2.2.10.4.1. *One-wavelength techniques*

Probability distributions of diffraction intensities and of selected functions of diffraction intensities for dispersive structures have been given by various authors [Parthasarathy & Srinivasan (1964), see also Srinivasan & Parthasarathy (1976) and relevant literature cited therein]. We describe here some probabilistic formulae for estimating invariants of low order.

(*a*) *Estimation of two-phase structure invariants.* The conditional probability distribution of $\Phi = \varphi_{\mathbf{h}} + \varphi_{-\mathbf{h}}$ given $R_{\mathbf{h}}$ and $G_{\mathbf{h}}$ (normalized moduli of $F_{\mathbf{h}}$ and $F_{-\mathbf{h}}$, respectively) (Hauptman, 1982*b*; Giacovazzo, 1983*b*) is

$$P(\Phi|R_{\mathbf{h}}, G_{\mathbf{h}}) \simeq [2\pi I_0(Q)]^{-1} \exp[Q\cos(\Phi - q)], \qquad (2.2.10.3)$$

where

$$Q = \frac{2R_{\mathbf{h}}G_{\mathbf{h}}}{\sqrt{c}}[c_1^2 + c_2^2]^{1/2},$$

$$\cos q = \frac{c_1}{[c_1^2 + c_2^2]^{1/2}}, \qquad \sin q = \frac{c_2}{[c_1^2 + c_2^2]^{1/2}},$$

$$c_1 = \sum_{j=1}^{N}(f_j'^2 - f_j''^2)/\textstyle\sum,$$

$$c_2 = 2\sum_{j=1}^{N} f_j' f_j''/\textstyle\sum,$$

$$c = [1 - (c_1^2 + c_2^2)]^2,$$

$$\textstyle\sum = \displaystyle\sum_{j=1}^{N}(f_j'^2 + f_j''^2).$$

q is the most probable value of Φ: a large value of the parameter Q suggests that the phase relation $\Phi = q$ is reliable. Large values of Q are often available in practice: q, however, may be considered an estimate of $|\Phi|$ rather than of Φ because the enantiomorph is not fixed in (2.2.10.3). A formula for the estimation of Φ in centrosymmetric structures has recently been provided by Giacovazzo (1987).

If the positions of the p anomalous scatterers are known *a priori* [let $F_{p\mathbf{h}} = |F_{p\mathbf{h}}|\exp(i\varphi_{p\mathbf{h}})$ be the structure factor of the partial structure], then an estimate of $\Phi' = \varphi_{\mathbf{h}} - \varphi_{p\mathbf{h}}$ is given (Cascarano & Giacovazzo, 1985) by

$$P(\Phi'|R_{\mathbf{h}}, R_{p\mathbf{h}}) \simeq [2\pi I_0(Q')]^{-1} \exp[Q'\cos\Phi'], \qquad (2.2.10.4)$$

where

$$Q' = 2R^+R_p^+ \bigg/ \left(1 - \sum_p \Big/ \sum\right), \qquad \sum_p = \sum_{j=1}^{p}(f_j'^2 + f_j''^2).$$

(2.2.10.4) may be considered the generalization of Sim's distribution (2.2.5.17) to dispersive structures.

(*b*) *Estimation of triplet invariants.* Kroon *et al.* (1977) first incorporated anomalous diffraction in order to estimate triplet invariants. Their work was based on an analysis of the complex double Patterson function. Subsequent probabilistic considerations (Heinermann *et al.*, 1978) confirmed their results, which can be so expressed:

$$\sin\bar{\Phi} = \frac{|\tau|^2 - |\bar{\tau}|^2}{4\tau''[\frac{1}{2}(|\tau|^2 + |\bar{\tau}|^2) - |\tau''|^2]^{1/2}}, \qquad (2.2.10.5)$$

where $(\mathbf{h} + \mathbf{k} + \mathbf{l} = 0)$

$$\tau = E_{\mathbf{h}}E_{\mathbf{k}}E_{\mathbf{l}} = R_{\mathbf{h}}R_{\mathbf{k}}R_{\mathbf{l}}\exp(i\Phi_{\mathbf{h},\mathbf{k}}),$$

$$\bar{\tau} = E_{-\mathbf{h}}E_{-\mathbf{k}}E_{-\mathbf{l}} = G_{\mathbf{h}}G_{\mathbf{k}}G_{\mathbf{l}}\exp(i\Phi_{\bar{\mathbf{h}},\bar{\mathbf{k}}}),$$

$$\bar{\Phi} = \tfrac{1}{2}(\Phi_{\mathbf{h},\mathbf{k}} - \Phi_{\bar{\mathbf{h}},\bar{\mathbf{k}}}),$$

and τ'' is the contribution of the imaginary part of τ, which may be approximated in favourable conditions by

$$\tau'' = 2f''[f_{\mathbf{h}}'f_{\mathbf{k}}' + f_{\mathbf{h}}'f_{\mathbf{l}}' + f_{\mathbf{k}}f_{\mathbf{l}}] \times [1 + S(R_{\mathbf{h}}^2 + R_{\mathbf{k}}^2 + R_{\mathbf{l}}^2 - 3)],$$

where S is a suitable scale factor.

