

1. GENERAL RELATIONSHIPS AND TECHNIQUES

$$\rho(r, \varphi + 2\pi k_1, z + k_2) = \rho(r, \varphi, z).$$

Suppose the fibre now has *helical* symmetry, with u copies of the same molecule in t turns, where $\text{g.c.d.}(u, t) = 1$. Using the Euclidean algorithm, write $u = \lambda t + \mu$ with λ and μ positive integers and $\mu < t$. The period lattice for the (φ, z) dependence of ρ may be defined in terms of the new basis vectors:

I, joining subunit 0 to subunit l in the same turn;

J, joining subunit 0 to subunit λ after wrapping around.

In terms of the original basis

$$\mathbf{I} = \frac{t}{u} \mathbf{i} + \frac{1}{u} \mathbf{j}, \quad \mathbf{J} = \frac{-\mu}{u} \mathbf{i} + \frac{\lambda}{u} \mathbf{j}.$$

If α and β are coordinates along **I** and **J**, respectively,

$$\begin{pmatrix} \varphi/2\pi \\ z \end{pmatrix} = \frac{1}{u} \begin{pmatrix} t & -\mu \\ 1 & \lambda \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

or equivalently

$$\begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \lambda & \mu \\ -1 & t \end{pmatrix} \begin{pmatrix} \varphi/2\pi \\ z \end{pmatrix}.$$

By Fourier transformation,

$$\begin{aligned} \left(\frac{\varphi}{2\pi}, z \right) &\Leftrightarrow (-n, l) \\ (\alpha, \beta) &\Leftrightarrow (m, p) \end{aligned}$$

with the transformations between indices given by the contragredients of those between coordinates, *i.e.*

$$\begin{pmatrix} n \\ l \end{pmatrix} = \begin{pmatrix} -\lambda & 1 \\ -\mu & t \end{pmatrix} \begin{pmatrix} m \\ p \end{pmatrix}$$

and

$$\begin{pmatrix} m \\ p \end{pmatrix} = \frac{1}{u} \begin{pmatrix} -t & 1 \\ \mu & \lambda \end{pmatrix} \begin{pmatrix} n \\ l \end{pmatrix}.$$

It follows that

$$l = tn + um,$$

or alternatively that

$$\mu n = up - \lambda l,$$

which are two equivalent expressions of the *selection rules* describing the decimation of the transform. These rules imply that only certain orders n contribute to a given layer l .

The 2D Fourier analysis may now be performed by analysing a *single* subunit referred to coordinates α and β to obtain

$$h_{m,p}(r) = \int_0^1 \int_0^1 \rho(r, \alpha, \beta) \exp[2\pi i(m\alpha + p\beta)] \, d\alpha \, d\beta$$

and then reindexing to get only the allowed g_{nl} 's by

$$g_{nl}(r) = uh_{-\lambda m + p, \mu m + tp}(r).$$

This is u times faster than analysing u subunits with respect to the (φ, z) coordinates.

1.3.4.5.2. Application to probability theory and direct methods

The Fourier transformation plays a central role in the branch of probability theory concerned with the limiting behaviour of sums of large numbers of independent and identically distributed random variables or random vectors. This privileged role is a consequence of the convolution theorem and of the 'moment-generating'

properties which follow from the exchange between differentiation and multiplication by monomials. When the limit theorems are applied to the calculation of joint probability distributions of structure factors, which are themselves closely related to the Fourier transformation, a remarkable phenomenon occurs, which leads to the saddlepoint approximation and to the maximum-entropy method.

1.3.4.5.2.1. Analytical methods of probability theory

The material in this section is not intended as an introduction to probability theory [for which the reader is referred to Cramér (1946), Petrov (1975) or Bhattacharya & Rao (1976)], but only as an illustration of the role played by the Fourier transformation in certain specific areas which are used in formulating and implementing direct methods of phase determination.

(a) Convolution of probability densities

The addition of independent random variables or vectors leads to the convolution of their probability distributions: if \mathbf{X}_1 and \mathbf{X}_2 are two n -dimensional random vectors independently distributed with probability densities P_1 and P_2 , respectively, then their sum $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$ has probability density \mathcal{P} given by

$$\begin{aligned} \mathcal{P}(\mathbf{X}) &= \int_{\mathbf{R}^n} P_1(\mathbf{X}_1) P_2(\mathbf{X} - \mathbf{X}_1) \, d^n \mathbf{X}_1 \\ &= \int_{\mathbf{R}^n} P_1(\mathbf{X} - \mathbf{X}_2) P_2(\mathbf{X}_2) \, d^n \mathbf{X}_2 \end{aligned}$$

i.e.

$$\mathcal{P} = P_1 * P_2.$$

This result can be extended to the case where P_1 and P_2 are singular measures (distributions of order zero, Section 1.3.2.3.4) and do not have a density with respect to the Lebesgue measure in \mathbb{R}^n .

