July 13, 2023

Measuring Lattices

Lawrence C. Andrews[a]* and Herbert J. Bernstein[b]

[a]*Ronin Institute, 9515 NE 137th St, Kirkland, WA, 98034-1820 USA, and* [b]*Ronin Institute, c/o NSLS-II, Brookhaven National Laboratory, Upton, NY, 11973-5000 USA. E-mail: lawrence.andrews@ronininstitute.org*

Delaunay; **Delone**; **Selling**; **lattices**; **unit cell**

Abstract

Unit cells are used to represent crystallographic lattices. Calculations measuring the differences between unit cells are used to provide metrics for measuring meaningful distances between three-dimensional crystallographic lattices. This is a surprisingly complex and computationally demanding problem. We present a review of the current best practice using Delaunay-reduced unit cells in the six-dimensional real space of Selling scalar cells $\mathbf{S^6}$ and the equivalent three-dimensional complex space $\mathbf{C^3}$. The process is a simplified version of the process needed when working with the more complex six-dimensional real space of Niggli-reduced unit cells $\mathbf{G^6}$. Obtaining a distance begins with identification of the fundamental region in the space, continues with conversion to primitive cells and reduction, analysis of distances to the boundaries of the fundamental unit, and is completed by a comparison of direct paths to boundary-interrupted paths, looking for a path of minimal length.

1. Forward

“Measuring Lattices” is intended to be a summary of an invited talk for satellite meeting, “Mathematics and Computer Science for Materials Innovation: Crystal Lattice Classifications” of August, 2022, ECM33 meeting (https://www.ecm33.fr/). The intent is describe in brief the basics of implementing a metric space that creates a measure of the “distances” between lattices.

2. History

Human fascination with crystals has a long history. 105,000 years ago, someone had a collection of calcite crystals (Iceland spar) (Wilkins *et* al., 2021). Theophrastus (ca. 372-287 B.C.), a student of Plato and successor to Aristotle, wrote the first known treatise on gems (“On Stones”) (Wikipedia contributors, 2022d).

Figure 1 notes a few key events in cataloging crystal properties. We start with Kepler (1611) (translated in Kepler *et* al. (1966)) and Steno (see Authier (2013)) who conjectured on the structures of crystals. Haüy (1800) created the first catalog of minerals.

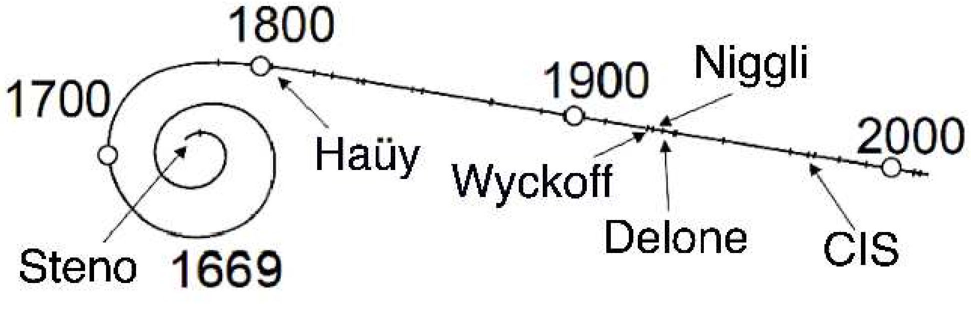


Fig. 1. Some key dates in the history of modern crystallography

With the creation of the 1978 version of the Chemical Information System (CIS) (Bernstein & Andrews, 1979), building on the earlier CIS (Feldmann *et* al., 1972) which lacked cell search capabilities, online computerized retrieval systems become available for searching on ranges of unit cell parameters from the Cambridge Structural Database (CSD) (Kennard *et* al., 1977). From a few tens of thousands in 1979, the community has progressed to collection of several hundreds of thousands of sets of unit cell parameters, with approximately one million in the CSD by 2019 (Taylor & Wood, 2019) alone. When placed in a common SAUC database (see below), the number of cells in the Protein Data Bank and the Crystallographic Open Database total more than 600,000 entries as of summer 2022. Below, we discuss algorithms for robust searching on ranges of unit cell parameters. The need for ranges can be due to uncertainty of measurements, misidentified symmetry classes, variable composition, and searches for related materials.

There have been multiple databases that allow searching on the values of unit cell parameters, including:

- Chemical Information System (CIS) (1978 version), transferred to private management in 1984 (Kadec & Jover, 1985).
- Protein Data Bank, Rutgers Univ. (Bernstein *et* al., 1977) (Berman *et* al., 2000).
- Cambridge Structural Database. The WebCSD (Hayward, 2019) cell search uses the (Hayward, 2019) version of $\mathbf{G^6}$.
- American Mineralogist Crystal Structure Database (Downs & Hall-Wallace, 2003).
- Crystal Lattice Structures. U.S. Naval Research Laboratory, now in the AFLOW library of crytallographic prototypes (Mehl *et* al., 2017).
- Crystallographic Database for Minerals and Their Structural Analogues, Russian Foundation of Basic Research (Chichagov *et* al., 2001).

- ICSD Web: the Inorganic Crystal Structure Database, FIZ Karlsruhe (Rühl, 2019).
- Crystallography Open Database, University of Cambridge (Gražulis *et* al., 2012)
- SAUC (Search of Alternate Unit Cells), (McGill *et* al., 2014)

Various paradigms are used by different databases, sometimes only simple ranges on unit cell edge length. SAUC uses the methods described herein.

3. Introduction

Crystallographers, in general, seem convinced that unit cells are simple; after all, we learn about them on the first day of studying crystallography.

Likewise, lattices are considered to be simple. Of course there are complications, such as space groups, but the concept of the simple lattice is felt to be almost as simple as unit cells.

In crystallography, a unit cell represents an entire lattice, displaying all the symmetries of a crystal except for the translations. The entire lattice is generated by assembling all the translational repetitions of the unit cell.

Stumbling blocks arise when the need to compare various unit cells is included in the tasks. For example, if the variations are due to temperature changes or site substitutions, it is not uncommon for the variations to be largely confined to a single parameter, such as one unit cell axial length.

A different issue arises with the need to compare the unit cells of an **a**rbitrary group of materials. For instance, one might want to search in a database of all known materials. In those cases, more sophisticated tools are needed.

What modern problems require comparing arbitrary groups of cells?

- database searches (approximately a million unit cells are known)
- clustering for serial crystallography (thousands to hundreds of thousands of

diffraction images in a single study)

- finding candidate materials for solving protein crystal structures (Nanao *et* al., 2022)
- searches for minerals or alloys (both of which may have large compositional variation)
- Bravais lattice determination
- Studies in epitaxy (Yang *et* al., 2014)

Crystallographic software delivered with diffractometers, used for processing beamline data, etc. has produced more than one and quarter million structures. The structures have had internally consistent lattice and symmetry identification, but there have been serious omissions in providing globally consistent lattice and symmetry identification because of structures being done in the incorrect space group. Le Trong & Stenkamp (2007) published the case of phospholipase A2, and more were found in McGill *et* al. (2014). Not having failed to be observed is quite different from always working. Two regions of the space of lattice determination come into play here. First, how much error do we want to consider? For protein work, especially for serial crystallography, the errors can be large. Second, the simple methods commonly used in diffractometer software are likely to fail for some of the less common cases: centered lattices commonly cause interesting issues; lattices that are needle- or pancake-shaped are often challenging; the five-fold ambiguity cited by Gruber (1973) also causes confusion. The cited five-fold ambiguity is only for the particular case that Gruber listed, but interchanges of the axes produce other, Buerger-reduced cells.

4. Techniques used to measure lattices

Several techniques other than common linear algebra methods will be used in measuring distances between unit cells.

- Metric vector spaces, see Section 6
- Simple topology, see Section 7
- “Fundamental region” (also called “fundamental unit”), see Section 10.1
- Unit cell reduction, see Section 10.6
- Projectors, see Appendix B

5. Unit cell parameters

If we are to measure the distance between two lattices, we will need to have a space where the measure can be performed. The conventional representation of a unit cell is the “unit cell parameters”, a, b, c, α, β, γ. Mathematically, this is a point in $\mathbf{R^3} \times \mathbf{R^3}$. That is the product of two different spaces; a point in one is not necessarily a point in the other; worse in this case is the fact that a point in the length space means nothing in the angle space. It is still possible to define a distance, but that would require inventing a unit that mixes the units of the length space and the angle space.

6. Spaces

In order to measure a meaningful distance between two objects, the objects need to be placed in a ”metric space”. For crystallographic unit cells, this will often be a vector space amenable to the tools and techniques of linear algebra, at least locally. Because six free parameters are needed to define a unit cell, the space will need to be at least six-dimensional. Depending on other criteria, more dimensions may be needed; gluing the edges of a manifold generates the need for more dimensions. The dimensionality of a unit cell itself should be distinguished from the higher dimensions of the spaces in which it may be useful to embed the cell.

7. A digression on topology

One aspect of topology is the study of objects and the relationships of their edges, which can be "glued" together to make new objects and disclose new relationships among the parts. Here we look at a simple example that demonstrates some of the issues that arise in measuring distances between lattices.

In the following discussion we will follow the common practice of viewing the circle as a two-dimensional object and the torus as a three-dimensional object. Formally, as a manifold, a circle is one-dimensional and a torus is a two-dimensional."

7.1. Create a simple, 1D unit cell

Fig. 2. This is just an empty single unit cell.

7.2. Put two atoms into the unit cell

Here we put atoms at the ends of the cell so that they will be close to their counterparts in the adjacent cell.

Fig. 3. Put an atom at or near each of two boundaries of the unit cell

7.3. Connect the unit cell with its neighbors

Here we see that the “correct” distance between A and B is quite small compared to the distance within the fundamental cell.

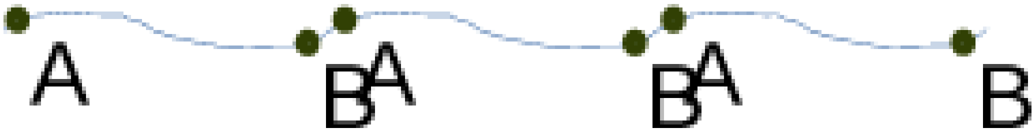


Fig. 4. Here we show just the nearest neighbor cells

7.4. Another choice: wrap around into a loop

A technique based in topology is to wrap one edge of a cell through a higher dimensional space to join another edge of the cell to make a loop. Doing this can make the problem seem simpler, since there is only one cell to be considered. However, the problem of distance calculation has not been simplified. Now the distance has to be measured in both directions around the loop, and the shorter distance reported. This process of joining is termed “gluing”. In addition, the distance must be measured within the curved surface.

Gluing two edges of the one-dimensional cell turns the cell into a circle, making the problem two dimensional. If we had started with a two-dimensional cell, then gluing opposite pairs of edges could be used to make a three-dimensional torus. In either case, the dimensionality of the system would be increased by one.

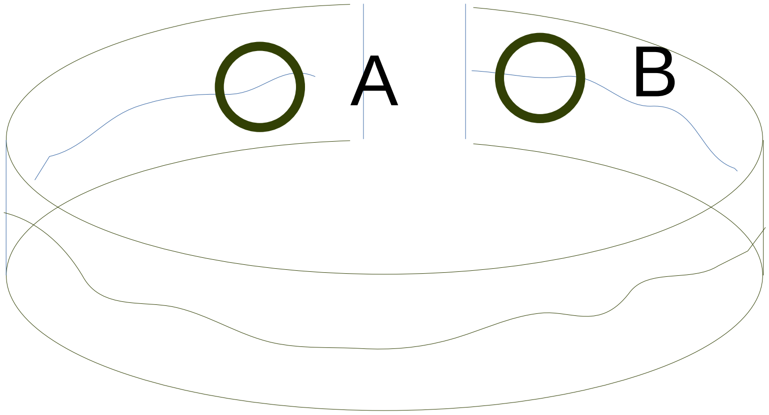


Fig. 5.

