with

$$\mathbf{n}_1 = \mathbf{t}_{i-1} \times \mathbf{t}_i$$

and

$$\mathbf{n}_2 = \mathbf{t}_i \times \mathbf{t}_{i+1}.$$

The method described will work in a mechanical manner without requiring elaborate geometrical constructions. A simple example perhaps will be useful.

In Fig. 1, the sample is at the origin of Cartesian coordinates in which components of vectors are expressed by (x,y,z). Suppose that point $\mathbf{m}_{s+1}$ is variable and is measured experimentally by the spherical coordinate system with components $[r,\psi,\eta]$ which is related to the Cartesian coordinate system by $x = r \sin \psi \cos \eta$, $y = r \sin \psi \sin \eta$, and $z = r \cos \psi$.

Let the scatterer $\mathbf{m}_{s-1}$ be located at (0,–3,0), and let the scatterer, or X-ray source, $\mathbf{m}_{s-2}$, be located at (0,–6,3).

Application of (A1) gives

$$\mathbf{t}_s = r \sin \psi \cos \eta \mathbf{i} + r \sin \psi \sin \eta \mathbf{j} + r \cos \psi \mathbf{k};$$

$$\mathbf{t}_{s-1} = 3\mathbf{j};\ \mathbf{t}_{s-2} = 3\mathbf{j} - 3\mathbf{k}.$$

The terms $(\mathbf{i},\mathbf{j},\mathbf{k})$ are unit vectors in the Cartesian coordinate system.

The term $\mathbf{n}_1$ is determined by

$$\mathbf{n}_1 = \mathbf{t}_{s-2} \times \mathbf{t}_{s-1} = \left\| \begin{matrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 0 & 3 & -3 \\ 0 & 3 & 0 \end{matrix} \right\| = 9\mathbf{i}.$$

In the same manner, $\mathbf{n}_2$ is found to be

$$\mathbf{n}_2 = 3r \cos \psi \mathbf{i} - 3r \sin \psi \cos \eta \mathbf{k}.$$

All of the quantities necessary to solve (A2) and (A3) have been determined. Use of (A2) results in the two total scattering angles $2\theta_{s-1} = \pi/4$ and $2\theta_s = \cos^{-1}(\sin \psi \sin \eta)$.

Use of (A3) results in the angle of rotation between the scattering planes

$$\gamma_{s,s-1} = \cos^{-1}\{(1 + \tan^2 \psi \cos^2 \eta)^{-1/2}\}.$$

Note added in proof. It is emphasized that inverse cosines and square roots result in both a positive and a negative solution. The definitions used allow the correct signs to be selected, but explicit rules may be helpful. The negative solution of (A3) is used if $\mathbf{n}_1 \times \mathbf{n}_2 = L\mathbf{t}_i$ with L positive. The angle φ is negative if $\mathbf{e} \times \mathbf{p} = N\mathbf{s}_0$ with N positive. Also, b must be positive if $2\theta < \pi/2$ and negative if $2\theta > \pi/2$.

References

Azaroff, L. V. (1955). *Acta Cryst.* **8**, 701–704.
Dwiggins, C. W. Jr (1972). *Acta Cryst.* A**28**, 158–163.
Dwiggins, C. W. Jr (1978). *J. Appl. Cryst.* **11**, 615–619.
Dwiggins, C. W. Jr (1980). *J. Appl. Cryst.* **13**, 572–576.
Dwiggins, C. W. Jr & Park, D. A. (1971). *Acta Cryst.* A**27**, 264–272.
Evans, E. H. M., Hine, R., Richards, J. P. G. & Tichy, K. (1980). *J. Appl. Cryst.* **13**, 163–167.
Jennings, L. D. (1981). *Acta Cryst.* A**37**, 584–593.
Le Page, Y., Gabe, E. J. & Calvert, L. D. (1979). *J. Appl. Cryst.* **12**, 25–26.
McMaster, W. H. (1954). *Am. J. Phys.* **22**, 351–362.
McMaster, W. H. (1961). *Rev. Mod. Phys.* **33**, 8–28.
Mathieson, A. McL. (1978). *Acta Cryst.* A**34**, 404–406.
Mathieson, A. McL. (1982). *Acta Cryst.* A**38**, 739–740.
Stokes, G. G. (1852). *Trans. Cambridge Philos. Soc.* **9**, 3, 399–416.
Templeton, D. H., Templeton, L. K., Phillips, J. C. & Hodgson, K. O. (1980). *Acta Cryst.* A**36**, 436–442.
Vincent, M. G. (1982). *Acta Cryst.* A**38**, 510–512.
Vincent, M. G. & Flack, H. D. (1980). *Acta Cryst.* A**36**, 614–620.

