Acta Cryst. (1980). A**36**, 891–897

On the Identity of the Enantiomorph-Sensitive Three-Phase Structure Seminvariants and Variants in $P2_1$ and Their Use in Procedures for Phase Determination

BY NIEK VAN DER PUTTEN AND HENK SCHENK

Laboratory for Crystallography, University of Amsterdam, Nieuwe Achtergracht 166, 1018 *WS Amsterdam, The Netherlands*

AND HERBERT HAUPTMAN

Medical Foundation of Buffalo Inc., 73 *High Street, Buffalo, New York* 14203, *USA*

(*Received* 1 *October* 1979; *accepted* 28 *April* 1980)

Abstract

The identity of the enantiomorph-sensitive three-phase structure seminvariants, $T = \varphi_{h_1k_1l_1} + \varphi_{h_2k_2l_2} + \varphi_{h_3k_3l_3}$, in space group $P2_1$ is derived *via* the sextet extension. It is shown, in addition, that the extension concept is applicable also to all three-phase sums (variants) $V = \varphi_{h_1k_1l_1} + \varphi_{h_2k_2l_2} + \varphi_{h_3k_3l_3}$, provided only that $k_1 + k_2 + k_3 = 0$. The T and V estimating procedures have been tested and are equally reliable, as expected. More importantly, it appears from the initial applications that the number of phase sums T and V, reliably estimated to be $\pm\pi/2$, is often quite large, and that they relate the phases of the largest E's. In the cases that there are many reliable such indications, these relations are very useful for the discrimination between the centrosymmetric and enantiomorph-sensitive phases preceding the final phase-determination step. Thus enantiomorph specification, a perennial problem in this space group, is facilitated.

1. Introduction

The problem of the definition and maintenance of the enantiomorph in space groups like $P2_1$ is often a difficult one to resolve in applications of direct methods. The first attempt to fix the enantiomorph definitively (strong enantiomorph discrimination) was made in 1972 (Hauptman & Duax, 1972) where, in effect, it was shown how to identify reliably a class of enantiomorph-sensitive (*i.e.* having the value $\pm\pi/2$) linear combinations of two phases *via* calculated values of appropriate three-phase structure invariants, in particular those having the value π. The initial applications of this procedure were made to the structure determination of aldosterone monohydrate and valinomycin (Duax & Hauptman, 1972*a*,*b*), structures for which enantiomorph specification had been difficult. With the development of the probabilistic theory of the higher-order structure invariants, in particular the means to identify reliably the so-called negative quartets, and the introduction of the extension concept (Hauptman, 1978; but see also Giacovazzo, 1977, for a related concept called representation theory) a new method emerged for identifying the enantiomorph-sensitive two-phase structure seminvariants by embedding the latter in suitable quartets (Hauptman & Green, 1978). Busetta (1976) was the first to demonstrate the usefulness of this technique in his solution of alborixin ($C_{48}H_{84}O_{14}^{-}.K^{+}$), a structure for which enantiomorph discrimination was particularly difficult.

Despite the successful applications of the enantiomorph-sensitive two-phase structure seminvariants, it is clearly important to develop methods for reliably identifying still larger classes of enantiomorph-sensitive structure seminvariants. In the present paper we shall show how to identify relatively large numbers of linear combinations of three phases in $P2_1$ (variants)

$$V = \varphi_{h_1k_1l_1} + \varphi_{h_2k_2l_2} + \varphi_{h_3k_3l_3}, \tag{1.1}$$

satisfying

$$k_1 + k_2 + k_3 = 0 \tag{1.2}$$

and having the value $\pm\pi/2$. If, in addition to (1.2), the conditions

$$h_1 + h_2 + h_3 \equiv l_1 + l_2 + l_3 \equiv 0 \pmod 2 \tag{1.3}$$

(*i.e.* $h_1 + h_2 + h_3$ and $l_1 + l_2 + l_3$ are even) are also satisfied, then (1.1) is a structure seminvariant (called T in the sequel).

2. The enantiomorph-sensitive three-phase structure seminvariants in $P2_1$

The probabilistic theory of the three-phase structure seminvariant T in $P2_1$ was recently initiated (Hauptman, 1978; Hauptman & Potter, 1979). By

0567-7394/80/060891-07$01.00

embedding T and its symmetry-related variants in suitable quintets Q to which T is related *via* the space-group-dependent relationships among the phases, one obtains the extensions Q of T. In this way the probabilistic theory of the structure seminvariant T is reduced to that of the structure invariants Q, which is well developed. In particular, the neighborhoods of T are defined in terms of the neighborhoods of the Q's. Thus the first neighborhood of T is the set-theoretic union of the first neighborhoods of all (four) of its extensions (seven magnitudes $|E|$ in all); the second neighborhood of T is the set-theoretic union of the second neighborhoods of all of its extensions (forty-one magnitudes $|E|$ in all); *etc.* As described in this earlier work, T has many first neighborhoods, many second neighborhoods, *etc.* By employing the quintet extensions Q of T, the theory led to methods for identifying and estimating reliably certain of the three-phase structure seminvariants T having the value 0 or π (the enantiomorph-insensitive ones).

