

2.3. Patterson and molecular-replacement techniques

BY M. G. ROSSMANN AND E. ARNOLD

2.3.1. Introduction

2.3.1.1. Background

Historically, the Patterson has been used in a variety of ways to effect the solutions of crystal structures. While some simple structures (Ketelaar & de Vries, 1939; Hughes, 1940; Speakman, 1949; Shoemaker *et al.*, 1950) were solved by direct analysis of Patterson syntheses, alternative methods have largely superseded this procedure. An early innovation was the heavy-atom method which depends on the location of a small number of relatively strong scatterers (Harker, 1936). Image-seeking methods and Patterson superposition techniques were first contemplated in the late 1930s (Wrinch, 1939) and applied sometime later (Beevers & Robertson, 1950; Clastre & Gay, 1950; Garrido, 1950*a*; Buerger, 1959). This experience provided the encouragement for computerized vector-search methods to locate individual atoms automatically (Mighell & Jacobson, 1963; Kraut, 1961; Hamilton, 1965; Simpson *et al.*, 1965) or to position known molecular fragments in unknown crystal structures (Nordman & Nakatsu, 1963; Huber, 1965). The Patterson function has been used extensively in conjunction with the isomorphous replacement method (Rossmann, 1960; Blow, 1958) or anomalous dispersion (Rossmann, 1961*a*) to determine the position of heavy-atom substitution. Pattersons have been used to detect the presence and relative orientation of multiple copies of a given chemical motif in the crystallographic asymmetric unit in the same or different crystals (Rossmann & Blow, 1962). Finally, the orientation and placement of known molecular structures ('molecular replacement') into unknown crystal structures can be accomplished *via* Patterson techniques.

The function, introduced by Patterson in 1934 (Patterson, 1934*a,b*), is a convolution of electron density with itself and may be defined as

$$P(\mathbf{u}) = \int_V \rho(\mathbf{x}) \cdot \rho(\mathbf{u} + \mathbf{x}) \, d\mathbf{x}, \quad (2.3.1.1)$$

where $P(\mathbf{u})$ is the 'Patterson' function at \mathbf{u} , $\rho(\mathbf{x})$ is the crystal's periodic electron density and V is the volume of the unit cell. The Patterson function, or F^2 series, can be calculated directly from the experimentally derived X-ray intensities as

$$P(\mathbf{u}) = \frac{2}{V^2} \sum_{\mathbf{h}}^{\text{hemisphere}} |\mathbf{F}_{\mathbf{h}}|^2 \cos 2\pi\mathbf{h} \cdot \mathbf{u}. \quad (2.3.1.2)$$

The derivation of (2.3.1.2) from (2.3.1.1) can be found in this volume (see Section 1.3.4.2.1.6) along with a discussion of the physical significance and symmetry of the Patterson function, although the principal properties will be restated here.

The Patterson can be considered to be a vector map of all the pairwise interactions between the atoms in a unit cell. The vectors in a Patterson correspond to vectors in the real (direct) crystal cell but translated to the Patterson origin. Their weights are proportional to the product of densities at the tips of the vectors in the real cell. The Patterson unit cell has the same size as the real crystal cell. The symmetry of the Patterson comprises the Laue point group of the crystal cell plus any additional lattice symmetry due to Bravais centring. The reduction of the real space group to the Laue symmetry is produced by the translation of all vectors to the Patterson origin and the introduction of a centre of symmetry. The latter is a consequence of the relationship between the vectors \mathbf{AB} and \mathbf{BA} . The Patterson symmetries for all 230 space groups are tabulated in *IT A* (1983).

An analysis of Patterson peaks can be obtained by considering N atoms with form factors f_i in the unit cell. Then

$$\mathbf{F}_{\mathbf{h}} = \sum_{i=1}^N f_i \exp(2\pi i\mathbf{h} \cdot \mathbf{x}_i).$$

Using Friedel's law,

$$\begin{aligned} |\mathbf{F}_{\mathbf{h}}|^2 &= \mathbf{F}_{\mathbf{h}} \cdot \mathbf{F}_{\mathbf{h}}^* \\ &= \left[\sum_{i=1}^N f_i \exp(2\pi i\mathbf{h} \cdot \mathbf{x}_i) \right] \left[\sum_{j=1}^N f_j \exp(-2\pi i\mathbf{h} \cdot \mathbf{x}_j) \right], \end{aligned}$$

which can be decomposed to

$$|\mathbf{F}_{\mathbf{h}}|^2 = \sum_{i=1}^N f_i^2 + \sum_{i \neq j}^N f_i f_j \exp[2\pi i\mathbf{h} \cdot (\mathbf{x}_i - \mathbf{x}_j)]. \quad (2.3.1.3)$$

On substituting (2.3.1.3) in (2.3.1.2), we see that the Patterson consists of the sum of N^2 total interactions of which N are of weight f_i^2 at the origin and $N(N-1)$ are of weight $f_i f_j$ at $\mathbf{x}_i - \mathbf{x}_j$.

The weight of a peak in a real cell is given by

$$w_i = \int_U \rho_i(\mathbf{x}) \, d\mathbf{x} = Z_i \quad (\text{the atomic number}),$$

where U is the volume of the atom i . By analogy, the weight of a peak in a Patterson (form factor $f_i f_j$) will be given by

$$w_{ij} = \int_U P_{ij}(\mathbf{u}) \, d\mathbf{u} = Z_i Z_j.$$

Although the maximum height of a peak will depend on the spread of the peak, it is reasonable to assume that heights of peaks in a Patterson are proportional to the products of the atomic numbers of the interacting atoms.

There are a total of N^2 interactions in a Patterson due to N atoms in the crystal cell. These can be represented as an $N \times N$ square matrix whose elements \mathbf{u}_{ij} , w_{ij} indicate the position and weight of the peak produced between atoms i and j (Table 2.3.1.1). The N vectors corresponding to the diagonal of this matrix are located at the Patterson origin and arise from the convolution of each atom with itself. This leaves $N(N-1)$ vectors whose locations depend on the relative positions of all of the atoms in the crystal cell and whose weights depend on the atom types related by the vector. Complete specification of the unique non-origin Patterson vectors requires description of only the $N(N-1)/2$ elements in either the upper or the lower triangle of this matrix, since the two sets of vectors represented by the two triangles are related by a centre of symmetry [$\mathbf{u}_{ij} \equiv \mathbf{x}_i - \mathbf{x}_j = -\mathbf{u}_{ji} \equiv -(\mathbf{x}_j - \mathbf{x}_i)$]. Patterson vector positions are usually represented as (uvw) , where u , v and w are expressed as fractions of the Patterson cell axes.

2.3.1.2. Limits to the number of resolved vectors

If we assume a constant number of atoms per unit volume, the number of atoms N in a unit cell increases in direct proportion with the volume of the unit cell. Since the number of non-origin peaks in the Patterson function is $N(N-1)$ and the Patterson cell is the same size as the real cell, the problem of overlapping peaks in the Patterson function becomes severe as N increases. To make matters worse, the breadth of a Patterson peak is roughly equal to the sum of the breadth of the original atoms. The effective width of a Patterson peak will also increase with increasing thermal motion, although this effect can be artificially reduced by sharpening techniques. Naturally, a loss of attainable resolution at high scattering angles