In conclusion it can be stated that the enantiomorph-specific refinements and extensions described here are capable of improving the electron-density maps of large molecules. Its limit of applicability, and the possible improvements suggested in this paragraph, will be pursued further.

The authors thank Professor B. O. Loopstra and Dr C. H. Stam for valuable comments on the manuscript. The Protein Data Bank is acknowledged for providing us with the data on metmyoglobin. We are grateful to Mr W. Gaykema of the Laboratory voor Chemische Fysica of the University of Groningen for assisting with the calcualtion of the *E* maps. One of us (GJO) is indebted to the Netherlands Organization for the Advancement of Pure Research (ZWO) for financial support.

References

BERNSTEIN, F. C., KOEZLE, T. F., WILLIAMS, G. J. B., MEYER, E. F. JR, BRICE, M. D., RODGERS, J. R., KENNARD, O., SHIMANOUCHI, T. & TASUMI, M. (1977). *J. Mol. Biol.* **112**, 535–542.

BUSETTA, B. (1976). *Acta Cryst.* A**32**, 139–143.

COULTER, C. L. & DEWAR, R. B. K. (1971). *Acta Cryst.* B**27**, 1730–1740.

DIAMOND, R. (1980). In *Computing in Crystallography*, edited by R. DIAMOND, S. RAMASESHAN & K. VENKATESAN. Bangalore: Indian Academy of Sciences.

OLTHOF, G. J. & SCHENK, H. (1981*a*). *Acta Cryst.* A**37**, 684–689.

OLTHOF, G. J. & SCHENK, H. (1981*b*). *Acta Cryst.* A**37**, 689–691.

OLTHOF, G. J., SINT, L. & SCHENK, H. (1979). *Acta Cryst.* A**35**, 941–946.

OVERBEEK, A. R. & SCHENK, H. (1978). In *Computing in Crystallography*, edited by H. SCHENK, R. OLTHOF-HAZEKAMP, H. VAN KONINGSVELD & G. BASSI, pp. 108–112. Delft Univ. Press.

SAYRE, D. (1974). *Acta Cryst.* A**30**, 180–184.

SAYRE, D. (1980). In *Theory and Practice of Direct Methods in Crystallography*, edited by M. F. C. LADD & R. A. PALMER, pp. 271–286. New York: Plenum.

SINT, L. & SCHENK, H. (1975). *Acta Cryst.* A**31**, S22.

STEWART, J. M. (1976). The XRAY system–version of 1976. Tech. Rep. TR-446. Computer Science Center, Univ. of Maryland, College Park, Maryland.

TAKANO, T. (1977). *J. Mol. Biol.* **110**, 537–568.

Acta Cryst. (1982). A**38**, 122–128

The Maximum Determinant Method and the Maximum Entropy Method

BY RAMESH NARAYAN AND RAJARAM NITYANANDA

Raman Research Institute, Bangalore-560080, India

(*Received* 9 *March* 1981; *accepted* 24 *July* 1981)

Abstract

A generalized maximum determinant rule is shown to be equivalent to maximizing the integral of the logarithm of the electron density, *i.e.* equivalent to the 'maximum entropy method' (MEM) of image reconstruction. Relations between the structure factors and the Fourier coefficients of the reciprocal of the electron density follow, leading to new algorithms for phase determination and refinement. Although structures with equal, spherical, resolved atoms automatically satisfy the MEM phase relations, the method really requires only positivity and 'peakiness' of the electron density.

1. Introduction

The maximum determinant method (MDM) of crystallography (Lajzerowicz & Lajzerowicz, 1966; Tsoucaris, 1970) has received fair attention as an alternative 'direct method' to the conventional approach of estimating low-order structure invariants and seminvariants (Sayre, 1952; Ladd & Palmer, 1980). Tsoucaris (1980) describes the use of the MDM to rederive low-order relations, determine phases *ab initio* (from medium-sized determinants), and refine and extend phases (from large determinants). In this paper we give a new interpretation of the MDM and use it to draw a number of conclusions of relevance to crystallography.

We introduce in § 2 a theorem relating a certain limit of the Karle–Hauptman (1950) determinant to the integral of the logarithm of the electron density in the

---

Editorial note: The similarity between this and the following paper, by Britten & Collins [*Acta Cryst.* (1982), A**38**, 129–132], has been recognized and, although they represent completely independent work, they have been published together to facilitate comparison.

