equivalent to a Patterson with coefficients $|\mathbf{F}''_{\mathbf{h}}(H)|^2$ which is proportional to $|\mathbf{F}_{\mathbf{h}}(H)|^2$. Such a Patterson (Rossmann, 1961*a*) will have vectors between all anomalous scatterers with heights proportional to the number of anomalous electrons $\Delta f''$. This 'anomalous dispersion' Patterson function has been used to find anomalous scatterers such as iron (Smith *et al*., 1983; Strahs & Kraut, 1968) and sulfur atoms (Hendrickson & Teeter, 1981).

It is then apparent that a Patterson with coefficients

$$\Delta F^2_{\rm ANO} = (|\mathbf{F}_{\mathbf{h}}| - |\mathbf{F}_{\bar{\mathbf{h}}}|)^2$$

(Rossmann, 1961*a*), as well as a Patterson with coefficients

$$\Delta F^2_{\rm ISO} = (|\mathbf{F}_{NH}| - |\mathbf{F}_H|)^2$$

(Rossmann, 1960; Blow, 1958), represent Pattersons of the heavy atoms. The $\Delta F^2_{\rm ANO}$ Patterson suffers from errors which may be larger than the size of the Bijvoet differences, while the $\Delta F^2_{\rm ISO}$ Patterson may suffer from partial lack of isomorphism. Hence, Kartha & Parthasarathy (1965) have suggested the use of the sum of these two Pattersons, which would then have coefficients $(\Delta F^2_{\rm ANO} + \Delta F^2_{\rm ISO})$.

However, given both SIR and anomalous-dispersion data, it is possible to make an accurate estimate of the $|\mathbf{F}_H|^2$ magnitudes for use in a Patterson calculation [Blundell & Johnson (1976, p. 340), Matthews (1966), Singh & Ramaseshan (1966)]. In essence, the Harker phase diagram can be constructed out of three circles: the native amplitude and each of the two isomorphous Bijvoet differences, giving three circles in all (Blow & Rossmann, 1961) which should intersect at a single point thus resolving the ambiguity in the SIR data and the anomalous-dispersion data. Furthermore, the phase ambiguities are orthogonal; thus the two data sets are cooperative. It can be shown (Matthews, 1966; North, 1965) that

$$F^2_N = F^2_{NH} + F^2_N \mp \frac{2}{k}(16k^2F^2_PF^2_H - \Delta I^2)^{1/2},$$

where $\Delta I = F^{+}_{NH}{}^2 - F^{-}_{NH}{}^2$ and $k = \Delta f''/f'$. The sign in the third-term expression is $-$ when $|(\alpha_{NH} - \alpha_H)| < \pi/2$ or + otherwise. Since, in general, $|\mathbf{F}_H|$ is small compared to $|\mathbf{F}_N|$, it is reasonable to assume that the sign above is usually negative. Hence, the heavy-atom lower estimate (HLE) is usually written as

$$F^2_{\rm HLE} = F^2_{NH} + F^2_H - \frac{2}{k}(16k^2F^2_PF^2_H - \Delta I^2)^{1/2},$$

which is an expression frequently used to derive Patterson coefficients useful in the determination of heavy-atom positions when both SIR and anomalous-dispersion data are available.

2.3.5. Noncrystallographic symmetry

2.3.5.1. *Definitions*

The interpretation of Pattersons can be helped by using various types of chemical or physical information. An obvious example is the knowledge that one or two heavy atoms per crystallographic asymmetric unit are present. Another example is the exploitation of a rigid chemical framework in a portion of a molecule (Nordman & Nakatsu, 1963; Burnett & Rossmann, 1971). One extremely useful constraint on the interpretation of Pattersons is noncrystallographic symmetry. Indeed, the structural solution of large biological assemblies such as viruses is only possible because of the natural occurrence of this phenomenon. The term 'molecular replacement' is used for methods that utilize noncrystallographic symmetry in the solution of structures [for earlier reviews see Rossmann (1972) and Argos & Rossmann (1980)]. These methods, which are only partially dependent on Patterson concepts, are discussed in Sections 2.3.6–2.3.8.

