Mathematical aspects of molecular replacement. II. Geometry of motion spaces

Gregory S. Chirikjian* and Yan Yan

Department of Mechanical Engineering, Johns Hopkins University, 3400 N. Charles Street, Baltimore, Maryland 21218, USA. Correspondence e-mail: gregc@jhu.edu

Molecular replacement (MR) is a well established computational method for phasing in macromolecular crystallography. In MR searches, spaces of motions are explored for determining the appropriate placement of rigid models of macromolecules in crystallographic asymmetric units. In the first paper of this series, it was shown that this space of motions, when endowed with an appropriate composition operator, forms an algebraic structure called a quasigroup. In this second paper, the geometric properties of these MR search spaces are explored and analyzed. This analysis includes the local differential geometry, global geometry and symmetry properties of these spaces.

1. Introduction

Molecular replacement (MR) is a computational method to phase macromolecular crystals (Rossmann & Blow, 1962; Rossmann, 2001; Vagin & Teplyakov, 2010). The inputs to MR computations are: (i) the electron density, $\rho(\mathbf{x})$, of a known rigid macromolecule (or fragment thereof) called the reference molecule; and (ii) the symmetry group of the crystal, Γ, which is a discrete subgroup of $G = \mathrm{SE}(n)$, the (continuous) Lie group of proper motions of rigid bodies in n-dimensional Euclidean space. While the three-dimensional case is of most interest in applications, much of the formulation presented here is applicable to the n-dimensional case, and $n = 2$ is used in some instances for illustration of concepts. In all cases, Γ has a normal subgroup of lattice translations, T.

The group operation for G and Γ is denoted as '$\circ$', and their action on Euclidean space, $\mathbb{R}^n$, is denoted as '$\cdot$'. Throughout this paper it often will be convenient to blur the distinction between the set of positions, $X \doteq \mathbb{R}^n$, and the continuous group of translations, $\mathcal{T} \doteq (\mathbb{R}^n, +)$, which contains T as a normal subgroup.

The reference molecule should be similar in structure to the one to be determined in order for the MR method to work. Such knowledge for proteins may come from prior knowledge of the similarity of the amino-acid sequences of the reference and actual molecules, and the many tens of thousands of existing structures in the Protein Data Bank (PDB) (Berman *et al.*, 2002).

In MR, the goal is to position and orient copies of the electron densities of the reference molecule in the crystallographic unit cell by some $g \in G$ to form a model density of the form

$$\rho_{\Gamma \backslash X}(\mathbf{x}; g) = \sum_{\gamma \in \Gamma} \rho[(\gamma \circ g)^{-1} \cdot \mathbf{x}]. \tag{1}$$

The density function $\rho(\mathbf{x})$ takes a non-negative value on the reference molecule and a zero value away from it.

Suppose that all of the dimensions of the reference molecule are smaller than all of the dimensions of $F_{\Gamma \backslash X}$, the fundamental domain X corresponding to $\Gamma \backslash X$. Then if the reference frame in which $\rho(\mathbf{x})$ is defined is centered at the origin of X, and $F_{\Gamma \backslash X}$ is defined to have its origin at the origin of X, then if g is a small motion, the body will still be fully contained in $F_{\Gamma \backslash X}$. In such a circumstance, the sum in equation (1) will only have nonzero contribution from $\gamma = e$. In contrast, if $F_{\Gamma \backslash X}$ were oddly shaped relative to the shape of the reference molecule, in such a way that does not allow motion of the reference molecule without it exiting one face of $F_{\Gamma \backslash X}$ and wrapping around another, then multiple terms in the sum in equation (1) would be required. This is one of many instances where the shape of $F_{\Gamma \backslash X}$ has computational implications.

For each fixed $g \in G$, $\rho_{\Gamma \backslash X}(\mathbf{x}; g)$ can be viewed either as a function on the asymmetric unit $F_{\Gamma \backslash X}$, or as a function on the unit cell $U = F_{T \backslash X}$. In the latter case, the function will have symmetry within the unit cell described by the finite factor group $\mathbb{F} = F_{T \backslash \Gamma} = F_{\Gamma / T}$. Here and throughout this paper, $F_{A \backslash B} \subset B$ denotes a fundamental domain from which the space B can be tiled or reconstructed by the left action of the group A on $F_{A \backslash B}$.

Using the notation established in Chirikjian (2011) (the first paper in this series), all candidate positions and orientations (called 'poses'), $g \in G$, can be chosen without loss of generality to be of the form $[g]_r \in F_{\Gamma \backslash G}$, or equivalently, each $[g]_r$ can be viewed as a representative of the coset $\Gamma g \in \Gamma \backslash G$.

These candidate poses can be evaluated and ranked according to the value of a cost function such as

$$C([g]_r) = \sum_{\mathbf{k} \in \hat{U}} (|\hat{\rho}_{\Gamma \backslash X}(\mathbf{k}; [g]_r)| - \hat{P}(\mathbf{k}))^2, \tag{2}$$

where $\hat{U}$ is its unitary dual (Fourier space) corresponding to U. The main goal of molecular replacement is to obtain a list of candidate poses $\{[g]_r\}$ rank ordered by the value of $C([g]_r)$.

This is a problem that requires numerical computation and involves a sampling scheme for $F_{\Gamma\backslash G}$ (which introduces a finite resolution). For any fixed resolution, it is desirable to obtain as rapidly as possible the list of the best candidate poses $\{[g]_r\}$ from among a large number of samples. To this end, in this paper the geometric properties of $\Gamma\backslash G$ are studied and related to $F_{\Gamma\backslash G}$ through the quasigroup operation $\hat{\circ}$ established in Chirikjian (2011), and an almost-uniform grid is established on these spaces for efficient sampling. Since the choice of how to define $F_{\Gamma\backslash G}$ is not unique, the relative merits of different choices are compared and contrasted in terms of how they affect the cost of MR computations.

An extensive list of references to the MR literature was provided in the first paper in this series, which is not repeated here. This second paper describes molecular replacement in terms of the local and global geometry of coset spaces of the form described above. The remainder of this paper is structured as follows. §2 reviews at a high level the motivation for studying the geometry of motion spaces and the objectives of this paper. §3 relates the geometry of motion spaces to the quasigroup operation defined in the first paper in this series. §4 discusses issues in the parametric representation of continuous rigid-body motions and crystallographic symmetry as they relate to establishing coordinate grids for MR computations. §5 discusses the geometry of unit cells and asymmetric units, including how to measure distance when faces are glued, and their relationship to $F_{\Gamma\backslash G}$. §6 develops the geometric properties of the motion space $(F_{\Gamma\backslash G}, \hat{\circ})$ which, as explained in the first paper in this series, is a quasigroup consisting of elements denoted as $[g]_r$, each of which is a representative of the coset Γg where $g \in G$. §7 addresses connections between group theory and the geometry of $F_{\Gamma\backslash G}$. §8 illustrates these concepts with planar examples and §9 develops efficient geometry-based sampling techniques for these spaces.

2. Motivation and objectives

This paper explores the geometry of the motion spaces $F_{\Gamma\backslash G}$ and how different realizations of these spaces impact on the cost of MR computations. Specifically, the following topics are addressed:

(a) The relationship between left-invariant metrics for $G = \mathrm{SE}(n)$ and several instantiations of $F_{\Gamma\backslash G}$ is established. Namely, $F_{\Gamma\backslash G}$ can be realized as Voronoi cells in G centered on elements of Γ.

(b) It is shown how the 'gluing' of boundary points of $F_{\Gamma\backslash G}$ can be described using the quasigroup operation $\hat{\circ}$ established in part I of this series.

(c) A new, almost-uniform sampling scheme for $\mathrm{SO}(n)$ (and hence for $F_{\Gamma\backslash G}$) is defined by dividing $\mathrm{SO}(n)$ up into Voronoi cells centered on symmetry operations of a Platonic solid and their n-dimensional analogs [which form finite groups $\Pi < \mathrm{SO}(n)$]. Sampling is achieved using the exponential map to parameterize each cell, and the deviation of this scheme from uniformity is analyzed.

(d) The fiber-bundle structure of $F_{\Gamma\backslash G}$ is established and algebraic definitions of this space are reconciled with geometric ones.

(e) This geometric information is related to the problem of MR computation, which has two facets: computational storage requirements and the number of computational operations.

A discussion of desirable properties for choices of $F_{\Gamma\backslash G}$ (such as symmetry, convexity and maximal closure under inversion of quasigroup elements) based on the above findings is also provided, leading to simplified data structures and associated MR computations.

3. Global geometry of $\Gamma\backslash G$ and gluing of $F_{\Gamma\backslash G}$ via the quasigroup operation

Whereas the emphasis of the first paper in this series was on algebraic properties of $(F_{\Gamma\backslash G}, \hat{\circ})$, the emphasis in the current paper is on geometry and its impact on the computational cost of the MR problem alluded to in §1.

The distance between points in $G = \mathrm{SE}(n)$ is measured by any number of distance metrics of the form $d_G : G \times G \to \mathbb{R}_{\geq 0}$ as reviewed in Chirikjian & Zhou (1998). For now, the discussion can be left general. The only additional property of importance beyond those in the definition of a metric is that these metrics will be taken to be left-invariant, i.e.

$$d_G(h \circ g_1, h \circ g_2) = d_G(g_1, g_2) \tag{3}$$

for any $h, g_1, g_2 \in G$.

The distance between points in $F_{\Gamma\backslash G}$ is then measured by the metric $d_{\Gamma\backslash G} : F_{\Gamma\backslash G} \times F_{\Gamma\backslash G} \to \mathbb{R}_{\geq 0}$ where

$$d_{\Gamma\backslash G}([g_1]_r, [g_2]_r) \doteq \min_{\gamma \in \Gamma} d_G(\gamma \cdot [g_1]_r, [g_2]_r)$$
$$= d_{\Gamma\backslash G}(g_1, g_2). \tag{4}$$

(The proof that this is a metric follows later in the paper.) According to this metric, disconnected regions on the boundary of the closure of $F_{\Gamma\backslash G}$, which is denoted as $\overline{F_{\Gamma\backslash G}}$, can have zero distance from each other, and identifying points in such regions with each other is precisely the sort of gluing operation alluded to in the field of low-dimensional geometry and topology.

The emphasis in MR is $F_{\Gamma\backslash G} \subset G$, whereas the emphasis in the pure mathematics literature is $\Gamma\backslash G$. These objects are related to each other as

$$F_{\Gamma\backslash G} + \text{gluing} \cong \Gamma\backslash G.$$

The 'gluing' operation, which is geometric in nature, can be related to the quasigroup operation $\hat{\circ}$ defined in Chirikjian (2011), which is algebraic. Consider two points $[g_1]_r \neq [g_2]_r \in \partial\overline{F_{\Gamma\backslash G}}$ (the boundary of $\overline{F_{\Gamma\backslash G}}$ in G). Let $[g_1]_r \sim [g_2]_r$ denote that they are glued and $[g_1]_r \not\sim [g_2]_r$ denote that they are not. Let $Y \in \mathcal{G}$ (the Lie algebra for G) such that $g_2(t) \doteq [g_1]_r \circ \exp(tY)$ where $t \in \mathbb{R}_{>0}$. Two possibilities exist as t approaches zero from above without reaching it: (i) either $[g_2(t)]_r$ approaches $\partial\overline{F_{\Gamma\backslash G}}$ at a point close to $[g_1]_r$ in the sense that $d_G([g_1]_r, [g_2(t)]_r) \to 0$ or (ii) $[g_2(t)]_r$ approaches $\partial\overline{F_{\Gamma\backslash G}}$ at

research papers

a point not close to $[g_1]_r$. Let $[g_2]_r = \lim_{t\to 0^+}[g_2(t)]_r = \lim_{t\to 0^+}[g_1]_r \hat{\circ} [\exp(tY)]_r$. Then, in the first case, $[g_1]_r$ and $[g_2]_r$ are proximal points in the same neighborhood of $\partial \overline{F_{\Gamma\backslash G}}$ requiring no gluing, and in the second case they are distant points [as measured with $d_G(\cdot,\cdot)$] on $\partial \overline{F_{\Gamma\backslash G}}$ that should be glued.

