

expanding (32) to first order in a Taylor series about some 'guessed' value of K , K_g . One obtains the following values to be used in (2).

$$c_{ij} = \frac{(\cos^2 2\theta_j + K_g)^2}{\cos^2 2\theta_j - \cos^2 2\theta_i},$$

$$d_{ij} = K_g - \frac{(\cos^2 2\theta_i + K_g)(\cos^2 2\theta_j + K_g)}{\cos^2 2\theta_j - \cos^2 2\theta_i}. \quad (33)$$

The weighting scheme was implemented in a Fortran program. The weights were refined in cycles according to (30) until the maximum relative shift was less than a specified quantity (a value of 0.5% was chosen). It was found that δ had to be small (0.03) to avoid divergence. A considerable number of cycles (~650) were necessary in order for the weights to converge but the calculation nevertheless is extremely rapid on a modern computer. We have not attempted to improve the rate of convergence. The value of K_g initially chosen was for

a kinematical monochromator and the whole calculation was rerun with K_g equal to \bar{K} obtained from the first run. The value of \bar{K} determined is highly insensitive to the value of K_g .

Some simulations were carried out to test the method and the program. These were based on the data used in paper II of this series but with intensities calculated from a model value of K , and with a pseudo-random Gaussian error added in to match the counting statistics of the experiment. From these simulations it is clear that the weighting scheme produces values of \bar{K} closer to K than the simple weighting scheme of x_{ij} and that the e.s.d. of \bar{K} is also reduced.

References

- HAMILTON, W. C. (1964). *Statistics in Physical Science*. New York: Ronald Press Co.
VINCENT, M. G. & FLACK, H. D. (1980). *Acta Cryst.* **A36**, 614–620.

Acta Cryst. (1980). **A36**, 624–632

On the Theory and Estimation of the Three-Phase Structure Seminvariant in $P\bar{1}$

BY HERBERT HAUPTMAN

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

(Received 16 October 1979; accepted 8 February 1980)

Abstract

By embedding the three-phase structure seminvariant T and its three symmetry-related variants in suitable quintets Q (five-phase structure invariants) one obtains the extensions Q of T to which T is related *via* the space-group symmetries. The neighborhoods of T are then defined in terms of the neighborhoods of its extensions. The conditional probability distribution of T , given the seven magnitudes $|E|$ in its first neighborhood, is derived. The distribution yields a reliable estimate (0 or π) for T in the favorable case that the variance of the distribution happens to be small.

1. Introduction

In recent years the basic concepts and mathematical formalism needed for the development of the probabilistic theory of the structure seminvariants have been elucidated. For example, it has long been known that, for fixed enantiomorph, the collection of observed magnitudes $|E|$ determines, in general, the values of all the structure seminvariants. A major recent advance is

the neighborhood principle: For fixed enantiomorph, the value of any structure seminvariant T is primarily determined, in favorable cases, by the values of one or more small sets of magnitudes $|E|$, the neighborhoods of T , and is relatively insensitive to the values of the great bulk of remaining magnitudes (Hauptman, 1975). The conditional probability distribution of T , given the magnitudes in any of its neighborhoods, yields an estimate for T which is particularly good in the favorable case that the variance of the distribution happens to be small. With the identification of systems of neighborhoods for the structure invariants, the probabilistic theory of the structure invariants developed rapidly, especially in space groups $P1$ and $P\bar{1}$, but a great deal of work still remains to be done in deriving accurate and readily computable probability distributions, particularly for the space-group-special structure invariants.

Again, with the formulation of the extension concept [Hauptman (1977*b*, 1978); but see Giacovazzo (1977) for the equivalent concept called representation theory], the probabilistic theory of the structure seminvariants was reduced to that of the structure invariants, which is more highly developed. In particular, the

neighborhoods of any structure seminvariant are defined in terms of those of its extensions, and associated probability distributions may then be derived. How this is to be done has been described in a few cases (Giacovazzo, 1978*a,b*; Green & Hauptman, 1976, 1978; Hauptman & Green, 1978; Hauptman & Potter, 1979). The importance of these developments of course is that the structure seminvariants, as certain well defined linear combinations of the phases, lead to unique values for the individual phases. Furthermore, the structure seminvariants are available in much larger numbers than the structure invariants, and the probabilistic theory of the structure seminvariants makes strong use of the space-group symmetries as well.

In the present paper the probabilistic theory of the three-phase structure seminvariant T in $P\bar{1}$, based on the first neighborhood of T , is described. Because of the heavy dependence on recent work, the present paper is greatly abbreviated, and only the briefest sketch of the derivation of the major result, equations (4.2)–(4.5), is given in Appendices IV–VIII. The derivation of the first neighborhood given in Appendices I–III is described in somewhat greater detail because it is expected that this work will serve as the prototype for deriving neighborhoods of the structure seminvariants in general.

2. The first neighborhood of the three-phase structure seminvariant T

The linear combination of three phases

$$T = \varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} + \varphi_{\mathbf{l}} \quad (2.1)$$

is a structure seminvariant in $P\bar{1}$ provided that

$$\mathbf{h} + \mathbf{k} + \mathbf{l} \equiv 0 \pmod{(222)}, \quad (2.2)$$

i.e. provided that the three components of $\mathbf{h} + \mathbf{k} + \mathbf{l}$ are even. Then the components of the eight reciprocal vectors

$$\frac{1}{2}(\pm\mathbf{h} \pm \mathbf{k} \pm \mathbf{l}) \quad (2.3)$$

are integers. Following techniques recently described (Hauptman, 1977*b*, 1978; Giacovazzo, 1977), one embeds the structure seminvariant T and its symmetry-related variants in suitable quintets (five-phase structure invariants) in order to obtain the extensions Q of T to which T is simply related. In this way, as described in Appendices I–III, one finds that the first neighborhood of T consists of the seven magnitudes

$$R_1 = |E_{\mathbf{h}}|, \quad R_2 = |E_{\mathbf{k}}|, \quad R_3 = |E_{\mathbf{l}}|; \quad (2.4)$$

$$r = |E_{\frac{1}{2}(\mathbf{h}+\mathbf{k}+\mathbf{l})}|, \quad r_1 = |E_{\frac{1}{2}(\mathbf{h}-\mathbf{k}+\mathbf{l})}|,$$

$$r_2 = |E_{\frac{1}{2}(\mathbf{h}-\mathbf{k}-\mathbf{l})}|, \quad r_3 = |E_{\frac{1}{2}(\mathbf{h}+\mathbf{k}-\mathbf{l})}|. \quad (2.5)$$