Crystallographic symmetry applies to the whole of the three-dimensional crystal lattice. Hence, the symmetry must be expressed both in the lattice and in the repeating pattern within the lattice. In contrast, noncrystallographic symmetry is valid only within a limited volume about the noncrystallographic symmetry element. For instance, the noncrystallographic twofold axes in the plane of the paper of Fig. 2.3.5.1 are true only in the immediate vicinity of each local dyad. In contrast, the crystallographic twofold axes perpendicular to the plane of the paper (Fig. 2.3.5.1) apply to every point within the lattice. Two types of noncrystallographic symmetry can be recognized: proper and improper rotations. A proper symmetry element is independent of the sense of rotation, as, for example, a fivefold axis in an icosahedral virus; a rotation either left or right by one-fifth of a revolution will leave all parts of a given icosahedral shell (but not the whole crystal) in equivalent positions. Proper noncrystallographic symmetry can also be recognized by the existence of a closed point group within a defined volume of the lattice. Improper rotation axes are found when two molecules are arbitrarily oriented relative to each other in the same asymmetric unit or when they occur in two entirely different crystal lattices. For instance, in Fig. 2.3.5.2, the object A_1B_1 can be rotated by $+\theta$ about the axis at P to orient it identically with A_2B_2. However, the two objects will not be coincident after a rotation of A_1B_1 by $-\theta$ or of A_2B_2 by $+\theta$. The envelope around each noncrystallographic object must be known in order to define an improper rotation. In contrast, only the volume about the closed point group need be defined for proper noncrystallographic operations. Hence, the boundaries of the repeating unit need not correspond to chemically covalently linked units in the presence of proper rotations.

Translational components of noncrystallographic rotation elements are said to be 'precise' in a direction parallel to the axis and

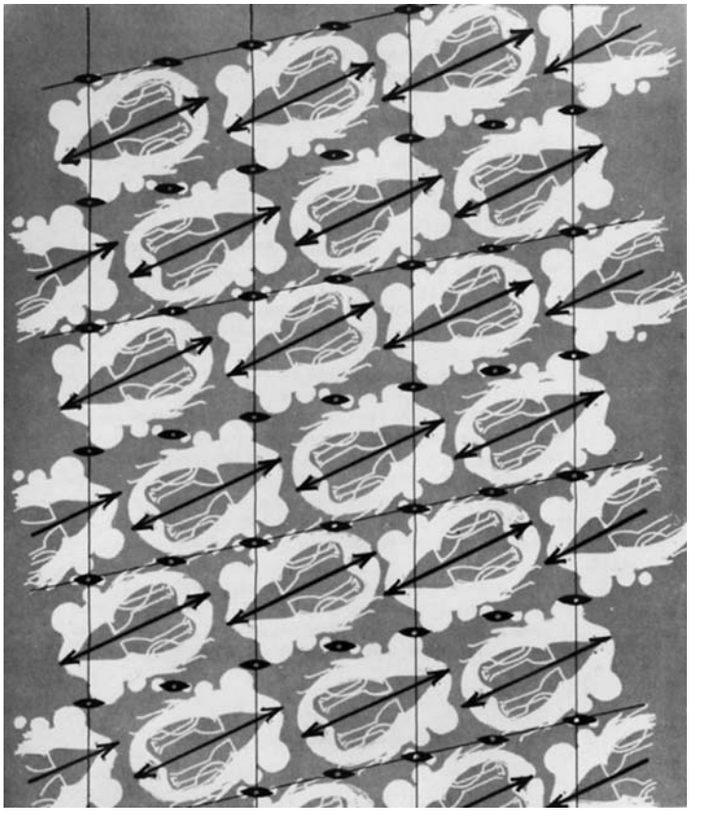

Fig. 2.3.5.1. The two-dimensional periodic design shows crystallographic twofold axes perpendicular to the page and local noncrystallographic rotation axes in the plane of the paper (design by Audrey Rossmann). [Reprinted with permission from Rossmann (1972, p. 8).]