0567-7394/82/010122-07$01.00

crystal unit cell. The proof of this theorem is presented in § 5. Using the above theorem, we show in § 3 that the MDM is closely related to one of the forms of the maximum entropy method (MEM) of image reconstruction. We apply the MEM to the phase problem and show that it leads to interesting phase relationships which can be made the basis of algorithms for phase refinement. We argue that such schemes will probably be significantly faster than current approaches to the MDM. We discuss in § 4 some of the properties of the MEM solution. We show that the solution may not be unique. However, all structures with equal, spherical, resolved atoms automatically satisfy the MEM phase relations. We make general statements on the shapes of peaks in the reconstructed electron density and suggest on this basis that a second form of the MEM may be more appropriate for crystallographic applications.

2. A theorem on Karle–Hauptman determinants

The structure factor $F_{\mathbf{H}}$ of a crystal with electron density $\rho(\mathbf{r})$ in a unit cell of volume V_r is defined as

$$F_{\mathbf{H}} = \int_{V_r} \rho(\mathbf{r}) \exp(2\pi i \mathbf{H}.\mathbf{r})\, \mathrm{d}V_r \tag{1}$$

with the inverse relation

$$\rho(\mathbf{r}) = \frac{1}{V_r} \sum_{\mathbf{H}_j} F_{\mathbf{H}_j} \exp(-2\pi i \mathbf{H}_j.\mathbf{r}), \tag{2}$$

where the $\mathbf{H}_j$ are reciprocal-lattice vectors. Let us choose d reciprocal vectors $\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_d$ arbitrarily and use the notation

$$F_{pq} = F_{\mathbf{H}_p - \mathbf{H}_q}; \quad p, q = 1, 2, \ldots, d. \tag{3}$$

The positivity of the electron density $\rho(\mathbf{r})$ for all $\mathbf{r}$ then implies that the Karle–Hauptman (1950) determinant D_d (of dimensionality $d \times d$) written below is non-negative

$$i.e. \qquad D_d = \begin{vmatrix} F_0 & F_{12} & \cdots & F_{1d} \\ F_{21} & F_0 & \cdots & F_{2d} \\ \vdots & \vdots & & \vdots \\ F_{d1} & F_{d2} & \cdots & F_0 \end{vmatrix} \geq 0. \tag{4}$$

This relation is rigorous for all d and for all choices of $\mathbf{H}_1, \ldots, \mathbf{H}_d$. In practice, D_d is written in terms of unitary structure factors U_{ij} $(= F_{ij}/Zf$, where Z is the total number of electrons per unit cell and f is an average atomic 'shape factor').

Without loss of generality we consider D_d to be defined in terms of F_{ij} [as in (4)] rather than the more usual U_{ij}. Let us choose $\mathbf{H}_1, \ldots, \mathbf{H}_d$ such that they entirely fill the following cube in reciprocal space

$$\mathbf{H}_j = (hkl); \quad h, k, l = -p, -p+1, \ldots, p. \tag{5}$$

We are in principle interested in the limit d $[= (2p+1)^3] \to \infty$, although, in practice, we would select some large value of p such that all the boundary structure factors have decayed sufficiently in amplitude. We now state the following theorem:

$$\operatorname*{Lim}_{d\to\infty} \frac{1}{d} \ln D_d = \frac{1}{V_r} \int_{V_r} \ln[V_r \rho(\mathbf{r})]\, \mathrm{d}V_r. \tag{6}$$

This is a special case of the remarkable theorem proved by Szegö (1920) for one-dimensional functions:

$$\operatorname*{Lim}_{d\to\infty} \frac{1}{d} \operatorname{Tr}[E(A_d)] = \frac{1}{V_r} \int_{V_r} E[V_r \rho(\mathbf{r})]\, \mathrm{d}V_r, \tag{7}$$

where E corresponds to any continuous function and A_d is the Karle–Hauptman matrix. Szegö's proof employs complex variable theory and cannot be generalized to higher dimensions. We give an alternative proof of (6) in § 5 which is valid for all dimensions.

3. Relationship between MDM and MEM

The MDM goes beyond the inequality (4) by stating that

$$D_d = \text{maximum} \tag{8}$$

and uses (8) in the applications mentioned in § 1. In deriving (8), D_d is written in terms of U_{ij} which are assumed to correspond to a crystal with equal point atoms. Moreover, a result derived for the *conditional* probability of one set of phases, given other phases, is generalized to one referring to the *joint* probability of all phases (Tsoucaris, 1980). In fact, Heinerman, Krabbendam & Kroon (1979) have derived the true joint probability distribution. Although their result shows some similarity to (8), it is clear that the MDM should, at present, be regarded as a heuristic principle. We present below an alternative approach to interpreting the method.

