Acta Cryst. (1977). A**33**, 694–696

Sextets in $P1$: The Joint Probability Distribution of Thirty-One Structure Factors

By Suzanne Fortier and Herbert Hauptman

Medical Foundation of Buffalo, 73 *High Street*, *Buffalo*, *NY* 14203, *USA*

(*Received* 16 *February* 1977; *accepted* 18 *April* 1977)

It is assumed that a crystal structure in $P1$ is fixed and that the random variables (vectors) **h**, **k**, **l**, **m**, **n**, **p** are uniformly and independently distributed over the subset of reciprocal space defined by $\mathbf{h}+\mathbf{k}+\mathbf{l}+\mathbf{m}+\mathbf{n}+\mathbf{p}=0$. Then the 31 structure factors $E_{\mathbf{h}}, E_{\mathbf{k}},\ldots,E_{\mathbf{p}}$; $E_{\mathbf{h}+\mathbf{k}}, E_{\mathbf{h}+\mathbf{l}},\ldots,E_{\mathbf{n}+\mathbf{p}}$; $E_{\mathbf{h}+\mathbf{k}+\mathbf{l}}, E_{\mathbf{h}+\mathbf{k}+\mathbf{m}},\ldots,E_{\mathbf{h}+\mathbf{n}+\mathbf{p}}$, as functions of the primitive random variables **h**, **k**, **l**, **m**, **n**, **p**, are themselves random variables, and their joint probability distribution is found. This distribution plays the central role in the theory and estimation of the six-phase structure invariant $\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{\mathbf{l}}+\varphi_{\mathbf{m}}+\varphi_{\mathbf{n}}+\varphi_{\mathbf{p}}$.

1. Introduction

Recent methods and results (Hauptman, 1975, 1976, 1977; Fortier & Hauptman, 1977) are here carried over without essential change to initiate the study of sextets, the six-phase structure invariants

$$\varphi=\varphi_{\mathbf{h}}+\varphi_{\mathbf{k}}+\varphi_{\mathbf{l}}+\varphi_{\mathbf{m}}+\varphi_{\mathbf{n}}+\varphi_{\mathbf{p}}\,, \tag{1.1}$$

where

$$\mathbf{h}+\mathbf{k}+\mathbf{l}+\mathbf{m}+\mathbf{n}+\mathbf{p}=0\,. \tag{1.2}$$

In analogy with this earlier work one defines the first neighborhood of φ to consist of the six magnitudes

$$|E_{\mathbf{h}}|, |E_{\mathbf{k}}|, |E_{\mathbf{l}}|, |E_{\mathbf{m}}|, |E_{\mathbf{n}}|, |E_{\mathbf{p}}|\,. \tag{1.3}$$

The second neighborhood consists of the six magnitudes (1.3) and the 25 additional magnitudes

$$\begin{aligned}&|E_{\mathbf{h}+\mathbf{k}}|, |E_{\mathbf{h}+\mathbf{l}}|, |E_{\mathbf{h}+\mathbf{m}}|, |E_{\mathbf{h}+\mathbf{n}}|, |E_{\mathbf{h}+\mathbf{p}}|,\\ &|E_{\mathbf{k}+\mathbf{l}}|, |E_{\mathbf{k}+\mathbf{m}}|, |E_{\mathbf{k}+\mathbf{n}}|, |E_{\mathbf{k}+\mathbf{p}}|, |E_{\mathbf{l}+\mathbf{m}}|,\\ &|E_{\mathbf{l}+\mathbf{n}}|, |E_{\mathbf{l}+\mathbf{p}}|, |E_{\mathbf{m}+\mathbf{n}}|, |E_{\mathbf{m}+\mathbf{p}}|, |E_{\mathbf{n}+\mathbf{p}}|;\end{aligned} \tag{1.4}$$

$$\begin{aligned}&|E_{\mathbf{h}+\mathbf{k}+\mathbf{l}}|, |E_{\mathbf{h}+\mathbf{k}+\mathbf{m}}|, |E_{\mathbf{h}+\mathbf{k}+\mathbf{n}}|, |E_{\mathbf{h}+\mathbf{k}+\mathbf{p}}|,\\ &|E_{\mathbf{h}+\mathbf{l}+\mathbf{m}}|, |E_{\mathbf{h}+\mathbf{l}+\mathbf{n}}|, |E_{\mathbf{h}+\mathbf{l}+\mathbf{p}}|, |E_{\mathbf{h}+\mathbf{m}+\mathbf{n}}|,\\ &|E_{\mathbf{h}+\mathbf{m}+\mathbf{p}}|, |E_{\mathbf{h}+\mathbf{n}+\mathbf{p}}|\,.\end{aligned} \tag{1.5}$$