7.5. Or you could twist the loop to make a Möbius strip

Consider the above case where the cell was on a 2-D surface. Instead of simply joining the ends, we can twist the loop before we glue the ends together, making a Möbius strip.

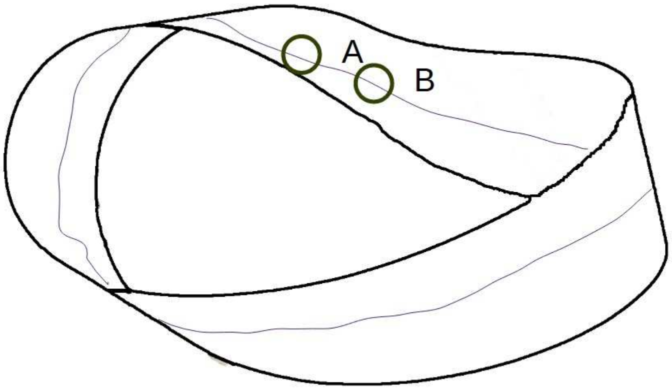


Fig. 6. Twisting before joining makes a Möbius strip

7.6. Summary of the properties on joining

One aspect of gluing is that the dimensionality increases. The issue of measuring distances also become more complex. With added dimensions, the directions in which to look for shortest paths may increase. Similarly the glued boundaries themselves add to the cases to be considered in looking for shortest paths.

8. Spaces, continued

In this paper, we use several spaces to represent unit cells. Other choices of spaces are possible, but this set will suffice for the current topics.

- Lengths in Ångstroms and angles in degrees, a space of $R^3 \times R^3$). Although the presentation of unit cell parameters as edge lengths and interaxial angles is familiar, it does not present an obvious metric (see Section 5). A possible alternative would be to treat each length and angle pair as a point in polar coordinates. This unexplored variation makes use of the association of each axis with its related angle, *e.g.* **a** and α (see appendix D).
- $\mathbf{G^6}$ (Niggli space) (see appendix E)
- $\mathbf{S^6}$ (Delone space) (see appendix F)
- $\mathbf{C^3}$ (see appendix G)

For future reference, we are in the process of extending this list by consideration of a seven-dimensional space of Wigner-Seitz cells (Bernstein *et* al., 2023).

9. Fundamental region

In working with a space in which there are multiple symmetry-related occurrences of items of interest (atoms, molecules, or, in this case, unit cells), it is conventional to choose a fundamental region. “Given a topological space and a group acting on it, the images of a single point under the group action form an orbit of the action. A fundamental domain or fundamental region is a subset of the space which contains exactly one point from each of these orbits." (Wikipedia contributors, 2022a). In a two-dimensional graph, it would often be chosen as the all-plus quadrant. In crystallography, it might be a unit cell, a particular asymmetric unit of a crystal, or even a molecule.

The conventional fundamental unit for Niggli space is the region of $\mathbf{G^6}$ that contains the Niggli-reduced unit cells. This choice has the disadvantage that the fundamental region is non-convex, and the boundaries are complex (Andrews & Bernstein, 2014).

For Delone space, the fundamental unit is conventionally chosen as the all-negative orthant of $\mathbf{S^6}$, which contains all reduced cells. In this orthant, all non-boundary points (that is, those having no zero scalars) correspond to a computable unit cell. All points with one or two zero scalars also correspond to valid unit cells. Boundaries that have four, five, or six zeros have no valid cells. For three zeros, there are valid cells except for those points with a prototype of $s_1, s_2, s_3, 0, 0, 0$ (Andrews *et* al., 2019b).

The all-negative orthant (excluding the boundaries) is an "open" manifold in topological terms; all points there correspond to exactly one lattice. The fact that certain boundaries do not correspond to valid unit cells complicates certain aspects of its topology.

10. Measuring

Measuring the distance between two lattices comes down to finding a geodesic between them, *i.e.* a minimal length path between them obeying the rules for points in the space. We first give a simple example of points in a 2-D manifold, which demonstrates some of the complexity of treating lattices.

10.1. Step 1, fundamental unit

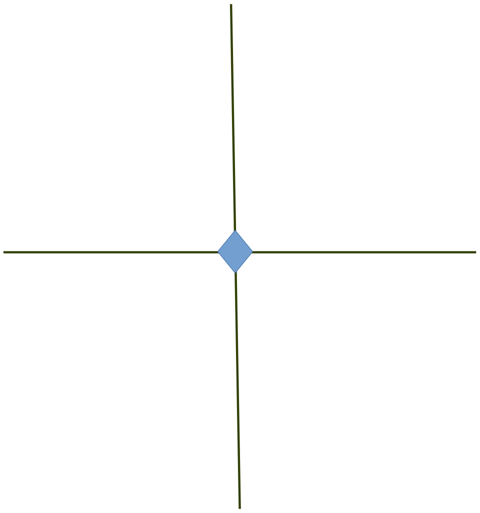
The obvious first step in finding the shortest distance between two lattices is to simply compute the Euclidean distance between them in the fundamental unit. That distance might not be the shortest that can be found, but no geodesic could be longer. It makes a useful starting point. Call this distance d_0.

10.2. Boundaries

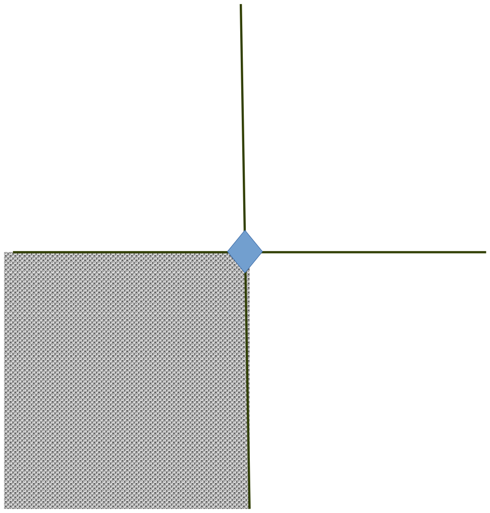
The next step is to consider the boundaries of the fundamental unit. If there is no boundary with a shorter distance to *either* of the lattices than d_0, then d_0 is the required minimum.

If either lattice does have a boundary closer to it than d_0, call the shortest distance from the first lattice to any boundary d_1 and the shortest distance from the second lattice to any boundary d_2. If $d_1 + d_2 > d_0$ then d_0 is still the required minimum and we are done. It is only when both lattices are closer than this to some boundaries that we need to engage in a more complex search for geodesics.

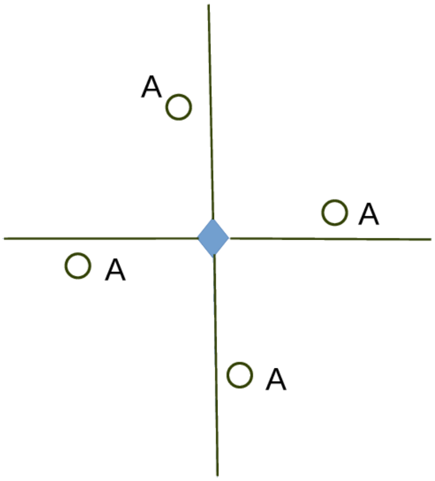
As a simple example of treating boundaries, consider a plane with a single symmetry element, two-dimensional point group 4.



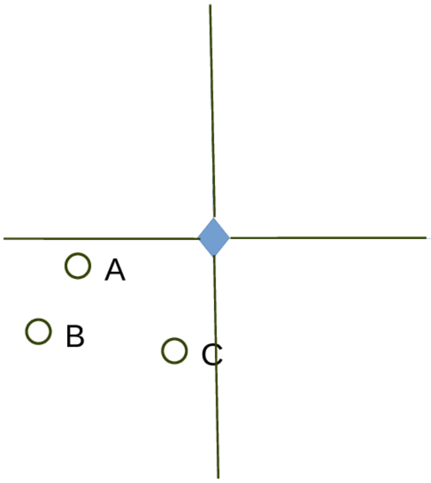
We choose the lower left quadrant to be the fundamental unit. This is an arbitrary choice.



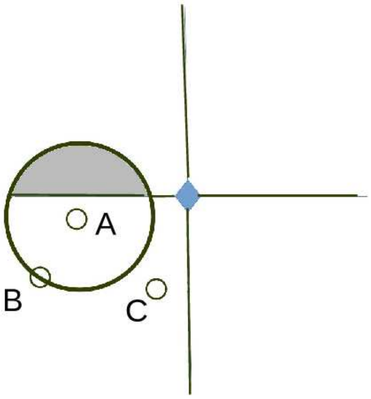
Now put a single atom, A, in the plane, and apply the 4-fold operation.



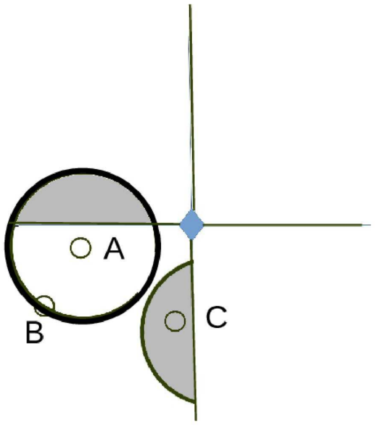
Now let us consider what happens if we place three atoms, A, B, and C in the fundamental unit.



It is clear that within the fundamental unit, ignoring all symmetry transforms and considering B and C, B is the atom closer to A. Draw a circle centered on A that touches B. We see that the circle crosses the boundary and a portion is outside of the fundamental unit. Any atom that is within that region outside the fundamental unit must be examined to see if it contains an atom.



Applying the 4-fold operation to the shaded segment, we see that the atom C (in the fundamental unit) is closer to A than B is to A.



The next part of the problem is how to compute the actual distance from A to C. The obvious answer is to apply the 4-fold symmetry operation either to A or to C and measure the Euclidean distance. We will see in the next section that such simplistic solutions are not sufficient for lattices.

10.3. Spaces for measuring

From this point on, only two spaces will be used for specific examples, $\mathbf{S^6}$, the space of Selling scalars as six real vector components, $(P, Q, R, S, T, U) = (\mathbf{b}\cdot\mathbf{c}$, $\mathbf{a}\cdot\mathbf{c}$, $\mathbf{a}\cdot\mathbf{b}$, $\mathbf{a}\cdot\mathbf{d}$, $\mathbf{b}\cdot\mathbf{d}$, $\mathbf{c}\cdot\mathbf{d})$ where $\mathbf{d} = -\mathbf{a}-\mathbf{b}-\mathbf{c}$, and $\mathbf{C^3}$, the same scalars as $\mathbf{S^6}$, but used as the real and imaginary parts of three complex vector components, $(P+iS, Q+iT, R+iU)$, with each complex number written as a column two-vector. See sections F and G in the appendices. $\mathbf{C^3}$ is simply an alter ego of $\mathbf{S^6}$, but some operations are simpler to describe in $\mathbf{C^3}$. The reason for limiting our consideration to these two spaces is their simplicity compared to other spaces. As stated above, the fundamental unit of $\mathbf{S^6}$ is a simply-connected, convex manifold, with only a few excluded regions (multiple zeros).

