Acta Cryst. (1982). A**38**, 129–132

Information Theory as a Basis for the Maximum Determinant

BY P. L. BRITTEN AND D. M. COLLINS

Department of Chemistry, Texas A & M University, College Station, Texas 77843, *USA*

(*Received* 14 *April* 1981; *accepted* 24 *July* 1981)

Abstract

The concepts of information theory have been applied to the analysis of spatial array data in the maximum entropy method. The development is similar to that of Tsoucaris's maximum determinant method for phase determination. It is shown that a maximum determinant (of structure factors) corresponds exactly to a function of maximum entropy. The squared modulus of its Fourier transform formally approximates the true electron density.

1. Introduction

The use of information theory and the concept of entropy in the analysis of time series was described by Shannon & Weaver (1949). Burg (1967) developed an algorithm for the calculation of an estimate for the power spectral density of a uniformly sampled random process, which he called maximum entropy spectral analysis (MESA). This algorithm is equivalent to the determination of an autoregressive process, and is related to information theory in that the spectral density obtained corresponds to a maximum entropy estimate.

McDonough (1974) has discussed the extension of the method to wavenumber estimation for spatial array data; its application to extrapolation of crystallographic structure-factor data and maximization of resolution has been investigated by Collins (1978). The fact that higher resolution is obtained by the maximum entropy method compared to classical methods of spectral analysis is one reason for its usefulness in data processing.

A method for phase determination in crystallography, introduced by Tsoucaris (1970), and known as the maximum determinant method, has a formulation similar to that of the maximum entropy method. It is the purpose of this paper to show that the formalisms of the two methods are, in fact, equivalent.

2. Information theory and the maximum entropy method

We begin the description of the maximum entropy method with a review of its connection to information theory. Shannon & Weaver (1949) refer to the uncertainty of a process as its entropy, H. For a univariate continuous process with probability distribution $p(x)$, the entropy is

$$H = -\int_{-\infty}^{\infty} p(x) \ln p(x)\, \mathrm{d}x. \tag{1}$$

For a function of n variables this becomes

$$H = -\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1, \ldots, x_n) \times \ln p(x_1, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n. \tag{2}$$

The form of the entropy expression is the same as that found in statistical mechanics, and it may be seen that, when the result of a process is certain, the entropy is zero; when all possible outcomes are equally likely (the most uncertain case), H is a maximum, equal to $\ln n$.

The distribution $p(x)$ which yields a maximum entropy when the standard deviation of x is fixed at σ is Gaussian in form (Shannon & Weaver, 1949). Such a distribution gives the entropy expression

$$H = \ln\,[(2\pi e)^{1/2}\, \sigma]. \tag{3}$$

In the multivariate case, with the second-order moments of $p(x_1, \ldots, x_n)$ fixed at A_{ij},

$$A_{ij} = \int \cdots \int x_i x_j p(x_1, \ldots, x_n)\, \mathrm{d}x_1 \ldots \mathrm{d}x_n, \tag{4}$$

the form of $p(x_1, \ldots, x_n)$ giving a maximum entropy is the Gaussian distribution having these moments:

$$p(x_1, \ldots, x_n) = (2\pi)^{-n/2}\, |\mathbf{A}|^{-1/2} \exp\{-\tfrac{1}{2}\mathbf{x}'\, [\mathbf{A}^{-1}]\mathbf{x}\}. \tag{5}$$

In (5), $\mathbf{A}$ and $|\mathbf{A}|$ are respectively the matrix of A_{ij} and the determinant with elements A_{ij}, $\mathbf{x}$ is the vector of x_n, and $'$ indicates transpose. The corresponding entropy expression is

$$H = \ln\,[(2\pi e)^{n/2}\, |\mathbf{A}|^{1/2}]. \tag{6}$$

These concepts of information theory are related to

Editorial note: The similarity between this and the preceding paper, by Narayan & Nityananda [*Acta Cryst.* (1982), A**38**, 122–128], has been recognized and, although they represent completely independent work, they have been published together to facilitate comparison.

0567-7394/82/010129-04$01.00

the development of Burg's maximum entropy method (Burg, 1967; McDonough, 1974), described as follows.

