



An Invertible 7-Dimensional Dirichlet Cell Characterization of Lattices

Herbert J. Bernstein^a, Lawrence C. Andrews^b Mario A. Xerri^{c,d}

^aRonin Institute for Independent Scholarship, c/o NSLS-II, BNL, Upton, NY USA, ^bRonin Institute for Independent Scholarship, Kirkland, WA USA, ^cCollege of Engineering, Cornell Univ., Ithaca, NY USA, ^dBrookhaven National Laboratory, Upton, NY USA

American Crystallographic Association Portland Meeting



Federov Polyhedra Tiling images from https://en.wikipedia.org/wiki/Parallelohedron

Introduction

- Characterization of crystallographic lattices is an important tool in structure solution, crystallographic database searches and clustering of diffraction images in serial crystallography.
- Characterization of lattices by Niggli-reduced cells (based on the three shortest non-coplanar lattice edge vectors) or by Delaunay-reduced cells (based on four edge vectors summing to zero and all meeting at obtuse or right angles) are commonly used.
- The Niggli cell derives from Minkowski reduction. The Delaunay cell derives from Selling reduction.



d = -a-b-c

Wigner-Seitz Cell

- In this talk we consider lattice representation based on the Wigner-Seitz [Wigner and Seitz, 1933] (or Dirichlet, or Voronoi) cell of the lattice, which consists of the points at least as close to a lattice point as they are to any other lattice point. Starting from a Niggli reduced cell, the Dirichlet cell is characterized by the planes determined by thirteen lattice half-edges: the midpoints of the three Niggli cell edges, the six Niggli cell face-diagonals and the four body-diagonals, but seven of the edge lengths are sufficient: three edge lengths, the three shorter face-diagonals and the shortest body-diagonal, from which the Nigglireduced cell may be recovered.
- The Wigner-Seitz cell is a polyhedron of six, eight, ten, twelve or fourteen faces. The general fourteen face case is a truncated octahedron.



- That crystals are built from some regular assembly of basic parts was clear in ancient times. In 1611 Kepler described this relationship [Kepler, 1611] translated in [Kepler *et al.*, 1966]. Steno was asked to prepare a catalog of a "cabinet of curiosities"; this is looked at as the first database of crystals (minerals in this case) [Steno, 1669]. In the 19th century indices were published with interaxial angles of crystals, specifically for the identification of minerals. Following the discovery of x-rays, catalogs of unit cell parameters started to be published. [Wyckoff, 1931].
- Often, these were arranged by crystal system and then sorted by some of the cell parameters. However, related minerals with distortions or deformed into another crystal system or incorrectly attributed to another could be difficult to find. Clearly a metric for relating unit cells was required.

Reduced Cells

- Niggli [Niggli, 1928] and Delaunay (aka Delone) [Delone, 1933] devised "reduced cells', which allowed for a more standard presentation of some of the crystal data.
- The Buerger-reduced cell is simpler than the Niggli-reduced cell, having fewer constraints [Buerger, 1957] [Azaroff and Buerger, 1958] [Buerger, 1960]. All Niggli-reduced cells are Buerger-reduced, but not all Buerger-reduced cells are Niggli reduced.
- Finally, in the 1970s, NIH and EPA joined to create the online searchable Chemical Information System (CIS) [Heller *et al.*, 1976] [Bernstein and Andrews, 1979]. Along with such physical measurements as NMR and IR, NIH/EPA wanted to include unit cell searching. At the time, there was no commonly accepted method to compute the "distance" between two unit cells (equivalently, lattices).
- There were two problems: experimental error and ambiguity of representation.

Experimental Error and Ambiguity of Representation

- The first problem was that measured unit cell parameters (conventionally, [**a**, **b**, **c**, **alpha**, **beta**, **gamma**] for the cell lengths and angles) always have experimental error in their determinations. Further, closely related compounds of interest might have slightly different cell edge parameters. That means that the problem to be solved is "the nearest neighbor problem" also known as "the post office problem". Exact match is too strict.
- The second issue relates to the problem of experimental error, but it manifests in a different way. It is well-known that for any given lattice, there is an infinity of unit cells that can be chosen.
- The problem is that two unit cells from the same lattice may not look the same. The cells: [10, 10, 10, 90, 90, 90] and [10, 10, 10, 120, 120, 90] are two different choices from a single crystal lattice. Only the first of these is Niggli reduced, but lattice variation may make the second the reduced one.