In view of the neighborhood principle, the value of T is, in favorable cases, primarily dependent on the seven

magnitudes (2.4) and (2.5) in the first neighborhood. As described in the sequel, the favorable cases here require first that the three magnitudes (2.4) all be large and second that either all four magnitudes (2.5) be large, in which case $T \simeq 0$, or that precisely two of the four magnitudes (2.5) be large and the remaining two be small, in which case $T \simeq \pi$.

3. The probabilistic background

Suppose that a crystal structure in $P\bar{1}$ is fixed and that the seven non-negative numbers $R_1, R_2, R_3; r, r_1, r_2, r_3$ are also specified. Denote reciprocal space by W and by $W \times W \times W$ the threefold Cartesian product which consists of all ordered triples $(\mathbf{h}, \mathbf{k}, \mathbf{l})$ of reciprocal vectors. Suppose finally that $(\mathbf{h}, \mathbf{k}, \mathbf{l})$ is the primitive random variable which is assumed to be uniformly distributed over the subset of $W \times W \times W$ defined by (2.2), (2.4) and (2.5). Then the structure seminvariant $T[(2.1)]$, as a function of the primitive random variable $(\mathbf{h}, \mathbf{k}, \mathbf{l})$, is itself a random variable. Denote by P^+ or P^- the conditional probability, given the seven magnitudes (2.4) and (2.5), that T be 0 or π , respectively, or, equivalently, that $\cos T = +1$ or -1 , respectively.

4. The conditional probability distribution P^\pm of the three-phase structure seminvariant T , given the seven magnitudes in its first neighborhood

Make the definition

$$\sigma_n = \sum_{j=1}^N f_j^n, \quad (4.1)$$

where N is the number of atoms in the whole unit cell and f_j is the zero-angle atomic scattering factor of the atom labeled j . 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. The formula for P^\pm , the major result of this paper, is an easy consequence of (VIII.6) and (VIII.10), Appendix VIII:

$$P^\pm = \frac{1}{K} Z^\pm, \quad K = Z^+ + Z^-, \quad (4.2)$$

$$Z^\pm = \exp(\pm V) \sum_{\eta_1, \eta_2, \eta_3 = \pm 1} \exp(W^\pm), \quad (4.3)$$

$$V = \left[\frac{1}{2\sigma_2^{9/2}} (6\sigma_3^3 - 6\sigma_2\sigma_3\sigma_4 + \sigma_2^2\sigma_5)(r^2 + r_1^2 + r_2^2 + r_3^2) \right. \\ \left. - \frac{2}{\sigma_2^{9/2}} (2\sigma_3^3 - 3\sigma_2\sigma_3\sigma_4 + \sigma_2^2\sigma_5) \right] R_1 R_2 R_3, \quad (4.4)$$

$$\begin{aligned}
W^\pm = & (\eta_1 r r_1 + \eta_2 \eta_3 r_2 r_3) \\
& \times \left[\pm \frac{\sigma_3}{\sigma_2^{3/2}} R_1 - \frac{1}{\sigma_2^3} (2\sigma_3^2 - \sigma_2 \sigma_4) R_2 R_3 \right] \\
& + (\eta_2 r r_2 + \eta_3 \eta_1 r_3 r_1) \\
& \times \left[\pm \frac{\sigma_3}{\sigma_2^{3/2}} R_2 - \frac{1}{\sigma_2^3} (2\sigma_3^2 - \sigma_2 \sigma_4) R_3 R_1 \right] \\
& + (\eta_3 r r_3 + \eta_1 \eta_2 r_1 r_2) \\
& \times \left[\pm \frac{\sigma_3}{\sigma_2^{3/2}} R_3 - \frac{1}{\sigma_2^3} (2\sigma_3^2 - \sigma_2 \sigma_4) R_1 R_2 \right], \quad (4.5)
\end{aligned}$$

where the summands of (4.3) consist of the eight exponentials obtained by permitting each of η_1, η_2, η_3 to take on the values $+1$ and -1 independently in (4.5); otherwise upper (lower) signs in (4.2)–(4.5) go together (Hauptman, 1979).

5. The discriminant of T

If one bases the derivation of P^\pm on (VII.1) of Appendix VII [rather than on (VII.3), as was done in Appendix VIII and § 4], employing the method of Appendix VIII and the identity

$$P_3 = \exp(\log P_3), \quad (5.1)$$

one obtains P^\pm in pure exponential form:

$$P^\pm \simeq \frac{1}{2 \cosh \Delta} \exp(\pm \Delta), \quad (5.2)$$

where Δ , the so-called discriminant of the structure seminvariant T , turns out to be

$$\begin{aligned}
\Delta = & \{(\sigma_3^3/\sigma_2^{9/2})(r^2 r_1^2 r_2^2 + r^2 r_2^2 r_3^2 + r^2 r_3^2 r_1^2 + r_1^2 r_2^2 r_3^2) \\
& - [(2\sigma_3^3 - \sigma_2 \sigma_3 \sigma_4)/\sigma_2^{9/2}] \\
& \times (r^2 r_1^2 + r^2 r_2^2 + r^2 r_3^2 + r_1^2 r_2^2 + r_2^2 r_3^2 + r_3^2 r_1^2) \\
& + (1/2\sigma_2^{9/2})(6\sigma_3^3 - 6\sigma_2 \sigma_3 \sigma_4 + \sigma_2^2 \sigma_5) \\
& \times (r^2 + r_1^2 + r_2^2 + r_3^2) \\
& - (2/\sigma_2^{9/2})(2\sigma_3^3 - 3\sigma_2 \sigma_3 \sigma_4 + \sigma_2^2 \sigma_5)\} R_1 R_2 R_3, \quad (5.3)
\end{aligned}$$

It is of course too much to expect that the pure exponential (5.2) can accurately represent the sum of the eight exponentials defined by (4.2)–(4.5), and some preliminary calculations confirm that (5.2) is in fact a rather poor approximation to the true distribution P^\pm .

