# Estimation of the One-Phase Structure Seminvariants in the Single Isomorphous Replacement Case: an Application of Hauptman's Distribution 

By Ning-Hai Hu and Yong-Sheng Liu<br>Applied Spectroscopy Laboratory, Changchun Institute of Applied Chemistry, Chinese Academy of Sciences, Changchun 130022, People's Republic of China

(Received 1 September 1994; accepted 18 January 1995)


#### Abstract

A method for estimating the one-phase structure seminvariants (OPSSs) having values of 0 or $\pi$ has been proposed on the basis of the probabilistic theory of the three-phase structure invariants for a pair of isomorphous structures [Hauptman (1982). Acta Cryst. A38, 289-294]. The test calculations using error-free diffraction data of protein cytochrome $\mathrm{c}_{550}$ and its $\mathrm{PtCl}_{4}^{2-}$ derivative show that reliable estimates of a number of the OPSSs can be obtained. The reliability of the estimation increases with the increase of the differences between diffraction intensities of the native protein and its heavyatom derivative. A means to estimate the parameters of the distribution from the diffraction ratio is suggested.


## Introduction

It has long been recognized that the combination of direct methods and isomorphous replacement facilitates the solution of the phase problem of macromolecular structures. Hauptman (1982) proposed the probabilistic theory of the three-phase structure invariants in the single isomorphous replacement (SIR) case. The theory was found to be quite effective in estimating several tens of thousands of the three-phase invariants having values close to 0 or $\pi$ with high reliability (Hauptman, Potter \& Weeks, 1982). In the further exploitation of the potential of Hauptman's distribution, Fortier, Moore \& Fraser (1985) obtained the estimates of cosine invariants in the full range of -1 to 1 and, subsequently, Hao \& Fan (1988) presented a method for the individual phase estimates by making use of heavy-atom structure information. These studies suggested that a combination of direct methods with SIR data might suffice to allow unique macromolecular structure determination.

A major application of the estimates of the three-phase invariants would be to extend and refine phases by use of, for example, a tangent or a least-squares procedure. In addition, the three-phase invariants could be used to evaluate some special individual phases without knowledge of initial phases. The OPSSs, having values of 0 or $\pi$, are a kind of individual phase whose value can be directly evaluated from structure-factor magnitudes. The accurate estimates of a number of the OPSSs may
play a significant role in the early stages of phase extension and refinement. The OPSS estimates in the presence of anomalous scattering have been studied (Velmurugan \& Hauptman, 1989; Velmurugan, Hauptman \& Potter, 1989; Liu \& Hu, 1994). In this paper, we suggest a method to obtain the estimates of the OPSSs in the SIR case by incorporating symmetry relations into Hauptman's distribution, which, in effect, is similar to the $\sum_{1}$ relation used for small-molecule structure determination (Karle \& Hauptman, 1956). The formulae have been tested with error-free diffraction data from protein cytochrome $\mathrm{c}_{550}$ and its $\mathrm{PtCl}_{4}^{2-}$ derivative.

## Theoretical basis

## 1. The three-phase invariant with an OPSS involved

Assuming that $\varphi_{\mathrm{H}}$ is an OPSS, the index of which satisfies

$$
\begin{equation*}
\mathbf{H}=\mathbf{K}-\mathbf{K}_{i} \equiv 0\left(\bmod \omega_{s}\right), \quad i=2, \ldots, m \tag{1}
\end{equation*}
$$

where $\omega_{s}$ is the seminvariant modulus and $m$ is the number of symmetry operators of the actual space group, which consist of the rotation components $\mathbf{R}_{i}$ and the translation components $\mathbf{T}_{i}$, we construct the three-phase structure invariant

$$
\begin{equation*}
\omega=\varphi_{\mathbf{H}}-\varphi_{\mathbf{K}}+\varphi_{\mathbf{K}_{i}} \tag{2}
\end{equation*}
$$

