Crystallographic map: A general lattice and basis formalism enabling efficient and discretized exploration of crystallographic phase space

David Mrdjenovich[1] and Kristin A. Persson[1,2,*]

[1]*Department of Materials Science & Engineering, 210 Hearst Mining Building, University of California, Berkeley, California 94720, USA*

[2]*Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, California 94720, USA*

(Received 3 August 2022; accepted 20 December 2023; published 13 March 2024)

Three-dimensional lattices are fundamental to solid-state physics. The description of a lattice with an atomic basis constitutes the necessary information to predict solid phase properties and evolution. Here, we present an algorithm for systematically exploring crystallographic phase space. Coupled with *ab initio* techniques, such as density functional theory, this algorithm offers an approach for exploring and tuning materials behavior, with a broad range of potential applications: particularly martensitic phase transformations and materials stability.

DOI: 10.1103/PhysRevMaterials.8.033401

I. INTRODUCTION

The fundamental description of any crystal consists of a lattice and an atomic basis. This core geometry dictates a myriad of materials properties, informing optical, electronic, mechanical, and chemical functionality. In materials science and solid-state physics, understanding how the energy varies as a function of its crystallography is crucial to explaining phase transformations and behaviors such as piezoelectricity, mechanical stability, ferroelectricity, etc. However, techniques for exploring this energetic landscape, particularly to find low-energy transformation pathways, have remained limited.

A common class of approaches have used Landau theory and symmetry group-subgroup relationships between starting and ending structures to guess possible transformation pathways [1]. However, these techniques rely on the crystal retaining a certain degree of symmetry throughout the transformation. This is not the case in general; further, linear interpolation and mapping of Wyckoff labels used in these approaches does not guarantee correspondence with a low-energy pathway. Other studies have employed specialized solid-state nudged elastic band techniques to search for transformation pathways [2]. These approaches show promise for surveying crystallographic phase space more generally; however, ultimately, the search is confined to a small region dictated by the relaxation of an initial transformation pathway ansatz.

Other approaches have focused on finding martensitic phase transformation pathways as a combination of a strain plus an atomic basis shuffle, including the work of Trinkle *et al.* Here various supercells of the starting and ending crystal structures are connected using a method similar to the magic strain method [3]. Basis atoms are then mapped by enumerating possible bijections between the relative coordinates of the crystal structures and filtering those where the atoms approach too closely or travel too long distances [3]. Each individual pathway is limited to a two-dimensional cut through crystallographic phase space, and computational work is required to screen out duplicate pathways. The low dimensionality of each pathway means that any transformation involving a series of microstrains or distinct atomic shuffles cannot be found by the formalism. Further, when calculating activation-energy barriers along a set of derived pathways, it is likely that many similar crystal geometries will be sampled. This represents redundant computational work.

As a final note, work on the crystal fingerprint algorithm [4], as a means of uniquely representing and grouping similar crystal structures, is very much aligned with the goals of this paper. In contrast to the use of radial distance functions to produce a string of structural coordinates, we pursue a different approach. By using the Voronoi representation of the lattice [5], detailed below, it is simpler to invert the mapping, going from a string of coordinates back to a physical crystal structure in terms of a lattice and basis. Further, each coordinate is a unique, nonredundant function of the underlying crystal geometry, and neighboring coordinates are related by a simple structural distortion. With these strengths in mind, we believe it important to put forward another approach to numerically represent crystal structures.

Here, we delineate a distinct algorithm for exhaustively exploring crystallographic phase space in a nonredundant, incremental, and rigorous fashion. Further, the approach is discrete, robust to numerical precision errors, and proceeds intuitively. The result is a crystallographic map: a means of uniquely representing a crystal structure's geometry as a series of coordinates in a way that immediately elucidates geometrically similar crystal structures and connects any two crystal structures, of equal numbers of atoms, by a series of small structural perturbations. Through the algorithm, we create an implicit graph data structure that can be used to explore arbitrary trajectories through phase space and identify low-energy transformation pathways. With this paper, we aim to simplify the process of exploring and studying how materials transform and how crystal structures are geometrically related, a fundamental part of solid-state physics.