By (6) we see that the problem of maximizing D_d, in the limit of large d, with respect to the unknowns occurring in it (which are the phases of the reflections with known amplitudes, and both the amplitudes and the phases of the rest), is equivalent to maximizing

$$S = \int_{V_r} \ln[\rho(\mathbf{r})]\, \mathrm{d}V_r. \tag{9}$$

This is identical to one of the forms of the MEM of image reconstruction, which has been widely discussed in the context of geophysics (Burg, 1967, 1975; Smylie, Clarke & Ulrych, 1973) and radio astronomy (Ables, 1974; Komesaroff, Narayan & Nityananda, 1981, and references therein).

The maximization of S leads to new relations which help in interpreting the MDM and also in its implementation. We write the structure factor $F_{\mathbf{H}}$ in the modulus phase form

$$F_{\mathbf{H}} = |F_{\mathbf{H}}| \exp(i\varphi_{\mathbf{H}}). \qquad (10)$$

We introduce the Fourier coefficients $G_{\mathbf{H}}$ of the reciprocal of the electron density

$$G_{\mathbf{H}} = |G_{\mathbf{H}}| \exp(i\theta_{\mathbf{H}}) = \int_{V_r} \frac{1}{\rho(\mathbf{r})} \exp(2\pi i \mathbf{H}.\mathbf{r})\, dV_r. \qquad (11)$$

We also rewrite (2) in the following explicit form which takes note of Friedel's law

$$\rho(\mathbf{r}) = \frac{1}{V_r} \left\{ F_0 + \sum_{\mathbf{H}_j}' |F_{\mathbf{H}_j}| [\exp(i\varphi_{\mathbf{H}_j} - 2\pi i \mathbf{H}_j.\mathbf{r}) + \exp(-i\varphi_{\mathbf{H}_j} + 2\pi i \mathbf{H}_j.\mathbf{r})] \right\}, \qquad (12)$$

where $\sum'$ implies summation over half the reciprocal space. If we substitute (12) into (9), it is possible to show that

$$\frac{\partial S}{\partial |F_{\mathbf{H}_j}|} = \frac{2}{V_r} |G_{\mathbf{H}_j}| \cos(\varphi_{\mathbf{H}_j} - \theta_{\mathbf{H}_j}), \qquad (13)$$

$$\frac{\partial S}{\partial \varphi_{\mathbf{H}_j}} = \frac{2}{V_r} |F_{\mathbf{H}_j}||G_{\mathbf{H}_j}| \sin(\varphi_{\mathbf{H}_j} - \theta_{\mathbf{H}_j}). \qquad (14)$$

The maximization of S therefore leads to the following results:

(*a*) For all **H** where both the amplitude and phase of $F_{\mathbf{H}}$ are unknown,

$$G_{\mathbf{H}} = 0. \qquad (15)$$

This result is well known (Burg, 1967).

(*b*) Where the modulus of $F_{\mathbf{H}}$ is known, but not its phase, then (Narayan & Nityananda, 1981)

$$\sin(\varphi_{\mathbf{H}} - \theta_{\mathbf{H}}) = 0. \qquad (16)$$

There are two standard approaches to the practical implementation of a variational scheme such as the maximization of S. (*a*) One could seek an iterative fixed-point scheme on the basis of (15) and (16). (*b*) One could develop gradient-type algorithms using (13) and (14). Willingale (1981) has a fixed-point scheme for applications of the MEM in X-ray astronomy where (15) alone occurs. The present authors have obtained encouraging results with the phase problem in simple one-dimensional simulations where the gradient method easily solves the 'structure' from initial random phases (Narayan & Nityananda, 1981). Although our experience with the MEM for phase refinement has so far been limited to small-scale problems, we are optimistic about its relevance and usefulness to 'real' problems.

A feature of both the fixed-point and gradient approaches is that they can be made very efficient by the incorporation of the fast Fourier transform algorithm. Since two FFT's will compute all the $G_{\mathbf{H}}$ simultaneously, it seems likely that these approaches will be much faster than current algorithms for the MDM which directly maximize D_d by determinant manipulations and matrix inversion. This expectation is yet to be verified.

The MEM is known to be sensitive to the value of F_0 assumed (Bhandari, 1978; Komesaroff *et al.*, 1981). For the pure Fourier extrapolation problem (equation 15), lower values of F_0 lead to flatter baselines and sharper peaks, but slower convergence (Nityananda & Narayan, 1982). For the pure phase refinement problem (equation 16), the effect of F_0 on convergence and baseline remains, though resolution is no longer a factor. There is another interesting effect of F_0. In experiments with simple one-dimensional structures such as in Narayan & Nityananda (1981), it was found that when large values of F_0 are used, the MEM sometimes leads to a centrosymmetric solution* even though the structure factors are computed from a non-centrosymmetric structure. For these reasons, one should not use the true value of F_0 in calculations with the MEM, but rather a higher value, selected on the above considerations. We may add that when there are errors in the measured $|F_{\mathbf{H}}|$, the electron density will go negative and the Karle–Hauptman inequality, (4), will itself be valid only if F_0 is increased sufficiently. The value of F_0 has to be carefully chosen in density modification schemes also (see, for example, Nixon & North, 1976).