Nevertheless, the applications also show that extreme values of the discriminant Δ are well correlated with the values of T in the sense that $T \simeq 0$ or π according as $\Delta \gg 0$ or $\Delta \ll 0$, respectively. Thus, as has been observed many times previously (e.g. Fortier & Hauptman, 1977) the discriminant proves to be a useful and readily computable indicator of the value of the seminvariant. In this particular case, however, little is gained by using the discriminant, since (4.2)–(4.5) are also readily computable.

Inspection of (5.3) shows that, on the assumption that the three magnitudes (2.4) are large, $\Delta \gg 0$ if all four magnitudes (2.5) are large; but $\Delta \ll 0$ if precisely two magnitudes (2.5) are large and the remaining two are small. In short $T \simeq 0$ or π in the respective cases, and the qualitative prediction of Table 3 (Appendix III) is in fact borne out. Inspection of the distribution (4.2)–(4.5) leads to the same conclusion, but in a less transparent way.

Finally, it is not difficult to show that if one of the magnitudes (2.5) is outside the observed range, then it is to be replaced by unity, the average value of $|E|^2$, in (5.3) or (4.2)–(4.5) (see, for example, Giovacazzo, 1975; Heinerman, Krabbendam & Kroon, 1977).

6. Concluding remarks

For the special case that all atoms are identical, Giovacazzo (1978a) has derived the conditional probability distribution of T , given the 17 magnitudes in the complete second neighborhood (called by him the first phasing shell), using a different mathematical technique. Although it is not possible to compare his result [Giovacazzo, 1978a, equation (10)] directly with (4.2)–(4.5), numerical comparisons will prove to be illuminating [see, e.g. Hauptman & Green (1976) where a similar comparison in the case of the four-phase structure invariant in $P\bar{I}$ revealed significant differences between the results of the two techniques].

In the present paper the conditional probability distribution of the three-phase structure seminvariant T , given the seven magnitudes in its first neighborhood, has been found for space group $P\bar{I}$. The distribution yields a reliable estimate (0 or π) for T in the favorable case that the variance of the distribution happens to be small. Because the structure seminvariants are available in much larger numbers than the structure invariants, it is anticipated that the results described here will find important application in the determination of real structures, and some initial calculations confirm this expectation. These applications will be reported shortly.

Of particular importance is the analysis leading to the definition of the first neighborhood given in Appendices I–III. This work, as well as the derivation of the distribution, has been written in such a way as to

permit easy generalization to the higher-order structure seminvariants, and the final results for the four- and five-phase structure seminvariants in $P\bar{1}$ have been derived and will be published in the near future. It is expected also that the work described here will serve as a useful guideline for the derivation of analogous distributions in other space groups.

The research described here was supported in part by Grant No. CHE79-11282 from the National Science Foundation.

APPENDIX I

The extensions of T

Following Hauptman [(1977*b*, 1978); but see also Giacovazzo (1977) for an equivalent procedure], one embeds the three-phase structure seminvariant.

$$T = \varphi_h + \varphi_k + \varphi_l \quad (\text{I.1})$$

and its three symmetry-related variants

$$T_1 = \varphi_{-h} + \varphi_k + \varphi_l \quad (\text{I.2})$$

$$T_2 = \varphi_h + \varphi_{-k} + \varphi_l \quad (\text{I.3})$$

$$T_3 = \varphi_h + \varphi_k + \varphi_{-l} \quad (\text{I.4})$$

in the respective quintets

$$Q_0 = T + \varphi_{-\frac{1}{2}(h+k+l)} + \varphi_{-\frac{1}{2}(h+k+l)}, \quad (\text{I.5})$$

$$Q_1 = T_1 + \varphi_{-\frac{1}{2}(-h+k+l)} + \varphi_{-\frac{1}{2}(-h+k+l)}, \quad (\text{I.6})$$

$$Q_2 = T_2 + \varphi_{-\frac{1}{2}(h-k+l)} + \varphi_{-\frac{1}{2}(h-k+l)}, \quad (\text{I.7})$$

$$Q_3 = T_3 + \varphi_{-\frac{1}{2}(h+k-l)} + \varphi_{-\frac{1}{2}(h+k-l)}. \quad (\text{I.8})$$

In view of (I.1)–(I.4) and the space-group-dependent relationships among the phases, it is readily verified that (I.5) to (I.8) are five-phase structure invariants (quintets) and

$$T = Q_0 = Q_1 = Q_2 = Q_3. \quad (\text{I.9})$$

Thus the probabilistic theory of the three-phase structure seminvariant T is reduced to that of quintets which is well developed. In particular, the neighborhoods of T are defined in terms of the neighborhoods of the quintet.

It should perhaps be stressed that each of the four quintets (I.5)–(I.8) is special since two phases in each quintet are equal to each other. For this reason, the reasoning of the following Appendices II and III is only plausible and should be regarded primarily as a heuristic device for identifying the first neighborhood of T . The analysis of the remaining Appendices IV to VIII however, leading to the major results of this paper, the

distribution P^\pm , (4.2)–(4.5), and the discriminant Δ , (5.3), is perfectly rigorous; the qualitative agreement among Table 3, P^\pm and Δ , as well as the initial applications to be reported shortly, vindicates the approach described here.

