we follow the derivation of his equation (13*c*), and concern ourselves not with the Patterson function, $P(\mathbf{r})$, but with the function $A(\mathbf{r})$, where

$$A(\mathbf{r})=\sum_{\mathbf{h}}|F_{\mathbf{h}}|^2\exp(i2\varphi_{\mathbf{h}})\exp(-2\pi i\mathbf{h}\,.\,\mathbf{r}), \quad (2.21)$$

in the noncentrosymmetric case we find

$$|E_{2\mathbf{h}}|\exp(i\varphi_{2\mathbf{h}})=N^{1/2}\{2|E_{\mathbf{h}}|^2\exp(i2\varphi_{\mathbf{h}})-N\langle|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|^2\exp[i(2\varphi_{\mathbf{k}}+2\varphi_{\mathbf{h}-\mathbf{k}})]\rangle_{\mathbf{k}}\} \quad (2.22)$$

from which (2.20) immediately follows. We know (2.9) to be useful when the vectors $\mathbf{k}$ are restricted such that

$|E_{\mathbf{k}}|$ is large, and we expect (2.14) and (2.19) to be similarly useful, but (2.20) as written is valid only when the averages extend over the complete set of vectors $\mathbf{k}$.

Finally, noting equation (2) of Cochran (1955) and the appropriate single-term special cases of (2.9) and (2.14), we have

$$\tan \varphi_{2\mathbf{h}} = T_2(\varphi_{2\mathbf{h}}) = \sin 2\varphi_{\mathbf{h}} / \cos 2\varphi_{\mathbf{h}} , \qquad (2.23)$$

which gives the most probable value of $\varphi_{2\mathbf{h}}$, provided $2\varphi_{\mathbf{h}}$ is known and there are no other contributors to the averages in (2.9) and (2.14).

A generalization that encompasses (2.9), (2.14), and (2.19) can be found in the following way:

$$\left\langle E_{\mathbf{k}}^p E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}^q \right\rangle_{\mathbf{k}} = \left\langle \left[ \frac{1}{N^{1/2}} \sum_{j=1}^{N} \exp\left(2\pi i \mathbf{k} \cdot \mathbf{r}_j\right) \right]^p \times \left\{ \frac{1}{N^{1/2}} \sum_{j=1}^{N} \exp\left[2\pi i \left(\mathbf{h} - \frac{p}{q}\mathbf{k}\right) \cdot \mathbf{r}_j\right] \right\}^q \right\rangle_{\mathbf{k}} . \qquad (2.24)$$

Therefore, when $p \equiv 0 \pmod q$, $p$ and $q$ being integers,

$$\left\langle E_{\mathbf{k}}^p E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}^q \right\rangle_{\mathbf{k}} = \frac{p!}{\left(\frac{p}{q}!\right) q N^{p/2}} \times \left[ \frac{1}{N^{1/2}} \sum_{j=1}^{N} \exp\left(2\pi i \mathbf{h} \cdot \mathbf{r}_j\right) \right]^q + \langle R \rangle_{\mathbf{k}} , \qquad (2.25)$$

which, if we neglect $\langle R \rangle_{\mathbf{k}}$ (see below), gives our generalization

$$N_{pq} E_{\mathbf{h}}^q = \left\langle E_{\mathbf{k}}^p E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}^q \right\rangle_{\mathbf{k}} , \qquad (2.26)$$

where

$$N_{pq} = \frac{p!}{\left(\frac{p}{q}!\right) q N^{p/2}} . \qquad (2.27)$$

As a tangent formula, (2.26) becomes

$$\tan(q\varphi_{\mathbf{h}}) = T_{pq}(q\varphi_{\mathbf{h}}) = \frac{\left\langle |E_{\mathbf{k}}|^p \left|E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q \sin\left(p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right) \right\rangle_{\mathbf{k}}}{\left\langle |E_{\mathbf{k}}|^p \left|E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q \cos\left(p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right) \right\rangle_{\mathbf{k}}} . \qquad (2.28)$$