This can be summarized as

$$\lim_{t\to 0^+}\frac{d_{\Gamma\backslash G}([g_1]_r,[g_2(t)]_r)}{d_G([g_1]_r,[g_2(t)]_r)} = \begin{cases} 0 & \text{if } [g_1]_r \sim [g_2]_r \\ 1 & \text{if } [g_1]_r \not\sim [g_2]_r \end{cases}.$$

Moreover, it is possible to use $d_G(\cdot,\cdot)$ to construct $F_{\Gamma\backslash G}$ geometrically and to decompose $\partial \overline{F_{\Gamma\backslash G}}$ into faces. In particular $\overline{F_{\Gamma\backslash G}}$ will be the region in G consisting of all g such that

$$d_G(e,g) \le d_G(g,\gamma) \;\forall\; \gamma \ne e \in \Gamma. \qquad (5)$$

This region, which is closed and convex in G relative to the metric $d_G(\cdot,\cdot)$, has a boundary $\partial \overline{F_{\Gamma\backslash G}}$ composed of faces. In particular, a face is defined to be the set of all $[g]_r \in \partial \overline{F_{\Gamma\backslash G}}$ such that $d_G(e,[g]_r) = d_G([g]_r,\gamma)$ for some fixed $\gamma \ne e \in \Gamma$. Two faces meet when this condition is met by more than one such γ. For our purposes, the distinction between $F_{\Gamma\backslash G}$ and $\overline{F_{\Gamma\backslash G}}$ can be blurred since they differ only by a set of measure zero, and $F_{\Gamma\backslash G}$ can be constructed from $\overline{F_{\Gamma\backslash G}}$ by removing appropriate facets of $\partial \overline{F_{\Gamma\backslash G}}$.

Very different versions of $F_{\Gamma\backslash G}$ can result from different choices of $d_G(\cdot,\cdot)$. The next section establishes families of metrics based on the exponential and logarithm maps, thereby parameterizing families of choices for $F_{\Gamma\backslash G}$.

4. Parameterizing continuous rigid-body motions

This section provides a brief review of the matrix exponential and its role in parameterizing continuous rigid-body motions in Euclidean space. This will be important for defining an important class of candidates for $d_G(\cdot,\cdot)$ and in establishing almost-uniform grids in $F_{\Gamma\backslash G}$. Uniformity in sampling has obvious computational implications in MR because for a fixed number of sample points (and corresponding data storage) a uniform sampling with respect to a given metric would provide the best resolution.

4.1. The special orthogonal and special Euclidean groups

Let Γ denote the chiral crystallographic space group of symmetries of a macromolecular crystal. Γ can be viewed as a subgroup of the group of rigid-body motions, $G = \mathrm{SE}(n)$. This relationship is written as $\Gamma < G$. The group G consists of all rotation–translation pairs $g = (R,\mathbf{t})$ where R is an $n \times n$ rotation matrix, the set of which forms the special orthogonal group $\mathrm{SO}(n)$ under the operation of matrix multiplication, and $\mathbf{t} \in \mathbb{R}^n$ is a translation vector. The group operation for G,

$$g_1 \circ g_2 = (R_1 R_2, R_1\mathbf{t}_2 + \mathbf{t}_1),$$

is equivalent to the multiplication

$$H(g_1)\,H(g_2) = H(g_1 \circ g_2) \qquad (6)$$

of $(n+1) \times (n+1)$ homogeneous transformation matrices of the form

$$H(g) = \begin{pmatrix} R & \mathbf{t} \\ \mathbf{0}^T & 1 \end{pmatrix}, \qquad (7)$$

where $\mathbf{0}^T = [0,0,0]$ is the transpose of the column vector $\mathbf{0}$. The distinction between the faithful matrix representation of the group G in equation (7) *versus* G itself is often blurred in the literature, and it will be here as well. G contains two important continuous subgroups: (i) pure translations consisting of elements of the form $(\mathbb{I},\mathbf{t})$; and (ii) pure rotations consisting of elements of the form $(R,\mathbf{0})$. These subgroups are, respectively, isomorphic to the groups $\mathcal{T} \doteq (\mathbb{R}^n,+)$ and $\mathcal{R} \doteq \mathrm{SO}(n)$. The group law for $G = \mathrm{SE}(n)$ above is that of a semi-direct product, so that

$$G = \mathcal{T} \rtimes \mathcal{R}. \qquad (8)$$

And any rigid-body motion can be decomposed as

$$(R,\mathbf{t}) = (\mathbb{I},\mathbf{t}) \circ (R,\mathbf{0}). \qquad (9)$$

The case $n = 3$ is of particular interest and specialized notation can be established. For example, let

$$\Omega = \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix} \qquad (10)$$

and let $\boldsymbol{\omega} = [\omega_1,\omega_2,\omega_3]^T \in \mathbb{R}^3$. The set of all such Ω is denoted as $so(3)$, which together with the operations of addition and the well known matrix commutator operation, $[\Omega_1,\Omega_2] \doteq \Omega_1\Omega_2 - \Omega_2\Omega_1$, is the Lie algebra associated with $\mathrm{SO}(3)$. Here $\mathbb{R}^3 \cong so(3)$ (as vector spaces), where the bijective mapping between them is established by bringing Ω and $\boldsymbol{\omega}$ into correspondence. For any vector $\mathbf{x} \in \mathbb{R}^3$, $\Omega\mathbf{x} = \boldsymbol{\omega} \times \mathbf{x}$, the cross product of $\boldsymbol{\omega}$ and $\mathbf{x}$.

4.2. The fiber-bundle structure of G and $\Gamma\backslash G$

Since every rigid-body motion $g = (R,\mathbf{t})$ can be decomposed as the product of a translation and a rotation $(\mathbb{I},\mathbf{t}) \circ (R,\mathbf{0})$ as in equation (9), $\mathrm{SO}(n)$ acts trivially on G (and on $\Gamma\backslash G$) from the right. That is, $(R_1,\mathbf{t}_1) \circ (R_2,\mathbf{0}) = (R_1 R_2, \mathbf{t}_1)$. A projection map $\pi : G \to X$ can be defined as $\pi(g_1) = \pi(R_1,\mathbf{t}_1) \doteq \mathbf{t}_1$. The trivial section mapping $\sigma : X \to G$ can be defined as $\sigma(\mathbf{t}_1) \doteq (\mathbb{I},\mathbf{t}_1)$, which obviously satisfies $\pi(\sigma(\mathbf{t}_1)) = \mathbf{t}_1$. Moreover, under right $\mathrm{SO}(n)$ actions $\pi(\sigma(\mathbf{t}_1) \circ (R_2,\mathbf{0})) = \pi(\sigma(\mathbf{t}_1))$, and under left Γ actions, $\pi(\gamma \circ \sigma(\mathbf{t}_1)) = \gamma \cdot \pi(\sigma(\mathbf{t}_1))$. This gives both G and $\Gamma\backslash G$ the structure of trivial $\mathrm{SO}(n)$ bundles. The base space of G is $X = \mathbb{R}^n$, and the base space of $\Gamma\backslash G$ is $\Gamma\backslash X$. Similarly, $F_{\Gamma\backslash G}$ is a trivial $\mathrm{SO}(n)$ bundle with base space $F_{\Gamma\backslash X}$. This means that we can always make the choice

$$F_{\Gamma\backslash G} = F_{\Gamma\backslash X} \times \mathcal{R}. \qquad (11)$$

But whether or not this is the best choice depends on a number of factors related to how uniformly $F_{\Gamma\backslash X}$ and $\mathcal{R}$ can be sampled in comparison with alternatives explored in §§7–9.

4.3. Matrix exponentials and logarithms in the three-dimensional case

Using the notation $\theta = \|\omega\|$ where $\|\cdot\|$ denotes the vector 2-norm, $\omega = \theta\mathbf{n}$ and $\Omega = \theta N$, it can be shown that every element of SO(3) can be expressed as the matrix exponential

$$\exp(\theta N) = \mathbb{I} + \sin\theta N + (1 - \cos\theta)N^2, \qquad (12)$$

where $\mathbb{I}$ is the 3×3 identity matrix and $\exp(\theta N)$ can also be written as $e^{\theta N}$. The above equation is essentially Rodrigues' rotation formula. This exponential together with the correspondence $\theta N \leftrightarrow \theta\mathbf{n}$ can be used to surjectively map the closed ball $B_{r \leq \pi} \subset \mathbb{R}^3 \cong so(3)$ onto SO(3), where $\theta \in [0, \pi]$ and $\mathbf{n} \in S^2$, the unit sphere. However, this exponential mapping is only invertible *via* the logarithm map when making the restriction $\theta < \pi$, and considering the set

$$SO_<(3) \doteq \{e^{\theta N} \mid \theta \in [0, \theta), \mathbf{n} \in S^2\}$$
$$= SO(3) - \{e^{\pi N} \mid \mathbf{n} \in S^2\}.$$

By defining $so_<(3) \cong B_{r<\pi}$ the mapping $so_<(3) \leftrightarrow SO_<(3)$ is bijective. The natural way to define distance between any two rotations R_1 and R_2 related as $R_2 = R_1 e^{\theta N}$ is to evaluate θ. This is equivalent to saying that $d_{SO(3)}(R_1, R_2) \doteq \|\log(R_1^T R_2)\|$. This is well defined for $R_1^T R_2 \in SO_<(3)$, and is extended to the set of measure zero in SO(3) where the matrix logarithm, $\log(\cdot)$, fails by defining $d_{SO(3)}(R_1, R_2)$ to be equal to π at all of those points. It is well known that $d_{SO(3)}(R_1, R_2)$ is a valid distance function, and it is the geodesic distance obtained when viewing SO(3) as a Riemannian manifold with a Riemannian metric tensor that is set to be the identity matrix. This distance function is both left- and right-invariant in the sense that

$$d_{SO(3)}(AR_1, AR_2) = d_{SO(3)}(R_1 A, R_2 A)$$
$$= d_{SO(3)}(R_1, R_2) \qquad (13)$$

for all $A, R_1, R_2 \in SO(3)$. This is not the only such metric for SO(3). For example, $\|R_1 - R_2\|$ is also a bi-invariant metric where $\|\cdot\|$ is the Frobenius norm.

A formula similar to, though somewhat more complicated than, equation (12) holds for the matrix exponential for the group of rigid-body motions:

$$g = \exp\begin{pmatrix} \Omega & \mathbf{b} \\ \mathbf{0}^T & 0 \end{pmatrix} = \begin{pmatrix} e^\Omega & J(\omega)\mathbf{b} \\ \mathbf{0}^T & 1 \end{pmatrix}, \qquad (14)$$

where

$$J(\omega) = \mathbb{I} + \frac{1 - \cos\|\omega\|}{\|\omega\|^2}\Omega + \frac{\|\omega\| - \sin\|\omega\|}{\|\omega\|^3}\Omega^2. \qquad (15)$$

The pair $(\Omega, \mathbf{b})$ can be identified with an element of the Lie algebra se(3), and the exponential map can be written as

$$\exp : se(3) \rightarrow SE(3).$$

This mapping is not bijective. But since $se(3) = \mathbb{R}^3 \times so(3) \cong \mathbb{R}^3 \times \mathbb{R}^3$, it is possible to restrict the discussion to $se_<(3) \doteq \mathbb{R}^3 \times B_{r<\pi}$ where $B_{r<\pi}$ denotes the open ball of radius π in $\mathbb{R}^3$, which is equivalent to $so_<(3)$. Then $\exp : se_<(3) \rightarrow SE_<(3)$ defines $SE_<(3) \subset SE(3)$, and these

differ by only the set of measure zero corresponding to $\exp(\mathbb{R}^3 \times S_\pi^2) \subset SE(3)$, where S_π^2 is the sphere of radius π. This follows from the exponential map for SO(3) and the Rodrigues formula [equation (12)]. With these definitions, the logarithm map $\log : SE_<(3) \rightarrow se_<(3)$ is well defined. This is applicable to discussions of how distance is defined in SE(3) and hence its geometry. Measuring distance in SE(3) is somewhat more involved than for SO(3), and it is not possible to define a bi-invariant metric. But a number of left-invariant metrics on SE(3) can be defined, which have been used in the design of mechanisms and machines (Chirikjian & Zhou, 1998; Park, 1995), including

$$d_G^{(0)}(g_1, g_2) = \|\log(g_1^{-1} \circ g_2)\|_W \qquad (16)$$

when $g_1^{-1} \circ g_2 \in SE_<(3)$. [The superscript (0) is to begin the enumeration of a number of different metrics for $G = SE(3)$ that will follow later.] Here the weighted Frobenius norm is $\|A\|_W = \text{tr}(AWA^T)$. The 4×4 weighting matrix can be chosen in a number of ways as discussed in the above references. One such choice is $W = (s^2\mathbb{I}_3) \oplus 1$ where $s \in \mathbb{R}_{>0}$ is a scale factor to reconcile translational displacements measured in units of length, and rotational displacements measured in radians. Then, even without leaving the class of metrics $d_G^{(0)}(g_1, g_2)$, a continuum of choices for $F_{\Gamma\backslash G}$ will result that are parameterized by s. For example, if $F_{\Gamma\backslash G}$ is the Voronoi cell centered on the identity of G constructed using $d_G^{(0)}(g_1, g_2)$, it will become the choice in equation (11) as $s \rightarrow 0$.