Copyright © 2006 International Union of Crystallography

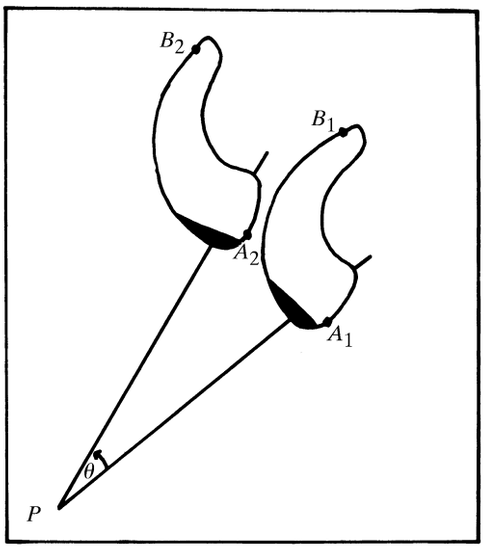

Fig. 2.3.5.2. The objects A_1B_1 and A_2B_2 are related by an improper rotation θ, since it is necessary to consider the sense of rotation to achieve superposition of the two objects. [Reprinted with permission from Rossmann (1972, p. 9).]

'imprecise' perpendicular to the axis (Rossmann *et al.*, 1964). The position, but not direction, of a rotation axis is arbitrary. However, a convenient choice is one that leaves the translation perpendicular to the axis at zero after rotation (Fig. 2.3.5.3).

Noncrystallographic symmetry has been used as a tool in structural analysis primarily in the study of biological molecules. This is due to the propensity of proteins to form aggregates with closed point groups, as, for instance, viruses with 532 symmetry. At best, only part of such a point group can be incorporated into the crystal lattice. Since biological materials cannot contain inversion elements, all studies of noncrystallographic symmetries have been limited to rotational axes. Reflection planes and inversion centres could also be considered in the application of molecular replacement to non-biological materials.

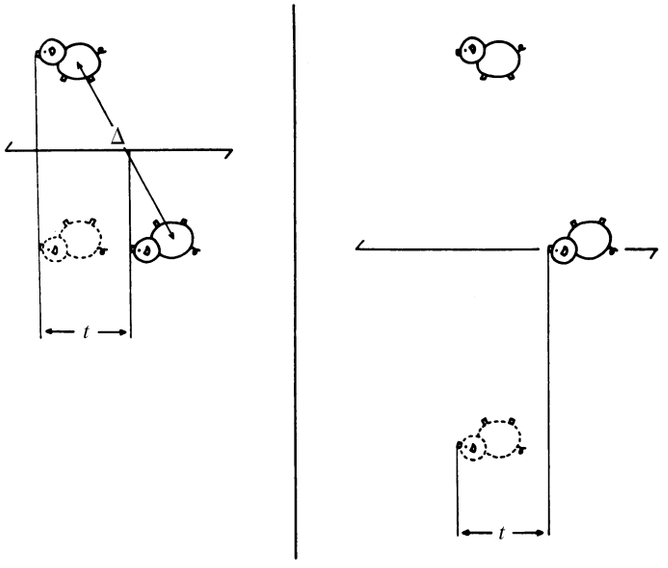

Fig. 2.3.5.3. The position of the twofold rotation axis which relates the two piglets is completely arbitrary. The diagram on the left shows the situation when the translation is parallel to the rotation axis. The diagram on the right has an additional component of translation perpendicular to the rotation axis, but the component parallel to the axis remains unchanged. [Reprinted from Rossmann *et al.* (1964).]

In this chapter, the relationship

$$\mathbf{x}' = [\boldsymbol{C}]\mathbf{x} + \mathbf{d}$$

will be used to describe noncrystallographic symmetry, where $\mathbf{x}$ and $\mathbf{x}'$ are position vectors, expressed as fractional coordinates, with respect to the crystallographic origin, $[\boldsymbol{C}]$ is a rotation matrix, and $\mathbf{d}$ is a translation vector. Crystallographic symmetry will be described as

$$\mathbf{x}' = [\boldsymbol{T}]\mathbf{x} + \mathbf{t},$$

where $[\boldsymbol{T}]$ and $\mathbf{t}$ are the crystallographic rotation matrix and translation vector, respectively. The noncrystallographic asymmetric unit will be defined as having n copies within the crystallographic asymmetric unit, and the unit cell will be defined as having m crystallographic asymmetric units. Hence, there are $L = nm$ noncrystallographic asymmetric units within the unit cell. Clearly, the n noncrystallographic asymmetric units cannot completely fill the volume of one crystallographic asymmetric unit. The remaining space must be assumed to be empty or to be occupied by solvent molecules which disobey the noncrystallographic symmetry.

