16. DIRECT METHODS

prejudice *m*. Indeed, it is a straightforward consequence of Shannon's second theorem (Section 16.2.2) as expressed in equation (16.2.2.3) that

$$\mathcal{P}^{\mathrm{ME}}(\mathbf{c}) \propto \exp(\mathcal{S}), \qquad (16.2.2.15)$$

where

$$\mathcal{S} = \log Z^N - \lambda \cdot \mathbf{c} = N\mathcal{S}_m(q^{\mathrm{ME}}) \qquad (16.2.2.16)$$

is the total entropy for *N* symbols.

16.2.3. Adaptation to crystallography

16.2.3.1. *The random-atom model*

The standard setting of probabilistic direct methods (Hauptman & Karle, 1953; Bertaut, 1955*a*,*b*; Klug, 1958) uses implicitly as its starting point a source of random atomic positions. This can be described in the terms introduced in Section 16.2.2.1 by using a continuous alphabet $\mathcal{A}$ whose symbols **s** are fractional coordinates **x** in the asymmetric unit of the crystal, the uniform measure μ being the ordinary Lebesgue measure $\mathrm{d}^3\mathbf{x}$. A message of length *N* generated by that source is then a random *N*-equal-atom structure.

16.2.3.2. *Conventional direct methods and their limitations*

The traditional theory of direct methods assumes a *uniform* distribution $q(\mathbf{x})$ of random atoms and proceeds to derive joint distributions of structure factors belonging to an *N*-atom random structure, using the asymptotic expansions of Gram–Charlier and Edgeworth. These methods have been described in Section 1.3.4.5.2.2 of *IT* B (Bricogne, 2001) as examples of applications of Fourier transforms. The reader is invited to consult this section for terminology and notation. These joint distributions of complex structure factors are subsequently used to derive conditional distributions of phases when the amplitudes are assigned their observed values, or of a subset of complex structure factors when the others are assigned certain values. In both cases, the *largest* structure-factor amplitudes are used as the conditioning information.

It was pointed out by the author (Bricogne, 1984) that this procedure can be problematic, as the Gram–Charlier and Edgeworth expansions have good convergence properties only in the vicinity of the expectation values of each structure factor: as the atoms are assumed to be uniformly distributed, these series afford an adequate approximation for the joint distribution $\mathcal{P}(\mathbf{F})$ only near the origin of structure-factor space, *i.e.* for *small* values of all the structure amplitudes. It is therefore incorrect to use these local approximations to $\mathcal{P}(\mathbf{F})$ near $\mathbf{F} = \mathbf{0}$ as if they were the global functional form for that function 'in the large' when forming conditional probability distributions involving large amplitudes.

16.2.3.3. *The notion of recentring and the maximum-entropy criterion*

These limitations can be overcome by recognizing that, if the locus $\mathcal{T}$ (a high-dimensional torus) defined by the large structure-factor amplitudes to be used in the conditioning data is too extended in structure-factor space for a single asymptotic expansion of $\mathcal{P}(\mathbf{F})$

to be accurate everywhere on it, then $\mathcal{T}$ should be broken up into sub-regions, and different local approximations to $\mathcal{P}(\mathbf{F})$ should be constructed in each of them. Each of these sub-regions will consist of a 'patch' of $\mathcal{T}$ surrounding a point $\mathbf{F}^* \neq \mathbf{0}$ located on $\mathcal{T}$. Such a point $\mathbf{F}^*$ is obtained by assigning 'trial' phase values to the known moduli, but these trial values do not necessarily have to be viewed as 'serious' assumptions concerning the true values of the phases: rather, they should be thought of as pointing to a patch of $\mathcal{T}$ and to a specialized asymptotic expansion of $\mathcal{P}(\mathbf{F})$ designed to be the most accurate approximation possible to $\mathcal{P}(\mathbf{F})$ on that patch. With a sufficiently rich collection of such constructs, $\mathcal{P}(\mathbf{F})$ can be accurately calculated anywhere on $\mathcal{T}$.

These considerations lead to the notion of *recentring*. Recentring the usual Gram–Charlier or Edgeworth asymptotic expansion for $\mathcal{P}(\mathbf{F})$ away from $\mathbf{F} = \mathbf{0}$, by making trial phase assignments that define a point $\mathbf{F}^*$ on $\mathcal{T}$, is equivalent to using a non-uniform prior distribution of atoms $q(\mathbf{x})$, reproducing the individual components of $\mathbf{F}^*$ among its Fourier coefficients. The latter constraint leaves $q(\mathbf{x})$ highly indeterminate, but Jaynes' argument given in Section 16.2.2.3 shows that there is a uniquely defined 'best' choice for it: it is that distribution $q^{\mathrm{ME}}(\mathbf{x})$ having *maximum entropy* relative to a uniform prior prejudice $m(\mathbf{x})$, and having the corresponding values $\mathbf{U}^*$ of the unitary structure factors for its Fourier coefficients. This distribution has the unique property that it rules out as few random structures as possible on the basis of the limited information available in $\mathbf{F}^*$.

In terms of the statistical mechanical language used in Section 16.2.1, the trial structure-factor values $\mathbf{F}^*$ used as constraints would be the macroscopic quantities that can be controlled externally; while the 3*N* atomic coordinates would be the internal degrees of freedom of the system, whose entropy should be a maximum under these macroscopic constraints.

16.2.3.4. *The crystallographic maximum-entropy formalism*

It is possible to solve explicitly the maximum-entropy equations (ME1) to (ME3) derived in Section 16.2.2.4 for the crystallographic case that has motivated this study, *i.e.* for the purpose of constructing $q^{\mathrm{ME}}(\mathbf{x})$ from the knowledge of a set of trial structure-factor values $\mathbf{F}^*$. These derivations are given in §3.4 and §3.5 of Bricogne (1984). Extensive relations with the algebraic formalism of traditional direct methods are exhibited in §4, and connections with the theory of determinantal inequalities and with the maximum-determinant rule of Tsoucaris (1970) are studied in §6, of the same paper. The reader interested in these topics is invited to consult this paper, as space limitations preclude their discussion in the present chapter.

16.2.3.5. *Connection with the saddlepoint method*

The saddlepoint method constitutes an alternative approach to the problem of evaluating the joint probability $\mathcal{P}(\mathbf{F}^*)$ of structure factors when some of the moduli in $\mathbf{F}^*$ are large. It is shown in §5 of Bricogne (1984), and in more detail in Section 1.3.4.5.2.2 of Chapter 1.3 of *IT* B (Bricogne, 2001), that there is complete equivalence between the maximum-entropy approach to the phase problem and the classical probabilistic approach by the method of joint distributions, provided the latter is enhanced by the adoption of the saddlepoint approximation.

348