5. Geometry in unit cells and asymmetric units

As is well known in crystallography, the basic unit that is translated to replicate the whole crystal is called the *unit cell*. The unit cell can be broken up into so-called *asymmetric units*. The union of these asymmetric units reconstitutes the unit cell, and translated copies of the unit cell completely tile space.

Let $X = \mathbb{R}^n$ where $n = 2$ or 3. The set of orbits $\Gamma\backslash X$ can be viewed as a region in X, denoted as $F_{\Gamma\backslash X}$. A point in $F_{\Gamma\backslash X}$ is denoted as $[\mathbf{x}]$ and serves as a representative for its orbit. Each point $\mathbf{x} \in X$ can be thought of as $\mathbf{x} = \gamma \cdot [\mathbf{x}]$ for some $\gamma \in \Gamma$. If $\Gamma = T \cong P1$, the resulting fundamental domain can be chosen as the crystallographic unit cell, and for a more general space group Γ, $F_{\Gamma\backslash X}$ can be chosen as the asymmetric unit.[1]

This asymmetric unit has faces 'glued' to adjacent copies to form the unit cell. Tiling with unit cells can then be described using the quasigroup operation as in §3, but with G restricted to the subgroup of translations, $\mathcal{T}$.

5.1. Measuring distance

Since $X = \mathbb{R}^n$, the Euclidean distance between points $\mathbf{x}_1, \mathbf{x}_2 \in X$ can be computed easily as $d_X(\mathbf{x}_1, \mathbf{x}_2) = \|\mathbf{x}_1 - \mathbf{x}_2\|$. However, a natural question to ask is how to measure distance in $F_{\Gamma\backslash X}$? For example, it can be that two points close to antipodal faces of $\overline{F_{\Gamma\backslash X}}$ (and hence have the largest possible

---

[1] In the planar case '$P1$' should be replaced with '$p1$', but the meaning should be clear from the context.

Euclidean distance) actually should be considered close to each other, since opposing faces are glued.

This can be reflected by defining a metric for $F_{\Gamma\backslash X}$ of the form

$$d_{\Gamma\backslash X}([\mathbf{x}_1], [\mathbf{x}_2]) \doteq \min_{\gamma \in \Gamma} d_X(\gamma \cdot [\mathbf{x}_1], [\mathbf{x}_2]). \qquad (17)$$

The fact that this is symmetric, $d_{\Gamma\backslash X}([\mathbf{x}_1], [\mathbf{x}_2]) = d_{\Gamma\backslash X}([\mathbf{x}_2], [\mathbf{x}_1])$, follows from the invariance of the Euclidean norm under rigid-body motions, $d_X(\mathbf{x}_1, \mathbf{x}_2) = d_X(g \cdot \mathbf{x}_1, g \cdot \mathbf{x}_2)$ for any $g \in \mathrm{SE}(n)$, and the symmetry property $d_X(\mathbf{x}_1, \mathbf{x}_2) = d_X(\mathbf{x}_2, \mathbf{x}_1)$. The fact that $d_X(\mathbf{x}_1, \mathbf{x}_2) \geq 0$ with equality implying $\mathbf{x}_1 = \mathbf{x}_2$ leads to the analogous property for $d_{\Gamma\backslash X}([\mathbf{x}_1], [\mathbf{x}_2])$. The triangle equality is proven as follows:

$$d_{\Gamma\backslash X}([\mathbf{x}_1], [\mathbf{x}_3]) = \min_{\gamma \in \Gamma} d_X(\gamma \cdot [\mathbf{x}_1], [\mathbf{x}_3])$$

$$\overset{(a)}{\leq} \min_{\gamma \in \Gamma} \min_{\dot{\gamma} \in \Gamma}(d_X(\gamma \cdot [\mathbf{x}_1], \dot{\gamma} \cdot [\mathbf{x}_2]) + d_X(\dot{\gamma} \cdot [\mathbf{x}_2], [\mathbf{x}_3]))$$

$$\overset{(b)}{=} \min_{\dot{\gamma} \in \Gamma} \min_{\gamma \in \Gamma}(d_X((\dot{\gamma}^{-1} \circ \gamma) \cdot [\mathbf{x}_1], \cdot[\mathbf{x}_2]) + d_X(\dot{\gamma} \cdot [\mathbf{x}_2], [\mathbf{x}_3]))$$

$$\overset{(c)}{=} \min_{\dot{\gamma} \in \Gamma} \min_{\ddot{\gamma} \in \Gamma}(d_X(\ddot{\gamma}^{-1} \cdot [\mathbf{x}_1], \cdot[\mathbf{x}_2]) + d_X(\dot{\gamma} \cdot [\mathbf{x}_2], [\mathbf{x}_3]))$$

$$= d_{\Gamma\backslash X}([\mathbf{x}_1], [\mathbf{x}_2]) + d_{\Gamma\backslash X}([\mathbf{x}_2], [\mathbf{x}_3]).$$

The inequality (*a*) above follows from the fact that

$$\min_{\gamma \in \Gamma} d_X(\gamma \cdot [\mathbf{x}_1], [\mathbf{x}_3]) \leq \min_{\gamma \in \Gamma}(d_X(\gamma \cdot [\mathbf{x}_1], \dot{\gamma} \cdot [\mathbf{x}_2]) + d_X(\dot{\gamma} \cdot [\mathbf{x}_2], [\mathbf{x}_3]))$$

for *any* $\dot{\gamma} \in \Gamma$ since $d_X(\cdot, \cdot)$ is a metric and thus satisfies the triangle inequality. The equality in (*b*) follows from switching the order of the minimizations and using the SE(3)-invariance of the Euclidean metric, which of course also makes it Γ-invariant. And (*c*) is related to the closure of the group Γ under inversion and multiplication of elements. The new group element $\ddot{\gamma} \in \Gamma$ introduced in (*c*) is defined by the equality $\ddot{\gamma}^{-1} \doteq \dot{\gamma}^{-1} \circ \gamma$.

Note that in practice if we want to compute $d_{\Gamma\backslash X}(\cdot, \cdot)$, we do not have to evaluate over all $\gamma \in \Gamma$, but only those that correspond to motions inside and between adjacent unit cells. Essentially the same proof as given above can be used to establish the metric properties of $d_{\Gamma\backslash G}(\cdot, \cdot)$ given in equation (4).

5.2. Symmetry of unit cells and asymmetric units

In protein crystallography, many copies of a protein molecule are coaxed into forming a crystal in the laboratory. Then X-ray diffraction experiments can be performed to gain information about the shape of these molecules. Atomic models are then fitted to these shapes.

Fig. 1 shows a unit cell with $P2_12_12_1$ symmetry with an articulated 3-body molecular model that looks like a 'rabbit' in an $L \times L \times L$ unit cell. Let the corner of the unit cell (the box) be located at $(0, 0, 0)$, then the coordinates of the center of the face of the first rabbit (the purple one) are $(L/4, L/4, L/4)$ where $L = 80$ represents the dimension of the unit cell. The transformations that produce replicas are then (x, y, z); $(L/2 - x, L - y, L/2 + z)$; $(L/2 + x, L/2 - y, L - z)$; $(L - x, L/2 + y, L/2 - z)$. Here we have modified slightly the standard choice of coset representatives so that all four rabbits fit

in this unit cell. In all subfigures the orientation of each 'ear' relative to the 'face' is kept constant, and in this figure the principal axes of the face are aligned with the axes of the unit cell.

If each rabbit is rigidly moved while maintaining $P2_12_12_1$ symmetry, a new configuration such as that in Fig. 2 will result. Here the translation of the purple copy of $= [0, 2, -2]^T$ and a rotation in *ZYZ* Euler angles of $\alpha = \pi/2$, $\beta = \pi/2$, $\gamma = \pi/12$ is shown. In this case, all rabbits remain within the unit cell and so $g = [g]_r$.

The crystallographic space groups and corresponding asymmetric units have been studied extensively, and were completely classified by the late nineteenth and early twentieth century. Thorough modern treatments can be found in Farmer (1996), Hahn (2002), Aroyo *et al.* (2010) and the references in the first paper in this series. More recently, the classification of 3-manifolds constructed as quotients of $\mathbb{R}^3$ by space groups (called orbifolds) was initiated in the early 1980s, as summarized in Thurston (1997). Orbifolds generalize the concept of a manifold in such a way as to allow points that do not locally 'look like' Euclidean space. Other treatments can be found (Dunbar, 1981; Bonahon & Siebenmann, 1985; Charlap, 1986; Johnson *et al.*, 1997; Montesinos, 1987; Weeks, 1985; Nikulin & Shafarevich, 1987; Conway *et al.*, 2001). A precursor to the orbifold concept is that of the *V-manifold* introduced by Satake (1956). Though the orbifold structure of $\Gamma\backslash G$ and $\Gamma\backslash X$ is not used here explicitly, it nevertheless is useful to make connections between different fields as additional new insights might be brought to bear in MR problems in the future.

6. Differential-geometric properties of G and $F_{\Gamma\backslash G}$

In this section several more choices for distance metrics $d_G(\cdot, \cdot)$ are provided, thereby defining distance and volume in $F_{\Gamma\backslash G}$. Then the global geometry and elementary topological properties of these spaces are illustrated with examples.

6.1. Measuring distance and volume

One metric for G was given in equation (16). Three other natural ways to define how distance is measured in G are given below. For example, if $f : G \to \mathbb{R}$ is a smooth L^2 function monotonically decreasing away from the identity, it can be shown that

$$d_G^{(1)}(g_1, g_2) \doteq \left[\int_G |f(g_1^{-1} \circ g) - f(g_2^{-1} \circ g)|^2 \, dg \right]^{1/2} \qquad (18)$$

satisfies the properties of a metric/distance function for G, where dg is the unique bi-invariant (Haar) measure for G. Similarly, given a function $\rho : X \to \mathbb{R}$ that is in $L^1(X)$, it is possible to compute

$$d_G^{(2)}(g_1, g_2) \doteq \left\{ \int_X [d_X(g_1 \cdot \mathbf{x}, g_2 \cdot \mathbf{x})]^2 \rho(\mathbf{x}) \, d\mathbf{x} \right\}^{1/2},$$