For any given lattice there are infinitely many choices of cells from that lattice that can be used to represent it. The point of reduction is to unambiguously choose a single one of those choices of cells. In the absence of experimental error that is what, for example, Niggli reduction does. In the presence of experimental error things are not so simple and any of several equally valid, but very different, unit cells may end up being chosen to represent the same lattice (Gruber, 1973) (McGill *et* al., 2014).

10.4. Reflections, a simple example

Reflections in $\mathbf{S^6}$ were defined in Andrews *et* al. (2019b). They are the operations that create isometric copies of a lattice, always within the fundamental unit. When comparing two lattices, it is necessary to examine the distances from a reduced cell in one lattice to each of the reflections of another reduced cell in the other lattice.

Returning to Figure 10.2 (where the distance to objects outside the fundamental unit were considered), consider instead the two-dimensional point group 4mm. Figure 7 shows a subset of the reflections of the basic set of three points. Immediately, it is clear that in this case, the distance between two copies of point A is the shortest.

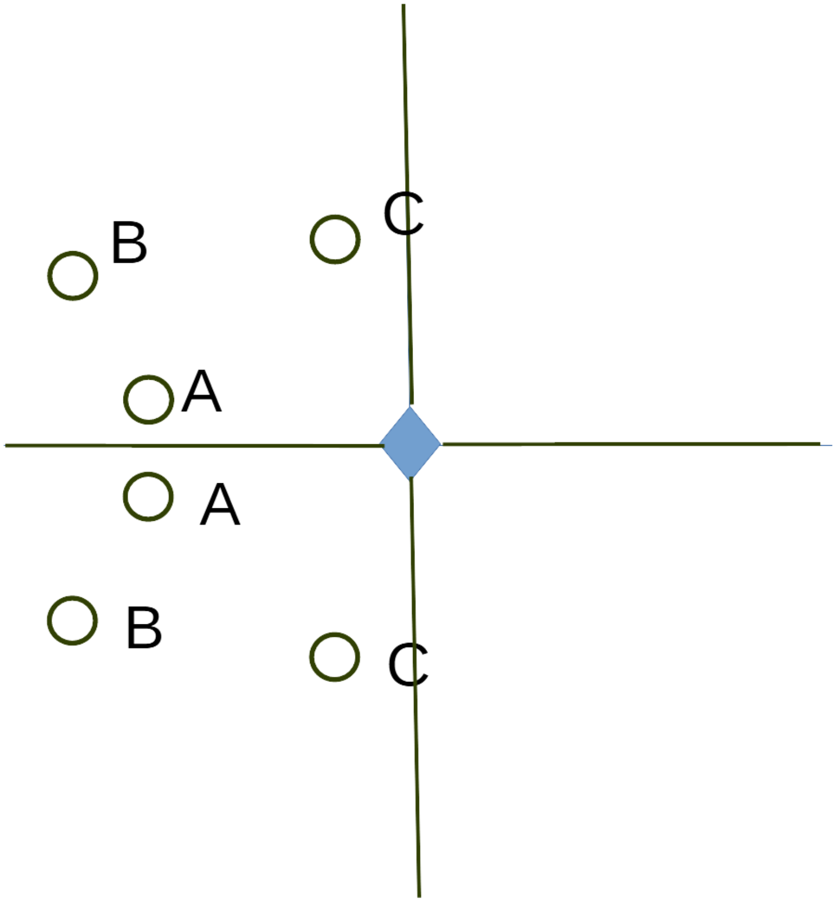


Fig. 7. Recreating the above example, but with point group 4mm.

10.5. Reflections in $\mathbf{S^6}$ *and* $\mathbf{C^3}$

"Reflections" were defined byAndrews *et* al. (2019b) as the 24 operations that create different, reduced versions of a reduced cell in $\mathbf{S^6}$ and $\mathbf{C^3}$ (isometric copies). By this definition, the 4-fold operation and the mirrors in point group 4mm (see Figure 7) are reflections.

An important difference between Niggli space and Delone space is that Niggli space has the vector scalars sorted by somewhat complex rules. $\mathbf{S^6}$ (that is, Delone space) being unsorted results in there being multiple points in the fundamental unit for each lattice. In general, there are 24 reflections, but equal scalars and zero scalars can lead to degeneracies.

The reflections are simpler to describe in $\mathbf{C^3}$, but, of course, the same ones exist in $\mathbf{S^6}$. Here we use the $\mathbf{C^3}$ vectors: c_1, c_2, and c_3, see Appendix G.

First there are the three interchanges of scalars: $c_1 \iff c_2$, $c_1 \iff c_3$, and $c_2 \iff c_3$.

There are also reflections that interchange elements within scalars.

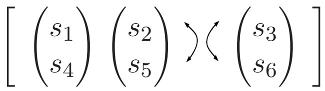


Fig. 8. Example of interchanging real and imaginary scalars for the case of reduction of scalar s_1.

And finally each pair of the reflections needs to be combined giving a total of 24.

10.6. Reduction

The objective of reduction in $\mathbf{S^6}$ (and in $\mathbf{C^3}$) is to make each Selling scalar less than or equal to zero. If any scalar is greater than zero, there are two stages to a reduction step. First, a vector is composed and added to the vector being reduced. Then two scalars are interchanged as shown in Figure 9.

While any scalar can be the one to be reduced, this figure shows the operations for s_1.

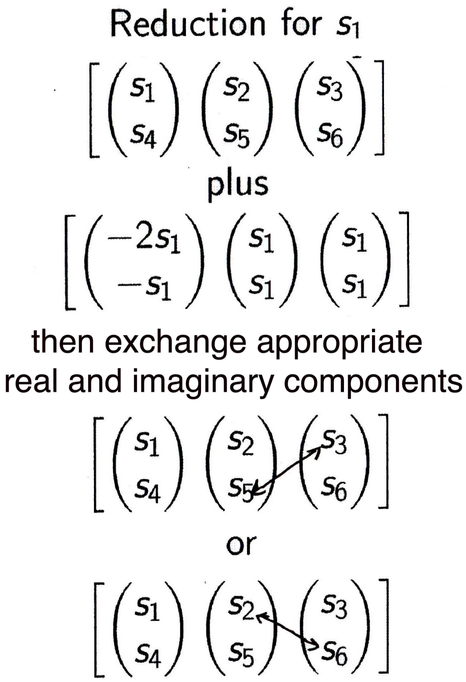

Reduction for s_1
$\left[\begin{pmatrix} s_1 \\ s_4 \end{pmatrix} \begin{pmatrix} s_2 \\ s_5 \end{pmatrix} \begin{pmatrix} s_3 \\ s_6 \end{pmatrix}\right]$
plus
$\left[\begin{pmatrix} -2s_1 \\ -s_1 \end{pmatrix} \begin{pmatrix} s_1 \\ s_1 \end{pmatrix} \begin{pmatrix} s_1 \\ s_1 \end{pmatrix}\right]$
then exchange appropriate
real and imaginary components
$\left[\begin{pmatrix} s_1 \\ s_4 \end{pmatrix} \begin{pmatrix} s_2 \\ s_5 \end{pmatrix} \begin{pmatrix} s_3 \\ s_6 \end{pmatrix}\right]$
or
$\left[\begin{pmatrix} s_1 \\ s_4 \end{pmatrix} \begin{pmatrix} s_2 \\ s_5 \end{pmatrix} \begin{pmatrix} s_3 \\ s_6 \end{pmatrix}\right]$

Fig. 9. Steps for reduction in $\mathbf{C^3}$

10.7. Boundary transformations

The transformations for crossing boundaries of the fundamental unit are the reduction operations.

The section on fundamental regions (see Section 9) shows the problem of points outside the fundamental unit and how boundary transformations were applied there. The boundary transformations for lattices are more complex than those of space groups.

1. The boundaries must be identified.

2. The transformations at the boundaries in $\mathbf{G^6}$ and $\mathbf{S^6}$ are not isometric, meaning that they do not preserve distance measure.

3. For a particular boundary, there may be more than one allowed transformation available.

4. Applying a transformation to a point near a boundary may generate a new point that is near another boundary, requiring further searching.

5. A point that is in the fundamental unit but close to a boundary might need to be transformed in order to find other nearby points.

6. All measurements must be done in the metric of the fundamental unit. A corollary is that when two points are being treated, one must be held fixed and within the fundamental unit while the possible boundary transformations and reflections of the other are treated.

7. If a point is near more than one boundary, then there are multiple transformations and reflections to consider.

The fundamental unit in $\mathbf{S^6}$ is the all-negative orthant. The boundaries are all where one (or more) of the scalars equals zero. Two points that are close together and have a corresponding scalar that is positive in one and negative in the other may have representations in the fundamental unit that are far apart.

In Figure 10a, the two points with $\gamma = 89.8^\circ$ and $\gamma = 90^\circ$ are far apart.

Figure 10b demonstrates a further complexity. Because the angles α and β were chosen to be 90°, all points in Figure 10a sit on two boundaries ($beta$=90° and γ=90°, s_4=0 and s_5=0). Perturbing those boundary points may cause them to be within the fundamental unit, but, on the other hand, may cause them to be outside and thus be far from the neighboring points (after reduction). All the gray lines correspond to members of lattice pairs that are close together but are far apart in the fundamental unit (that is, Delone-reduced).

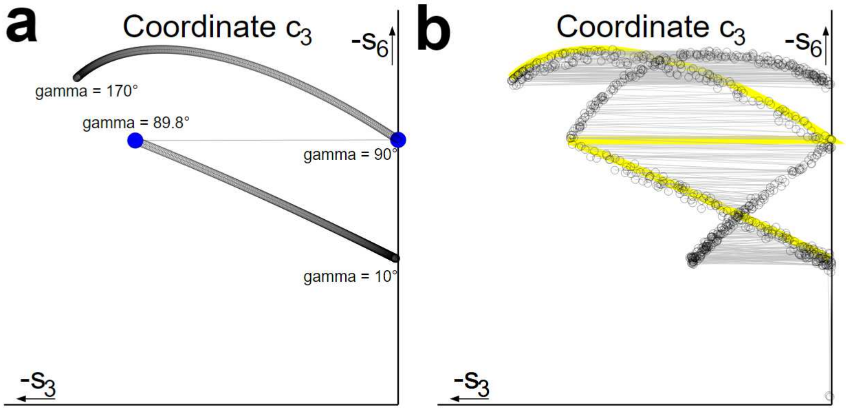


Fig. 10.
In both figures (a) and (b), a series of points has been generated in $\mathbf{S^6}$. The basic cell is a=10, b=10, c=10, $\alpha = 90°$, $\beta = 90°$, and γ ranges sequentially from 10° to 170°, in 0.2° increments. Each cell was Delone-reduced. Each was then placed in the $\mathbf{C^3}$ asymmetric unit (Andrews et al., 2019b). Then, in sequence from the first cell, for the succeeding cell, all 24 reflections of the succeeding cell were examined, and the reflection that is closest to the first cell chosen. Calculation then proceeds to the next pair (where the new first is the second chosen in the previous step). A line (shown in gray) was drawn between each successive pair of points. The individual points are shown as circles.

This figure shows the points converted to $\mathbf{C^3}$, and the c3 coordinate is displayed. In figure (a) on the left the points are generated as described. The gray connection lines, except for the one connection $\gamma = 90°$ to $\gamma = 89.8°$, are hidden by the circles. That single long line represents a case when distance measured in the fundamental unit is quite long, although the originally generated points are quite close together. In figure (b) on the right, initially, three copies of each point were generated, and all points were perturbed by 1% in a random direction orthogonal to the generated vector (point). Then (in the order of generation), the above sequencing was performed.

