

symmetric two-phase s.s.'s of first rank *via* their generalized first representations (Burla, Nunzi, Giacobazzo & Polidori, 1980); (b) The estimation of the non-centrosymmetric two-phase s.s.'s of first rank *via* their generalized first representations (Busetta, Giacobazzo, Spagna & Viterbo, 1980); (c) The estimation of the one-phase s.s.'s of first rank *via* their generalized second representations (Giacobazzo, 1980b).

The results obtained in (a) and (b) improve previous results (Giacobazzo, Spagna, Vicković & Viterbo, 1979). It may be expected that the application of the probabilistic theory in (c) will be successful too. The application of the generalized upper representations to two- and three-phase s.s.'s is a difficult but not prohibitive task.

## APPENDIX

### Symbols and abbreviations

$m$  = number of symmetry operators in the space group  
 $N$  = number of atoms in the unit cell  
 $E_h$  = normalized structure factor  
 $R_h$  = magnitude of the normalized structure factor  
 $C_p \equiv (R_p, T_p)$  =  $p$ th symmetry operator  
 $R_p$  =  $p$ th rotation matrix of the point group

$T_p$  = translation vector associated with the  $p$ th rotation matrix of the point group  
 $I$  = identity  $3 \times 3$  matrix  
s.i. = structure invariant  
s.s. = structure seminvariant

### References

- BURLA, M. C., NUNZI, A., GIACOVAZZO, C. & POLIDORI, G. (1980). *Acta Cryst.* Submitted.  
BURLA, M. C., NUNZI, A., POLIDORI, G., Busetta, B. & GIACOVAZZO, C. (1980). *Acta Cryst.* **A36**, 573–578.  
BUSETTA, B., GIACOVAZZO, C., SPAGNA, R. & VITERBO, D. (1980). In preparation.  
GIACOVAZZO, C. (1975). *Acta Cryst.* **A31**, 602–609.  
GIACOVAZZO, C. (1977). *Acta Cryst.* **A33**, 933–944.  
GIACOVAZZO, C. (1978a). *Acta Cryst.* **A34**, 562–574.  
GIACOVAZZO, C. (1978b). *Acta Cryst.* **A34**, 27–30.  
GIACOVAZZO, C. (1979). *Acta Cryst.* **A35**, 296–305.  
GIACOVAZZO, C. (1980a). *Acta Cryst.* **A36**, 362–372.  
GIACOVAZZO, C. (1980b). *Acta Cryst.* **A36**, 711–715.  
GIACOVAZZO, C., SPAGNA, R., VICKOVIĆ, I. & VITERBO, D. (1979). *Acta Cryst.* **A35**, 401–412.  
GIACOVAZZO, C. & VICKOVIĆ, I. (1980). In preparation.  
HAUPTMAN, H. (1975). *Acta Cryst.* **A31**, 680–687.  
HAUPTMAN, H. (1976). *Acta Cryst.* **A32**, 934–940.  
HAUPTMAN, H. & GREEN, E. A. (1978). *Acta Cryst.* **A34**, 224–229.  
HAUPTMAN, H. & POTTER, S. (1979). *Acta Cryst.* **A35**, 371–381.

*Acta Cryst.* (1980). **A36**, 711–715

## The Estimation of the One-Phase Structure Seminvariants of First Rank by means of their Generalized Second Representations

BY CARMELO GIACOVAZZO

*Istituto di Mineralogia, Palazzo Ateneo, Università, 70121 Bari, Italy*

(Received 9 January 1980; accepted 18 February 1980)

### Abstract

The concept of generalized second representation [Giacobazzo (1980). *Acta Cryst.* **A36**, 704–711] has been used in order to estimate the one-phase structure seminvariants of first rank.