*Corresponding author: kapersson@lbl.gov

II. REPRESENTATION OF 3D LATTICES

Any 3D lattice can be represented by a set of three linearly independent generating vectors; however, an equivalent and alternative specification exists. Every 3D lattice possesses at least one nonstrictly obtuse superbasis $\{\vec{v}_0, \vec{v}_1, \vec{v}_2, \vec{v}_3\}$, calculable using the selling algorithm [5]. Here $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ are a generating set for the lattice, $\vec{v}_0 := -(\vec{v}_1 + \vec{v}_2 + \vec{v}_3)$, and the angles between the vectors are all at least 90°. Additionally, the seven square vector lengths $\{v_0^2, v_1^2, v_2^2, v_3^2, (\vec{v}_0 + \vec{v}_1)^2, (\vec{v}_0 + \vec{v}_2)^2, (\vec{v}_0 + \vec{v}_3)^2\}$, termed here as squared vonorms, also represent the lattice. The matrix equation below [5] relates these vector lengths to the dot products of the superbasis: $\{\vec{v}_i \cdot \vec{v}_j\}$. Thus, by knowing these seven vector lengths, one may calculate the dot products of an obtuse superbasis of the lattice, thereby reconstructing the lengths and angles of the 3D lattice's generating set with some algebra [Eqs. (4)–(11)]:

$$\begin{pmatrix} -1 & -1 & 1 & 1 & 1 & -1 & -1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} v_0^2 \\ v_1^2 \\ v_2^2 \\ v_3^2 \\ (\vec{v}_0 + \vec{v}_1)^2 \\ (\vec{v}_0 + \vec{v}_2)^2 \\ (\vec{v}_0 + \vec{v}_3)^2 \end{pmatrix} = 4 \begin{pmatrix} \vec{v}_0 \cdot \vec{v}_1 \\ \vec{v}_0 \cdot \vec{v}_2 \\ \vec{v}_0 \cdot \vec{v}_3 \\ \vec{v}_1 \cdot \vec{v}_2 \\ \vec{v}_1 \cdot \vec{v}_3 \\ \vec{v}_2 \cdot \vec{v}_3 \\ 0 \end{pmatrix}. \tag{1}$$

The set of seven squared vonorms and the aforementioned dot products are a geometric invariant of the lattice: any two lattices related by an affine transformation will have the same set of square vonorms and dot products. By definition, Eq. (1) must be satisfied by all obtuse superbases that exist for a given lattice. However, there may be several distinct obtuse superbases for the same lattice geometry, either related by affine transformations or a reordering of the vonorms and dot products.

Studying the reordering of distinct superbases, we term the lengths $\{v_0^2, v_1^2, v_2^2, v_3^2\}$ primary squared vonorms and the lengths $\{(\vec{v}_0 + \vec{v}_1)^2, (\vec{v}_0 + \vec{v}_2)^2, (\vec{v}_0 + \vec{v}_3)^2\}$ secondary squared vonorms. It can be shown by brute force calculation that it is impossible to interchange primary and secondary squared vonorms while preserving the validity of Eq. (1) and the algebraic structure of a superbasis. However, it is possible to permute the labels of the primary squared vonorms arbitrarily. These manipulations constitute the S_4 permutation group and signify that there can be up to 24 distinct obtuse superbases representing the same lattice geometry, each superbasis having the same set of lengths and angles in a distinct order.

Calculating the orbit of the squared vonorms under the group action of S_4 yields a collection of reordered primary and secondary squared vonorms. We choose a unique representative (L) from this orbit by choosing the permutation for which the seven ordered lengths, $\{v_0^2, v_1^2, v_2^2, v_3^2, (\vec{v}_0 + \vec{v}_1)^2, (\vec{v}_0 + \vec{v}_2)^2, (\vec{v}_0 + \vec{v}_3)^2\}$, are maximally ascending. In a precise sense, we say that for each orbit member (M), there exists an index j such that the first (j)–(1) elements of (L) are $\leqslant$ the first (j)–(1) elements of (M) *and* the *j*th element of (L) is strictly less than the *j*th element of (M). Thus, for any lattice, we can select a unique ordered string of lengths to represent its geometry unambigouously.

As an example, consider the rhombohedral crystal structure of the element, antimony, where the squared vonorms are as follows (units of Å^2, rounded to the nearest 0.1) [6]:

$$\{v_0^2 = 19.2,\ v_1^2 = 21.3,\ v_2^2 = 19.2,\ v_3^2 = 21.3,\ (\vec{v}_0 + \vec{v}_1)^2 = 40.5,\ (\vec{v}_0 + \vec{v}_2)^2 = 19.2,\ (\vec{v}_0 + \vec{v}_3)^2 = 21.3\}.$$

