Acta Cryst. (1957). **10**, 515

A Unified Algebraic Approach to the Phase Problem. II. Space Group *P*1

By J. Karle and H. Hauptman

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

(*Received* 5 *March* 1957)

The unified algebraic approach to the phase problem yields, under suitable circumstances, explicit formulas for the magnitudes of certain structure invariants. These lead to a procedure for phase determination in the space group *P*1. Least-squares adjustments for refining the phases so determined are also described. Some simple examples are given which illustrate the application of the procedure.

1. Introduction

The unified algebraic method described in a previous paper (Hauptman & Karle, 1957) hereafter referred to as I, is here applied to space group *P*1. Formulas which lead to a procedure for phase determination are thus obtained. The relationship between this approach and that of the joint probability distributions previously described (Hauptman & Karle, 1953) will be studied in subsequent papers.

Although, as in I, the algebraic approach yields formulas having exact validity only in the case that the structure consist of N identical point atoms, this restriction can be readily removed. In fact the joint probability distribution plays an important role in modifying the formulas in the case that the structure contains unequal atoms. However, in this paper we treat only the case of N identical point atoms so that the method and formulas appear in their simplest forms. It is to be emphasized that in this case the formulas have exact, not merely probable, validity. As in I, the restriction that no six atoms be rationally dependent is required for our formulas to be rigorously true. However, it will be apparent that the procedure to be described is ordinarily not limited by this requirement.

Since the magnitudes of the structure factors determine only the magnitudes of the structure invariants (Hauptman & Karle, 1956, Main Theorem 7·3) the fundamental formula relates the magnitude of a particular structure invariant to the observed intensities. In space group *P*1 no single phase is a structure invariant. Hence an important problem to be considered in this paper is the evaluation of single phases when only the magnitudes of certain structure invariants are known. To this end a simple but important identity will be employed.

2. Phase-determining formulas

We list here all formulas which are derived in this paper.

$$|E_{\mathbf{h}}|^2-1 = N\langle(|E_{\mathbf{k}}|^2-1)(|E_{\mathbf{h}-\mathbf{k}}|^2-1)\rangle_{\mathbf{k}} . \quad (2\cdot1)$$

$$E_{\mathbf{h}_1}E_{\mathbf{h}_2} = N\langle(|E_{\mathbf{k}}|^2-1)E_{\mathbf{h}_1+\mathbf{k}}E_{\mathbf{h}_2-\mathbf{k}}\rangle_{\mathbf{k}}+(1/N^{\frac{1}{2}})E_{\mathbf{h}_1+\mathbf{h}_2} . \quad (2\cdot1a)$$

$$|E_{\mathbf{h}_1}E_{\mathbf{h}_2}E_{\mathbf{h}_3}|\cos(\varphi_1+\varphi_2+\varphi_3) = (N^{\frac{3}{2}}/2)\langle(|E_{\mathbf{k}}|^2-1)(|E_{\mathbf{h}'+\mathbf{k}}|^2-1)(|E_{\mathbf{h}_1+\mathbf{h}_2+\mathbf{k}}|^2-1)\rangle_{\mathbf{k}} + (1/N^{\frac{1}{2}})(|E_{\mathbf{h}_1}|^2+|E_{\mathbf{h}_2}|^2+|E_{\mathbf{h}_3}|^2-2), \quad \mathbf{h}' = \mathbf{h}_1 \text{ or } \mathbf{h}_2,\ \mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3 = 0 . \quad (2\cdot2)$$

$$|E_{\mathbf{h}_1}E_{\mathbf{h}_2}E_{\mathbf{h}_3}|\cos(\varphi_1+\varphi_2+\varphi_3) = 2N^{\frac{3}{2}}\langle|E_{\mathbf{k}}E_{\mathbf{k}'}E_{\mathbf{k}''}E_{\mathbf{h}_1-\mathbf{k}}E_{\mathbf{h}_2-\mathbf{k}'}E_{\mathbf{h}_3-\mathbf{k}''}| \times\cos(\varphi_{\mathbf{k}}+\varphi_{\mathbf{k}'}+\varphi_{\mathbf{k}''}) \times\cos(\varphi_{\mathbf{h}_1-\mathbf{k}}+\varphi_{\mathbf{h}_2-\mathbf{k}'}+\varphi_{\mathbf{h}_3-\mathbf{k}''})\rangle_{\mathbf{k},\mathbf{k}',\mathbf{k}''}-2N^{\frac{1}{2}}, \quad \mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3 = \mathbf{k}+\mathbf{k}'+\mathbf{k}'' = 0 . \quad (2\cdot3)$$

$$E_{\mathbf{h}} = N^{\frac{1}{2}}\langle E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}} . \quad (2\cdot4)$$

$$\langle|E_{\mathbf{k}}|^2\rangle_{\mathbf{k}} = 1 . \quad (2\cdot5)$$

In these equations the φ's are the phases of the corresponding E's, the normalized structure factors. Of these formulas, (2·4) has been previously derived by Hughes (1953) and (2·1) by Hauptman & Karle (1955). Evidently (2·5) is the special case $\mathbf{h} = 0$ of (2·4). The remaining equations are new, although Vaughan (1956) has obtained, by means of the Patterson superposition method, an approximate formula which resembles (2·2). It should be noted that (2·3) and (2·4) are supplementary formulas in the sense that (2·3) gives an improved value for the magnitude of a structure invariant once the approximate values of a large number of structure invariant magnitudes are known, while (2·4) gives an improved value for a phase once the approximate values of a large number of phases are known. Although (2·3) and (2·4) are rigorously correct it must be emphasized that, as a practical matter, they are not likely to be useful in any scheme for phase determination except possibly in the very final stages. The reason for this lies in the fact that initially (2·2) can be expected to yield reliable values for the magnitudes of those structure invariants $\varphi_1+\varphi_2+\varphi_3$ corresponding to the larger products $|E_1E_2E_3|$. Hence the required averages in (2·3) and (2·4) can be carried out not over all vectors $\mathbf{k}$, $\mathbf{k}'$, $\mathbf{k}''$ (as required by (2·3) and (2·4)), but rather over only such vectors for which the corresponding $|E_{\mathbf{k}}|$'s are

large. However, it can be shown (Karle & Hauptman, 1956) that in this case (2·3) and (2·4) may be replaced by

$$|E_{\mathbf{h}_1}E_{\mathbf{h}_2}E_{\mathbf{h}_3}|\cos(\varphi_1+\varphi_2+\varphi_3)\approx$$
$$2N^{\frac{3}{2}}\langle|E_{\mathbf{k}}E_{\mathbf{k}'}E_{\mathbf{k}''}E_{\mathbf{h}_1-\mathbf{k}}E_{\mathbf{h}_2-\mathbf{k}'}E_{\mathbf{h}_3-\mathbf{k}''}|$$
$$\times\cos(\varphi_{\mathbf{k}}+\varphi_{\mathbf{k}'}+\varphi_{\mathbf{k}''})\cos(\varphi_{\mathbf{h}_1-\mathbf{k}}+\varphi_{\mathbf{h}_2-\mathbf{k}'}+\varphi_{\mathbf{h}_3-\mathbf{k}''})\rangle_{\mathbf{k},\mathbf{k}',\mathbf{k}''}$$
$$\langle|E_{\mathbf{k}}E_{\mathbf{k}'}E_{\mathbf{k}''}E_{\mathbf{h}_1-\mathbf{k}}E_{\mathbf{h}_2-\mathbf{k}'}E_{\mathbf{h}_3-\mathbf{k}''}|^2\rangle_{\mathbf{k},\mathbf{k}',\mathbf{k}''}$$
$$-2N^{\frac{1}{2}},\quad \mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3=\mathbf{k}+\mathbf{k}'+\mathbf{k}''=0\,, \qquad (2\cdot3a)$$

$$E_{\mathbf{h}}\approx\frac{N^{\frac{1}{2}}\langle E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}\rangle_{\mathbf{k}}}{\langle|E_{\mathbf{k}}E_{\mathbf{h}-\mathbf{k}}|^2\rangle_{\mathbf{k}}}\,. \qquad (2\cdot4a)$$

Equation (2·3*a*) still has the disadvantage that it requires enormous computing facilities. In this paper we shall present a least-squares procedure which yields formulas that appear to supersede (2·3), (2·3*a*), (2·4), and (2·4*a*). This least-squares approach is particularly valuable since it is equally applicable to the case that the structure contains unequal atoms.

