Acta Cryst. (1980). A**36**, 573–578

Practical Aspects and Applications of Some Probabalistic Formulae which Estimate One-Phase Seminvariants

BY M. C. BURLA AND A. NUNZI

Istituto di Mineralogia dell'Università, 06100 *Perugia, Italy*

G. POLIDORI

Centro Calcolo Elettronico dell'Università, 06100 *Perugia, Italy*

AND B. BUSETTA* AND C. GIACOVAZZO

Istituto di Mineralogia dell'Università, 70121 *Bari, Italy*

(*Received* 25 *May* 1979; *accepted* 1 *January* 1980)

Abstract

Some algorithms have been derived to calculate the expected values of one-phase seminvariants of first rank in space groups up to orthorhombic by a probabilistic approach [Giacovazzo (1978). *Acta Cryst.* A**34**, 562–576]. The method has been tested on several known structures. The results show how the method can secure a very reliable estimate for a limited number of one-phase seminvariants which can be used from the initial stages of phase determination. In addition, the estimated values can be used as a good figure of merit to select the correct $\sum_2$ solution in a multisolution procedure.

1. Introduction

A one-phase structure seminvariant is the phase, $\varphi_{\mathbf{H}}$, of a reflexion, the indices of which satisfy the condition $\mathbf{H} \equiv 0(\text{mod } \boldsymbol{\omega}_s)$ where $\boldsymbol{\omega}_s$ is the seminvariant modulus of the given space group. According to representations theory (Giacovazzo, 1977), $\varphi_{\mathbf{H}}$ is a structure seminvariant of first rank if there is at least one vector $\mathbf{h}$ ($E_{\mathbf{h}}$ may be measured or not) and a rotation matrix $\mathbf{R}_n$ for which

$$\mathbf{H} = \mathbf{h}\mathbf{D}_n = \mathbf{h}(\mathbf{I} - \mathbf{R}_n). \qquad (1)$$

In (1), $\mathbf{I}$ denotes the identity matrix.

Probabilistic theories for estimation of one-phase seminvariants of first rank have been supplied by various authors (*e.g.* Hauptman & Karle, 1953, 1957; Cochran & Woolfson, 1954, 1955; Naya, Nitta & Oda, 1964; Weeks & Hauptman, 1970; Hauptman, 1972; Giacovazzo, 1975). Unfortunately, $\sum_1$ formulae for representative space groups only have been worked out; moreover, the frequence of failures suggested that one-phase seminvariants should not be used in the early stages of direct procedures.

An alternative use of one-phase seminvariants has been suggested by Overbeek & Schenk (1976) who formulated a criterion, based on the $\sum_1$ relations, which can be used to select the correct $\sum_2$ solution in multisolution procedures. The criterion led, in some artificial monoclinic structures, to a successful prediction of the correct structure and proved by far superior to the negative-quartet criterion, most notably so in the case of large structures. However, Overbeek & Schenk (1976) made full use of the space-group symmetry for the $\sum_1$ criterion, while they applied to monoclinic structures the usual negative-quartet criterion as derived in the triclinic system. On the other hand, a recent paper by Busetta *et al.* (1980) shows that the effectiveness of the negative-quartet criteria strongly increases when full use of the space-group symmetry is made [see Giacovazzo (1976) for the theoretical background]. The results previously mentioned strongly suggest that $\sum_1$ and negative-quartet criteria are two of the most effective figures of merit for selecting the correct solution from a large number of possible candidates often produced by multisolution techniques.

More recently, the theory of representations has given new insights into probabilistic methods for obtaining accurate estimates of one-phase seminvariants of first rank (Giacovazzo, 1978*a*). The formulae yield an estimate for any seminvariant $\varphi_{\mathbf{H}}$ *via* the magnitudes contained in the first and second representations of $\varphi_{\mathbf{H}}$. Since a large amount of information can be exploited, it is expected that one-phase seminvariants can be reliably estimated. In this paper we deal with practical aspects and appli-

* Present address: Laboratoire de Cristallographie (LA 144 CNRS), Université de Bordeaux I, 33405 Talence, France.