Because the basic structure of the point generation requires that all points be on multiple boundaries, there are many cases where the distances in the fundamental unit are long. Every gray line is a case where the simple Euclidean distance in the fundamental unit is long, but the generated points are close together.

As seen in Figure 10, points that are a small distance apart may be far apart in the fundamental unit. As an example, we choose a primitive cubic unit cell, $\mathbf{S^6}$ vector $\left[0, 0, 0, 10, 10, 10\right]$. In Figure 11, 1000 copies have all been perturbed by 1% in a direction orthogonal to that vector.

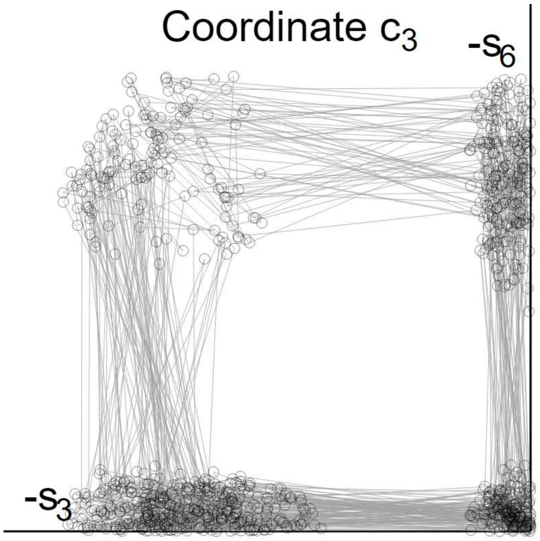


Fig. 11. 1000 copies of $\mathbf{S^6}$ vector $\left[0, 0, 0, 10, 10, 10\right]$ were perturbed by 1% orthogonal to the vector. As in Figure 10, for each pair of points, the second point was chosen to be the reflection that is closest to the first point. The choice of reflection pairs was done by treating the list of generated points as a sequence. Long lines indicate pairs that are far apart within the fundamental unit.

11. Boundaries (and reduction) in $\mathbf{S^6}$ and $\mathbf{C^3}$

11.1. Multiple boundaries

If a point is near two boundaries, all the cases for each boundary need to be considered (recall that in $\mathbf{S^6}$ there are four possible boundary transformations at each of the six boundaries). In addition, applying the transformations of one of the boundaries and then those of the second is necessary. That means that there are 16 possible results. Further still, the case of applying them in the other order must be considered (resulting in another 16 cases).

For the case of being near three boundaries, similar complexity to the case for two boundaries must be considered: all single boundaries, all boundary pairs (in both orders) and all triples (in all possible orders).

11.2. Repeated boundaries

In some cases, applying a boundary transformation results in one of the modified scalars being near a boundary that it was not near before. Such a case requires that the new neighbor boundary also be considered.

11.3. Iteration

Much of the complexity of determining the shortest distance between two lattices is due to the need to repeat steps of reduction and reflection. Among the places in the process where this becomes necessary is when a reduction step creates a new point that is near a boundary.

11.4. Measuring lattice distances in $\mathbf{S^6}$

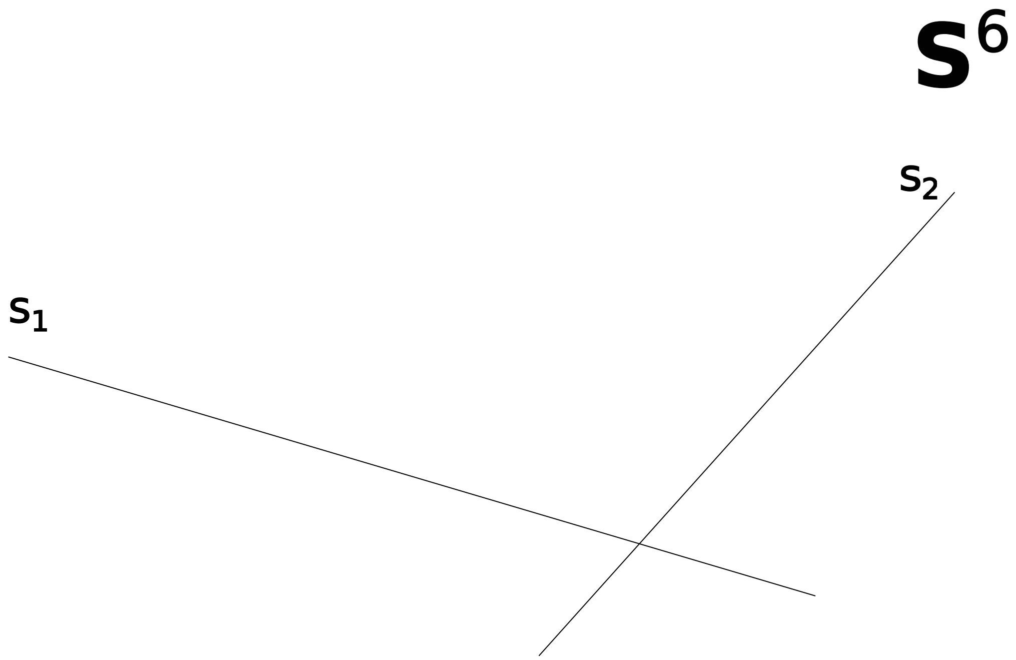


Fig. 12. In space $\mathbf{S^6}$, choose two axes. The operations shown will be a cartoon of the operations in six-space.

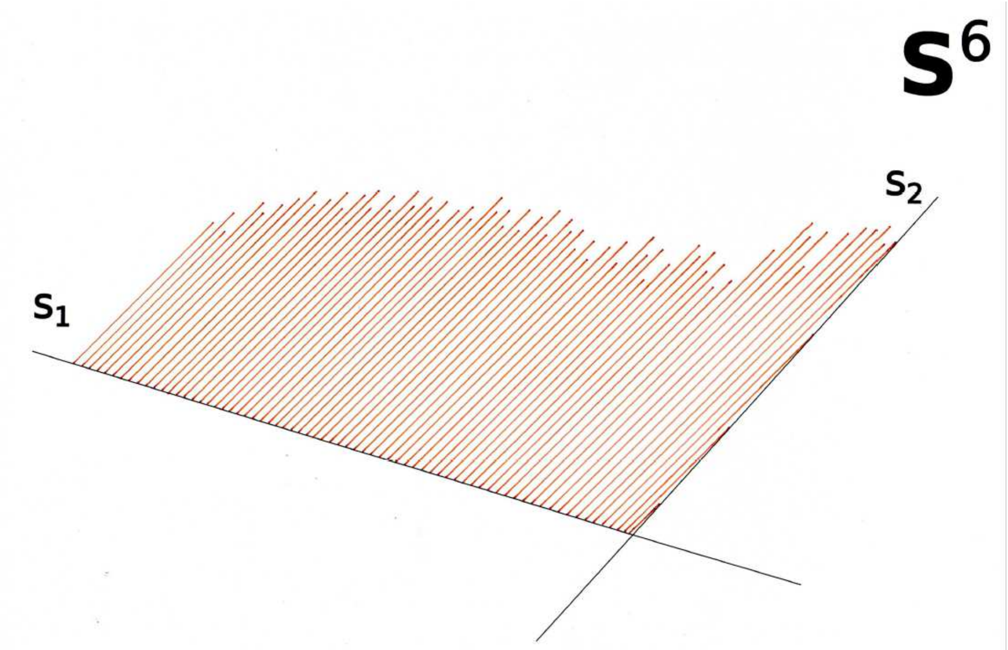


Fig. 13. Define a fundamental unit (all non-positive region).

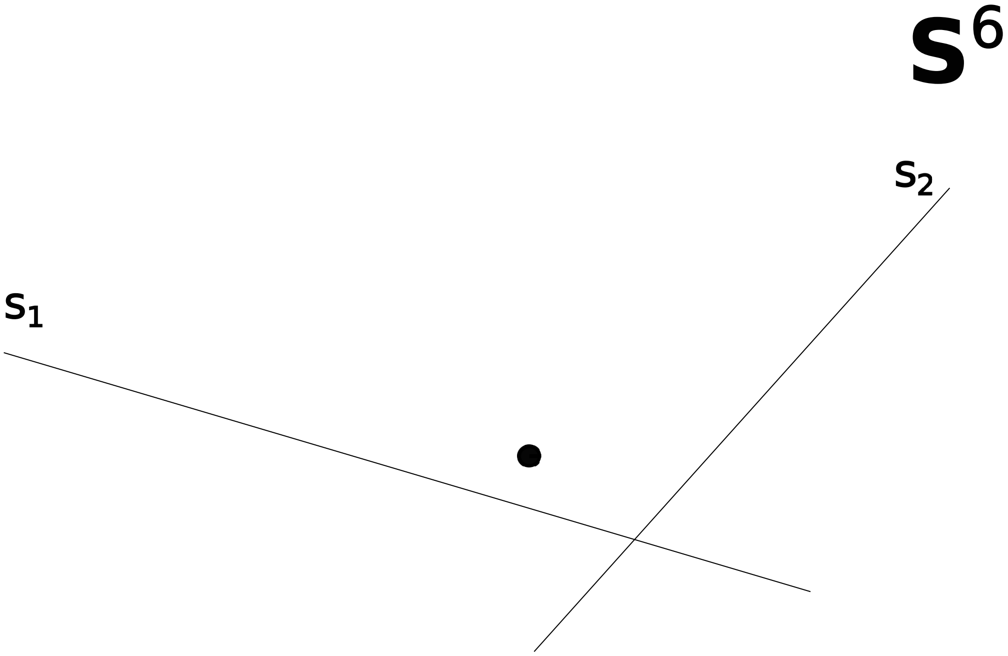


Fig. 14. Put in a unit cell (a point in $\mathbf{S^6}$).

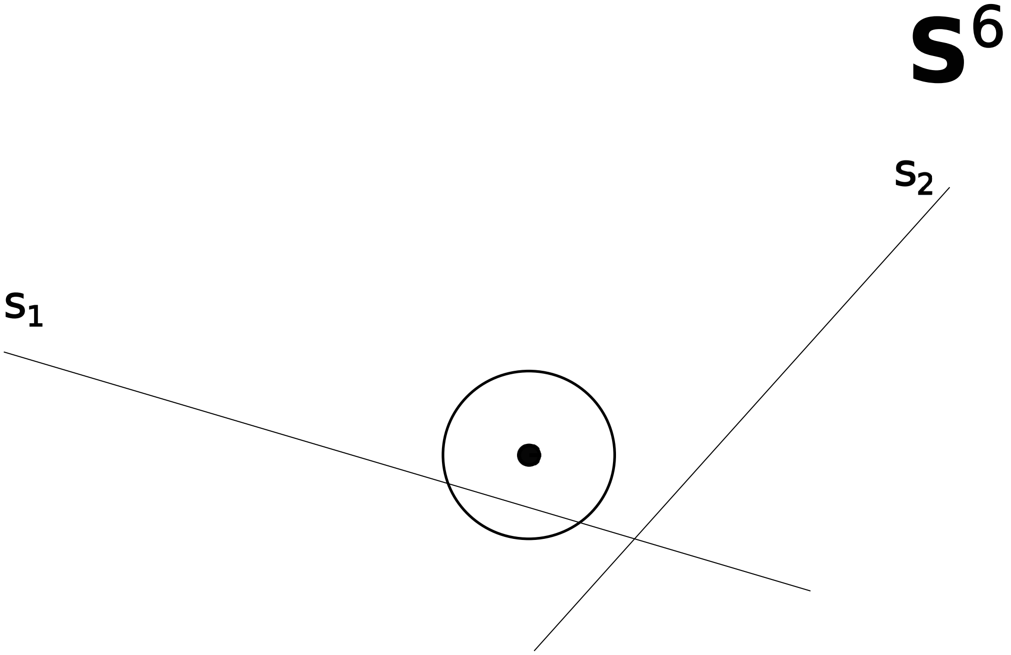


Fig. 15. Define a search radius. In this case, it crosses a boundary and a region is outside the fundamental unit. Cells there are *not* reduced.