APPENDIX II

The first two neighborhoods of the extensions

Clearly only four of the five ‘main terms’ (Schenk, 1975; Hauptman, 1977*a*) of the special quintet Q_0 (I.5) are distinct. The first neighborhood of Q_0 is accordingly defined to consist of the four magnitudes

$$|E_h|, |E_k|, |E_l|, |E_{\frac{1}{2}(h+k+l)}|. \quad (\text{II.1})$$

Referring to the earlier work it is readily verified that only seven of the ten ‘cross terms’ of the special quintet Q_0 are in fact distinct. Thus the second neighborhood of Q_0 is defined to consist of the four main terms (II.1) and the seven cross terms

$$|E_{\frac{1}{2}(-h+k+l)}|, |E_{\frac{1}{2}(h-k+l)}|, |E_{\frac{1}{2}(h+k-l)}|; \\ |E_{h+k}|, |E_{k+l}|, |E_{l+h}|; |E_{h+k+l}|, \quad (\text{II.2})$$

or eleven magnitudes $|E|$ in all. (In contrast, the second neighborhood of the general quintet consists of fifteen magnitudes $|E|$.)

Not only does the earlier work serve to identify the first two neighborhoods of the special quintet Q_0 but Table 2 of Hauptman (1977*a*) shows in a qualitative way what the relationship between Q_0 and the magnitudes $|E|$ in its second neighborhood must be. Thus the 1977 Table 2 leads, by suitable specialization, to rows 1–8 of Table 1. Although the 1977 Table 2 has 16 rows, only eight of these yield entries in Table 1 which are internally consistent; the remaining eight rows of the 1977 Table 2, which would lead to contradictory entries in Table 1 (arising from the fact that only seven of the cross terms of the special quintet Q_0 are distinct), are therefore omitted from Table 1. Thus the 16 rows of the 1977 Table 2 yield only rows 1–8 of Table 1 for Q_0 .

In a similar way it may be shown that the first neighborhood of the special quintet Q_1 consists of the four magnitudes

$$|E_h|, |E_k|, |E_l|, |E_{\frac{1}{2}(-h+k+l)}| \quad (\text{II.3})$$

and the second neighborhood of the four main terms (II.3) plus the additional seven cross terms

$$|E_{\frac{1}{2}(h+k+l)}|, |E_{\frac{1}{2}(h-k+l)}|, |E_{\frac{1}{2}(h+k-l)}|; \\ |E_{h-k}|, |E_{k+l}|, |E_{l+h}|; |E_{-h+k+l}|, \quad (\text{II.4})$$

i.e. eleven magnitudes $|E|$ in all. Now the 16 rows of the 1977 Table 2 for Q_1 lead to the eight rows 9–16 of

precisely two of the magnitudes (2.5) are large and the remaining two are small.

It should be noted in passing that a more detailed study of Table 2 leads in a similar way to the definition of the second neighborhoods of T , but this work is beyond the scope of the present paper.

APPENDIX IV

The integral formula for P , the joint probability distribution of the seven structure factors whose magnitudes constitute the first neighborhood of T

Suppose that (h,k,l) is the primitive random variable which is assumed to be uniformly distributed over the subset of $W \times W \times W$ defined by (2.2). In view of (2.2), the eight reciprocal vectors

$$\frac{1}{2}(\pm h \pm k \pm l) \tag{IV.1}$$

have integer components; but only four of the eight magnitudes

$$|E_{\frac{1}{2}(\pm h \pm k \pm l)}| \tag{IV.2}$$

are distinct. The seven structure factors

$$S_1 = E_h, \quad S_2 = E_k, \quad S_3 = E_l; \tag{IV.3}$$

$$s_1 = E_{\frac{1}{2}(h+k+l)}, \quad s_2 = E_{\frac{1}{2}(-h+k+l)},$$

$$s_3 = E_{\frac{1}{2}(h-k+l)}, \quad s_4 = E_{\frac{1}{2}(h+k-l)} \tag{IV.4}$$

as functions of the primitive random variable (h,k,l) , are themselves random variables. Denote by

$$P = P(S_1, S_2, S_3; s_1, s_2, s_3) \tag{IV.5}$$

the joint probability distribution of the seven structure factors (IV.3) and (IV.4). Then, following Karle & Hauptman (1958), P is given by the sevenfold integral

$$P = \frac{1}{(2\pi)^7} \int_{X_1, X_2, X_3, x_1, x_2, x_3 = -\infty}^{\infty} \exp \{ -i [S_1 X_1 + S_2 X_2 + S_3 X_3 + s_1 x_1 + s_2 x_2 + s_3 x_3] \} \times \prod_{j=1}^{N/2} g_j dX_1 dX_2 dX_3 dx_1 dx_2 dx_3, \tag{IV.6}$$

Table 2. The probable values of the structure seminvariant T , given the magnitudes $|E|$ in the second neighborhoods of its extensions; L means large; S means small; obtained by combining reinforcing rows in Table 1