0567-7394/80/040573-06$01.00

cations of the theory described by Giacovazzo (1978*a*) (hereafter referred to as paper I), and with the new role that one-phase seminvariants can play in direct procedures.

2. An algorithm for the estimation of one-phase seminvariants in space groups up to orthorhombic

Let $\varphi_{\mathbf{H}}$ be a first-rank seminvariant. Because of (1), the first representation of $\varphi_{\mathbf{H}}$ is the collection of the special triplets

$$\psi_1 = \varphi_{\mathbf{H}} - \varphi_{\mathbf{h}} + \varphi_{\mathbf{hR}_n}. \qquad (2)$$

The information contained in the first representation of $\varphi_{\mathbf{H}}$ will be fully exploited when all the **h** in (1) are known. Property 2 of paper I shows that the use of the generalized inverse matrices is the most general way to search for vectors **h**. Then,

$$\mathbf{h} = \mathbf{H}\mathbf{D}_n^* + \mathbf{Z}(\mathbf{I} - \mathbf{D}_n^*\mathbf{D}_n), \qquad (3)$$

where $\mathbf{D}_n^*$ is the generalized inverse of $\mathbf{D}_n$ and **Z** is a free vector in reciprocal space. When considering only the space groups up to orthorhombic, the calculation of the right-hand side of (3) is rather simple. In fact, any $\mathbf{D}_n = [d_{ij}]$ is a matrix for which $d_{ij} = 0$ for $i \neq j$. Then, $\mathbf{D}_n^* = [d_{ij}^*]$ is a matrix for which $d_{ij}^* = 0$ for $i \neq j$; $d_{ii}^* = 0$ if $d_{ii} = 0$, otherwise $d_{ii}^* = 1/d_{ii}$.

As an example, let us consider, in class 222, the seminvariant $\varphi_{\mathbf{H}} = \varphi_{400}$. The matrices for which (1) is satisfied are

$$\mathbf{R}_p = \begin{vmatrix} \bar{1} & 0 & 0 \\ 0 & \bar{1} & 0 \\ 0 & 0 & 1 \end{vmatrix}, \quad \mathbf{R}_q = \begin{vmatrix} \bar{1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \bar{1} \end{vmatrix},$$

so that

$$\mathbf{D}_p = \begin{vmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 0 \end{vmatrix}, \quad \mathbf{D}_q = \begin{vmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{vmatrix}.$$

If $n = p$, then (3) gives $\mathbf{h} = 20l$; if $n = q$ then (3) gives $\mathbf{h} = 2k0$, where l and k are free indices. Assuming that the maximum values of the indices k and l are known (*i.e.* MAXK and MAXL), we are immediately able to construct the first representation of φ_{400}:

$$\{\psi\}_1 = \{\varphi_{400} - \varphi_{20l} + \varphi_{\bar{2}0l}\} \cup \{\varphi_{400} - \varphi_{2k0} + \varphi_{\bar{2}k0}\},$$

where $0 \leq l \leq \text{MAX}L$, $0 < k \leq \text{MAX}K$. The set of magnitudes $|E_{400}|$, $|E_{20l}|$, $|E_{2k0}|$ is called the first phasing shell of φ_{400}.

A computer program was written to generate the one-phase seminvariants of first rank in all the space groups up to orthorhombic and to define the magnitudes in the first phasing shell as described above. As the whole reciprocal space is stored in the central core of the computer, the search for the **h** vectors is very fast and does not require the previous search of the $\sum_2$ relations as in systems such as *MULTAN* (Germain, Main & Woolfson, 1971). A further advantage of the algorithm is that the contribution of the small $|E_{\mathbf{h}}|$'s to the probabilistic estimation of $\varphi_{\mathbf{H}}$ can be considered too.