4. Properties of the MEM solution

Equation (16) leads to the following non-trivial† ambiguity for each **H**:

$$\varphi_{\mathbf{H}} = \theta_{\mathbf{H}}$$

or

$$\varphi_{\mathbf{H}} = \theta_{\mathbf{H}} + \pi. \qquad (17)$$

It is true that many of these possibilities may not be realized or may result in minima and saddle points of S in 'phase space'. However, one cannot rule out the existence of many maxima. The MEM solution to the phase problem is therefore not unique. This is in contrast to the Fourier extrapolation case (equation 15), where the solution can be shown to be unique (Burg, 1975; Nityananda & Narayan, 1982).

We now demonstrate the special relevance of the MEM to equal-atom structures. Consider a structure consisting of equal, spherical, resolved atoms. By

---

* It will be noted that all centrosymmetric structures automatically satisfy equation (16).

† As against trivial ambiguities related to the choice of origin and enantiomorph.

arguments similar to those employed by Sayre (1952), the electron density would satisfy the following relation

$$\frac{1}{\rho(\mathbf{r})} = \rho(\mathbf{r}) * s(\mathbf{r}), \tag{18}$$

where $s(\mathbf{r})$ is a spherically symmetric 'shape' function and $*$ denotes convolution. Equation (18) leads to

$$G_{\mathbf{H}_j} = F_{\mathbf{H}_j} S(|\mathbf{H}_j|), \tag{19}$$

where $S(|\mathbf{H}_j|)$ is real. We note that (19) leads to the MEM phase relations (16) or (17). We thus see that structures with equal, spherical, resolved atoms are automatically solutions of the MEM. Moreover, the sign of $S(|\mathbf{H}_j|)$ can be used to resolve the ambiguity in (17). Although the discussion here is based on maximizing S as defined in (9), the result is clearly true for maximizing any general function of the form (7).

Despite the above arguments, it should be noted that the MEM does not *require* the assumption of equal atoms. Nityananda & Narayan (1982) have shown in another context that the basic requirement for the successful application of the MEM is that the function being reconstructed should be 'peaky' and have a flat baseline. Therefore, in crystallographic applications, the MEM (or MDM) should be useful so long as there are separate, well-defined atoms, though they may be unequal.

The form of $\rho(\mathbf{r})$ reconstructed by maximizing (9) has the following properties. Firstly, $\rho(\mathbf{r})$ will be positive at all $\mathbf{r}$ because of the occurrence of $\ln[\rho(\mathbf{r})]$. Secondly, the maxima in $\rho(\mathbf{r})$ arise from the minima in $[\rho(\mathbf{r})]^{-1}$. By (15), $[\rho(\mathbf{r})]^{-1}$ has only a finite number of non-vanishing Fourier coefficients clustered around the origin and is therefore Taylor expandable about its minima. The peaks in $\rho(\mathbf{r})$ are therefore of the approximate form:

$$\rho(\mathbf{r}) \simeq \rho(\mathbf{r}_0)[1 + \tfrac{1}{2}(\mathbf{r}-\mathbf{r}_0)^T Q(\mathbf{r}-\mathbf{r}_0)]^{-1}, \tag{20}$$

which are generalized Lorentzians (Q is a matrix of second derivatives and $\mathbf{r}_0$ is the position of the peak). Now, the true electron density distributions of atoms are closer to Gaussians than Lorentzians. It can be shown by a similar argument to (20) (Nityananda & Narayan, 1982) that maximizing the alternative form of 'entropy' S' defined below (Gull & Daniell, 1978) leads to Gaussian peaks

$$S' = -\int_{V_r} \rho(\mathbf{r}) \ln[\rho(\mathbf{r})]\, \mathrm{d}V_r. \tag{21}$$

It might therefore be preferable to maximize (21) in crystallographic applications.

It should be noted that the maximization of S' is in general *not* equivalent to the MDM. In this regard we disagree with Gassmann's (1976) claim that the MDM is equivalent to maximizing any general function

$$\int_{V_r} \rho(\mathbf{r}) f[\rho(\mathbf{r})]\, \mathrm{d}V_r.$$

As shown by (7), for the one-dimensional case, maximizing integrals of different functions of $\rho(\mathbf{r})$ is equivalent to maximizing different functions of the Karle–Hauptman matrix. It is only the form (9) that leads to the MDM. However, when the structure has equal, spherical, resolved atoms, then all functions are equivalent and agree with the MDM in the large-d limit.