3. Analysis

3·1. *The preliminary formulas*

We start with the definition of the normalized structure factor $E_{\mathbf{h}}$* (Karle & Hauptman, 1956, equation (3·12)). For space group $P1$ with N identical point atoms per unit cell, this reduces to

$$E_{\mathbf{h}}=\frac{1}{N^{\frac{1}{2}}}\sum_{j=1}^{N}\exp(2\pi i\mathbf{h}.\mathbf{r}_j)\,, \qquad (3\cdot01)$$

where $\mathbf{r}_j$ is the position vector of the jth atom. Equation (3·01) is the case $q=1$ of the more general definition†

$$E_{\mathbf{h}_1\mathbf{h}_2\cdots\mathbf{h}_q}=\frac{1}{N^{q/2}}\sum_{j_1\neq j_2\neq1\cdots\neq j_q}^{N}\prod_{i=1}^{q}\exp(2\pi i\mathbf{h}_i.\mathbf{r}_j)\,. \qquad (3\cdot02)$$

Employing (3·01) to compute $E_{\mathbf{h}_1}E_{\mathbf{h}_2}$ and $E_{\mathbf{h}_1}E_{\mathbf{h}_2}E_{\mathbf{h}_3}$, we readily derive in turn the two preliminary formulas

$$E_{\mathbf{h}_1\mathbf{h}_2}=E_{\mathbf{h}_1}E_{\mathbf{h}_2}-\frac{1}{N^{\frac{1}{2}}}E_{\mathbf{h}_1+\mathbf{h}_2}\,, \qquad (3\cdot03)$$

$$E_{\mathbf{h}_1\mathbf{h}_2\mathbf{h}_3}=E_{\mathbf{h}_1}E_{\mathbf{h}_2}E_{\mathbf{h}_3}$$
$$-(1/N^{\frac{1}{2}})(E_{\mathbf{h}_1+\mathbf{h}_2}E_{\mathbf{h}_3}+E_{\mathbf{h}_1+\mathbf{h}_3}E_{\mathbf{h}_2}+E_{\mathbf{h}_2+\mathbf{h}_3}E_{\mathbf{h}_1})$$
$$+(2/N)E_{\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3}\,. \qquad (3\cdot04)$$

3·2. *The final formulas*

Equation (2·4) is an immediate consequence of (3·03). In (3·03) we make the substitutions

$$\mathbf{h}_1\to\mathbf{k},\quad \mathbf{h}_2\to\mathbf{h}-\mathbf{k} \qquad (3\cdot05)$$

and average both sides of (3·03) over $\mathbf{k}$. From (3·02),

$$\langle E_{\mathbf{k}(\mathbf{h}-\mathbf{k})}\rangle_{\mathbf{k}}=0\,, \qquad (3\cdot06)$$

and solving (3·03) for $E_{\mathbf{h}}$ yields (2·4).

In order to prove (2·1) we compute

$$\prod_{i=1}^{4}E_{\mathbf{h}_i}$$

from (3·01) as follows:

$$\prod_{i=1}^{4}E_{\mathbf{h}_i}=\frac{1}{N^2}\sum_{j,j',j'',j'''}^{N}$$
$$\times\exp[2\pi i(\mathbf{h}_1.\mathbf{r}_j+\mathbf{h}_2.\mathbf{r}_{j'}+\mathbf{h}_3.\mathbf{r}_{j''}+\mathbf{h}_4.\mathbf{r}_{j'''})] \qquad (3\cdot07)$$
$$=R_1+R_2+R_2'+R_3+R_4\,, \qquad (3\cdot08)$$

where

$$R_1=\frac{1}{N^2}\sum_{j=1}^{N}\exp[2\pi i(\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3+\mathbf{h}_4).\mathbf{r}_j]\,, \qquad (3\cdot09)$$

$$R_2=\frac{1}{N^2}\sum_{j\neq j'}^{N}\{\exp(2\pi i[(\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3).\mathbf{r}_j+\mathbf{h}_4.\mathbf{r}_{j'}])$$
$$+3\text{ similar terms}\}\,, \qquad (3\cdot10)$$

$$R_2'=\frac{1}{N^2}\sum_{j\neq j'}^{N}\{\exp(2\pi i[(\mathbf{h}_1+\mathbf{h}_2).\mathbf{r}_j+(\mathbf{h}_3+\mathbf{h}_4).\mathbf{r}_{j'}])$$
$$+2\text{ similar terms}\}\,, \qquad (3\cdot11)$$

$$R_3=\frac{1}{N^2}\sum_{j\neq j'\neq j''}^{N}\{\exp(2\pi i[(\mathbf{h}_1+\mathbf{h}_2).\mathbf{r}_j+\mathbf{h}_3.\mathbf{r}_{j'}+\mathbf{h}_4.\mathbf{r}_{j''}])$$
$$+5\text{ similar terms}\}\,, \qquad (3\cdot12)$$

$$R_4=\frac{1}{N^2}\sum_{j\neq j'\neq j''\neq j'''}^{N}\exp[2\pi i(\mathbf{h}_1.\mathbf{r}_j+\mathbf{h}_2.\mathbf{r}_{j'}$$
$$+\mathbf{h}_3.\mathbf{r}_{j''}+\mathbf{h}_4.\mathbf{r}_{j'''})]\,. \qquad (3\cdot13)$$

Next we specialize (3·08) by means of the substitutions

$$\mathbf{h}_1\to\mathbf{k},\quad \mathbf{h}_2\to-\mathbf{k},\quad \mathbf{h}_3\to\mathbf{h}-\mathbf{k},\quad \mathbf{h}_4\to-\mathbf{h}+\mathbf{k}\,, \qquad (3\cdot14)$$

and average both sides over all vectors $\mathbf{k}$. By imposing the condition that no four different vectors $\mathbf{r}_j$, $\mathbf{r}_{j'}$, $\mathbf{r}_{j''}$, $\mathbf{r}_{j'''}$ be rationally dependent*, i.e. that there do not exist four integers m_i, $i=1,2,3,4$, not all zero, such that

$$m_1\mathbf{r}_j+m_2\mathbf{r}_{j'}+m_3\mathbf{r}_{j''}+m_4\mathbf{r}_{j'''}=\mathbf{r},\quad j\neq j'\neq j''\neq j'''\,, \qquad (3\cdot15)$$

where the three components of $\mathbf{r}$ are integers, it is readily verified that

$$\langle R_2\rangle_{\mathbf{k}}=\langle R_3\rangle_{\mathbf{k}}=\langle R_4\rangle_{\mathbf{k}}=0\,. \qquad (3\cdot16)$$

Furthermore, it is easily seen that

---

* In practice $E_{\mathbf{h}}$ is obtained from $F_{\mathbf{h}}$ by means of

$$E_{\mathbf{h}}=F_{\mathbf{h}}\Big/\left(\sum_{j=1}^{N}f_{j\mathbf{h}}^2\right)^{\frac{1}{2}}\,,$$

where $F_{\mathbf{h}}$ is the crystal structure factor and $f_{j\mathbf{h}}$ is the atomic scattering factor.

† This definition and the following analysis are similar to those already described in I.

* This condition may be replaced by the less stringent requirement that there do not exist two integers m_1, m_2, not both zero, such that $m_1\mathbf{r}_j+m_2\mathbf{r}_{j'}-(m_1+m_2)\mathbf{r}_{j''}=\mathbf{r}$, $j\neq j'\neq j''$ where the three components of $\mathbf{r}$ are integers; and that no two interatomic vectors be rationally dependent.

$$\langle R_1\rangle_{\mathbf{k}} = 1/N \tag{3·17}$$

and

$$\langle R_2'\rangle_{\mathbf{k}} = \frac{N(N-1)}{N^2} + \frac{1}{N^2}\sum_{\substack{j,j'=1\\ j\neq j'}}^{N} \exp\left[2\pi i\mathbf{h}.(\mathbf{r}_j-\mathbf{r}_{j'})\right]. \tag{3·18}$$

Substituting from (3·16)–(3·18) into (3·07) and making use of (3·03) we find

$$\langle |E_{\mathbf{k}}|^2|E_{\mathbf{h}-\mathbf{k}}|^2\rangle_{\mathbf{k}} = 1+(1/N)(|E_{\mathbf{h}}|^2-1)\,. \tag{3·19}$$

Employing (2·5), we immediately deduce (2·1).

Evidently (2·1) is the special case $\mathbf{h}_1 = -\mathbf{h}_2$ of (2·1*a*). We omit the proof of the latter since it parallels closely that of (2·1).

The proof of (2·2) follows the same lines as that of (2·1). From (3·01)

$$\prod_{i=1}^{6} E_{\mathbf{h}_i} = \frac{1}{N^3}\sum_{j,j',j'',j''',j'''',j'''''=1}^{N} \exp\left[2\pi i(\mathbf{h}_1.\mathbf{r}_j+\mathbf{h}_2.\mathbf{r}_{j'}+\mathbf{h}_3.\mathbf{r}_{j''}+\mathbf{h}_4.\mathbf{r}_{j'''}+\mathbf{h}_5.\mathbf{r}_{j''''}+\mathbf{h}_6.\mathbf{r}_{j'''''})\right] \tag{3·20}$$

$$= R_1+R_2+R_2'+R_2''+R_3+R_3'+R_3''+R\,, \tag{3·21}$$

where

$$R_1 = \frac{1}{N^3}\sum_{j=1}^{N} \exp\left[2\pi i(\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3+\mathbf{h}_4+\mathbf{h}_5+\mathbf{h}_6).\mathbf{r}_j\right], \tag{3·22}$$

$$R_2 = \frac{1}{N^3}\sum_{\substack{j,j'=1\\ j\neq j'}}^{N} \{\exp\left(2\pi i[(\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3+\mathbf{h}_4+\mathbf{h}_5).\mathbf{r}_j+\mathbf{h}_6.\mathbf{r}_{j'}]\right) + 5 \text{ similar terms}\}\,, \tag{3·23}$$

$$R_2' = \frac{1}{N^3}\sum_{\substack{j,j'=1\\ j\neq j'}}^{N} \{\exp\left(2\pi i[(\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3+\mathbf{h}_4).\mathbf{r}_j + (\mathbf{h}_5+\mathbf{h}_6).\mathbf{r}_{j'}]\right)+14 \text{ similar terms}\}\,, \tag{3·24}$$