Paper I gives also probabilistic formulae which estimate $\varphi_{\mathbf{H}}$ *via* its second representation. We recall here that the second representation of $\varphi_{\mathbf{H}}$ is the collection of special quintets:

$$\psi_2 = \varphi_{\mathbf{H}} - \varphi_{\mathbf{h}} + \varphi_{\mathbf{hR}_n} - \varphi_{\mathbf{kR}_j} + \varphi_{\mathbf{kR}_j}, \quad j = 1, \ldots, r \qquad (4)$$

where **h** varies over the set of vectors defined in $\{\psi\}_1$ and $\mathbf{R}_n$ over the subset of matrices which satisfy (1). According to (I.34) and (I.39), $\mathbf{R}_j$ varies over the subset of $r = m/2$ matrices not related by the centre of symmetry when the space group is centrosymmetric (m

Table 1. *Abbreviations, references, space groups, chemical formulae (obs meaning E's computed from the observed intensities and calc E's calculated from the atomic coordinates) and minimum values for* $|E_{\mathbf{H}}|$, $|E_{\mathbf{h}}|$ *and* $|E_{\mathbf{k}}|$

	Reference	Space group	Formula	Z	E	$E_{H\min}$	$E_{h\min}$	$E_{k\min}$
TETRA	Spagna & Vaciago (1978)	$P\bar{1}$	$C_{30}H_{32}N_4$	1	obs	1·50	1·50	1·80
RIBO	James & Stevens (1977)	$P2_1$	$C_{13}H_{18}O_9$	4	calc	1·70	1·30	1·83
METHOX	Hanson & Nordman (1975)	$P2_1/c$	$C_{18}H_{22}O_2$	4	calc	1·68	1·30	1·80
TOLY	Brufani, Cellai, Cerrini, Fedeli & Vaciago (1978)	$P2_12_12_1$	$C_{37}H_{43}NO_{13}$	4	obs	1·70	1·30	1·85
TOXE	Cerrini, Fedeli, Gavuzzo & Mazza (1975)	$P2_12_12_1$	$C_{21}H_{35}O_5$	4	obs	1·57	1·50	1·90
KARLE	Karle, Karle & Estlin (1967)	$P2_12_12_1$	$C_{12}H_{13}NO_4$	4	obs	1·40	1·30	1·50
AZET	Colens, Declercq, Germain, Putzeys & Van Meerssche (1974)	$Pca2_1$	$C_{21}H_{16}ClNO$	8	obs	1·35	1·50	1·65

is the order of the space group), whereas $r = m$ for a non-centrosymmetric space group. **k** is a free vector which sweeps over the asymmetric region of reciprocal space.

For every **h** the program calculates indices and magnitudes of the basis and cross reflexions of the special quintets (4). The set of these magnitudes is called the second phasing shell of $\varphi_{\mathbf{H}}$.

The probabilistic method described in paper I was tested on seven known structures of different complexity covering five of the seven symmetry classes of the first three crystallographic systems in which one-phase seminvariants can be found. Table 1 shows the references and the most relevant features of the test structures whose results are discussed in this paper.

With the aim of comparing the reliability of different probabilistic formulae, we constructed the following criterion in which use is made of all P^+ values associated with the signs of the seminvariants:

$$\mathrm{ARI} = \frac{2}{n} \sum_l S_l \Delta_l^+,$$

where ARI means 'averaged reliability index', n is the number of seminvariants considered in the calculations, S_l is the true sign of the lth seminvariant and $\Delta_l^+ = P_l^+ - 0{\cdot}5$. If all the seminvariants are correctly defined with the highest reliability, then $\sum_l S_l \Delta_l^+ = n/2$ and ARI = 1. The worst situation occurs when all the seminvariants are incorrectly defined with the highest reliability; then ARI = −1.

3. The estimation of $\varphi_{\mathbf{H}}$ from its first representation: practical aspects

Three types of formulae were given in paper I in order to estimate $\varphi_{\mathbf{H}}$ *via* its first representation: (I.14), (I.17) and (I.28). The first one holds in centrosymmetric as well as in non-centrosymmetric space groups, the second was specifically derived in centrosymmetric space groups and the third for centrosymmetric seminvariants in non-centrosymmetric space groups. Moreover, (I.14) arises from the Gram–Charlier expansion of the characteristic function of the joint probability distribution of the E's in the first phasing shell, while (I.17) and (I.28) are directly obtained from the Fourier transform of the characteristic function. In agreement with theoretical expectation, (I.14), (I.17) and (I.28) give nearly identical results for all space groups. Formula (I.14) can thus be chosen for subsequent use also because it coincides with the classical $\sum_1$ relationship, currently used to estimate one-phase seminvariants in direct procedures for phase solution (*e.g.* in *MULTAN*).