In view of (1.2),

$$|E_{\mathbf{k}+\mathbf{l}+\mathbf{m}}|=|E_{\mathbf{h}+\mathbf{n}+\mathbf{p}}| \tag{1.6}$$

etc. so that nothing is gained by adjoining the additional ten magnitudes $|E_{\mathbf{k}+\mathbf{l}+\mathbf{m}}|, |E_{\mathbf{k}+\mathbf{l}+\mathbf{n}}|, \ldots$ to the second neighborhood.

The following usual definition is made

$$\sigma_n=\sum_{j=1}^{N} f_j^n \tag{1.7}$$

where N is the number of atoms, not necessarily identical, in the unit cell, and f_j is the zero-angle atomic structure factor. 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.

2. The probabilistic background

It is assumed that a crystal structure consisting of N atoms per unit cell in $P1$ is fixed. The sixfold Cartesian product $W\times W\times W\times W\times W\times W$ of reciprocal space W consists of all ordered sextuples (**h**, **k**, **l**, **m**, **n**, **p**) of reciprocal vectors **h**, **k**, **l**, **m**, **n**, **p**. Suppose that the ordered sextuple (**h**, **k**, **l**, **m**, **n**, **p**) is a random variable (vector) which is uniformly distributed over the subset of $W\times W\times W\times W\times W\times W$ defined by

$$\mathbf{h}+\mathbf{k}+\mathbf{l}+\mathbf{m}+\mathbf{n}+\mathbf{p}=0\,. \tag{2.1}$$

Then the 31 structure factors

$$E_{\mathbf{h}}, E_{\mathbf{k}}, E_{\mathbf{l}}, E_{\mathbf{m}}, E_{\mathbf{n}}, E_{\mathbf{p}}\,; \tag{2.2}$$

$$E_{\mathbf{h}+\mathbf{k}}, E_{\mathbf{h}+\mathbf{l}}, \ldots, E_{\mathbf{n}+\mathbf{p}}\,; \tag{2.3}$$

$$E_{\mathbf{h}+\mathbf{k}+\mathbf{l}}, E_{\mathbf{h}+\mathbf{k}+\mathbf{m}}, \ldots, E_{\mathbf{h}+\mathbf{n}+\mathbf{p}}\,; \tag{2.4}$$

whose magnitudes (1.3)–(1.5) constitute the second neighborhood of φ, are functions of the primitive random variable (**h**, **k**, **l**, **m**, **n**, **p**) and are therefore themselves random variables. Denote by

$$\begin{aligned}P_{31}=P(&R_1, \ldots, R_6; R_{12}, R_{13}, \ldots, R_{56};\\ &R_{123}, R_{124}, \ldots, R_{156};\\ &\Phi_1, \ldots, \Phi_6; \Phi_{12}, \Phi_{13}, \ldots, \Phi_{56};\\ &\Phi_{123}, \Phi_{124}, \ldots, \Phi_{156})\end{aligned} \tag{2.5}$$

the joint probability distribution of the magnitudes $|E_{\mathbf{h}}|, \ldots$ and phases $\varphi_{\mathbf{h}}, \ldots$ of the 31 structure factors (2.2)–(2.4). The major result of this paper is the derivation of P_{31}, described next.