### 3. A general procedure for finding tangent formulae

The procedures of the preceding section are ill-suited to developing a complete set of phase relationships. We now show that it is possible, in principle, to find as large a variety of tangent formulae as might be desired. We follow exactly the work of Vaughan (1959), except the quantities $E_{\mathbf{k}_j}$ are taken as observables and the quantity to be estimated is not assumed to be an origin invariant. Vaughan's equation (12) then becomes

$$G \simeq \sum_{a_1,\ldots,a_n}^{v=m} C_{a_1,\ldots,a_n}^{(v)} (E_{\mathbf{k}_1})^{a_1} \ldots (E_{\mathbf{k}_n})^{a_n} , \qquad (3.1)$$

where $v = a_1 + \ldots + a_n$ is the total order, $m$ is the selected maximum order, $C_{a_1,\ldots,a_n}^{(v)}$ are constants which may be determined by least-squares methods, and $G$ is any product of normalized structure factors. Again, following Vaughan we make the following approximations:

(*a*) All terms will be omitted from (3.1) for which the covariances

$$\overline{G(E_{-\mathbf{k}_1})^{a_1} \ldots (E_{-\mathbf{k}_n})^{a_n}}$$

are zero.

(*b*) All terms that have a common value for the above covariances have the same coefficient in (3.1).

We limit ourselves, for the present, to considering equations of one coefficient in accordance with (*a*) and (*b*). For example, to estimate $E_{\mathbf{h}}^r$ in space group $P1$, we consider for our first case

$$\overline{E_{\mathbf{h}}^r E_{-\mathbf{k}_j}^p} > 0 .$$

It immediately follows that $\mathbf{k}_j = (r/p)\mathbf{h}$ and

$$E_{\mathbf{h}}^r \simeq C E_{\frac{r}{p}\mathbf{h}}^p , \qquad (3.2)$$

$$\tan(r\varphi_{\mathbf{h}}) \simeq \frac{\sin p\varphi_{\frac{r}{p}\mathbf{h}}}{\cos p\varphi_{\frac{r}{p}\mathbf{h}}} ,$$

which is equivalent to (2.23) when $r=1$ and $p=2$. For our second case, we consider

$$\overline{E_{\mathbf{h}}^r E_{-\mathbf{k}_i}^p E_{-\mathbf{k}_j}^q} > 0 . \qquad (3.3)$$

It immediately follows that $\mathbf{k}_j = (r/q)\mathbf{h} - (p/q)\mathbf{k}_i$ and

$$E_{\mathbf{h}}^r \simeq C \sum_{\mathbf{k}} E_{\mathbf{k}}^p E_{r\,\mathbf{h}-\frac{p}{q}\mathbf{k}}^q , \qquad (3.4)$$

$$\tan(r\varphi_{\mathbf{h}}) \simeq \frac{\sum_{\mathbf{k}} |E_{\mathbf{k}}|^p \left|E_{\frac{r}{q}\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q \sin\left(p\varphi_{\mathbf{k}} + q\varphi_{q\,\mathbf{h}-\frac{p}{q}\mathbf{k}}\right)}{\sum_{\mathbf{k}} |E_{\mathbf{k}}|^p \left|E_{\frac{r}{q}\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q \cos\left(p\varphi_{\mathbf{k}} + q\varphi_{q\,\mathbf{h}-\frac{p}{q}\mathbf{k}}\right)} , \qquad (3.5)$$

which is equivalent to (2.28) if $r=q$.

It is clear that a single-term formula obtained by the above procedure is valid, in the sense that any coefficient in (3.1) may be evaluated. Nevertheless, if we seek a reliable determination of some particular phase information it will, in general, be necessary to construct a multiple-term formula, *e.g.* (2.20). In the general case, we may have to evaluate the relative importance of each of many terms and arbitrarily select the most significant few for use in the estimation of phases. The problem then is to determine each coefficient in (3.1) and compute $G$.

We can readily justify (2.26) along these lines. In its derivation, the term $\langle R \rangle_{\mathbf{k}}$ was ignored, subject to the condition that $p$ be an integral multiple of $q$. In the context of this section, it is necessary to evaluate co-

variances of the type (3.3) to validate our neglect of $\langle R\rangle_{\mathbf{k}}$. Recalling the form of the left-hand side of (2.25), we must evaluate

$$\overline{E^r_{-\frac{q}{r}\mathbf{h}} E^p_{\mathbf{k}} E^q_{\mathbf{h}-\frac{p}{q}\mathbf{k}}} \tag{3.6}$$

for integral $r$ and fixed $\mathbf{h}$, $\mathbf{k}$, $p$, and $q$. Straightforward, but tedious, computation reveals that if $r=q$, expression (3.6) is of order $N^{-p/2}$, otherwise it is of order $N^{-(p+1)/2}$ or smaller. Since the term retained in (2.26) corresponds to $r=q$, $\langle R\rangle_{\mathbf{k}}$, if it is nonzero, contains terms that correspond to $r\neq q$ and may be neglected, provided $N$ is sufficiently large.