A more interesting test is that of comparing the reliability of the formulae when the following conditions are used.

(*a*) Only the largest $|E_{\mathbf{h}}|$ values are used. In usual direct procedures only $|E|$'s larger than a fixed threshold are introduced for the search of $\sum_2$ relationships. Depending on the structural complexity, on the symmetry class and on the particular structure, the value of this threshold can vary between about 1·30 and 1·70. This value fixes also the minimum magnitudes for the $E_{\mathbf{h}}$'s which can be phased and for the

Table 2. *AZET: indices and E's of the one-phase seminvariants of first rank together with the corresponding probabilities*

h k l	E	$P^+(E)$ (I.14)$^+$	$P^+(E)$ (I.14)	$P^+(E)$ (I.17)	$P^+(E)$ (I.28)
0 4 0	3·46	1·00	1·00	1·00	1·00
8 2 0	−2·78	0·50	0·34	0·33	0·34
0 6 0	−2·67	0·0	0·0	0·01	0·01
12 2 0	−2·36	0·44	0·35	0·35	0·35
8 4 0	2·31	0·69	0·80	0·81	0·80
28 4 0	2·20	0·50	0·42*	0·42*	0·42*
22 6 0	1·71	0·68	0·67	0·67	0·67
24 6 0	1·61	0·50	0·48*	0·48*	0·48*
30 2 0	−1·56	0·40	0·33	0·33	0·33
14 0 0	−1·56	0·21	0·12	0·11	0·11
30 0 0	1·55	0·73	0·90	0·90	0·90
12 6 0	−1·45	0·56*	0·58*	0·58*	0·58*
24 4 0	1·39	0·64	0·66	0·66	0·66
0 8 0	1·38	0·97	0·98	0·97	0·97

* Incorrect evaluations.

Table 3. *KARLE: indices and E's of the one-phase seminvariants of first rank together with the corresponding probabilities*

h k l	E	$P^+(E)$ (I.14)$^+$	$P^+(E)$ (I.14)	$P^+(E)$ (I.17)	$P^+(E)$ (I.28)
4 2 0	−2·54	0·17	0·04	0·03	0·04
0 0 12	−2·42	0·50	0·44	0·44	0·44
0 4 8	2·22	0·50	0·65	0·67	0·66
0 2 8	−2·05	0·50	0·69*	0·70*	0·70*
0 4 14	1·95	0·79	0·82	0·83	0·82
4 0 4	−1·87	0·50	0·47	0·47	0·47
0 6 18	−1·82	0·31	0·24	0·23	0·23
4 0 0	1·81	0·50	0·71	0·72	0·71
0 6 16	1·76	0·50	0·54	0·54	0·54
4 0 12	−1·53	0·37	0·36	0·35	0·36
0 4 4	−1·51	0·51*	0·51*	0·51*	0·51*
0 4 12	−1·46	0·39	0·27	0·27	0·27

* Incorrect evaluations.

Table 4. *The values of ARI for* (I.14)$^+$ *and* (I.14)

	(I.14)$^+$	(I.14)
TETRA	0·03	0·05
RIBO	0·15	0·20
METHOX	0·05	0·07
TOLY	0·12	0·23
TOXE	0·27	0·37
KARLE	0·17	0·28
AZET	0·38	0·41

$|E_h|$'s which can be exploited. The threshold values used for our tests are given in Table 1 and are labelled as $E_{H\min}$.

(*b*) All the $|E_h|$'s contained in the first phasing shell of φ_H are used.