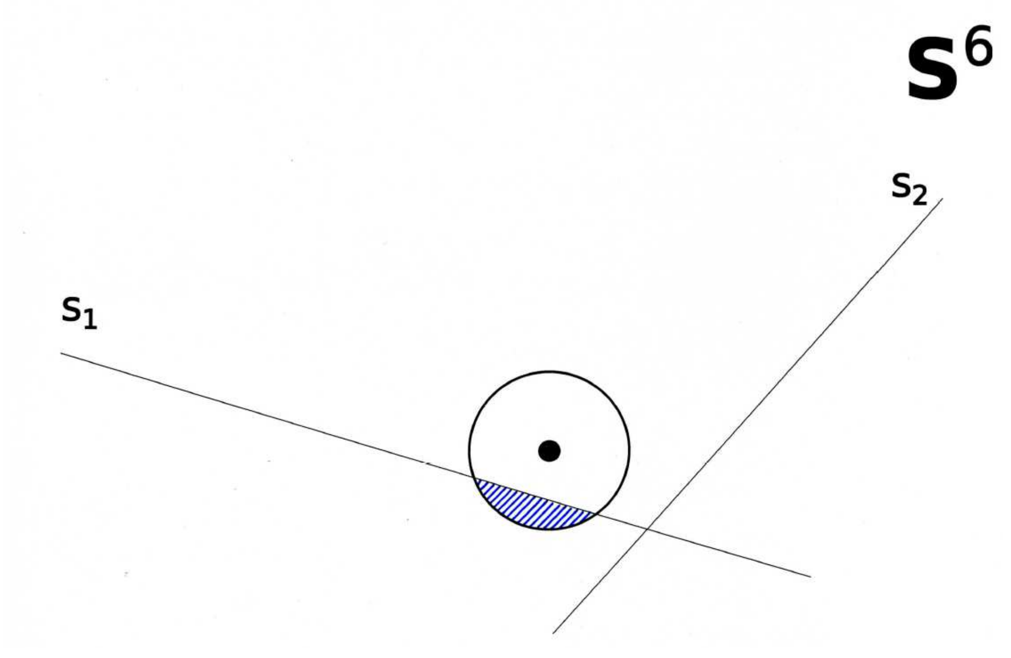


Fig. 16. The non-reduced region is marked in crosshatch.

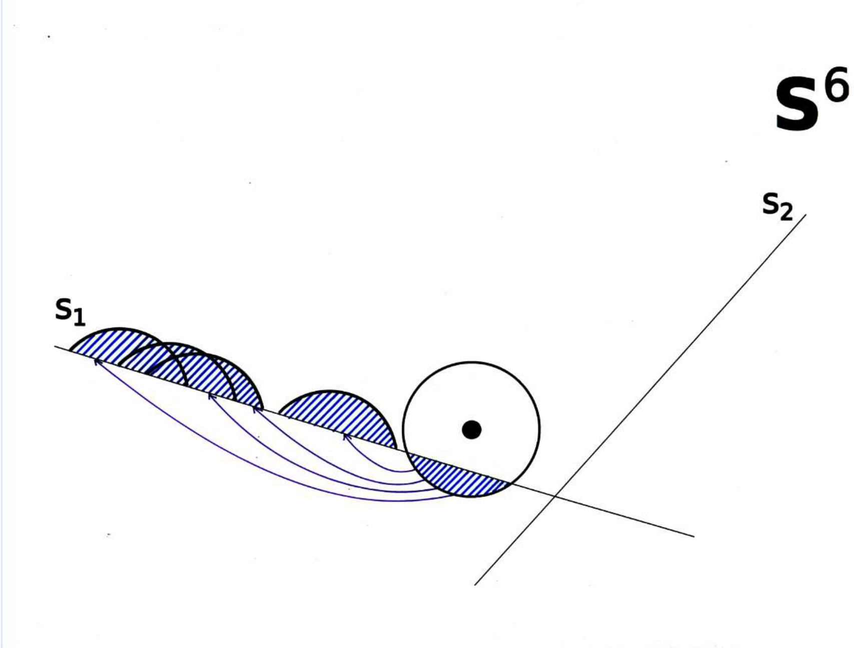


Fig. 17. Unlike in the 4-fold example, there are multiple (up to 4) places to go. A further complication is that the transformations at the boundaries do **not** preserve distances. Distances must be measured in the frame of the fundamental unit.

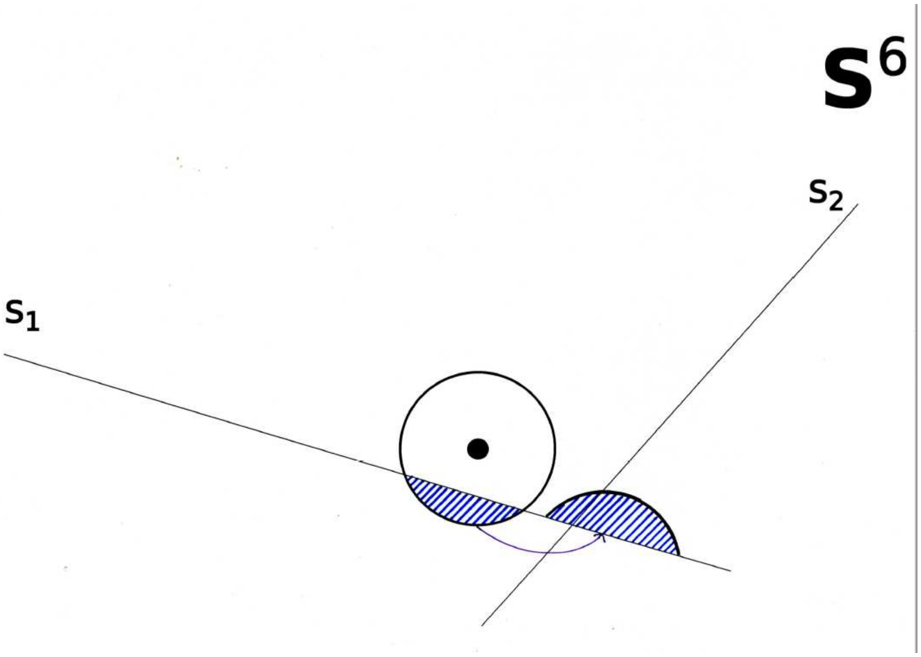


Fig. 18. Sometimes, in the process, another boundary is crossed. Then further iterations may be required.

The distance between two lattices is determined by first computing d_0, the minimum of all the Euclidean distances between reflections of those two points within the fundamental unit. If either lattice is closer to any boundary than d_0, then we need to iterate through all possible paths that may cross a boundary.

If the second point is near a boundary, then the four transforms for that boundary are applied. We need to compute the tunneled distances. If the generated points are not near another boundary, then we are done. If they are near a new boundary, then we need to iteratively apply the boundary transforms.

1. Reduce both unit cells.

2. Choose one point ("first point") to be the fixed point, and the other ("second point") to be the one to which transformations are applied.

3. Compute d_0, the minimum of all the Euclidean distances between the first point and all reflections of the second point (the reflected points are all within the fundamental unit).

4. Compute d_1, the minimum of all the Euclidean distances between the first point and all boundaries.

5. Compute d_2, the minimum of all the Euclidean distances between the second point and all boundaries.

6. If $d_1 + d_2 < d_0$, then there may be a path between some reflection of the first point and some reflection of the second point that is shorter than d_0 that crosses a boundary. It is sufficient to explore possibly shorter paths by leaving the first point fixed and only considering reflections of the second point. For each boundary to be considered there are at most four possible paths from whichever of the two points is close enough to that boundary to satisfy $d_1 + d_2 < d_0$. For

such a case we need to iterate through all four possible paths that may cross the boundary.

7. If the second point is near two or three boundaries, then each boundary needs to be treated by itself. Then all pairs (and all triples) in all permuted orders must be transformed through the boundaries by all four paths of each boundary. For a given $\mathbf{S^6}$ boundary there are four possible boundary transforms to consider which may take a reflection of the second point outside of the fundamental unit (see Figure 17).

8. If the points produced in the above steps are near new boundaries, then the above steps need to be iterated using the new points.

9. N.B., the distance must be measured in the metric of the fundamental unit; see section 11.5.

In the above discussion of gluing in topology, mention was made of the increase in dimensions as edges are glued and more if there are twists. Here, the boundary transformation operations interchange two scalars, and that is a twist operation. Counting up the gluings and twists leads to the conclusion that the required dimension of the space in which to embed the glued manifold is quite high; immediately, we understand that there are potentially going to be many paths to examine.

11.5. Measuring with the metric of the fundamental unit

The problem that arises in computing distances between lattices is that it is necessary to transform points to be outside the fundamental unit. The boundary transformations are not isometric; distance measure is not preserved on crossing out of the fundamental unit. Here we describe how to create a measure for those points that is isometric.

If the second point of a measurement is near a boundary, then project the point onto the boundary (the dashed arrow in Figure 19) and extend the line segment from the second point to the boundary to twice the length creating a transformed point creating an artificial mirror point. The "Tunneled Mirrored Boundaries" distance is found by locating the crossing point on the boundary where the line from the first point to the artificial mirror point crosses the boundary and then finding the four boundary transforms of the crossing point. Applying a boundary transform to the crossing point provides an alternate boundary point to be used for a broken path (first point to crossing point point – equivalent to boundary-mapped crossing point, followed by mapped crossing point to second point). This broken path is sometimes shorter than the direct d_0 path and is entirely in the fundamental unit. If two boundaries are close, then similar actions on both lattice points are needed.

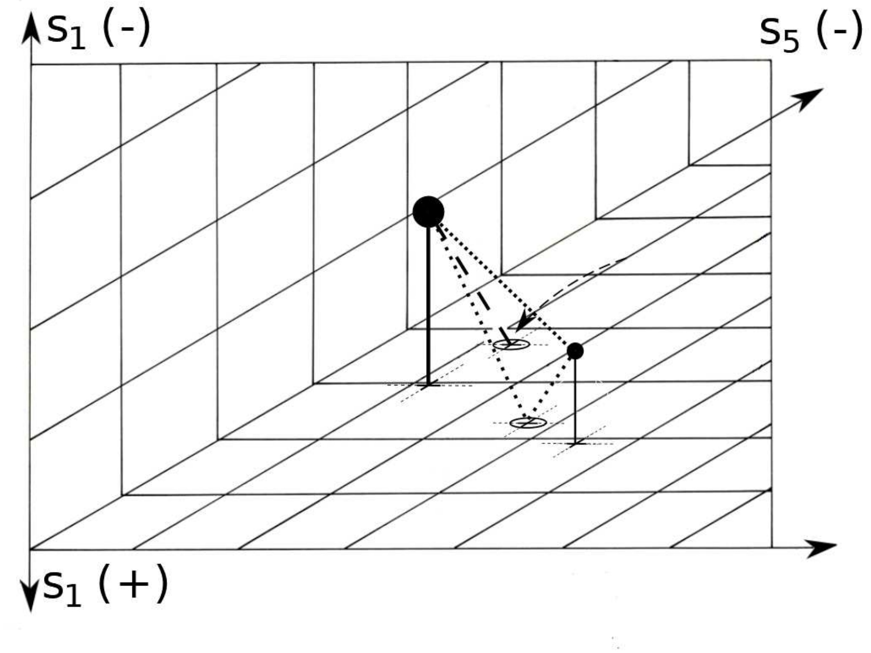


Fig. 19. Tunneled Mirrored Boundary

11.6. Algorithm

For finding the shortest distance between two reduced lattices, the first step is to compute the simple Euclidean distance between them. Each lattice is considered as a point in a space of six or more dimensions with a metric derived from the cell reduction rules for that space.