(b) Characteristic functions

This convolution can be turned into a simple multiplication by considering the Fourier transforms (called the *characteristic functions*) of P_1 , P_2 and \mathcal{P} , defined with a slightly different normalization in that there is no factor of 2π in the exponent (see Section 1.3.2.4.5), *e.g.*

$$C(\mathbf{t}) = \int_{\mathbf{R}^n} P(\mathbf{X}) \exp(i\mathbf{t} \cdot \mathbf{X}) \, d^n \mathbf{X}.$$

Then by the convolution theorem

$$\mathcal{C}(\mathbf{t}) = C_1(\mathbf{t}) \times C_2(\mathbf{t}),$$

so that $\mathcal{P}(\mathbf{X})$ may be evaluated by Fourier inversion of its characteristic function as

$$\mathcal{P}(\mathbf{X}) = \frac{1}{(2\pi)^n} \int_{\mathbf{R}^n} C_1(\mathbf{t}) C_2(\mathbf{t}) \exp(-i\mathbf{t} \cdot \mathbf{X}) \, d^n \mathbf{t}$$

(see Section 1.3.2.4.5 for the normalization factors).

It follows from the differentiation theorem that the partial derivatives of the characteristic function $C(\mathbf{t})$ at $\mathbf{t} = \mathbf{0}$ are related to the moments of a distribution P by the identities

$$\begin{aligned} \mu_{r_1 r_2 \dots r_n} &\equiv \int_D P(\mathbf{X}) X_1^{r_1} X_2^{r_2} \dots X_n^{r_n} \, d^n \mathbf{X} \\ &= i^{-(r_1 + \dots + r_n)} \frac{\partial^{r_1 + \dots + r_n} C}{\partial t_1^{r_1} \dots \partial t_n^{r_n}} \Big|_{\mathbf{t}=\mathbf{0}} \end{aligned}$$

for any n -tuple of non-negative integers (r_1, r_2, \dots, r_n) .

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

(c) Moment-generating functions

The above relation can be freed from powers of i by defining (at least formally) the *moment-generating function*:

$$M(\mathbf{t}) = \int_{\mathbb{R}^n} P(\mathbf{X}) \exp(\mathbf{t} \cdot \mathbf{X}) d^n \mathbf{X}$$

which is related to $C(\mathbf{t})$ by $C(\mathbf{t}) = M(i\mathbf{t})$ so that the inversion formula reads

$$\mathcal{P}(\mathbf{X}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} M_1(i\mathbf{t}) M_2(i\mathbf{t}) \exp(-i\mathbf{t} \cdot \mathbf{X}) d^n \mathbf{t}.$$

The moment-generating function is well defined, in particular, for any probability distribution with compact support, in which case it may be continued analytically from a function over \mathbb{R}^n into an entire function of n complex variables by virtue of the Paley–Wiener theorem (Section 1.3.2.4.2.10). Its moment-generating properties are summed up in the following relations:

$$\mu_{r_1 r_2 \dots r_n} = \left. \frac{\partial^{r_1 + \dots + r_n} M}{\partial t_1^{r_1} \dots \partial t_n^{r_n}} \right|_{\mathbf{t}=\mathbf{0}}.$$

(d) Cumulant-generating functions

The multiplication of moment-generating functions may be further simplified into the addition of their logarithms:

$$\log \mathcal{M} = \log M_1 + \log M_2,$$

or equivalently of the coefficients of their Taylor series at $\mathbf{t} = \mathbf{0}$, viz:

$$\kappa_{r_1 r_2 \dots r_n} = \left. \frac{\partial^{r_1 + \dots + r_n} (\log M)}{\partial t_1^{r_1} \dots \partial t_n^{r_n}} \right|_{\mathbf{t}=\mathbf{0}}.$$