Equation (2.2.10.5) gives two possible values for $\bar{\Phi}$ (Φ and $\pi - \Phi$). Only if $R_{\mathbf{h}}R_{\mathbf{k}}R_{\overline{\mathbf{h}+\mathbf{k}}}$ is large enough may this phase ambiguity be resolved by choosing the angle nearest to zero.

The evaluation of triplet phases by means of anomalous dispersion has been further pursued by Hauptman (1982*b*) and Giacovazzo (1983*b*). Owing to the breakdown of Friedel's law there are eight distinct triplet invariants which can contemporaneously be exploited:

$$\Phi_1 = \varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} + \varphi_{\mathbf{l}}, \qquad \Phi_2 = -\varphi_{-\mathbf{h}} + \varphi_{\mathbf{k}} + \varphi_{\mathbf{l}}$$

$$\Phi_3 = \varphi_{\mathbf{h}} - \varphi_{-\mathbf{k}} + \varphi_{\mathbf{l}}, \qquad \Phi_4 = \varphi_{\mathbf{h}} + \varphi_{\mathbf{k}} - \varphi_{-\mathbf{l}}$$

$$\Phi_5 = \varphi_{-\mathbf{h}} + \varphi_{-\mathbf{k}} + \varphi_{-\mathbf{l}}, \qquad \Phi_6 = -\varphi_{\mathbf{h}} + \varphi_{-\mathbf{k}} + \varphi_{-\mathbf{l}}$$

$$\Phi_7 = \varphi_{-\mathbf{h}} - \varphi_{\mathbf{k}} + \varphi_{-\mathbf{l}}, \qquad \Phi_8 = \varphi_{-\mathbf{h}} + \varphi_{-\mathbf{k}} - \varphi_{\mathbf{l}}.$$

The conditional probability distribution for each of the eight triplet invariants, given $R_{\mathbf{h}}$, $R_{\mathbf{k}}$, $R_{\mathbf{l}}$, $G_{\mathbf{h}}$, $G_{\mathbf{k}}$, $G_{\mathbf{l}}$, is

$$P_j(\Phi_j) \simeq \frac{1}{L_j}\exp[A_j\cos(\Phi_j - \omega_j)].$$

The definitions of A_j, L_j and ω_j are rather extensive and so the reader is referred to the published papers. A_j and L_j are positive values, so ω_j is the expected value of Φ_j. It may lie anywhere between 0 and 2π.

An algebraic analysis of triplet phase invariants coupled with probabilistic considerations has been carried out by Karle (1984, 1985). The rules permit the qualitative selection of triple phase invariants that have values close to $\pi/2$, $-\pi/2$, 0, and other values in the range from $-\pi$ to π.

Let us now describe some practical aspects of the integration of direct methods with OAS techniques.

Anomalous difference structure factors

$$\Delta_{\rm iso} = |F^+| - |F^-|$$

can be used for locating the positions of the anomalous scatterers (Mukherjee *et al.*, 1989). Tests prove that accuracy in the difference magnitudes is critical for the success of the phasing process.

Suppose now that the positions of the heavy atoms have been found. How do we estimate the phase values for the protein? The phase ambiguity strictly connected with OAS techniques can be overcome by different methods: we quote the Qs method by Hao & Woolfson (1989), the Wilson distribution method and the MPS method by Ralph & Woolfson (1991), and the Bijvoet–Ramachandran–Raman method by Peerdeman & Bijvoet (1956), Raman (1959) and Moncrief & Lipscomb (1966). More recently, a probabilistic method by Fan & Gu (1985) gained additional insight into the problem.

2.2.10.4.2. *The SIRAS, MIRAS and MAD cases*

Isomorphous replacement and anomalous scattering are discussed in Chapter 2.4 and in *IT* F (2001). We observe here only that the SIRAS case can lead algebraically to unambiguous phase determination provided the experimental data are sufficiently good.

references