where $d_X(\mathbf{x}, \mathbf{y})$ is a G-invariant metric such as the Euclidean distance between $\mathbf{x}, \mathbf{y} \in X$. Finally, metrics of the form
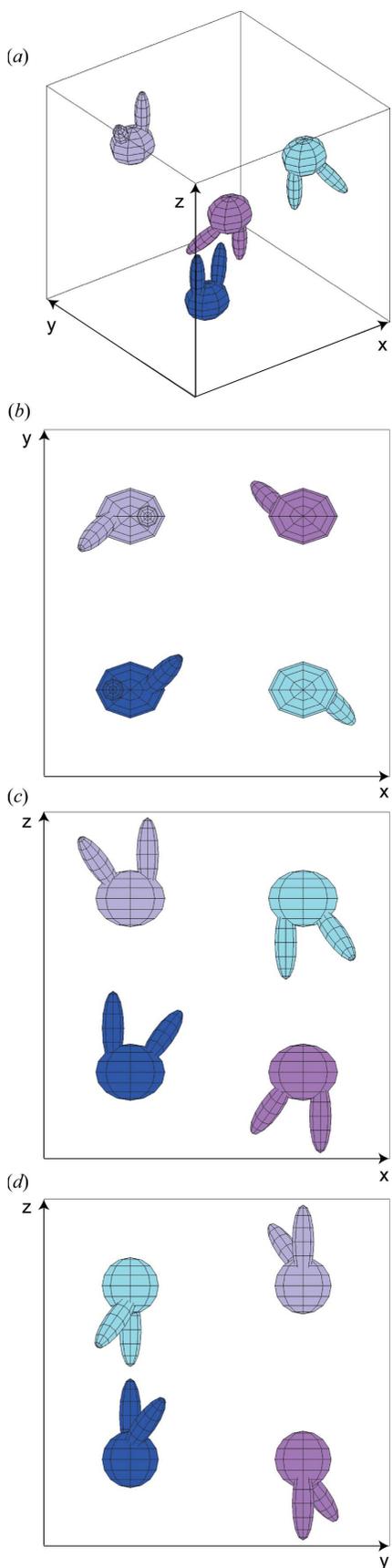
Figure 1
Objects arranged with $P2_12_12_1$ space-group symmetry. (a) Three-dimensional view; (b)–(d) projections.
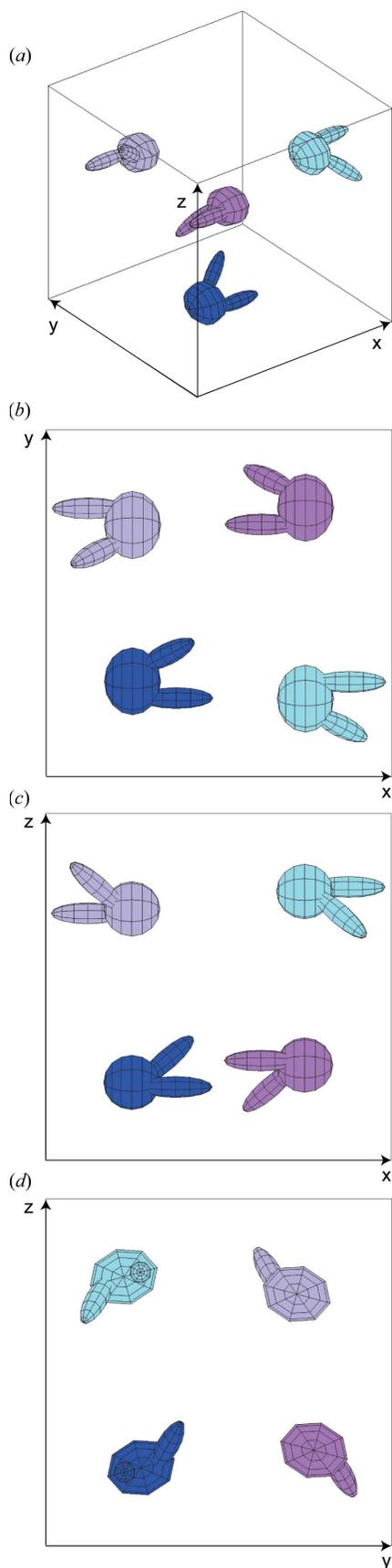

Figure 2
Coordinated movement of these objects that maintains $P2_12_12_1$ symmetry. (a) Three-dimensional view; (b)–(d) projections.

Chirikjian and Yan · Mathematical aspects of molecular replacement. II **213**

$$d_G^{(3)}(g_1, g_2) \doteq \left\{ \|\mathbf{t}_1 - \mathbf{t}_2\|^p + [s \cdot d_{SO(3)}(R_1, R_2)]^p \right\}^{1/p}$$

for $p = 1$ or 2 and a scale factor s are in use.

All of the metric functions listed above are left-invariant in the sense that $d_G^{(i)}(h \circ g_1, h \circ g_2) = d_G^{(i)}(g_1, g_2)$ for any $h \in G$.

Given a generic left-invariant metric function for G, metrics on $\Gamma \backslash G$ can be constructed as in equation (4). This is analogous to how the metric $d_{\Gamma \backslash X}([\mathbf{x}_1], [\mathbf{x}_2])$ in equation (17) was generated from a metric on X, and essentially the same proof of the triangle inequality applies.

Volume in G is defined using the Haar measure. For $G = SE(3)$ with translations and rotations, respectively, parameterized by Cartesian coordinates and Euler angles, the volume element is of the form

$$dg = (dx\,dy\,dz)(\sin \beta\, d\alpha\, d\beta\, d\gamma)$$

which is the product of volume elements for $\mathbb{R}^3$ and $SO(3)$.

Using this, the volume of the six-dimensional fundamental region $F_{\Gamma \backslash G}$ can be related to the volume of $SO(3)$ and the three-dimensional volumes of the asymmetric unit and unit cell as

$$V(F_{\Gamma \backslash G}) = 8\pi^2\, V(F_{\Gamma \backslash X}) = \frac{8\pi^2}{|\mathbb{F}|}\, V(F_{T \backslash X}), \qquad (19)$$

where $|\mathbb{F}| = |T \backslash \Gamma|$.

6.2. Minimization on the quasigroup $(F_{\Gamma \backslash G}, \hat{\circ})$

Functions on the quasigroup $(F_{\Gamma \backslash G}, \hat{\circ})$ such as $C([g]_r)$ in equation (2) were shown to arise naturally in the context of molecular replacement in the first paper in this series, where the goal is to minimize the error (or maximize the correlation) between a model of the crystallographic unit cell and the actual one described by an X-ray diffraction pattern.

Another function of relevance that can be used to pre-screen candidate values of $[g]_r \in F_{\Gamma \backslash G}$ is

$$\cap([g]_r) \doteq \int_{F_{\Gamma \backslash X}} [\rho_{\Gamma \backslash X}([\mathbf{x}]; [g]_r)]^2\, d[\mathbf{x}] - \int_X [\rho(\mathbf{x})]^2\, d\mathbf{x}. \qquad (20)$$

This function will be zero when the reference molecule does not intersect its symmetry mates and it will be positive when it does intersect. Since such intersections are not physically realizable, regions in $F_{\Gamma \backslash G}$ for which $\cap([g]_r) \gg 0$ can be immediately removed from consideration when performing a search for minima in $C([g]_r)$. Alternatively, a composite cost function can be constructed from the sum of $C([g]_r)$ and $\cap([g]_r)$, or a minimization procedure can interweave minimization steps alternating between $C([g]_r)$ and $\cap([g]_r)$. For this reason it makes sense to consider the general problem of how to minimize over $(F_{\Gamma \backslash G}, \hat{\circ})$.

Minimization of a function $f(\cdot)$ with argument in $F_{\Gamma \backslash G}$ can be performed much in the same way as minimization on G. In particular, the components of a gradient can be defined as

$$(\tilde{E}_i^r f)([g]_r) \doteq \frac{d}{dt} f([g]_r \circ e^{E_i t})|_{t=0}$$
$$= \frac{d}{dt} f([[g]_r \circ e^{E_i t}]_r)|_{t=0},$$

where $\{E_i\}$ is a basis for the Lie algebra $se(3)$ and $\tilde{E}_i^r$ is the corresponding directional (or Lie) derivative. The second of the above equalities holds because $f(g) = f(\gamma \circ g)$ for all $\gamma \in \Gamma$ and $g \in G$, and so $f(g) = f([g]_r)$ for any $g \in G$. A numerical optimization procedure for gradient descent is then achieved by iteratively computing $\alpha_i = (\tilde{E}_i^r f)([g]_r)$ and performing updates

$$[g]_r \to \left[[g]_r \circ \exp\left(-\varepsilon \sum_{i=1}^n \alpha_i E_i \right) \right]_r$$

that lead to lower values of $f([g]_r)$. This is simply a gradient descent procedure in which the current value is updated by following the negative of the gradient by a small amount ε.

From a practical point of view, when minimizing $C([g]_r)$ on a discrete grid of values of $[g]_r$, it would be convenient to have as uniform a grid as possible so that the result of gradient descent updates can be instantly rounded to the nearest grid point in a consistent manner without having to consult a look-up table or complicated data structure, or performing computations associated with multivariate (in this case six-dimensional) interpolation. This is one of several reasons why it makes sense to examine the different choices available for $F_{\Gamma \backslash G}$.

7. Algebraic-geometric properties of $F_{\Gamma \backslash G}$

The fundamental region $F_{\Gamma \backslash G} \subset G$ that is formed by collecting one representative group element for each coset $\Gamma g \in \Gamma \backslash G$ has interesting global geometric properties that can be expressed in the language of algebraic geometry. When $G = SE(3)$ and Γ is one of the 65 chiral space groups, $F_{\Gamma \backslash G}$ is a six-dimensional region that is difficult to visualize. And to the authors' knowledge, the geometry and topology of these have not been fully explored, even in the pure mathematics literature, let alone in crystallography.

In this section, different ways of choosing $F_{\Gamma \backslash G}$ are explored. Let $[g]_r \in \Gamma g \in \Gamma \backslash G$ be a coset representative which also is contained in the fundamental region $F_{\Gamma \backslash G} \subset G$. G is reconstructed from $F_{\Gamma \backslash G}$ by the union

$$G = \bigcup_{\gamma \in \Gamma} \gamma \cdot F_{\Gamma \backslash G}$$

where the action $\cdot$ of Γ on $F_{\Gamma \backslash G}$ is the product of γ with each element of $F_{\Gamma \backslash G}$ via the operation $\circ$ for G. Though there is no unique way to define $F_{\Gamma \backslash G}$, given that the cost function in equation (2) can be written via Parseval's equality as

$$C([g]_r) = \int_{\Gamma \backslash X} \left(|\rho_{\Gamma \backslash X}([\mathbf{x}]; [g]_r)| - P([\mathbf{x}]) \right)^2\, d[\mathbf{x}], \qquad (21)$$

it would be desirable to choose for a given $\rho(\mathbf{x})$ a pair $(F_{\Gamma \backslash G}, F_{\Gamma \backslash X})$ so that

$$K_\rho(F_{\Gamma\backslash G}, F_{\Gamma\backslash X}) \doteq$$
$$\int_{\Gamma\backslash G} \int_{\Gamma\backslash X} \left(\rho([g]_r^{-1} \cdot [\mathbf{x}]) - \rho_{\Gamma\backslash X}([\mathbf{x}]; [g]_r) \right)^2 \, \mathrm{d}[\mathbf{x}] \, \mathrm{d}[g]_r$$

is as small as possible. This means that, relative to the alternative choices, there is more room for the reference molecule to move in $F_{\Gamma\backslash X}$ under the action of $F_{\Gamma\backslash G}$ without penetrating $\partial F_{\Gamma\backslash X}$ than otherwise. In this way, the overall number of times that multiple terms in the summation in equation (1) will be required will be smaller than for other choices. In other words, $(F_{\Gamma\backslash G}, F_{\Gamma\backslash X})$ is better than $(F'_{\Gamma\backslash G}, F'_{\Gamma\backslash X})$ if

$$K_\rho(F_{\Gamma\backslash G}, F_{\Gamma\backslash X}) < K_\rho(F'_{\Gamma\backslash G}, F'_{\Gamma\backslash X})$$

because this will lead to fewer calculations when computing minima in $C([g]_r)$ by gradient descent since $\rho([g]_r^{-1} \cdot [\mathbf{x}])$ can be used in place of $\rho_{\Gamma\backslash X}([\mathbf{x}]; [g]_r)$ for a greater fraction of the space $F_{\Gamma\backslash G} \times F_{\Gamma\backslash X}$.

But it would be impractical to design a different pair $(F_{\Gamma\backslash G}, F_{\Gamma\backslash X})$ for each new test molecule. Therefore, it would be useful to establish some rules of thumb for good pairs in general.