$$R_2'' = \frac{1}{N^3}\sum_{\substack{j,j'=1\\ j\neq j'}}^{N} \{\exp\left(2\pi i[(\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3).\mathbf{r}_j + (\mathbf{h}_4+\mathbf{h}_5+\mathbf{h}_6).\mathbf{r}_{j'}]\right)+9 \text{ similar terms}\}\,, \tag{3·25}$$

$$R_3 = \frac{1}{N^3}\sum_{\substack{j,j',j''=1\\ j\neq j'\neq j''}}^{N} \{\exp\left(2\pi i[(\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3+\mathbf{h}_4).\mathbf{r}_j + \mathbf{h}_5.\mathbf{r}_{j'}+\mathbf{h}_6.\mathbf{r}_{j''}]\right)+14 \text{ similar terms}\}\,, \tag{3·26}$$

$$R_3' = \frac{1}{N^3}\sum_{\substack{j,j',j''=1\\ j\neq j'\neq j''}}^{N} \{\exp\left(2\pi i[(\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3).\mathbf{r}_j+(\mathbf{h}_4+\mathbf{h}_5).\mathbf{r}_{j'} + \mathbf{h}_6.\mathbf{r}_{j''}]\right)+59 \text{ similar terms}\}\,, \tag{3·27}$$

$$R_3'' = \frac{1}{N^3}\sum_{\substack{j,j',j''=1\\ j\neq j'\neq j''}}^{N} \{\exp\left(2\pi i[(\mathbf{h}_1+\mathbf{h}_2).\mathbf{r}_j+(\mathbf{h}_3+\mathbf{h}_4).\mathbf{r}_{j'} + (\mathbf{h}_5+\mathbf{h}_6).\mathbf{r}_{j''}]\right)+14 \text{ similar terms}\}\,, \tag{3·28}$$

and R is a sum of various fourth-, fifth-, and sixth-order sums. Next, we introduce the substitutions

$$\mathbf{h}_1 \to \mathbf{k},\ \mathbf{h}_2 \to -\mathbf{k},\ \mathbf{h}_3 \to \mathbf{h}_1+\mathbf{k},\ \mathbf{h}_4 \to -\mathbf{h}_1-\mathbf{k},\ \mathbf{h}_5 \to -\mathbf{h}_3+\mathbf{k},\ \mathbf{h}_6 \to \mathbf{h}_3-\mathbf{k}\,, \tag{3·29}$$

where the new $\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3$ satisfy

$$\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3 = 0\,, \tag{3·30}$$

into (3·21) and average over $\mathbf{k}$. By imposing the condition that no six position vectors be rationally dependent* it is easily seen that

$$\langle R_2\rangle_{\mathbf{k}} = \langle R_2''\rangle_{\mathbf{k}} = \langle R_3\rangle_{\mathbf{k}} = \langle R_3'\rangle_{\mathbf{k}} = \langle R\rangle_{\mathbf{k}} = 0\,. \tag{3·31}$$

Furthermore, it is readily verified that

$$\langle R_1\rangle_{\mathbf{k}} = 1/N^2\,, \tag{3·32}$$

$$\langle R_2'\rangle_{\mathbf{k}} = \frac{1}{N^3}\sum_{\substack{j,j'=1\\ j\neq j'}}^{N} \left\{3+2\sum_{\nu=1}^{3}\exp\left[2\pi i\mathbf{h}_\nu(\mathbf{r}_j-\mathbf{r}_{j'})\right]\right\}\,, \tag{3·33}$$

$$\langle R_3''\rangle_{\mathbf{k}} = \frac{1}{N^3}\sum_{\substack{j,j',j''=1\\ j\neq j'\neq j''}}^{N} \left\{1+\sum_{\nu=1}^{3}\exp\left[2\pi i\mathbf{h}_\nu.(\mathbf{r}_j-\mathbf{r}_{j'})\right] + \exp\left[2\pi i(\mathbf{h}_1.\mathbf{r}_j+\mathbf{h}_2.\mathbf{r}_{j'}+\mathbf{h}_3.\mathbf{r}_{j''})\right] + \exp\left[-2\pi i(\mathbf{h}_1.\mathbf{r}_j+\mathbf{h}_2.\mathbf{r}_{j'}+\mathbf{h}_3.\mathbf{r}_{j''})\right]\right\}\,. \tag{3·34}$$

Next, substituting from (3·31)–(3·34) into (3·21) and employing (3·02)–(3·04) we obtain

$$\langle |E_{\mathbf{k}}E_{\mathbf{h}_1+\mathbf{k}}E_{\mathbf{h}_1+\mathbf{h}_2+\mathbf{k}}|^2\rangle_{\mathbf{k}} = 1-\frac{3}{N}+\frac{4}{N^2}+\frac{N-2}{N^2}(|E_{\mathbf{h}_1}|^2+|E_{\mathbf{h}_2}|^2+|E_{\mathbf{h}_3}|^2) + \frac{2}{N^{3/2}}|E_{\mathbf{h}_1}E_{\mathbf{h}_2}E_{\mathbf{h}_3}|\cos(\varphi_1+\varphi_2+\varphi_3)\,. \tag{3·35}$$

Evidently

$$\langle(|E_{\mathbf{k}}|^2-1)(|E_{\mathbf{h}_1+\mathbf{k}}|^2-1)(|E_{\mathbf{h}_1+\mathbf{h}_2+\mathbf{k}}|^2-1)\rangle_{\mathbf{k}} = \langle|E_{\mathbf{k}}E_{\mathbf{h}_1+\mathbf{k}}E_{\mathbf{h}_1+\mathbf{h}_2+\mathbf{k}}|^2\rangle_{\mathbf{k}} -\langle|E_{\mathbf{k}}E_{\mathbf{h}_1+\mathbf{k}}|^2\rangle_{\mathbf{k}}-\langle|E_{\mathbf{k}}E_{\mathbf{h}_1+\mathbf{h}_2+\mathbf{k}}|^2\rangle_{\mathbf{k}} -\langle|E_{\mathbf{h}_1+\mathbf{k}}E_{\mathbf{h}_1+\mathbf{h}_2+\mathbf{k}}|^2\rangle_{\mathbf{k}}+2\,, \tag{3·36}$$

where use has been made of (2·5). Substituting from (3·19) and (3·35) into (3·36) we find

$$\langle(|E_{\mathbf{k}}|^2-1)(|E_{\mathbf{h}_1+\mathbf{k}}|^2-1)(|E_{\mathbf{h}_1+\mathbf{h}_2+\mathbf{k}}|^2-1)\rangle_{\mathbf{k}} = 4/N^2-(2/N^2)(|E_{\mathbf{h}_1}|^2+|E_{\mathbf{h}_2}|^2+|E_{\mathbf{h}_3}|^2) +(2/N^{3/2})|E_{\mathbf{h}_1}E_{\mathbf{h}_2}E_{\mathbf{h}_3}|\cos(\varphi_1+\varphi_2+\varphi_3)\,, \tag{3·37}$$

from which (2·2) is immediately deduced.

By multiplying

$$|E_{\mathbf{h}_1}E_{\mathbf{h}_2}E_{\mathbf{h}_3}|\cos(\varphi_1+\varphi_2+\varphi_3) = \frac{1}{N^{3/2}}\sum_{j,j',j''=1}^{N}\cos 2\pi(\mathbf{h}_1.\mathbf{r}_j+\mathbf{h}_2.\mathbf{r}_{j'}+\mathbf{h}_3.\mathbf{r}_{j''}) \tag{3·38}$$

and

$$|E_{\mathbf{h}_4}E_{\mathbf{h}_5}E_{\mathbf{h}_6}|\cos(\varphi_4+\varphi_5+\varphi_6) = \frac{1}{N^{3/2}}\sum_{j,j',j''=1}^{N}\cos 2\pi(\mathbf{h}_4.\mathbf{r}_j+\mathbf{h}_5.\mathbf{r}_{j'}+\mathbf{h}_6.\mathbf{r}_{j''})\,, \tag{3·39}$$

employing the substitutions

* As before this requirement may be weakened somewhat.

$$\mathbf{h}_1 \to \mathbf{k},\ \mathbf{h}_2 \to \mathbf{k}',\ \mathbf{h}_3 \to \mathbf{k}'', \tag{3·40}$$

$$\mathbf{h}_4 \to \mathbf{h}_1-\mathbf{k},\ \mathbf{h}_5 \to \mathbf{h}_2-\mathbf{k}',\ \mathbf{h}_6 \to \mathbf{h}_3-\mathbf{k}'', \tag{3·41}$$

where the new $\mathbf{h}_1$, $\mathbf{h}_2$, $\mathbf{h}_3$ satisfy

$$\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3 = 0\,, \tag{3·42}$$

and averaging over all $\mathbf{k}$, $\mathbf{k}'$, $\mathbf{k}''$ subject to

$$\mathbf{k}+\mathbf{k}'+\mathbf{k}'' = 0\,, \tag{3·43}$$

we obtain (2·3).