In Tables 2 and 3 the results are described in some detail for AZET and KARLE respectively. The asterisks indicate incorrect evaluations. The values of ARI for the seven test structure are shown in Table 4. We have labelled in the tables by $(I.14)^+$ the results obtained by (I.14) when only the $|E_h|$'s larger than $E_{H\min}$ are used.

The results suggest that, when possible, information contained in small $|E_h|$ reflexions should not be neglected. In particular, seminvariants which are estimated positive when only the largest $|E_h|$'s are used can become negative if all the $|E_h|$'s are used, and *vice versa*. Moreover, it is worth noticing that the ARI values are always better for (I.14) than for $(I.14)^+$.

A first conclusion may be that computer-based procedures which store the whole reciprocal space are in principle more able to exploit information contained in diffraction data. On the other hand, only those procedures can estimate one-phase seminvariants *via* their second representations.

4. The estimation of φ_H from its second representation: practical aspects

Two types of procedure were described in paper I to estimate φ_H *via* the magnitudes contained in its second phasing shell. The first uses the joint probability distribution of the complete set of E's belonging to the second phasing shell of φ_H and leads to (I.36).

The second procedure uses the idea of complementary invariants described in general terms by Giacovazzo (1977). The complementary quartet invariants

$$\varphi_4 = \varphi_H - \varphi_h - \varphi_k + \varphi_{hR_n+k}$$

are estimated *via* their first representations. If $\mathbf{k}$ is allowed to vary over the region of reciprocal space for which $|E_{hR_n}E_kE_{hR_n+k}|$ is large, then $\varphi_3 = \varphi_{hR_n} + \varphi_k - \varphi_{hR_n+k}$ can be assumed to be nearly zero. Since $\varphi_{hR_n} = \varphi_h - 2\pi\mathbf{hT}_n$, then $\varphi_4 \simeq \varphi_H - 2\pi\mathbf{hT}_n$, from which φ_H is fixed if φ_4 is correctly estimated. The usefulness of the method is increased by the fact that several reliable quartets can be found for a given φ_H.

Expression (I.44) gives the probability that E_H is positive when the first representation of the jth complementary quartet is used. A measure of the overall probability that E_H has a positive sign when more complementary quartets are used is given by

$$P^+ \simeq \left(1 + \frac{\Pi_l P_j^-}{\Pi_j P_j^+}\right)^{-1}. \qquad (5)$$

A series expansion of (5) leads to

$$P^+ \simeq 0{\cdot}5 + 0{\cdot}5 \tanh\left\{\frac{R_H}{2\sqrt{N}}\left[\sum_{n,\mathbf{h}}\left(\varepsilon_h + \frac{2}{\sqrt{N}} \times \frac{\sum_{\mathbf{k}} A_{h,k}}{1 + \sum_{\mathbf{k}} B_{h,k}}\right)(-1)^{2\mathbf{hT}_n}\right]\right\}, \qquad (6)^*$$

where the symbols have the same meaning as in paper I.

The calculation of (I.36), (I.44) and (I.45) are time consuming if $\mathbf{k}$ sweeps over all reciprocal space. In our tests, $\mathbf{k}$ is restricted to range over the subset of reflexions with $E_k \geq E_{k\min}$. The chosen values of $E_{k\min}$ for the various structures are shown in the last column of Table 1. They were arbitrarily fixed depending on the structure complexity and the symmetry class.

Some other measurements were taken in order to improve the estimation process. For (I.36), the use of Hermite polynomials of order four proved not to be essential; so they were not used in the calculations. On the other hand, in accordance with §I.8, precautions were taken to avoid duplications of contributions arising when $\mathbf{k}$ swept over the subset of largest E_k's. This condition proved essential in order to obtain good estimates of the seminvariants. As regards (I.44) and (I.45), those quartets which, singly considered, estimated φ_H with a reliability smaller than a threshold value (*i.e.* $0{\cdot}35 < P^+ < 0{\cdot}80$) were excluded from the calculations. In this way, φ_H was not determined from too unreliable quartets. However, in some cases no quartet with the required qualifications were available for a given φ_H.