3. The joint probability distribution of the 31 structure factors (2.2)–(2.4)

Use the notation and terminology described in §§1 and 2 and refer to earlier work (Hauptman, 1975, 1976; Fortier & Hauptman, 1977) to infer that P_{31}, correct up to and including terms of order $1/N^2$, is an exponential function whose argument is a linear combination of all three-phase, four-phase, five-phase and six-phase cosine invariants which can be constructed from the 31 phase variables $\Phi_1, \ldots, \Phi_{156}$, with coefficients which are polynomials in the 31 R's, $R_1, \ldots, R_{156}$. Furthermore, this earlier work establishes the pattern of results two typical examples of which are described in some detail in the Appendix for the present case. Complete details are described in a Technical Report of the Medical Foundation of Buffalo which is available from the authors. Only those terms are written down here which are needed to derive the conditional distribution of φ described in the accompanying paper (Hauptman & Fortier, 1977). Thus P_{31}, the major result of this paper, is given by

$$
\begin{aligned}
P_{31} = {} & \frac{R_1R_2\ldots R_{156}}{\pi^{31}} \exp\{-R_1^2-R_2^2-\ldots-R_{156}^2\} \\
& \times \exp\left\{\frac{2\sigma_3}{\sigma_2^{3/2}}\left[\sum_{15} R_1R_2R_{12}\cos(\Phi_1+\Phi_2-\Phi_{12})\right.\right. \\
& + \sum_{30} R_1R_{23}R_{123}\cos(\Phi_1+\Phi_{23}-\Phi_{123}) \\
& + \sum_{30} R_4R_{56}R_{123}\cos(\Phi_4+\Phi_{56}+\Phi_{123}) \\
& \left. + \sum_{15} R_{12}R_{34}R_{56}\cos(\Phi_{12}+\Phi_{34}+\Phi_{56})\right] \\
& - \frac{2(3\sigma_3^2-\sigma_2\sigma_4)}{\sigma_2^3}\left[\sum_{10} R_1R_2R_3R_{123}\right. \\
& \times \cos(\Phi_1+\Phi_2+\Phi_3-\Phi_{123}) \\
& + \sum_{10} R_4R_5R_6R_{123}\cos(\Phi_4+\Phi_5+\Phi_6+\Phi_{123}) \\
& \left. + \sum_{45} R_1R_2R_{34}R_{56}\cos(\Phi_1+\Phi_2+\Phi_{34}+\Phi_{56})\right] \\
& + \frac{2(15\sigma_3^3-10\sigma_2\sigma_3\sigma_4+\sigma_2^2\sigma_5)}{\sigma_2^{9/2}}\sum_{15} R_1R_2R_3R_4R_{56} \\
& \times \cos(\Phi_1+\Phi_2+\Phi_3+\Phi_4+\Phi_{56}) \\
& - \frac{2}{\sigma_2^6}(105\sigma_3^4-105\sigma_2\sigma_3^2\sigma_4 \\
& + 15\sigma_2^2\sigma_3\sigma_5+10\sigma_2^2\sigma_4^2-\sigma_2^3\sigma_6)R_1R_2R_3R_4R_5R_6 \\
& \left. \times \cos(\Phi_1+\Phi_2+\Phi_3+\Phi_4+\Phi_5+\Phi_6)\right\} \\
& \times \left\{1+O\left(\frac{1}{N}\right)\right\}, \qquad (3.1)
\end{aligned}
$$

where $O(1/N)$ consists of those terms of order $1/N$ or higher which affect only terms of order higher than $1/N^2$ in the conditional distribution of φ, given the 31 magnitudes of the second neighborhood, to be derived in the accompanying paper (Hauptman & Fortier, 1977). In (3.1)

$$\sum_{15} R_1R_2R_{12}\cos(\Phi_1+\Phi_2-\Phi_{12}), \qquad (3.2)$$

for example, is an abbreviation for the sum of the 15 terms

$$R_1R_2R_{12}\cos(\Phi_1+\Phi_2-\Phi_{12}) + R_1R_3R_{13}\cos(\Phi_1+\Phi_2-\Phi_{13})+\ldots \qquad (3.3)$$

obtained by using all similar combinations of the six numbers 1, ..., 6 represented by the single term exhibited in (3.2), *etc.*

As shown in the accompanying paper, (3.1) leads to the conditional probability distribution of the sextet φ, given the 31 magnitudes (1.3)–(1.5) in its second neighborhood.