To illustrate the general need for multiple-term formulae, we consider (3.2) and (3.4) and the possibility of setting $r=1$ and $p=q=2$ therein. This corresponds to obtaining an estimate of $\varphi_{2\mathbf{h}}$ from values of $2\varphi$ for other structure factors. The equation in which the coefficients would be determined is

$$E_{2\mathbf{h}} = C'_1 E^2_{\mathbf{h}} + C'_2 \sum_{\mathbf{k}} E^2_{\mathbf{k}} E^2_{\mathbf{h}-\mathbf{k}} \tag{3.7a}$$

or

$$E_{2\mathbf{h}} = C_1 E^2_{\mathbf{h}} + C_2 \langle E^2_{\mathbf{k}} E^2_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}} \,. \tag{3.7b}$$

Equations (3.7*a*&*b*) are to be compared with (2.22). Consideration of the covariances $\overline{E_{2\mathbf{h}} E^2_{-\mathbf{h}}} = N^{-1/2}$ and $\overline{E_{2\mathbf{h}} \langle E^2_{-\mathbf{k}} E^2_{-\mathbf{h}+\mathbf{k}}\rangle_{\mathbf{k}}} = N^{-3/2}$ makes it clear that if one of the terms in (3.7*b*) is to be omitted it must be the second. But if we follow reasoning analogous to that presented by Vaughan for the centrosymmetric case, it is seen that the permissible estimation of $E_{2\mathbf{h}}$, using only $E^2_{\mathbf{h}}$, is of quite modest reliability as measured by its formal variance. On the other hand, use of both terms in (3.7*b*) leads to a significantly improved reliability for estimating $E_{2\mathbf{h}}$, provided the second term contains a large number of contributors. Since the physical background of the problem demands priority of inclusion in (3.1) for lowest-order terms, we conclude that a realistic estimation of $E_{2\mathbf{h}}$ from the set $\{E^2\}$ will require an equation of at least two terms, and one of them will be of the form $CE^2_{\mathbf{h}}$.

The connection between phase-determining formulae for the centrosymmetric and the noncentrosymmetric cases is clear. Although in writing (3.1) we have taken $(E_{\mathbf{k}j})$ as observables to be raised to the powers $a_j$, we may equally well take as observables $(E^{a_j}_{\mathbf{k}j})$. The observables, as well as the quantity to be estimated, must of course be reduced by their mean values. The most general form of (3.4) is, then,

$$E^r_{\mathbf{h}} - \overline{E^r} \simeq C \sum_{\mathbf{k}} (E^p_{\mathbf{k}} - \overline{E^p}) \left(E^q_{\frac{r}{q}\mathbf{h}-\frac{p}{q}\mathbf{k}} - \overline{E^q}\right), \tag{3.8}$$

and it applies to both centrosymmetric and noncentrosymmetric structures. For example, appropriate choice of $r$, $p$, and $q$ in the centrosymmetric case makes (3.8) equivalent to the $\sum_2$, $\sum_3$, or $\sum_4$ formula of Hauptman & Karle (1953).

## 4. Approximate formulae and figures of merit

In order to develop additional relationships useful at the start of phase determination, we follow portions of §3 of Karle & Karle (1966), but start with

$$E^q_{\mathbf{h}} = N^{-1}_{pq} \left\langle E^p_{\mathbf{k}} E^q_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right\rangle_{\mathbf{k}}, \tag{4.1}$$

a restatement of (2.26). We assume* the quantity $\left(-q\varphi_{\mathbf{h}} + p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right)$ is distributed about zero and that for the largest $|E|$ values it is small. Then,

$$1 \simeq Q N^{-1}_{pq} \left\langle E^p_{\mathbf{k}} E^q_{\mathbf{h}-\frac{p}{q}\mathbf{k}} E^{-q}_{\mathbf{h}} \right\rangle_{\mathbf{k}}, \tag{4.2}$$

where we may restrict $\mathbf{k}$ such that $|E_{\mathbf{k}}|$ is large; this restriction will be denoted by averaging over the vectors $\mathbf{k}_r$. Replacement of each exponential implied in (4.2) by the first term of its series expansion leads to

$$Q \simeq N_{pq} \left\langle |E_{\mathbf{k}}|^p \left|E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q |E_{\mathbf{h}}|^{-q} \right\rangle^{-1}_{\mathbf{k}_r}. \tag{4.3}$$