If the original second point is near one (or more) boundaries, then the the boundary searches (above) are performed.

If new near-boundary-contacts are detected, then the same kind of searches is performed.

Distances must be measured in the metric of the fundamental unit *not in the metric* *of any region outside the fundamental unit.*

11.7. Nearest neighbors, the "Post Office Problem"

Up to this point, the discussion has targeted finding the shortest distance between two lattices. Sometimes, the problem is to find the collection of points that are near a particular probe. One important use is for creating and for searching within databases of crystals. Known as the nearest neighbor problem, it is also known as the post-office problem (Andrews, 2001), and (Andrews & Bernstein, 2016). (see Appendix C.)

12. Determining Bravais lattice types

12.1. Notation

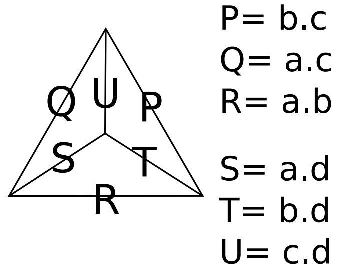


Fig. 20. Bravais organized the scalars into a "tetrahedron"(Delaunay, 1932).

The International Tables for Crystallography define this assignment of the scalars (PQRSTU) in the tetrahedron (Donnay & Donnay, 1953). Other authors used alternative assignments.

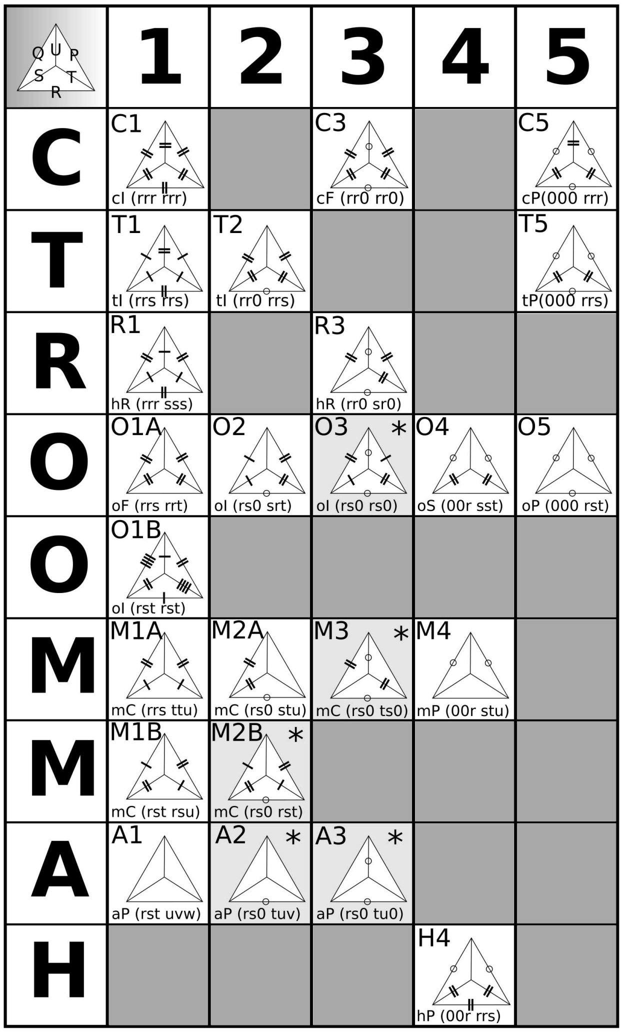
As $\mathbf{S^6}$: $\mathbf{s} = \Big(P, Q, R, S, T, U\Big)$
$= \Big(s_1, s_2, s_3, s_4, s_5, s_6\Big)$.

As $\mathbf{C^3}$: $\left[\begin{pmatrix} P \\ S \end{pmatrix} \begin{pmatrix} Q \\ T \end{pmatrix} \begin{pmatrix} R \\ U \end{pmatrix}\right]$

12.2. Bravais lattice types in $\mathbf{S^6}$

Delaunay (1932) and Delone *et* al. (1975) produced a table classifying the reduced forms of the Bravais types as 24 cases. Figure 21 shows the same classification (the letter codes for crystal systems have been changed to modern notation. For some types, revised names have been given to Delone's types where a cell included two types (for example, Delone's "M1" is split into "M1A" and "M1B"). In Figure 21, the Bravais type is listed, and the restrictions on $\mathbf{S^6}$ scalars is shown graphically as the Bravais tetrahedron and listed as the lattice character.

Figure 22 describes the contents of each table entry in Figure 21.



	1	2	3	4	5
C	C1 cI (rrr rrr)		C3 cF (rr0 rr0)		C5 cP(000 rrr)
T	T1 tI (rrs rrs)	T2 tI (rr0 rrs)			T5 tP(000 rrs)
R	R1 hR (rrr sss)		R3 hR (rr0 sr0)		
O	O1A oF (rrs rrt)	O2 oI (rs0 srt)	O3 * oI (rs0 rs0)	O4 oS (00r sst)	O5 oP (000 rst)
O	O1B oI (rst rst)				
M	M1A mC (rrs ttu)	M2A mC (rs0 stu)	M3 * mC (rs0 ts0)	M4 mP (00r stu)	
M	M1B mC (rst rsu)	M2B * mC (rs0 rst)			
A	A1 aP (rst uvw)	A2 * aP (rs0 tuv)	A3 * aP (rs0 tu0)		
H				H4 hP (00r rrs)	

* The right angles have no relationship to symmetry.

Fig. 21. The table of Delaunay (1932) describing the 24 Bravais types in $\mathbf{S^6}$. It has been redone removing the images of the Dirichlet cells, the non-reduced cells, and adding the “lattice character”, which describes the linear manifold of each type. The crystal family types have been renamed to modern usage: Q changed to T for tetrahedral, K changed to C for cubic, and T changed to A for anorthic. Where Delone in some places included two types in one table cell, they have been split into two (for example: “M1” is split into “M1A” and “M1B”. Note that five types (O3, M3, M2B, A2, and A3) are not normal crystallographic types. They are boundary types, and they have fewer free parameters than the generic type requires. For instance, O3 (character: rs0 rs0) has only two free parameters (r and s), whereas an ordinary orthorhombic type requires three variables.

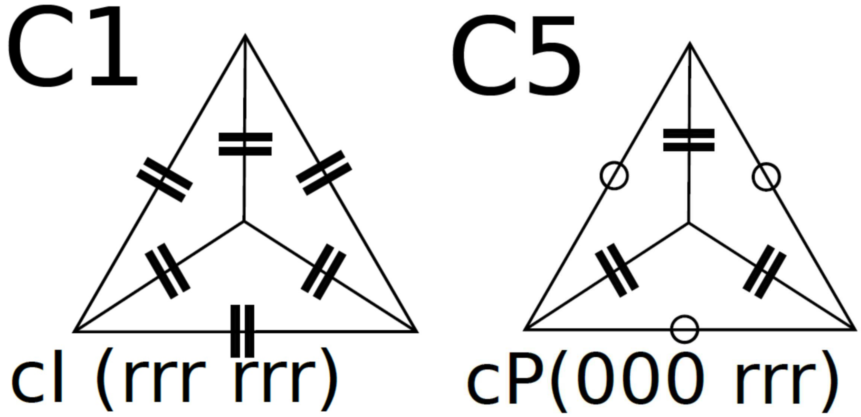


Fig. 22. Explaining the information in Figure 21; it includes the name of the Delone type (*e.g.*: C1),the Bravais type (*e.g.*: cI for body-centered cubic), and the lattice character that describes the linear manifold of each Delone type (*e.g.*: 000 rrr). The markings on the tetrahedra describe the restriction on the $\mathbf{S^6}$ scalars; they encode the same information that is in the lattice character. A zero (circle) means that the value of that scalar is zero, several with the same lines means they are constrained to be equal; in Figure 22, they are a pair of double line hatches.

12.3. How to determine Bravais lattice types

First, the projectors for each type are necessary. There are two ways to determine the Bravais lattice type.

1. The program BGAOL uses methods essentially the same as those used for distance calculations. For a given input unit cell, the reflections and boundary transformations are applied. The original cell and each of the transformed cells, is projected onto each of the manifolds of the lattice types (24 in the cases of $\mathbf{S^6}$). The distance from the probe cell to the projected cell is the agreement factor. For each lattice type, the cell with the shortest distance is reported.

2. The program SELLA (Andrews & Bernstein, 2022a) starts from vectors representing each of the lattice types and transforms them using the reflections and first layer boundary transformations. (In the case of multiple boundaries, all combinations must be used as in distance calculations). Then the distance from

the reduced probe to each of the transformed manifolds of each type is computed. The shortest distance to one of the transformed manifolds is reported. If needed, the inverse reduction of the input (unreduced) unit cell can be applied to the best projected cell to get the best approximation of the input.

12.4. What do you need to compute distances?

- Reflections
  - $\mathbf{G^6}$ (Andrews & Bernstein, 2014)
  - $\mathbf{S^6}$ (Andrews *et* al., 2019b)
- Boundary transforms
  - $\mathbf{G^6}$ (Andrews & Bernstein, 2014)
  - $\mathbf{S^6}$ (Andrews *et* al., 2019a) (only two, following Delone, but reflections generate a total of 4)
- Bravais lattice projectors
  - $\mathbf{G^6}$ (Paciorek & Bonin, 1992)
  - $\mathbf{S^6}$ (Andrews & Bernstein, 2022b)
- Lattice characters
  - $\mathbf{G^6}$ (Andrews & Bernstein, 2014)
  - $\mathbf{S^6}$ see Section 12.5

12.5. $\mathbf{S^6}$ *Bravais lattice characters*

Each Bravais type defined by Delaunay (1932) corresponds to a linear manifold in $\mathbf{S^6}$. The definitions of each type were described by markings on the elements of the tetrahedra for each type in Figure 21. For implementation in software, it is more convenient to use the "lattice characters" that describe the restrictions.

type	Bravais type	type	Bravais type	type	Bravais type
	& character		& character		& character
C1	cI (rrr rrr)	O1A	oF (rrs rrt)	M1B	mC (rst rsu)
C3	cF (rr0 rr0)	O1B	oI (rst rst)	M2A	mC (rs0 stu)
C5	cP (000 rrr)	O2	oI (rs0 srt)	*M2B*	*mC (rs0 rst)* *
T1	tI (rrs rrs)	*O3*	*oI (rs0 rs0)* *	*M3*	*mC (rs0 ts0)* *
T2	tI (rr0 rrs)	O4	oS (00r sst)	M4	mP (00r stu)
T5	tP (000 rrs)	O5	oP (000 rst)	A1	aP (rst uvw)
R1	hR (rrr sss)	O4	oS (00r sst)	*A2*	*aP (rs0 tuv)* *
R3	hR (rr0 sr0)	O5	oP (000 rst)	A3	aP (rs0 tu0) *
H4	hP (00r rrs)			M1A	mC (rrs ttu)

* cases of non-crystallographic $\mathbf{S^6}$ scalars (90° angles)

12.6. Lattice Matching

A common problem in crystallography is to provide a list of the unit cells of several (or many) crystals so that they can be visually compared, making it easier to identify meaningful clusters of crystals of related morphology. Data collections of unit cells have been created based on similarity of morphology (for example, see Donnay *et* al. (1963). In recent years, the clustering of unit cells from the myriad of images in serial crystallography has been increasingly important.