Row	Derived from Rows of Table 1	T	Magnitudes $ E $																
			h	k	l	$\frac{1}{2}(h+k+l)$	$\frac{1}{2}(-h+k+l)$	$\frac{1}{2}(h-k+l)$	$\frac{1}{2}(h+k-l)$	$h+k$	$k+l$	$l+h$	$h-k$	$k-l$	$l-h$	$h+k+l$	$-h+k+l$	$h-k+l$	$h+k-l$
1	1,9,17,25	0	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L	L
2	2,10	π	L	L	L	L	L	S	S	L	S	L	L	L	S	S			
3	3,19	π	L	L	L	L	L	S	L	S	L	L	L	S			S		
4	4,28	π	L	L	L	L	L	S	S	L	S	L	L	L	S				S
5	11,27	π	L	L	L	S	L	L	S	L	L	L	L	S		S			S
6	12,20	π	L	L	L	S	L	L	S	L	L	S	L	L		S	S		
7	18,26	π	L	L	L	S	S	L	L	L	L	S	L	L			S	S	
8	2,14	π	L	L	L	L	L	L	S	S	L	S	L	S	S	L			
9	3,23	π	L	L	L	L	L	S	L	S	L	L	S	S	S		L		
10	4,32	π	L	L	L	L	L	S	S	L	S	L	L	S	S	S			L
11	6,10	π	L	L	L	L	L	L	S	S	S	L	L	L	L	S			
12	11,31	π	L	L	L	S	L	L	S	L	S	L	S	S		S			L
13	12,24	π	L	L	L	S	L	L	S	L	S	S	S	L		S	L		
14	18,30	π	L	L	L	S	S	L	L	S	L	L	S	S			S	L	
15	7,19	π	L	L	L	L	L	S	L	S	S	S	L	L	L		S		
16	16,20	π	L	L	L	S	L	L	S	L	S	L	S	L	S		L	S	
17	22,26	π	L	L	L	S	S	L	L	L	S	S	L	L			L	L	S
18	15,27	π	L	L	L	S	L	S	L	L	S	L	S	L	S		L		S
19	8,28	π	L	L	L	L	L	S	S	L	S	S	L	L	L				S
20	6,14	π	L	L	L	L	L	L	S	S	S	S	L	S	L	L			
21	7,23	π	L	L	L	L	L	S	L	S	S	S	S	S	L		L		
22	8,32	π	L	L	L	L	L	S	S	L	S	S	S	S	L				L
23	15,31	π	L	L	L	S	L	L	S	L	S	S	S	S		L			L
24	16,24	π	L	L	L	S	L	L	S	L	S	S	S	S		L	L		
25	22,30	π	L	L	L	S	S	L	L	L	S	S	S	S			L	L	

where

$$g_j = \left\langle \exp \left\{ \frac{2if_j}{\sigma_2^{1/2}} [X_1 \cos 2\pi\mathbf{h} \cdot \mathbf{r}_j + X_2 \cos 2\pi\mathbf{k} \cdot \mathbf{r}_j + X_3 \cos 2\pi\mathbf{l} \cdot \mathbf{r}_j + x \cos \pi(\mathbf{h} + \mathbf{k} + \mathbf{l}) \cdot \mathbf{r}_j + x_1 \cos \pi(-\mathbf{h} + \mathbf{k} + \mathbf{l}) \cdot \mathbf{r}_j + x_2 \cos \pi(\mathbf{h} - \mathbf{k} + \mathbf{l}) \cdot \mathbf{r}_j + x_3 \cos \pi(\mathbf{h} + \mathbf{k} - \mathbf{l}) \cdot \mathbf{r}_j] \right\} \right\rangle_{\mathbf{h}, \mathbf{k}, \mathbf{l}} \quad (IV.7)$$

and the average in (IV.7) is taken over all reciprocal vectors satisfying (2.2).

APPENDIX V

The derivation of g_j

Expanding the exponential and carrying out the indicated average in (IV.7), one finds after a short calculation that

$$g_j = \{1 - (f_j^2/\sigma_2)(X_1^2 + X_2^2 + X_3^2 + x^2 + x_1^2 + x_2^2 + x_3^2) - (2if_j^3/\sigma_2^{3/2})(X_1 xx_1 + X_2 xx_2 + X_3 xx_3 + X_1 x_2 x_3 + X_2 x_3 x_1 + X_3 x_1 x_2) + (2f_j^4/\sigma_2^2)(X_1 X_2 xx_3 + X_2 X_3 xx_1 + X_3 X_1 xx_2 + X_1 X_2 x_1 x_2 + X_2 X_3 x_2 x_3 + X_3 X_1 x_3 x_1 + xx_1 x_2 x_3) + (if_j^5/\sigma_2^{5/2}) X_1 X_2 X_3 \times (x^2 + x_1^2 + x_2^2 + x_3^2)\} \{1 + O(1/N^2)\}, \quad (V.1)$$

where $O(1/N^2)$ consists of all terms of order $1/N^2$ or higher which contribute only to terms of order higher than $1/N^{3/2}$ in the final result [(4.2)–(4.5)].

APPENDIX VI

The derivation of $\prod_{j=1}^{N/2} g_j$

Using (V.1), one readily calculates $\log g_j$, $\sum_{j=1}^{N/2} \log g_j$ and, finally, from the identity

$$\prod_{j=1}^{N/2} g_j = \exp \left\{ \sum_{j=1}^{N/2} \log g_j \right\}, \quad (VI.1)$$

one finds

$$\prod_{j=1}^{N/2} g_j = \exp \left\{ -\frac{1}{2}(X_1^2 + X_2^2 + X_3^2 + x^2 + x_1^2 + x_2^2 + x_3^2) \times \left[1 - (i\sigma_3/\sigma_2^{3/2})[X_1(xx_1 + x_2 x_3) + X_2(xx_2 + x_3 x_1) + X_3(xx_3 + x_1 x_2)] + (\sigma_4/\sigma_2^2)[X_1 X_2(xx_3 + x_1 x_2) + X_2 X_3(xx_1 + x_2 x_3) + X_3 X_1(xx_2 + x_3 x_1) + xx_1 x_2 x_3] - (\sigma_3^2/\sigma_2^3)\{X_1 X_2[xx_3(x_1^2 + x_2^2) + x_1 x_2(x_2^2 + x_3^2)] + X_2 X_3[xx_1(x_2^2 + x_3^2) + x_2 x_3(x_2^2 + x_1^2)] + X_3 X_1[xx_2(x_3^2 + x_1^2) + x_3 x_1(x_2^2 + x_3^2)] + (X_1^2 + X_2^2 + X_3^2)xx_1 x_2 x_3\} + (i\sigma_3/2\sigma_2^{5/2})X_1 X_2 X_3(x_2^2 + x_1^2 + x_2^2 + x_3^2) - (i\sigma_3 \sigma_4/\sigma_2^{7/2})X_1 X_2 X_3(x_2^2 x_1^2 + x_2^2 x_2^2 + x_2^2 x_3^2 + x_1^2 x_2^2 + x_2^2 x_3^2 + x_3^2 x_1^2) + (i\sigma_3^3/\sigma_2^{9/2})X_1 X_2 X_3(x_2^2 x_1^2 x_2^2 + x_2^2 x_2^2 x_3^2 + x_2^2 x_3^2 x_1^2 + x_1^2 x_2^2 x_3^2) \} \right] \{1 + O(1/N)\}, \quad (VI.2)$$