Acta Cryst. (1983). A**39**, 777–783

Extrapolative Filtering.
II. Formalism for a Maximum-Entropy Enhancement of Resolution in *n* Dimensions

By Douglas M. Collins and Martin C. Mahar

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

(*Received* 3 *January* 1983; *accepted* 10 *May* 1983)

Abstract

An extrapolative filtering formalism is used to determine a high-resolution positive-definite density estimator. In an example, comparison with results from Fourier transformation of structure factors shows a resolution enhancement factor of 1·3 for the density estimator. The density estimator satisfies a maximum-entropy criterion and is the direct-space representation of a maximum determinant of structure factors as used in phase determination.

1. Introduction

In a previous paper (Collins, 1978), hereafter referred to as paper I, a foundation was laid for extraction of

0108-7673/83/050777-07$01.50

enhanced-resolution density functions from a given set of X-ray diffraction data. As in paper I, the phase problem is not at issue in this work and it is assumed that available data are in the form of structure factors, both magnitudes and phases. It is also assumed, as before, that the functions in which resolution is to be maximized are density functions which are everywhere positive. This assumption has been discussed recently by Collins & Mahar (1983).

Paper I left as open the question of applicability of extrapolative filtering in n dimensions. Here we show that usable n-dimensional forms can be found and also how the corresponding densities relate to determinant maximization (Tsoucaris, 1970). The crucial problems in developing a three-dimensional filtering formalism arise upon going from one to two dimensions. Thus the two-dimensional worked problem serves as an illustration for any number of dimensions beyond one. This corresponds to Helson & Lowdenslager's (1958) statement that 'analytic function theory divides into two distinct disciplines in higher dimensions', where reference is made to function theory on the unit circle as contrasted with function theory on the unit bicylinder. This state of affairs is obvious in practice as well, most notably in the loss of Toeplitz-form matrices corresponding to Karle–Hauptman determinants, in passing from one dimension to two. Nevertheless, Helson & Lowdenslager proved there exists a solution to the prediction (extrapolation) problem in several variables, and it is their half-lattice construction which underlies the present work.

2. A density estimator

The following construction and theorem are taken from Helson & Lowdenslager (1958) who also provide the appropriate proofs and analysis. On a two-dimensional primitive net whose points are represented by integral coordinates (m,n), S is a half-lattice if

(*a*) (0,0) does not exist in S,

(*b*) (m,n) exists in S if and only if $(-m,-n)$ does not exist in S unless $m = n = 0$,

(*c*) (m,n) exists in S and (m',n') exists in S imply $(m + m', n + n')$ exists in S.

The half-lattice so defined is employed in a slightly weakened form of Helson and Lowdenslager's theorem 2:

Let p be summable on the bicylinder and given by the Fourier series

$$p(e^{ix}, e^{iy}) = P_{00} + \sum_S P_{mn} \exp\{-2\pi i(mx + ny)\} \quad (2.1)$$

where S is any half-lattice. Then

$$\int \ln|p| \mathrm{d}x\, \mathrm{d}y \geq \ln|P_{00}|. \quad (2.2)$$

For present purposes, the important result is that $|p| > 0$ provided it is composed from a half-lattice of coefficients and $|P_{00}| > 0$. Furthermore, if p is summable on the unit bicylinder and $|p| > 0$, then p^{-1} is summable and $|p^{-1}| > 0$ as well. This allows a formal two-dimensional analogy to the one-dimensional density estimator given in paper I. The one-dimensional estimator is

$$^{c}\rho_x = (\sigma_n^2/L)\left|C_0 + \sum_{k=1}^{n} C_k \exp\{-2\pi ikx\}\right|^{-2}, \quad (2.3)$$

where (σ_n^2/L) is constant for a given order, and the coefficients C_k correspond to an extrapolation-error filter. The two-dimensional formal analog to (2.3) is

$$^{c}\rho_{xy} = (\sigma_n^2/A)|C_{00} + \sum_S C_{hk} \exp\{-2\pi i(hx + ky)\}|^{-2} \quad (2.4)$$

in which the coefficients C_{hk} (yet to be determined) must provide a finite nonzero modulus in (2.4). It follows from (2.1) and (2.2) that the density estimation $^{c}\rho_{xy}$ is both positive definite and well behaved if as in the one-dimensional case the computation of C_{hk} is constrained to give $C_{00} = 1{\cdot}0$.

In the course of proving another theorem, Helson & Lowdenslager (1958) also show that if g (real) is positive and summable on the bicylinder then there is a unique function H given by the series

$$H(x,y) = \sum_S C_{mn} \exp\{-2\pi i(hx + ky)\} \quad (2.5)$$

such that

$$g = \theta|1 + H|^{-2}, \quad (2.6)$$

where θ is a constant. Of course H is not unique in the trivial sense that S may be any half-lattice, but, for definite S, it follows that a density estimator of the desired form exists and is unique. Evidently the search for an estimator of electron density is restricted only by the requirement that an estimator be of the form of (2.4), $C_{00} = 1{\cdot}0$. Moreover, if in agreement with the data such a form is found by any means, it is in fact the unique desired result.