For each transformed vector, one must store the transformation that will bring it back to the original presentation.

Start with a collection of experimental unit cells. From among them, we select or create the "reference" cell; that is the one to which all the rest will be matched as closely as possible.

We transform the reference cell by many operations in the course of exploring alternative lattice representations. For each newly generated lattice representation,we accumulate the transformations needed to convert back to the original reference cell. All of these operations are be performed in $\mathbf{S^6}$.

Only accumulate transformations that have not already been found (to avoid duplication).

Next all the accumulated, transformed representations of the reference cell must be scaled to a common length and the saved transformations inverted which allows one to return a vector to its original presentation. For improved searching in the final step, it is convenient to store the the vectors in a nearest neighbor search application such as NearTree (Andrews & Bernstein, 2016). This ends the preprocessing of the reference lattice.

To match a probe lattice to the reference cell, first the probe lattice should be Selling reduced. Then the nearest point of the collection of transformed lattice points to the reduced probe is found. The stored transform of the nearest reference is applied to the reduced probe to generate the best approximation.

As an example, consider the case of krait toxin phospholipase A2 (Le Trong & Stenkamp, 2007). Table 1 shows 6 reported forms from the Protein Data Bank. The lattice type and unit cell parameters are as initially reported.

Table 2 shows the same data in the same order after lattice matching to the first entry (1DPY). The immediate conclusion is that the crystals are likely the same crystal form.

Table 1. *Unit cells of phospholipase A2 from the Protein Data Bank.*

PDB		a	b	c	α	β	γ
1DPY	R	57.98	57.98	57.98	92.02	92.02	92.02
1FE5	R	57.98	57.98	57.98	92.02	92.02	92.02
1G0Z	H	80.36	80.36	99.44	90	90	120
1G2X	C	80.95	80.57	57.10	90	90.35	90
1U4J	H	80.36	80.36	99.44	90	90	120
2OSN	R	57.10	57.10	57.10	89.75	89.75	89.75

Table 2. *The data of Table 1 matching a rhombohedral reference. The reference cell is marked in boldface.*

PDB	a	b	c	α	β	γ
1DPY	**57.98**	**57.98**	**57.98**	**92.02**	**92.02**	**92.02**
1FE5	57.98	57.98	57.98	92.02	92.02	92.02
1G0Z	57.02	57.02	57.02	90.39	90.39	89.61
1G2X	57.10	57.11	57.11	89.25	90.25	90.25
1U4J	57.02	57.02	57.02	90.39	90.39	89.61
2OSN	57.10	57.10	57.10	90.25	90.25	89.75

13. Various available programs

- SAUC: Search for related unit cells
- BGAOL: Determine likely cells of higher symmetry
- Follower: Compute distances between cells along a path
- LatticeMatching: Try to match a reference cell
- SELLA: Determine likely cells of higher symmetry
- various cmd programs–

13.1. SAUC (Search of Alternative Unit Cells)

SAUC is software plus an online database of the Protein Data Bank unit cells and the Crystallographic Open Database unit cells.

`https://iterate.sourceforge.net/sauc/`

SAUC provides searching for several alternate distance measures, including $\mathbf{S^6}$ and $\mathbf{G^6}$. Searching by nearest neighbor allows specifying various criteria, such as percent range, nearest neighbors, or within a sphere of specified radius.

A useful example is that of krait toxin phospholipase A2 (Le Trong & Stenkamp, 2007).

Starting from PDB entry 1u4j with cell parameters: (80.36, 80.36, 99.44, 90, 90,

120) in space group R3 with a distance cutoff of 3.5Å using the $\mathbf{G^6}$ metric, four structures of phospholipase A2 are found.

PDB code	G6 distance	protein name
1u4j	0.0	Phospholipase A2 isoform 2
1g0z	0.0	Phospholipase A2
1g2x	0.9	Phospholipase A2
2osn	0.9	Phospholipase A2 isoform 3

13.2. BGAOL - Bravais General Analysis of Lattices

BGAOL is an online program for seeking possible Bravais lattice types for a given unit cell.

`https://iterate.sourceforge.net/bgaol/`

Consider 1,8-terpin (Mighell, 2001) with unit cell 10.912(2), 22.79(4), 10.705(2), 90, 120.64, 90.

The BGAOL results (only the best 4, which are also 1,8-terpins) are:

lattice	distance	unit cell					
oF	**0.025**	18.419	22.790	10.912	90	90	90
mC	**0.025**	10.912	18.419	12.634	90	115.585	90
mI	**0.000**	10.704	22.790	10.705	90	118.713	90
mI	**0.018**	14.652	10.912	14.651	90	102.109	90

The unit cells reported in Table 13.2 are then matched to the lattice type in column 1. If the discovered lattice type is monoclinic (in this case mI), then the unit cell is projected into the manifold of the discovered type.

A web site on which searches may be done is located at `http://iterate.sourceforge.net/bgaol/`.

14. Summary

Unit cells are used to represent crystallographic lattices. Calculations measuring the differences between unit cells are used to provide metrics for measuring meaningful distances between three-dimensional crystallographic lattices. This is a surprisingly complex and computationally demanding problem. We present a review of the current best practice using Delaunay-reduced unit cells in the six-dimensional real space of Selling scalar cells $\mathbf{S^6}$ and the equivalent three-dimensional complex space $\mathbf{C^3}$. The process is a simplified version of the process needed when working with the more complex six-dimensional real space of Niggli-reduced unit cells $\mathbf{G^6}$. Obtaining a distance begins with identification of the fundamental region in the space, continues with conversion to primitive cells and reduction, analysis of distances to the boundaries of the fundamental unit, and is completed by a comparison of direct paths to boundary-interrupted paths, looking for a path of minimal length.

Appendix A
Availability of code

Code supporting this paper is available on github at `http://github.com/duck20/LatticeRepLib.git` and `http://github.com/yayahjb/ncdist.git`.

Appendix B
Projectors

Orthogonal projectors onto linear subspaces can be calculated from any algebraic descriptor of that subspace that can provide a matrix X with vectors that at least span that subspace. Redundant vectors can help to improve the accuracy of the calculation. If A is the columnspace of X, then $P = A(A^TA)^{-1}A^T$ is a projector onto that linear subspace, and $PP = P$ and $P^T = P$.

The application of a projector yields the least-squares best fit.

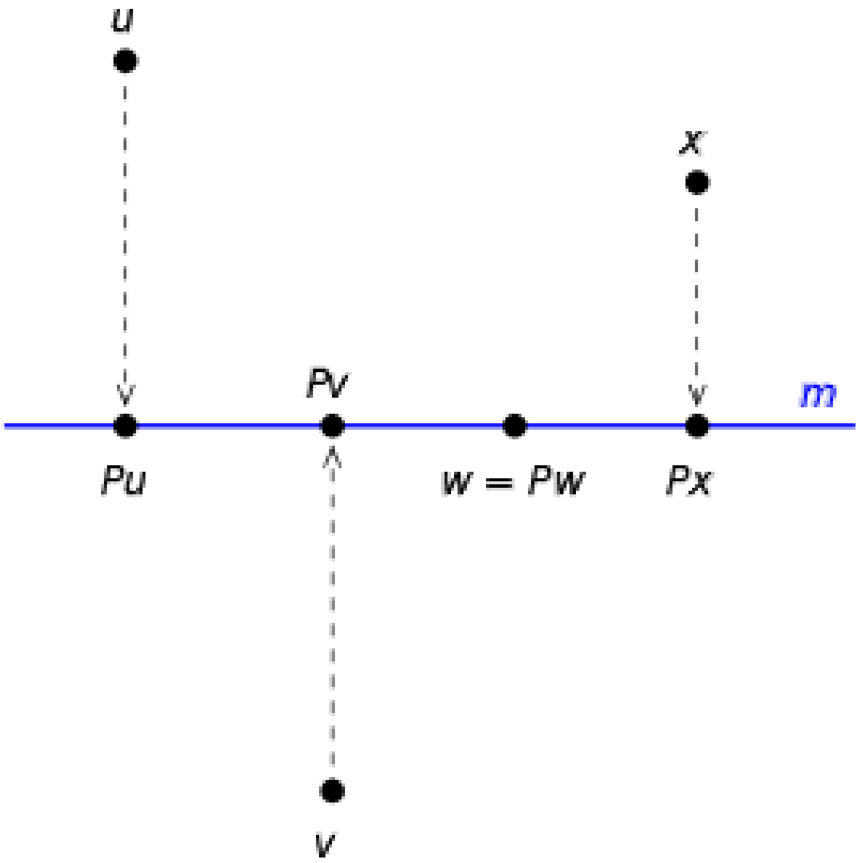


Fig. 23. Example of projectors: they are the operators that do orthogonal projection.

Appendix C
Nearest neighbor search

Wikipedia has a useful description of nearest neighbor searching (Wikipedia contributors, 2022c).

Code for nearest neighbor searching using NearTree is available on GitHub at `http://github.com/yayahjb/neartree.git`. NearTree uses the algorithm of Kalantari & McDonald (1983) with enhancements for other kinds of searches.

Appendix D
Polar coordinates

Expressing unit cell parameters as edge lengths and angles does not lead to a easily interpreted metric: lengths and angles are not commensurate. An unexplored alternative is to express them in polar coordinates, which is equivalent to expressing them in complex coordinates.

Delone pointed out that the "opposite" scalars form special pairs. The Bravais tetrahedron is the origin of his use of the term "opposite". Another way to view this relationship is to organize the six scalars into three pairs, treating the first member of the each pair as the real component of a complex number and the second member of each pair as the imaginary component. Write the definition of an $\mathbf{S^6}$ vector $(s_1, s_2, s_3, s_4, s_5, s_6)$ in $\mathbf{C^3}$:

$$\left[\ \begin{pmatrix} s_1 \\ s_4 \end{pmatrix} \begin{pmatrix} s_2 \\ s_5 \end{pmatrix} \begin{pmatrix} s_3 \\ s_6 \end{pmatrix}\ \right]$$

Rewriting the definitions of the $\mathbf{S^6}$ scalars, we see that for each $\mathbf{C^3}$ scalar, the angle in the real position, is related to the edge vector used in the imaginary position. For instance for c_1, alpha is associated with a. For reflections, these associations are not changed.

$$\left[\ \begin{pmatrix} \mathbf{b}\cos\alpha\ \mathbf{c} \\ \mathbf{a}\cdot\mathbf{d} \end{pmatrix} \begin{pmatrix} \mathbf{a}\cos\beta\ \mathbf{c} \\ \mathbf{b}\cdot\mathbf{d} \end{pmatrix} \begin{pmatrix} \mathbf{a}\cos\gamma\ \mathbf{b} \\ \mathbf{c}\cdot\mathbf{d} \end{pmatrix}\ \right]$$

The stability of the angle/edge length association implies that a potentially useful new coordinate system for describing unit cells might be $\Big(a, \alpha\Big)$, $\Big(b, \beta\Big)$,

$\left(c, \gamma\right)$.