These coefficients are called *cumulants*, since they add when the independent random vectors to which they belong are added, and $\log M$ is called the *cumulant-generating function*. The inversion formula for \mathcal{P} then reads

$$\mathcal{P}(\mathbf{X}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \exp[\log M_1(i\mathbf{t}) + \log M_2(i\mathbf{t}) - i\mathbf{t} \cdot \mathbf{X}] d^n \mathbf{t}.$$

(e) Asymptotic expansions and limit theorems

Consider an n -dimensional random vector \mathbf{X} of the form

$$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \dots + \mathbf{X}_N,$$

where the N summands are independent n -dimensional random vectors identically distributed with probability density P . Then the distribution \mathcal{P} of \mathbf{X} may be written in closed form as a Fourier transform:

$$\begin{aligned} \mathcal{P}(\mathbf{X}) &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} M^N(i\mathbf{t}) \exp(-i\mathbf{t} \cdot \mathbf{X}) d^n \mathbf{t} \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \exp[N \log M(i\mathbf{t}) - i\mathbf{t} \cdot \mathbf{X}] d^n \mathbf{t}, \end{aligned}$$

where

$$M(\mathbf{t}) = \int_{\mathbb{R}^n} P(\mathbf{Y}) \exp(\mathbf{t} \cdot \mathbf{Y}) d^n \mathbf{Y}$$

is the moment-generating function common to all the summands.

This an exact expression for \mathcal{P} , which may be exploited analytically or numerically in certain favourable cases. Supposing for instance that P has compact support, then its characteristic function $M(i\mathbf{t})$ can be sampled finely enough to accommodate the bandwidth of the support of $\mathcal{P} = P^{*N}$ (this sampling rate clearly depends on n) so that the above expression for \mathcal{P} can be used for its

numerical evaluation as the *discrete* Fourier transform of $M^N(i\mathbf{t})$. This exact method is practical only for small values of the dimension n .

In all other cases some form of approximation must be used in the Fourier inversion of $M^N(i\mathbf{t})$. For this purpose it is customary (Cramér, 1946) to expand the cumulant-generating function around $\mathbf{t} = \mathbf{0}$ with respect to the carrying variables \mathbf{t} :

$$\log[M^N(i\mathbf{t})] = \sum_{\mathbf{r} \in \mathbb{N}^n} \frac{N \kappa_{\mathbf{r}}}{\mathbf{r}!} (i\mathbf{t})^{\mathbf{r}},$$

where $\mathbf{r} = (r_1, r_2, \dots, r_n)$ is a multi-index (Section 1.3.2.2.3). The first-order terms may be eliminated by recentring \mathcal{P} around its vector of first-order cumulants

$$\langle \mathbf{X} \rangle = \sum_{j=1}^N \langle \mathbf{X}_j \rangle,$$

where $\langle \cdot \rangle$ denotes the mathematical expectation of a random vector. The second-order terms may be grouped separately from the terms of third or higher order to give

$$\begin{aligned} M^N(i\mathbf{t}) &= \exp(-\frac{1}{2} N \mathbf{t}^T \mathbf{Q} \mathbf{t}) \\ &\times \exp \left\{ \sum_{|\mathbf{r}| \geq 3} \frac{N \kappa_{\mathbf{r}}}{\mathbf{r}!} (i\mathbf{t})^{\mathbf{r}} \right\}, \end{aligned}$$

where $\mathbf{Q} = \nabla \nabla^T (\log M)$ is the covariance matrix of the multivariate distribution P . Expanding the exponential gives rise to a series of terms of the form

$$\exp(-\frac{1}{2} N \mathbf{t}^T \mathbf{Q} \mathbf{t}) \times \text{monomial in } t_1, t_2, \dots, t_n,$$

each of which may now be subjected to a Fourier transformation to yield a Hermite function of \mathbf{t} (Section 1.3.2.4.4.2) with coefficients involving the cumulants κ of P . Taking the transformed terms in natural order gives an asymptotic expansion of P for large N called the *Gram–Charlier series* of \mathcal{P} , while grouping the terms according to increasing powers of $1/\sqrt{N}$ gives another asymptotic expansion called the *Edgeworth series* of \mathcal{P} . Both expansions comprise a leading Gaussian term which embodies the *central-limit theorem*:

$$\mathcal{P}(\mathbf{E}) = \frac{1}{\sqrt{\det(2\pi\mathbf{Q})}} \exp(-\frac{1}{2} \mathbf{E}^T \mathbf{Q}^{-1} \mathbf{E}), \quad \text{where } \mathbf{E} = \frac{\mathbf{X} - \langle \mathbf{X} \rangle}{\sqrt{N}}.$$