Construction of an electron-density estimator in crystallographic application is dependent on the inherent positivity of the electron density ρ. Positivity may be imposed by the assumption

$$\rho_{xy} = |g_{xy}|^2 > 0. \quad (2.7)$$

It is immediately clear that for some positive constant κ there is c such that

$$\kappa|g_{xy}|^{-1} = |c_{xy}| \quad (2.8)$$

and

$$gc = \kappa \exp\{i\varphi\} \quad (2.9)$$

provided only that c has the form of (2.1), $C_{00} = 1{\cdot}0$.

The Fourier transform of (2.9) may be represented in matrix notation as

$$\mathbf{G}\mathbf{C} = \boldsymbol{\varepsilon}; \tag{2.10}$$

where $G_{\mathbf{h}}$, the transform of $g_{\mathbf{x}}$, is a complete set of coefficients; $C_{\mathbf{h}}$, the transform of $c_{\mathbf{x}}$, vanishes off the half-lattice S, $C_{00} = 1{\cdot}0$; and $\varepsilon_{\mathbf{h}}$ may be interpreted as the extrapolation error, that is, the failure of the filter C to combine elements of $\mathbf{G}$ on a shifted half-lattice to yield $G_{\mathbf{h}}$, the element at the shifted origin. Equation (2.7) requires that F, the transform of ρ, be related to $\mathbf{G}$ and its Hermitian transpose, $\mathbf{G}^{\dagger}$, by

$$\mathbf{F} = \frac{1}{A}\,\mathbf{G}^{\dagger}\mathbf{G}, \tag{2.11}$$

where A is the area of a unit cell. Evidently

$$\frac{1}{A}\,\mathbf{G}^{\dagger}\mathbf{G}\mathbf{C} = \frac{1}{A}\,\mathbf{G}^{\dagger}\boldsymbol{\varepsilon} \tag{2.12}$$

is the same as

$$\mathbf{F}\mathbf{C} = \mathbf{B}, \tag{2.13}$$

in which, according to the analysis of Appendix B in paper I, the only nonzero element $\boldsymbol{\beta}$ is $\beta_0 = \sigma_n^2$, the variance given by

$$\frac{1}{A}\,\boldsymbol{\varepsilon}^{\dagger}\boldsymbol{\varepsilon}. \tag{2.14}$$

The computation of $C_{\mathbf{h}}$ follows immediately as the solution of

$$\begin{bmatrix} F_{00}\,F_{01} & \cdot & \cdot & \cdot & F_{0n} \\ F_{10}\,F_{11} & & & & \\ \vdots & & & & \\ F_{n0}\,F_{n1} & \cdot & \cdot & \cdot & F_{nn} \end{bmatrix} \begin{bmatrix} C_0 = 1 \\ C_1 \\ \vdots \\ C_n \end{bmatrix} = \begin{bmatrix} \sigma_n^2 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{2.15}$$

in which $C_{\mathbf{h}_i}$ has been replaced by C_i and F_{ij} is in the place of $F_{h_i - h_j}$.

It is clear from (2.9) and (2.10) that the autocorrelation of ε, apart from its value of σ_n^2 at the origin, is zero. Its transform then is σ_n^2/A, and upon replacement of c by its Fourier representation, the squared modulus of (2.9) becomes

$${}^{c}\rho_{\mathbf{x}} = \frac{\sigma_n^2/A}{\left|1 + \sum\limits_{S} C_{\mathbf{h}} \exp\{-2\pi i \mathbf{h}.\mathbf{x}\}\right|^2}, \tag{2.16}$$

the desired density estimator in which ${}^{c}\rho$ has been written for $|g|^2$. A value for n is not fixed and it corresponds in some sense to the degree of approximation. For $n = 1$, ${}^{c}\rho = F_{00}/A$ everywhere; but for point atoms at rest and n somewhere in the neighborhood of N, the number of atoms in a unit cell, $\mathbf{F}$ becomes semidefinite (Goedkoop, 1950) and the whole formalism breaks down.

3. Construction of an estimator

It is not immediately evident how a half-lattice is to be selected. If the special quality of a function on a half-lattice is termed one-sidedness, then the problem is to impose one-sidedness on C with due regard for experimental conditions. The problem is clarified by consideration of the Hermitian form

$$\mathbf{C}^{\dagger}\mathbf{F}\,\mathbf{C} = \sigma_n^2 = \sum_{i,j} C_i^{*} F_{ij} C_j. \tag{3.1}$$

In addition to the requirement that $\{\mathbf{h}_i\}$ describe a half-lattice on which C_i is to be nonvanishing, the experimental bound is that every $\mathbf{h}_{ij} = \mathbf{h}_i - \mathbf{h}_j$ be in the range $0 \le |\mathbf{h}| \le 2\sin\theta_{\max}/\lambda$, and consequently that $F_{\mathbf{h}_i - \mathbf{h}_j} = F_{ij}$ lie in the range of experimental observation.