4. Concluding remarks

Joint probability distributions of three, seven, fifteen and thirty-one structure factors are now known. They constitute the foundation on which the probabilistic theories of the three, four, five and six-phase structure invariants are built. Now that the pattern of these distributions is established, it is feasible, but still time consuming to write down the analogous distributions in 63, 127, ... structure factors, and these lead in turn to conditional distributions of the seven and eight-phase structure invariants, *etc.* In view of the applications which have already been made, it is clear that the quartets, and almost certainly the quintets and sextets as well, will have an important role to play in procedures of phase determination. It is too soon to say whether the higher invariants will be equally useful in the applications.

This research was supported by Ministère de l'Education, Gouvernement du Québec, DHEW Grant No. RR05716 and HL15378 and NSF Grant No. MPS73-04992.

APPENDIX
Construction of (3.1)

A description of the recipe used to derive two representative terms in (3.1) is given.

I.1. *The coefficient of* $\cos(\Phi_1+\Phi_2+\Phi_3-\Phi_{123})$

The order is $1/N$ so that the coefficient is composed of two parts, σ_4/σ_2^2 and $-\sigma_3^2/\sigma_2^3$, each of order $1/N$. Associated with the term σ_4/σ_2^2 (the numerator of which is the single σ_4) is the *quartet* $\Phi_1+\Phi_2+\Phi_3-\Phi_{123}$ itself so that σ_4/σ_2^2 is simply multiplied by two. Associated with the term $-\sigma_3^2/\sigma_2^3$ (the numerator of which is the product of two σ_3's) are the three ways of combining two *triples* to obtain the quartet $\Phi_1+\Phi_2+\Phi_3-\Phi_{123}$:

$$\left.\begin{matrix}\Phi_1+\Phi_{23}-\Phi_{123}\\ \Phi_2+\Phi_3\ \ -\Phi_{23}\end{matrix}\right\}, \qquad (I.1)$$

$$\left.\begin{matrix}\Phi_2+\Phi_{13}-\Phi_{123}\\ \Phi_1+\Phi_3\ \ -\Phi_{13}\end{matrix}\right\}, \qquad (I.2)$$

$$\left.\begin{array}{l}\Phi_3+\Phi_{12}-\Phi_{123}\\ \Phi_1+\Phi_2\ -\Phi_{12}\end{array}\right\}, \qquad (I.3)$$

the sum of the two triples in each pair, (I.1)–(I.3), being equal to the quartet $\Phi_1+\Phi_2+\Phi_3-\Phi_{123}$. Hence the term $-\sigma_3^2/\sigma_2^3$ is multiplied by $3\times2=6$. Addition of the two terms $2\sigma_4/\sigma_2^2$ and $-6\sigma_3^2/\sigma_2^3$ gives the coefficient $-2(3\sigma_3^2-\sigma_2\sigma_4)/\sigma_2^3$ of $R_1R_2R_3R_{123}\cos(\Phi_1+\Phi_2+\Phi_3-\Phi_{123})$ shown in (3.1), and the coefficient $R_1R_2R_3R_{123}$ corresponding to $\Phi_1+\Phi_2+\Phi_3-\Phi_{123}$, is self-explanatory.

I.2. *The coefficient of* $\cos(\Phi_1+\Phi_2+\Phi_3+\Phi_4+\Phi_{56})$

The order is $1/N^{3/2}$ so that the coefficient is composed of three parts, $\sigma_5/\sigma_2^{5/2}$, $-\sigma_3\sigma_4/\sigma_2^{7/2}$ and $\sigma_3^3/\sigma_2^{9/2}$, each of order $1/N^{3/2}$. Associated with the term $\sigma_5/\sigma_2^{5/2}$ (the numerator of which is the single σ_5) is the *quintet* $\Phi_1+\Phi_2+\Phi_3+\Phi_4+\Phi_{56}$ itself so that $\sigma_5/\sigma_2^{5/2}$ is simply multiplied by two. Associated with the term $-\sigma_3\sigma_4/\sigma_2^{7/2}$ (the numerator of which consists of one σ_3 and one σ_4) are the ten ways of combining a *triple* and a *quartet* to obtain the quintet $\Phi_1+\Phi_2+\Phi_3+\Phi_4+\Phi_{56}$:

$$\left.\begin{array}{l}\Phi_1+\Phi_2+\Phi_3-\Phi_{123}\\ \Phi_4+\Phi_{56}+\Phi_{123}\end{array}\right\}, \qquad (I.4)$$