A stationary random process will be given by $x(t)$, and its random variables or samples by x_n, $n = 1, \ldots, N$. Then the correlation coefficients between these random variables are

$$R_k = E(x_{n+k} x_n^*), \tag{7}$$

where E denotes the expectation value and * denotes the complex conjugate. Since the data set $x_1, \ldots, x_N$ is finite, R_k must be approximated:

$$\hat{R}_k = \frac{1}{N} \sum_{n=1}^{N-k} (x_{n+k} - \mu)(x_n - \mu)^*, \tag{8a}$$

where

$$\mu = \frac{1}{N} \sum_{n=1}^{N} x_n \tag{8b}$$

$$\hat{R}_k = \hat{R}^*_{-k}, \quad k = -(N-1), \ldots, -1. \tag{8c}$$

Thus, the samples $x_1, \ldots, x_N$ generate correlation samples $\hat{R}_0, \ldots, \hat{R}_K$; $K = N - 1$.

The sequence of correlation samples may be extended beyond the given data; that is, values for R_{K+1}, R_{K+2}, ... may be obtained, even though the corresponding $x_{K+2}, x_{K+3}, \ldots$ are not given. The joint density of the x_n and the hypothetical variable x_{K+2}, for instance, is $p(x_1, \ldots, x_{K+2})$; its entropy is

$$H = -\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1, \ldots, x_{K+2}) \times \ln p(x_1, \ldots, x_{K+2}) \mathrm{d}x_1 \ldots \mathrm{d}x_{K+2}. \tag{9}$$

The density distribution $p(x_1, \ldots, x_{K+2})$ is chosen so that its first $K + 1$ second-order moments are the known $\hat{R}_0, \ldots, \hat{R}_K$, and which maximizes the entropy (9). For real random variables $x_1, \ldots, x_{K+2}$, this density is

$$p(x_1, \ldots, x_{K+2}) = (2\pi)^{-(K+2)/2} |\mathbf{R}_{K+1}|^{-1/2} \times \exp[-\tfrac{1}{2}\mathbf{x}' \mathbf{R}_{K+1}^{-1} \mathbf{x}], \tag{10}$$

where $\mathbf{x}$ is the vector of the x_n, ′ indicates transpose, and

$$\mathbf{R}_{K+1} = \begin{bmatrix} \hat{R}_0 & \hat{R}_1^* & \ldots & \hat{R}_K^* & R_{K+1}^* \\ \hat{R}_1 & \hat{R}_0 & \ldots & & \hat{R}_K^* \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \hat{R}_K & & \ldots & & \hat{R}_1^* \\ R_{K+1} & & \ldots & & \hat{R}_0 \end{bmatrix}. \tag{11}$$

This density has entropy

$$H = \ln [(2\pi e)^{(K+2)/2} |\mathbf{R}_{K+1}|^{1/2}]. \tag{12}$$

In the complex case, the corresponding density and entropy expressions are

$$p(x_1, \ldots, x_{K+2}) = \pi^{-(K+2)} |\mathbf{R}_{K+1}|^{-1} \times \exp[-\mathbf{x}' \mathbf{R}_{K+1}^{-1} \mathbf{x}] \tag{13}$$

where ′ now indicates conjugate transpose, and

$$H = \ln [(\pi e)^{K+2} |\mathbf{R}_{K+1}|]. \tag{14}$$

Maximizing either entropy expression (12) or (14) implies that the determinant of $\mathbf{R}_{K+1}$ must be maximized with respect to R_{K+1}. For any matrix $\mathbf{M}$, Bodewig (1959) has shown that

$$\mathrm{d}|\mathbf{M}| = |\mathbf{M}| \operatorname{trace} (\mathbf{M}^{-1} \mathrm{d}\mathbf{M}). \tag{15}$$

