in the first approach and the reciprocal-lattice vectors in the second see Heinerman (1977*a*; 1977*b*, ch. IV) and Heinerman, Krabbendam & Kroon (1977). As discussed earlier (Heinerman, 1977*a*) the approach in which the atomic position vectors are the primitive random variables opens the possibility of including structural information.

We are very much indebted to Professor Dr F. van der Blij of the Mathematical Institute of the Rijksuniversiteit Utrecht for discussions on mathematical problems.

References

HEINERMAN, J. J. L. (1975). *Acta Cryst.* A**31**, 727–730.
HEINERMAN, J. J. L. (1977*a*). *Acta Cryst.* A**33**, 100–106.
HEINERMAN, J. J. L. (1977*b*). Thesis, Utrecht.
HEINERMAN, J. J. L., KRABBENDAM, H. & KROON, J. (1977). *Acta Cryst.* A**33**, 873–878.
KARLE, J. & HAUPTMAN, H. (1950). *Acta Cryst.* **3**, 181–187.
KARLE, J & HAUPTMAN, H. (1958). *Acta Cryst.* **11**, 264–269.
RANGO, C. DE (1969). Thesis, Paris.
TSOUCARIS, G. (1970). *Acta Cryst.* A**26**, 492–499.
WATSON, G. N. (1966). *A Treatise on the Theory of Bessel Functions.* Cambridge Univ. Press.

Acta Cryst. (1979). A**35**, 105–107

Conditional Phase Probability Distributions of Structure Factors in a Karle–Hauptman Matrix*

BY J. J. L. HEINERMAN,† J. KROON‡ AND H. KRABBENDAM

Laboratorium voor Structuurchemie, Rijksuniversiteit, Padualaan 8, *Utrecht, The Netherlands*

(*Received* 24 *February* 1978; *accepted* 4 *July* 1978)

Abstract

From the joint probability distribution of all structure factors in a Karle–Hauptman matrix new phase probability distributions are obtained. These calculations lead to a reformulation of the maximum-determinant rule for phase determination. In addition a new function is derived whose maximum corresponds to the most probable values for the phases of an arbitrary subset of the structure factors in a Karle–Hauptman matrix. This function accounts for the interaction among phases in a Karle–Hauptman matrix through triple products and quartets simultaneously.

Introduction

The maximum-determinant rule for phase determination (de Rango, 1969; Tsoucaris, 1970) has been derived from the joint probability distribution of the normalized structure factors $E_{\mathbf{h}_0-\mathbf{h}_m}$, $E_{\mathbf{h}_1-\mathbf{h}_m}$, ... and $E_{\mathbf{h}_{m-1}-\mathbf{h}_m}$ where $\mathbf{h}_0, \ldots, \mathbf{h}_{m-1}$ are fixed and $\mathbf{h}_m$ is the primitive random variable. Since the $\mathbf{h}_i$ $(i = 0, \ldots, m-1)$ are fixed the $E_{\mathbf{h}_i-\mathbf{h}_j}$ $(i, j = 0, \ldots, m-1)$, which enter into the joint probability distribution of $E_{\mathbf{h}_i-\mathbf{h}_m}$ $(i = 0, \ldots, m-1)$, are also fixed. Therefore their magnitudes and phases should be known before the maximum-determinant rule can be applied. In practice, if some of the $E_{\mathbf{h}_i-\mathbf{h}_j}$ are unknown, they are set equal to zero (Castellano, Podjarny & Navaza, 1973; de Rango, Mauguen & Tsoucaris, 1975; Podjarny, Yonath & Traub, 1976).

The present paper gives the derivation of a new function whose maximum, by analogy with the maximum-determinant rule, corresponds to the most probable values for structure-factor phases, but which does not require knowledge of all $E_{\mathbf{h}_i-\mathbf{h}_j}$ $(i, j = 0, \ldots, m-1)$. The basis of all our calculations is the result of the preceding paper, *viz* the joint probability distribution of all structure factors in a Karle–Hauptman matrix (Karle & Hauptman, 1950) for structures in space group $P1$. With this distribution we shall perform the following calculations. (i) By fixing the magnitudes of the structure factors we obtain a function of phases which is closely related to the phase-dependent terms that appear in the evaluation of a Karle–Hauptman determinant. (ii) Integrations with respect to an arbitrary set of phases are performed; next the magnitudes of the structure factors are fixed. This leads to a function of phases which is related to that in (i) but which depends only on an arbitrary subset of the phases in a Karle–Hauptman matrix. (iii) In addition to the integrations in (ii) we perform the integrations with respect to the structure-factor magnitudes that correspond to an arbitrary subset of the phases for which the integrations have been performed; next, the remaining magnitudes are fixed. This leads to a function of phases

* Presented at the Eleventh International Congress of Crystallography, Warsaw, Poland, 3–12 August 1978, Abstract 03.2-14.