### 1. Introduction\*

In a preceding paper (Giacobazzo, 1978; from now on paper I), the estimation of the one-phase s.s.'s of first

rank was carried out by means of the joint probability distribution method. The *a priori* information exploited in the calculations was chosen according to the theory of representations of the s.s.'s (Giacobazzo, 1977). In particular, any one-phase s.s.  $\Phi$  was estimated *via* its second representation: that is to say, the knowledge of the diffraction magnitudes belonging to the second phasing shell of  $\Phi$  was exploited in order to give a probabilistic estimate of  $\Phi$ .

Burla, Nunzi, Polidori, Busetta & Giacobazzo (1980) showed that the estimates of the one-phase s.s.'s *via* their second representation are in general considerably more accurate than the corresponding estimates *via* the  $\sum_1$  relationships (Hauptman & Karle, 1953; Cochran

\* Symbols and abbreviations are defined in the Appendix.

& Woolfson, 1955; Weeks & Hauptman, 1970). This result is expected to be very useful in direct procedures for phase solution because it makes practicable both the passive (Overbeek & Schenk, 1976) and the active use of the one-phase s.s.'s.

The concept of the generalized representation (Giacovazzo, 1980*a*, *b*) enables one to exploit an amount of information (*i.e.* the knowledge of the diffraction magnitudes contained in the generalized phasing shells) larger than that accessible to the mere representations (*i.e. via* the phasing shells). The first aim of this paper is to obtain, by means of the joint probability distribution method, formulae estimating one-phase s.s.'s in all space groups up to orthorhombic *via* their generalized second phasing shells. It is expected that these formulae will be more accurate than the previous ones.

## 2. The mathematical approach

We assume that the reciprocal vectors are fixed and that the atomic coordinates are the primitive random variables. The mathematical device of the joint probability distribution functions will be used. In accordance with preceding papers, the density function  $P(E_1, \dots, E_r)$  will be calculated *via* the Gram-Charlier expansion (Klug, 1958) of the characteristic function:

$$\exp(S_2/t) \{1 + S_3/t^{3/2} + S_4/t^2 + S_5^2/2t^3 + S_3 S_4/t^{7/2} + S_5/t^{5/2} + S_3^3/t^{9/2} + \dots\},$$

where

$$S_\nu = t \sum_{r+s+\dots+w=\nu} \frac{\lambda_{rs\dots w}}{r!s!\dots w!} (iu_1)^r (iu_2)^s \dots (iu_n)^w$$

in centrosymmetric space groups, and

$$S_\nu = t \sum_{r+s+\dots+w=\nu} \frac{1}{2^{\nu/2}} \frac{\lambda_{rs\dots w}}{r!s!\dots w!} (iu_1)^r (iu_2)^s \dots (iv_n)^w$$

in non-centrosymmetric space groups.  $u_j$ ,  $j = 1, \dots, r$  are carrying variables associated with the  $E_j$ 's in the centrosymmetric space groups, and  $u_j$ ,  $v_j$ ,  $j = 1, \dots, r$  are carrying variables associated with the real and the imaginary parts of the  $E_j$ 's in the non-centrosymmetric space groups.

$\lambda_{rs\dots w}$  are the standardized cumulants of the distribution. Their values depend, for chosen values  $r, s, \dots, w$ , on the actual space group, on the statistical weights of the reflections and on their mutual correlation (see Appendices A and C of paper I). Though the algebraic expressions of the conclusive probabilistic formulae depend on the values of the standardized cumulants, their exact estimation will not be introduced into our calculations because: (a) an automatic procedure which estimates one-phase s.s.'s *via* the calculation of

the exact values of the cumulants would be too slow; (b) our practical experience (Giacovazzo, Spagna, Vicković & Viterbo, 1979) suggests that in the probabilistic formulae approximate values of the cumulants can be used without lowering their reliability too much. These approximate values can be obtained by neglecting the effects of the statistical weights of the reflections and their mutual correlation.

## 3. The expected sign of $E_{2h}$ in $P\bar{1}$ from its generalized second representation

The second phasing shell of the one-phase s.s.  $\Phi = \varphi_{2h}$  is

$$\{B\}_2 = \{R_{2h}, R_h, R_k, R_{h\pm k}, R_{2h\pm k}\}.$$

(I.32) gives the probability that the sign of  $E_{2h}$  is plus when the magnitudes in  $\{B\}_2$  are known.