$$\left.\begin{array}{l}\Phi_2+\Phi_3+\Phi_4+\Phi_{156}\\ \Phi_1+\Phi_{56}-\Phi_{156}\end{array}\right\}, \qquad (I.5)$$

$$\left.\begin{array}{l}\Phi_1+\Phi_2-\Phi_{12}\\ \Phi_3+\Phi_4+\Phi_{12}+\Phi_{56}\end{array}\right\}, \qquad (I.6)$$

$$\left.\begin{array}{l}\Phi_3+\Phi_4-\Phi_{34}\\ \Phi_1+\Phi_2+\Phi_{34}+\Phi_{56}\end{array}\right\}, \qquad (I.7)$$

the sum of the two invariants in each pair, (I.4)–(I.7), being equal to the quintet $\Phi_1+\Phi_2+\Phi_3+\Phi_4+\Phi_{56}$. Hence the term $-\sigma_3\sigma_4/\sigma_2^{7/2}$ is multiplied by $10\times2=20$. Associated with the term $\sigma_3^3/\sigma_2^{9/2}$ (the numerator of which consists of three σ_3's) are the 15 ways of combining three *triples* to obtain the quintet $\Phi_1+\Phi_2+\Phi_3+\Phi_4+\Phi_{56}$:

$$\left.\begin{array}{l}\Phi_1+\Phi_2-\Phi_{12}\\ \Phi_3+\Phi_{12}-\Phi_{123}\\ \Phi_4+\Phi_{123}+\Phi_{56}\end{array}\right\}, \qquad (I.8)$$

$$\left.\begin{array}{l}\Phi_3+\Phi_4-\Phi_{34}\\ \Phi_2+\Phi_{34}+\Phi_{156}\\ \Phi_1-\Phi_{156}+\Phi_{56}\end{array}\right\}, \qquad (I.9)$$

$$\left.\begin{array}{l}\Phi_1+\Phi_2-\Phi_{12}\\ \Phi_3+\Phi_4-\Phi_{34}\\ \Phi_{12}+\Phi_{34}+\Phi_{56}\end{array}\right\}, \qquad (I.10)$$

$$\left.\begin{array}{l}\Phi_1+\Phi_3-\Phi_{13}\\ \Phi_2+\Phi_4-\Phi_{24}\\ \Phi_{13}+\Phi_{24}+\Phi_{56}\end{array}\right\}, \qquad (I.11)$$

$$\left.\begin{array}{l}\Phi_1+\Phi_4-\Phi_{14}\\ \Phi_2+\Phi_3-\Phi_{23}\\ \Phi_{14}+\Phi_{23}+\Phi_{56}\end{array}\right\}, \qquad (I.12)$$

the sum of the three invariants in each triple of invariants, (I.8)–(I.12), being equal to the quintet $\Phi_1+\Phi_2+\Phi_3+\Phi_4+\Phi_{56}$. Hence the term $\sigma_3^3/\sigma_2^{9/2}$ is multiplied by $15\times2=30$. Addition of the three terms $2\sigma_5/\sigma_2^{5/2}$, $-20\sigma_3\sigma_4/\sigma_2^{7/2}$, and $30\sigma_3^3/\sigma_2^{9/2}$ gives the coefficient

$$2(15\sigma_3^3-10\sigma_2\sigma_3\sigma_4+\sigma_2^2\sigma_5)/\sigma_2^{9/2}$$

of

$$R_1R_2R_3R_4R_{56}\cos(\Phi_1+\Phi_2+\Phi_3+\Phi_4+\Phi_{56})$$

in (3.1), and the coefficient $R_1R_2R_3R_4R_{56}$, which corresponds to $\Phi_1+\Phi_2+\Phi_3+\Phi_4+\Phi_{56}$, is self-explanatory.

References

FORTIER, S. & HAUPTMAN, H. (1977). *Acta Cryst.* A**33**, 572–575.

HAUPTMAN, H. (1975). *Acta Cryst.* A**31**, 671–679.

HAUPTMAN, H. (1976). *Acta Cryst.* A**32**, 877–882.

HAUPTMAN, H. (1977). *Acta Cryst.* A**33**, 553–555.

HAUPTMAN, H. & FORTIER, S. (1977). *Acta Cryst.* A**33**, 697–701.