† Present address: Koninklijke/Shell-laboratorium, Badhuisweg 3, Amsterdam-N, The Netherlands.

‡ To whom correspondence should be addressed.

0567-7394/79/010105-03$01.00

which is related to that in (ii) but which does not take into account the magnitudes of structure factors that correspond to an arbitrary subset of the excluded phases.

For notation which is not explained, see the preceding paper (Heinerman, Krabbendam & Kroon, 1979).

$P(\Phi_{lm}|R_{lm})$

The conditional joint probability distribution $P(\Phi_{lm}|R_{lm})$ is easily obtained from $P(R_{lm};\ \Phi_{lm})$ in the preceding paper by fixing the R's. The result, correct up to and including terms of order $1/N$, is

$$P(\Phi_{lm}|R_{lm}) = C_1 \exp(M_1), \qquad (1)$$

in which

$$M_1 = 2\frac{\sigma_3}{\sigma_2^{3/2}} \sum^{m}{}'' \, T_{i_1i_2i_3} - 2\frac{\sigma_3^2}{\sigma_2^3} \times \sum^{m}{}''' \, (Q_{i_1i_2i_3i_4} + Q_{i_1i_2i_4i_3} + Q_{i_1i_3i_2i_4}) \qquad (2)$$

and C_1 is a normalizing constant.

$P(\Phi_{lm}$ except n Φ's$|R_{lm})$

The conditional joint probability distribution $P(\Phi_{lm}$ except n Φ's$|R_{lm})$ is calculated from $P(R_{lm};\ \Phi_{lm})$ in the preceding paper: first the integrations with respect to the $\Phi_{p_iq_i}$ (in which p_iq_i stands for $\mathbf{h}_{p_i} - \mathbf{h}_{q_i}$) $(i = 1, \ldots, n)$ are performed and next the R's are fixed.

Employing the sum of cosines formula [preceding paper, formula (5)] and

$$\int_0^{2\pi} \exp(2z\cos\theta)\,\mathrm{d}\theta = 2\pi[\exp(z^2)][1 + O(z^4)] \qquad (3)$$

(Heinerman, 1977, p. 33) the following formula for $P(R_{lm};\ \Phi_{lm}$ except n Φ's) is obtained (the proof is given by mathematical induction):

$$P(R_{lm};\ \Phi_{lm} \text{ except } n\ \Phi\text{'s}) = \frac{(2\pi)^n \prod^{m}{}' R_{i_1i_2}}{\pi^{m(m+1)/2}} \times \exp\left(-\sum^{m}{}' R^2_{i_1i_2} + M_2\right)\left[1 + O'\left(\frac{1}{N}\right)\right], \qquad (4)$$

where

$$\begin{aligned} M_2 = {} & 2\frac{\sigma_3}{\sigma_2^{3/2}} \sum^{m}_{A_{p_1q_1},\ldots,A_{p_nq_n}}{}'' \, T_{i_1i_2i_3} \\ & - 2\frac{\sigma_3^2}{\sigma_2^3}\Bigg[\sum^{m}_{B'_{p_1q_1},\ldots,B'_{p_nq_n}}{}''' \, Q_{i_1i_2i_3i_4} - \sum_{i=1}^{n} R^2_{p_iq_i} \\ & \times \Bigg(\sum^{m}_{\substack{i_2=0\\ i_1=p_i<i_2<i_3=q_i<i_4\\ B'_{p_1q_1},\ldots,B'_{p_nq_n}}} \sum^{m}_{i_4=0} Q_{p_ii_2q_ii_4} + \sum^{m}_{\substack{i_1=0\\ i_1<i_2=p_i<i_3<i_4=q_i\\ B'_{p_1q_1},\ldots,B'_{p_nq_n}}} \sum^{m}_{i_3=0} Q_{i_ip_ii_3q_i}\Bigg) \\ & + \sum^{m}_{B''_{p_1q_1},\ldots,B''_{p_nq_n}}{}''' \, Q_{i_1i_2i_4i_3} - \sum_{i=1}^{n} R^2_{p_iq_i} \\ & \times \Bigg(\sum^{m}_{\substack{i_2=0\\ i_1=p_i<i_2<i_3<i_4=q_i\\ B''_{p_1q_1},\ldots,B''_{p_nq_n}}} \sum^{m}_{i_3=0} Q_{p_ii_2q_ii_3} + \sum^{m}_{\substack{i_1=0\\ i_1<i_2=p_i<i_3=q_i<i_4\\ B''_{p_1q_1},\ldots,B''_{p_nq_n}}} \sum^{m}_{i_4=0} Q_{i_1p_ii_4q_i}\Bigg) \\ & + \sum^{m}_{B'''_{p_1q_1},\ldots,B'''_{p_nq_n}}{}''' \, Q_{i_1i_3i_2i_4} - \sum_{i=1}^{n} R^2_{p_iq_i} \\ & \times \Bigg(\sum^{m}_{\substack{i_3=0\\ i_1=p_i<i_2=q_i<i_3<i_4\\ B'''_{p_1q_1},\ldots,B'''_{p_nq_n}}} \sum^{m}_{i_4=0} Q_{p_ii_3q_ii_4} + \sum^{m}_{\substack{i_1=0\\ i_1<i_2<i_3=p_i<i_4=q_i\\ B'''_{p_1q_1},\ldots,B'''_{p_nq_n}}} \sum^{m}_{i_2=0} Q_{i_1p_ii_2q_i}\Bigg)\Bigg], \end{aligned} \qquad (5)$$