The generalized second phasing shell of  $\varphi_{2h}$  (Giacovazzo, 1980*b*) is

$$\{B\}_2 = \{R_{2h}, R_h, R_k, R_{h\pm k}, R_{2h\pm k}, R_{2k}, R_{2h\pm 2k}\}.$$

In order to keep the notation of this paragraph similar to that used in paper I and in the following paragraphs, where we deal with symmetries higher than in  $P\bar{1}$ , we denote

$$E_1 = E_{2h}; \quad E_2 = E_h; \quad E_3 = E_k; \quad E_{4,1} = E_{h+k};$$

$$E_{4,2} = E_{h-k}; \quad E_{5,1} = E_{2h+k}; \quad E_{5,2} = E_{2h-k};$$

$$E_6 = E_{2k}; \quad E_{8,1} = E_{2h+2k}; \quad E_{8,2} = E_{2h-2k}.$$

By means of the mathematical approach described in § 2 we obtain

$$P_+(E_{2h}) \simeq 0.5 \times 0.5 \tanh \frac{|E_1|}{2\sqrt{N}} \left\{ \varepsilon_2 + \frac{\sum_k' A_k}{1 + \sum_k' B_k} \right\} \quad (1)$$

where

$$\begin{aligned} A_k = & \{(2R_2^2 - 1) \varepsilon_3 (\varepsilon_{4,1} \varepsilon_{5,1} + \varepsilon_{4,2} \varepsilon_{5,2} + \varepsilon_{4,1} \varepsilon_{4,2}) \\ & - 0.5 \varepsilon_3 (\varepsilon_{4,1} + \varepsilon_{4,2}) \\ & - 0.5 (\varepsilon_{4,1} \varepsilon_{5,1} + \varepsilon_{4,2} \varepsilon_{5,2}) \\ & + \frac{1}{2} (2R_3^2 - 1) (\varepsilon_{4,1} \varepsilon_{5,1} \varepsilon_{8,1} + \varepsilon_{4,2} \varepsilon_{5,2} \varepsilon_{8,2}) \\ & + \varepsilon_3 \varepsilon_6 \varepsilon_{4,1} \varepsilon_{4,2} \\ & + \frac{1}{2} \varepsilon_3 (\varepsilon_6 + 1) (\varepsilon_{4,1} \varepsilon_{8,1} + \varepsilon_{4,2} \varepsilon_{8,2}) \\ & + \frac{1}{2} \varepsilon_3 \varepsilon_6 (\varepsilon_{4,1} + \varepsilon_{4,2})\} / N, \end{aligned}$$

$$B_{\mathbf{k}} = \left\{ \frac{1}{4} \varepsilon_1 H_4(E_2) + \varepsilon_1 \varepsilon_3 (\varepsilon_{5,1} + \varepsilon_{5,2}) + \varepsilon_2 \varepsilon_3 (\varepsilon_{4,1} + \varepsilon_{4,2}) \right. \\ \left. + \varepsilon_2 (\varepsilon_{4,1} \varepsilon_{5,1} + \varepsilon_{4,2} \varepsilon_{5,2}) + \varepsilon_1 \varepsilon_{4,1} \varepsilon_{4,2} \right. \\ \left. + [\varepsilon_1 \varepsilon_6 (\varepsilon_{8,1} + \varepsilon_{8,2}) + \varepsilon_3 (\varepsilon_{5,1} \varepsilon_{8,1} + \varepsilon_{5,2} \varepsilon_{8,2}) \right. \\ \left. + \varepsilon_6 (\varepsilon_{4,1} \varepsilon_{4,2} + \varepsilon_{5,1} \varepsilon_{5,2}) \right\} / 2N, \\ \varepsilon_i = R_i^2 - 1, \quad H_4(E_i) = E_i^4 - 6E_i^2 + 3.$$

We have enclosed in square brackets the terms which are not in (I.32). The prime to the summations warns the reader that precautions have to be taken in order to avoid duplications in the contributions when  $\mathbf{k}$  sweeps over reciprocal space.