In the MEM with S, (16) is the relation which determines the phases of reflections whose amplitudes are known. Since $\theta_{\mathbf{H}}$ is a function of all the $F_{\mathbf{H}}$, (16) is a higher-order relation among the phases of the structure factors (which involves their amplitudes as well). It is thus a generalization of the low-order relations such as triplets, quartets, *etc.* (Ladd & Palmer, 1980) which are extensively used.

5. Proof of the theorem

It is proved below that the Karle-Hauptman (KH) determinant is related to the integral over the unit cell of the logarithm of the electron density. The proof is valid in all dimensions. However, for notational convenience, it is presented for the two-dimensional case. The fractional coordinates x and y run from $-\frac{1}{2}$ to $+\frac{1}{2}$.

$$\rho(x,y) = \sum_{m,n=-\infty}^{\infty}\sum F_{m,n} \exp[2\pi i(mx+ny)]. \tag{22}$$

If we let m, n, r and s run from $-p$ to $+p$, the KH matrix A reads

$$A_{mn,rs} = F_{m-r,n-s}. \tag{23}$$

The index pairs mn and rs which label the rows and columns take $(2p+1)^2$ values and this defines the dimension d of the matrix A. The relation to be proved reads

$$\lim_{p\to\infty} \frac{1}{d} \ln \det A = \int\int_{-1/2}^{1/2} \ln[\rho(x,y)]\, \mathrm{d}x\, \mathrm{d}y. \tag{24}$$

The proof goes as follows. The first step is to truncate the Fourier series (22) to the range $-N \le m,n \le N$. This defines a truncated density $\rho_t(x,y)$ (equation 25). In the limit $N \to \infty$, $\rho_t(x,y)$ is equivalent to $\rho(x,y)$ in (24) (see equations 26, 27). Next the integral in (24) is approximated by a discrete sum with d terms (equation 28) which is again permissible in the limit $d \to \infty$. The sum of sampled values of $\ln \rho_t$ is then written (equations 29, 30) as the logarithm of a diagonal determinant $\det A'''$. By a unitary transformation (equations 31 to 34), this becomes a determinant made up of structure factors. Two further modifications (equations 39, 40) are needed to obtain the left-hand side of (24). It is essential to verify that the errors in these modifications vanish as $d \to \infty$. This requires representing the determinant as an integral, as shown in the Appendix.

The set of m,n such that $-N \leq m,n \leq N$ is called R_N. The truncated density is then given by

$$\rho_t(x,y) = \sum_{m,n \in R_N}\sum F_{m,n} \exp[2\pi i(mx + ny)]. \tag{25}$$

We assume that $\rho(x,y)$ has a minimum value ρ_m which is non-zero, and also that the amplitudes of the Fourier coefficients are bounded by $C \exp[-(m^2 + n^2)/N_S^2]$. Both these are assured by zero-point and thermal vibrations. In practice, the existence of a positive ρ_m is further ensured since F_0 will be adjusted, as discussed in § 3. From the uniform convergence of the Fourier series it follows that, for any ε however small, it is possible to find an N_ε such that

$$|\rho_t(x,y) - \rho(x,y)| \leq \varepsilon \quad \text{for all } x,y;\ N \geq N_\varepsilon. \tag{26}$$

Choosing ε less than ρ_m, we get a positive lower bound ρ_{tm} for ρ_t. The error in replacing $\ln \rho$ by $\ln \rho_t$ is therefore controllable, *i.e.*

$$\left|\int\int \ln \rho \, dx \, dy - \int\int \ln \rho_t \, dx \, dy\right| \leq \varepsilon/\rho_{tm}. \tag{27}$$

Likewise, the error on replacing the integral of the continuous function $\ln \rho_t$ by a sum is clearly controllable, *i.e.*

$$\left|\int\int \ln \rho_t \, dx \, dy - \frac{1}{(2p+1)^2} \times \sum_{m,n=-p}^{p}\sum \ln\left\{\rho_t\left(\frac{m}{2p+1}, \frac{n}{2p+1}\right)\right\}\right| \leq \delta;\ p > p_\delta. \tag{28}$$

We choose $p \geq N$.