In this context it is straightforward to make C suitably one-sided. Identify a subset of $\{\mathbf{h}\}$ by $\mathbf{s}$ which is defined by $0 \le |\mathbf{s}| \le \sin\theta_{\max}/\lambda$. Suppose that one vector $\mathbf{s}'$ is larger in magnitude than all others. Now

$$\mathbf{h} = \mathbf{s} - \mathbf{s}' \tag{3.2}$$

describes a lattice portion on which any function is one-sided, it includes $\mathbf{h} = (0,0)$, and the crucial requirement

$$0 \le |\mathbf{h}_{ij}| = |\mathbf{s}_i - \mathbf{s}_j| \le 2\sin\theta_{\max}/\lambda \tag{3.3}$$

is fulfilled.

At any order n in (2.15), the following construction of a density estimator systematically builds the structure-factor array of entries for which the maximum of $|\mathbf{h}_{ij}|$ is as small as possible. Suppose the sets $\{\mathbf{s}_i\}$ and $\{\mathbf{h}_i\}$ have been ordered by their magnitudes, smallest first so that $|\mathbf{s}_0| = |\mathbf{h}_0| = 0$. For given n construct (2.15) with $C_{\mathbf{s}_i}$, $i = 0, 1, \ldots, n$. It turns out that (2.15) as constructed is satisfactory, but needs rearrangement because C is not one-sided. This condition may be satisfied by replacing $\mathbf{s}$ according to (3.2). The structure-factor array is unaffected by this change inasmuch as

$$\mathbf{h}_{ij} = \mathbf{s}_i - \mathbf{s}_j = (\mathbf{s}_i - \mathbf{s}') - (\mathbf{s}_j - \mathbf{s}'), \tag{3.4}$$

but the identity of the elements of $\mathbf{C}$ is changed with the last element becoming C_{00}. Now (2.15) may be put in final form by reordering of $\mathbf{C}$ so that new $\mathbf{h}$ obtained by (3.2) forms a nondecreasing sequence

$$|\mathbf{h}_i| \le |\mathbf{h}_j|,\ i \le j \tag{3.5}$$

and by subsequent appropriate interchange of rows and columns in $\mathbf{F}$.

For applications, construction of (2.15) is quite simple. From the ordered set $\{\mathbf{s}_i\}$ select the first $n + 1$ elements, find the set $\{\mathbf{h}_i\}$ by (3.2), reorder it to satisfy (3.5), and make the required entries in $\mathbf{F}$. To solve the equations, set $\sigma^2 = 1{\cdot}0$, compute $\mathbf{C}$, then rescale the result so that $C_0 = 1\cdot0$ and σ_n^2 has its proper value. Density estimation follows immediately by (2.16).

4. An example

The two-dimensional centrosymmetric projection of hexamethylbenzene reported by Brockway & Robertson (1939) was used as the basis for test computations illustrating the use of our formalism. For hexamethylbenzene the space group is $P\bar{1}$ with one half the molecule in the asymmetric unit and unit-cell constants $a = 8\cdot92$ (2), $b = 8\cdot86$ (2), $c = 5\cdot30$ (1) Å, $\alpha = 44\cdot45$, $\beta = 116\cdot72$, $\gamma = 119\cdot57°$. After least-squares adjustment of atomic coordinates for the projection and determination of an overall isotropic thermal parameter to be $B = 3\cdot0\,(3)$ Å^2, the resulting calculated structure factors were taken as error-free data and, except as noted, were used without change in the calculations. For this model hydrogen atoms were ignored and $R = 0\cdot21$.

Procedures necessary to the formation and solution of (2.15) and (2.16) were reduced to a computer program and density estimates were obtained for most values of n in the range $8 \le n \le 64$. In this case, as in any for which structure factors have not been altered to correspond to point atoms at rest, n may be much greater than N, the number of atoms in a unit cell. The smallest value of n for which ${}^{c}\rho$ gives straightforward resolution of the (carbon) atoms is 52; $n = 54$ leads to similar but subjectively better results for little additional computation. Order 54 corresponds to use of $\mathbf{h}_n = (\bar{4},3,0)$ and all other $\mathbf{h}$ with smaller magnitude in the summation of (2.16); $\sigma_{54}^2 = 37$. The projection density obtained by normal Fourier synthesis of the complete set of (calculated) two-dimensional data is given in Fig. 1(*a*). The density obtained as ${}^{c}\rho$ for order $n = 54$ is given in Fig. 1(*b*). For comparison with Fig. 1(*b*), the density ρ given in Fig. 1(*c*) was composed by standard Fourier synthesis from the 54 unique data of lowest resolution.