4. Procedure for phase determination

4·1. *Initial determination of the phases*

In a previous paper (Hauptman & Karle, 1956) it was shown that the magnitudes of the structure factors determine the magnitudes of the structure invariants. Equation (2·2) shows how to find the magnitudes of all structure invariants of the type $\varphi_1+\varphi_2+\varphi_3$, where $\mathbf{h}_1+\mathbf{h}_2+\mathbf{h}_3 = 0$, in terms of the known magnitudes of the structure factors. There remains the problem of extracting the values of the individual phases from the known magnitudes of these structure invariants. To this end it is necessary to fix the origin and to select one of the two enantiomorphous structures permitted by the given set of structure factor magnitudes. A program for doing this has already been described in our previous paper (Hauptman & Karle, 1956). We here carry out the details of this program.

First, (2·2) is used to compute the magnitudes of all the structure invariants $\varphi_{\mathbf{h}_1}+\varphi_{\mathbf{h}_2}+\varphi_{-\mathbf{h}_1-\mathbf{h}_2}$. In practice, those invariants for which the corresponding products $|E_{\mathbf{h}_1}E_{\mathbf{h}_2}E_{-\mathbf{h}_1-\mathbf{h}_2}|$ are large will be, in general, the ones most accurately determined.

Next, a suitable primitive set $\mathbf{h}_\alpha$, $\mathbf{h}_\beta$, $\mathbf{h}_\gamma$ is chosen, i.e.

$$\begin{vmatrix} h_\alpha & k_\alpha & l_\alpha \\ h_\beta & k_\beta & l_\beta \\ h_\gamma & k_\gamma & l_\gamma \end{vmatrix} = \pm 1\,, \tag{4·01}$$

and each of $|E_\alpha|$, $|E_\beta|$, $|E_\gamma|$ must be quite large. We need to choose a primitive triple $\varphi_{\mathbf{h}_\alpha}$, $\varphi_{\mathbf{h}_\beta}$, $\varphi_{\mathbf{h}_\gamma}$ because we shall want to specify the values of these phases arbitrarily in order to fix the origin uniquely (Hauptman & Karle, 1956). In order to determine whether this triple is suitable we proceed as follows. We examine 19 related structure invariants arranged in ten sets of four (some invariants appearing several times):*

$$\left.\begin{array}{llll} \varphi_\alpha & +\varphi_\beta & +\varphi_{-\alpha-\beta} & \\ \varphi_{\bar\alpha} & +\varphi_\gamma & +\varphi_{\alpha-\gamma} & \\ \varphi_{\bar\beta} & +\varphi_{\bar\gamma} & +\varphi_{\beta+\gamma} & \\ \varphi_{\alpha+\beta} & +\varphi_{-\alpha+\gamma} & +\varphi_{-\beta-\gamma} & \end{array}\right\} \tag{4·02}$$

$$\left.\begin{array}{llll} \varphi_\alpha & +\varphi_\beta & +\varphi_{-\alpha-\beta} & \\ \varphi_{\bar\alpha} & +\varphi_{\bar\gamma} & +\varphi_{\alpha+\gamma} & \\ \varphi_{\bar\beta} & +\varphi_\gamma & +\varphi_{\beta-\gamma} & \\ \varphi_{\alpha+\beta} & +\varphi_{-\alpha-\gamma} & +\varphi_{-\beta+\gamma} & \end{array}\right\} \tag{4·03}$$

* A similar procedure involving identities has been suggested by Vaughan (1956).

$$\left.\begin{array}{llll} \varphi_\alpha & +\varphi_{\bar\beta} & +\varphi_{-\alpha+\beta} & \\ \varphi_{\bar\alpha} & +\varphi_\gamma & +\varphi_{\alpha-\gamma} & \\ \varphi_\beta & +\varphi_{\bar\gamma} & +\varphi_{-\beta+\gamma} & \\ \varphi_{\alpha-\beta} & +\varphi_{-\alpha+\gamma} & +\varphi_{\beta-\gamma} & \end{array}\right\} \tag{4·04}$$

$$\left.\begin{array}{llll} \varphi_\alpha & +\varphi_{\bar\beta} & +\varphi_{-\alpha+\beta} & \\ \varphi_{\bar\alpha} & +\varphi_{\bar\gamma} & +\varphi_{\alpha+\gamma} & \\ \varphi_\beta & +\varphi_\gamma & +\varphi_{-\beta-\gamma} & \\ \varphi_{\alpha-\beta} & +\varphi_{-\alpha-\gamma} & +\varphi_{\beta+\gamma} & \end{array}\right\} \tag{4·05}$$

$$\left.\begin{array}{llll} \varphi_\alpha & +\varphi_\alpha & +\varphi_{2\bar\alpha} & \\ \varphi_{\bar\alpha} & +\varphi_\beta & +\varphi_{\alpha-\beta} & \\ \varphi_{\bar\alpha} & +\varphi_{\bar\beta} & +\varphi_{\alpha+\beta} & \\ \varphi_{2\alpha} & +\varphi_{-\alpha+\beta} & +\varphi_{-\alpha-\beta} & \end{array}\right\} \tag{4·06}$$

$$\left.\begin{array}{llll} \varphi_\alpha & +\varphi_\alpha & +\varphi_{2\bar\alpha} & \\ \varphi_{\bar\alpha} & +\varphi_\gamma & +\varphi_{\alpha-\gamma} & \\ \varphi_{\bar\alpha} & +\varphi_{\bar\gamma} & +\varphi_{\alpha+\gamma} & \\ \varphi_{2\alpha} & +\varphi_{-\alpha+\gamma} & +\varphi_{-\alpha-\gamma} & \end{array}\right\} \tag{4·07}$$

$$\left.\begin{array}{llll} \varphi_\beta & +\varphi_\beta & +\varphi_{2\bar\beta} & \\ \varphi_{\bar\beta} & +\varphi_\alpha & +\varphi_{\beta-\alpha} & \\ \varphi_{\bar\beta} & +\varphi_{\bar\alpha} & +\varphi_{\beta+\alpha} & \\ \varphi_{2\beta} & +\varphi_{-\beta+\alpha} & +\varphi_{-\beta-\alpha} & \end{array}\right\} \tag{4·08}$$

$$\left.\begin{array}{llll} \varphi_\beta & +\varphi_\beta & +\varphi_{2\bar\beta} & \\ \varphi_{\bar\beta} & +\varphi_\gamma & +\varphi_{\beta-\gamma} & \\ \varphi_{\bar\beta} & +\varphi_{\bar\gamma} & +\varphi_{\beta+\gamma} & \\ \varphi_{2\beta} & +\varphi_{-\beta+\gamma} & +\varphi_{-\beta-\gamma} & \end{array}\right\} \tag{4·09}$$

$$\left.\begin{array}{llll} \varphi_\gamma & +\varphi_\gamma & +\varphi_{2\bar\gamma} & \\ \varphi_{\bar\gamma} & +\varphi_\alpha & +\varphi_{\gamma-\alpha} & \\ \varphi_{\bar\gamma} & +\varphi_{\bar\alpha} & +\varphi_{\gamma+\alpha} & \\ \varphi_{2\gamma} & +\varphi_{-\gamma+\alpha} & +\varphi_{-\gamma-\alpha} & \end{array}\right\} \tag{4·10}$$

$$\left.\begin{array}{llll} \varphi_\gamma & +\varphi_\gamma & +\varphi_{2\bar\gamma} & \\ \varphi_{\bar\gamma} & +\varphi_\beta & +\varphi_{\gamma-\beta} & \\ \varphi_{\bar\gamma} & +\varphi_{\bar\beta} & +\varphi_{\gamma+\beta} & \\ \varphi_{2\gamma} & +\varphi_{-\gamma+\beta} & +\varphi_{-\gamma-\beta} & \end{array}\right\} \tag{4·11}$$

where the abbreviations $\varphi_{2\alpha} = \varphi_{2\mathbf{h}_\alpha}$, $\varphi_{\alpha+\beta} = \varphi_{\mathbf{h}_\alpha+\mathbf{h}_\beta}$, etc. have been used. In order for the primitive triple $\mathbf{h}_\alpha$, $\mathbf{h}_\beta$, $\mathbf{h}_\gamma$ to be suitable, not only must $|E_\alpha|$, $|E_\beta|$, and $|E_\gamma|$ be moderately large, but so must many of the $|E|$'s corresponding to the φ's appearing in the sets (4·02)–(4·11). Furthermore, at least one of these ten sets of four must be suitable for distinguishing the two enantiomorphous structures permitted by the given set of structure factor magnitudes. Ideally, this means that in at least one set of four the magnitudes of three of the structure invariants should be approximately in the range 0·5–1·5 radians while the magnitude of the fourth should be approximately in the range 2·0–2·6 radians and also about equal to the sum of the other three. In order to specify one of the two possible enantiomorphs we specify arbitrarily the sign of the fourth and then choose the opposite sign for the remaining three. This is a consequence of the fact that the ten sets (4·02)–(4·11) have been so constructed that the sum of the four invariants in any set vanishes while any structure invariant has one sign for one enantiomorph and the opposite sign for the other.