In the present paper one employs the (unique) sextet extension S of T and is thus led to a method for identifying those seminvariants T having the value $\pm\pi/2$, *i.e.* the enantiomorph-sensitive ones. Although the method is also capable of identifying those seminvariants T having the value 0 or π, it is not able to make a unique estimate in this way (but the method of the preceding paragraph does have this capability).

2.1. *The extension*

The linear combination of three phases

$$T = \varphi_{h_1 k_1 l_1} + \varphi_{h_2 k_2 l_2} + \varphi_{h_3 k_3 l_3} \tag{2.1}$$

is a structure seminvariant in $P2_1$ if and only if (1.2) and (1.3) are satisfied.

Following methods previously described (Hauptman, 1978; Hauptman & Green, 1978), the structure seminvariant T is embedded in the six-phase structure invariant (sextet) S, called an extension of T, by means of

$$S = \varphi_{h_1 k_1 l_1} + \varphi_{h_2 k_2 l_2} + \varphi_{h_3 k_3 l_3} + \varphi_{\bar{h}_1 k_1 \bar{l}_1} + \varphi_{\bar{h}_2 k_2 \bar{l}_2} + \varphi_{\bar{h}_3 k_3 \bar{l}_3} \tag{2.2}$$

so that, because of (1.2) and the space-group-dependent relationships among the phases,

$$S = 2T. \tag{2.3}$$

Hence, if

$$S \simeq \pi, \tag{2.4}$$

then

$$T \simeq \pm\pi/2. \tag{2.5}$$

2.2. *The neighborhoods of T*

The second sequence [to distinguish it from the sequence defined earlier (Hauptman, 1978)] of nested neighborhoods of the structure seminvariant T is defined to be the sequence of nested neighborhoods of the special sextet S (Fortier & Hauptman, 1977). Thus the first neighborhood of the second kind of T is defined to consist of the three magnitudes

$$R_i = |E_{h_i k_i l_i}|, \quad i = 1, 2, 3. \tag{2.6}$$

Again, the second neighborhood of T of the second kind consists of 19 magnitudes $|E|$, the three 'main terms' (2.6) and the additional 16 'cross terms':

$$\begin{aligned}
&R_{12} = |E_{h_1+h_2, k_3, l_1+l_2}|, \quad R_{23} = |E_{h_2+h_3, k_1, l_2+l_3}|,\\
&R_{31} = |E_{h_3+h_1, k_2, l_3+l_1}|;\\
&R_{1\bar{2}} = |E_{h_1-h_2, k_3, l_1-l_2}|, \quad R_{2\bar{3}} = |E_{h_2-h_3, k_1, l_2-l_3}|,\\
&R_{3\bar{1}} = |E_{h_3-h_1, k_2, l_3-l_1}|;\\
&R_{0i} = |E_{02k_i0}|, \quad i = 1, 2, 3;
\end{aligned} \tag{2.7}$$

$$\begin{aligned}
R_{123} &= |E_{h_1+h_2+h_3, 0, l_1+l_2+l_3}|,\\
R_{12\bar{3}} &= |E_{h_1+h_2-h_3, 0, l_1+l_2-l_3}|,\\
R_{1\bar{2}3} &= |E_{h_1-h_2+h_3, 0, l_1-l_2+l_3}|,\\
R_{\bar{1}23} &= |E_{-h_1+h_2+h_3, 0, -l_1+l_2+l_3}|;\\
R_{1\bar{2}\bar{3}} &= |E_{h_1, k_2-k_3, l_1}|,\\
R_{\bar{1}2\bar{3}} &= |E_{h_2, k_3-k_1, l_2}|,\\
R_{\bar{1}\bar{2}3} &= |E_{h_3, k_1-k_2, l_3}|.
\end{aligned} \tag{2.8}$$

2.3. *The discriminant of T*

Following Hauptman & Fortier (1977), the discriminant Δ of T of the second kind is defined by

$$\begin{aligned}
\Delta = (2/\sigma_2^6)(R_1^2 - 1)(R_2^2 - 1)(R_3^2 - 1)\Big\{ &\sigma_3^4 \Big[\sum_{15} + \sum_{90}\Big]\\
&- \sigma_3^2(3\sigma_3^2 - \sigma_2\sigma_4)\Big[\sum_{45} + 2\sum_{30}\Big]\\
&+ \Big[\sigma_3(15\sigma_3^3 - 10\sigma_2\sigma_3\sigma_4 + \sigma_2^2\sigma_5)\sum_{15}'\\
&+ (3\sigma_3^2 - \sigma_2\sigma_4)^2 \sum_{10}\Big] - \Big[105\sigma_3^4 - 105\sigma_2\sigma_3^2\sigma_4\\
&+ 15\sigma_2^2\sigma_3\sigma_5 + 10\sigma_2^2\sigma_2^4 - \sigma_2^3\sigma_6\Big]\Big\},
\end{aligned} \tag{2.9}$$