where $A_{p_iq_i}$ denotes $i_1i_2,\ i_2i_3, i_1i_3 \neq p_iq_i$; $B'_{p_iq_i}$ denotes $i_1i_2,\ i_2i_3,\ i_3i_4,\ i_1i_4 \neq p_iq_i$; $B''_{p_iq_i}$ denotes $i_1i_2,\ i_2i_4,\ i_3i_4,\ i_1i_3 \neq p_iq_i$; and $B'''_{p_iq_i}$ denotes $i_1i_3,\ i_2i_3,\ i_2i_4,\ i_1i_4 \neq p_iq_i$. The conditions $A_{p_iq_i}$, $B'_{p_iq_i}$, $B''_{p_iq_i}$ and $B'''_{p_iq_i}$ in (5) exclude the triple products and quartets that depend on the Φ's for which the integrations have been performed. Next, fixing the R's we obtain, correct up to and including terms of order $1/N$,

$$P(\Phi_{lm} \text{ except } n\ \Phi\text{'s}|R_{lm}) = C_2 \exp(M_2), \qquad (6)$$

where C_2 is a normalizing constant.

$P(\Phi_{lm}$ except n Φ's$|R_{lm}$ except p R's)

The conditional joint probability distribution $P(\Phi_{lm}$ except n Φ's$|R_{lm}$ except p R's) with $p \leq n$ and the p R's corrsponding to Φ's that are excluded is calculated from (4): first the integrations with respect to the p R's are performed and next the remaining R's are fixed.

Employing

$$\int_0^{\infty} R\exp[-(1-a)R^2]\,\mathrm{d}R = \frac{1}{2(1-a)} = \tfrac{1}{2}(\exp a)[1 + O(a^2)] \qquad (7)$$

for the integrations with respect to the p R's it can easily be seen that

$$P(R_{lm} \text{ except } p\ R\text{'s};\ \Phi_{lm} \text{ except } n\ \Phi\text{'s}) = \frac{2^{n-p}\pi^n \prod^{m}_{\text{except } p\ R\text{'s}}{}' R_{i_1i_2}}{\pi^{m(m+1)/2}} \times \exp\left(-\sum^{m}_{\substack{\text{except}\\ p \text{ terms}}}{}' R^2_{i_1i_2} + M_3\right)\left[1 + O'\left(\frac{1}{N}\right)\right], \qquad (8)$$

in which M_3 is given by the expression M_2 with the $R^2_{p_iq_i}$ corresponding to excluded R's replaced by 1 [which is the mean value of $|E|^2$; *cf.* Heinerman (1977, ch. IV) and Heinerman, Krabbendam & Kroon (1977)]. Next, fixing the remaining R's we obtain, correct up to and including terms of order $1/N$,

$$P(\Phi_{lm} \text{ except } n\ \Phi\text{'s} | R_{lm} \text{ except } p\ R\text{'s}) = C_3 \exp(M_3), \quad (9)$$

where C_3 is a normalizing constant.

Example

As an illustrative example we consider the following case: $m = 3$; $n = 2$, the excluded Φ's are Φ_{02} and Φ_{13}; $p = 1$, the excluded R being R_{13}. We first calculate M_2:

$$\sum_{A_{02}, A_{13}}^{3}{}'' T_{i_1 i_2 i_3} = 0,$$

$$\sum_{B'_{02}, B'_{13}}^{3}{}''' Q_{i_1 i_2 i_3 i_4} = Q_{0123},$$

$$\sum_{i=1}^{2} R^2_{p_i q_i} \sum_{\substack{i_2=0 \\ i_1=p_i<i_2<i_3=q_i<i_4 \\ B'_{02}, B'_{13}}}^{3} \sum_{i_4=0}^{3} Q_{p_i i_2 q_i i_4} = R^2_{02} Q_{0123},$$

$$\sum_{i=1}^{2} R^2_{p_i q_i} \sum_{\substack{i_1=0 \\ i_1<i_2=p_i<i_3<i_4=q_i \\ B'_{02}, B'_{13}}}^{3} \sum_{i_3=0}^{3} Q_{i_1 p_i i_3 q_i} = R^2_{13} Q_{0123}$$

and the remaining terms in (5) are equal to zero. Then

$$M_2 = -2 \frac{\sigma_3^2}{\sigma_2^3} (1 - R^2_{02} - R^2_{13}) Q_{0123}. \quad (10)$$