Since

$$
\begin{equation*}
\varphi_{\mathbf{K}}-\varphi_{\mathbf{K}_{i}}=n \pi, \tag{3}
\end{equation*}
$$

where

$$
\begin{equation*}
n=2 \mathbf{K T}_{i} \tag{4}
\end{equation*}
$$

is a constant when the origin is restricted among the permitted origins, the value of $\varphi_{\mathrm{H}}$ depends only on the invariant $\omega$. The types of $\varphi_{\mathrm{H}}$ are just the same as those defined by the $\sum_{1}$ relation if they have the extreme values of 0 or $\pi$. For the space group $P 2_{1} 2_{1} 2_{1}$, for example, there are three types of $\varphi_{\mathrm{H}}: \varphi_{2 h, 2 k, 0} ; \varphi_{0,2 k, 22} ;$ $\varphi_{2 h, 0,21}$. For the space groups with higher symmetry, some vectors $\mathbf{H}$ that are not independent of each other may occur when $i$ changes from 2 to $m$. Only the independent $\mathbf{H}$ vectors should be retained. The types of
the independent $\varphi_{\mathbf{H}}$ for all space groups can be found in the papers of Hašek (1977) and Liu, Jin \& Guo (1982).
2. The probability distributions of the OPSSS in the SIR case

According to Hauptman (1982), for a triplet of reciprocal-lattice vectors $H, K, L$ satisfying $H+K+L=$ 0 , there exist four kinds of three-phase structure invariants:

$$
\begin{align*}
& \omega_{0}=\varphi_{\mathbf{H}}+\varphi_{\mathbf{K}}+\varphi_{\mathbf{L}} \\
& \omega_{1}=\varphi_{\mathbf{H}}+\varphi_{\mathbf{K}}+\psi_{\mathbf{L}}  \tag{5}\\
& \omega_{2}=\varphi_{\mathbf{H}}+\psi_{\mathbf{K}}+\psi_{\mathbf{L}} \\
& \omega_{3}=\psi_{\mathbf{H}}+\psi_{\mathbf{K}}+\psi_{\mathbf{L}}
\end{align*}
$$

where $\varphi$ and $\psi$ are the phases of the structure factors from the native protein and the heavy-atom derivative, respectively, and the corresponding magnitudes are $|E|$ and $|G|$. The conditional probability distributions of the structure invariants $\omega_{\mu}$, assuming as known the six magnitudes $\left|E_{\mathbf{H}}\right|,\left|E_{\mathbf{K}}\right|,\left|E_{\mathbf{L}}\right|,\left|G_{\mathbf{H}}\right|,\left|G_{\mathbf{K}}\right|,\left|G_{\mathbf{L}}\right|$, in their first neighborhood are given by

$$
\begin{align*}
& P_{\mu}\left(\Omega_{\mu}\right)=\left[2 \pi I_{0}\left(A_{\mu}\right)\right]^{-1} \exp \left(A_{\mu} \cos \Omega_{\mu}\right) \\
& \mu=0,1,2,3 \tag{6}
\end{align*}
$$

The notation used here is the same as that of Hauptman (1982). The distribution (6) leads to unique estimates of $\omega_{\mu}$ at 0 or $\pi$ when $A_{\mu}>0$ or $A_{\mu}<0$, respectively.

Now, substituting (2) into (5), we have four particular structure invariants:

$$
\begin{align*}
& \omega_{0}=\varphi_{\mathbf{H}}-\varphi_{\mathbf{K}}+\varphi_{\mathbf{K}_{i}}, \\
& \omega_{1}=\psi_{\mathbf{H}}-\varphi_{\mathbf{K}}+\varphi_{\mathbf{K}_{i}},  \tag{7}\\
& \omega_{2}=\varphi_{\mathbf{H}}-\psi_{\mathbf{K}}+\psi_{\mathbf{K}_{i}}, \\
& \omega_{3}=\psi_{\mathbf{H}}-\psi_{\mathbf{K}}+\psi_{\mathbf{K}_{i}} .
\end{align*}
$$

The conditional probability distribution of $\omega_{0}$, for example, can be written as

$$
\begin{equation*}
P\left(\Omega_{0}\right)=\left[2 \pi I_{0}\left(A_{0}\right)\right]^{-1} \exp \left[A_{0} \cos \left(\varphi_{\mathbf{H}}-\varphi_{\mathbf{K}}+\varphi_{\mathbf{K}_{i}}\right)\right] \tag{8}
\end{equation*}
$$

and $A_{0}$ is given by

$$
\begin{align*}
A_{0}= & 2\left[\beta_{0} R_{1} R_{2}^{2}+\beta_{1}\left(2 R_{1} R_{2} S_{2} T_{2}+R_{2}^{2} S_{1} T_{1}\right)\right. \\
& \left.+\beta_{2}\left(R_{1} S_{2}^{2} T_{2}^{2}+2 R_{2} S_{1} S_{2} T_{1} T_{2}\right)+\beta_{3} S_{1} S_{2}^{2} T_{1} T_{2}^{2}\right] \tag{9}
\end{align*}
$$

where

$$
\begin{equation*}
R_{1}=\left|E_{\mathbf{H}}\right|, \quad R_{2}=\left|E_{\mathbf{K}}\right|, \quad S_{1}=\left|G_{\mathbf{H}}\right|, \quad S_{2}=\left|G_{\mathbf{K}}\right| \tag{10}
\end{equation*}
$$