(f) The saddlepoint approximation

A limitation of the Edgeworth series is that it gives an accurate estimate of $\mathcal{P}(\mathbf{X})$ only in the vicinity of $\mathbf{X} = \langle \mathbf{X} \rangle$, i.e. for small values of \mathbf{E} . These convergence difficulties are easily understood: one is substituting a *local* approximation to $\log M$ (viz a Taylor-series expansion valid near $\mathbf{t} = \mathbf{0}$) into an integral, whereas integration is a *global* process which consults values of $\log M$ far from $\mathbf{t} = \mathbf{0}$.

It is possible, however, to let the point \mathbf{t} where $\log M$ is expanded as a Taylor series depend on the particular value \mathbf{X}^* of \mathbf{X} for which an accurate evaluation of $\mathcal{P}(\mathbf{X})$ is desired. This is the essence of the *saddlepoint method* (Fowler, 1936; Khinchin 1949; Daniels, 1954; de Bruijn, 1970; Bleistein & Handelsman, 1986), which uses an analytical continuation of $M(\mathbf{t})$ from a function over \mathbb{R}^n to a function over \mathbb{C}^n (see Section 1.3.2.4.2.10). Putting then $\mathbf{t} = \mathbf{s} - i\tau$, the \mathbb{C}^n version of Cauchy's theorem (Hörmander, 1973) gives rise to the identity

1. GENERAL RELATIONSHIPS AND TECHNIQUES

$$\mathcal{P}(\mathbf{X}^*) = \frac{\exp(-\boldsymbol{\tau} \cdot \mathbf{X}^*)}{(2\pi)^n} \times \int_{\mathbb{R}^n} \exp \left\{ N \left[\log M(\boldsymbol{\tau} + i\mathbf{s}) - i\mathbf{s} \cdot \frac{\mathbf{X}^*}{N} \right] \right\} d^n \mathbf{s}$$

for any $\boldsymbol{\tau} \in \mathbb{R}^n$. By a convexity argument involving the positive-definiteness of covariance matrix \mathbf{Q} , there is a unique value of $\boldsymbol{\tau}$ such that

$$\nabla(\log M)|_{\mathbf{t}=\mathbf{0}-i\boldsymbol{\tau}} = \frac{\mathbf{X}^*}{N}.$$

At the *saddlepoint* $\mathbf{t}^* = \mathbf{0} - i\boldsymbol{\tau}$, the modulus of the integrand above is a maximum and its phase is stationary with respect to the integration variable \mathbf{s} : as N tends to infinity, all contributions to the integral cancel because of rapid oscillation, except those coming from the immediate vicinity of \mathbf{t}^* where there is no oscillation. A Taylor expansion of $\log M^N$ to second order with respect to \mathbf{s} at \mathbf{t}^* then gives

$$\log M^N(\boldsymbol{\tau} + i\mathbf{s}) \approx \log M^N(\boldsymbol{\tau}) + i\mathbf{s} \cdot \mathbf{X}^* - \frac{N}{2} [\mathbf{s}^T \mathbf{Q} \mathbf{s}]$$

and hence

$$\mathcal{P}(\mathbf{X}^*) \approx \exp[\log M^N(\boldsymbol{\tau}) - \boldsymbol{\tau} \cdot \mathbf{X}^*] \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \exp(-\frac{1}{2} \mathbf{s}^T \mathbf{Q} \mathbf{s}) d^n \mathbf{s}.$$

The last integral is elementary and gives the ‘saddlepoint approximation’:

$$\mathcal{P}^{\text{SP}}(\mathbf{X}^*) = \frac{\exp(\mathbf{S})}{\sqrt{\det(2\pi \mathbf{Q})}},$$

where

$$\mathbf{S} = \log M^N(\boldsymbol{\tau}) - \boldsymbol{\tau} \cdot \mathbf{X}^*$$

and where

$$\mathbf{Q} = \nabla \nabla^T (\log M^N) = N \mathbf{Q}.$$

This approximation scheme amounts to using the ‘conjugate distribution’ (Khinchin, 1949)

$$P_{\boldsymbol{\tau}}(\mathbf{X}_j) = P(\mathbf{X}_j) \frac{\exp(\boldsymbol{\tau} \cdot \mathbf{X}_j)}{M(\boldsymbol{\tau})}$$

instead of the original distribution $P(\mathbf{X}_j) = P_0(\mathbf{X}_j)$ for the common distribution of all N random vectors \mathbf{X}_j . The exponential modulation results from the analytic continuation of the characteristic (or moment-generating) function into \mathbb{C}^n , as in Section 1.3.2.4.2.10. The saddlepoint approximation \mathcal{P}^{SP} is only the leading term of an asymptotic expansion (called the *saddlepoint expansion*) for \mathcal{P} , which is actually the Edgeworth expansion associated with $P_{\boldsymbol{\tau}}^*$.