Since, in the above discussion, $[g]_r^{-1} \cdot [\mathbf{x}]$ appears, we could seek $(F_{\Gamma\backslash G}, F_{\Gamma\backslash X})$ such that the mapping $m : F_{\Gamma\backslash G} \times F_{\Gamma\backslash X} \to X$ defined by $m([g]_r, [\mathbf{x}]) = [g]_r^{-1} \cdot [\mathbf{x}]$ maximizes the amount of points in its range that fall in $F_{\Gamma\backslash X}$. The solution to this problem is not known to the authors. A more manageable problem would be to design $F_{\Gamma\backslash G}$ so that it has nice closure properties on its own and so that it interacts with $F_{\Gamma\backslash X}$ in a predictable way. In general $[g]_r^{-1} \notin F_{\Gamma\backslash G}$ and it is not even possible to construct $F_{\Gamma\backslash G}$ in such a way that it is closed under inversion. This can be seen from the left-invariant metrics $d_G(g_1, g_2)$, such as equation (16), which satisfy $d_G^{(0)}(e, g) = d_G^{(0)}(e, g^{-1})$, but since they are not right-invariant, $d_G^{(0)}(\gamma, g) \neq d_G^{(0)}(\gamma^{-1}, g^{-1})$. And so even when $F_{\Gamma\backslash G}$ is constructed as a Voronoi cell centered on the identity $e \in G$, it will not be closed under inversion. An attempt to address this lack of closure is that $F_{\Gamma\backslash G}$ could be chosen so as to maximize the volume within it that is closed under inversion. That is, if $\delta(g_1, g_2)$ is the Kronecker delta (equal to unity when $g_1 = g_2$ and zero otherwise) for $g_1, g_2 \in G$, it may be desirable to design $F_{\Gamma\backslash G}$ so as to maximize

$$I(F_{\Gamma\backslash G}) \doteq \int_{F_{\Gamma\backslash G}} \delta([g]_r^{-1}, [[g]_r^{-1}]_r) \, \mathrm{d}[g]_r. \tag{22}$$

Second if $\mathbb{F} = F_{T\backslash\Gamma}$, and $\mathbb{S} < \mathbb{F}$ is the largest subgroup of purely rotational symmetry elements, then when $F_{T\backslash X}$ is chosen to be a Wigner–Seitz cell and $F_{\Gamma\backslash X}$ is a fraction thereof, it may be desirable for

$$p \, F_{\Gamma\backslash G} \, p^{-1} = F_{\Gamma\backslash G} \quad \forall \; p \in \mathbb{S} \tag{23}$$

in order to facilitate internal cancellation in the product $[g]_r^{-1} \cdot [\mathbf{x}]$ so that the result lands in $F_{\Gamma\backslash X}$. Alternatively, if $F_{\Gamma\backslash G}$ is defined as a Voronoi cell in G with respect to a metric such that $d_G(g_1, g_2) = d_G(pg_1p^{-1}, pg_2p^{-1})$, then equation (23) will also hold. The metric in equation (16) with $W = (s^2 \mathbb{I}) \oplus 1$ is one such example of this.

In the next section various ways of viewing $F_{\Gamma\backslash G}$ are explored from a group-theoretic perspective in order to inform future MR software design.

7.1. Viewing $F_{\Gamma\backslash G}$ as $F_{\Gamma\backslash X} \times \mathcal{R}$

Recall that $\mathcal{R} = SO(n)$ and $X = \mathbb{R}^n$ and $\mathcal{T} = (\mathbb{R}^n, +)$. In the symmorphic case, $\Gamma = T \rtimes \mathbb{P}$ (where T is the lattice translation group) and point group $\mathbb{P} < \Gamma$. In this case equation (23) can be achieved by choosing

$$F'_{\Gamma\backslash G} \doteq (F_{T\backslash X}) \times (F_{\mathbb{P}\backslash\mathcal{R}}) \tag{24}$$

with $F_{T\backslash X}$ taken to be the Wigner–Seitz unit cell centered at the origin. In contrast, $T\backslash\mathcal{T}$ would be the prismatic unit cell with opposing faces glued to form the torus $\mathbb{T}^n$.

In the nonsymmorphic case, equation (24) cannot be assumed. However, in both the symmorphic and nonsymmorphic cases it is possible to take

$$F_{\Gamma\backslash G} = (F_{\Gamma\backslash X}) \times \mathcal{R} \tag{25}$$

instead of equation (24). On the other hand, if one does computations in the unit cell rather than the asymmetric unit, then Γ is replaced with $P1 \cong T$, which is trivially symmorphic, and $F = F'$ and equation (23) will hold trivially.

The definition of the fundamental region is not unique for several reasons. This is demonstrated above in the symmorphic case where $F_{\Gamma\backslash G}$ exists and $F_{\Gamma\backslash G}$ and $F'_{\Gamma\backslash G}$ are usually different. Second, even if we limit the discussion to the convention in equation (25), the choice of the asymmetric unit $F_{\Gamma\backslash X}$ is not unique, as is showed by the many different works of M. C. Escher. Third, even when the shape of the asymmetric unit is fixed, it is possible to redefine the whole unit cell, and all asymmetric units that constitute it, by shifting by an arbitrary continuous translation since the choice of origin is not unique. And lastly, for any choice of $F_{\Gamma\backslash X}$, another choice $\gamma \cdot F_{\Gamma\backslash X}$ is equally valid for any $\gamma \in \Gamma$. However, choosing $F_{\Gamma\backslash X}$ to include the origin of X and to be convex reduces the freedom significantly. And in the case when $\Gamma = T$, placing the origin at the center of the unit cell and choosing $F_{T\backslash X}$ as the Wigner–Seitz cell has nice aesthetic properties.

7.2. Viewing $F_{\Gamma\backslash G}$ as $(T\backslash\Gamma)\backslash(\mathbf{F}_{T\backslash G})$

Since $T \triangleleft \Gamma$, it follows that $T\backslash\Gamma = \Gamma/T \cong \mathbb{F}$ is a factor group with elements that are cosets of the form $T\gamma_i \in T\backslash\Gamma$ where $i = 1, \ldots, |\mathbb{F}|$ and $\mathbb{F} = F_{T\backslash\Gamma}$. In the symmorphic case $\mathbb{F} = \mathbb{P}$, the point group of the lattice, and in both the symmorphic and nonsymmorphic cases $|\mathbb{F}| = |\mathbb{P}|$. $\mathbb{F}$ is the group that can be constructed from representatives $\Gamma' \doteq \{\gamma_1, \ldots, \gamma_{|\mathbb{P}|}\}$ with operation such that

$$\gamma_i \, \square \, \gamma_j \doteq \gamma_i \circ \gamma_j \bmod T.$$

This simply means that γ_i and γ_j are multiplied as usual, and then lattice translations are removed from the translational part of the product until the result is within the set Γ'. Each element of this group can be written as $\gamma_i = (R_i, \mathbf{v}_{R_i})$. Then

$$\gamma_i \, \square \, \gamma_j = (R_i R_j, [R_i \mathbf{v}_{R_j} + \mathbf{v}_{R_i}]_T),$$

research papers

where $[\cdot]_T$ brings the translational parts of the product back into the set $\{\mathbf{v}_{R_k}\}$.[2] In the symmorphic case it is always possible to choose $\mathbf{v}_{R_i} = \mathbf{0}$ for all $i = 1, \ldots, |\mathbb{P}|$, otherwise not.

For example, in the case of $P2_12_12_1$, the standard four symmetry elements are $\{e, \gamma_1, \gamma_2, \gamma_3\}$ where their actions on $X = \mathbb{R}^3$ are $e \cdot (x, y, z) = (x, y, z)$; $\gamma_1 \cdot (x, y, z) = (-x + 1/2, -y, z + 1/2)$; $\gamma_2 \cdot (x, y, z) = (-x, y + 1/2, -z + 1/2)$; and $\gamma_3 \cdot (x, y, z) = (x + 1/2, -y + 1/2, -z)$. If we first compute $(\gamma_i \circ \gamma_j) \cdot (x, y, z) = \gamma_i \cdot (\gamma_j \cdot (x, y, z))$ and then 'mod out' translations corresponding to positions that are outside of the unit cell, this defines the operation $\square$. The following group table summarizes the group $(F_{T \backslash \Gamma}, \square) \cong T \backslash \Gamma = P1 \backslash P2_12_12_1$:

$\square$	e	γ_1	γ_2	γ_3
e	e	γ_1	γ_2	γ_3
γ_1	γ_1	e	γ_3	γ_2
γ_2	γ_2	γ_3	e	γ_1
γ_3	γ_3	γ_2	γ_1	e

More generally, the space $T \backslash G$ containing cosets with representatives that can be taken to be pairs of the form $(R, [\mathbf{t}]_T) \in F_{T \backslash G} \cong \mathbb{T}^n \times SO(n)$ is not a group, since T is not normal in G. However, a valid action of $T \backslash \Gamma$ on $F_{T \backslash G}$ can be defined using $F_{T \backslash \Gamma}$ as

$$(R_i, \mathbf{v}_{R_i}) \blacksquare (R, [\mathbf{t}]_T) \doteq (R_i R, [R_i \mathbf{t} + \mathbf{v}_{R_i}]_T) \qquad (26)$$

when the unit cell $F_{T \backslash X}$ is taken to be the Wigner–Seitz cell, since in this case

$$(R_i, \mathbf{v}_{R_i}) \blacksquare \{(R_j, \mathbf{v}_{R_j}) \blacksquare (R, [\mathbf{t}]_T)\} =$$
$$(R_i R_j R, [R_i [R_j [\mathbf{t}]_T + \mathbf{v}_{R_j}]_T + \mathbf{v}_{R_i}]_T)$$

and

$$\{(R_i, \mathbf{v}_{R_i}) \square (R_j, \mathbf{v}_{R_j})\} \blacksquare (R, [\mathbf{t}]_T) =$$
$$(R_i R_j R, [R_i R_j [\mathbf{t}]_T + [R_i \mathbf{v}_{R_j} + \mathbf{v}_{R_i}]_T]_T)$$

and these can be equated using the properties

$$[R_i [\mathbf{t}]_T]_T = [R_i \mathbf{t}]_T \qquad (27)$$

$$[[\mathbf{t}]_T + [\mathbf{t}']_T]_T = [[\mathbf{t}]_T + \mathbf{t}']_T = [\mathbf{t} + \mathbf{t}']_T \qquad (28)$$

where $\mathbf{t}, \mathbf{t}' \in X$ and $[\mathbf{t}]_T \in F_{T \backslash X}$.

Since $(T \backslash \Gamma)$ acts on $F_{T \backslash G}$, it is then possible to write the set of equivalence classes into which $F_{T \backslash G}$ is divided by $(T \backslash \Gamma)$ as $(T \backslash \Gamma) \backslash F_{T \backslash G}$. And as a result,

$$F_{\Gamma \backslash G} \cong (T \backslash \Gamma) \backslash (F_{T \backslash G}). \qquad (29)$$

This statement should not be confused with the third isomorphism theorem from group theory. Rather, it is a statement of the equivalence (as sets) of $F_{\Gamma \backslash G}$ and the equivalence classes of $F_{T \backslash G}$ under the action of $T \backslash \Gamma$ defined in equation (26).

Since $\mathbb{F} \cong T \backslash \Gamma$ and $F_{T \backslash G} \cong F_{T \backslash X} \times \mathcal{R}$, it is possible to write equation (29) as

[2] Here the argument of $[\cdot]_T$ is the translational part of a crystallographic screw symmetry operation, but more generally could include any position or translation in X. In this light, $[\mathbf{x}]$ from the first paper in this series could be called $[\mathbf{x}]_\Gamma$.

$$F_{\Gamma \backslash G} \cong \mathbb{F} \backslash (F_{T \backslash X} \times \mathcal{R}). \qquad (30)$$

Moreover, in the symmorphic case where $\mathbb{F} = \mathbb{P}$ and $\Gamma = T \rtimes \mathbb{P}$, equation (30) can be written as

$$F_{\Gamma \backslash G} \cong F_{T \backslash X} \times F_{\mathbb{P} \backslash \mathcal{R}} \qquad (31)$$

because in this case the action of $\mathbb{P}$ on the Wigner–Seitz cell leaves it invariant, which effectively allows $\mathbb{P}$ to 'pass through' the translations $T \backslash X$.

In analogy with equation (29), if Γ has a subgroup Γ_S such that $\mathbb{S} = F_{T \backslash \Gamma_S}$ is a normal subgroup of the factor group $\mathbb{F} = F_{T \backslash \Gamma}$, then

$$\Gamma_S \backslash \Gamma \cong \mathbb{S} \backslash \mathbb{F}.$$

In particular, if Γ is symmorphic with point group $\mathbb{P}$, and $\mathbb{S} \triangleleft \mathbb{P}$, then it is easy to show that

$$\Gamma_S = T \rtimes \mathbb{S} \triangleleft T \rtimes \mathbb{P} = \Gamma.$$

This means that $\mathbb{S} \backslash \mathbb{P}$ is a finite group, and with an action of this group on $F_{\Gamma_S \backslash G}$ defined,

$$F_{\Gamma \backslash G} \cong (\Gamma_S \backslash \Gamma) \backslash (F_{\Gamma_S \backslash G}) \cong (\mathbb{S} \backslash \mathbb{P}) \backslash (F_{\Gamma_S \backslash G}) \qquad (32)$$

and so

$$F_{\Gamma \backslash G} \cong (\mathbb{S} \backslash \mathbb{P}) \backslash (\mathbb{S} \backslash (F_{T \backslash X} \times \mathcal{R}))$$
$$\cong ((\mathbb{S} \backslash \mathbb{P}) \backslash (F_{T \backslash X})) \times F_{\mathbb{S} \backslash \mathcal{R}} \qquad (33)$$

or

$$F_{\Gamma \backslash G} \cong (\mathbb{S} \backslash \mathbb{P}) \backslash (F_{\Gamma_S \backslash X} \times \mathcal{R})$$
$$\cong F_{\Gamma_S \backslash X} \times F_{(\mathbb{S} \backslash \mathbb{P}) \backslash \mathcal{R}}. \qquad (34)$$

Altogether, this means that $F_{\Gamma \backslash G}$ can be viewed in a variety of ways ranging from the translational part being defined by the asymmetric unit and the rotational part being the whole rotation group, to the other extreme of the translational part being the unit cell and the rotational part being a coset space $\mathbb{P} \backslash SO(n)$. And there are intermediate descriptions in which the translational part is constructed from several asymmetric units and the rotational part is a coset space $\mathbb{S} \backslash SO(n)$ that is larger than $\mathbb{P} \backslash SO(n)$ and smaller than the whole of $SO(n)$. These concepts are illustrated in the following section where $n = 2$ and $\mathbb{P}$ is C_2 or C_4 and the resulting fundamental domains are depicted graphically.