Next M_3 is found from M_2 by substituting $R^2_{13} = 1$,

$$M_3 = 2 \frac{\sigma_3^2}{\sigma_2^3} R^2_{02} Q_{0123}. \quad (11)$$

Finally, from (9) to (11) we obtain

$$P(\Phi_{01}, \Phi_{12}, \Phi_{03}, \Phi_{23} | R_{01}, R_{02}, R_{12}, R_{03}, R_{23}) = C_3 \exp\left(2 \frac{\sigma_3^2}{\sigma_2^3} R^2_{02} Q_{0123}\right). \quad (12)$$

We note that by integrating (12) with respect to Φ_{01}, Φ_{12}, Φ_{03}, and Φ_{23}, in such a way that $\Phi_{01} + \Phi_{12} + \Phi_{23} - \Phi_{03} = \Phi$, we obtain the conditional probability distribution of the phase of a quartet given the magnitudes of the structure factors that form the quartet and the magnitude of one cross term,

$$P(\Phi | R_{01}, R_{02}, R_{12}, R_{03}, R_{23}) = C'_3 \exp\left(2 \frac{\sigma_3^2}{\sigma_2^3} R^2_{02} R_{01} R_{12} R_{23} R_{03} \cos \Phi\right). \quad (13)$$

For equal-atom structures (13) reduces to (13) of Heinerman (1977, p. 30).

Discussion

The function M_1 depends on the phases of all structure factors in a Karle–Hauptman matrix. The most probable values for the phases are those for which M_1 takes on its maximum value. From (15) of the preceding paper it can be seen that M_1 contains all triple products and all quartets which appear in the evaluation of a Karle–Hauptman determinant. [Strictly speaking U_n in (15) of the preceding paper is not a Karle–Hauptman determinant but reduces to it for equal-atom structures.] Moreover, for large N and $m \ll N$, which has been assumed throughout our calculations, these terms are the most important, so the maximum-determinant rule for phase determination may now be reformulated as: the most probable values for the phases of all structure factors in a Karle–Hauptman matrix are those for which the determinant of the Karle–Hauptman matrix takes on its maximum value. In fact, this formulation has been used by Woolfson (1977) in finding the most probable values for the phases in a 'magic determinant'.

If among the $\varphi_{\mathbf{h}_i - \mathbf{h}_j}$ $(i, j = 0, \ldots, m)$ there are phases that we do not wish to consider we may use M_2 or M_3. M_2 is a function of all phases in a Karle–Hauptman matrix except an arbitrary subset of them. To be able to employ M_2 one may still need magnitudes corresponding to excluded phases. This implies that unknown magnitudes may hamper the use of M_2. However, the function M_3 shows how to deal with these magnitudes, *viz* put their squares (only their squares appear in M_2) equal to 1. In fact, M_1 and M_2 are special cases of M_3.

Finally we would like to point out that if one starts with the structure factors from which a subset of the triple products in a Karle–Hauptman determinant is formed, M_3 also accounts for the associated quartets (quartets that are formed from the same set of structure factors as the triple products and that appear in the Karle–Hauptman determinant) and may even contain magnitudes of those structure factors in a Karle–Hauptman matrix which are not used to form the subset of triple products.

We are very much indebted to Professor Dr F. van der Blij of the Mathematical Institute of the Rijksuniversiteit Utrecht for discussions on mathematical problems.

References

CASTELLANO, E. E., PODJARNY, A. D. & NAVAZA, J. (1973). *Acta Cryst.* A**29**, 609–615.

HEINERMAN, J. J. L. (1977). Thesis, Utrecht.

HEINERMAN, J. J. L., KRABBENDAM, H. & KROON, J. (1977). *Acta Cryst.* A**33**, 873–878.

HEINERMAN, J. J. L., KRABBENDAM, H. & KROON, J. (1979). *Acta Cryst.* A**35**, 101–105.

KARLE, J. & HAUPTMAN, H. (1950). *Acta Cryst.* **3**, 181–187.

PODJARNY, A. D., YONATH, A. & TRAUB, W. (1976). *Acta Cryst.* A**32**, 281–292.

RANGO, C. DE (1969). Thesis, Paris.

RANGO, C. DE, MAUGUEN, Y. & TSOUCARIS, G. (1975). *Acta Cryst.* A**31**, 227–233.

TSOUCARIS, G. (1970). *Acta Cryst.* A**26**, 492–499.

WOOLFSON, M. M. (1977). *Acta Cryst.* A**33**, 219–225.