In order to avoid overestimations,  $\sum_{\mathbf{k}} B_{\mathbf{k}}$  should be assumed to be zero if it is experimentally found negative. Some properties of (1) deserve to be stressed and compared with those enjoyed by (I.32) and by the corresponding  $\sum_1$  relationship. Unlike the  $\sum_1$  formula, (I.32) is able in principle to give information about the sign of  $E_{2\mathbf{h}}$  even when  $|E_{\mathbf{h}}| = 1$ . On the other hand, less favourable situations are expected when  $|E_{\mathbf{h}}| < 1$ . In (1) some contributions, *a priori* not negligible, do not depend on the value of  $|E_{\mathbf{h}}|$ . This situation suggests that the sign of any  $E_{2\mathbf{h}}$  reflection can in principle be evaluated by (1). This is due to the fact that whereas the mere second representation exploits only quintets of which  $E_{\mathbf{h}}$  is a basis reflection, the generalized second representation exploits also quintets of which  $E_{\mathbf{h}}$  is a cross reflection.

**4. The expected sign of  $E_{\mathbf{H}} = E_{\mathbf{h}(I-\mathbf{R}_n)}$  in any centrosymmetric space group up to orthorhombic from its generalized second representation**

Let  $\Phi = \varphi_{\mathbf{H}} = \varphi_{\mathbf{h}(I-\mathbf{R}_n)}$  be our one-phase s.s. of first rank. For a fixed  $\mathbf{H}$ ,  $\mathbf{h}$  is a free vector ( $\mathbf{h} \in \{\mathbf{h}\}$ ) under the condition

$$\mathbf{h}(I - \mathbf{R}_n) = \mathbf{H}. \quad (2)$$

More matrices  $\mathbf{R}_n$  can satisfy (2) and consequently more sets  $\{\mathbf{h}\}$  can exist.

When  $\Phi$  is estimated *via* its second representation the set of special quintets

$$\{\Psi\}_2 = \varphi_{\mathbf{H}} - \varphi_{\mathbf{h}} + \varphi_{\mathbf{hR}_n} - \varphi_{\mathbf{kR}_j} + \varphi_{\mathbf{kR}_j}, \quad j = 1, \dots, m/2 \quad (3)$$

is exploited, from which

$$\{B\}_2 = \{R_{\mathbf{H}}, R_{\mathbf{h}}, R_{\mathbf{k}}, R_{\mathbf{h}+\mathbf{kR}_j}, R_{\mathbf{H}+\mathbf{kR}_j}, \quad j = 1, \dots, m\}. \quad (4)$$

Since  $\mathbf{h}$  under certain conditions and  $\mathbf{k}$  are free vectors in (3), the joint probability distribution

$P(E_{\mathbf{H}}, \{E_{\mathbf{h}}\}, \{E_{\mathbf{k}}\}, \{E_{\mathbf{h}+\mathbf{kR}_j}\}, \{E_{\mathbf{H}+\mathbf{kR}_j}\}, \quad j = 1, \dots, m)$  was studied in paper I from which the conclusive formula (I.36) arises.

When  $\Phi$  is estimated *via* its generalized second representation, besides (3) the special quintets (Giacovazzo, 1980b)

$$\{\Psi'\}_2 = \{\varphi_{\mathbf{H}} - \varphi_{\mathbf{h}+\mathbf{kR}_j} + \varphi_{(\mathbf{h}+\mathbf{kR}_j)\mathbf{R}_n} + \varphi_{\mathbf{kR}_j} - \varphi_{\mathbf{kR}_j\mathbf{R}_n}, \\ j = 1, \dots, m/2\} \quad (5)$$

are also exploited. Then the generalized second phasing shell is

$$\{B\}'_2 = \{R_{\mathbf{H}}, R_{\mathbf{h}}, R_{\mathbf{k}}, R_{\mathbf{h}+\mathbf{kR}_j}, R_{\mathbf{H}+\mathbf{kR}_j}, R_{\mathbf{k}(I-\mathbf{R}_n)}, \\ R_{\mathbf{h}+\mathbf{kR}_j(I+\mathbf{R}_n)}, R_{\mathbf{H}+\mathbf{kR}_j(I-\mathbf{R}_n)}, \quad j = 1, \dots, m\}. \quad (6)$$