From (4.2) and (4.3) we find

$$\exp(iq\varphi_{\mathbf{h}}) \simeq \frac{\left\langle |E_{\mathbf{k}}|^p \left|E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q \exp\left[i\left(p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right)\right]\right\rangle_{\mathbf{k}_r}}{\left\langle |E_{\mathbf{k}}|^p \left|E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q \right\rangle_{\mathbf{k}_r}}, \tag{4.4}$$

which gives us two new formulae:

$$\cos(q\varphi_{\mathbf{h}}) \simeq C_{pq}(q\varphi_{\mathbf{h}}) = \frac{\left\langle |E_{\mathbf{k}}|^p \left|E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q \cos\left(p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right)\right\rangle_{\mathbf{k}_r}}{\left\langle |E_{\mathbf{k}}|^p \left|E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q \right\rangle_{\mathbf{k}_r}}, \tag{4.5}$$

$$\sin(q\varphi_{\mathbf{h}}) \simeq S_{pq}(q\varphi_{\mathbf{h}}) = \frac{\left\langle |E_{\mathbf{k}}|^p \left|E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q \sin\left(p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right)\right\rangle_{\mathbf{k}_r}}{\left\langle |E_{\mathbf{k}}|^p \left|E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q \right\rangle_{\mathbf{k}_r}}. \tag{4.6}$$

If we consider the imaginary part of (4.2) we have

$$\left\langle |E_{\mathbf{k}}|^p \left|E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q \sin\left(-q\varphi_{\mathbf{h}} + p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right)\right\rangle_{\mathbf{k}_r} \simeq 0 \,. \tag{4.7}$$

Expansion of the sine function in (4.7) and retention of the first term gives

$$\left\langle |E_{\mathbf{k}}|^p \left|E_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right|^q \left(-q\varphi_{\mathbf{h}} + p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}}\right)\right\rangle_{\mathbf{k}_r} \simeq 0 \,, \tag{4.8}$$

* This assumption and the validity of allowing only large $|E|$ values in (4·2) may be justified by following § 3·2 of Karle & Karle (1966) and replacing each equation by the corresponding more general form; our equation (4·12) is to be compared with (3·25) of Karle and Karle. These authors have pointed out the approximate nature of this reasoning and the inherent loss of accuracy in formulae dependent upon restriction of $\mathbf{k}$.

which upon rearrangement yields a general angle-sum formula:

$$q\varphi_{\mathbf{h}} \simeq \sum_{pq}^{wa} = \frac{\left\langle |E_{\mathbf{k}}|^p \left| E_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right|^q \left( p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right) \right\rangle_{\mathbf{k}_r}}{\left\langle |E_{\mathbf{k}}|^p \left| E_{\mathbf{h}-\frac{o}{q}\mathbf{k}} \right|^q \right\rangle_{\mathbf{k}_r}} . \tag{4.9}$$

If for the vectors $\mathbf{k}_r$ the values of $|E|$ are about the same, then (4.9) becomes

$$q\varphi_{\mathbf{h}} \simeq \sum_{pq}^{a} = \left\langle p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right\rangle_{\mathbf{k}_r} , \tag{4.10}$$

and under the same conditions (2.28) can be written

$$\tan(q\varphi_{\mathbf{h}}) \simeq \frac{\left\langle \sin\left( p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right) \right\rangle_{\mathbf{k}_r}}{\left\langle \cos\left( p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right) \right\rangle_{\mathbf{k}_r}} . \tag{4.11}$$