8. Examples: $F_{p1 \backslash SE(2)}$, $F_{p2 \backslash SE(2)}$, $F_{p4 \backslash SE(2)}$

When $G = SE(2)$ the geometric structure of $F_{\Gamma \backslash G}$, which is three dimensional, has been studied in some special cases and lends itself more generally to intuitive understanding. It is instructive to examine this case to develop intuition about the three-dimensional case.

$SE(2)$ can be viewed as the set of all triplets (x, y, θ) where x and y form the Euclidean plane and θ can be taken from the closed interval $[0, 2\pi]$ with the points 0 and 2π identified. Elements of $SE(2)$ are expressed as homogeneous transformation matrices of the form

$$H(g(x, y, \theta)) = \begin{pmatrix} \cos\theta & -\sin\theta & x \\ \sin\theta & \cos\theta & y \\ 0 & 0 & 1 \end{pmatrix},$$

with matrix multiplication serving as the group operation.

Five classes of crystallographic groups are subgroups of SE(2): $p1$, $p2$, $p3$, $p4$ and $p6$. Here $\Gamma = p2$ and $p4$ are used to illustrate $F_{\Gamma\backslash SE(2)}$, and $p1$ was used in the first paper in this series. These groups have elements of the following form:

$$p1 = \{g(m, n, 0) \,|\, m, n \in \mathbb{Z}\},$$
$$p2 = \{g(m, n, 0), g(m+1, n+1, \pi) \,|\, m, n \in \mathbb{Z}\},$$
$$p4 = \{g(m, n, 0), g(m+1, n, \pi/2), g(m+1, n+1, \pi),$$
$$g(m, n+1, 3\pi/2) \,|\, m, n \in \mathbb{Z}\}.$$

The point groups C_2 and C_4, which in the planar case are discrete subgroups of SO(2), are written in terms of elements as

$$C_2 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}$$

and

$$C_4 = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} \right\}.$$

8.1. The case when $\Gamma = p2$

If $\Gamma = p2$, several ways to visualize the fundamental region $F_{\Gamma\backslash G}$ are shown in Figs. 3, 4 and 5.

In Fig. 3 the gluings are defined by the points in the following sets being equivalent: $\{(x, y, 0), (x, y, 2\pi)\}$, $\{(0, y, \theta),$
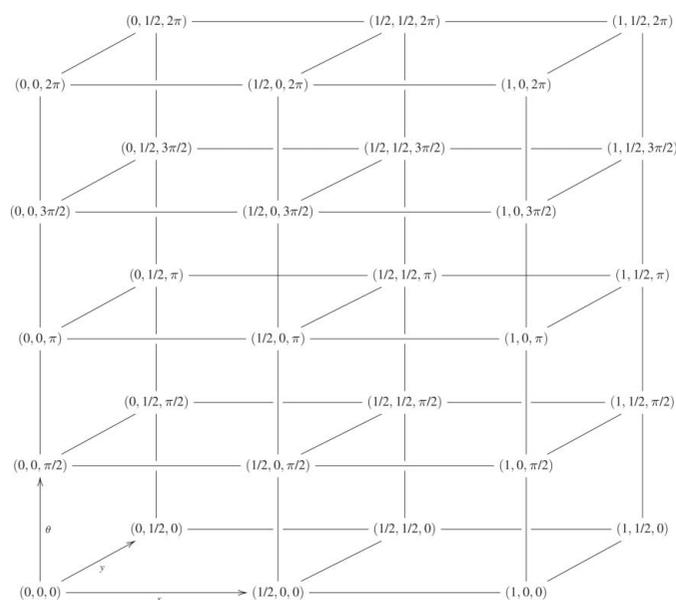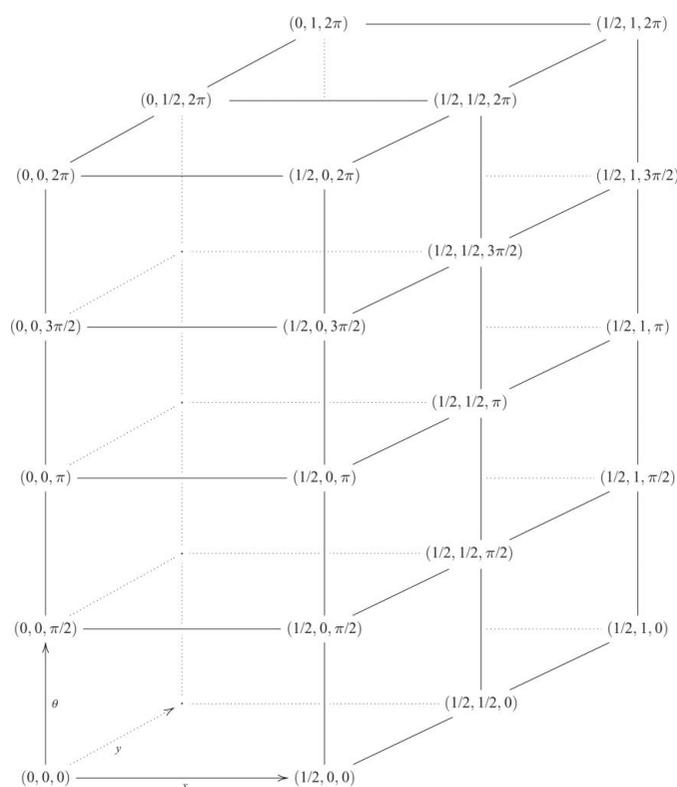

Figure 4
The space $F_{p2\backslash SE(2)}$ identified with the set $F_{p2\backslash\mathbb{R}^2} \times$ SO(2) in a different way than in Fig. 3.

$(1, y, \theta)\}$, $\{(x, 0, \theta), (1 - x, 0, \theta + \pi \bmod 2\pi)\}$, $\{(x, 1/2, \theta), (1 - x, 1/2, \theta + \pi \bmod 2\pi)\}$, where $(x, y, \theta) \in [0, 1] \times [0, 1/2] \times [0, 2\pi]$. Here and in the sequel, the notation '$a + b \bmod 2\pi$' means that the sum $a + b$ is computed and then replaced by the unique number in the range $[0, 2\pi)$ that is congruent to $a + b$ modulo 2π.

In Fig. 4 the equivalent points are $\{(x, y, 0), (x, y, 2\pi)\}$, $\{(x, 0, \theta), (x, 1, \theta)\}$, $\{(0, y, \theta), (0, 1 - y, \theta + \pi \bmod 2\pi)\}$, $\{(1/2, y, \theta), (1/2, 1 - y, \theta + \pi \bmod 2\pi)\}$, where $(x, y, \theta) \in [0, 1/2] \times [0, 1] \times [0, 2\pi]$. Fig. 5 shows other alternative choices for $F_{\Gamma\backslash G}$ in which the sets of equivalent points are $\{(0, y, \theta), (1, y, \theta)\}$, $\{(x, 0, \theta), (x, 1, \theta)\}$ and $\{(x, y, 0), (1 - x, 1 - y, \pi)\}$ for $(x, y, \theta) \in [0, 1] \times [0, 1] \times [0, \pi]$.
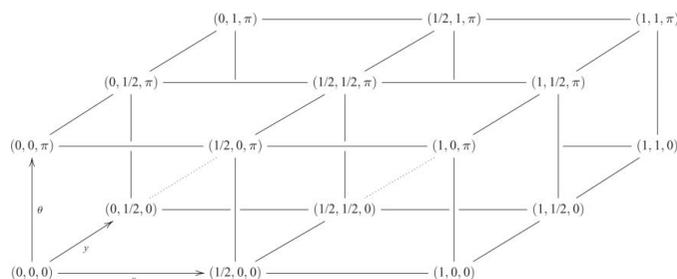

Figure 3
The space $F_{p2\backslash SE(2)}$ identified with the set $F_{p2\backslash\mathbb{R}^2} \times$ SO(2).

Figure 5
The space $F_{p2\backslash SE(2)}$ identified with the set $F_{p1\backslash\mathbb{R}^2} \times F_{C_2\backslash SO(2)}$.

Chirikjian and Yan · Mathematical aspects of molecular replacement. II
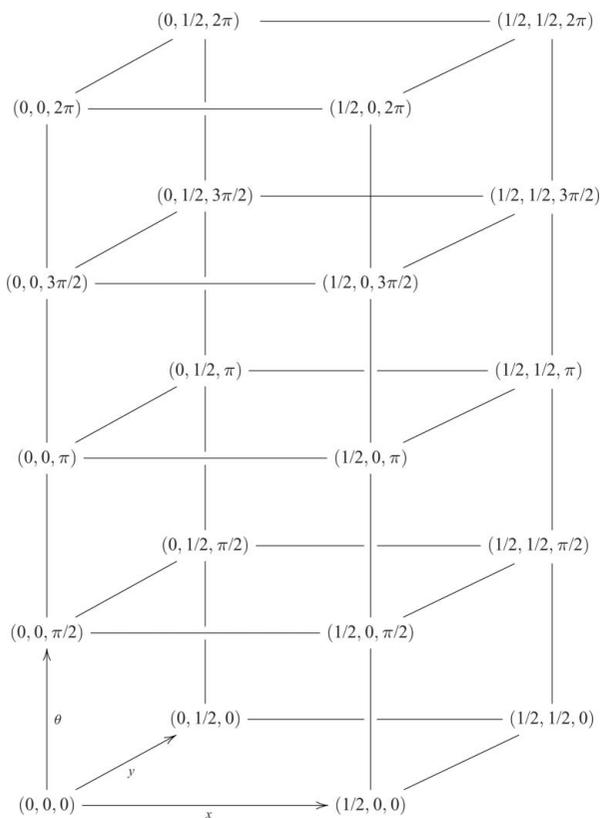
Figure 6
The space $F_{p4\backslash SE(2)}$ identified with the set $F_{p4\backslash \mathbb{R}^2} \times SO(2)$.

8.2. The case when $\Gamma = p4$

If $\Gamma = p4$, there are also several ways to visualize $F_{\Gamma\backslash G}$. One of these is shown in Fig. 6 where the fundamental region corresponding to the coset space $p4\backslash SE(2)$ is identified with the set $(p4\backslash \mathbb{R}^2) \times SO(2)$ with faces, edges and vertices glued in the proper way. In particular, since in $p4$ symmetry four asymmetric units compose the unit cell, each rotated by $\pi/2$ relative to each other around the center of the cell, the gluings of faces are defined by identifying each point in the following sets with each other: $\{(x, y, 0), (x, y, 2\pi)\}$, $\{(x, 0, \theta), (0, x, \theta + \pi/2 \bmod 2\pi)\}$, $\{(x, 1/2, \theta), (1/2, x, \theta + \pi/2 \bmod 2\pi)\}$,
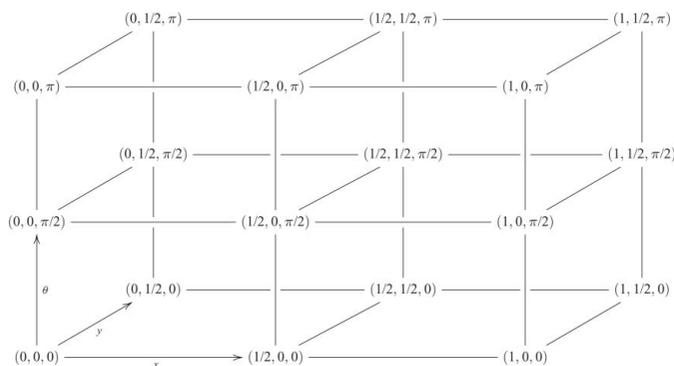
$\{(0, y, \theta), (y, 0, \theta + 3\pi/2 \bmod 2\pi)\}$, $\{(1/2, y, \theta), (y, 1/2, \theta + 3\pi/2 \bmod 2\pi)\}$, where $(x, y, \theta) \in [0, 1/2] \times [0, 1/2] \times [0, 2\pi]$. The gluings in Figs. 7 and 8 can be computed easily as well. In all cases this involves simply applying transformations from $F_{p1\backslash p4}$ to arbitrary points on a face of $F_{p4\backslash SE(2)}$ and observing which points on another face these map to.