Since  $\mathbf{h}$  under certain conditions and  $\mathbf{k}$  are free vectors in (3) and (6), the joint probability distribution

$$P(E_{\mathbf{H}}, \{E_{\mathbf{h}}\}, \{E_{\mathbf{k}}\}, \{E_{\mathbf{h}+\mathbf{kR}_j}\}, \{E_{\mathbf{H}+\mathbf{kR}_j}\}, \{E_{\mathbf{k}(I-\mathbf{R}_n)}\}, \\ \{E_{\mathbf{h}+\mathbf{kR}_j(I+\mathbf{R}_n)}\}, \{E_{\mathbf{H}+\mathbf{kR}_j(I-\mathbf{R}_n)}\}, \quad j = 1, \dots, m) \quad (7)$$

has to be studied. In (7),  $\{E_{\mathbf{h}}\}$  is the set of structure factors whose indices belong to  $\{\mathbf{h}\}$ ,  $\{E_{\mathbf{k}}\}$  is any chosen set in the asymmetric region of reciprocal space and  $\{E_{\mathbf{h}+\mathbf{kR}_j}, \dots, \{E_{\mathbf{H}+\mathbf{kR}_j(I-\mathbf{R}_n)}\}$  are sets obtainable from the specified condition on  $\mathbf{h}$  and  $\mathbf{k}$ .

In order to describe in a simple way our results whatever the space group may be we generalize the notation introduced in paper I in the following way:

$$E_1 = E_{\mathbf{H}}; \quad E_2 = E_{\mathbf{h}}; \quad E_3 = E_{\mathbf{k}}; \quad E_{4,j} = E_{\mathbf{h}+\mathbf{kR}_j}; \\ E_{5,j} = E_{\mathbf{H}+\mathbf{kR}_j}; \quad E_6 = E_{\mathbf{k}(I-\mathbf{R}_n)}; \quad E_{7,j} = E_{\mathbf{h}+\mathbf{kR}_j(I+\mathbf{R}_n)}; \\ E_{8,j} = E_{\mathbf{H}+\mathbf{kR}_j(I-\mathbf{R}_n)}.$$

We obtain

$$P_+(E_{\mathbf{H}}) \simeq 0.5 \\ + 0.5 \tanh \left\{ \frac{|E_1|}{2\sqrt{N}} \left( \sum_{n,\mathbf{k}} \varepsilon_2 + \frac{\sum_{n,\mathbf{h},\mathbf{k}} A_{\mathbf{h},\mathbf{k}}}{1 + \sum_{n,\mathbf{h},\mathbf{k}} B_{\mathbf{h},\mathbf{k}}} \right) (-1)^{2\mathbf{hT}_n} \right\}, \quad (8)$$