Restrictions must be placed on the angles in (4.8), (4.9), and (4.10) to make the equations meaningful. In (4.8), it is sufficient to impose the condition:

$$-\pi < -q\varphi_{\mathbf{h}} + p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \leq \pi$$

for each vector $\mathbf{k}$. In (4.9) and (4.10) it is necessary to impose procedural restrictions. As in (4.8), each individual angle sum is adjusted to be greater than $-\pi$ and not greater than $\pi$. We may interpret the right-hand sides of (4.9) and (4.10) to be weighted averages; application of either then begins with the calculation (subject to the foregoing condition) of four sums: (1*a*) the sum of positive contributors, (1*b*) the sum of the corresponding weights, (2*a*) the sum of nonpositive contributors, and (2*b*) the sum of the corresponding weights. If $A=(1a)/(1b)$ is greater than $B=(2a)/(2b)$ by more than $\pi$, $2\pi$ is added to $B$. The desired result is then given by

$$q\varphi_{\mathbf{h}} = [A \times (1b) + B \times (2b)]/[(1b) + (2b)] .$$

It remains to find a means of estimating the goodness of a phase determination as obtained from (2.28). In following work by Cochran (1955), and the extension of it by Karle & Karle (1966), we take as the probability distribution for $q\varphi_{\mathbf{h}}$:

$$P(q\varphi_{\mathbf{h}}) = \{2\pi I_o[\alpha_{pq}(\mathbf{h})]\}^{-1} \times \exp\{\alpha_{pq}(\mathbf{h}) \cos[q\varphi_{\mathbf{h}} - \beta_{pq}(\mathbf{h})]\} , \tag{4.12}$$

where

$$\alpha_{pq}(\mathbf{h}) = \left\{ \left[ \sum_{\mathbf{k}_r} \kappa_{pq}(\mathbf{h},\mathbf{k}) \cos\left( p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right) \right]^2 + \left[ \sum_{\mathbf{k}_r} \kappa_{pq}(\mathbf{h},\mathbf{k}) \sin\left( p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right) \right]^2 \right\}^{1/2} , \tag{4.13}$$

$$\kappa_{pq}(\mathbf{h},\mathbf{k}) = 2\frac{N_{pq}}{p!\langle |E_{\mathbf{h}}|^{2q}_{\mathbf{h}} \rangle} |E_{\mathbf{h}}|^q |E_{\mathbf{k}}|^p \left| E_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right|^q , \tag{4.14}$$

$$\beta_{pq}(\mathbf{h}) = \tan^{-1} \frac{\left\langle |E_{\mathbf{k}}|^p \left| E_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right|^q \sin\left( p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right) \right\rangle_{\mathbf{k}_r}}{\left\langle |E_{\mathbf{k}}|^p \left| E_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right|^q \cos\left( p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right) \right\rangle_{\mathbf{k}_r}} , \tag{4.15}$$

and $I_o$ is a modified Bessel function (see Abramowitz & Stegun, 1965). It is emphasized that (4.14) is for equal atoms. Inspection of (4.12) shows that the strength of any phase calculation can be estimated by the value of $\alpha$. However, the variation of $\alpha$ with the number of contributors in the calculation of a phase leads us to propose a normalized form of $\alpha$, a consistency index given by

$$\gamma_{pq}(\mathbf{h}) = \alpha_{pq}(\mathbf{h}) / \sum_{\mathbf{k}_r} \kappa_{pq}(\mathbf{h},\mathbf{k}) , \tag{4.16}$$

which can be simplified to

$$\gamma_{pq}(\mathbf{h}) = (A_{pq}^2 + B_{pq}^2)^{1/2} / \sum_{\mathbf{k}_r} |E_{\mathbf{k}}|^p \left| E_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right|^q , \tag{4.17}$$

and is equivalent to

$$\gamma_{pq}(\mathbf{h}) = (S_{pq}^2 + C_{pq}^2)^{1/2} . \tag{4.18}$$

In (4.17)

$$A_{pq} = \sum_{\mathbf{k}_r} |E_{\mathbf{k}}|^p \left| E_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right|^p \cos\left( p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right) , \tag{4.19}$$

and

$$B_{pq} = \sum_{\mathbf{k}_r} |E_{\mathbf{k}}|^p \left| E_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right|^q \sin\left( p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-\frac{p}{q}\mathbf{k}} \right) . \tag{4.20}$$