1.3.4.5.2.2. The statistical theory of phase determination

The methods of probability theory just surveyed were applied to various problems formally similar to the crystallographic phase problem [e.g. the ‘problem of the random walk’ of Pearson (1905)] by Rayleigh (1880, 1899, 1905, 1918, 1919) and Kluyver (1906). They became the basis of the statistical theory of communication with the classic papers of Rice (1944, 1945).

The Gram–Charlier and Edgeworth series were introduced into crystallography by Bertaut (1955*a,b,c*, 1956*a*) and by Klug (1958), respectively, who showed them to constitute the mathematical basis of numerous formulae derived by Hauptman & Karle (1953). The saddlepoint approximation was introduced by Bricogne (1984) and was shown to be related to variational methods involving the

maximization of certain entropy criteria. This connection exhibits most of the properties of the Fourier transform at play simultaneously, and will now be described as a final illustration.

(a) Definitions and conventions

Let H be a set of unique non-origin reflections \mathbf{h} for a crystal with lattice Λ and space group G . Let H contain n_a acentric and n_c centric reflections. Structure-factor values attached to all reflections in H will comprise $n = 2n_a + n_c$ real numbers. For \mathbf{h} acentric, $\alpha_{\mathbf{h}}$ and $\beta_{\mathbf{h}}$ will be the real and imaginary parts of the complex structure factor; for \mathbf{h} centric, $\gamma_{\mathbf{h}}$ will be the real coordinate of the (possibly complex) structure factor measured along a real axis rotated by one of the two angles $\theta_{\mathbf{h}}$, π apart, to which the phase is restricted modulo 2π (Section 1.3.4.2.2.5). These n real coordinates will be arranged as a column vector containing the acentric then the centric data, i.e. in the order

$$\alpha_1, \beta_1, \alpha_2, \beta_2, \dots, \alpha_{n_a}, \beta_{n_a}, \gamma_1, \gamma_2, \dots, \gamma_{n_c}.$$

(b) Vectors of trigonometric structure-factor expressions

Let $\boldsymbol{\xi}(\mathbf{x})$ denote the vector of trigonometric structure-factor expressions associated with $\mathbf{x} \in D$, where D denotes the asymmetric unit. These are defined as follows:

$$\begin{aligned} \alpha_{\mathbf{h}}(\mathbf{x}) + i\beta_{\mathbf{h}}(\mathbf{x}) &= \Xi(\mathbf{h}, \mathbf{x}) \quad \text{for } \mathbf{h} \text{ acentric} \\ \gamma_{\mathbf{h}}(\mathbf{x}) &= \exp(-i\theta_{\mathbf{h}}) \Xi(\mathbf{h}, \mathbf{x}) \quad \text{for } \mathbf{h} \text{ centric,} \end{aligned}$$

where

$$\Xi(\mathbf{h}, \mathbf{x}) = \frac{1}{|G_{\mathbf{x}}|} \sum_{g \in G} \exp\{2\pi i \mathbf{h} \cdot [S_g(\mathbf{x})]\}.$$

According to the convention above, the coordinates of $\boldsymbol{\xi}(\mathbf{x})$ in \mathbb{R}^n will be arranged in a column vector as follows:

$$\begin{aligned} \xi_{2r-1}(\mathbf{x}) &= \alpha_{\mathbf{h}_r}(\mathbf{x}) \quad \text{for } r = 1, \dots, n_a, \\ \xi_{2r}(\mathbf{x}) &= \beta_{\mathbf{h}_r}(\mathbf{x}) \quad \text{for } r = 1, \dots, n_a, \\ \xi_{n_a+r}(\mathbf{x}) &= \gamma_{\mathbf{h}_r}(\mathbf{x}) \quad \text{for } r = n_a + 1, \dots, n_a + n_c. \end{aligned}$$