Computing Distances Among Lattices

- For crystallography, there are many alternative representations to choose from as a basis for distance calculations.
- V⁷: [Andrews *et al.*, 1980] discussed V⁷, a perturbation-stable space in which, using real and reciprocal space Niggli reduction, a lattice is represented by three cell edge lengths, three reciprocal cell edge lengths and the cell volume, which was proposed for cell database searches, but which has difficulties when used for lattice determination because of poor sensitivity near 90°.
- **G**⁶: [Andrews and Bernstein, 1988]: discussed **G**⁶ that uses a modified metric tensor and iterative searches through 25 alternative reduction boundary transforms [Gruber, 1973] to work in a satisfactory manner both for database searches and lattice identification in the presence of experimental error.

Computing Distances Among Lattices II

- **NCDist**: [Andrews and Bernstein, 2014] discussed sewing together regions of the fundamental region of G⁶ under Niggli reduction at 15 boundaries.
- S⁶: [Andrews *et al.*, 2019] presented the simplest and 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 a, b, c, and d (d = -a b c), the scalars are b ⋅ c, a ⋅ c, a ⋅ b, a ⋅ d, b ⋅ d, c ⋅ d where, *e.g.*, b ⋅ c represents the dot product of the b and c axes). For the purpose of organizing these six quantities as a vector space in which one can compute simple Euclidean distances, we describe the set of scalars as a vector, s, with components, s₁, s₂, s₃, s₄, s₅, s₆.
- A cell is Selling-reduced if all six components are negative or zero [Delone, 1933]. Reversing Allman's 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 Niggli-reduced cell.

Why Consider the Wigner-Seitz Cell

- All of V⁷, G⁶ and S⁶ work for comparing lattices, but differ in speed of comparison and smoothness of the space. Speed becomes important when the number of lattices to be compared becomes large, as when clustering thousands of samples [Soares *et al.*, 2022]. Smoothness of the space is a more subtle issue, but smoother spaces help in recognizing particularly good candidates for molecular replacement in macromolecular studies. Until the present work, except for V⁷, S⁶ had been the fastest, smoothest candidate, and S⁶ is much more accurate. We are investigating the Wigner-Seitz cell to see how well it performs in comparison to S⁶.
- Define the Wigner-Seitz cell to consist of the points which are no farther from a given lattice point than they are from any other lattice point. As [Hart *et al.*, 2019] has shown the Wigner-Seitz cell centered on a given lattice point is contained entirely within the convex envelope of the eight immediate neighbors of a given lattice point, *i.e.* in terms of thirteen Miller indices and their negatives:

100, 010, 001, 011, 101, 110, 011, 101, 110, 111, 1

Wigner-Seitz Cell Definition

- We organize the lattice in terms of a basis of the three shortest non-coplanar distances, *i.e.* on the basis of Minkowski reduction.
- Formally, the definition of the Wigner-Seitz cell is: Let L be an R³ lattice with Minkowski basis a, b, c, *i.e.* such that ha + kb + lc, h ∈ Z, k ∈ Z, l ∈ Z spans L and ||a||, ||b||, ||c|| are minimal. We define the Wigner-Seitz cell of L as WS(L) = {w ∈ R³ | ∀x ∈ L, x ≠ 0, ||w|| ≤ ||x - w|| }
- If we translate this cell to each element of the lattice L, we tile the space and have a Voronoi decomposition.
- Niggli-reduction provides an unambiguous Minkowski reduction. Assume the cell formed by a, b, c is Niggli reduced, with ||a|| ≤ ||b|| ≤ ||c||. Define

$$\mathbf{r}, \mathbf{s}, \mathbf{t}, \mathbf{u}, \mathbf{v}, \mathbf{w} = \mathbf{a} \cdot \mathbf{a}, \mathbf{b} \cdot \mathbf{b}, \mathbf{c} \cdot \mathbf{c}, 2\mathbf{b} \cdot \mathbf{c}, 2\mathbf{a} \cdot \mathbf{c}, 2\mathbf{a} \cdot \mathbf{b}$$

As a Niggli-reduced cell, we either have all of $\mathbf{u}, \mathbf{v}, \mathbf{w} \le 0$ or all of $\mathbf{u}, \mathbf{v}, \mathbf{w} > 0$.