in which $O(1/N)$ consists of all terms of order $1/N$ or higher which contribute only to terms of order higher than $1/N^{3/2}$ in the final result [(4.2)–(4.5)].

Table 3. The probable value of the structure seminvariant T , given the seven magnitudes in its first neighborhood; L means large; S means small; obtained by combining reinforcing rows from the first seven columns (under magnitudes $|E|$) of Table 2

Row	Derived from rows of table 2	T	Magnitudes $ E $						
			\mathbf{h}	\mathbf{k}	\mathbf{l}	$\frac{1}{2}(\mathbf{h} + \mathbf{k} + \mathbf{l})$	$\frac{1}{2}(-\mathbf{h} + \mathbf{k} + \mathbf{l})$	$\frac{1}{2}(\mathbf{h} - \mathbf{k} + \mathbf{l})$	$\frac{1}{2}(\mathbf{h} + \mathbf{k} - \mathbf{l})$
1	1	0	L	L	L	L	L	L	L
2	2, 8, 11, 20	π	L	L	L	L	L	S	S
3	3, 9, 15, 21	π	L	L	L	L	S	L	S
4	4, 10, 19, 22	π	L	L	L	L	S	S	L
5	5, 12, 18, 23	π	L	L	L	S	L	S	L
6	6, 13, 16, 24	π	L	L	L	S	L	L	S
7	7, 14, 17, 25	π	L	L	L	S	S	L	L

APPENDIX VII

The derivation of P

Substituting from (VI.2) into (IV.6) and carrying out the indicated sevenfold integration, one finds

$$\begin{aligned}
P = & \frac{1}{(2\pi)^{7/2}} \exp \left\{ -\frac{1}{2}(S_1^2 + S_2^2 + S_3^2 \right. \\
& + s^2 + s_1^2 + s_2^2 + s_3^2) \} \\
& \times \left\{ 1 + (\sigma_3/\sigma_2^{3/2})[S_1(ss_1 + s_2 s_3) \right. \\
& + S_2(ss_2 + s_3 s_1) + S_3(ss_3 + s_1 s_2)] \\
& + (\sigma_4/\sigma_2^2)[S_1 S_2(ss_3 + s_1 s_2) + S_2 S_3(ss_1 + s_2 s_3) \\
& + S_3 S_1(ss_2 + s_3 s_1) + ss_1 s_2 s_3] \\
& + (\sigma_3^2/\sigma_2^3) \{ S_1 S_2[ss_3(s_1^2 + s_2^2 - 2) \\
& + s_1 s_2(s_2^2 + s_3^2 - 2)] \\
& + S_2 S_3[ss_1(s_2^2 + s_3^2 - 2) + s_2 s_3(s_2^2 + s_1^2 - 2)] \\
& + S_3 S_1[ss_2(s_3^2 + s_1^2 - 2) + s_3 s_1(s_2^2 + s_2^2 - 2)] \\
& + ss_1 s_2 s_3(S_1^2 + S_2^2 + S_3^2 - 3) \} \\
& + (\sigma_3/2\sigma_2^{5/2})S_1 S_2 S_3(s^2 + s_1^2 + s_2^2 + s_3^2 - 4) \\
& + (\sigma_3 \sigma_4/\sigma_2^{7/2})S_1 S_2 S_3[(s^2 - 1)(s_1^2 + s_2^2 + s_3^2 - 3) \\
& + (s_1^2 - 1)(s_2^2 - 1) + (s_2^2 - 1)(s_3^2 - 1) \\
& + (s_3^2 - 1)(s_1^2 - 1)] \\
& + (\sigma_3^3/\sigma_2^9)S_1 S_2 S_3[(s^2 - 1)(s_1^2 - 1)(s_2^2 - 1) \\
& + (s^2 - 1)(s_2^2 - 1)(s_3^2 - 1) \\
& + (s^2 - 1)(s_3^2 - 1)(s_1^2 - 1) \\
& + (s_1^2 - 1)(s_2^2 - 1)(s_3^2 - 1)] \left. \right\} \{ 1 + O(1/N) \}, \tag{VII.1}
\end{aligned}$$

where $O(1/N)$ consists of all terms of order $1/N$ or higher which contribute only to terms of order higher than $1/N^{3/2}$ in the final result [(4.2)–(4.5)].