We now rewrite the sum in (28) as follows

$$\frac{1}{(2p+1)^2}\sum_{m,n=-p}^{p}\sum \ln\left\{\rho_t\left(\frac{m}{2p+1}, \frac{n}{2p+1}\right)\right\} = \frac{1}{d} \ln \det A''', \tag{29}$$

where

$$A'''_{mn,rs} = \delta_{mr}\,\delta_{ns}\,\rho_t\left(\frac{m}{2p+1}, \frac{n}{2p+1}\right). \tag{30}$$

Let us introduce A'' defined as follows:

$$A'' = U^{-1} A''' U, \tag{31}$$

where the unitary matrix U is given by

$$U_{mn,rs} = \frac{1}{(2p+1)} \exp\left[\frac{2\pi i}{(2p+1)}(mr + ns)\right] \tag{32}$$

$$U^{-1}_{kl,mn} = \frac{1}{(2p+1)} \exp\left[-\frac{2\pi i}{(2p+1)}(km + ln)\right]. \tag{33}$$

From (31), (32) and (33)

$$A''_{kl,rs} = \frac{1}{(2p+1)^2}\sum_{m,n=-p}^{p}\sum \rho_t\left(\frac{m}{2p+1}, \frac{n}{2p+1}\right) \times \exp\left[-\frac{2\pi i}{(2p+1)}\{m(k-r) + n(l-s)\}\right]. \tag{34}$$

Substituting (25) into (34), we find that the elements of A'' are just the structure factors. Since $k-r$ and $l-s$ both range from $-2p$ to $+2p$, we reduce these to the range $-p$ to $+p$ by adding or subtracting $(2p+1)$. Denoting the result of this operation by $\text{red}(k-r)$ and $\text{red}(l-s)$ (red for reduced), we then find that

$$A''_{kl,rs} = F_{\text{red}(k-r),\,\text{red}(l-s)};\ \{\text{red}(k-r), \text{red}(l-s)\} \in R_N \tag{35}$$

$$= 0 \text{ otherwise.} \tag{36}$$

We finally prove the desired result, (24), by showing that we can replace A by A''. Let us define the following matrix A'.

$$A'_{mn,rs} = F_{m-r,n-s};\ (m-r, n-s) \in R_N \tag{37}$$

$$= 0 \text{ otherwise.} \tag{38}$$

We therefore have to show that for N and p sufficiently large,

$$\left|\frac{1}{d} \ln \det A' - \frac{1}{d} \ln \det A''\right| < \varepsilon_2 \tag{39}$$

$$\left|\frac{1}{d} \ln \det A - \frac{1}{d} \ln \det A'\right| < \varepsilon_1. \tag{40}$$

The full proof of (39) and (40) is involved and is hence given in the Appendix. However, we point out that (39) and (40) are similar to certain well-known results in the statistical mechanics of crystals. Let us regard the Hermitian forms (see equation 41 in the Appendix) associated with the matrices A, A' and A'' as Hamiltonians with x'_{mn}, x''_{mn} playing the role of coordinates and momenta at the site mn of a two-dimensional crystal. A describes a situation in which every site is coupled to every other, with the interaction $F_{mn,rs}$ falling exponentially (we have assumed an exponential bound for the structure factors). A' describes a lattice in which the interaction is cut off outside a square of side N. A'' describes the same cut-off interaction but with periodic boundary conditions. $1/d \ln \det A$ is the free energy per site. Equations (39) and (40) express the intuitively obvious results that, in the limit of a large crystal, the free energy per site is unchanged by introducing periodic boundary conditions or by truncating the tail of the interaction.

APPENDIX

The proof of equations (39) and (40) involves the following identity for a positive definite Hermitian $d \times d$ matrix B. Let x be a $d \times 1$ column vector with complex elements $x'_1 + ix''_1, \ldots, x'_d + ix''_d$. The Hermitian form $H(B)$ is defined by

$$H(B) = x^\dagger Bx. \tag{41}$$

Denoting integration with respect to $x'_1, x''_1, \ldots, x'_d, x''_d$ by $\mathrm{d}\mathbf{x}$, we define $F(B)$ by

$$F(B) = -\frac{1}{d} \ln[\textstyle\int \exp\{-H(B)\}\, \mathrm{d}\mathbf{x}]. \tag{42}$$

One can introduce a new set of variables by a unitary transformation V which diagonalizes H. The Jacobian is unity:

$$x = Vy; \quad x^\dagger Bx = y^\dagger \Lambda y = \lambda_1 |y_1|^2 + \ldots + \lambda_d |y_d|^2, \tag{43}$$

where $\Lambda = V^\dagger BV$ is diagonal and we choose the y's so that $\lambda_1, \lambda_2, \ldots, \lambda_d$, the real positive eigenvalues of B, are in ascending order. Inserting (43) into (42) and integrating we obtain

$$F(B) = -\ln \pi + \frac{1}{d} \ln(\lambda_1 \lambda_2 \ldots \lambda_d) = -\ln \pi + \frac{1}{d} \ln \det B. \tag{44}$$