where

$$\begin{aligned}
\sum_{15} = {} & R_{01}^2 R_{02}^2 R_{03}^2 + R_{01}^2(R_{23}^4 + R_{2\bar{3}}^4) + R_{02}^2(R_{31}^4 + R_{3\bar{1}}^4)\\
&+ R_{03}^2(R_{12}^4 + R_{1\bar{2}}^4) + 2R_{12}^2 R_{23}^2 R_{3\bar{1}}^2\\
&+ 2R_{1\bar{2}}^2 R_{2\bar{3}}^2 R_{31}^2 + 2R_{1\bar{2}}^2 R_{23}^2 R_{3\bar{1}}^2\\
&+ 2R_{12}^2 R_{2\bar{3}}^2 R_{3\bar{1}}^2;
\end{aligned} \tag{2.10}$$

$$\sum_{90} = R_{123}^2(R_{12}^2 + R_{23}^2 + R_{31}^2)^2 + R_{\bar{1}23}^2(R_{\bar{1}2}^2 + R_{23}^2 + R_{3\bar{1}}^2)^2 + R_{1\bar{2}3}^2(R_{1\bar{2}}^2 + R_{2\bar{3}}^2 + R_{31}^2)^2 + R_{12\bar{3}}^2(R_{12}^2 + R_{2\bar{3}}^2 + R_{3\bar{1}}^2)^2 + 2R_{12\bar{3}}(R_{02}^2 + R_{12}^2 + R_{1\bar{2}}^2) \times (R_{03}^2 + R_{31}^2 + R_{3\bar{1}}^2) + 2R_{\bar{1}2\bar{3}}^2(R_{01}^2 + R_{12}^2 + R_{1\bar{2}}^2) \times (R_{03}^2 + R_{23}^2 + R_{2\bar{3}}^2) + 2R_{\bar{1}\bar{2}3}^2(R_{01}^2 + R_{31}^2 + R_{3\bar{1}}^2) \times (R_{02}^2 + R_{23}^2 + R_{2\bar{3}}^2); \quad (2.11)$$

$$\sum_{45} = R_{01}^2 R_{02}^2 + R_{02}^2 R_{03}^2 + R_{03}^2 R_{01}^0 + 2R_{01}^2(R_{23}^2 + R_{2\bar{3}}^2) + 2R_{02}^2(R_{31}^2 + R_{3\bar{1}}^2) + 2R_{03}^2(R_{12}^2 + R_{1\bar{2}}^2) + R_{12}^4 + R_{23}^4 + R_{31}^4 + R_{1\bar{2}}^4 + R_{2\bar{3}}^4 + R_{3\bar{1}}^4 + 2R_{12}^2 R_{23}^2 + 2R_{23}^2 R_{31}^2 + 2R_{31}^2 R_{12}^2 + 2R_{12}^2 R_{2\bar{3}}^2 + 2R_{2\bar{3}}^2 R_{3\bar{1}}^2 + 2R_{31}^2 R_{1\bar{2}}^2 + 2R_{1\bar{2}}^2 R_{23}^2 + 2R_{23}^2 R_{31}^2 + 2R_{31}^2 R_{12}^2 + 2R_{12}^2 R_{2\bar{3}}^2 + 2R_{2\bar{3}}^2 R_{3\bar{1}}^2 + 2R_{3\bar{1}}^2 R_{1\bar{2}}^2; \quad (2.12)$$

$$\sum_{30} = R_{123}^2(R_{12}^2 + R_{23}^2 + R_{31}^2) + R_{\bar{1}23}^2(R_{1\bar{2}}^2 + R_{23}^2 + R_{3\bar{1}}^2) + R_{1\bar{2}3}^2(R_{1\bar{2}}^2 + R_{2\bar{3}}^2 + R_{31}^2) + R_{12\bar{3}}^2(R_{12}^2 + R_{2\bar{3}}^2 + R_{3\bar{1}}^2) + R_{\bar{1}\bar{2}3}^2(R_{02}^2 + R_{03}^2 + R_{31}^2 + R_{12}^2 + R_{3\bar{1}}^2 + R_{1\bar{2}}^2) + R_{\bar{1}2\bar{3}}^2(R_{03}^2 + R_{01}^2 + R_{12}^2 + R_{23}^2 + R_{1\bar{2}}^2 + R_{2\bar{3}}^2) + R_{1\bar{2}\bar{3}}^2(R_{01}^2 + R_{02}^2 + R_{23}^2 + R_{31}^2 + R_{2\bar{3}}^2 + R_{3\bar{1}}^2); \quad (2.13)$$

$$\sum_{15}' = R_{01}^2 + R_{02}^2 + R_{03}^2 + 2(R_{12}^2 + R_{23}^2 + R_{31}^2 + R_{1\bar{2}}^2 + R_{2\bar{3}}^2 + R_{3\bar{1}}^2); \quad (2.14)$$

$$\sum_{10} = R_{123}^2 + R_{\bar{1}23}^2 + R_{1\bar{2}3}^2 + R_{12\bar{3}}^2 + 2(R_{1\bar{2}\bar{3}}^2 + R_{\bar{1}2\bar{3}}^2 + R_{\bar{1}\bar{2}3}^2); \quad (2.15)$$

and

$$\sigma_n = \sum_{j=1}^{N} f_j^n, \quad (2.16)$$

where f_j is the zero-angle atomic scattering factor of the jth atom and N is the number of atoms in the whole unit cell. In the X-ray diffraction case the f_j are equal to the atomic numbers Z_j and are therefore all positive; in the neutron diffraction case some of the f_j may be negative.