ACA Portland 29 July – 2 August 2022

\mathbf{DC}^{13} and sorted \mathbf{DC}^7

• We define a **DC**¹³ cell as the three edges, the six face diagonals and the four body diagonals: **dc**_{13,raw}(**r**, **s**, **t**, **u**, **v**, **w**) =

```
[r, s, t,
s+t-u, s+t+u, r+t-v, r+t+v, r+s-w, r+s+w,
r+s+t+u+v+w, r+s+t+u-v-w, r+s+t-u+v-w, r+s+t-u-v+w]
when presented in terms of the G<sup>6</sup> components as the squares of the
lengths.
```

- If we sort the elements of DC¹³ in ascending numerical order and only present the first seven elements, we have DC⁷ as discussed in [Bernstein and Andrews, 2021], which is a smooth, but ambiguous, characterization of lattices.
- **DC**⁷ is not invertible in some cases unless the symmetry is known *a priori* or if some elements after the seventh are retained.

The Bright **DC**⁷ Ambiguity Example

- Bright [Bright, 2021] has demonstrated the DC⁷ ambiguity with the cells i: [2.8284, 3.162277, 3.4641, 117.157, 107.8295, 116.5651] and ii: [2.8284, 3.162277, 3.4641, 123.211, 107.8295, 109.59748] as [a, b, c, α, β, γ], or
 - i: [8, 10, 12, -10, -6, -4] and
 - ii: [8, 10, 12, -12, -6, -3]

as the \mathbf{G}^6 vectors [\mathbf{r} , \mathbf{s} , \mathbf{t} , \mathbf{u} , \mathbf{v} , \mathbf{w}] for which the sorted \mathbf{DC}^{13} elements after Niggli reduction are

i: 2.44949, 2.82843, 3.16228, 3.16228, 3.4641, 3.4641, 3.74166, **4**,

4.24264, 4.47214, 5.09902, 5.09902, 6.16441

and

ii: 2.44949, 2.82843, 3.16228, 3.16228, 3.4641, 3.4641, 3.74166, **4.24264**,

4.24264, 4.89898, 4.89898, 5.2915, 5.65685

respectively, which do not differ until the eighth element.

ACA Portland 29 July – 2 August 2022

\mathbf{DC}^7 unsorted

- If all of u, v, w ≤ 0, then s+t+u ≤ s+t-u, r+t+v ≤ r+t-v, r+s+w ≤ r+s-w, *i.e.* the three summed squares of the face diagonals are no larger than the corresponding differences, and r+s+t+u+v+w ≤ { r+s+t+u-v-w, r+s+t-u-v+w } , *i.e.* the summed squared main body diagonal is no larger than the three remaining body diagonals.
- On the other hand, If all of u, v, w > 0, then s+t+u > s+t-u, r+t+v > r+t-v, r+s+w > r+s-w, *i.e.* the three summed squares of the face diagonals are strictly greater than the corresponding differences, and r+s+t+u+v+w > each of {r+s+t+u-v-w, r+s+t-u+v-w, r+s+t-u-v+w}, *i.e.* the smallest squared body diagonal is strictly less than r+s+t+u+v+w, and strictly greater than r+s+t-|u|-|v|-|w|.
- Therefore we can recover r, s, t, u, v, w from the three cell edges, the three shorter face diagonals, and the shortest squared body diagonal defining DC⁷_{unsrt} as follows:

$$dc7_{unsrt} = [dc7_{unsrt,1}, dc7_{unsrt,2}, dc7_{unsrt,3}, dc7_{unsrt,4}, dc7_{unsrt,5}, dc7_{unsrt,6}, dc7_{unsrt,7}] = [r, s, t, s+t-|u|, r+t-|v|, r+s-|w|,$$

min(**r**+**s**+**t**+**u**+**v**+**w**, **r**+**s**+**t**+**u**-**v**-**w**, **r**+**s**+**t**-**u**+**v**-**w**, **r**+**s**+**t**-**u**-**v**+**w**)]

ACA Portland 29 July – 2 August 2022

\mathbf{DC}^7 unsorted II

 If we subtract the face diagonal from the matching pairs of edges, we get the absolute values of u, v, w

$$\begin{aligned} & \mathsf{dc7}_{unsrt,2} + \mathsf{dc7}_{unsrt,3} - \mathsf{dc7}_{unsrt,4} = \mathsf{s+t-}(\mathsf{s+t-}|\mathsf{u}|) = |\mathsf{u}| \\ & \mathsf{dc7}_{unsrt,1} + \mathsf{dc7}_{unsrt,3} - \mathsf{dc7}_{unsrt,5} = \mathsf{r+t-}(\mathsf{r+t-}|\mathsf{v}|) = |\mathsf{v}| \\ & \mathsf{dc7}_{unsrt,1} + \mathsf{dc7}_{unsrt,2} - \mathsf{dc7}_{unsrt,6} = \mathsf{r+s-}(\mathsf{r+s-}|\mathsf{w}|) = |\mathsf{w}| \\ & \text{from which we can compute } \boldsymbol{\tau} = \mathsf{r+s+t-}|\mathsf{u}|-|\mathsf{v}|-|\mathsf{w}|. \end{aligned}$$