Thus, differentiating $|\mathbf{R}_{K+1}|$ with respect to R^*_{K+1} yields

$$\mathrm{d}|\mathbf{R}_{K+1}| = (-1)^{K+1} |\mathbf{B}| \mathrm{d}R^*_{K+1}, \tag{16}$$

where

$$\mathbf{B} = \begin{bmatrix} \hat{R}_1 & \hat{R}_0 & \ldots & \hat{R}^*_{K-1} \\ \hat{R}_2 & \hat{R}_1 & \ldots & \hat{R}^*_{K-2} \\ \cdots & \cdots & \cdots & \cdots \\ \hat{R}_K & & \ldots & \hat{R}_0 \\ R_{K+1} & \hat{R}_K & \ldots & \hat{R}_1 \end{bmatrix}. \tag{17}$$

If the derivative (16) is set equal to zero to obtain the maximum, then this implies that the equation

$$|\mathbf{B}| = 0 \tag{18}$$

is solved for R_{K+1}.

Once R_{K+1} is found, the set $\hat{R}_0, \ldots, \hat{R}_K, R_{K+1}$ may be used to determine R_{K+2} in a similar manner, by considering the entropy of the joint density $p(x_1, \ldots, x_{K+2}, x_{K+3})$.

3. The maximum determinant method

Tsoucaris (1970) introduced a method for phase determination which he called the maximum determinant method, and which is derived from the conditional joint probability of a set of structure factors in a Karle–Hauptman determinant. The procedure allows the simultaneous determination of many structure-factor phases.

Inequality theory (Karle & Hauptman, 1950) shows that the non-negativity of the electron density function $\rho(\mathbf{r})$ imposes the condition of non-negativity of a Karle–Hauptman determinant composed of the structure factors. The maximum determinant rule states that the most probable combination of phases satisfying this condition is the one which maximizes this determinant (Tsoucaris, 1970).

For this discussion we consider equal atoms in space group $P1$. A normalized structure factor, $E_\mathbf{h}$, is given by

$$E_\mathbf{h} = N^{-1/2} \sum_{j=1}^{N} \exp(2\pi i \mathbf{h}.\mathbf{r}_j), \tag{19}$$

where $\mathbf{h}$ is a reciprocal-lattice vector, $\mathbf{r}_j$ is the position vector of the jth atom in the crystal, and N is the number of atoms in the unit cell. The corresponding unitary structure factors are

$$U_{\mathbf{h}} = N^{-1/2} E_{\mathbf{h}}. \tag{20}$$

The set of m normalized structure factors $E_1, \ldots, E_p, \ldots, E_m$ are considered as random variables. Correlation coefficients between these random variables are generated according to the Sayre (1952) equation:

$$\overline{E_{\mathbf{k}} E_{\mathbf{h}-\mathbf{k}}} = U_{\mathbf{h}}, \tag{21}$$

where **h** is a fixed vector, and the bar indicates average over **k**.

We now let $E_p = E_{\mathbf{L}+\mathbf{h}_p}$, where the $\mathbf{h}_p$ are fixed vectors and **L** is a random vector which sweeps out all reciprocal space. Then, for each pair of structure factors E_p and E_q, we have a fixed vector $\mathbf{h}_p - \mathbf{h}_q$, such that $(\mathbf{L} + \mathbf{h}_p) - (\mathbf{L} + \mathbf{h}_q) = \mathbf{h}_p - \mathbf{h}_q$. The unitary structure factors corresponding to these vectors are $U_{\mathbf{h}_p - \mathbf{h}_q} = U_{pq}$.

With the change of notation $\mathbf{k} = \mathbf{L} + \mathbf{h}_p$, $\mathbf{h} = \mathbf{h}_p - \mathbf{h}_q$, equation (21) may be rewritten

$$\overline{E_p E_q^*} = U_{pq}; \quad p, q = 1, \ldots, m. \tag{22}$$

The unitary structure factors U_{pq} are the correlation coefficients between the random variables E_p and E_q, and are assumed to be fixed and known. They form a matrix whose determinant is the non-negative Karle–Hauptman determinant

$$D_m = \det(U_{pq}) \geq 0. \tag{23}$$

The following determinant, Δ_{m+1}, may be formed from D_m by appending the structure factors $E_1, \ldots, E_m$ as a last row and column, the last element being N.