The magnitude of any structure invariant is of course the same for both enantiomorphs (Hauptman & Karle, 1956). Less ideal sets are also permissible. For example the magnitudes of two of the structure invariants in a set of four may be about $\frac{1}{2}\pi$ radians while each of the two others may almost vanish (or both approximate π). In either case the signs of the first two will be opposite and either of the two combinations of signs may be chosen in order to distinguish the enantiomorphs. As another example, the magnitudes of two of the structure invariants may be about $\frac{1}{2}\pi$ radians, a third may almost vanish, and the fourth may be about π radians. In this case the first two invariants will have the same sign (either both plus or both minus) and either sign may be chosen in order to distinguish the enantiomorphs. It should be noted that it is not possible (or significant) to assign a sign to an invariant the magnitude of which is π (or 0) since such an invariant has the same value for both enantiomorphs. The sign of an invariant which is close to 0 or π will be obtained automatically as a consequence of the refinement process to be described.

Let us assume that the primitive triple φ_α, φ_β, φ_γ is suitable in the sense just described and that the first set (4·02) is suitable for distinguishing the enantiomorphs. Write

$$\left.\begin{aligned} &\varphi_\alpha+\varphi_\beta+\varphi_{-\alpha-\beta}\approx b_{10},\quad \varphi_{\bar{\alpha}}+\varphi_\gamma+\varphi_{\alpha-\gamma}\approx b_{20}\,,\\ &\varphi_{\bar{\beta}}+\varphi_{\bar{\gamma}}+\varphi_{\beta+\gamma}\approx b_{30},\quad \varphi_{\alpha+\beta}+\varphi_{-\alpha+\gamma}+\varphi_{-\beta-\gamma}\approx b_{40}\,, \end{aligned}\right\} \quad (4\cdot12)$$

where the $|b_{i0}|$, $i = 1, 2, 3, 4$ are initial values of the magnitudes of the structure invariants (4·02) obtained by using (2·2). Our first task is to determine the signs of the four numbers b_{i0}. Let us assume an ideal situation (although a similar argument applies if the situation is less than ideal), so that $|b_{i0}|$ is in the range 0·5–1·5 radians for $i = 1, 2, 3$ while $|b_{40}|$ is in the range 2·0–2·6 radians. Furthermore $|b_{40}| \approx |b_{10}|+|b_{20}|+|b_{30}|$. We arbitrarily attach a sign (either plus or minus) to b_{40}, thus selecting one of the two permissible enantiomorphs. Once this is done then the values, not merely the magnitudes, of all the remaining structure invariants are uniquely determined by the magnitudes of the structure factors. In order to determine the signs of b_{10}, b_{20}, b_{30} for example, we observe that the sum of the four invariants (4·12) must vanish, since $\varphi_{\mathbf{h}} = -\varphi_{\mathbf{h}}$. In short

$$b_{10}+b_{20}+b_{30}+b_{40}\approx 0\,, \quad (4\cdot13)$$

an equation which at this stage can be only approximately realized since the $|b_{i0}|$'s, as obtained from (2·2), only approximate to the true values of the corresponding invariants. Nevertheless (4·13) enables us to fix the signs of b_{10}, b_{20}, b_{30} once the sign of b_{40} has been specified. Assuming for example that the set (4·12) is ideal and that b_{40} has been chosen to be positive, then b_{10}, b_{20}, and b_{30} will all be negative.

In practice, since (4·13) cannot be expected to be exactly fulfilled, an initial adjustment of the values of the b_{i0}, $i = 1, 2, 3, 4$, in order to satisfy (4·13) exactly, is called for. We have

$$\sum_{i=1}^{4} b_{i0} = \varepsilon\,, \quad (4\cdot14)$$

where ε is close to zero. We wish to add small increments ε_{i0} to the b_{i0}, such that

$$\sum_{i=1}^{4} (b_{i0}+\varepsilon_{i0}) = 0\,, \quad (4\cdot15)$$

in order to obtain improved values $b'_{i0} = b_{i0}+\varepsilon_{i0}$ to replace the initially computed values b_{i0}. Hence, subject to the equation of restraint (4·13), we seek to minimize

$$\Phi = \sum_{i=1}^{4} w_{i0}\varepsilon_{i0}^2\,, \quad (4\cdot16)$$

where w_{i0} is a suitable weighting function. Taking w_{i0} to be a reasonable function which takes into account the number of contributors to the determination of b_{i0}, the size of the triple product of E's associated with b_{i0}, and the location of the value b_{i0} in the angle range, we find

$$w_{i0} = n_i^{1/2}|V_i|\,.\,|\sin b_{i0}| \quad (4\cdot17)$$

where

$$\left.\begin{aligned} &|V_1| = |E_\alpha E_\beta E_{-\alpha-\beta}|,\quad |V_2| = |E_{\bar{\alpha}} E_\gamma E_{\alpha-\gamma}|\,,\\ &|V_3| = |E_{\bar{\beta}} E_{\bar{\gamma}} E_{\beta+\gamma}|,\quad |V_4| = |E_{\alpha+\beta} E_{-\alpha+\gamma} E_{-\beta-\gamma}|\,, \end{aligned}\right\} \quad (4\cdot18)$$

and n_i is the number of terms contributing to the corresponding average in (2·2). The problem of minimizing (4·16) subject to (4·15) is readily solved in the usual way and leads to

$$\varepsilon_{i0} = -\varepsilon/w_{i0}\sum_{j=1}^{4}\frac{1}{w_{j0}}\,, \quad (4\cdot19)$$

where ε is defined by (4·14). Hence, revised values b'_{i0} of b_{i0} are given by

$$b'_{i0} = b_{i0}-\varepsilon/w_{i0}\sum_{j=1}^{4}\frac{1}{w_{j0}}\,,\quad i = 1, 2, 3, 4\,. \quad (4\cdot20)$$

Since the values, as well as the magnitudes, of four invariants (or perhaps only three or two if the set (4·12) is somewhat less than ideal) are now available, we now proceed to the remaining nine sets (4·03)–(4·11) in whatever order proves to be expedient, making use of previously determined invariants to determine the signs of the remaining ones. We thus determine the signs of those of the 19 distinct invariants appearing in (4·02)–(4·11) which differ from 0 and π. In addition to determining the signs of these invariants, improved values b'_{i0} are also obtained by means of the least squares adjustment just described. Since many of the invariants appear more than once in (4·02)–(4·11) we obtain several determinations for these invariants.

Thus far we have been considering a specific group of ten sets of invariants generated by the primitive

origin determining triple. Similarly, there is associated with any linearly independent triple of phases (primitive or not) a corresponding group of ten sets. However, it is useful to restrict attention to linearly independent triples chosen from among the phases occurring in the initial set (4·02)–(4·11). As a result, many of the initial invariants recur in the newly formed sets thus facilitating the determination of the signs of new invariants. Furthermore, since the invariants will reoccur so frequently, we are enabled to make a more accurate determination of their values.

By taking a weighted average of these several determinations for a given invariant we obtain new values b_{i1} instead of b'_{i0} for the corresponding invariants. At this stage any reasonable weight w'_{i0} can be attached to the determination b'_{i0} of the corresponding invariant. For example, we may take

$$\left.\begin{aligned} w'_{10} &= (2w_{10}^2+w_{20}^2+w_{30}^2+w_{40}^2)^{\frac{1}{2}} \\ w'_{20} &= (w_{10}^2+2w_{20}^2+w_{30}^2+w_{40}^2)^{\frac{1}{2}} \end{aligned}\right\} \quad (4\cdot21)$$

etc. In this way we obtain

$$b_{i1} = \Sigma w'_{i0}b'_{i0}/\Sigma w'_{i0}\,, \quad (4\cdot22)$$

where the average (4·22) is taken over all determinations b'_{i0} of a particular invariant. Furthermore, we can now attach a reasonable weight w_{i1} to the determination b_{i1} by means of

$$w_{i1} = (\Sigma w_{i0}'^2)^{\frac{1}{2}}\,, \quad (4\cdot23)$$

where the sum in (4·23) is taken over all those w'_{i0} which appear in (4·22). In short, starting with crude values b_{i0} and corresponding weights w_{i0} we have obtained, by means of a least-squares refinement and subsequent averaging, improved values b_{i1} and corresponding weights w_{i1} for these invariants. This cycle of least-squares refinement and subsequent averaging may now be applied to the new values b_{i1} with corresponding weights w_{i1} to obtain a third set of values b_{i2} with corresponding weights w_{i2} for these invariants.

Finally, among the nineteen distinct invariants (two invariants in which the indices of one are the negatives of the corresponding indices of the other are not counted as distinct) appearing in (4·02)–(4·11) only nine involve two of the primitive triple φ_α, φ_β, φ_γ (or their negatives). Specifying arbitrarily the values of the phases φ_α, φ_β, φ_γ (thus uniquely fixing the origin), these nine invariants lead to values of the nine additional phases $\varphi_{\alpha\pm\beta}$, $\varphi_{\alpha\pm\gamma}$, $\varphi_{\beta\pm\gamma}$, $\varphi_{2\alpha}$, $\varphi_{2\beta}$, $\varphi_{2\gamma}$.