A straightforward rearrangement of (4.16) leads to

$$\gamma = \left[ \frac{\sum_{\mathbf{k},\mathbf{k}'} \kappa(\mathbf{k})\kappa(\mathbf{k}') \cos(\xi_{\mathbf{k}} - \xi_{\mathbf{k}'})}{\sum_{\mathbf{k},\mathbf{k}'} \kappa(\mathbf{k})\kappa(\mathbf{k}')} \right]^{1/2} , \tag{4.21}$$

where $\xi_{\mathbf{k}} = [p\varphi_{\mathbf{k}} + q\varphi_{\mathbf{h}-(p/q)\mathbf{k}}]$. Clearly $\gamma$ is a function of the agreement between individual indications of a phase; and its value may range between zero and unity. The value of $\gamma$ will tend towards unity when the contributors to the right-hand side of (2.28) are in substantial agreement among themselves, and it will tend towards zero when these contributors are in substantial disagreement, that is, the value of $\gamma$ is a measure of the internal consistency of (2.28) in its application.

## 5. Conclusion

Several phase-determining relationships for space group $P1$ have been presented. As demanded by the requirements of practicality, the forms we have found are quite generally applicable to all noncentrosymmetric problems, provided equations (3.4) and (3.8) are equivalent in application, *i.e.*, provided the mean values of (3.8) are sufficiently near zero. This condition may not be met in particular cases due to special distributions of atoms, those having a tendency toward an

inversion center at the origin being the most frequently encountered. In any case, it is easy to show that all phase relationships of the preceding sections have their analogous forms, which include correction terms to compensate for nonzero mean values, as in (3.8); but in normal noncentrosymmetric problems, these correction terms will be small and can be neglected safely.

Unfortunately, in certain space groups, the special reflection subsets that provide *a priori* phase information at the start of phase determination yield nonzero values for $\overline{E^p}$. For example, in space group $P2_12_12_1$ the three principal projections have $\overline{E^2}=0$, but the corresponding value for space group $P222$ is $\overline{E^2}=1$. In our further remarks, we consider only problems where $\overline{E^2}\simeq 0$, both for the three-dimensional data set and for the subsets of special symmetry; the general case will be dealt with elsewhere.

We are experimenting with a multiformula procedure for the application of our phase relationships, and excellent preliminary results indicate the following outline as remaining the likely core of our computer formalism for the implementation of the procedure.

(i) *The calculation of values for* $2\varphi$

Two formulae are used to generate values for all $2\varphi$ from the subset values of known $2\varphi$:

$$2\varphi_{\mathbf{h}}{}^{*}\simeq \sum_{22}^{wa}=\frac{\langle |E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|^2(2\varphi_{\mathbf{k}}+2\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}_r}}{\langle |E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|^2\rangle_{\mathbf{k}_r}}, \tag{5.1}$$

$$\cos 2\varphi_{\mathbf{h}}\dagger\simeq C_{22}=\frac{\langle |E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|^2\cos(2\varphi_{\mathbf{k}}+2\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}_r}}{\langle |E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|^2\rangle_{\mathbf{k}_r}}. \tag{5.2}$$

A more useful approximate‡ form of $C_{22}$ is used in practice:

$$\cos 2\varphi_{\mathbf{h}}\simeq C_{22}\simeq \frac{\langle |E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|^2\cos 2\varphi_{\mathbf{k}}\cos 2\varphi_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}_r}}{\langle |E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|^2\cos^2 2\varphi_{\mathbf{k}}\rangle_{\mathbf{k}_r}}. \tag{5.3}$$

(ii) *Phase assignment*

From among the reflections for which $2\varphi$ is unambiguously known, such reflections as may be used to specify the origin and enantiomorph are chosen, and each one is arbitrarily assigned one of the two possible values for $\varphi$. Also included in the table of initial phases, $\varphi_{\mathbf{h}}$, are those determined by other means, *e.g.*, the application of (2.20) or the $B_{3,0}$ formula of Karle & Hauptman (Karle & Hauptman, 1958; Karle, 1970).