where

$$A_{\mathbf{h},\mathbf{k}} = \left\{ (2E_2^2 - 1) \varepsilon_3 \left( \sum_{\substack{\mathbf{R}_j=\mathbf{R}_i \\ \mathbf{R}_j=-\mathbf{R}_i\mathbf{R}_n}} \varepsilon_{4,i} \varepsilon_{5,j} + \sum_{\substack{\mathbf{R}_j=\mathbf{R}_i\mathbf{R}_n \\ \mathbf{R}_j=\mathbf{R}_i\mathbf{R}_n}} \varepsilon_{4,i} \varepsilon_{4,j} \right) \right. \\ \left. - \frac{1}{2} \varepsilon_3 \sum_j \varepsilon_{4,j} - \frac{1}{2} \sum_{\substack{\mathbf{R}_j=\mathbf{R}_i \\ \mathbf{R}_j=-\mathbf{R}_i\mathbf{R}_n}} \varepsilon_{4,i} \varepsilon_{5,j} \right. \\ \left. + \left[ 2\varepsilon_3 \left( \sum_{\substack{\mathbf{R}_j=\mathbf{R}_i \\ \mathbf{R}_j=-\mathbf{R}_i\mathbf{R}_n}} \varepsilon_{4,i} \varepsilon_{5,j} \varepsilon_{7,i} + \sum_{\substack{\mathbf{R}_j=\mathbf{R}_i\mathbf{R}_n \\ \mathbf{R}_j=\mathbf{R}_i\mathbf{R}_n}} \varepsilon_{4,i} \varepsilon_{4,j} \varepsilon_{7,i} \right) \right] \right.$$

$$\begin{aligned}
& + \varepsilon_3 \left( \varepsilon_{4,1} \varepsilon_{4,R_n} \varepsilon_6 + \varepsilon_{4,-1} \varepsilon_{4,-R_n} \varepsilon_6 \right. \\
& \left. + \frac{1}{2} \varepsilon_6 \sum_i \varepsilon_{4,i} \varepsilon_{8,i} + \frac{1}{2} \varepsilon_6 \sum_i \varepsilon_{4,i} + \frac{1}{2} \sum_i \varepsilon_{4,i} \varepsilon_{8,i} \right) \\
& \left. + (E_3^2 - \frac{1}{2}) \sum_i \varepsilon_{4,i} \varepsilon_{5,i} \varepsilon_{8,i} \right] \Big/ N, \\
B_{h,k} = & \left\{ -\frac{1}{4} \sum_j H_4(E_j) + \varepsilon_1 \varepsilon_3 \sum_j \varepsilon_{5,j} + \varepsilon_1 \sum_{\substack{R_j=R,R_n \\ R_j=R,R_n}} \varepsilon_{4,i} \varepsilon_{4,j} \right. \\
& + \varepsilon_2 \varepsilon_3 \sum_j \varepsilon_{4,j} + \varepsilon_2 \sum_{\substack{R_j=R_i \\ R_j=-R_i R_n}} \varepsilon_{4,i} \varepsilon_{5,j} + \frac{1}{4} \varepsilon_1 H_4(E_2) \\
& + \left[ \varepsilon_1 \varepsilon_6 \sum_j \varepsilon_{8,j} + \varepsilon_3 \sum_j (\varepsilon_{4,j} \varepsilon_{7,j} + \varepsilon_{5,j} \varepsilon_{8,j}) \right. \\
& + \sum_{\substack{R_j=R_i \\ R_j=-R_i R_n}} \varepsilon_{4,i} \varepsilon_{5,j} \varepsilon_{7,i} + \varepsilon_6 (\varepsilon_{4,1} \varepsilon_{4,R_n} \\
& + \varepsilon_{4,-1} \varepsilon_{4,-R_n} + \varepsilon_{5,1} \varepsilon_{5,R_n} + \varepsilon_{5,-1} \varepsilon_{5,-R_n} \\
& + \frac{1}{4} \varepsilon_1 \sum_j H_4(E_{7,j}) + \frac{1}{4} \varepsilon_6 H_4(E_3) \\
& \left. \left. + \frac{1}{4} \sum_j \varepsilon_{8,j} H_4(E_{4,j}) \right] \right\} \Big/ 2N.
\end{aligned}$$

We have enclosed in square brackets the terms which are not in (I.36). The summation over the index  $n$  in (8) takes into account the fact that more matrices  $\mathbf{R}_n$  can satisfy (2). It may be noted that in some terms in  $A_{h,k}$  and  $B_{h,k}$  we have replaced the index  $j$  by specific rotation matrices. As in (I.36) precautions have to be taken in (8) to avoid duplication of the contributions to  $\sum_{n,h,k} A_{h,k}$  and  $\sum_{n,h,k} B_{h,k}$  when  $n$ ,  $h$  and  $k$  vary over their allowed values.