In the applications one lists the discriminants Δ in ascending algebraic order, the most negative ones appearing first. It is known (Hauptman & Fortier, 1977) that the values of the discriminant are correlated with the values of the sextet S in the sense that $S \simeq 0$ or $S \simeq \pi$ according as $\Delta \gg 0$ or $\Delta \ll 0$ respectively. In view of (2.4) and (2.5) it follows that the structure seminvariant $T \simeq \pm\pi/2$ when $\Delta \ll 0$. [When $\Delta \gg 0$, $T \simeq 0$ or π, in view of (2.3).]

In calculating the discriminant Δ the full complement of 16 cross terms (2.7) and (2.8) will, in general, not be available; in such cases the missing cross terms, provided there are not more than six of them, are replaced by unity, the average value of $|E|^2$.

3. Three-phase variants by sextet extensions

The formalism described in the preceding paragraph can be applied more generally to all variants

$$V = \varphi_{h_1k_1l_1} + \varphi_{h_2k_2l_2} + \varphi_{h_3k_3l_3} \quad (3.1)$$

with $k_1 + k_2 + k_3 = 0$ and no condition for the other indices. The sextet structure invariant in which (3.1) is embedded is again

$$S = \varphi_{h_1k_1l} + \varphi_{h_2k_2l_2} + \varphi_{h_3k_3l_2} + \varphi_{h_1k_1l_1} + \varphi_{h_2k_2l_2} + \varphi_{h_3k_3l_3}. \quad (3.2)$$

In this case too, it follows from the space-group symmetries that $S = 2V$ [compare with (2.3)]. Thus V is estimated to be $\pm\pi/2$ if $\Delta \ll 0$ and 0 or π if $\Delta \gg 0$. It should be noted that the three-phase seminvariants T form a subset of the set V. The advantage of using the estimate $V = \pm\pi/2$ by means of (2.9) instead of only the estimate $T = \pm\pi/2$ is that approximately four times as many phase relationships become available for the applications. It should be noted finally that choosing arbitrarily the sign of a single enantiomorph-sensitive structure seminvariant $T = \pm\pi/2$ is equivalent to specifying the enantiomorph.

4. Tests of T and V estimates

The T and V estimating procedures using (2.9) have been applied to the following four structures.

(1) Diethylmalonic acid (DIEMAL) (van der Putten, 1979), $C_7H_{12}O_4$, $Z = 4$, $N = 44$, $P2_1$.

(2) Tribenzamide (TRIBEN) (Olthof, 1979), $C_{21}H_{15}NO_3$, $Z = 2$, $N = 50$, $P2_1$.

(3) Aldosterone monohydrate (ALDO) (Duax & Hauptman, 1972*a*), $C_{21}H_{28}O_5.H_2O$, $Z = 2$, $N = 54$, $P2_1$.

(4) Valinomycin (VALI) (Smith, Duax, Langs, DeTitta, Edmonds, Rohrer & Weeks, 1975), $C_{54}H_{90}N_6O_{18}$, $Z = 2$, $N = 156$, $P2_1$.

In order to be able to judge whether the T and V estimates are equally reliable all variants $V = \varphi_{h_1k_1l_1} + \varphi_{h_2k_2l_2} + \varphi_{h_3k_3l_3}$ were divided into four groups in accordance with the scheme:

$$\text{group } A: h_1 + h_2 + h_3 \equiv 0 \pmod 2, \quad k_1 + k_2 + k_3 = 0, \quad l_1 + l_2 + l_3 \equiv 0 \pmod 2; \quad (4.1)$$

$$\text{group } B: h_1 + h_2 + h_3 \equiv 0 \pmod 2$$
$$k_1 + k_2 + k_3 = 0$$
$$l_1 + l_2 + l_3 \equiv 1 \pmod 2; \qquad (4.2)$$

$$\text{group } C: h_1 + h_2 + h_3 \equiv 1 \pmod 2$$
$$k_1 + k_2 + k_3 = 0$$
$$l_1 + l_2 + l_3 \equiv 0 \pmod 2; \qquad (4.3)$$

$$\text{group } D: h_1 + h_2 + h_3 \equiv 1 \pmod 2$$
$$k_1 + k_2 + k_3 = 0$$
$$l_1 + l_2 + l_3 \equiv 1 \pmod 2. \qquad (4.4)$$