(iii) *The calculation of values for* $\varphi$

Two formulae are used in the calculation of values for all $\varphi$:

$$\tan\varphi_{\mathbf{h}}\simeq T'_{21}=\frac{\langle w_{\mathbf{k}}w_{\mathbf{h}-2\mathbf{k}}|E_{\mathbf{k}}|^2|E_{\mathbf{h}-2\mathbf{k}}|\sin(2\varphi_{\mathbf{k}}+\varphi_{\mathbf{h}-2\mathbf{k}})\rangle_{\mathbf{k}_r}}{\langle w_{\mathbf{k}}w_{\mathbf{h}-2\mathbf{k}}|E_{\mathbf{k}}|^2|E_{\mathbf{h}-2\mathbf{k}}|\cos(2\varphi_{\mathbf{k}}+\varphi_{\mathbf{h}-2\mathbf{k}})\rangle_{\mathbf{k}_r}}, \tag{5.4}$$

$$\tan\varphi_{\mathbf{h}}\simeq T'_{11}=\frac{\langle w_{\mathbf{k}}w_{\mathbf{h}-\mathbf{k}}|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|\sin(\varphi_{\mathbf{k}}+\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}_r}}{\langle w_{\mathbf{k}}w_{\mathbf{h}-\mathbf{k}}|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|\cos(\varphi_{\mathbf{k}}+\varphi_{\mathbf{h}-\mathbf{k}})\rangle_{\mathbf{k}_r}}. \tag{5.5}$$

Equations (2.9) and (2.14) have been modified to give (5.4) and (5.5), where $w$ is a discrimination function. In accord with the customary usage of $w$, albeit implicit, $w$ is taken to be unity unless by reason of uncertainty in the value of $\varphi$, or one of its multiples, it should be zero. As an alternative to this step function, a continuous function such as

$$w_{\mathbf{h}}=\tanh[\tfrac{1}{2}\gamma(\mathbf{h})\alpha(\mathbf{h})] \tag{5.6}$$

probably would be computationally more efficient. A phase, the average of values obtained from the tangent formulae (5.4) and (5.5), is taken as determined when there is good agreement between the results. After a suitably large number of phases are thus determined, the results of (5.5) are taken as the set of tangent-refined phases.

This multiformula approach to the phase problem provides an unambiguous route to a single-solution set of phases. The explicit use of *a priori* phase information, in effect, introduces new phase constraints that serve to extend the range of problems amenable to solution by direct methods. Indeed, we expect this approach to yield the correct phases in cases where solution of the phase problem would be uncertain or, at best, difficult with only the basic product relationships (*i.e.*, those where $p=q=1$).

We wish to express our thanks to Professors J. L. Hoard and R. E. Hughes for their support and cooperation while this work was being done. One of us (C.T.) wishes to thank Drs H. Hauptman, J. Karle, and I. L. Karle, and Professor M. M. Woolfson for many helpful discussions and for their encouragement during the NATO Summer School in Parma, Italy, September 1970. This work was supported in part by Grants No. GM09370 and GM14832 from the National Institutes of Health, General Medical Sciences, by National Science Foundation Grants Nos. GP-13351 and 6710X, and by the Advanced Research Projects Agency through the Cornell University Materials Science Center.

* It may be necessary to arbitrarily restrict $2\varphi$ to the range $0\leq 2\varphi\leq\pi$, for example, as in the case where the *a priori* phase information is limited to $2\varphi=0$ or $\pi$. Clearly, the restricted values of $2\varphi$ then correspond to $|2\varphi|$ or $\cos 2\varphi$.

† In the case where the space group is one of an enantiomorphous pair, $S_{22}$ also may be applicable, but then the straightforward use of $T_{22}$ is indicated.

‡ This approximation to $C_{22}$ is rigorously correct if $\langle\cos^2 2\varphi_{\mathbf{k}}\rangle_{\mathbf{k}_r}=1$; it becomes invalid as $\langle\cos^2 2\varphi_{\mathbf{k}}\rangle_{\mathbf{k}_r}\to 0$.

### References

ABRAMOWITZ, M. & STEGUN, I. (1965). Editors of *Handbook of Mathematical Functions*. New York: Dover Publications.
COCHRAN, W. (1954). *Acta Cryst.* **7**, 581.
COCHRAN, W. (1955). *Acta Cryst.* **8**, 473.
HAUPTMAN, H. (1970). *Acta Cryst.* B**26**, 531.
HAUPTMAN, H. & KARLE, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal.* A.C.A. Monograph No. 3. Pittsburgh: Polycrystal Book Service.
HUGHES, E. W. (1953). *Acta Cryst.* **6**, 871.
KARLE, J. (1970). *Acta Cryst.* B**26**, 1614.
KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635.
KARLE, J. & HAUPTMAN, H. (1958). *Acta Cryst.* **11**, 264.
KARLE, J. & KARLE, I. L. (1966). *Acta Cryst.* **21**, 849.
SAYRE, D. (1952). *Acta Cryst.* **5**, 60.
VAUGHAN, P. A. (1959). *Acta Cryst.* **12**, 981.