The derivation of (8) requires the application of space-group algebra to the joint probability distribution method. The reader will find a short description of the procedure in the appendices of paper I.

Equation (8) has been specifically derived for space groups with symmetry up to orthorhombic but it is expected to hold approximately for space groups with higher symmetry.

The following may be noted. (a) If  $\mathbf{R}_n$  corresponds to a symmetry operator of order two then the condition  $\mathbf{R}_i = \mathbf{R}_j \mathbf{R}_n$  coincides with  $\mathbf{R}_j = \mathbf{R}_i \mathbf{R}_n$ . Furthermore,  $R_{7,j} \in \{R_h\}$ . In fact,  $h' = h + kR_j(\mathbf{I} + \mathbf{R}_n)$  satisfies (2). (b) When  $\mathbf{R}_n = -\mathbf{I}$  the condition  $\mathbf{R}_j = -\mathbf{R}_i \mathbf{R}_n$  coincides

with  $\mathbf{R}_j = \mathbf{R}_i$ ; furthermore,  $E_{7,j} \equiv E_h$  whatever  $j$ . (c) In  $P\bar{1}$ , (8) coincides with (1).

### 5. Expected value of $\varphi_H = \varphi_{h(1-R_n)}$ in the non-centrosymmetric space groups up to orthorhombic via its generalized second representation

In the space groups up to orthorhombic, all the one-phase s.s.'s of first rank have restricted phase values' (*i.e.* 0 or  $\pi$ ). Therefore, one has only to estimate their signs. In accordance with paper I we introduce the fictitious (not belonging to the space group) symmetry operators  $\mathbf{C}_{m+j} = (-\mathbf{R}_j, -\mathbf{T}_j)$ ,  $j = 1, \dots, m$ . Then the generalized second phasing shell of  $\varphi_H$  is again given by (6) provided  $m' = 2m$  replaces  $m$ . That enables us to use the same notation as in § 4. Our result is that (8) holds for non-centrosymmetric groups too.

In non-centrosymmetric space groups with symmetry higher than orthorhombic the one-phase s.s.'s are in general non-centrosymmetric. It may be expected that

$$\cos \varphi_H \simeq I_1(2G)/I_0(2G) \quad (9)$$

approximately estimates  $\cos \varphi_H$ , where  $G$  is the argument of the hyperbolic tangent in (8).

### 6. Conclusions

We have obtained some probabilistic expressions [*i.e.* equations (8) and (9)] which estimate one-phase s.s.'s of first rank *via* their generalized second representations. If these expressions are compared with those obtained *via* the mere second representations [*i.e.* equations (I.36) and (I.40)], some new terms can be recognized. The contribution of these new terms seems not negligible and can remarkably improve (in the probabilistic sense) the accuracy of the estimates.

### APPENDIX Symbols and abbreviations

$m$	= number of symmetry operators in the space group
$N$	= number of atoms in the unit cell
$E_h$	= normalized structure factor
$R_h$	= magnitude of the normalized structure factor
$\mathbf{C}_p$	= $(\mathbf{R}_p, \mathbf{T}_p)$ = $p$ th symmetry operator
$\mathbf{R}_p$	= $p$ th rotation matrix of the point group
$\mathbf{T}_p$	= translation vector associated with the $p$ th rotation matrix of the point group
$\mathbf{I}$	= identity $3 \times 3$ matrix
s.s.	= structure seminvariant
$t$	= number of atoms in the asymmetric unit