(c) Distributions of random atoms and moment-generating functions

Let position \mathbf{x} in D now become a random vector with probability density $m(\mathbf{x})$. Then $\boldsymbol{\xi}(\mathbf{x})$ becomes itself a random vector in \mathbb{R}^n , whose distribution $p(\boldsymbol{\xi})$ is the image of distribution $m(\mathbf{x})$ through the mapping $\mathbf{x} \rightarrow \boldsymbol{\xi}(\mathbf{x})$ just defined. The locus of $\boldsymbol{\xi}(\mathbf{x})$ in \mathbb{R}^n is a compact algebraic manifold \mathcal{V} (the multidimensional analogue of a Lissajous curve), so that p is a singular measure (a distribution of order 0, Section 1.3.2.3.4, concentrated on that manifold) with compact support. The average with respect to p of any function Ω over \mathbb{R}^n which is infinitely differentiable in a neighbourhood of \mathcal{V} may be calculated as an average with respect to m over D by the ‘induction formula’:

$$\langle p, \Omega \rangle = \int_D m(\mathbf{x}) \Omega[\boldsymbol{\xi}(\mathbf{x})] d^3 \mathbf{x}.$$

In particular, one can calculate the moment-generating function M for distribution p as

$$M(\mathbf{t}) \equiv \langle p_{\boldsymbol{\xi}}, \exp(\mathbf{t} \cdot \boldsymbol{\xi}) \rangle = \int_D m(\mathbf{x}) \exp[\mathbf{t} \cdot \boldsymbol{\xi}(\mathbf{x})] d^3 \mathbf{x}$$

and hence calculate the moments μ (respectively cumulants κ) of p by differentiation of M (respectively $\log M$) at $\mathbf{t} = \mathbf{0}$:

1.3. FOURIER TRANSFORMS IN CRYSTALLOGRAPHY

$$\begin{aligned}\mu_{r_1 r_2 \dots r_n} &\equiv \int_D m(\mathbf{x}) \xi_1^{r_1}(\mathbf{x}) \xi_2^{r_2}(\mathbf{x}) \dots \xi_n^{r_n}(\mathbf{x}) d^3 \mathbf{x} \\ &= \frac{\partial^{r_1 + \dots + r_n} (M)}{\partial t_1^{r_1} \dots \partial t_n^{r_n}} \\ \kappa_{r_1 r_2 \dots r_n} &= \frac{\partial^{r_1 + \dots + r_n} (\log M)}{\partial t_1^{r_1} \dots \partial t_n^{r_n}}.\end{aligned}$$

The structure-factor algebra for group G (Section 1.3.4.2.2.9) then allows one to express products of ξ 's as linear combinations of other ξ 's, and hence to express all moments and cumulants of distribution $p(\xi)$ as linear combinations of real and imaginary parts of Fourier coefficients of the prior distribution of atoms $m(\mathbf{x})$. This plays a key role in the use of non-uniform distributions of atoms.

(d) The joint probability distribution of structure factors

In the random-atom model of an equal-atom structure, N atoms are placed randomly, independently of each other, in the asymmetric unit D of the crystal with probability density $m(\mathbf{x})$. For point atoms of unit weight, the vector \mathbf{F} of structure-factor values for reflections $\mathbf{h} \in H$ may be written

$$\mathbf{F} = \sum_{l=1}^N \xi^{[l]},$$

where the N copies $\xi^{[l]}$ of random vector ξ are independent and have the same distribution $p(\xi)$.

The joint probability distribution $\mathcal{P}(\mathbf{F})$ is then [Section 1.3.4.5.2.1(e)]

$$\mathcal{P}(\mathbf{X}) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \exp[N \log M(it) - it \cdot \mathbf{X}] d^n t.$$

For low dimensionality n it is possible to carry out the Fourier transformation numerically after discretization, provided $M(it)$ is sampled sufficiently finely that no aliasing results from taking its N th power (Barakat, 1974). This exact approach can also accommodate heterogeneity, and has been used first in the field of intensity statistics (Shmueli *et al.*, 1984, 1985; Shmueli & Weiss, 1987, 1988), then in the study of the Σ_1 and Σ_2 relations in triclinic space groups (Shmueli & Weiss, 1985, 1986). Some of these applications are described in Chapter 2.1 of this volume. This method could be extended to the construction of any joint probability distribution (j.p.d.) in any space group by using the generic expression for the moment-generating function (m.g.f.) derived by Bricogne (1984). It is, however, limited to small values of n by the necessity to carry out n -dimensional FFTs on large arrays of sample values.