From this set of twelve phases any set of three linearly independent ones may be chosen to replace the set φ_α, φ_β, φ_γ in the preceding discussion. As a consequence of having carried out the above procedure, we have already obtained the new group of ten sets of invariants corresponding to this triple of phases. Thus we have the values of nineteen invariants, some of which coincide with those in the first set of nineteen and the remainder of which are new. By repeating the procedure described when starting with the triple φ_α, φ_β, φ_γ we obtain the values of nine more phases, some of which will coincide with ones previously determined. This process may be repeated indefinitely to yield the values of as many phases as desired. The fact that any phase is ultimately accessible in this way is a consequence of the choice of the triple φ_α, φ_β, φ_γ as primitive. In fact, any vector $\mathbf{h}$ is a linear combination, with *integer* coefficients, of the primitive set $\mathbf{h}_\alpha$, $\mathbf{h}_\beta$, $\mathbf{h}_\gamma$ (an immediate consequence of (4·01)). For example, among the second group of ten sets of four structure invariants, the set

$$\left.\begin{aligned} &\varphi_{\alpha+\beta}+\varphi_{-\alpha+\gamma}+\varphi_{-\beta-\gamma},\ \varphi_{-\alpha-\beta}+\varphi_{\bar{\beta}}+\varphi_{\alpha+2\beta}\,, \\ &\varphi_{\alpha-\gamma}+\varphi_{\beta}+\varphi_{-\alpha-\beta+\gamma},\ \varphi_{\beta+\gamma}+\varphi_{-\alpha-2\beta}+\varphi_{\alpha+\beta-\gamma} \end{aligned}\right\} \quad (4\cdot24)$$

may occur. The value of the first invariant of (4·24) has been previously determined while only the magnitudes of the remaining three are known from (2·2). However, as already described, the values of the latter three have also been determined, thus leading to the values of the new phases $\varphi_{\alpha+2\beta}$ and $\varphi_{\alpha+\beta-\gamma}$. It is easily seen that by continuing in this way initial values of all the phases may be determined. It is to be noted that at any stage sets of three suitable phases, leading to nineteen corresponding structure invariants, may be chosen in many ways. Hence the value of a particular phase may be obtained several times, thus affording a check on the self-consistency of the procedure. In fact the value of a particular phase will be obtained by averaging all the different values obtained for that phase.

4·2. *Second least-squares refinement*

We assume that initial values of a large number of the phases have been determined and that refined values b of the various structure invariants are known (as described in § 4·1). Our purpose in this section is to find improved values for the phases by means of a least-squares adjustment. It is important to note that the least-squares refinement to be described is valid also in the case that the structure contains unequal atoms. We introduce the notation

$$E = E_{\mathbf{h}},\ E_i = E_{\mathbf{h}_i},\ E_j = E_{\mathbf{h}_j}\,, \quad (4\cdot25)$$

$$\varphi_i = \varphi_{\mathbf{h}_i},\ \varphi_j = \varphi_{\mathbf{h}_j}\,, \quad (4\cdot26)$$

$$E_{ij} = E_iE_j\,, \quad (4\cdot27)$$

$$\varphi_{ij} = \varphi_i+\varphi_j\,, \quad (4\cdot28)$$

and impose the condition

$$\mathbf{h}+\mathbf{h}_i+\mathbf{h}_j = 0\,. \quad (4\cdot29)$$

We seek an improved value for the phase φ of $E_{\mathbf{h}}$, making use of initial values φ_i, φ_j of those phases whose indices satisfy (4·29). We start with

$$|E_{ij}|\cos(\varphi+\varphi_i+\varphi_j) = c_{ij}\,, \quad (4\cdot30)$$

$$|E_{ij}|\sin(\varphi+\varphi_i+\varphi_j) = s_{ij}\,, \quad (4\cdot31)$$

Table 1. *The* 19 *structure invariants associated with the primitive triple* (5·1), *their values in radians as computed from* (2·2) *and the initial least-squares refinements*

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
In-variant	$\mathbf{h}_{i1}$	$\mathbf{h}_{i2}$	$\mathbf{h}_{i3}$	$\lvert E_{i1}E_{i2}E_{i3}\rvert$	cos b_{i0}	b_{i0}	Stage (or set)	w_{i0}	ε_{i0}	b'_{i0}	w'_{i0}	$b_{i1} = \langle b'_{i0} \rangle$	w_{i1}	b_{i2}	b_i
1	$4\bar{1}0$	$2\bar{2}\bar{5}$	$\bar{6}35$	3·33	+0·9159	−0·413		88	+0·094	−0·319	294	−0·361	614	−0·330	−0·124
2	$\bar{4}10$	$5\bar{1}1$	$\bar{1}0\bar{1}$	2·51	+0·4502	−1·104*	2	179	+0·046	−1·058	332	−1·115	789	−1·116	−1·098
3	$\bar{2}25$	$\bar{5}1\bar{1}$	$7\bar{3}\bar{4}$	4·33	+0·7344	−0·746		187	+0·044	−0·702	337	−0·726	717	−0·721	−0·685
4	$6\bar{3}5$	101	$\bar{7}34$	1·04	−0·3654	+1·945*		61	+0·134	+2·079	287	+2·079	287	+2·134	+1·907
1	$4\bar{1}0$	$2\bar{2}\bar{5}$	$\bar{6}35$	3·33	+0·9159	−0·413		88	+0·102	−0·311	330	−0·361	614	−0·330	−0·124
5	$\bar{4}10$	$5\bar{1}\bar{1}$	$9\bar{2}1$	4·90	+0·7939	−0·654*	3	201	+0·045	−0·609	376	−0·648	821	−0·644	−0·636
6	$\bar{2}25$	$5\bar{1}1$	$3\bar{1}\bar{6}$	3·31	+0·2538	−1·314		216	+0·042	−1·272	384	−1·312	766	−1·309	−1·237
7	$6\bar{3}\bar{5}$	$\bar{9}2\bar{1}$	316	1·55	−0·4839	+2·076*		77	+0·116	+2·192	327	+2·192	327	+2·247	+1·997
8	$4\bar{1}0$	$\bar{2}25$	$2\bar{1}\bar{5}$	8·81	+1·0159	0·000		0	−0·072	−0·072	282	−0·069*	567	−0·072	−0·092
2	$\bar{4}10$	$5\bar{1}1$	$\bar{1}0\bar{1}$	2·51	+0·4502	−1·104	4	179	0·000	−1·104	334	−1·115	789	−1·116	−1·098
$\bar{6}$	$2\bar{2}\bar{5}$	$5\bar{1}\bar{1}$	$3\bar{1}\bar{6}$	3·31	+0·2538	+1·314		216	0·000	+1·314	355	+1·312	766	+1·309	+1·237
9	215	101	$\bar{3}\bar{1}\bar{6}$	2·10	+0·9905	−0·138*		21	0·000	−0·138	282	−0·138	282	−0·134	−0·047
8	$4\bar{1}0$	$\bar{2}25$	$2\bar{1}\bar{5}$	8·81	+1·0159	0·000		0	−0·031	−0·031	275	−0·069	567	−0·072	−0·092
5	$\bar{4}10$	$5\bar{1}\bar{1}$	$9\bar{2}1$	4·90	+0·7939	−0·654	5	201	0·000	−0·654	340	−0·648	821	−0·644	−0·636
$\bar{3}$	$2\bar{2}\bar{5}$	$5\bar{1}1$	$\bar{7}34$	4·33	+0·7344	+0·746		187	0·000	+0·746	332	+0·726	717	+0·721	+0·685
10	215	$\bar{9}2\bar{1}$	$7\bar{3}\bar{4}$	5·36	+1·0261	0·000		0	−0·061	−0·061	275	−0·061*	275	−0·037	+0·043
11	$4\bar{1}0$	$4\bar{1}0$	$\bar{8}20$	5·86	+0·5188	+1·025		361	0·000	+1·025	527	+1·014	788	+0·997	+0·932
$\bar{8}$	$\bar{4}10$	$2\bar{2}\bar{5}$	215	8·81	+1·0159	0·000	7	0	−0·081	−0·081	383	+0·069	567	+0·072	+0·092
$\bar{1}$	$\bar{4}10$	$\bar{2}25$	$6\bar{3}\bar{5}$	3·33	+0·9159	+0·413*		88	0·000	+0·413	393	+0·361	614	+0·330	+0·124
12	$8\bar{2}0$	$2\bar{1}\bar{5}$	$\bar{6}35$	1·60	+0·2125	−1·357		96	0·000	−1·357	394	−1·357	394	−1·388	−1·148
11	$4\bar{1}0$	$4\bar{1}0$	$\bar{8}20$	5·86	+0·5188	+1·025*		361	−0·021	+1·004	586	+1·014	788	+0·997	+0·932
2	$\bar{4}10$	$5\bar{1}1$	$\bar{1}0\bar{1}$	2·51	+0·4502	−1·104	6	179	−0·042	−1·146	495	−1·115	789	−1·116	−1·098
5	$\bar{4}10$	$5\bar{1}\bar{1}$	$9\bar{2}1$	4·90	+0·7939	−0·654		201	−0·037	−0·691	503	−0·648	821	−0·644	−0·636
13	$8\bar{2}0$	101	$\bar{9}2\bar{1}$	1·90	+0·6158	+0·907*		101	−0·074	+0·833	472	+0·833	472	+0·803	+0·802
14	$2\bar{2}\bar{5}$	$2\bar{2}\bar{5}$	$\bar{4},4,10$	3·46	+0·8439	−0·566		105	0·000	−0·566	176	−0·566	369	−0·610	−0·470
8	$\bar{2}25$	$4\bar{1}0$	$2\bar{1}\bar{5}$	8·81	+1·0159	0·000	8	0	−0·544	−0·544	141	−0·069	567	−0·072	−0·092
$\bar{1}$	$\bar{2}25$	$\bar{4}10$	$6\bar{3}\bar{5}$	3·33	+0·9159	+0·413		88	0·000	+0·413	166	+0·361	614	+0·330	+0·124
15	$4,\bar{4},\bar{1}0$	215	$\bar{6}35$	1·03	+0·7670	+0·697*		35	0·000	+0·697	145	+0·697	145	+0·473	+0·438
14	$2\bar{2}\bar{5}$	$2\bar{2}\bar{5}$	$\bar{4},4,10$	3·46	+0·8439	−0·566*		105	−0·000	−0·566	324	−0·566	369	−0·610	−0·470
6	$\bar{2}25$	$5\bar{1}1$	$3\bar{1}\bar{6}$	3·31	+0·2538	−1·314*	1	216	−0·000	−1·314	375	−1·312	766	−1·309	−1·237
3	$\bar{2}25$	$\bar{5}1\bar{1}$	$7\bar{3}\bar{4}$	4·33	+0·7344	−0·746*		187	−0·000	−0·746	359	−0·726	717	−0·721	−0·685
16	$4,\bar{4},\bar{1}0$	316	$\bar{7}34$	1·55	−0·8710	+2·628*		38	−0·001	+2·627	309	+2·627	309	+2·618	+2·392
17	$5\bar{1}1$	$5\bar{1}1$	$\bar{1}0,2,\bar{2}$	3·76	+0·5213	−1·022*		202	+0·027	−0·995	405	−0·992	576	−0·982	−0·884
$\bar{2}$	$\bar{5}1\bar{1}$	$4\bar{1}0$	101	2·51	+0·4502	+1·104	9	179	+0·030	+1·134	394	+1·115	789	+1·116	+1·098
5	$\bar{5}1\bar{1}$	$\bar{4}10$	$9\bar{2}1$	4·90	+0·7939	−0·654		201	+0·027	−0·627	405	−0·648	821	−0·644	−0·636
18	$10,\bar{2},2$	$\bar{1}0\bar{1}$	$\bar{9}2\bar{1}$	3·75	+0·9067	+0·435*		100	+0·055	+0·490	365	+0·490	365	+0·504	+0·422
17	$5\bar{1}1$	$5\bar{1}1$	$\bar{1}0,2,\bar{2}$	3·76	+0·5213	−1·022		202	+0·033	−0·989	409	−0·992	576	−0·982	−0·884
$\bar{6}$	$\bar{5}1\bar{1}$	$2\bar{2}\bar{5}$	316	3·31	+0·2538	+1·314	10	216	+0·031	+1·345	416	+1·312	766	+1·309	+1·237
3	$\bar{5}1\bar{1}$	$\bar{2}25$	$7\bar{3}\bar{4}$	4·33	+0·7344	−0·746		187	+0·036	−0·710	402	−0·726	717	−0·721	−0·685
19	$10,\bar{2},2$	$\bar{3}\bar{1}\bar{6}$	$\bar{7}34$	4·76	+0·9685	+0·252*		65	+0·103	+0·355	362	+0·355	362	+0·375	+0·332