$$\Delta_{m+1} = \frac{1}{N} \begin{vmatrix} 1 & U_{12} & \cdots & U_{1m} & E_1 \\ U_{21} & 1 & \cdots & & E_2 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ U_{m1} & & \cdots & 1 & E_m \\ E_{-1} & & \cdots & E_{-m} & N \end{vmatrix}. \tag{24}$$

The probability law $p(E_1, \ldots, E_m)$ under the condition that the U_{pq} are fixed and known is the Gaussian distribution (Tsoucaris, 1970):

$$p(E_1, \ldots, E_m) = (2\pi)^{-m/2} D_m^{-1/2} \exp(-\tfrac{1}{2} Q_m) \tag{25a}$$

or

$$p(E_1, \ldots, E_m) = (2\pi)^{-m} D_m^{-1/2} \exp(-Q_m) \tag{25b}$$

for the centrosymmetric and non-centrosymmetric cases, respectively. Here, Q_m is given by

$$Q_m = \mathbf{E}' \, [\mathbf{U}^{-1}] \, \mathbf{E} = N \frac{D_m - \Delta_{m+1}}{D_m}, \tag{26}$$

where $\mathbf{E}$ is the vector of the E_p and $[\mathbf{U}^{-1}]$ is the inverse of the matrix of the U_{pq}.

When the magnitudes of the E_p are known, the probability law (25) becomes an expression for the probability of the phases or signs of the structure factors. The maximum probability is reached when Δ_{m+1} is a maximum; this, in turn, leads to the most probable combination of the phases.

4. Discussion

Comparison of the maximum entropy and maximum determinant methods clearly shows their similar developments. From the probability distributions of the random variables to the construction of the determinants which are maximized, the formulations of the two methods are completely analogous, save for one step—the calculation of the correlation coefficients.

The maximum determinant method, as put forth by Tsoucaris, makes use of correlation coefficients obtained from the Sayre equation (21). The approximation inherent in this equation, however, is not crucial to the development of the maximum determinant method, for there is an exact foundation upon which it may be based. Specifically, the determinant maximum corresponds exactly to a maximum of entropy in a function related to electron density.

The imposition of positivity upon the electron density could take many forms, but it is from the perfectly general form

$$\rho(\mathbf{r}) = |g(\mathbf{r})|^2 > 0 \tag{27}$$

that $G(\mathbf{h})$, the function of maximum entropy, is found. Equation (27) allows us to write

$$F_{\mathbf{h}} = \frac{1}{V} \sum_{\mathbf{k}} G_{\mathbf{k}} G_{\mathbf{k}-\mathbf{k}}^*, \tag{28}$$

where $G_{\mathbf{h}}$ and $F_{\mathbf{h}}$ are the Fourier transforms of $g(\mathbf{r})$ and $\rho(\mathbf{r})$, respectively. Then $G_{\mathbf{h}}$, corresponding to x_n of §2, represents the random variables, whose existence is ensured by the generalized Fejér–Riesz theorem (Papoulis, 1973).

No distinction is intended here between $F_{\mathbf{h}}$ and its normalized equivalent $E_{\mathbf{h}}$, but $F_{\mathbf{h}}$ is more convenient to use. This arises from the fact that structure factors obtained from the completely general equation (28) are exact correlation coefficients, analogous to R_k of the maximum entropy method. Hence, the covariance matrix constructed of $F_{\mathbf{h}}$ corresponds exactly to the matrix of estimates $\hat{R}_k$, whose determinant is analogous to Tsoucaris's D_m. Phase determination or extrapolation, then, corresponds to maximization of the determinant of the extended matrix **F**, order $m + 1$, just as for $|\mathbf{R}_{K+1}|$ and Δ_{m+1}.