• If $\tau = dc7_{unsrt,7}$, we can be certain that all of $\mathbf{u}, \mathbf{v}, \mathbf{w} \le 0$. If $\tau \ne dc7_{unsrt,7}$ and the difference is larger than the possible experimental or rounding errors, we can be certain that all of $\mathbf{u}, \mathbf{v}, \mathbf{w} > 0$. Thus the Niggli cell can be recovered from $dc7_{unsrt}$.

DC⁷ unsorted III

For example, the Niggli-reduced G⁶ versions of the cells in the Bright example are i: [6, 8, 10, 8, 4, 2] and ii: [6, 8, 10, -6, -2, -4], Note that the former has all of u, v, w positive and the latter negative. The resulting squared dc7_{unsrt} vectors,

[**r**, **s**, **t**, **s** + **t** - |**u**|, **r**+**t**-|**v**|, **r**+**s**-|**w**|,

min(**r**+**s**+**t**+**u**+**v**+**w**, **r**+**s**+**t**+**u**-**v**-**w**, **r**+**s**+**t**-**u**+**v**-**w**, **r**+**s**+**t**-**u**-**v**+**w**)

are then i: [6, 8, 10, 10, 12, 12, 14] and ii: [6, 8, 10, 12, 14, 10, 12], so $|\mathbf{u}|, |\mathbf{v}|, |\mathbf{w}|$ are 8 + 10 - 10 = 8, 6 + 10 - 12 = 4, 6 + 8 - 12 = 2 in the first case and 8 + 10 - 12 = 6, 6 + 10 - 14 = 2, 6 + 8 - 10 = 4 in the second case. $\boldsymbol{\tau} = 6 + 8 + 10 - 8 - 4 - 2 = 10 \neq 14$ in the first case, indicating all-acute angles $\boldsymbol{\tau} = 6 + 8 + 10 - 6 - 2 - 4 = 12$, indicating all right or obtuse angles in the second case, correctly recovering the signs.



\mathbf{G}^{6} :	r	S	t	u	V	V
i	6	8	10	8	4	2
ii	6	8	10	-6	-2	-4

DC ⁷ unsrt	r	S	t	s + t	r + t	r + s	min body diag
				- u 	- v 	- w 	(MBD)
i	6	8	10	10	12	12	14
ii	6	8	10	12	14	10	12

recover \mathbf{G}^6 :	r	S	t	u	 v 	 w 	<i>τ</i> = r + s + t − u − v − w
i	6	8	10	8	4	2	τ 10 \neq MBD 14 (disagree + + +)
ii	8	8	10	6	2	4	$\tau 12 = MBD 12$ (agree)

Theory and Practice

- As we have shown, the unsorted version of DC⁷ is easily computed from a primitive Niggli-reduced cell and, at least in theory, is accurately invertible. Extensive testing needs to be done to compare database searches and clustering using NCDIST versus S⁶ versus sorted DC⁷ versus the new unsorted DC⁷. At present the invertibility of unsorted DC⁷ has been tested against all the X-ray diffraction unit cells in the Protein Data Bank as of mid-June 2022. All but one cell in 166278 inverted to better than six digit accuracy. The one exception only inverted to five digit accuracy.
- A preprint of a paper on unsorted DC⁷ is available at http://www.bernstein-plus-sons.com/TMM/dc7unsrt-10.pdf
- The code is available on github in http://github.com/yayahjb/ncdist

Acknowlegements

- Our thanks to Frances C. Bernstein for careful copy-editing and helpful suggestions.
- Our thanks to Elizabeth Kincaid for the timeline figure.
- Our thanks to Robert M. Sweet for helpful comments.
- Work supported in part by
 - US Dept. of Energy, Office of Science, DOE Office of Biological and Environmental Research, grant KP1607011 and DOE Office of Basic Energy Sciences Program, contract DE-SC0012704.
 - US NIH National Institute of General Medical Sciences P30GM133893.
 - U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists (WDTS) under the Science Undergraduate Laboratory Internships Program (SULI)

References I

[Allmann, 1968] Allmann, R. (1968).