First, we sort the primary square vonorms. This can be accomplished by swapping labels 1 and 2. Notice both secondary square vonorms $(\vec{v}_0 + \vec{v}_1)^2$ and $(\vec{v}_0 + \vec{v}_2)^2$ also swap:

$$\{v_0^2 = 19.2,\ v_1^2 = 19.2,\ v_2^2 = 21.3,\ v_3^2 = 21.3,\ (\vec{v}_0 + \vec{v}_1)^2 = 19.2,\ (\vec{v}_0 + \vec{v}_2)^2 = 40.5,\ (\vec{v}_0 + \vec{v}_3)^2 = 21.3\}.$$

Seeing that square vonorms v_0^2 and v_1^2 are now equal, we can swap indices 0 and 1 without changing the order of the primary square vonorms. By swapping 0 and 1, the secondary square vonorms are reordered. The result is maximally ascending:

$$\{v_0^2 = 19.2,\ v_1^2 = 19.2,\ v_2^2 = 21.3,\ v_3^2 = 21.3,\ (\vec{v}_0 + \vec{v}_1)^2 = 19.2,\ (\vec{v}_0 + \vec{v}_2)^2 = 21.3,\ (\vec{v}_0 + \vec{v}_3)^2 = 40.5\}.$$

Then the final unambiguous representation of this lattice is the following string:

{19.2, 19.2, 21.3, 21.3, 19.2, 21.3, 40.5}.

From this standardized representation, it is easy to calculate similar lattices. Any two square vonorms out of the seven can be varied by a small amount $\pm\xi$, such that Eq. (2) remains true:

$$-\left(v_0^2 + v_1^2 + v_2^2 + v_3^2\right) + (\vec{v}_0 + \vec{v}_1)^2 + (\vec{v}_0 + \vec{v}_2)^2 + (\vec{v}_0 + \vec{v}_3)^2 = 0. \tag{2}$$

This is a necessary invariant of all 3D lattices [5]. Where ξ is small, the new lengths correspond to a lattice, one that is related to the original by a small strain (Appendix, Proof 1). Using this covariation technique, a lattice can have up to 42 different neighboring lattices. Each neighbor will be distinct but geometrically similar to the original lattice.

It is possible to use a measuring unit of small size ξ. Then, by rounding each vonorm to the nearest ξ, while preserving Eq. (2), we can store each lattice as a series of integer coordinates. By varying these coordinates ± 1 it is then possible to calculate the neighbors as before. Referencing Eq. (1), we observe that, with the given discretization, each dot product must be an integer multiple of $\xi/4$. By storing the quantities $4\vec{v}_i \cdot \vec{v}_j$, it is possible to represent the dot products also as integers. For any nonstrictly obtuse superbasis, each dot product must be less than or equal to zero. Thus, out of the 42 total possible lattice neighbors, we keep only those who are distinct with qualifying nonpositive dot products. In floating-point

arithmetic, due to limited numerical precision, it can be difficult to determine whether a dot product is positive [7]. Thus, the decision to use an integer representation has the added bonus of avoiding any ambiguities in this determination.

It can be shown that, by a series of lattice neighbors, any integer 3D lattice can be connected to any other integer 3D lattice (Appendix, Proof 2). Further, by design, if lattice A is a neighbor of lattice B, it follows that lattice B is a neighbor of lattice A (Appendix, Proof 3). Keeping these properties in mind, this discretization scheme creates a grid over the entire configuration space of 3D lattices and provides a straightforward means of exploring the space. In general, neighbors are similar to one another if the discretization parameter is sufficiently small, and two lattices become less similar the more neighbors that separate them. It is shown in the Appendix, Proof 1, that ξ/v_i^2 is roughly equal to the square root of the magnitude of the principal strains connecting two neighbors.

III. REPRESENTATION OF CRYSTALLINE ATOMIC BASES

In the representation of a crystal's atomic basis (n atoms), there are many redundancies in the specification, namely, arbitrary origin, identical atoms, and lattice translations. By choosing relative coordinates in the interval [0,1), the ambiguity of arbitrary lattice translations is resolved. By confining a particular atom to the origin, the ambiguity of arbitrary origin is accounted for, and only $n-1$ atomic positions need be specified. However, where there are identical atoms, there remains possible ambiguity as to which one of the identical atoms is confined to the origin. Further, in an ordered listing of the $n-1$ basis vectors, $\{\vec{v}_1, ..., \vec{v}_{n-1}\}$, swapping identical nonorigin atoms will not change the crystal structure; however, the listing will be reordered.

Using the same technique employed with the lattice length strings above, we can calculate the full orbit of basis vector strings under their permutation group and then pick a unique representative that is maximally ascending. Specifically, we order atoms primarily by species and, secondarily, by first, second, and then third relative coordinates, in ascending numeric order. This uniquely specifies the representation of the crystal basis in terms of a string of numbers.