In order to reduce the number of these cases the threshold for $|E_h|$ can be suitably lowered. In Table 1 the chosen values of the threshold for the various structures are shown in the column labelled $E_{h\min}$. In spite of these precautions, for a given E_h no quartet may, on occasion, be available. Then in (I.44) and (I.45) only the contribution arising from the $\sum_1$ relationship and corresponding to that $|E_h|$ is used.

In Tables 5 and 6 the results of our calculations are described in some detail for AZET and KARLE respectively. The column labelled NQ gives the number of quartets used in (I.44) and (I.45). As in Tables 2 and 3, the asterisks indicate incorrect evaluations. The values of ARI for the test structures are shown in Table 7.

The following conclusions can be drawn.

(a) The estimates of one-phase seminvariants *via* the second representation are in general considerably more

* Equation (6) is the correct form of equation (I.45) which was wrongly reported in paper I.

Table 5. *AZET: indices and E's of the one-phase seminvariants of first rank together with the corresponding probabilities*

The seminvariants are arranged in descending order of expected accuracy according to (I.36).

h k l	E	$P^+(E)$ (I.14)	$P^+(E)$ (I.36)	$P^+(E)$ (I.44)	$P^+(E)$ (I.45)	NQ
0 6 0	−2·67	0·00	0·00	0·00	0·00	156
0 4 0	3·46	1·00	1·00	1·00	1·00	115
0 8 0	1·38	0·98	1·00	1·00	0·99	13
14 0 0	−1·56	0·12	0·00	0·00	0·00	45
30 0 0	1·55	0·90	0·98	0·90	0·92	4
8 4 0	2·31	0·80	0·93	0·97	0·95	3
24 4 0	1·39	0·66	0·91	0·66	0·66	0
22 6 0	1·71	0·67	0·85	0·89	0·88	2
30 2 0	−1·56	0·33	0·18	0·18	0·16	4
12 2 0	−2·36	0·35	0·26	0·35	0·35	0
8 2 0	−2·78	0·34	0·33	0·34	0·34	0
28 4 0	2·20	0·42*	0·60	0·42*	0·42*	0
12 6 0	−1·45	0·58*	0·42	0·58*	0·58*	0
24 6 0	1·61	0·48*	0·43*	0·48*	0·48*	0

* Incorrect evaluations.

Table 6. *KARLE: indices and E's of the one-phase seminvariants of first rank together with the corresponding probabilities*

The seminvariants are arranged in descending order of expected accuracy according to (I.36).

h k l	E	$P^+(E)$ (I.14)	$P^+(E)$ (I.36)	$P^+(E)$ (I.44)	$P^+(E)$ (I.45)	NQ
0 4 14	1·95	0·82	0·99	0·94	0·95	1
4 2 0	−2·54	0·04	0·01	0·00	0·00	55
4 0 12	−1·53	0·36	0·07	0·09	0·08	4
0 4 12	−1·46	0·27	0·09	0·22	0·21	1
4 0 0	1·81	0·71	0·91	0·71	0·71	0
0 4 4	−1·51	0·51*	0·11	0·22	0·19	3
0 6 18	−1·82	0·24	0·23	0·24	0·24	0
0 2 8	−2·05	0·69*	0·68*	0·69*	0·69*	0
4 0 4	−1·87	0·47	0·39	0·47	0·47	0
0 4 8	2·22	0·65	0·60	0·65	0·65	0
0 6 16	1·76	0·54	0·59	0·54	0·54	0
0 0 12	−2·42	0·44	0·44	0·44	0·44	0

* Incorrect evaluations.