$T_{j}$ is the ratio of two modified Bessel functions,

$$
\begin{equation*}
T_{j}=I_{1}\left(2 \beta R_{j} S_{j}\right) / I_{0}\left(2 \beta R_{j} S_{j}\right), \quad j=1,2 \tag{11}
\end{equation*}
$$

and $\beta$ parameters have the same meaning as defined by

Hauptman (1982). When a reflection K, together with its equivalent reflection $\mathbf{K}_{i}$, satisfying (1) is given, the probability distribution for a fixed $\varphi_{\mathbf{H}}$ is obtained from (8):

$$
\begin{equation*}
P_{\mathbf{K}}\left(\varphi_{\mathbf{H}}\right)=\left[2 \pi I_{0}\left(A_{0}\right)\right]^{-1} \exp \left[A_{0} \cos \left(\varphi_{\mathbf{H}}-n \pi\right)\right] \tag{12}
\end{equation*}
$$

Then we multiply the individual probability distributions to obtain the approximate form

$$
\begin{align*}
P\left(\varphi_{\mathbf{H}}\right) & =\prod_{\mathbf{K}} P_{\mathbf{K}}\left(\varphi_{\mathbf{H}}\right) \simeq K_{0} \exp \left[\sum_{\mathbf{K}} A_{0} \cos \left(\varphi_{\mathbf{H}}-n \pi\right)\right] \\
& =K_{0} \exp \left[\sum_{\mathbf{K}}(-1)^{n} A_{0} \cos \varphi_{\mathbf{H}}\right] \tag{13}
\end{align*}
$$

where $K_{0}$ is a normalizing constant and the summation involves all the reflections $\mathbf{K}$ that satisfy (1).

A similar expression is derived from the $\omega_{2}$ invariant with

$$
\begin{align*}
A_{2}= & 2\left[\beta_{0} R_{1} R_{2}^{2} T_{2}^{2}+\beta_{1}\left(2 R_{1} R_{2} S_{2} T_{2}+R_{2}^{2} S_{1} T_{1} T_{2}^{2}\right)\right. \\
& \left.+\beta_{2}\left(R_{1} S_{2}^{2}+2 R_{2} S_{1} S_{2} T_{1} T_{2}\right)+\beta_{3} S_{1} S_{2}^{2} T_{1}\right] \tag{14}
\end{align*}
$$

Using the average of $A_{0}$ and $A_{2}$, therefore, the probability distribution of $\varphi_{\mathbf{H}}$ may be rewritten in the form

$$
\begin{equation*}
P\left(\varphi_{\mathbf{H}}\right) \simeq\left[2 \pi I_{0}\left(B_{\varphi}\right)\right]^{-1} \exp \left(B_{\varphi} \cos \varphi_{\mathbf{H}}\right) \tag{15}
\end{equation*}
$$

where

$$
\begin{equation*}
B_{\varphi}=\frac{1}{2} \sum_{\mathbf{K}}(-1)^{n}\left(A_{0}+A_{2}\right) \tag{16}
\end{equation*}
$$

Clearly, (15) has a unique maximum at $\varphi_{\mathrm{H}}=0$ or $\pi$ when $B_{\varphi}>0$ or $B_{\varphi}<0$, respectively, and the larger the value of $\left|B_{\varphi}\right|$, the smaller the variance of the distribution. Thus, one obtains the estimate of the cosine seminvariant, $\cos \varphi_{\mathrm{H}}$, by calculating the sign of $B_{\varphi}$ when $\left|B_{\varphi}\right|$ is large.