It is clear that a coefficient set of given size has potentially greater resolving power when used in determining ${}^{c}\rho$ in contrast to ρ. Heuristic interpretation of the densities and their differences suggests that resolution gain in ${}^{c}\rho$ is obtained, at least in part, through an imposed peakiness. In a design to force greater resolution through increased peakiness, the diagonal elements of $\mathbf{F}$ in (2.15) were reduced in successive iterations by arbitrary fractions of its minimum eigenvalue until the ratio of largest to smallest eigenvalues was ~100. To be certain the results were not misleadingly conditioned by data perfection, the experimental data (Brockway & Robertson, 1939) with calculated phases were used without modification for this test except for sharpening through overcorrection for thermal motion; the correction factor was $\exp\{B \sin^2\theta/\lambda^2\}$ with $B = 6\cdot0$ Å^2.

A series of calculations were carried out for each of $n = 8, 12, 16, 20, 24$. Although ${}^{c}\rho$ displays substantial information even for $n = 12$, adequate qualitative representation of the structure as peaks near atomic positions does not occur until $n = 20$. At $n = 20$ the best representation of several was for F_{00} reduced from $90\cdot0$ to $69\cdot5$ electrons and maximum and minimum eigenvalues of $\mathbf{F}$ at 180 and $1\cdot37$ respectively. Each of the six independent atoms was represented by a peak

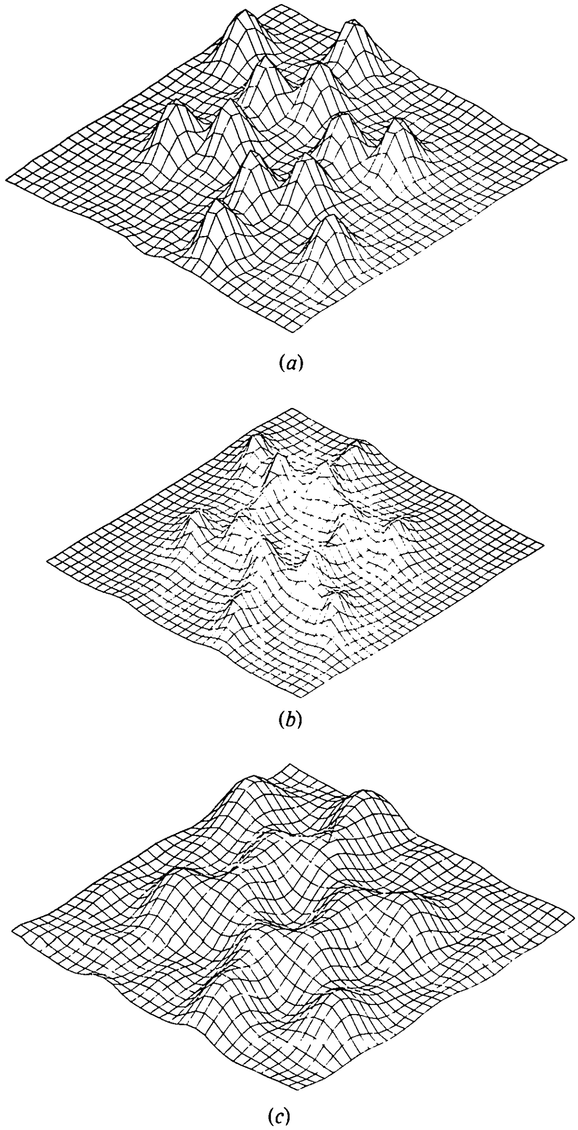

Fig. 1. Electron density for the hexamethylbenzene structure projected along its c axis computed by (*a*) standard Fourier synthesis of (calculated) structure factors, (*b*) evaluation of order 54 ${}^{c}\rho$, and (*c*) standard Fourier synthesis of the 54 lowest-order unique structure factors.

not more than 0·4 Å from the correct location, but in three cases the peaks were double. The map was excessively spiky with a peak maximum of 77 e $Å^{-2}$ in a map which averaged to 2 e $Å^{-2}$; the principal peak maxima lay in the range 20–77 e $Å^{-2}$ and only 22% of the map was above 1 e $Å^{-2}$. For $n = 20$ the range of $\sin\theta/\lambda$ represented in $\mathbf{F}$ is [0, 0·36 $Å^{-1}$] for which the minimum interplanar spacing is 1·37 Å. The smallest interatomic separation resolved in the projection is a foreshortened 1·06 Å. This may be compared with the failure of a standard Fourier synthesis of perfect data to resolve the same interatomic separation with data extending to a minimum interplanar spacing of 1·19 Å as shown in Fig. 1(*c*).