*Acta Cryst.* (1972). B**28**, 1607

# The Crystal and Molecular Structure of Carbonatodiamminecopper(II), $Cu(NH_3)_2CO_3$

BY MARTHA H. MEYER, PHIRTU SINGH, WILLIAM E. HATFIELD AND DEREK J. HODGSON*

*Department of Chemistry, University of North Carolina, Chapel Hill, North Carolina* 27514, *U.S.A.*

(*Received* 13 *September* 1971).

The crystal and molecular structure of carbonatodiamminecopper(II), $Cu(NH_3)_2CO_3$, has been determined from three-dimensional X-ray data collected by counter methods. The material crystallizes in the space group $P2_1/c$ of the monoclinic system, with four molecules in a cell of dimensions $a=5{\cdot}640$ (2), $b=10{\cdot}579$ (3), $c=7{\cdot}462$ (3) Å, and $\beta=97°49$ (1)′. The observed and calculated densities are 2·35 (2) and 2·372 g.cm$^{-3}$. Least-squares refinement of the structure has yielded a final value for the conventional $R$ factor, on $F$, of 0·032 for 934 data greater than their estimated standard deviations. The geometry at the copper atom is approximately a tetragonal pyramid whose basal plane consists of the two ammine nitrogen atoms and two oxygen atoms from a single carbonate. The copper atom of each formula unit is bonded to the adjacent unit through the terminal carbonate oxygen atom, with a Cu–O bond length of 2·303 (2) Å. The Cu–N distances are 1·971 (2) and 1·984 (2) Å. The Cu–O distances to the oxygen atoms in the bidentate carbonate are 1·986 (2) and 1·989 (2) Å. The N–Cu–N angle is 97·80 (12)°, while the chelated O–Cu–O angle is 66·38 (8)°. There is hydrogen-bonding between the ammine hydrogen atoms and the carbonate oxygen atoms of adjacent units.

## Introduction

The structure of carbonatodiamminecopper(II), $Cu(NH_3)_2CO_3$, has been previously investigated by Hanic (1962, 1963), using two-dimensional X-ray data, and the spectroscopic properties of the complex have been discussed by Tomlinson & Hathaway (1968). Carbonatodiamminecopper(II) is also of considerable magnetic interest, as described in our recent examination of its low-temperature magnetic properties (Jeter, Hodgson & Hatfield, 1972).

The crystal structure reported by Hanic (1963) indicates that apparently chemically equivalent bonds, such as the two Cu–N bonds, have significantly different lengths. While it is true that two-dimensional data cannot be expected to give great precision in the atomic parameters, the reported values of these bond lengths are so disparate as to be either of enormous chemical interest or to cast doubt on the validity of the model. We have, therefore, undertaken an accurate three-dimensional single-crystal X-ray structural determination of the complex.

* Author to whom correspondence should be addressed.

## Experimental

Suitable crystals were obtained by adding 10 ml of 95% ethanol to a saturated solution of basic copper carbonate, $Cu(CO_3).Cu(OH)_2$, in 35 ml concentrated ammonium hydroxide solution and removing the resulting turquoise precipitate by filtration. The filtrate was allowed to stand for six days, after which the dark-blue, rhombic crystals were separated by filtration and were air-dried. The crystals were found to decompose slowly, turning green after a week's exposure to moist air; they could be preserved indefinitely in a desiccator.

On the basis of precession and Weissenberg photography, the crystals were assigned to the monoclinic system, and examination of the $0kl$, $hk0$, $h0l$, and $h1l$ zones demonstrated systematic absences of $l$ odd in the $h0l$ zone and $k$ odd for $0k0$ reflections; hence, the previous (Hanic, 1963) space group assignment of $P2_1/c$ was confirmed. The cell constants, obtained by the least-squares procedure described below, are $a=5{\cdot}640$ (2), $b=10{\cdot}579$ (3), $c=7{\cdot}462$ (3) Å, and $\beta=97{\cdot}82$ (1)°. The observations were made at 25°C, with the wavelength assumed to be $\lambda(\text{Mo } K\alpha_1)=0{\cdot}7093$ Å. A density of 2·372 g.cm$^{-3}$ calculated for four formula