Table 7. *The values of ARI for* (I.14), (I.36), (I.44) *and* (I.45)

	(I.14)	(I.36)	(I.44)	(I.45)
TETRA	0·05	0·07	0·06	0·12
RIBO	0·20	0·41	0·43	0·41
METHOX	0·07	0·17	0·12	0·12
TOLY	0·23	0·30	0·45	0·45
TOXE	0·37	0·52	0·49	0·47
KARLE	0·28	0·51	0·41	0·42
AZET	0·41	0·64	0·55	0·56

accurate than the corresponding estimates *via* the first representation. For example, in Table 5, $\varphi_{28,4,0}$ and $\varphi_{12,6,0}$ are wrongly estimated by the $\sum_1$ relationship while they are correctly estimated *via* their second representations. Furthermore, in Table 6, φ_{044} was erroneously estimated as 2π with very small reliability by the $\sum_1$ relationship while its value is correctly indicated to be π with high probability by the second representation. Similar cases can be found for the other structures tested in this paper, *e.g.* in RIBO the phase $\varphi_{20\bar{6}}$ is wrongly estimated as π by the $\sum_1$ relationship with sign probability $P^+ = 0{\cdot}30$ (the second value in order of reliability) and is correctly estimated as 2π by the second representation ($P^+ = 0{\cdot}52$). In the same structure $\varphi_{18,0,\bar{10}}$ is wrongly estimated as π ($P^+ = 0{\cdot}47$) by the $\sum_1$ relationship while it is correctly estimated as 2π by the second representation ($P^+ = 0{\cdot}86$).

(*b*) The number of seminvariants estimated with high probability value is larger than that of the older $\sum_1$ estimate. As a rule, phases associated with high probability are correctly estimated. In some cases the estimates are so reliable that one-phase seminvariants can be used in an active way from the first stages of the direct procedure. The availability of some reliably estimated one-phase seminvariants at the beginning of a multisolution procedure speeds up the phase-determination process. Furthermore, the need to introduce a large number of variable phases in the starting set is reduced, and consequently it reduces the number of trial sets of phases to be investigated to find the correct solution.

(*c*) The ARI values corresponding to (I.36), (I.44) and (I.45) are always higher than those corresponding to (I.14). This is not a mathematical result only, but has also a physical meaning because in our tests the second representation shows a strong tendency to correct the wrong estimates of the $\sum_1$ relationships. Thus the use of ARI for selecting the correct $\sum_2$ solution when the one-phase seminvariants are calculated *via* the second representation, seems able to improve analogous criteria based on $\sum_1$ relationships.

(*d*) Table 7 clearly shows for TETRA the negligible improvement in phase estimation of the second representation with respect to the first one. Further tests on other $P\bar{1}$ structures [$C_{30}H_{37}NO_5$, $Z = 2$ (Shakked & Kennard, 1977); $C_{18}H_{12}N_5O_6$, $Z = 4$ (Kiers, de Boer, Olthof & Spek, 1976) show a similar trend. A careful examination of our results leads us to conclude that, as for the first, the use of the second representation seems the more effective the more the symmetry increases. This seems to hold for the method of complementary invariants too.

(*e*) The computer-time requirement for estimating useful seminvariant phases *via* their second representation is quite modest: typical times are 20–30 s per structure by an IBM 370/158. Relations (I.44) or (I.45) are less time consuming; however, their use is not

recommended in practice. In fact, unlike representation, the complementary invariant method is too sensitive to changes in the parameters $E_{h\min}$, $E_{k\min}$, *etc.* If the values of these parameters are too high, no quartet may, on occasion, be available for a given φ_H; if they are too low, too many unreliable quartets may be calculated with waste of computer time.

It is anticipated that the estimation of one-phase structure seminvariants *via* the second representation will play an important role in a new procedure which exploits also the information contained in the two-phase structure seminvariants (Burla, Giacovazzo, Nunzi & Polidori, 1980). It is expected that a similar role may be played by a procedure using the three-phase structure seminvariants as well (Giacovazzo, 1978*b*, Hauptman & Potter, 1979).

References

BRUFANI, M., CELLAI, L., CERRINI, S., FEDELI, W. & VACIAGO, A. (1978). *Mol. Pharmacol.* **16**, 693–703.

BURLA, M. C., GIACOVAZZO, C., NUNZI, A. & POLIDORI, G. (1980). In preparation.

BUSETTA, B., GIACOVAZZO, C., BURLA, M. C., NUNZI, A., POLIDORI, G. & VITERBO, D. (1980). *Acta Cryst.* A**36**, 68–74.