* Indicates stage at which the sign of the corresponding invariant was determined.

where

$$c_{ij} = |E_{ij}| \cos b \,, \tag{4·32}$$

$$s_{ij} = |E_{ij}| \sin b \,, \tag{4·33}$$

and the b are the values of the respective structure invariants $\varphi+\varphi_i+\varphi_j$ determined in § 4·1. Setting

$$c = \cos\varphi \,, \tag{4·34}$$

$$s = \sin\varphi \,, \tag{4·35}$$

we therefore need to minimize

$$\Phi = \sum_{i,j} \{[|E_{ij}|(c\cos\varphi_{ij} - s\sin\varphi_{ij}) - c_{ij}]^2 + [|E_{ij}|(c\sin\varphi_{ij} + s\cos\varphi_{ij}) - s_{ij}]^2\} \,, \tag{4·36}$$

subject to the condition

$$c^2+s^2-1 = 0 \,. \tag{4·37}$$

Using Lagrange's method of undetermined multipliers we are led to the solution

$$c = \cos\varphi = C/(C^2+S^2)^{\frac{1}{2}} \,, \tag{4·38}$$

$$s = \sin\varphi = S/(C^2+S^2)^{\frac{1}{2}} \,, \tag{4·39}$$

where

$$C = \sum_{i,j} |E_{ij}|(c_{ij}\cos\varphi_{ij} + s_{ij}\sin\varphi_{ij}) \,, \tag{4·40}$$

$$S = \sum_{i,j} |E_{ij}|(-c_{ij}\sin\varphi_{ij} + s_{ij}\cos\varphi_{ij}) \,, \tag{4·41}$$

and the sums in (4·40) and (4·41) are taken over all i, j satisfying (4·29). In this way refined values for the phases may be obtained, and the process may be reiterated if necessary.

5. Illustrative example

In order to illustrate the methods described in § 4·1, a structure consisting of ten identical point atoms was artificially constructed. The magnitudes of 4630 E's and their phases were then computed. This would correspond roughly to five or six times the number of intensities contained within the copper sphere. The set $\mathbf{h}_\alpha$, $\mathbf{h}_\beta$, $\mathbf{h}_\gamma$ was chosen from these 4630 $E_{\mathbf{h}}$'s to be

$$\mathbf{h}_\alpha = 4, \bar{1}, 0, \quad \mathbf{h}_\beta = 2, \bar{2}, \bar{5}, \quad \mathbf{h}_\gamma = 5, \bar{1}, 1 \,. \tag{5·1}$$

It is readily verified that the triple $\varphi_\alpha, \varphi_\beta, \varphi_\gamma$ is primitive. The values of the nine associated triples $\mathbf{h}_\alpha \pm \mathbf{h}_\beta$, $\mathbf{h}_\alpha \pm \mathbf{h}_\gamma$, $\mathbf{h}_\beta \pm \mathbf{h}_\gamma$, $2\mathbf{h}_\alpha$, $2\mathbf{h}_\beta$, $2\mathbf{h}_\gamma$ are readily found. The corresponding $|E|$ and φ values are listed in columns 2 and 3 of Table 2. From these values of the phases the true values of the nineteen structure invariants associated with the given primitive triple $\varphi_\alpha, \varphi_\beta, \varphi_\gamma$ were computed. These are shown in column 16 of Table 1. The invariants are labeled by column 1, where e.g. invariant $\bar{1}$ is obtained from invariant 1 by reversing the signs of all the indices. Two such invariants are the negatives of each other and therefore not essentially distinct. Since we have listed all ten sets of four invariants each, many of the invariants occur several times in Table 1.

Columns 2, 3, 4 of Table 1 list the indices of the sets of invariants (4·02) to (4·11). Column 6 lists the values of $\cos b_{i0}$ obtained from (2·2). The magnitudes of the corresponding b_{i0} were then obtained and listed in column 7. There remained the problem of specifying their signs. We note that the cosines of invariants 8 and 10 slightly exceed unity, and the values of these invariants were set equal to zero. Each of the ten sets of four invariants is labeled by means of column 8 which indicates also the order in which each set was examined. The set labeled 1 in column 8 was considered to be ideal as seen from the magnitudes of these invariants and

$$2·628 \approx 0·566+1·314+0·746 \,.$$

Also the values of the corresponding products $|E_{i1}E_{i2}E_{i3}|$ for this set were fairly large. The sign of invariant 16 was chosen to be plus in agreement with the known sign (+2·392) for this invariant obtained from column 16. In this way we select that one of the two enantiomorphs permitted by the structure factor magnitudes which coincides with the given structure, in order that comparison between computed and true phases may later be made. We conclude that the signs of the remaining three invariants 14, 6, and 3 in set 1 must be minus. No other combination of signs is even approximately consistent with the requirement (4·13).

Table 2. *The values of the phases in radians computed from the invariants of Table* 1, *showing successive refinements*

1	2	3	4	5	6	7	8	9
h	$\|E\|$	φ	φ_0	φ_1	φ_2	$\|\varphi-\varphi_0\|$	$\|\varphi-\varphi_1\|$	$\|\varphi-\varphi_2\|$
$4\bar{1}0$	2·21	+0·123	+0·123	+0·123	+0·123	0·000	0·000	0·000
$2\bar{2}\bar{5}$	2·12	−0·248	−0·248	−0·248	−0·248	0·000	0·000	0·000
$5\bar{1}1$	1·26	+1·387	+1·387	+1·387	+1·387	0·000	0·000	0·000
$\bar{6}35$	0·71	−0·001	−0·298	−0·236	−0·205	0·297	0·235	0·204
101	0·90	+2·362	+2·368	+2·379	+2·380	0·006	0·017	0·018
$7\bar{3}4$	1·62	+0·454	+0·393	+0·414	+0·418	0·061	0·040	0·036
$9\bar{2}1$	1·76	+0·874	+0·856	+0·862	+0·866	0·018	0·012	0·008
316	1·24	+2·872	+2·949	+2·947	+2·944	0·077	0·075	0·072
215	1·88	+0·463	+0·371	+0·440	+0·443	0·092	0·023	0·020
$8\bar{2}0$	1·20	−0·686	−0·779	−0·768	−0·751	0·093	0·082	0·065
$4,\bar{4},\bar{1}0$	0·77	−0·026	+0·070	+0·070	+0·114	0·096	0·096	0·140
$10,\bar{2},2$	2·37	−2·625	−2·487	−2·517	−2·527	0·138	0·108	0·098

(We may note at this stage that this assignment of signs is also consistent with the known signs as shown in column 16. It is to be emphasized however that no use of this information was made in determining these signs.)