5. Maxima in entropy and determinants

Equation (2.9) leads immediately to

$$\int_{\text{cell}} \ln|g_{\mathbf{x}}|^2\,d\mathbf{x} = \ln\kappa^2 + \int_{\text{cell}} \ln\left|\frac{1}{c_{\mathbf{x}}}\right|^2 d\mathbf{x}. \tag{5.1}$$

This corresponds exactly to Shannon's theorem 14 (Shannon & Weaver, 1949) concerning entropy per degree of freedom for ensembles of functions. In (5.1) the ensembles are related to the unspecified nature of φ, the phase to be assigned to κ, and the similarly arbitrary phase of g. For the ensemble of functions G related to ε by passing the ε ensemble through the filter C^{Inv}, (5.1) gives the entropy per degree of freedom for ensemble F (Shannon & Weaver). Because the second moments of the n-dimensional distribution of G are fixed at F, the maximization of entropy for the ensemble of synthetic processes G is, within a constant factor, achieved by formal maximization of the left-hand side of (5.1) subject to satisfaction of fixed F.

The constraint to be imposed is F, the autocorrelation of G, defined as

$$\int_{\text{cell}} {}^{c}\rho_{\mathbf{x}} \exp\{2\pi i\mathbf{h}.\mathbf{x}\}\,d\mathbf{x} = F_{\mathbf{h}}, \tag{5.2}$$

where ${}^{c}\rho$ is the density to be determined and $\mathbf{h}$ is drawn from the set T which is of the type $\{\mathbf{h}_{ij}\}$ as it appears in (2.15). The expression to be maximized is

$$\int_{\text{cell}} \left\langle \ln {}^{c}\rho_{\mathbf{x}} - \sum_{h\in T} \lambda_{\mathbf{h}}\,[{}^{c}\rho_{\mathbf{x}} \exp(2i\mathbf{h}.\mathbf{x}) - F_{\mathbf{h}}] \right\rangle d\mathbf{x}, \tag{5.3}$$

in which Lagrange multipliers have been introduced. Maximization with respect to ${}^{c}\rho_{\mathbf{x}}$ results in

$${}^{c}\rho_{\mathbf{x}} = \frac{1}{\sum_{\mathbf{h}\in T} \lambda_{\mathbf{h}} \exp\{2\pi i\mathbf{h}.\mathbf{x}\}}. \tag{5.4}$$

Evidently the entropy maximization requires that density be the inverse of a summable positive-definite function. This condition is already satisfied by the estimator given in (2.16) which consequently is a maximum-entropy estimator of electron density.

The distribution of random variables which yields a maximum of entropy given by

$$-\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} p(x_1,\dots,x_n) \times \ln p(x_1,\dots,x_n)\,dx_1\dots dx_n \tag{5.5}$$

when the second moments are fixed is the Gaussian distribution with those moments (Britten & Collins, 1982; McDonough, 1974; Shannon & Weaver, 1949). Evidently ${}^{c}\rho$ corresponds to the Gaussian distribution of G which has $F_{\mathbf{h}}$, $\mathbf{h} \in T$, as its second moments. But it is also the case that with

$$F_{ij} = E\{G_{ik}G^{*}_{kj}\}, \tag{5.6}$$

F a complex expectation, the entropy for the ensemble of G realizations is given both by (5.1) and, within a constant, by

$$\ln[(\pi e)^{n+1}|\mathbf{F}|] \tag{5.7}$$

(Britten & Collins, 1982; McDonough, 1974; Shannon & Weaver, 1949), $\mathbf{F}$ and n as given in (2.15). Thus a maximum in the determinant $|\mathbf{F}|$ corresponds to a maximum-entropy distribution for G as represented by ${}^{c}\rho$. A similar result is given by Narayan & Nityananda (1982).

The maximum-determinant method for phase determination (Tsoucaris, 1970) does not use a densely sampled space of structure factors. This is in sharp contrast with the construction of ${}^{c}\rho$ as given in §3. The contrast represents not a difference of principle, but a difference in strategy. Whatever the motivation or strategy in the construction of a particular determinant for maximization and phase determination, the reciprocal-lattice vectors identifying the first-column entries may be identified as $\{\mathbf{h}'_i\}$. By the following trivial rearrangement these may be changed to delimit a half-lattice, origin ignored: find $\mathbf{h}'_i$ of largest magnitude, say h'', and calculate $\mathbf{h}_i = \mathbf{h}'_i - \mathbf{h}''$. After relabeling $\mathbf{h}$ to form a nondecreasing sequence as given by inequality (3.5) then (2.15) can be constructed. The elements of $\mathbf{F}$ are rearranged but identical and the corresponding density estimate follows directly by solution of the equation and evaluation of ${}^{c}\rho$ as given by (2.16).