CERRINI, S., FEDELI, W., GAVUZZO, E. & MAZZA, F. (1975). *Gazz. Chim. Ital.* **105**, 651–655.

COCHRAN, W. & WOOLFSON, M. M. (1954). *Acta Cryst.* **7**, 450–451.

COCHRAN, W. & WOOLFSON, M. M. (1955). *Acta Cryst.* **8**, 1–12.

COLENS, A., DECLERCQ, J. P., GERMAIN, G., PUTZEYS, J. P. & VAN MEERSSCHE, M. (1974). *Cryst. Struct. Commun.* **3**, 119–122.

GERMAIN, G., MAIN, P. & WOOLFSON, M. M. (1971) *Acta Cryst.* A**27**, 368–376.

GIACOVAZZO, C. (1975). *Acta Cryst.* A**31**, 602–609.

GIACOVAZZO, C. (1976). *Acta Cryst.* A**32**, 958–967.

GIACOVAZZO, C. (1977). *Acta Cryst.* A**33**, 933–944.

GIACOVAZZO, C. (1978*a*). *Acta Cryst.* A**34**, 562–574.

GIACOVAZZO, C. (1978*b*). *Acta Cryst.* A**34**, 27–30.

HANSON, J. C. & NORDMAN, C. E. (1975). *Acta Cryst.* B**31**, 493–501.

HAUPTMAN, H. (1972). *Crystal Structure Determination: The Role of the Cosine–Seminvariants.* NY/London: Plenum Press.

HAUPTMAN, H. & KARLE, J. (1953). *Solution of the Phase Problem. 1. The Centrosymmetric Crystal.* ACA Monograph N.3. Pittsburgh: Polycrystal Book Service.

HAUPTMAN, H. & KARLE, J. (1957). *Acta Cryst.* **10**, 267–270.

HAUPTMAN, H. & POTTER, S. (1979). *Acta Cryst.* A**35**, 371–381.

JAMES, V. J. & STEVENS, J. D. (1977). *Cryst. Struct. Commun.* **6**, 241–246.

KARLE, I. L., KARLE, J. & ESTLIN, J. A. (1967). *Acta Cryst.* **23**, 494–500.

NAYA, S., NITTA, I. & ODA, T. (1964). *Acta Cryst.* **17**, 421–433.

KIERS, C. TH., DE BOER, J. L., OLTHOF, R. & SPEK, A. L. (1976). *Acta Cryst.* B**32**, 2297–2305.

OVERBEEK, A. R. & SCHENK, H. (1976). *Proc. K. Ned. Akad. Wet.* **B79**, 341–343.

SHAKKED, Z. & KENNARD, O. (1977). *Acta Cryst.* B**33**, 516–522.

SPAGNA, R. & VACIAGO, A. (1978). *Acta Cryst.* B**34**, 993–995.

WEEKS, C. M. & HAUPTMAN, H. (1970). *Z. Kristallogr.* **131**, 437–442.

Acta Cryst. (1980). A**36**, 578–581

The Influence of Twinning by Merohedry on Intensity Statistics

BY DOUGLAS C. REES

Gibbs Chemical Laboratory, Harvard University, 12 Oxford Street, Cambridge, MA 02138, USA

(*Received 7 August* 1979; *accepted 7 January* 1980)

Abstract

A simple test based on intensity statistics is presented for the detection of twinning by merohedry. Using relationships derived in the text, the twinning fraction of a crystal may be estimated from the intensity probability distribution. Unlike most methods for the detection of twinning, application of this test does not require knowledge of the twinning operation. Two possible mechanisms for increasing the apparent diffraction symmetry of a crystal, twinning by merohedry and crystal disorder, may be distinguished in certain cases by these procedures.

Crystals twinned by merohedry present special problems in X-ray crystal-structure determinations since the reciprocal lattices of the twins have identical orientations (Buerger, 1960). This class of twinning may occur in space groups of tetragonal or higher symmetry whenever the point symmetry of the crystal is lower than that of the lattice (Catti & Ferraris, 1976). Since the twinning operation exactly superimposes non-

0567-7394/80/040578-04$01.00