We now introduce the following notation for averages

$$\langle X \rangle_B = \frac{\int X \exp[-H(B)]\, \mathrm{d}\mathbf{x}}{\int \exp[-H(B)]\, \mathrm{d}\mathbf{x}}, \tag{45}$$

where X is a function of the x_i's. We then have

$$\langle x_i^* x_i \rangle = \sum_j |V_{ij}|^2/\lambda_i \le 1/\lambda_1, \tag{46}$$

where we have used the fact that V is unitary and that λ_1 is the smallest eigenvalue of B. Using the Schwartz inequality we also have

$$|\langle x_i^* x_j \rangle \le [\langle x_i^* x_i \rangle \langle x_j^* x_j \rangle]^{1/2} \le 1/\lambda_1. \tag{47}$$

To prove (39) and (40) we need an upper bound for $|F(B_1) - F(B_2)|$. We define

$$B(\lambda) = (1 - \lambda) B_1 + \lambda B_2; \quad 0 \le \lambda \le 1 \tag{48}$$

$$B(0) = B_1; \quad B(1) = B_2. \tag{49}$$

We further define $f(\lambda)$ which interpolates between $F(B_1)$ and $F(B_2)$:

$$f(\lambda) = -\frac{1}{d} \ln \textstyle\int \exp\{-H[B(\lambda)]\}\, \mathrm{d}\mathbf{x}. \tag{50}$$

Differentiating (50), we have

$$\frac{\mathrm{d}f}{\mathrm{d}\lambda} = -\frac{1}{d} \langle H(B_1 - B_2) \rangle_{B(\lambda)} \tag{51}$$

$$\frac{\mathrm{d}^2 f}{\mathrm{d}\lambda^2} = -\frac{1}{d} \{\langle [H(B_1 - B_2)]^2 \rangle_{B(\lambda)} - [\langle H(B_1 - B_2) \rangle_{B(\lambda)}]^2\} \le 0. \tag{52}$$

The inequality in (52) arises because the mean square is larger than the square of the mean. Since $\mathrm{d}^2 f/\mathrm{d}\lambda^2 \le 0$, $\mathrm{d}f/\mathrm{d}\lambda$ clearly assumes its greatest absolute value G at either $\lambda = 0$ or $\lambda = 1$. We thus obtain G by substituting either B_1 or B_2 for $B(\lambda)$ in (51).

$$G = \max \left| \frac{\mathrm{d}f}{\mathrm{d}\lambda} \right| = \left| -\frac{1}{d} \langle H(B_1 - B_2) \rangle_{B_{1,2}} \right|. \tag{53}$$

We use G to bound

$$|F(B_1) - F(B_2)| \le \left| \int_0^1 \frac{\mathrm{d}f}{\mathrm{d}\lambda} \mathrm{d}\lambda \right| \le G.$$

Therefore,

$$\left| \frac{1}{d} \ln \det B_1 - \frac{1}{d} \ln \det B_2 \right| \le \left| -\frac{1}{d} \langle H(B_1 - B_2) \rangle_{B_{1,2}} \right|. \tag{54}$$

To prove (39), we use (54) for the case $B_1 = A'$, $B_2 = A''$. Therefore,

$$\left| \frac{1}{d} \ln \det A' - \frac{1}{d} \ln \det A'' \right| \le \frac{1}{d} \langle x^\dagger (A' - A'')\, x \rangle_{A',''}\,. \tag{55}$$

Comparing the definitions of A' and A'' in (35) to (38), a typical non-zero term on the right-hand side of (55) reads

$$|T_1| = |F_{\mathrm{red}(m-r),\, \mathrm{red}(n-s)}|\, |\langle x_{mn}^* x_{rs} \rangle|;$$

$$(m - r, n - s) \notin R_N,$$

$$\{\mathrm{red}(m - r), \mathrm{red}(n - s)\} \in R_N. \tag{56}$$

The geometry of terms of the type (56) is shown in Fig. 1. P is the point (m,n) and Q is (r,s) in a $(2p + 1) \times (2p + 1)$ grid of points. We require P to be within a belt of width N around the boundary and a periodically repeated copy of Q to lie within a $(2N + 1) \times (2N + 1)$ square drawn around P. Counting all such terms and introducing the exponential bound $|F_{m,n}| < C \exp[-(m^2 + n^2)/N_s^2]$ on the structure factors, (55) leads to

$$\left| \frac{1}{d} \ln \det A' - \frac{1}{d} \ln \det A'' \right| < \frac{K_1 C (2p + 1) N_s^3}{\lambda_1 d}, \tag{57}$$

where K_1 is a numerical factor.