With a similar approach, we have the probability distribution of $\psi_{\mathrm{H}}$ from the $\omega_{1}$ and $\omega_{3}$ invariants:

$$
\begin{equation*}
P\left(\psi_{\mathbf{H}}\right) \simeq\left[2 \pi I_{0}\left(B_{\psi}\right)\right]^{-1} \exp \left(B_{\psi} \cos \psi_{\mathbf{H}}\right) \tag{17}
\end{equation*}
$$

where

$$
\begin{equation*}
B_{\psi}=\frac{1}{2} \sum_{\mathbf{K}}(-1)^{n}\left(A_{1}+A_{3}\right) \tag{18}
\end{equation*}
$$

$$
\begin{align*}
A_{1}= & 2\left[\beta_{0} R_{1} R_{2}^{2} T_{1}+\beta_{1}\left(2 R_{1} R_{2} S_{2} T_{1} T_{2}+R_{2}^{2} S_{1}\right)\right. \\
& \left.+\beta_{2}\left(R_{1} S_{2}^{2} T_{1} T_{2}^{2}+2 R_{2} S_{1} S_{2} T_{2}\right)+\beta_{3} S_{1} S_{2}^{2} T_{2}^{2}\right] \tag{19}
\end{align*}
$$

$$
\begin{align*}
A_{3}= & 2\left[\beta_{0} R_{1} R_{2}^{2} T_{1} T_{2}^{2}+\beta_{1}\left(2 R_{1} R_{2} S_{2} T_{1} T_{2}+R_{2}^{2} S_{1} T_{2}^{2}\right)\right. \\
& \left.+\beta_{2}\left(R_{1} S_{2}^{2} T_{1}+2 R_{2} S_{1} S_{2} T_{2}\right)+\beta_{3} S_{1} S_{2}^{2}\right] \tag{20}
\end{align*}
$$

Similarly, the cosine seminvariant $\cos \psi_{\mathrm{H}}$ can be estimated from (17).

Further accuracy of the estimates will accrue if the heavy-atom structure is known. As suggested by Fortier,

Moore \& Fraser (1985), the calculated values of the cosine of the phase difference,

$$
\begin{equation*}
\cos \left(\varphi_{j}-\psi_{j}\right)=\left(\left|F_{j P H}\right|^{2}+\left|F_{j p}\right|^{2}-\left|F_{j H}\right|^{2}\right) / 2\left|F_{j P H}\right|\left|F_{j P}\right|, \tag{21}
\end{equation*}
$$

can be substituted for its expected values $T_{j}$ in (9), (14), (19) and (20). The symbols $\left|F_{j P H}\right|,\left|F_{j P}\right|$ and $\left|F_{j H}\right|$ are the $j$ th-reflection structure-factor magnitudes of the heavyatom derivative, native protein and heavy-atom structure, respectively.

## 3. Estimation of the $\beta$ parameters

The $\beta$ parameters in the expressions of $A$ are expressed in terms of $\alpha_{20}, \alpha_{02}, \alpha_{30}, \alpha_{03}, \alpha_{21}, \alpha_{12}$ and $\alpha_{11}$ [see equations (3.5)-(3.9) of Hauptman (1982)], which are defined by

$$
\begin{equation*}
\boldsymbol{\alpha}_{m n}=\sum_{j=1}^{N} f_{j}^{m} g_{j}^{n}, \tag{22}
\end{equation*}
$$

where $N$ is the number of atoms in the unit cell, $f_{j}$ and $g_{j}$ are zero-angle atomic scattering factors for a pair of isomorphous structures and therefore equal to the atomic number $Z_{j}$ in the X-ray diffraction case. In order to calculate the values of $\alpha_{m n}$, information concerning the chemical identity, the number and occupancy factors of heavy atoms is required. If the $f$ structure is a native protein and the $g$ structure a heavy-atom derivative, $\alpha_{0 n}$ consists of contributions from the native protein and the heavy atoms. In such a case, it is possible to make estimates of $\alpha_{0 n}$ from the diffraction ratio

$$
\begin{equation*}
\left.r=\left\langle\left(|E|^{2}-|G|^{2}\right)^{2}\right\rangle^{1 / 2} /\left.\langle | E\right|^{2}\right\rangle, \tag{23}
\end{equation*}
$$

which is a measure of the average change in intensity due to the addition of heavy atoms.