It is obvious that group A consists of the three-phase structure seminvariants T. For all test structures the strongest $|E|$ values were used to generate in all groups the three-phase variants V with the largest values of

$$E_3 = N^{-1/2} |E_{h_1k_1l_1} E_{h_2k_2l_2} E_{h_3k_3l_3}|. \qquad (4.5)$$

Then the sextet discriminants Δ [equation (2.9)] were calculated and within the different groups the variants were sorted in order of increasing value of Δ, the negative ones appearing first. Since $\Delta \ll 0$ corresponds to $S = \pi$ and $\Delta \gg 0$ to $S = 0$, the top of the list comprises the 'enantiomorph-sensitive' variants, $V = \pm\pi/2$ and the bottom the 'enantiomorph-insensitive' ones, $V = 0$ or π.

4.1. *DIEMAL*

The strongest 200 $|E|$ values were used to generate approximately 42 000 three-phase variants V per group with an E_3 lower limit of 0·7. Summaries of the estimates of $V = \pm\pi/2$ and $V = 0$ or π are given in Tables 1 and 2 respectively. It can be seen from the tables that the reliability of the V estimates in all groups is approximately equal. There is no significant difference between the results for the three-phase seminvariants (group A) and the three-phase variants (groups B, C and D). In Table 1, column 5, the Δ values range from −121·5 to −17·2 for all variants $V \simeq \pm\pi/2$ with E_3 values > 1·5.

In the above experiment about 168 000 variants were calculated in all, which is quite a job, even for the modern computer. Therefore the test was repeated with approximately 23 000 variants having $E_3 > 1{\cdot}5$, and similar results were obtained; the only difference was that the number of $V \simeq 0, \pi$ estimates was smaller. From these and other experiments it is expected that normally the calculation of the 20 000–50 000 strongest variants suffices to get all strong Δ indications for $V = \pm\pi/2$ and the strongest Δ indications for $V = 0, \pi$.

It should be noted that the top and the bottom of the discriminant list give estimates for those variants which relate the phases of the strongest reflections. In particular, this applies to the variants with $E_3 > 1{\cdot}5$ and $E_3 > 3{\cdot}0$ (columns 5 and 6 in Table 1). The average errors of these variants are comparable to those of the $\sum_2$ relationships with $E_3 > 1{\cdot}5$ and $E_3 > 3{\cdot}0$ respectively [44 millicycles for 316 $\sum_2$'s and 23 millicycles for 22 $\sum_2$'s (1000 millicycles = 2 π rad)].

Table 2. *Average error for DIEMAL of the estimates of $V = 0, \pi$ in the four groups A, B, C and D defined by* (4.1), (4.2), (4.3) *and* (4.4), *respectively*

The variants are listed in decreasing order of Δ [equation (2.9)]. The average errors ⟨DEV⟩ are given in millicycles for the number of variants (NR) with largest Δ, where NR has the different values shown.

NR	Group A ⟨DEV⟩	Group B ⟨DEV⟩	Group C ⟨DEV⟩	Group D ⟨DEV⟩
10	23	29	19	26
25	28	35	42	31
50	40	36	46	37
75	43	37	42	41
100	45	32	43	48
125	47	35	47	52
150	47	38	45	54
175	50	37	45	55

Table 1. *Average error for DIEMAL of the estimates of the three-phase variants $V = \pm\pi/2$ in the four groups A, B, C and D defined by* (4.1), (4.2), (4.3) *and* (4.4) *respectively when $E_3 > 0{\cdot}7$ and of the estimates of all variants $V = \pm\pi/2$ with $E_3 > 1{\cdot}5$ and 3·0*

The variants are listed in increasing order of the Δ values [equation (2.9)], *i.e.* starting with the most negative ones. The average errors ⟨DEV⟩ are given in millicycles for the number of variants (NR) with smallest (most negative) Δ, NR having the different values shown.

NR	Group A ⟨DEV⟩	Group B ⟨DEV⟩	Group C ⟨DEV⟩	Group D ⟨DEV⟩	All $V \simeq \pm\pi/2$ ($E_3 > 1{\cdot}5$) ⟨DEV⟩	All $V \simeq \pm\pi/2$ ($E_3 > 3{\cdot}0$) ⟨DEV⟩
5	53	23	39	39	29	29
10	32	27	49	48	28	28
20	32	27	61	56	32	30
40	37	30	59	41	43	39
60	46	29	63	52	39	

4.2. *TRIBEN*

The strongest 200 $|E|$ values were used to generate about 19 000 three-phase variants V with $E_3 > 1{\cdot}0$. Summaries of the estimates of $V \simeq \pm\pi/2$ are given in Table 3. The Δ values of the variants used in column 5 range from $-3{\cdot}6$ to $-1{\cdot}5$. For this structure, which is difficult to solve by the usual methods, the most reliable estimates of $V \simeq \pm\pi/2$ with $E_3 > 1{\cdot}5$ are of comparable quality to those of the 38 triplets with $E_3 > 1{\cdot}5$ (average error 64 millicycles). The small number of variants is due to the small maxima for h and l and the large one for k.