Next, using (VII.1) to calculate $\log P$, and the identity

$$P = \exp(\log P), \tag{VII.2}$$

one finally finds the joint probability distribution, in pure exponential form, of the seven structure factors (IV.3) and (IV.4):

$$\begin{aligned}
P = & \frac{1}{(2\pi)^{7/2}} \exp \left\{ -\frac{1}{2}(S_1^2 + S_2^2 + S_3^2 \right. \\
& + s^2 + s_1^2 + s_2^2 + s_3^2) \} \exp \{ (\sigma_3/\sigma_2^{3/2})[S_1(ss_1 + s_2 s_3) \\
& + S_2(ss_2 + s_3 s_1) + S_3(ss_3 + s_1 s_2)] \\
& - [(2\sigma_3^2 - \sigma_2 \sigma_4/\sigma_2^3)[S_2 S_3(ss_1 + s_2 s_3) \\
& + S_3 S_1(ss_2 + s_3 s_1) + S_1 S_2(ss_3 + s_1 s_2)] \\
& + (1/2\sigma_2^9)(6\sigma_3^3 - 6\sigma_2 \sigma_3 \sigma_4 + \sigma_2^2 \sigma_5) \}
\end{aligned}$$

$$\begin{aligned}
& \times S_1 S_2 S_3(s^2 + s_1^2 + s_2^2 + s_3^2) \\
& - (2/\sigma_2^{9/2})(2\sigma_3^3 - 3\sigma_2 \sigma_3 \sigma_4 + \sigma_2^2 \sigma_5)S_1 S_2 S_3 \} \\
& \times \{ 1 + O(1/N) \}, \tag{VII.3}
\end{aligned}$$

where again $O(1/N)$ consists of terms of order $1/N$ or higher which contribute only to terms of order higher than $1/N^{3/2}$ in the final result [(4.2)–(4.5)]. Presumably (VII.3) is a more accurate representation of P than (VII.1) since it is non-negative everywhere [a property not shared by (VII.1)] and agrees with (VII.1), except possibly for an irrelevant normalizing factor, up to and including terms of order $1/N^{3/2}$.

APPENDIX VIII

The joint conditional probability distribution of the three phases ϕ_h, ϕ_k, ϕ_l , given the seven magnitudes (2.4) and (2.5)

With the probabilistic background described in §3, the phases ϕ_h, ϕ_k and ϕ_l of the structure factors E_h, E_k and E_l , respectively, are functions of the primitive random variable $(\mathbf{h}, \mathbf{k}, \mathbf{l})$, and therefore are themselves random variables. Denote by

$$P_3 = P(\Phi_1, \Phi_2, \Phi_3 | R_1, R_2, R_3; r, r_1, r_2, r_3) \tag{VIII.1}$$

the joint conditional probability distribution of the three phases ϕ_h, ϕ_k, ϕ_l , given (2.2), (2.4) and (2.5). Then P_3 is obtained from (VII.3) by fixing the magnitudes of $S_1, S_2, S_3; s, s_1, s_2, s_3$ in accordance with the scheme

$$|S_1| = R_1, \quad |S_2| = R_2, \quad |S_3| = R_3; \tag{VIII.2}$$

$$|s| = r, \quad |s_1| = r_1, \quad |s_2| = r_2, \quad |s_3| = r_3, \tag{VIII.3}$$

i.e.

$$S_1 = R_1 \cos \Phi_1, \quad S_2 = R_2 \cos \Phi_2, \quad S_3 = R_3 \cos \Phi_3; \tag{VIII.4}$$

$$s = r \cos \phi, \quad s_1 = r_1 \cos \phi_1, \quad s_2 = r_2 \cos \phi_2, \quad s_3 = r_3 \cos \phi_3, \tag{VIII.5}$$

where ϕ is the variable associated with the phase $\phi_{\frac{1}{2}(h+k+l)}$ ϕ_1 with $\phi_{\frac{1}{2}(-h+k+l)}$ etc., summing with respect to s, s_1, s_2, s_3 over their two possible signs (+ and -) or, equivalently, summing with respect to $\phi, \phi_1, \phi_2, \phi_3$ over their two possible values (0 and π), and multiplying the result by a suitable normalizing factor. Carrying out these summations one finally obtains, correct up to and including terms of order $1/N^{3/2}$,

$$P_3 = \frac{1}{K_3} Z_3, \tag{VIII.6}$$

where

$$\begin{aligned}
 Z_3 = & \exp \{V \cos (\Phi_1 + \Phi_2 + \Phi_3)\} \\
 & \times \sum_{\varepsilon, \varepsilon_1, \varepsilon_2, \varepsilon_3 = \pm 1} \exp \{(\sigma_3 / \sigma_2^{3/2}) \\
 & \times [(\varepsilon \varepsilon_1 r r_1 + \varepsilon_2 \varepsilon_3 r_2 r_3) R_1 \cos \Phi_1 \\
 & + (\varepsilon \varepsilon_2 r r_2 + \varepsilon_3 \varepsilon_1 r_3 r_1) R_2 \cos \Phi_2 \\
 & + (\varepsilon \varepsilon_3 r r_3 + \varepsilon_1 \varepsilon_2 r_1 r_2) R_3 \cos \Phi_3] \\
 & - (1 / \sigma_2^3)(2 \sigma_3^2 - \sigma_2 \sigma_4) \\
 & \times [(\varepsilon \varepsilon_1 r r_1 + \varepsilon_2 \varepsilon_3 r_2 r_3) R_2 R_3 \cos (\Phi_2 + \Phi_3) \\
 & + (\varepsilon \varepsilon_2 r r_2 + \varepsilon_3 \varepsilon_1 r_3 r_1) R_3 R_1 \cos (\Phi_3 + \Phi_1) \\
 & + (\varepsilon \varepsilon_3 r r_3 + \varepsilon_1 \varepsilon_2 r_1 r_2) R_1 R_2 \cos (\Phi_1 + \Phi_2)]\}.
 \end{aligned} \tag{VIII.7}$$

V is defined by (4.4) and (VIII.7) is a sum of the 16 exponentials obtained by permitting each of $\varepsilon, \varepsilon_1, \varepsilon_2, \varepsilon_3$ to take on the values $+1$ and -1 independently. The normalizing parameter K_3 is not relevant for the present purpose and is therefore not derived here explicitly.