The asymptotic expansions of Gram–Charlier and Edgeworth have good convergence properties only if $F_{\mathbf{h}}$ lies in the vicinity of $\langle F_{\mathbf{h}} \rangle = N \mathcal{F}[m](\mathbf{h})$ for all $\mathbf{h} \in H$. Previous work on the j.p.d. of structure factors has used for $m(\mathbf{x})$ a uniform distribution, so that $\langle \mathbf{F} \rangle = \mathbf{0}$; as a result, the corresponding expansions are accurate only if all moduli $|F_{\mathbf{h}}|$ are small, in which case the j.p.d. contains little phase information.

The saddlepoint method [Section 1.3.4.5.2.1(f)] constitutes the method of choice for evaluating the joint probability $\mathcal{P}(\mathbf{F}^*)$ of structure factors when some of the moduli in \mathbf{F}^* are large. As shown previously, this approximation amounts to using the ‘conjugate distribution’

$$p_{\tau}(\xi) = p(\xi) \frac{\exp(\tau \cdot \xi)}{M(\tau)}$$

instead of the original distribution $p(\xi) = p_0(\xi)$ for the distribution of random vector ξ . This conjugate distribution p_{τ} is induced from

the modified distribution of atoms

$$q_{\tau}(\mathbf{x}) = m(\mathbf{x}) \frac{\exp[\tau \cdot \xi(\mathbf{x})]}{M(\tau)}, \quad (\text{SP1})$$

where, by the induction formula, $M(\tau)$ may be written as

$$M(\tau) = \int_D m(\mathbf{x}) \exp[\tau \cdot \xi(\mathbf{x})] d^3 \mathbf{x} \quad (\text{SP2})$$

and where τ is the unique solution of the saddlepoint equation:

$$\nabla_{\tau}(\log M^N) = \mathbf{F}^*. \quad (\text{SP3})$$

The desired approximation is then

$$\mathcal{P}^{\text{SP}}(\mathbf{F}^*) = \frac{\exp(\mathbf{S})}{\sqrt{\det(2\pi \mathcal{Q})}},$$

where

$$\mathbf{S} = \log M^N(\tau) - \tau \cdot \mathbf{F}^*$$

and where

$$\mathcal{Q} = \nabla \nabla^T (\log M^N) = \mathbf{N} \mathbf{Q}.$$

Finally, the elements of the Hessian matrix $\mathbf{Q} = \nabla \nabla^T (\log M)$ are just the trigonometric second-order cumulants of distribution p , and hence can be calculated *via* structure-factor algebra from the Fourier coefficients of $q_{\tau}(\mathbf{x})$. All the quantities involved in the expression for $\mathcal{P}^{\text{SP}}(\mathbf{F}^*)$ are therefore effectively computable from the initial data $m(\mathbf{x})$ and \mathbf{F}^* .

(e) Maximum-entropy distributions of atoms

One of the main results in Bricogne (1984) is that the modified distribution $q_{\tau}(\mathbf{x})$ in (SP1) is the unique distribution which has maximum entropy $\mathcal{J}_m(q)$ relative to $m(\mathbf{x})$, where

$$\mathcal{J}_m(q) = - \int_D q(\mathbf{x}) \log \left[\frac{q(\mathbf{x})}{m(\mathbf{x})} \right] d^3 \mathbf{x},$$

under the constraint that \mathbf{F}^* be the centroid vector of the corresponding conjugate distribution $\mathcal{P}_{\tau}(\mathbf{F})$. The traditional notation of maximum-entropy (ME) theory (Jaynes, 1957, 1968, 1983) is in this case (Bricogne, 1984)

$$q^{\text{ME}}(\mathbf{x}) = m(\mathbf{x}) \frac{\exp[\lambda \cdot \xi(\mathbf{x})]}{Z(\lambda)} \quad (\text{ME1})$$

$$Z(\lambda) = \int_D m(\mathbf{x}) \exp[\lambda \cdot \xi(\mathbf{x})] d^3 \mathbf{x} \quad (\text{ME2})$$

$$\nabla_{\lambda}(\log Z^N) = \mathbf{F}^* \quad (\text{ME3})$$

so that Z is identical to the m.g.f. M , and the coordinates τ of the saddlepoint are the Lagrange multipliers λ for the constraints \mathbf{F}^* .