By use of the result of Crick \& Magdoff (1956), it readily follows that

$$
\begin{equation*}
Z_{H} \simeq r\left(\alpha_{20} / 2 N_{H}\right)^{1 / 2}, \tag{24}
\end{equation*}
$$

where $N_{H}$ is the number of heavy atoms in the unit cell and $Z_{H}$ can be regarded as an effective atomic number of the heavy atoms, which is an average effect over the different atoms and occupancy factors. Although one cannot uniquely determine both $Z_{H}$ and $N_{H}$ from (24), it is possible to obtain an approximate value of $N_{H}$ by making $Z_{H}$ reasonable. A similar way to estimate the content of heavy atoms has been suggested by Fortier, Weeks \& Hauptman (1984). Once $N_{H}$ has been determined, those $\alpha_{0 n}$ containing the contributions from the heavy atoms can be evaluated as follows:

$$
\begin{align*}
\alpha_{02} & \simeq \alpha_{20}+\frac{1}{2} r^{2} \alpha_{20},  \tag{25}\\
\alpha_{03} & \simeq \alpha_{30}+\frac{1}{2}\left(2 N_{H}\right)^{-1 / 2} r^{3} \alpha_{20}^{3 / 2} \tag{26}
\end{align*}
$$

and thus the $\beta$ parameters are obtained without detailed knowledge of the heavy-atom content.

## Applications

1. The test calculations for a pair of isomorphous structures

The method for estimating the values ( 0 or $\pi$ ) of the OPSSs described above was tested with error-free data of protein cytochrome $\mathrm{c}_{550}$, molecular weight about 14500 , space group $P 2_{1} 2_{1} 2_{1}$ and a single $\mathrm{PtCl}_{4}^{2-}$ isomorphous derivative (Timkovich \& Dickerson, 1976). The normalized structure factors were calculated from the atomic coordinates to a resolution of $2.5 \AA$ ( $4159 E$ and $4159 G$ values). For each structure, there are 236 OPSSs ( $\varphi_{\mathrm{H}}$ or $\psi_{H}$ ) having indices in the forms $2 h, 2 k, 0 ; 0,2 k, 2 l ; 2 h, 0,2 l$. The cosines of $\varphi_{\mathrm{H}}$ and $\psi_{\mathrm{H}}$ were estimated by calculating $B_{\varphi}$ and $B_{\psi}$ using (16) and (18) according to three protocols:
I. $\alpha_{m n}$ containing the contributions of the heavy atoms was estimated from (25) and (26) and $T_{j}$ was calculated from (11). This means no knowledge concerning the heavy atoms was used.
II. An exact $\alpha_{m n}$ was calculated from (22) and $T_{j}$ from (11).
III. $\alpha_{m n}$ was calculated as in II but $T_{j}$ was substituted by the calculated value (21) of $\cos \left(\varphi_{j}-\psi_{j}\right)$. This means that the heavy-atom information including the number, occupancy and positions was utilized.
The calculated results were arranged in descending order of $|B|$ values. Table 1 was constructed by accumulation into the four groups shown according to the given minimum $|B|$ values. The top 200 divided into five groups are listed in Table 2. It is clearly shown that, as expected, the larger the $|B|$ value the more reliable is the cosine seminvariant estimate. In fact, as shown in Table 1, almost all the signs of the cosine seminvariants are correct for $|B|>1.5$. Even when the $|B|$ criterion is lowered to 0.5 , the percentage of the seminvariants correctly estimated is still over $90 \%$. Protocol I gives the results using the estimated values of $\alpha_{0 n}$ with $N_{H}=4$. For a given $\left|B_{\min }\right|$, the number of seminvariants in the group for I is a little smaller than that for II and the number of incorrect signs is also smaller for I. This indicates that employing approximate values of $\alpha_{0 n}$ does not interfere with the signs of the cosine seminvariants but causes an overall reduction of the $|B|$ values in this example. Comparison of I with II and III shows that the use of knowledge of the heavy-atom structure markedly improves the estimation while the use of the heavy-atom content does little. It is noted that the reliably estimated seminvariants are not necessarily the reflections having large $|E|$ (or $|G|$ ) values but correspond instead to those having large $|E-G|$ values. Fig. 1 shows the relation between the change of the normalized structure-factor magnitudes and the reliability of the estimates. For Fig. 1 , the results for the native protein $\left(\varphi_{\mathrm{H}}\right)$ were arranged in descending order of $|E-G|$ and then the percentages of the seminvariants with $|B|>1.5$ and the averages $\langle | E-G| \rangle$ were computed for the ranked groups; each