Next, we proceed to the set labeled 2 in column 8. Since invariant 3 is known to be minus (from set 1) we conclude that invariants 2 and 4 are minus and plus respectively (in order to fulfil (4·13)). At this stage it also appears likely that invariant 1 is negative but we delay this decision since the evidence is not yet conclusive.

Proceeding to the third set we make use of the known sign of invariant 6 to infer that invariants 5 and 7 are minus and plus respectively. Again invariant 1 appears to be negative, but the evidence is still inconclusive.

In the fourth stage the known signs for invariants 2 and $\bar{6}$ yield the negative sign for invariant 9. Due to the small magnitude of this invariant this assignment might have turned out to be wrong, an error which the subsequent least-squares adjustment would presumably have corrected. In a more complete analysis, involving the examination perhaps of hundreds of sets, this invariant would occur many times and its sign could be specified with greater certainty. Here the invariant occurs only once and the decision concerning its sign had to be made at this stage although the evidence is hardly conclusive.

Proceeding to the sixth set we employ the known signs of invariants 2 and 5 to deduce the signs of invariants 11 and 13. Again, using the known invariants 11 and 12, we finally determine unambiguously the sign of invariant $\bar{1}$ to be positive in set 7. This is in agreement with the slight indications in sets 2 and 3 that invariant 1 was negative. The signs of the remaining invariants 15, 17, 18, and 19 in sets 8, 9, and 10 are now readily determined in the same way. We note that in set 8 the criterion for minimum ε is not fulfilled owing to the fact that the signs of invariants 14 and $\bar{1}$ had already been determined. The combination of signs obtained for this set yielded $\varepsilon = +0{\cdot}544$ whereas the minimum that could be obtained is $\varepsilon = -0{\cdot}282$. We observe that set 5 was not useful in determining the signs of any invariants, but served merely to confirm the previously obtained signs for invariants 5 and 3.

Next, (4·17) was used to compute the weights w_{i0} listed in column 9. By means of (4·19) and (4·20) the ε_{i0} and b'_{i0} listed in columns 10 and 11 were computed. Finally (4·21) was used to compute the weights w'_{i0} (column 12) which are needed in (4·22) to compute the b_{i1} listed in column 13. The weights w_{i1} associated with the improved values b_{i1} of the invariants were obtained from (4·23). Finally the results of a second cycle of refinement are shown in column 15 which may be compared with the true values as shown in column 16.

From columns 7, 13, and 15 of Table 1, the values of the b_{i0}, b_{i1} and b_{i2} respectively, we readily computed the values of the phases φ_0, φ_1, and φ_2 corresponding to the successive least-squares refinement as described in § 4. These are listed in columns 4, 5, and 6 of Table 2, in which it is to be noted that the first three phases are arbitrarily assigned their true values as computed from the structure. This assignment is permissible since we thereby uniquely specify the origin to agree with that selected in the original structure, and is made in order to facilitate comparison between the computed and true values of the remaining nine phases listed in Table 2. The successive improvements in the computed values of the phases resulting from the least-squares refinement is apparent from inspection of the deviations listed in columns 7, 8, and 9 of Table 2. The weighted average deviations (excluding the initial primitive triple) are 0·091, 0·067, and 0·063 radians respectively.

6. Concluding remarks

It is in order to compare the ideal example described in § 5 with the experimental situation. Clearly the accuracy of the calculated phases in the illustrative example is improved as a consequence of using exact data, whereas in practice a certain amount of error in the $|E|$ values is unavoidable. The experimental situation is, however, improved as a consequence of the initial least-squares adjustment of the b_{i0}'s, especially since in practice many sets of invariants would be employed rather than the ten sets used in the illustrative example. In practice many hundreds of invariants would be available for examination, and it is therefore to be expected that a more suitable primitive triple for fixing the origin would be available than the one actually chosen in our example. Finally, owing to the fact that the precision with which (2·2) can be computed is approximately proportional to the square-root of the number of data, it is possible to estimate roughly the effect of using fewer data.

Our formulas have exact validity provided that the structure consist of N identical point atoms and that no six position vectors be rationally dependent. Even if these requirements are only approximately fulfilled it is clear that these formulas have approximate validity and are useful for determining phases. Since both least-square procedures here described are valid whether or not these conditions are fulfilled, the phases finally determined are not likely to be much affected if the above assumed conditions do not hold rigorously.

Our formulas imply that a structure in $P1$ consisting of N identical point atoms is uniquely determined (except for its enantiomorph) by the magnitudes of its structure factors, provided that no six position vectors are rationally dependent. In short we have proven the following:

THEOREM. A $P1$ structure consisting of N identical point atoms (where $N \geq 6$) and such that no six

position vectors are rationally dependent, has precisely one homometric mate, its enantiomorph.

We wish to express our appreciation to Mr Peter O'Hara of the National Bureau of Standards for performing the calculations contained in this paper. He has programmed the calculation of equation (2·2) for the digital computer, SEAC, thus making a high-speed computer available for the computation of phases.

References

HAUPTMAN, H. & KARLE, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal.* A.C.A. Monograph No. 3. Wilmington: The Letter Shop.
HAUPTMAN, H. & KARLE, J. (1955). *Acta Cryst.* **8**, 355.
HAUPTMAN, H. & KARLE, J. (1956). *Acta Cryst.* **9**, 45.
HAUPTMAN, H. & KARLE, J. (1957). *Acta Cryst.* **10**, 267.
HUGHES, E. W. (1953). *Acta Cryst.* **6**, 871.
KARLE, J. & HAUPTMAN, H. (1956). *Acta Cryst.* **9**, 635.
VAUGHAN, P. A. (1956). American Crystallographic Association Annual Meeting, French Lick, Ind.

Acta Cryst. (1957). **10**, 524

A Neutron-Diffraction Study of Potassium Hydrogen Bis-Phenylacetate

BY G. E. BACON AND N. A. CURRY

Atomic Energy Research Establishment, Harwell, Berks, England

(*Received* 7 *February* 1957)

Fourier projections derived from single-crystal intensity data give the positions and thermal vibrations of all the atoms in the structure. The mean length of the C–H bonds to the benzene ring is 1·13 Å. The acidic hydrogen atom at the centre of symmetry shows no measurable anisotropy but it is very diffuse, with a r.m.s. amplitude of displacement of 0·29 Å.

Introduction

In two earlier papers (Bacon & Curry, 1956*a*, *b*) we have shown how neutron-diffraction data can supplement existing knowledge of the molecular structure derived using X-rays in the cases of sodium sesquicarbonate and α-resorcinol. In each instance the supplementary information depended on the detection of hydrogen atoms. In the former substance we were able to provide detailed information about the hydrogen bond across a symmetry centre, which links two carbonate groups, and also about the hydrogen atoms within the molecules of water of crystallization; in resorcinol we were able to show detail both for the intermolecular hydrogen bonds and for the C–H bonds to the benzene ring. The present study is a natural sequel to these two earlier ones. Potassium hydrogen bis-phenylacetate, $(C_6H_5CH_2COO)_2KH$, is an acid salt of phenylacetic acid and its structure was studied, using X-ray diffraction methods, by Speakman (1949). This X-ray work showed that phenylacetate residues in the structure were related by a centre of symmetry in such a way that the O–O distance between their carboxyl groups was 2·55 Å, suggesting a short hydrogen bond in which the hydrogen atom was either centrally located at the centre of symmetry or statistically distributed between two possible positions, one on each side of the centre. These two possible structures may usefully be considered in relation to the structure of KHF_2, of which $(C_6H_5CH_2COO)_2KH$ might be regarded as an aromatic analogue, where there is a very short F–H–F bond, of length only 2·26 Å, and a variety of physical evidence that the hydrogen atom is indeed centrally located (Peterson & Levy, 1952).

Experimental details

Potassium hydrogen bis-phenylacetate has monoclinic symmetry with space group $C2/c$ and

$$a = 28{\cdot}4,\ b = 4{\cdot}50,\ c = 11{\cdot}97\ \text{Å},\ \beta = 90{\cdot}4°.$$

It was prepared as described by Speakman and recrystallized from ethyl alcohol, from which it develops as laths elongated along the b axis. This habit is of the desired type, since the aim of the experimental work is to produce a projection on the (010) plane of the neutron-scattering density, but it was found very difficult to grow crystals sufficiently large to give adequately intense diffracted beams. As a very exceptional occurrence, a crystal measuring $3{\cdot}7 \times 0{\cdot}7 \times 0{\cdot}3$ cm. was produced and from this was cut a very suitable crystal of dimensions $3{\cdot}7 \times 0{\cdot}3 \times 0{\cdot}3$ cm., which was used for most of the work, together with two smaller specimens. The crystal was mounted with the b axis vertical, and 118 ($h0l$) reflexions were measured using a monochromatic neutron beam of wavelength 1·09 Å. About 100 of these reflexions were first measured with the small simple spectrometer designed for single-crystal work and described by Bacon & Dyer (1955). Subsequent checks of the intensities were car-