Jaynes's ME theory also gives an estimate for $\mathcal{P}(\mathbf{F}^*)$:

$$\mathcal{P}^{\text{ME}}(\mathbf{F}^*) \approx \exp(\mathcal{J}),$$

where

$$\mathcal{J} = \log Z^N - \lambda \cdot \mathbf{F}^* = N \mathcal{J}_m(q^{\text{ME}})$$

is the total entropy and is the counterpart to \mathbf{S} under the equivalence just established.

\mathcal{P}^{ME} is identical to \mathcal{P}^{SP} , but lacks the denominator. The latter, which is the normalization factor of a multivariate Gaussian with covariance matrix \mathcal{Q} , may easily be seen to arise through Szegő's theorem (Sections 1.3.2.6.9.4, 1.3.4.2.1.10) from the extra logarithmic term in Stirling's formula

$$\log(q!) \approx q \log q - q + \frac{1}{2} \log(2\pi q)$$

1. GENERAL RELATIONSHIPS AND TECHNIQUES

(see, for instance, Reif, 1965) beyond the first two terms which serve to define entropy, since

$$\frac{1}{n} \log \det (2\pi \mathbf{Q}) \approx \int_{\mathbb{R}^3/\mathbb{Z}^3} \log 2\pi q^{\text{ME}}(\mathbf{x}) d^3\mathbf{x}.$$

The relative effect of this extra normalization factor depends on the ratio

$$\frac{n}{N} = \frac{\text{dimension of } \mathbf{F} \text{ over } \mathbb{R}}{\text{number of atoms}}.$$

The above relation between entropy maximization and the saddlepoint approximation is the basis of a Bayesian statistical approach to the phase problem (Bricogne, 1988) where the assumptions under which joint distributions of structure factors are sought incorporate many new ingredients (such as molecular boundaries, isomorphous substitutions, known fragments, noncrystallographic symmetries, multiple crystal forms) besides trial phase choices for basis reflections. The ME criterion intervenes in the construction of $q^{\text{ME}}(\mathbf{x})$ under these assumptions, and the distribution $q^{\text{ME}}(\mathbf{x})$ is a very useful computational intermediate in obtaining the approximate joint probability $\mathcal{P}^{\text{SP}}(\mathbf{F}^*)$ and the associated conditional distributions and likelihood functions.

(f) Role of the Fourier transformation

The formal developments presented above make use of the following properties of the Fourier transformation:

- (i) the convolution theorem, which turns the convolution of probability distributions into the multiplication of their characteristic functions;
- (ii) the differentiation property, which confers moment-generating properties to characteristic functions;
- (iii) the reciprocity theorem, which allows the retrieval of a probability distribution from its characteristic or moment-generating function;
- (iv) the Paley–Wiener theorem, which allows the analytic continuation of characteristic functions associated to probability

distributions with compact support, and thus gives rise to conjugate families of distributions;

(v) Bertaut’s structure-factor algebra (a discrete symmetrized version of the convolution theorem), which allows the calculation of all necessary moments and cumulants when the dimension n is small;

(vi) Szegő’s theorem, which provides an asymptotic approximation of the normalization factor when n is large.

This multi-faceted application seems an appropriate point at which to end this description of the Fourier transformation and of its use in crystallography.

Acknowledgements

Many aspects of the theory of discrete Fourier transform algorithms and of its extension to incorporate crystallographic symmetry have been the focus of a long-standing collaborative effort between Professor Louis Auslander, Professor Richard Tolimieri, their co-workers and the writer. I am most grateful to them for many years of mathematical stimulation and enjoyment, for introducing me to the ‘big picture’ of the discrete Fourier transform which they have elaborated over the past decade, and for letting me describe here some of their unpublished work. In particular, the crystallographic extensions of the Rader/Winograd algorithms presented in Section 1.3.4.3.4.3 were obtained by Richard Tolimieri, in a collaboration partially supported by NIH grant GM 32362 (to the writer).

I am indebted to the Editor for many useful and constructive suggestions of possible improvements to the text, only a few of which I have been able to implement. I hope to incorporate many more of them in the future. I also wish to thank Dr D. Sayre for many useful comments on an early draft of the manuscript.

This contribution was written during the tenure of a Visiting Fellowship at Trinity College, Cambridge, with partial financial support from Trinity College and the MRC Laboratory of Molecular Biology. I am most grateful to both institutions for providing ideal working conditions.