9. Applications to MR: efficient sampling strategies

The previous sections of this paper established that when $G = SE(3)$, the fundamental domain $F_{\Gamma\backslash G}$ can be taken to be $F_{\Gamma\backslash\mathbb{R}^3} \times SO(3)$. And if the factor group $\mathbb{F}$ has a subgroup of purely rotational symmetry elements, $\mathbb{S}$, this effectively can be transferred over to the rotation part of $F_{\Gamma\backslash G}$ as $F_{(\mathbb{S}\backslash\Gamma)\backslash X} \times F_{\mathbb{S}\backslash SO(3)}$. When numerical computations are considered, $F_{\Gamma\backslash G}$ is replaced with finite sets of points $\{[g_i]_r\}$. Such sampling can be deterministic or stochastic. Either way, the desire is to sample functions such as $\{C([g_i]_r)\}$ in equation (2) in an efficient and uniform way so as to find the best candidate poses. Optimization over this discrete set is a proxy for optimizing over the original space $F_{\Gamma\backslash G}$. As the sampling becomes finer, the optima that are observed can be expected to be closer to the true optima. For given fixed resolution, we desire to minimize the amount of computing effort by not drawing more points than required. Sampling efficiently on the torus/unit cell is simple because uniform resolution is obtained by discretizing in each coordinate direction independently. In contrast, on the sphere parameterized by the classical polar and azimuthal angles, (θ, φ), or on $SO(3)$ parameterized by Euler angles, (α, β, γ), the discretization is not uniform in the sense that the intrinsically measured distance between nearest neighbors varies widely depending on location in the space. This is inefficient in the sense that achieving a desired resolution at the equator carries the baggage of a large number of points accumulating at the poles. A smart scheme would seek to spread the sample points as evenly as possible under constraints on computing resources. For example, the concentration of points at the poles can be circumvented by sampling $\theta = \cos^{-1}(x)$ where x is uniformly sampled on the interval $[-1, 1]$ rather than sampling θ uniformly on $[0, \pi]$. But the resulting samples will still not be uniformly sampled in the sense that the distribution of distances between nearest points will not be the same at all points. Despite these problems, Euler angles and their variants are the standard in MR.
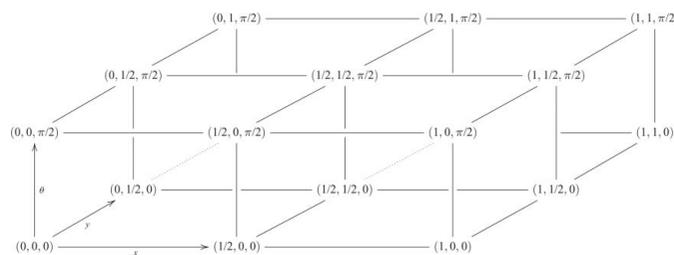

Figure 7
The space $F_{p4\backslash SE(2)}$ identified with the set $F_{p2\backslash\mathbb{R}^2} \times F_{C_2\backslash SO(2)}$.

Figure 8
The space $F_{p4\backslash SE(2)}$ identified with the set $F_{p1\backslash\mathbb{R}^2} \times F_{C_4\backslash SO(2)}$.

9.1. Previous efforts at almost-uniform sampling on SO(3)

The topic of 'uniform' sampling on spheres and rotation groups has been addressed in several different research communities over the past half century. There are two very different versions of this problem: (i) the generation of samples drawn uniformly at random with respect to the natural integration measures on these spaces; (ii) the deterministic sampling of equally spaced samples. It is easy to sample uniformly at random by: either (*a*) sampling from a nonuniform distribution in a parameter space using *a priori* knowledge of the Jacobian to map to uniformly random samples in the space itself; or (*b*) sampling uniformly in an ambient Euclidean space, rejecting samples that do not meet certain conditions (such as being inside a unit ball) and projecting the remaining samples onto the space of interest. [See Avro (1992), Shoemake (1992), for a discussion of such methods.] In contrast, problem (ii) is much harder (and in fact not possible) to solve at arbitrary resolution in the sense of every sample point being surrounded by an identical neighborhood of points that are all equidistant from each other. And so several different variations have been developed.

One area is that of spherical codes and designs (Bannai & Damerell, 1979; Delsarte *et al.*, 1977; Neutsch, 1983, 1996; Sloane *et al.*, 2003) which, for a given number of sample points, is concerned with maximizing the minimal distance between points. A related (though different) problem is that of packing equally sized circles of given size on a sphere of given size (Chirikjian & Stein, 1999; Conway & Sloane, 1999; Clare & Kepert, 1991; Fejes-Toth, 1985; Kottwitz, 1991; Tarrnai, 1984). A third problem is that of sampling on the sphere or rotation group in such a way that integrals of band-limited functions can be expressed exactly as finite sums using quadrature/cubature formulas (Sobolev, 1962; Sobolev & Vaskevich, 1997). Yet another approach is to seek equivolumetric partitioning (Yang & Chen, 2006), but the aspect ratios of such partitions can be very anisotropic.

In crystallography and materials science, symmetries in the rotation function and orientational distribution function have been studied extensively (Heinz & Neumann, 1991; Moss, 1985; Yeates, 1993; Rao *et al.*, 1980). In most works on this subject, the rotation function is viewed as a function of Euler angles, rather than as a function of rotations, which, of course, can be parameterized with Euler angles. Other alternatives that have been examined in this literature are the Lattman angles (Lattman, 1972) and Rodrigues parameters (Neumann, 1991).

Though related in spirit to the goal of this section, none of these capture the concept of uniformity of sampling and simplicity of associated data structures that would be beneficial in MR calculations. The most closely related literature pertaining to deterministically sampling points as equally as possible on spheres and rotation groups is Saff & Kuijlaars (1997), Mitchell (2008) and Yershova *et al.* (2010). While those methods build on sampling methods for spheres and extend them to rotation groups, the method presented here directly addresses sampling on rotation groups. And it does so keeping with the general theme of this series of papers devoted to $F_{\Gamma \backslash G}$. Namely, SO(3) is divided into Voronoi cells [with distance measured using $d_{SO(3)}(\cdot, \cdot)$] centered around elements of a finite group of rotations, $\Pi < SO(3)$, corresponding to rotational symmetry operations for the Platonic solids.

9.2. Voronoi cells in SO(3) and sampling *via* exponential coordinates

In analogy with the way that Wigner–Seitz cells can be used as the fundamental domain for the unit cell in $\mathbb{R}^3$, it is also the case that SO(3) can be divided into cells centered around the elements of $\Pi < SO(3)$. Given a metric such as $d_{SO(3)}(R_1, R_2) = \| \log(R_1^T R_2) \|$, the Voronoi cell centered around the identity consists of all $R \in SO(3)$ such that $d_{SO(3)}(\mathbb{I}, R) \leq d_{SO(3)}(R, A_i)$ where $A_i \neq \mathbb{I} \in \Pi$. It is this Voronoi cell that will be taken as $\overline{F_{\Pi \backslash SO(3)}}$. Since $d_{SO(3)}(R_1, R_2)$ is a bi-invariant metric, it follows that this cell is invariant under conjugation by any element of $\mathbb{P} < \Pi$. And if the Wigner–Seitz cell also is invariant under the action of $\mathbb{P}$, then the combination of these facts means that equation (23) will hold.

In the case when the whole of SO(3) needs to be sampled reasonably uniformly, the properties of the exponential map $\exp : so(3) \rightarrow SO(3)$ can be used. Namely, if $\Omega = -\Omega^T \in so(3)$, the exponential will be that given by equation (12) with $\theta = (\omega_1^2 + \omega_2^2 + \omega_3^2)^{1/2} = \|\omega\|$ and $\Omega = \theta N$. Near the identity, the metric tensor $G(\omega) = J^T(\omega)J(\omega)$ is approximately the identity matrix, where $J(\omega)$ is the same as that defined in equation (15). The determinant of the Jacobian matrix is $|J(\omega)| \simeq 1$ when $\Omega \simeq \mathbb{O}$, and the invariant volume element is $dR = |J(\omega)|d\omega_1 d\omega_2 d\omega_3$. It is only as the distance away from the identity increases that the volumetric distortion effects of

$$|J(\omega)| = 2(1 - \cos\|\omega\|)/\|\omega\|^2$$

cause significant deviation from unity. But our goal is more restricting than that of equivolumetric partitioning. We want $d_{SO(3)}(R_i, R_j) = \delta \ll 1$, a fixed sample distance, if R_i and R_j are any two adjacent rotation samples. A measure of how much the exponential map deviates from this goal is

$$\Delta(\exp(B_{r \leq \pi})) \doteq \int_{\omega \in B_{r \leq \pi}} \| J^T(\omega)J(\omega) - \mathbb{I} \|^2 \cdot |J(\omega)| \, d\omega.$$

Ideally, we would like $\Delta \rightarrow 0$, or what is the same for finely spaced finite samples is that $\sum_{i,j}(d_{SO(3)}(R_i, R_j) - \delta)^2$ is driven to a very small value.

Since contributions to the above integral from around $\omega \simeq \mathbf{0}$ do in fact contribute close to zero to this integral, a better strategy than using the exponential map for the whole solid ball is to divide up SO(3) into Voronoi cells that are related to each other by the action of Π (left and right actions have the same effect). The cell centered on the identity will be taken as $\overline{F_{\Pi \backslash SO(3)}} = \exp(F_{\Pi \backslash B_{r \leq \pi}})$, and the distortion between samples in this cell generated by exponentiating points on a uniform grid in the Voronoi cell $F_{\Pi \backslash B_{r \leq \pi}} \subset B_{r \leq \pi}$ will be very small if Π has many elements. A distortion integral such as the one above can be computed for each cell. The result is that

$$|\Pi| \cdot \Delta(\exp(F_{\Pi \backslash B_{r \leq \pi}})) \ll \Delta(\exp(B_{r \leq \pi})).$$

And if we choose the group Π to be as large as possible (*i.e.* the icosohedral group, Π_{icos}) the Voronoi cell containing the identity will be as small as possible, maximizing the benefits.

Each such Voronoi cell can be sampled almost uniformly by exponentiating points drawn from a Cartesian grid about the origin in so(3) so as to cover $F_{\Pi_{\mathrm{icos}} \backslash \mathrm{SO}(3)}$. Then, by left translation, SO(3) can be tiled with copies of $F_{\Pi_{\mathrm{icos}} \backslash \mathrm{SO}(3)}$. In the case when $\mathbb{S} < \Pi_{\mathrm{icos}} \cap \mathbb{F}$ is not trivial, the same procedure can be used to cover $F_{\mathbb{S} \backslash \mathrm{SO}(3)}$, but with fewer tiles. Or, if one is willing to live with some distortion, $\mathbb{S}$ can be used in place of the icosahedral group and a single tile can be used. Hence, the exponential map together with the decomposition of SO(3) into right coset spaces $\mathbb{S} \backslash \mathrm{SO}(3)$ and the choice of the corresponding fundamental domain with desirable geometric properties provide a means for efficiently sampling both the rotation function, and functions on the motion space $F_{\Gamma \backslash G}$.

Fig. 9 shows the relative size of each cell $F_{\Pi \backslash B_{r \leq \pi}} = \log(\overline{F_{\Pi \backslash \mathrm{SO}(3)}})$ for the three symmetry groups of the Platonic solids. In these figures the yellow faces correspond to the plane of intersection between the identity element and the closest element of Π. The blue planes indicate that a plane between the identity element and next-nearest neighbors in Π clips the cells generated by considering only nearest neighbors. In the icosahedral case, there is no such clipping.