Now, λ_1 is the lowest eigenvalue of A' or A'', whichever is lower. The eigenvalues of A'' clearly have a lower bound ρ_{tm} since they are the values of $\rho_t(x,y)$ on a grid (see equation 30). Since A' is the KH matrix associated with $\rho_t(x,y)$, $A' - \rho_{tm} I_d$ (where I_d is the $d \times d$ unit matrix) is the matrix corresponding to $\rho_t - \rho_{tm}$ and this too has non-negative eigenvalues. Thus ρ_{tm} is a lower bound to the eigenvalues of A' as well. Consequently, the right side of (57) clearly tends to zero as $N \to \infty$ ($p \geq N$ also $\to \infty$) and (39) is proved.

We now apply the bound (54) with $B_1 = A'$, $B_2 = A$. We obtain a result similar to (55) with a typical term now reading

$$|T_2| = |F_{m-r,n-s}|\,|\langle x^*_{mn} x_{rs} \rangle|; \quad (m-r, n-s) \notin R_N. \tag{58}$$

In Fig. 1, this implies that P can lie anywhere in the square, but Q should be outside a $(2N+1) \times (2N+1)$ square drawn around P. Summing over the terms and introducing the exponential bound on the structure factors, we obtain

$$\left|\frac{1}{d} \ln \det A - \frac{1}{d} \ln \det A'\right| < \frac{K_2 N_s^2\, C \exp(-N^2/N_s^2)}{\lambda_1}, \tag{59}$$

which again tends to zero as $N \to \infty$, thus proving (40).

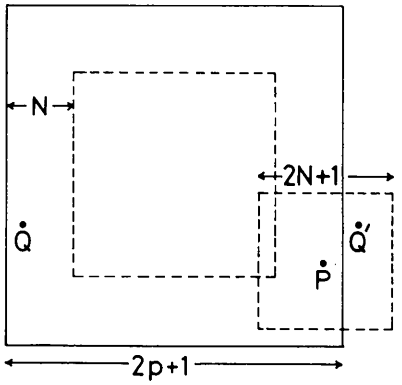

Fig. 1. P and Q are the points (m,n) and (r,s) in a $(2p+1) \times (2p+1)$ square grid. The elements of A' and A'' (see equation 56) differ when P lies within the indicated belt around the boundary and a periodic copy of Q, *viz* Q', lies within the square centred on P. The elements of A and A' (see equation 58) differ for all positions of P provided Q lies outside the square around P.

References

Ables, J. G. (1974). *Astron. Astrophys. Suppl. Ser.* **15**, 383–393.

Bhandari, R. (1978). *Astron. Astrophys.* **70**, 331–333.

Burg, J. P. (1967). Annual meeting of the International Society of Exploration Geophysicists, Oklahoma City. Reprinted in *Modern Spectrum Analysis* (1978), edited by D. G. Childers, pp. 34–41. New York: IEEE Press.

Burg, J. P. (1975). Thesis. Department of Geophysics, Stanford Univ.

Gassmann, J. (1976). *Acta Cryst.* A**32**, 274–280.

Gull, S. F. & Daniell, G. J. (1978). *Nature* (*London*), **272**, 686–690.

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.

Komesaroff, M. M., Narayan, R. & Nityananda, R. (1981). *Astron. Astrophys.* **93**, 269–281.

Ladd, M. F. C. & Palmer, R. A. (1980). *Theory and Practice of Direct Methods in Crystallography*. New York & London: Plenum Press.

Lajzerowicz, J. & Lajzerowicz, J. (1966). *Acta Cryst.* **21**, 8–12.

Narayan, R. & Nityananda, R. (1981). *Curr. Sci.* **50**, 168–170.

Nityananda, R. & Narayan, R. (1982). In preparation.

Nixon, P. E. & North, A. C. T. (1976). *Acta Cryst.* A**32**, 325–333.

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

Smylie, D. E., Clarke, G. K. C. & Ulrych, T. J. (1973). *Methods in Computational Physics*, Vol. 13, pp. 391–430. New York: Academic Press.

Szegö, G. (1920). *Math. Z.* **6**, 167–202 (Theorem 18).

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

Tsoucaris, G. (1980). *Theory and Practice of Direct Methods in Crystallography*, edited by M. F. C. Ladd & R. A. Palmer, pp. 287–360. New York & London: Plenum Press.

Willingale, R. (1981). *Mon. Not. R. Astron. Soc.* **194**, 359–364.