It is clear that in $\{\mathbf{h}_i\}$ constructed from $\{\mathbf{h}'_i\}$ any number or pattern of potential members may be missing since the original determinant construction need not involve one-sidedness. Moreover, it is not necessary that even one of $\{\mathbf{h}'_i\}$ be found in $\{\mathbf{h}_i\}$, except possibly for $\mathbf{h}_0 = \mathbf{h}'_0 = (0,0)$. Nevertheless, the equivalence of entropy expressions (5.1) and (5.7) makes it clear that for any determinant maximized in the sense of phase determination (Tsoucaris, 1970), ${}^{C}\rho$ is the correct corresponding density. Both the determinant and the density correspond to maximization of entropy for an ensemble of synthetic functions in reciprocal space with second moments fixed at F.

6. Discussion

We have presented an estimator of electron density at enhanced resolution. It was derived using an extrapolative filtering formalism but shown to be a maximum-entropy estimator as well. Elsewhere (Collins, 1982) a different estimator of electron density was described as computed by iterative entropy maximization. This too is a maximum-entropy estimator, but in regard to a different objective function, and may be identified as ${}^{\mathrm{MEM}}\rho$.

For both density estimators the Shannon (Shannon & Weaver, 1949) formulation of entropy has been used as given in (5.5). In the case of ${}^{\mathrm{MEM}}\rho$ the entropy actually maximized (Collins, 1982) is a modified form of (5.5), which, in any event, is set in three-dimensional Cartesian space and uses the probability density for an electron at any point, that is, the electron density itself. In the case of ${}^{C}\rho$, the entropy expression is set in an n-dimensional space and the probability density is for the n random variables G, a synthetic process in reciprocal space.

So long as the information sought in an X-ray diffraction experiment is to be found in the electron density, a density to be preferred is one which is as smooth as possible in agreement with the data. This describes the maximum-entropy criterion for entropy calculated on the electron distribution itself and ${}^{\mathrm{MEM}}\rho$ is such a function. On the other hand, if a set of structure factors is to be studied while conceding a maximum of ignorance concerning those not members of the set, the corresponding proper density is ${}^{C}\rho$.

There are two serious shortcomings in the computation of ${}^{C}\rho$. In contrast to many calculations involving structure factors, there is no evident way to introduce weights in recognition of prior or experimental uncertainties. Secondly, the common profusion of diffraction data cannot be exploited because of the limitation on the order of **F** in (2.15). While the inability to include weighting has no discernible related advantage, the limitation on the order of **F** is directly related to one of the distinctive properties of ${}^{C}\rho$.

If a set of structure factors completely describes a structure in the sense that a density of n distinct δ functions is entirely specified by **F** of order n (Goedkoop, 1950), then a density estimation which adequately represents this must be the true density. As Goedkoop points out, all $|\mathbf{F}|$ of higher order are zero and thus they could not be used to generate a new ${}^{C}\rho$. This corresponds to the property of ${}^{C}\rho$ which concedes maximum ignorance concerning samples of F unused in its computation (Ponsonby, 1973). Of course, there is no ignorance to concede after a structure is completely specified and ${}^{C}\rho$ therefore should be not subject to change by additional samples of F. The advantage to this is clearly reflected in the maximum-determinant method of phase determination (Tsoucaris, 1970) which is based on a (relatively) small subset of structure factors. In this method it is desired that a phase determination be subject to as little change as possible upon consideration of additional structure factors not included in **F**. This is just the specification for ${}^{C}\rho$ representing the maximum-entropy distribution of a synthetic process G, and ${}^{C}\rho$ is the proper maximum-entropy density representation of the maximum determinant as it is used in phase determination.

Resolution enhancement or super-resolution was the primary aim of this work and it occurs in ${}^{C}\rho$. It is not clear how super-resolution is to be measured but a number corresponding to the effectiveness of structure factors for the location of atoms is the ratio (shortest interatomic separation resolved in a density): (minimum interplanar spacing of the Fourier coefficients used in its computation). For the example of §4 this ratio is 1·01 for the standard Fourier synthesis shown in Fig. 1(*c*). For the computation of ${}^{C}\rho$ at forced high resolution, $n = 20$, the ratio is a more favorable 0·77. These ratios suggest that in favourable cases ${}^{C}\rho$ could afford an enhancement of resolution over conventional synthesis of structure factors by a factor of 1·3 in linear measure or 2·2 in the cubic measure of the effectiveness of a single Fourier coefficient. Note that similar super-resolution, although subjectively judged, was observed for the reported calculation of ${}^{\mathrm{MEM}}\rho$ for a protein crystal structure (Collins, 1982).