In fact, in all cases the cells $F_{\Pi \backslash B_{r \leq 0}}$ are not exactly polyhedral. This is because for $P_i \in \Pi$ the equation

$$d_{\mathrm{SO}(3)}(\mathbb{I}, \exp \Omega) = d_{\mathrm{SO}(3)}(\exp \Omega, P_i)$$

does not define a plane in the space of values $\boldsymbol{\omega} \in \mathbb{R}^3 \cong \mathrm{so}(3)$. Rather, it defines a surface passing through the point $\log[(P_i)^{1/2}]$ that curves inward toward the identity in comparison to the plane passing through the same point with the normal given by the direction from the origin to the vector corresponding to $\log P_i$. This 'curving-in' effect can be seen in Fig. 10 where (a) and (b) correspond to yellow and blue faces that are, respectively, at distances of $\pi/3$ and $\pi/4$ from the origin, and (c) and (d) which are for faces that are $\pi/4$ and $\pi/3$ from the origin. In Fig. 10(e) the centers of all faces are $\pi/5$ from the origin. Each of these plots starts with the ordinate at these respective values and decreases as the abscissa takes values increasing from zero to the distance of the furthest vertex on the face from its center.

What this means is that the polyhedral cells shown in Fig. 9 are conservative in that they contain even more points than the true Voronoi cells. And hence some additional curvilinear clipping should be applied so as to reduce redundancy in sampling when tiling SO(3) with the true Voronoi cells.

10. Conclusions

The geometric structure of the molecular-replacement problem in macromolecular crystallography has been articulated here. This builds on the algebraic properties of the motion space $(F_{\Gamma \backslash G}, \hat{\circ})$ that were articulated in the first paper in this series, where Γ is the space group of the crystal and G is

the continuous group of rigid-body motions. Equipped with these properties of this space, it becomes possible to formulate codes for searching the space of motions of macromolecules in asymmetric units in a way that is not subject to the arbitrariness of a choice of coordinates such as Euler angles, and the inescapable distortions and singularities that result from coordinate-dependent approaches. Numerical aspects of the
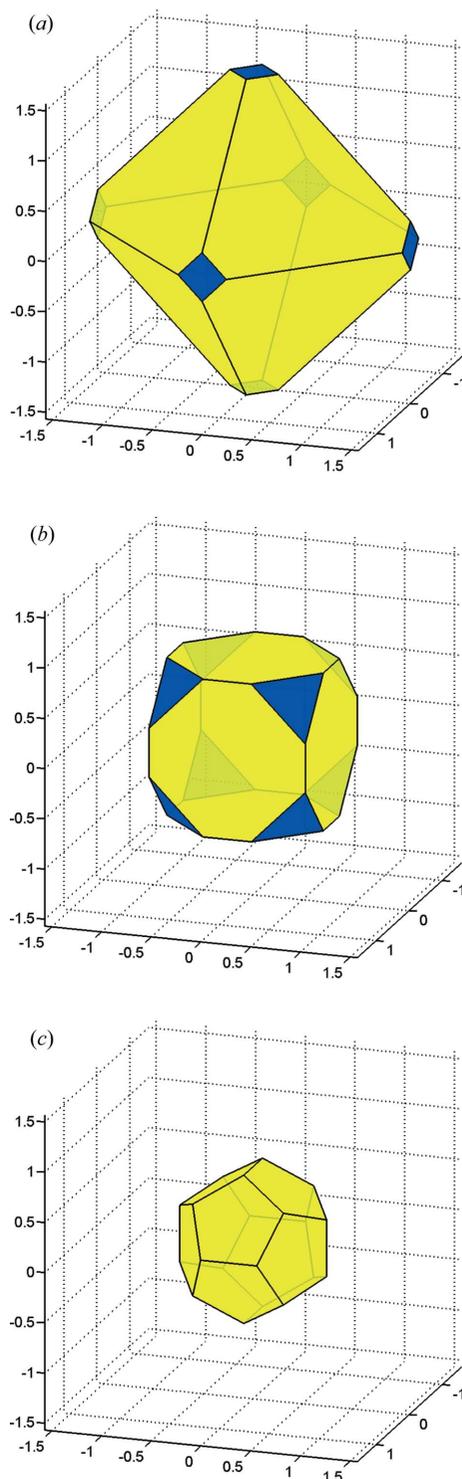

Figure 9
Polyhedra in $\mathbb{R}^3 \cong \mathrm{so}(3)$ depicting Voronoi cells in SO(3) corresponding to the (a) tetrahedral, (b) octahedral and (c) icosahedral groups.
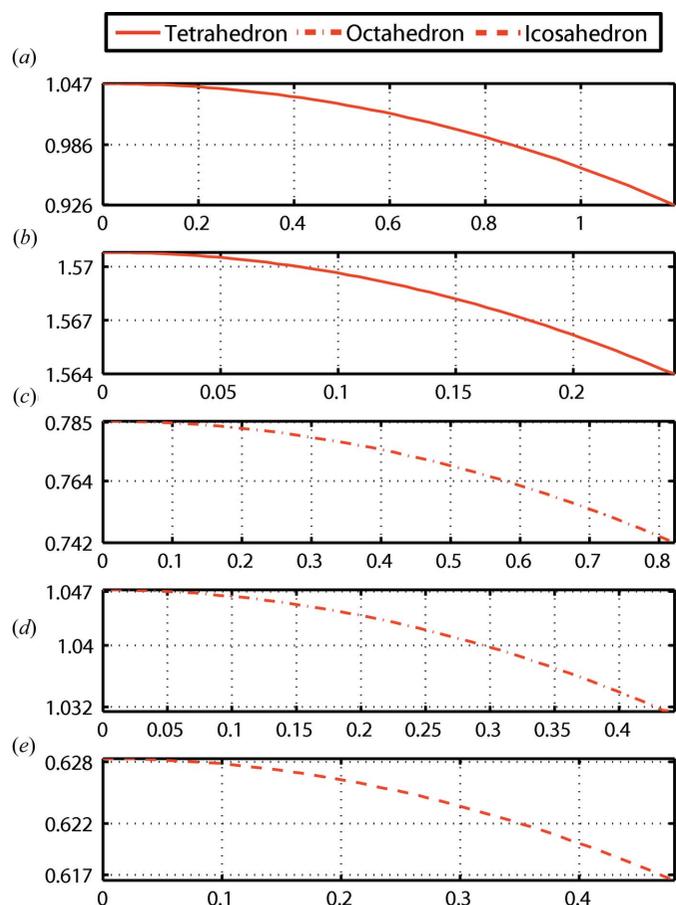
Figure 10
Inward bending of faces for Voronoi cells: (*a*) and (*b*) tetrahedral; (*c*) and (*d*) octahedral; (*e*) icosahedral.

coordinate-free formulation presented here will be investigated in a follow-on paper, as will methods of harmonic (Fourier) analysis on these motion spaces.

References

Aroyo, M. I. *et al.* (2010). *Representations of Crystallographic Space Groups*, Commission on Mathematical and Theoretical Crystallography, Nancy, France, 28 June – 2 July 2010.
Avro, J. (1992). *Graphics Gems III*, edited by D. Kirk, pp. 117–120. San Diego: Academic Press.
Bannai, E. & Damerell, R. (1979). *J. Math. Soc. Jpn*, **31**, 199–207.
Berman, H. M. *et al.* (2002). *Acta Cryst.* D**58**, 899–907.
Bonahon, F. & Siebenmann, L. (1985). *The Classification of Seifert Fibred 3-Orbifolds*. In *Low Dimensional Topology*, edited by R. Fenn. London Mathematical Society Lecture Notes, 95, Cambridge University Press.
Charlap, L. S. (1986). *Bieberbach Groups and Flat Manifolds*. New York: Springer-Verlag.
Chirikjian, G. S. (2011). *Acta Cryst.* A**67**, 435–446.
Chirikjian, G. S. & Stein, D. (1999). *IEEE/ASME Trans. Mechatronics*, **4**, 342–353.
Chirikjian, G. S. & Zhou, S. (1998). *ASME J. Mech. Des.* **120**, 252–261.
Clare, B. W. & Kepert, D. L. (1991). *J. Math. Chem.* **6**, 325–349.
Conway, J. H., Delgado Friedrichs, O., Huson, D. H. & Thurston, W. P. (2001). *Beitr. Algebr. Geom.* **42**, 475–507.
Conway, J. H. & Sloane, N. J. A. (1999). *Sphere Packings, Lattices and Groups*, 3rd ed. New York: Springer.
Delsarte, P., Goethals, J.-M. & Seidel, J. J. (1977). *Geometriae Dedicata*, **6**, 363–388.
Dunbar, W. D. (1981). PhD dissertation, Department of Mathematics, Princeton University, USA.
Farmer, D. W. (1996). *Groups and Symmetry*. Providence: American Mathematical Society.
Fejes-Toth, L. (1985). *Struct. Topol.* **11**, 9–14.
Hahn, Th. (2002). Editor. Brief Teaching Edition of *International Tables for Crystallography*, Vol. A, *Space-Group Symmetry*. Dordrecht: Kluwer.
Heinz, A. & Neumann, P. (1991). *Acta Cryst.* A**47**, 780–789.
Johnson, C. K., Burnett, M. N. & Dunbar, W. D. (1997). *Crystallographic Topology and Its Applications*. In *Crystallographic Computing 7, Macromolecular Crystallographic Data*, edited by P. E. Bourne & K. D. Watenpaugh. Oxford University Press, http://www.ornl.gov/Sci/ortep/topology/preprint.html.
Kottwitz, D. A. (1991). *Acta Cryst.* A**47**, 158–165.
Lattman, E. E. (1972). *Acta Cryst.* B**28**, 1065–1068.
Mitchell, J. C. (2008). *SIAM J. Sci. Comput.* **30**, 525–547.
Montesinos, J. M. (1987). *Classical Tessellations and Three-Manifolds*. Berlin: Springer-Verlag.
Moss, D. S. (1985). *Acta Cryst.* A**41**, 470–475.
Neumann, P. (1991). *Textures Microstruct.* **14–18**, 53–58.
Neutsch, W. (1983). *J. Comput. Phys.* **51**, 313–325.
Neutsch, W. (1996). *Coordinates*. Berlin: de Gruyter.
Nikulin, V. V. & Shafarevich, I. R. (1987). *Geometries and Groups*, translated by M. Reid. New York: Springer.
Park, F. C. (1995). *Trans. ASME J. Mech. Des.* **117**, 48–54.
Rao, S. N., Hih, J.-H. & Hartsuck, J. A. (1980). *Acta Cryst.* A**36**, 878–884.
Rossmann, M. G. (2001). *Acta Cryst.* D**57**, 1360–1366.
Rossmann, M. G. & Blow, D. M. (1962). *Acta Cryst.* **15**, 24–31.
Saff, E. B. & Kuijlaars, A. B. J. (1997). *Math. Intell.* **19**, 5–11.
Satake, I. (1956). *Proc. Natl Acad. Sci. USA*, **42**, 359–363.
Shoemake, K. (1992). *Graphics Gems III*, edited by D. Kirk, pp. 124–132. San Diego: Academic Press.
Sloane, N. J. A., Hardin, R. H. & Cara, P. (2003). *Spherical Designs in Four Dimensions. Proc. 2003 IEEE Inf. Theory Workshop*, pp. 253–258. Piscataway: IEEE.
Sobolev, S. L. (1962). *Dokl. Akad. Nauk SSSR*, **146**, 310–313.
Sobolev, S. L. & Vaskevich, V. L. (1997). *The Theory of Cubature Formulas*. Dordrecht: Kluwer Academic Publishers.
Tarrnai, T. (1984). *Struct. Topol.* **9**, 39–58.
Thurston, W. P. (1997). *Three-Dimensional Geometry and Topology*, edited by S. Levy. Princeton University Press.
Vagin, A. & Teplyakov, A. (2010). *Acta Cryst.* D**66**, 22–25.
Weeks, J. R. (1985). *The Shape of Space*. New York: Marcel Dekker, Inc.
Yang, G. & Chen, I.-M. (2006). *IEEE Trans. Robotics*, **22**, 869–879.
Yeates, T. O. (1993). *Acta Cryst.* A**49**, 138–141.
Yershova, A., Jain, S., LaValle, S. & Mitchell, J. C. (2010). *Int. J. Robotics Res.* **29**, 810–812.