12. ISOMORPHOUS REPLACEMENT
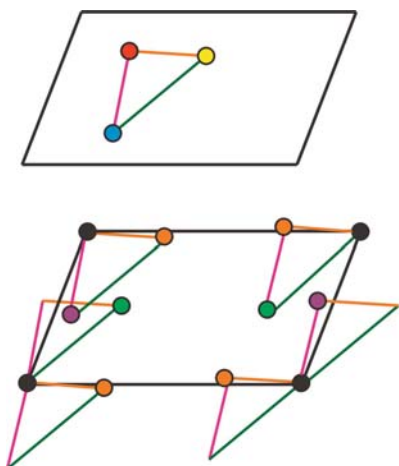

Fig. 12.2.2.2. The vector superposition method. The Patterson map of Fig. 12.1.1.1 can be regarded as the superposition of the structure (and its inverse), with each of its atoms placed alternately at the origin. By shifting each peak of the Patterson function to the origin and calculating the correlation of all remaining peaks with the unshifted map, it is possible to deconvolute the Patterson function.

i.e. there are three Harker sections: $u = \frac{1}{2}$, $v = \frac{1}{2}$ and $w = \frac{1}{2}$. Peaks occurring on the Harker sections must reduce to a self-consistent set of coordinates (x, y, z), allowing reconstruction of the atomic positions.

If we have two isomorphous (see below) data sets F_{PH} and F_P, then the difference in the two Patterson functions,

$$P_{PH} - P_P = \int [F_{PH}^2(\mathbf{S}) - F_P^2(\mathbf{S})] \exp\{-2\pi i \mathbf{r} \cdot \mathbf{S}\} \, d^3\mathbf{S},$$

will deliver information about the heavy-atom structure. Such a difference function gives rise to non-negligible peaks arising from interference between the F_H and F_P terms, however (Perutz, 1956). Rossmann (1960) showed that these interference terms could be reduced through calculation of the modified Patterson function

$$P_H = \int [F_{PH}(\mathbf{S}) - F_P(\mathbf{S})]^2 \exp\{-2\pi i \mathbf{r} \cdot \mathbf{S}\} \, d^3\mathbf{S}.$$

In the case of a single-site derivative, peaks should occur only at the Harker vectors corresponding to the heavy-atom position. Even so, there is a choice of positions for the heavy atom: *e.g.*, in the $P2_12_12_1$ case, coordinates $(\pm x + \xi, \pm y + \nu, \pm z + \zeta)$, where ξ, ν and ζ can each take the value 0 or 1/2, will all give rise to the same Harker vectors. This in itself is not a problem, relating to equivalent choices of origin and of handedness, but has important ramifications for multisite derivatives or multiple isomorphous replacement (see below).

If there is more than one site, then there will be two sets of peaks: one set corresponding to the Harker sections (self-vector set) and one set corresponding to the difference vectors between different heavy-atom sites (the cross-vector set). In this case, the choice of one heavy-atom position (x_{H1}, y_{H1}, z_{H1}) determines the origin and the handedness to which all other peaks *must* correspond. Thus, in the $P2_12_12_1$ example, only one cross vector will occur for

$$(x_{h1} \pm x_{h2} + \xi, y_{h1} \pm y_{h2} + \nu, z_{h1} \pm z_{h2} + \zeta).$$

An alternative to the Harker-vector approach is Patterson-vector superposition (Sheldrick *et al.*, 1993; Richardson & Jacobson, 1987). The Patterson map contains several images of the structure that have been shifted by interatomic vectors (Fig. 12.2.2.2). If this structure is relatively simple (as is to be hoped for in a 'normal' heavy-atom derivative), then it should be possible to deconvolute the superimposed structures by vector shifts (Buerger, 1959).

12.2.3. The difference Fourier

Once the heavy-atom positions have been found, they can be used to calculate approximate phases and Fourier maps. Ideally, difference Fourier maps calculated with phases from a single site should reveal the other positions determined from the Harker search procedure. This ensures that all heavy-atom positions correspond to a single origin and hand. Similarly, phases calculated from derivative *H*1 should reveal the heavy-atom structure for derivative *H*2. Merging and refinement of all phase information will result in a phase set that can be used to solve the structure.

12.2.4. Reality

12.2.4.1. *Treatment of errors*

Until now, we have dealt with cases involving perfect data. Although this ideal may now be attainable using MAD techniques, this is not necessarily the usual laboratory situation. In the first place, it is necessary to scale the derivative data F_{PH} to the native F_P. One of the most common scaling procedures is based on the expected statistical dependence of intensity on resolution (Wilson, 1949). This may not be particularly accurate when only low-resolution data are available, in which case a scaling through equating the Patterson origin peaks of native and derivative sets may provide better results (Rogers, 1965).

A model to account for errors in the data, determination of heavy-atom positions *etc.* was proposed by Blow & Crick (1959), in which all errors are associated with $|F_{PH}|_{obs}$ (Fig. 12.2.4.1); a more detailed treatment has been provided by Terwilliger & Eisenberg
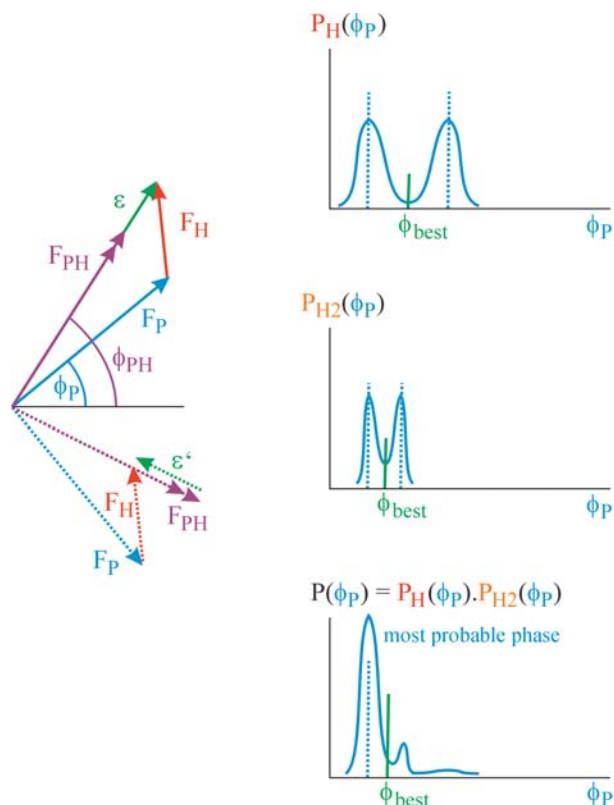

Fig. 12.2.4.1. The treatment of phase errors. The calculated heavy-atom structure results in a calculated value for both the phase and magnitude of F_H (red). According to the value of φ_P, the triangle $F_P - F_H - F_{PH}$ will fail to close by an amount ε, the lack of closure (green). This gives rise to a phase distribution which is bimodal for a single derivative. The combined probability from a series of derivatives has a most probable phase (the maximum) and a best phase (the centroid of the distribution), for which the overall phase error is minimum.

258

references