Table 1. Estimates of the one-phase cosire seminvariants accumulated in groups according to $\left|B_{\text {min }}\right|$ for cytochrome $c_{550}$ and its $\mathrm{PtCl}_{4}^{2-}$ derivative
$\left|B_{\min }\right|$ : the minimum $|B|$ value in the group. Nwr: the number of signs of the cosine seminvariants incorrectly estimated.

|  |  | 1 |  |  |  | II |  |  |  | III |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\left\|B_{\text {min }}\right\|$ | Number in group | $\langle \| E\rangle$ | $\langle \| E-G\| \rangle$ | Nwr | Number in group | $\langle \| E\rangle$ | $\langle \| E-G\| \rangle$ | Nwr | Number in group | $\langle \| E\rangle$ | $\langle \| E-G\| \rangle$ | Nwr |
| Native | 5.0 | 70 | 0.99 | 0.49 | 0 | 79 | 0.97 | 0.48 | 0 | 112 | 0.87 | 0.44 | 0 |
| $\left(\varphi_{\mathbf{H}}\right)$ | 3.0 | 93 | 0.94 | 0.46 | 0 | 106 | 0.92 | 0.44 | 0 | 132 | 0.82 | 0.42 | 0 |
|  | 1.5 | 128 | 0.89 | 0.42 | 0 | 137 | 0.89 | 0.41 | 2 | 159 | 0.81 | 0.39 | 0 |
|  | 0.5 | 173 | 0.85 | 0.37 | 13 | 176 | 0.85 | 0.36 | 15 | 190 | 0.82 | 0.37 | 4 |
| Deriva- | 5.0 | 76 | 0.95 | 0.47 | 0 | 86 | 0.94 | 0.45 | 0 | 112 | 0.87 | 0.44 | 0 |
| tive | 3.0 | 96 | 0.95 | 0.45 | 0 | 104 | 0.94 | 0.44 | 0 | 132 | 0.82 | 0.42 | 0 |
| $\left(\psi_{H}\right)$ | 1.5 | 118 | 0.95 | 0.42 | 0 | 127 | 0.95 | 0.42 | 2 | 159 | 0.79 | 0.39 | 0 |
|  | 0.5 | 159 | 0.88 | 0.38 | 10 | 168 | 0.85 | 0.37 | 13 | 190 | 0.79 | 0.37 | 4 |

Table 2. Estimate results of the top 200 one-phase cosine seminvariants arranged in descending order of $|B|$ and divided into five groups, each containing 40 , for cytochrome $c_{550}$ and its $\mathrm{PtCl}_{4}^{2-}$ derivative

Nwr has the same meaning as in Table 1.

contains 40 data but the last contains 36 data. It can be seen that at least $95 \%$ of the seminvariants are correctly estimated when the change of the normalized structurefactor magnitudes is larger than 0.6 and this percentage rapidly decreases with decreasing $\langle | E-G| \rangle$. Fig. 1 also shows that the incorporation of the heavy-atom structure information (protocol III) leads to an obvious increase in the number of seminvariants with large $|B|$ values, especially when $\langle | E-G| \rangle<0.45$. In other words, a remarkable gain in reliability for those seminvariants having smaller $|E-G|$ values can be obtained by the use of the calculated values of $\cos \left(\varphi_{j}-\psi_{j}\right)$ instead of its expected values.

## 2. Comparison with the non-isomorphous case

In the non-isomorphous case, the cosine seminvariants can be estimated by

$$
\begin{equation*}
P\left(\varphi_{\mathbf{H}}\right)=\left[2 \pi I_{0}(B)\right]^{-1} \exp \left(B \cos \varphi_{\mathbf{H}}\right), \tag{27}
\end{equation*}
$$

where

$$
\begin{equation*}
B=2\left(\sum_{j} Z_{j}^{3}\right)\left(\sum_{j} Z_{j}^{2}\right)^{-3 / 2}\left|E_{\mathbf{H}}\right| \sum_{\mathbf{K}}(-1)^{n}\left|E_{\mathbf{K}}\right|^{2} \tag{28}
\end{equation*}
$$

and $Z_{j}$ is the atomic number of the $j$ th atom in the unit cell. This is almost the same as the $\sum_{1}$ formula.