In the estimation of R_{K+1} from the finite data record no new information is used. However, in the crystallographic case, new information is employed, since the structure-factor magnitudes are obtained from experiment. Nevertheless, the extrapolation of data, in both cases, corresponds to determination of an analytic function whose squared modulus is a function of maximum entropy measure. The definition for entropy measure is that given by Ponsonby (1973),

$$H' = \int_V \ln \rho(\mathbf{r})\, \mathrm{d}V_r, \qquad (29)$$

V being the unit-cell volume. Ponsonby emphasizes that this is strictly only a measure of relative entropy of the real power spectrum of an underlying signal. In this sense maximum H' corresponds to maximum entropy measure $\rho(\mathbf{r}) = |g(\mathbf{r})|^2$, and a maximum-entropy distribution function for $G_{\mathbf{h}}$, the transform of $g(\mathbf{r})$. Equation (6) shows that such a distribution requires a maximum determinant D_m. Moreover, Ponsonby points out that entropy maximization provides a unique, smooth autocorrelation function in exact agreement with its observed values which concedes greatest possible ignorance of its unobserved values. This corresponds to the liberty of choosing by unspecified criteria the vectors $(\mathbf{L}' + \mathbf{h}_p)$ which direct the formulation of D_m and allow as many different appended rows and columns in Δ_{m+1} as there may be reasonable choices of $\mathbf{L}'$.

The formalisms of the maximum entropy and maximum determinant methods have been cast in parallel form and shown to be equivalent. With only the restrictions that (electron) density be positive, and that matrices and determinants be of small enough order to be positive definite, the two equivalent formulations yield identical results and provide an information-theoretic interpretation of the maximum determinant.

This work was supported in part by the Robert A. Welch Foundation through Grant A-742, and by the Research Corporation through a Cottrell Research Grant.

References

Bodewig, E. (1959). *Matrix Calculus*. Amsterdam: North-Holland.

Burg, J. P. (1967). *Maximum-Entropy Spectral Analysis*. 37th Annual International SEG Meeting, October 31, Oklahoma City. Reprinted in *Modern Spectrum Analysis* (1978), edited by D. G. Childers, pp. 34–41. New York: IEEE Press.

Collins D. M. (1978). *Acta Cryst.* A**34**, 533–541.

Karle, J. & Hauptman, H. (1950). *Acta Cryst.* **3**, 181–187.

McDonough, R. N. (1974). *Geophysics*, **39**, 843–851.

Papoulis, A. (1973). *IEEE Trans. Inf. Theory*, **19**, 9–12.

Ponsonby, J. E. B. (1973). *Mon. Not. R. Astron. Soc.* **163**, 369–380.

Sayre, D. (1952). *Acta Cryst.* **5**, 60–65.

Shannon, C. E. & Weaver, W. (1949). *The Mathematical Theory of Communication*. Univ. of Illinois Press.

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

Acta Cryst. (1982). A**38**, 132–137

The Layered Perovskites $(C_3H_7NH_3)_2MnCl_4$ (PAMC)* and $(C_2H_5NH_3)_2MnCl_4$ (EAMC):† Birefringence Studies and the Symmetry of the Commensurately Modulated ε Phase of PAMC

By Ian Hedley Brunskill and Wulf Depmeier

Université de Genève, Chimie Appliquée, 30 *quai Ernest Ansermet, CH* 1211 *Genève, Switzerland*

(*Received* 21 *May* 1981; *accepted* 4 *August* 1981)

Abstract

By measuring the birefringence on (001) platelets of PAMC as a function of temperature, previously determined phase transitions have been confirmed. Large effects were measured on passing through the normal–incommensurate $\delta \rightarrow \gamma$ transition and through the normal–commensurate $\zeta \rightarrow \varepsilon$ transition, whereas the commensurate–normal $\varepsilon \rightarrow \delta$ and the incommensurate–normal $\gamma \rightarrow \beta$ transitions are less pronounced, in accordance with former thermoanalytic studies. A comparison is made with a similar study on EAMC. The optical anisotropy is discussed for both compounds in terms of the structural changes of the perovskite-type layers. Power laws of the form $\Delta n \sim (1 - T/T_c)^{2\beta}$ are used to describe the birefringence where appropriate. By analyzing systematic extinctions amongst main and satellite reflexions, the ε phase of PAMC is proposed to be a (3 + 1)-dimensional commensurately modulated version of the δ phase. The superspace group is determined to be $P^{Abma}_{s\bar{1}1}$. With the

---

* Bis(*n*-propylammonium) tetrachloromanganate.

† Bis(ethylammonium) tetrachloromanganate.

0567-7394/82/010132-06$01.00