2.3.5.2. *Interpretation of Pattersons in the presence of noncrystallographic symmetry*

If noncrystallographic symmetry is present, an atom at a general position within the relevant volume will imply the presence of others within the same crystallographic asymmetric unit. If the noncrystallographic symmetry is known, then the positions of equivalent atoms may be generated from a single atomic position. The additional vector interactions which arise from crystallographically and noncrystallographically equivalent atoms in a crystal may be predicted and exploited in an interpretation of the Patterson function.

An object in real space which has a closed point group may incorporate some of its symmetry in the crystallographic symmetry. If there are l such objects in the cell, then there will be mn/l equivalent positions within each object. The 'self-vectors' formed between these positions within the object will be independent of the position of the objects. This distinction is important in that the self-vectors arising from atoms interacting with other atoms within a single particle may be correctly predicted without the knowledge of the particle centre position. In fact, this distinction may be exploited in a two-stage procedure in which an atom may be first located relative to the particle centre by use of the self-vectors and subsequently the particle may be positioned relative to crystallographic symmetry elements by use of the 'cross-vectors' (Table 2.3.5.1).

The interpretation of a heavy-atom difference Patterson for the holo-enzyme of lobster glyceraldehyde-3-phosphate dehydrogenase (GAPDH) provides an illustration of how the known noncrystallographic symmetry can aid the solution (Rossmann *et al.*, 1972; Buehner *et al.*, 1974). The GAPDH enzyme crystallized in a $P2_12_12_1$ cell (a = 149.0, b = 139.1, c = 80.7 Å) containing one tetramer per asymmetric unit. A rotation-function analysis had indicated the presence of three mutually perpendicular molecular twofold axes which suggested that the tetramer had 222 symmetry, and a locked rotation function determined the precise orientation of the tetramer relative to the crystal axes (see Table 2.3.5.2). Packing considerations led to assignment of a tentative particle centre near $\frac{1}{2}, \frac{1}{4}, Z$.

An isomorphous difference Patterson was calculated for the K_2HgI_4 derivative of GAPDH using data to a resolution of 6.8 Å. From an analysis of the three Harker sections, a tentative first

Table 2.3.5.1. *Possible types of vector searches*

Self-vectors	Cross-vectors	Dimension of search, n
(1) Locate single site relative to particle centre		$n = 3$
(2)	Use information from (1) to locate particle centre	$n \leq 3$
(3) Simultaneous search for both (1) and (2). In general this is a six-dimensional search but may be simplified when particle is on a crystallographic symmetry axis		$3 \leq n \leq 6$
(4) Given (1) for more than one site, find all vectors within particle		$n = 3$
(5) Given information from (3), locate additional site using complete vector set		$n = 3$

heavy-atom position was assigned (atom A_2 at x, y, z). At this juncture, the known noncrystallographic symmetry was used to obtain a full interpretation. From Table 2.3.5.2 we see that molecular axis 2 will generate a second heavy atom with coordinates roughly $\frac{1}{4}+y, -\frac{1}{4}+x, 2Z-z$ (if the molecular centre was assumed to be at $\frac{1}{2}, \frac{1}{4}, Z$). Starting from the tentative coordinates of site A_2, the site A_1 related by molecular axis 1 was detected at about the predicted position and the second site A_1 generated acceptable cross-vectors with the earlier determined site A_2. Further examination enabled the completion of the set of four noncrystallographically related heavy-atom sites, such that all predicted Patterson vectors were acceptable and all four sites placed the molecular centre in the same position. Following refinement of these four sites, the corresponding SIR phases were used to find an additional set of four sites in this compound as well as in a number of other derivatives. The multiple isomorphous replacement phases, in conjunction with real-space electron-density averaging of the noncrystallographically related units, were then sufficient to solve the GAPDH structure.