In order to make a comparison with the nonisomorphous case, the $236 \varphi_{\mathrm{H}}$ values for the native protein were estimated from (28) as protocol IV. These were arranged in descending order of $|B|$ values and the top 200 are listed in Table 2. The same calculations were also carried out for the derivative structure. The average


Fig. 1. Percentage of the seminvariants $\varphi_{\mathrm{H}}$ with $|B|>1.5$ as a function of $\langle | E-G| \rangle$, showing the reliability of the estimates dependent on the differences between the structure-factor magnitudes of the native protein and the derivative.
values of $|B|$ in the non-isomorphous case are much smaller than those in the isomorphous case, as is shown in Table 2, resulting in the rather poor estimates. Nearly half of the seminvariants are incorrectly estimated. This implies that the traditional $\sum_{1}$ formula is not applicable in the macromolecular case and the comparison confirms that the ability to combine direct methods and isomorphous replacement is powerful.

## Concluding remarks

The distribution of Hauptman (1982) employing a combination of direct methods and the SIR technique has been developed to estimate the values ( 0 or $\pi$ ) of the OPSSs. The method proved to be effective with the errorfree data of a pair of isomorphous structures. No heavyatom information, neither the positions nor the content of heavy atoms, is necessary to obtain the estimates of the cosine seminvariants but, if the heavy atoms are located, better results can be obtained by making use of the heavy-atom structure information. The test calculations agree with the prediction of Fortier, Weeks \& Hauptman (1984) that, even when the normalized structure factors themselves are small, reliable estimates can be obtained provided that the differences beween the structure-factor magnitudes of the native protein and the derivative are large.

The method presented here is actually equivalent to the $\sum_{1}$ formula combining with SIR data and provides a supplementary technique of finding individual phases in the initial stages of the phasing procedure. An obvious practical application lies in the possibility of enhancing the starting set in the standard tangent refinement
by incorporating a number of the OPSSs with high reliability. In view of the fact that at least two thirds of the OPSSs can be accurately determined for the chosen example, it is not unreasonable to expect that the method may play a more important role in the solution of macromolecular structures than the $\sum_{1}$ formula does in the small molecule case. This work also shows that Hauptman's distribution is very promising in solving the phase problem in the SIR case, although further theoretical and experimental studies are needed for applying it to unknown structures.

This work was supported in part by Jilin Aodong Pharmaceuticals Ltd.

## References

Crick, F. H. C. \& Magdoff, B. S. (1956). Acta Cryst. 9, 901-908.
Fortier, S., Moore, N. J. \& Fraser, M. E. (1985). Acta Cryst. A41, 571-577.
Fortier, S., Weeks, C. M. \& Hauptman, H. (1984). Acta Cryst. A40, 544-548.
Hao, Q. \& Fan, H.-F. (1988). Acta Cryst. A44, 379-382.
HAšek, J. (1977). Z. Kristallogr. 145, 263-270.
Hauptman, H. (1982). Acta Cryst. A38, 289-294.
Hauptman, H., Potter, S. \& Weeks, C. M. (1982). Acta Cryst. A38, 294-300.
Karle, J. \& Hauptman, H. (1956). Acta Cryst. 9, 635-651.
Liu, Y.-S. \& Hu, N.-H. (1994). Acta Cryst. A50, 595-601.
LiU, Y.-S., Jin, Z.-S. \& Guo, D.-Y. (1982). Sci. Sin. (Ser. B), 25, 800 814.

Timkovich, R. \& Dickerson, R. E. (1976). J. Biol. Chem. 251, 40334046.

Velmurugan, D. \& Hauptman, H. (1989). Acta Cryst. A45, 158-163.
Velmurugan, D., Hauptman, H. \& Potier, S. A. (1989). Acta Cryst. A45, 163-165.

Acta Cryst. (1995). A51, 524-529

# Magnetic Completely Transposable Twin Laws and Tensor Distinction 

By D. B. Litvin and S. Y. Litvin<br>Department of Physics, The Pennsylvania State University, Penn State Berks Campus, PO Box 7009, Reading, PA 19610-6009, USA

and V. Janovec
Institute of Physics, Academy of Sciences of the Czech Republic, Na Slovance 2, 18040 Prague 8, Czech Republic
(Received 25 July 1994; accepted 17 January 1995)


#### Abstract

Macroscopic tensorial physical properties that are different in two domains of a ferroic crystal provide a tensor distinction of the two domains. This tensor distinction is determined from a symmetry relationship,


called a twin law, between the bulk structures, the domain states, of the two domains. The simplest type of twin law is the so-called completely transposable twin law. We extend here the concept of completely transposable twin laws from non-magnetic to magnetic completely transposable twin laws. We establish the