Appendix E
The linear space $\mathbf{G^6}$ (Niggli space)

Label the cell axes $\mathbf{a}, \mathbf{b}, \mathbf{c}$. The scalars are $\mathbf{a} \cdot \mathbf{a},\ \mathbf{b} \cdot \mathbf{b},\ \mathbf{c} \cdot \mathbf{c},\ \ 2\mathbf{b} \cdot \mathbf{c},\ 2\mathbf{a} \cdot \mathbf{c},\ 2\mathbf{a} \cdot \mathbf{b}$ where, *e.g.*, $\mathbf{b} \cdot \mathbf{c}$ represents the dot product of the $\mathbf{b}$ and $\mathbf{c}$ axes. In order to organize these six quantities as a vector space in which one can compute simple Euclidean distances, we describe this set of scalars as a vector, $\mathbf{g}$, with components, $g_1, g_2, g_3, g_4, g_5, g_6$. (Andrews & Bernstein, 2014) It should be noted that the first three scalars of $\mathbf{G^6}$ are the diagonal elements of the metric tensor. The second three are the off-diagonal elements, doubled. One reason for doubling is that those elements occur twice within the metric tensor.

Appendix F
The linear space $\mathbf{S^6}$ (Delone space)

$\mathbf{S^6}$: Andrews *et* al. (2019b) presented a simple and the fastest currently known representation of lattices as the six Selling scalars obtained from the dot products of the unit cell axes in addition to the negative of their sum (a body diagonal). Labeling the cell axes $\mathbf{a}, \mathbf{b}, \mathbf{c}$, and $\mathbf{d}$ ($\mathbf{d} = -\mathbf{a} - \mathbf{b} - \mathbf{c}$), The scalars are $\mathbf{b} \cdot \mathbf{c},\ \mathbf{a} \cdot \mathbf{c},\ \mathbf{a} \cdot \mathbf{b},\ \ \mathbf{a} \cdot \mathbf{d},\ \mathbf{b} \cdot \mathbf{d},\ \mathbf{c} \cdot \mathbf{d}$ where, *e.g.*, $\mathbf{b} \cdot \mathbf{c}$ represents the dot product of the $\mathbf{b}$ and $\mathbf{c}$ axes. For the purpose of organizing these six quantities as a vector space in which one can compute simple Euclidean distances, we describe this set of scalars as a vector, $\mathbf{s}$, with components, $s_1, s_2, s_3, s_4, s_5, s_6$.

A cell is Selling-reduced if all six components are negative or zero (Delaunay, 1932).

- Reversing the observation that a Buerger-reduced cell is a good stepping stone to a Selling-reduced cell (Allmann, 1968), a Selling-reduced cell can be a very efficient stepping stone to a Buerger-reduced cell from which a Niggli-reduced cell can quickly be derived.
- Selling showed that the all-obtuse (Selling parameters all-negative) case is unique (Selling, 1874). Every $\mathbf{S^6}$ vector with all scalars non-positive corresponds to a valid unit cell.

Appendix G
The linear space $\mathbf{C^3}$

$\mathbf{C^3}$ is a space created from the scalars of $\mathbf{S^6}$.

$$\left[\begin{pmatrix} s_1 \\ s_4 \end{pmatrix} \begin{pmatrix} s_2 \\ s_5 \end{pmatrix} \begin{pmatrix} s_3 \\ s_6 \end{pmatrix}\right]$$

This associates each axis with its associated angle (Andrews *et* al., 2019b).

$$\left[\begin{pmatrix} \alpha \\ \mathbf{a} \end{pmatrix} \begin{pmatrix} \beta \\ \mathbf{b} \end{pmatrix} \begin{pmatrix} \gamma \\ \mathbf{c} \end{pmatrix}\right]$$

Appendix H
Geodesics

A geodesic is a curve representing in some sense the shortest path between two points in a surface (Wikipedia contributors, 2022b).

Acknowledgements

Careful copy-editing and corrections by Frances C. Bernstein are gratefully acknowledged. Elizabeth Kincaid is to be thanked for artwork.

Funding information

Funding for this research was provided in part by: US Department of Energy Offices of Biological and Environmental Research and of Basic Energy Sciences (grant No. DE-AC02-98CH10886; grant No. E-SC0012704); U.S. National Institutes of Health (grant No. P41RR012408; grant No. P41GM103473; grant No. P41GM111244; grant No. R01GM117126, grant No. 1R21GM129570); Dectris, Ltd.

References

Allmann, R. (1968). Z. Kryst.-Crystalline Materials, **126**(1-6), 272 – 276.

Andrews, L. (2001). C/C++ Users Journal, **19**, 40 – 49. `http://sf.net/projects/neartree`.

Andrews, L. & Bernstein, H. (2022a). Acta Cryst. A**78**, a81.

Andrews, L. C. & Bernstein, H. J. (2014). J. Appl. Cryst. **47**(1), 346 – 359.

Andrews, L. C. & Bernstein, H. J. (2016). J. Appl. Cryst. **49**(3), 756 – 761.

Andrews, L. C. & Bernstein, H. J. (2022b). J. Appl. Cryst. **55**(4).

Andrews, L. C., Bernstein, H. J. & Sauter, N. K. (2019a). Acta Cryst. A**75**, 115 – 120. **Discusses lesser complexity of Selling reduction and includes pseudocode.**

Andrews, L. C., Bernstein, H. J. & Sauter, N. K. (2019b). Acta Cryst. A**75**(3), 593 – 599.

Authier, A. (2013). Early days of X-ray crystallography. OUP Oxford.

Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N. & Bourne, P. E. (2000). Nucleic Acids Res. **28**, 235 – 242.

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, Jr., E. F., Brice, M. D., Rodgers, J. R., Kennard, O., Shimanouchi, T. & Tasumi, M. (1977). J. Mol. Biol. **112**, 535 – 542.

Bernstein, H. J. & Andrews, L. C. (1979). Database, **2**(1).

Bernstein, H. J., Andrews, L. C. & Xerri, M. (2023). In preparation.

Chichagov, A. V., Varlamov, D. A., Dilanyan, R. A., Dokina, T. N., Drozhzhina, N. A., Samokhvalova, O. L. & Ushakovskaya, T. V. (2001). Crystallography Reports, **46**(5), 876 – 879.

Delaunay, B. N. (1932). Z. Krist. **84**, 109 – 149.

Delone, B. N., Galiulin, R. V. & Shtogrin, M. I. (1975). J. Sov. Math. **4**(1), 79 – 156.

Donnay, J. & Donnay, G. (1953). Science, **118**(3060), 222 – 223.

Donnay, J. D. H., Donnay, G., Cox, E. X., Kennard, O. & King, M. V. (1963). Crystal data: Determinative tables, vol. 547. The American Crystallographic Association, The University of California.

Downs, R. T. & Hall-Wallace, M. (2003). American Mineralogist, **88**(1), 247 – 250.

Feldmann, R. J., Heller, S. R., Shapiro, K. P. & Heller, R. S. (1972). J. Chem. Doc. **12**(1), 41 – 47.

Gražulis, S., Daškevič, A., Merkys, A., Chateigner, D., Lutterotti, L., Quiros, M., Serebryanaya, N. R., Moeck, P., Downs, R. T. & Le Bail, A. (2012). Nucl. Acids Res. **40**(D1), D420 – D427.

Gruber, B. (1973). Acta Cryst. A**29**, 433 – 440.

Haüy, R. J. (1800). Addition au mémoire sur l'arragonite: inséré dans le tome XI des annnales (p. 241 et suiv.).

Hayward, M. (2019). `https://rrpress.utsa.edu/bitstream/handle/20.500.12588/121/Cambridge%20Structural%20Database%20Review%20with%20Screenshots%2020190418.pdf`.

Kadec, S. T. & Jover, A. (1985). Information Services and Use, **5**(5), 259 – 68.

Kalantari, I. & McDonald, G. (1983). IEEE Transactions on Software Engineering, **SE-9**(5), 631 – 634.

Kennard, O., Allen, F, H., Brice, M. D., Hummelink, T. W. A., Motherwell, W. D. S., Rodgers, J. R. & Watson, D. G. (1977). Pure Appl. Chem. **49**(12), 1807 – 1816.

Kepler, J. (1611). Strena Seude Niue Sexangula. Godefridum Tampach.

Kepler, J., Hardie, C. G., Mason, B. J. & Whyte, L. L. (1966). The Six-cornered Snowflake.[Edited and Translated by Colin Hardie. With Essays by LL Whyte and BJ Mason. With Illustrations.] Lat. & Eng. Clarendon Press.

Le Trong, I. & Stenkamp, R. E. (2007). Acta Cryst. D**63**, 548 – 549.

McGill, K. J., Asadi, M., Karakasheva, M. T., Andrews, L. C. & Bernstein, H. J. (2014). J. Appl. Cryst. **47**(1), 360 – 364.

Mehl, M. J., Hicks, D., Toher, C., Levy, O., Hanson, R. M., Hart, G. & Curtarolo, S. (2017). Comput. Mat. Sci. **136**, S1 – S828.

Mighell, A. D. (2001). J. Res. Natl. Inst. Stand. Technol. **106**(6), 983 – 996.

Nanao, M., Basu, S., Zander, U., Giraud, T., Surr, J., Guijarro, M., Lentini, M., Felisaz, F., Sinoir, J., Morawe, C. et al. (2022). J. Sync. Rad. **29**(2).

Paciorek, W. & Bonin, M. (1992). J. Appl. Cryst. **25**(5), 632 – 637. **Projectors for G6 and Niggli reduction**.

Rühl, S. (2019). Materials Informatics: Methods, Tools and Applications, pp. 41 – 54.

Selling, E. (1874). J. reine angew. Math. **1874**(77), 143 – 229.

Taylor, R. & Wood, P. A. (2019). Chem. Rev. **119**(16), 9427 – 9477.

Wikipedia contributors, (2022a). Fundamental domain — Wikipedia, the free encyclopedia. [Online; accessed 15-July-2022]. https://en.wikipedia.org/w/index.php?title=Fundamental_domain&oldid=1028832752

Wikipedia contributors, (2022b). Geodesic — Wikipedia, the free encyclopedia. [Online; accessed 31-October-2022]. https://en.wikipedia.org/w/index.php?title=Geodesic&oldid=1116672577

Wikipedia contributors, (2022c). Nearest neighbor search — Wikipedia, the free encyclopedia. [Online; accessed 31-October-2022]. https://en.wikipedia.org/w/index.php?title=Nearest_neighbor_search&oldid=1103897119

Wikipedia contributors, (2022d). Theophrastus — Wikipedia, the free encyclopedia. [Online; accessed 17-October-2022]. https://en.wikipedia.org/w/index.php?title=Theophrastus&oldid=1114534722

Wilkins, J., Schoville, B. J., Pickering, R., Gliganic, L., Collins, B., Brown, K. S., von der Meden, J., Khumalo, W., Meyer, M. C., Maape, S. et al. (2021). Nature, **592**(7853), 248 – 252.

Yang, P., Liu, H., Chen, Z., Chen, L. & Wang, J. (2014). J. Appl. Cryst. **47**(1), 402 – 413.

Synopsis

Unit cells are used to represent crystallographic lattices. Calculations measuring the differences between unit cells are used to provide metrics for measuring meaningful distances between three-dimensional crystallographic lattices. This is a surprisingly complex and computationally demanding problem. We present a review of the current best practice using Delaunay-reduced unit cells in the six-dimensional real space of Selling scalar cells $\mathbf{S^6}$ and the equivalent three-dimensional complex space $\mathbf{C^3}$.