4.3. *ALDO*

The strongest 200 $|E|$ values were used to generate about 21 000 three-phase variants V with $E_3 > 1{\cdot}5$. Summaries of the estimates of $V \simeq \pm\pi/2$ are given in Table 4. In this table the Δ values range from $-12{\cdot}0$ to $-2{\cdot}0$ for the overall group (column 5). Probably because of the large number of restricted phases for the strongest reflections the number and goodness of the estimates of $V \simeq \pm\pi/2$ were relatively poor.

4.4. *VALI*

About 50 000 three-phase variants with $E_3 > 1{\cdot}5$ were generated using the strongest 500 $|E|$ values. Summaries of the $V \simeq \pm\pi/2$ estimates are given in Table 5. The Δ values range from $-13{\cdot}4$ to $-3{\cdot}1$. Again the errors of the most reliable estimates of $V \simeq \pm\pi/2$ with $E_3 > 2{\cdot}0$ and $1{\cdot}5$ are comparable to those of 34 and 51 millicycles for the 34 and 190 triplets with E_3 values $> 2{\cdot}0$ and $1{\cdot}5$ respectively.

Table 3. *Average error for TRIBEN of the estimates of the variants* $V = \pm\pi/2$ *in the four groups A, B, C and D defined by* (4.1), (4.2), (4.3) *and* (4.4) *respectively with* $E_3 > 1{\cdot}0$ *and of the estimates of all variants* $V = \pm\pi/2$ *with* $E_3 > 1{\cdot}0$ *and* $1{\cdot}5$

The variants are listed in increasing order of their Δ values [equation (2.9)]. The average errors $\langle$DEV$\rangle$ are given in millicycles for the number of variants (NR) with smallest (most negative) Δ, NR having the different values shown.

NR	Group *A* ⟨DEV⟩	Group *B* ⟨DEV⟩	Group *C* ⟨DEV⟩	Group *D* ⟨DEV⟩	All $V \simeq \pm\pi/2$ ($E_3 > 1{\cdot}0$) ⟨DEV⟩	All $V \simeq \pm\pi/2$ ($E_3 > 1{\cdot}5$) ⟨DEV⟩
5	49	145	113	81	129	77
10	90	132	94	82	106	88
20	108	129	97	93	108	88
40	117	124	85	95	93	113
60					104	

Table 4. *Average error for ALDO of the estimates of the variants* $V \simeq \pm\pi/2$ *in the four groups A, B, C and D with* $E_3 > 1{\cdot}5$ *and of the estimates of all variants* $V \simeq \pm\pi/2$ *with* $E_3 > 1{\cdot}5$ *and* $2{\cdot}5$

The variants are listed in increasing order of their Δ values [equations (2.9)]. The average errors $\langle$DEV$\rangle$ are given in millicycles for the number of variants (NR) with smallest (most negative) Δ, NR having the different values shown.

NR	Group *A* ⟨DEV⟩	Group *B* ⟨DEV⟩	Group *C* ⟨DEV⟩	Group *D* ⟨DEV⟩	All $V \simeq \pm\pi/2$ ($E_3 > 1{\cdot}5$) ⟨DEV⟩	All $V \simeq \pm\pi/2$ ($E_3 > 2{\cdot}5$) ⟨DEV⟩
5	47	139	95	84	106	107
10	88	134	83	104	98	95
20	98	123	68	91	102	80
40	101	117	91	104	91	
60	106			104	103	

Table 5. *Average error for VALI of the estimates of the variants* $V \simeq \pm\pi/2$ *in the four groups A, B, C and D with* $E_3 > 1{\cdot}5$ *and of the estimates of all variants* $V \simeq \pm\pi/2$ *with* $E_3 > 1{\cdot}5$ *and* $2{\cdot}0$

The variants are listed in increasing order of their Δ value [equation (2.9)]. The average errors $\langle$DEV$\rangle$ are given in millicycles for the number of variants (NR) with smallest (most negative) Δ, NR having the different values shown.

NR	Group *A* ⟨DEV⟩	Group *B* ⟨DEV⟩	Group *C* ⟨DEV⟩	Group *D* ⟨DEV⟩	All $V \simeq \pm\pi/2$ ($E_3 > 1{\cdot}5$) ⟨DEV⟩	All $V \simeq \pm\pi/2$ ($E_3 > 2{\cdot}0$) ⟨DEV⟩
5	51	105	159	38	67	30
10	75	125	90	67	46	41
20	73	88	77	77	78	73
40	75	94	68	85	90	69
60	89	96	82	95	88	75

5. Possible applications of the estimates of $V \simeq \pm\pi/2$

In our opinion the $V \simeq \pm\pi/2$ estimates may be applied in three different ways.