Some remarks on reduced cells.

Z. Kristallogr. – Crystalline Materials, 126(1-6):272 – 276.

[Andrews and Bernstein, 1988] Andrews, L. C. and Bernstein, H. J. (1988). Lattices and reduced cells as points in 6-space and selection of Bravais lattice type by projections. *Acta Cryst.*, A44:1009 – 1018.

[Andrews and Bernstein, 2014] Andrews, L. C. and Bernstein, H. J. (2014).

The geometry of niggli reduction: Bgaol–embedding niggli reduction and analysis of boundaries.

J Appl. Cryst., 47(1):346 – 359.

[Andrews *et al.*, 1980] Andrews, L. C., Bernstein, H. J., and Pelletier, G. A. (1980).

A perturbation stable cell comparison technique.

Acta Cryst., A36:248 - 252.

References II

[Andrews *et al.*, 2019] Andrews, L. C., Bernstein, H. J., and Sauter, N. K. (2019).

Selling reduction versus niggli reduction for crystallographic lattices. *Acta Cryst.*, A75:115 – 120.

[Azaroff and Buerger, 1958] Azaroff, L. V. and Buerger, M. J. (1958).

The powder method in X-ray crystallography, chapter Chapter 11, Reduced cells and their application, pages 124 – 159. McGraw-Hill.

[Bernstein and Andrews, 1979] Bernstein, H. J. and Andrews, L. C. (1979). The NIH-EPA Chemical Information System. *Database*, 2(1):35 – 43.

[Bernstein and Andrews, 2021] Bernstein, H. J. and Andrews, L. C. (2021). DC7, a very efficient lattice comparison metric. *Acta Cryst.*, A77:C809.

References III

[Bright, 2021] Bright, M. (2021).

Personal communication.

[Buerger, 1957] Buerger, M. J. (1957). Reduced cells.

Z. Kristallogr., 109:42 – 60.

[Buerger, 1960] Buerger, M. J. (1960).

Note on reduced cells.

Z. Kristallogr., 113:52 – 56.

[Delone, 1933] Delone, B. N. (1933).Neue Darstellung der geometrices Kristallographie.*Z. Kristallogr.*, 84:109 – 149.

[Gruber, 1973] Gruber, B. (1973).

The relationship between reduced cells in a general bravais lattice. *Acta Cryst.*, A29:433 – 440.

References IV

[Hart *et al.*, 2019] Hart, G. L., Jorgensen, J. J., Morgan, W. S., and Forcade, R. W. (2019).

A robust algorithm for k-point grid generation and symmetry reduction.

J. Phys. Comm., 3(6):065009.

[Heller *et al.*, 1976] Heller, S. R., Milne, G. W. A., and Feldmann, R. J. (1976).

Quality control of chemical data bases.

J. Chem. Info. and Comp. Sci., 16(4):232 – 233.

[Kepler, 1611] Kepler, J. (1611).

Strena Seude Niue Sexangula.

Godefridum Tampach.

[Kepler *et al.*, 1966] Kepler, J., Hardie, C. G., Mason, B. J., and 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.

References V

[Niggli, 1928] Niggli, P. (1928).

Krystallographische und Strukturtheoretische Grundbegriffe, Handbuch der Experimentalphysik, Vol. 7, part 1.

Akademische Verlagsgesellschaft, Leipzig.

[Soares et al., 2022] Soares, A., Yamada, Y., Jakoncic, J., McSweeney, S., Sweet, R. M., Skinner, J., Foadi, J., Fuchs, M. R., Schneider, D., Shi, W., Andi, B., Andrews, L. C., and Bernstein, H. J. (2022). Serial Crystallography with Multi-stage Merging oi 1000s of Images. *Acta Cryst.*, F77(to appear).

[Steno, 1669] Steno, N. (1669).

De solido intra solidum naturaliter contento dissertationis prodromus. Florence Insegna della Stella.

[Wigner and Seitz, 1933] Wigner, E. and Seitz, F. (1933). On the constitution of metallic sodium. *Physical Review*, 43(10):804 – 810.

References VI

[Wyckoff, 1931] Wyckoff, R. W. G. (1931). *The structure of crystals number 19*. The Chemical Catalog Company, Inc.