When investigators studied larger macromolecular aggregates such as the icosahedral viruses, which have 532 point symmetry, systematic methods were developed for utilizing the noncrystallographic symmetry to aid in locating heavy-atom sites in isomorphous heavy-atom derivatives. Argos & Rossmann (1974, 1976) introduced an exhaustive Patterson search procedure for a single heavy-atom site within the noncrystallographic asymmetric unit which has been successfully applied to the interpretation of both virus [satellite tobacco necrosis virus (STNV) (Lentz *et al.*, 1976), southern bean mosaic virus (Rayment *et al.*, 1978), alfalfa mosaic virus (Fukuyama *et al.*, 1983), cowpea mosaic virus (Stauffacher *et al.*, 1987)] and enzyme [catalase (Murthy *et al.*, 1981)] heavy-atom difference Pattersons. A heavy atom is placed in turn at all plausible positions within the volume of the noncrystallographic asymmetric unit and the corresponding vector set is constructed from the resulting constellation of heavy atoms. Argos & Rossmann (1976) found a spherical polar coordinate search grid to be convenient for spherical viruses. After all vectors for the current search position are predicted, the vectors are allocated to the nearest grid point and the list is sorted to eliminate recurring ones. The criterion used by Argos & Rossmann for selecting a solution is that the sum

$$S = \sum_{i=1}^{N} P_i - NP_{\rm av}$$

of the lookup Patterson density values P_i achieves a high value for a correct heavy-atom position. The sum is corrected for the carpet of cross-vectors by the second term in the sum.

An additional criterion, which has been found useful for discriminating correct solutions, is a unit vector density criterion

$$U = \sum_{i=1}^{N} (P_i/n_i) \Big/ N,$$

where n_i is the number of vectors expected to contribute to the Patterson density value P_i (Arnold *et al.*, 1987). This criterion can be especially valuable for detecting correct solutions at special search positions, such as an icosahedral fivefold axis, where the number of vector lookup positions may be drastically reduced owing to the higher symmetry. An alternative, but equivalent, method for locating heavy-atom positions from isomorphous difference data is discussed in Section 2.3.3.5.

Even for a single heavy-atom site at a general position in the simplest icosahedral or $(T = 1)$ virus, there are 60 equivalent heavy atoms in one virus particle. The number of unique vectors corresponding to this self-particle vector set will depend on the crystal symmetry but may be as many as $(60)(59)/2 = 1770$ for a virus particle at a general crystallographic position. Such was the case for the STNV crystals which were in space group $C2$ containing four virus particles at general positions. The method of Argos & Rossmann was applied successfully to a solution of the K_2HgI_4 derivative of STNV using a 10 Å resolution difference Patterson. Application of the noncrystallographic symmetry vector search procedure to a $K_2Au(CN)_2$ derivative of human rhinovirus 14 (HRV14) crystals (space group $P2_13, Z = 4$) has succeeded in establishing both the relative positions of heavy atoms within one particle and the positions of the virus particles relative to the crystal symmetry elements (Arnold *et al.*, 1987). The particle position was established by incorporating interparticle vectors in the search and varying the particle position along the crystallographic threefold axis until the best fit for the predicted vector set was achieved.

Table 2.3.5.2. *Orientation of the glyceraldehyde-3-phosphate dehydrogenase molecular twofold axis in the orthorhombic cell*

Rotation axes	Polar coordinates (°)		Cartesian coordinates (direction cosines)		
	ψ	φ	u	v	w
1	45.0	−7.0	0.7018	0.7071	−0.0862
2	180.0–55.0	38.6	0.6402	−0.5736	0.5111
3	180.0–66.0	−70.6	0.3035	−0.4067	−0.8616

2.3.6. Rotation functions

2.3.6.1. *Introduction*

The rotation function is designed to detect noncrystallographic rotational symmetry (see Table 2.3.6.1). The normal rotation function definition is given as (Rossmann & Blow, 1962)

$$R = \int_U P_1(\mathbf{u}) \cdot P_2(\mathbf{u}')\, {\rm d}\mathbf{u}, \qquad (2.3.6.1)$$

where P_1 and P_2 are two Pattersons and U is an envelope centred at the superimposed origins. This convolution therefore measures the degree of similarity, or 'overlap', between the two Pattersons when P_2 has been rotated relative to P_1 by an amount defined by

references