(1) Preceding the symbolic addition or multi-solution procedure the $V \simeq \pm\pi/2$ estimates can be used together with the strongest triplets, the $V \simeq 0, \pi$ estimates and the $T \simeq 0$ and $T \simeq \pi$ estimates (Hauptman & Potter, 1979) for the discrimination between centrosymmetric and enantiomorph-sensitive phases. This will facilitate the final phasing procedure enormously. An example will be given in the next section for the structure of DIEMAL.

(2) The $V \simeq \pm\pi/2$ and $T \simeq 0$ or π estimates may be used in an enantiomorph-specific figure of merit (see van der Putten, Schenk & Hauptman, 1980).

(3) In the stage of numerical refinements there are two possible applications. Firstly, preceding tangent refinement, a modulo π phase refinement is carried out using the variants $V \simeq \pm\pi/2$, the variants $V \simeq 0, \pi$ and the seminvariants $T \simeq 0$ or π:

$$\tan\alpha_{\mathbf{h}} \simeq \frac{\sum_{\mathbf{k}} W \sin\{[\varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} - (V, T, \varphi)] \bmod \pi\}}{\sum_{\mathbf{k}} W \cos\{[\varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} - (V, T, \varphi)] \bmod \pi\}},$$

where (V,T,φ) means the estimated value (0, π, or $\pm\pi/2$) of a structure seminvariant (or variant) $\varphi_{\mathbf{h}_1} + \varphi_{\mathbf{h}_2} + \varphi_{\mathbf{h}_3}$ with $\mathbf{h}_1 + \mathbf{h}_2 + \mathbf{h}_3 = \mathbf{k}$, and all phases are limited to $0 < \alpha_{\mathbf{h}} < \pi$. This refinement divides the reflections into centrosymmetric ones ($\alpha_{\mathbf{h}} \simeq 0$) and enantiomorph-sensitive ones ($\alpha_{\mathbf{h}} \simeq \pi/2$). In the next step the $\alpha_{\mathbf{h}}$ values are used to weight the contributions of phases to the usual tangent refinement:

$$\tan\varphi_{\mathbf{h}} = \frac{\sum_{\mathbf{k}} W_{\mathbf{k}} W_{\mathbf{h}-\mathbf{k}} E_3 \sin(\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}})}{\sum_{\mathbf{k}} W_{\mathbf{k}} W_{\mathbf{h}-\mathbf{k}} E_3 \cos(\varphi_{\mathbf{k}} + \varphi_{\mathbf{h}-\mathbf{k}})}$$

with $W_{\mathbf{k}} \propto 1/[|(\varphi_{\mathbf{k}} \bmod \pi) - \alpha_{\mathbf{k}}|]$. Thus phases which do not get their expected centrosymmetric or enantiomorph-sensitive values get a low weight in the tangent refinement, and this will help to maintain the enantiomorph during refinement.

Secondly, the $V \simeq \pm\pi/2$ estimates may be used in an adapted tangent refinement, similar to the one successfully used with quartet and quintet estimates (van der Putten & Schenk, 1979) and described there in detail. Here we will confine ourselves to remarking that the $V \simeq \pm\pi/2$ estimates are mixed with the triplet invariants and seminvariants in an enantiomorph-specific extension and refinement procedure.

6. Discrimination between centrosymmetric and enantiomorph-sensitive phases

If a sufficient number of reliable $V \simeq \pm\pi/2$ estimates (*e.g.* $\Delta < -10{\cdot}0$) is available, then it is possible to use them together with the most reliable triplets, $V \simeq 0, \pi$ estimates and $T \simeq 0$ and $T \simeq \pi$ estimates for the discrimination between centrosymmetric [$\varphi \simeq 0 \pmod\pi$] and noncentrosymmetric phases [$\varphi \simeq \pi/2 \pmod\pi$] before starting the final phasing procedure. We have applied this to DIEMAL. Firstly, the origin was defined with two phase-restricted reflections and one general reflection with k index equal to 1. The strongest triplet and the second $T \simeq \pm\pi/2$ estimate assigned a reflection with large $|E|$ to be enantiomorph sensitive. Thereafter, from the 125 strongest reflections we could assign 49 reflections to be centrosymmetric and 18 reflections to be enantiomorph sensitive using the $V \simeq \pm\pi/2$ estimates with $\Delta < -10{\cdot}0$, the strongest $T \simeq 0$ or π estimates and the strongest $V \simeq 0, \pi$ estimates. A post-mortem examination showed that all assignments were correct. Using these values in our symbolic addition program *SIMPEL* (Overbeek, van der Putten, Olthof & Schenk, 1977), we got numerous indications that two of the symbolic phases in the starting set were centrosymmetric and one was enantiomorph sensitive. In this way the number of possible solutions was reduced by a factor of eight. Eventually the correct solution was found to be the one with the highest $\sum_2$ consistency.