Under the transformation

$$\eta = \varepsilon \cos (\Phi_1 + \Phi_2 + \Phi_3), \quad \eta_j = \varepsilon_j \cos \Phi_j, \quad j = 1, 2, 3, \tag{VIII.8}$$

(VIII.7) becomes

$$\begin{aligned}
 Z_3 = & \exp \{V \cos (\Phi_1 + \Phi_2 + \Phi_3)\} \\
 & \times \sum_{\eta, \eta_1, \eta_2, \eta_3 = \pm 1} \exp \{(\sigma_3 / \sigma_2^{3/2}) \\
 & \times [(\eta \eta_1 r r_1 + \eta_2 \eta_3 r_2 r_3) R_1 + (\eta \eta_2 r r_2 + \eta_3 \eta_1 r_3 r_1) R_2 \\
 & + (\eta \eta_3 r r_3 + \eta_1 \eta_2 r_1 r_2) R_3] \\
 & \times \cos (\Phi_1 + \Phi_2 + \Phi_3) - (1 / \sigma_2^3)(2 \sigma_3^2 - \sigma_2 \sigma_4) \\
 & \times [(\eta \eta_1 r r_1 + \eta_2 \eta_3 r_2 r_3) R_2 R_3 \\
 & + (\eta \eta_2 r r_2 + \eta_3 \eta_1 r_3 r_1) R_3 R_1 \\
 & + (\eta \eta_3 r r_3 + \eta_1 \eta_2 r_1 r_2) R_1 R_2]\},
 \end{aligned} \tag{VIII.9}$$

so that the dependence on the structure seminvariant $\Phi_1 + \Phi_2 + \Phi_3$ is explicit. In view of the transformation (VIII.8), (VIII.9) is also a sum of 16 exponentials obtained by permitting each of $\eta, \eta_1, \eta_2, \eta_3$ to take on the values $+1$ and -1 independently.

Replacing $\eta, \eta_1, \eta_2, \eta_3$ simultaneously by their negatives leaves the argument of the exponential in (VIII.9) unchanged. Hence only eight of the 16 summands in (VIII.9) are distinct, the summands being equal in pairs. Thus η may be set equal to $+1$ in

(VIII.9), and Z_3 finally becomes simply

$$\begin{aligned}
 Z_3 = & \exp \{V \cos (\Phi_1 + \Phi_2 + \Phi_3)\} \\
 & \times \sum_{\eta_1, \eta_2, \eta_3 = \pm 1} \exp \{(\sigma_3 / \sigma_2^{3/2}) \\
 & \times [(\eta_1 r r_1 + \eta_2 \eta_3 r_2 r_3) R_1 \\
 & + (\eta_2 r r_2 + \eta_3 \eta_1 r_3 r_1) R_2 \\
 & + (\eta_3 r r_3 + \eta_1 \eta_2 r_1 r_2) R_3] \cos (\Phi_1 + \Phi_2 + \Phi_3) \\
 & - (1 / \sigma_2^3)(2 \sigma_3^2 - \sigma_2 \sigma_4) \\
 & \times [(\eta_1 r r_1 + \eta_2 \eta_3 r_2 r_3) R_2 R_3 \\
 & + (\eta_2 r r_2 + \eta_3 \eta_1 r_3 r_1) R_3 R_1 \\
 & + (\eta_3 r r_3 + \eta_1 \eta_2 r_1 r_2) R_1 R_2]\},
 \end{aligned} \tag{VIII.10}$$

in which the summands consist of the eight exponentials obtained by permitting each of η_1, η_2, η_3 to take on the values $+1$ and -1 independently.

In view of (VIII.10), P_3 [(VIII.6)] is seen to be a function of the structure seminvariant $\Phi_1 + \Phi_2 + \Phi_3$. Hence (VIII.6) leads immediately to the major result of this paper [(4.2)–(4.5)], the conditional probability distribution, P^\pm , of the three-phase structure seminvariant T , given the seven magnitudes $|E|$ in its first neighborhood.

References

- FORTIER, S. & HAUPTMAN, H. (1977). *Acta Cryst.* **A33**, 829–833.
- GIACOVAZZO, C. (1975). *Acta Cryst.* **A31**, 252–259.
- GIACOVAZZO, C. (1977). *Acta Cryst.* **A33**, 933–944.
- GIACOVAZZO, C. (1978a). *Acta Cryst.* **A34**, 27–30.
- GIACOVAZZO, C. (1978b). *Acta Cryst.* **A34**, 562–564.
- GREEN, E. & HAUPTMAN, H. (1976). *Acta Cryst.* **A32**, 940–944.
- GREEN, E. & HAUPTMAN, H. (1978). *Acta Cryst.* **A34**, 216–223.
- HAUPTMAN, H. (1975). *Acta Cryst.* **A31**, 680–687.
- HAUPTMAN, H. (1977a). *Acta Cryst.* **A33**, 568–571.
- HAUPTMAN, H. (1977b). Michigan State Univ. Meeting of Am. Crystallogr. Soc., August 7–12, Abstract H3.
- HAUPTMAN, H. (1978). *Acta Cryst.* **A34**, 525–528.
- HAUPTMAN, H. (1979). *Proc. Natl Acad. Sci. USA*, **76**, No. 10, 4747–4749.
- HAUPTMAN, H. & GREEN, E. (1976). *Acta Cryst.* **A32**, 45–49.
- HAUPTMAN, H. & GREEN, E. (1978). *Acta Cryst.* **A34**, 224–229.
- HAUPTMAN, H. & POTTER, S. (1979). *Acta Cryst.* **A35**, 371–381.
- HEINERMAN, J. J. L., KRABBENDAM, H. & KROON, J. (1977). *Acta Cryst.* **A33**, 873–878.
- KARLE, J. & HAUPTMAN, H. (1958). *Acta Cryst.* **11**, 264–269.
- SCHENK, H. (1975). *Acta Cryst.* **A31**, S14.