## References

- BURLA, M. C., NUNZI, A., POLIDORI, G., Busetta, B. & GIACOVAZZO, C. (1980). *Acta Cryst.* **A35**, 573–578.
- COCHRAN, W. & WOOLFSON, M. M. (1955). *Acta Cryst.* **8**, 1–12.
- GIACOVAZZO, C. (1977). *Acta Cryst.* **A33**, 933–944.
- GIACOVAZZO, C. (1978). *Acta Cryst.* **A34**, 562–574.
- GIACOVAZZO, C. (1980a). *Acta Cryst.* **A36**, 362–372.
- GIACOVAZZO, C. (1980b). *Acta Cryst.* **A36**, 704–711.
- GIACOVAZZO, C., SPAGNA, R., VICKOVIĆ, I. & VITERBO, D. (1979). *Acta Cryst.* **A35**, 401–412.
- HAUPTMAN, H. & KALRE, J. (1953). *Solution of the Phase Problem. I. The Centrosymmetric Crystal*. ACA Monograph No. 3. Pittsburgh: Polycrystal Book Service.
- KLUG, A. (1958). *Acta Cryst.* **11**, 513–543.
- OVERBEEK, A. R. & SCHENK, H. (1976). *Proc. K. Ned. Akad. Wet.* **B79**, 341–343.
- WEEKS, C. M. & HAUPTMAN, H. (1970). *Z. Kristallogr.* **131**, 437–442.

*Acta Cryst.* (1980). **A36**, 715–723

## Calculated Energy and Conformation of Clusters of Benzene Molecules and Their Relationship to Crystalline Benzene

BY DONALD E. WILLIAMS

*Department of Chemistry, University of Louisville, Louisville, KY 40208, USA*

(Received 18 July 1979; accepted 4 March 1980)

### Abstract

The detailed conformations of benzene clusters containing 2, 3, 5, 7, 9, 11, 13 and 15 molecules were calculated. The nonbonded potential energy of the clusters was minimized by the Newton–Raphson method with exp–6–1 potential functions. All of the clusters exhibited a predominating edge-to-side or herringbone pattern of packing. The concept of intershell coordination, as contrasted to ligand coordination, was introduced and illustrated with the undecamer and larger clusters. The pentadecamer clearly showed the beginning of a second coordination shell. The tridecamer conformation was related to a 13-molecule fragment from crystalline orthorhombic benzene. This crystal fragment has a higher energy than the tridecamer. The fragment can convert to the tridecamer conformation by a process of plane slippage with cooperative molecular motion. Two examples, an isoheptamer and an isotridecamer conformation, are given of clusters with lower total cluster energy but with a higher energy for the reference molecule. In neither case do the conformations follow a smooth trend with increasing cluster size. The isotridecamer has approximate threefold symmetry and has a conformation quite different from the crystal fragment.

### Introduction

Molecular complexation or association and subsequent crystallization are caused by the action of weak

nonbonded forces between molecules. For hydrocarbons considerable progress has been made in the elucidation of the quantitative nature of these forces in the crystal. The procedure used is to assume a reasonably simple and theoretically justifiable model for the nonbonded potential energy. The coefficients of the nonbonded energy functions are then adjusted to give the best fit to several crystal structures (Williams, 1966; Kitaigorodsky, 1973).

The model for the nonbonded potential energy used here is referred to as exp–6–1:

$$V_{jk} = B \exp(-Cr_{jk}) - Ar_{jk}^{-6} + q_j q_k r_{jk}^{-1}.$$

$V_{jk}$  is a nonbonded pair potential between atoms  $j$  and  $k$  in different molecules separated by distance  $r$ . The first term is the short-range repulsion energy caused by overlap of filled electron shells. The values of  $C$  are usually estimated from theory. Values of  $B$  for  $C \cdots C$ ,  $C \cdots H$  and  $H \cdots H$  interactions are found by fitting observed crystal structures. The second term is the dispersion attraction energy and values for the coefficients  $A$  may also be fitted from observed crystal structures. For hydrocarbons, the values of the net atomic charges,  $q$ , are small but not negligible. In this study we use the nonbonded parameters of Williams & Starr (1977), set II. These nonbonded parameters are given in Table 1; the geometric mean combining law was used for  $C \cdots H$  interactions. These parameters give a good fit to the crystal structures of a set of nine aromatic plus nine saturated hydrocarbon crystal structures. It seems reasonable to assume that these