Computation of ${}^{C}\rho$ requires no preliminary representations of density. This is in contrast to the iterative computation of ${}^{\mathrm{MEM}}\rho$ (Collins, 1982) which requires repeated Fourier syntheses on grids of fineness suitable to the resolution sought in the final density function. Computation of ${}^{C}\rho$ need involve only one Fourier synthesis and at as few or as many points as chosen. The corresponding disadvantage is matrix inversion for **F** which can become an unreasonable computation for large n or a too small signal/noise ratio. It is never certain that any method can rescue an experiment from the latter problem. But so long as n is smaller than the number of atoms in a unit cell the inversion of **F** should be straightforward, except possibly for the very largest crystal structures, and ${}^{C}\rho$ would follow upon routine calculations.

The range of problems for which ${}^{C}\rho$ has practical utility will ultimately be determined in experimental application. It seems clear in advance, however, that ${}^{C}\rho$

can be useful as a density function primarily for cases in which all available (complex) structure factors can be incorporated into a well conditioned matrix of the type given in (2.15). If there are significantly more structure factors available than can be used in one matrix, ${}^{\mathrm{MEM}}\rho$ (Collins, 1982) will give a more authoritative result, but not necessarily at a higher resolution than could be realized in ${}^{C}\rho$. Although the present work has not been predicated on any particular modification of experimental structure-factor moduli, the normalized structure factor corresponding to point atoms at rest is most consistent with a goal of resolution maximization.

Phase determination by the maximum-determinant method (Tsoucaris, 1970) leads to a different kind of application for ${}^{C}\rho$. In this case, ${}^{C}\rho$ has the role of the correct density representation of a maximum determinant irrespective of the size or completeness of a set of structure moduli. Because ${}^{C}\rho$ is the direct-space representation of the information carried by the corresponding matrix of (phased) structure factors, it is the correct initial prior distribution for extending the maximum-determinant phases through an entire data set by the iterative entropy maximization which results in ${}^{\mathrm{MEM}}\rho$ (Collins, 1982).

7. Conclusion

Maximum-entropy methodology and an extrapolative filtering formalism lead to the same density estimate with respect to a synthetic reciprocal-space process having as samples of its autocorrelation the crystallographic structure factors. This density ${}^{C}\rho$ may be compared with a maximum-entropy electron density ${}^{\mathrm{MEM}}\rho$ (Collins, 1982) for which entropy is iteratively maximized on the electron density itself. When electron density is itself the desired object of experimental determination, the principles of information physics (Jaynes, 1957) require the latter use of entropy to define an electron density least likely to be changed on consideration of data outside the experimental range (Benjamin, 1980).

On the other hand, the same principles of information physics (Jaynes, 1957) lead to entropy maximization on a reciprocal-space process to find ${}^{C}\rho$, the density corresponding to a uniquely smooth function F (Ponsonby, 1973) constrained by the observed (complex) structure-factor samples. Although in this case it is the smooth function F which is the object to change as little as possible upon consideration of additional data, its Fourier transform must also be smooth and in some sense subject to minimal change. This latter smoothness corresponds to the extrapolative filtering formalism and enhanced resolution in ${}^{C}\rho$ while the smoothness of F corresponds to phase determination by the maximum-determinant method (Tsoucaris, 1970). So long as a determinant order is less than the number of atoms in a crystallographic unit cell, ${}^{C}\rho$ is routinely calculable to obtain an enhanced-resolution density function which is the direct-space representation 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

Benjamin, R. (1980). *IEE Proc.* F**127**, 341–353.
Britten, P. L. & Collins, D. M. (1982). *Acta Cryst.* A**38**, 129–132.
Brockway, L. O. & Robertson, J. M. (1939). *J. Chem. Soc.* pp. 1324–1332.
Collins, D. M. (1978). *Acta Cryst.* A**34**, 533–541.
Collins, D. M. (1982). *Nature (London)*, **298**, 49–51.
Collins, D. M. & Mahar, M. C. (1983). *Acta Cryst.* A**39**, 252–256.
Goedkoop, J. A. (1950). *Acta Cryst.* **3**, 374–378.
Helson, H. & Lowdenslager, D. (1958). *Acta Math.* **99**, 165–202.
Jaynes, E. T. (1957). *Phys. Rev.* **106**, 620–630.
McDonough, R. N. (1974). *Geophysics*, **39**, 843–851.
Narayan, R. & Nityananda, R. (1982). *Acta Cryst.* A**38**, 122–128.
Ponsonby, J. E. B. (1973). *Mon. Not. R. Astron. Soc.* **163**, 369–380.
Shannon, C. E. & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, Illinois: Univ. of Illinois Press.
Tsoucaris, G. (1970). *Acta Cryst.* A**26**, 492–499.