7. Concluding remarks

The discriminant Δ [equation (2.9)] used throughout this work is based on the general sextet discriminant dependent on 31 magnitudes $|E|$. Here, because the sextet S is special, only 19 magnitudes $|E|$ are distinct. Thus a better discriminant, presumably leading to a more reliable identification of the enantiomorph-sensitive seminvariants, remains to be derived in a theoretical, more satisfactory way. This work has not yet been done because of the enormous size of the undertaking. It is nevertheless noteworthy that, in view of the applications described here [and others (van der Putten, Schenk & Hauptman, 1980)], the discriminant Δ [equation (2.9)] is good enough to identify reliably a sufficient number of enantiomorph-sensitive seminvariants to be useful in the determination of crystal structures.

It should be noted that discriminants, while useful, easily computed and well suited for identifying those structure seminvariants having the extreme values 0, π, or $\pm\pi/2$, are not the most powerful expressions for estimating the values of the structure seminvariants. The best method is to derive accurate conditional

probability distributions of the structure seminvariants, assuming as known the magnitudes $|E|$ in their neighborhoods; in the present case the 19 magnitudes $|E|$ in the second neighborhood of T. This enormous undertaking has been carried out in only a few cases so far, and a great deal of additional work in this direction remains to be done.

HH acknowledges partial support from the National Science Foundation, Grant CHE79—11282.

References

BUSETTA, B. (1976). *Acta Cryst.* A**32**, 139–143.
DUAX, W. & HAUPTMAN, H. (1972*a*). *J. Am. Chem. Soc.* **94**, 5467–5471.
DUAX, W. & HAUPTMAN, H. (1972*b*). *Acta Cryst.* B**28**, 2912–2916.
FORTIER, S. & HAUPTMAN, H. (1977). *Acta Cryst.* A**33**, 694–696.
GIACOVAZZO, C. (1977). *Acta Cryst.* A**33**, 933–944.
HAUPTMAN, H. (1978). *Acta Cryst.* A**34**, 525–528.
HAUPTMAN, H. & DUAX, W. (1972). *Acta Cryst.* A**28**, 393–395.
HAUPTMAN, H. & FORTIER, S. (1977). *Acta Cryst.* A**33**, 697–701.
HAUPTMAN, H. & GREEN, E. (1978). *Acta Cryst.* A**34**, 224–229.
HAUPTMAN, H. & POTTER, S. (1979). *Acta Cryst.* A**35**, 371–381.
OLTHOF, G. J. (1979). Unpublished.
OVERBEEK, O. (1980). Thesis, Amsterdam.
PUTTEN, N. VAN DER (1979). Unpublished.
PUTTEN, N. VAN DER & SCHENK, H. (1979). *Acta Cryst.* A**35**, 381–387.
PUTTEN, N. VAN DER, SCHENK, H. & HAUPTMAN, H. (1980). *Acta Cryst.* A**36**, 897–903.
SMITH, G. D., DUAX, W. L., LANGS, D. A., DETITTA, G. T., EDMONDS, J. W., ROHRER, D. C. & WEEKS, C. M. (1975). *J. Am. Chem. Soc.* **97**, 7242–7247.

Acta Cryst. (1980). A**36**, 897–903

Application of the Three-Phase Seminvariants to Phase Determination in $P2_1$

BY NIEK VAN DER PUTTEN AND HENK SCHENK

Laboratory for Crystallography, University of Amsterdam, Nieuwe Achtergracht 166, 1018 *WS Amsterdam, The Netherlands*

AND HERBERT HAUPTMAN

Medical Foundation of Buffalo Inc., 73 *High Street, Buffalo, New York* 14203, *USA*

(*Received* 5 *November* 1979; *accepted* 28 *April* 1980)

Abstract

Formulas for estimating the values (0 or π) of the enantiomorph-insensitive two- and three-phase structure seminvariants in $P2_1$ and for identifying the enantiomorph-sensitive ones (*i.e.* those having the values $\pm\pi/2$) have recently been secured. These procedures are here integrated in order to identify still more reliably those seminvariants having the values 0, π or $\pm\pi/2$. The results are employed in two ways to strengthen direct-methods procedures. The first makes active use of the three-phase structure seminvariants alone and the second relies heavily on the three- and four-phase structure invariants but employs the seminvariant enantiomorph-specific figure of merit CRISEM to select the best of a number of possible solutions. The heavy dependence on the enantiomorph-sensitive seminvariants facilitates the selection and maintenance of the enantiomorph, often a source of difficulty in this space group.

1. Introduction

Recently a number of probability distributions for two- and three-phase structure seminvariants were derived on the basis of their quartet, quintet and sextet extensions (Hauptman & Green, 1978; Hauptman & Potter, 1979; van der Putten, Schenk & Hauptman, 1980, respectively). Examples were given which showed that the reliability of the estimates of the seminvariants is good and, particularly in the case of the enantiomorph-sensitive seminvariants, applications to phase determining procedures in space group $P2_1$ were indicated as well. The first object of this paper is to show that integrating the different approaches leads to still more reliable estimates. Then, secondly, the estimates are employed in two direct-methods procedures: *ab initio* phase determination using the three-phase seminvariants alone and phase determination *via* the three-phase and four-phase invariants, employing the seminvariant enantiomorph-specific

0567-7394/80/060897-07$01.00