Finally, by dividing the interval [0,1) into δ small intervals, it's possible to round each basis coordinate to the nearest interval boundary. Then, the basis can be completely specified as a string of integers.

As an example, consider the compound Sn_2O_4, a unit cell of the experimentally observed tin-oxide rutile phase in the canonical setting [8]. This crystal has the following cell parameters: $a = 3.24$ Å, $b = 4.83$ Å, $c = 4.83$ Å, $\alpha = 90^o$, $\beta = 90^o$ $\gamma = 90^o$, when rounded to the nearest hundredth of an angstrom. It also has the following atomic basis, when rounded to the nearest tenth in relative coordinates:

Two atoms: Sn @0, 0, 0 and @0.5, 0.5, 0.5.

Four atoms: O @0, 0.7, 0.7; @0, 0.3, 0.3; @0.5; 0.8, 0.2; and @0.5, 0.2, 0.8.

We decide here to divide the interval [0,1) into 10 small intervals, and we choose atom Sn to be at the origin. Notice that there are two distinct choices for which atom Sn is at the origin.

With the first choice, as written above, the basis can be represented by the following coordinate string:

5,5,5, 0,3,3, 0,7,7, 5,2,8 5,8,2.

The first Sn atom at the origin is omitted. The second Sn atom is listed first and its coordinates are discretized. Next, the four O atoms are listed. They are sorted by first relative coordinate, then second relative coordinate, then third relative coordinate.

Choosing the second Sn atom to be at the origin, a distinct coordinate string results, following the same ordering protocol:

5,5,5, 0,3,7, 0,7,3, 5,3,3 5,8,8.

Now comparing the two possible basis strings, element by element, we find that the first string is less than the second string. Thus, the standard basis representation is 5,5,5, 0,3,3, 0,7,7, 5,2,8 5,8,2, the maximally ascending element of the orbit of equivalent basis atom coordinates, given the discretization parameter of 10.

To generate similar crystal atomic bases, each integer relative coordinate of each atomic vector can be individually varied ± 1. Further, we include the vector $\pm\langle 1, 1, 1\rangle$ as an additional allowed variation (explicit enumeration in the Appendix, motivation in Proof 8). Together, these generate $8n$ neighbors for any crystal basis. If δ is sufficiently large, these neighbors are similar geometrically to the original basis. In general, $1/\delta$ is roughly related to the phonon mode amplitudes connecting two neighbors.

IV. REPRESENTATION OF COMPLETE CRYSTAL STRUCTURE

By combining the integer specifications of the lattice and the crystal atomic basis, it is possible to completely specify a crystal using only integers. However, the representation of the crystal atomic basis, being in relative coordinates, requires a particular unit cell of the lattice to be defined. Because each lattice is represented by seven integers that are in an unambiguous order, the following formulas [Eqs. (3)–(10)] will always define three unambiguous generators for the lattice. Further, this mapping of vonorms to generators is injective: two distinct strings of vonorms leads to two distinct generating sets (Appendix, Proof 4). Since the crystal's orientation is arbitrary in this paper, a variable orientation arises naturally when using these formulas:

$$\vec{v}_0 = \xi v_0 \langle 1, 0, 0\rangle, \tag{3}$$

$$\vec{v}_1 = \xi v_1 \langle x, y, 0\rangle, \tag{4}$$

$$x = \frac{\vec{v}_0 \cdot \vec{v}_1}{\xi^2 v_0 v_1}, \tag{5}$$

$$y = (1 - x^2)^{1/2}, \tag{6}$$

$$\vec{v}_2 = \xi v_2 \langle a, b, c\rangle, \tag{7}$$

$$a = \frac{\vec{v}_0 \cdot \vec{v}_2}{\xi^2 v_0 v_2}, \tag{8}$$

$$b = \frac{1}{y}\left(\frac{\vec{v}_1 \cdot \vec{v}_2}{\xi^2 v_1 v_2} - xa\right), \tag{9}$$

$$c = (1 - a^2 - b^2)^{1/2}. \tag{10}$$

For conciseness, we group these relations together in a function $\mathbb{L} : \mathbb{Z}^7 \to \mathbb{R}^{3\times3}$. $\mathbb{L}$ takes in the seven integer vonorms as a vector, ordered as v_0^2, v_1^2, v_2^2, v_3^2, $(\vec{v}_0 + \vec{v}_1)^2$, $(\vec{v}_0 + \vec{v}_2)^2$, $(\vec{v}_0 + \vec{v}_3)^2$. $\mathbb{L}$ uses these seven values to calculate the corresponding dot products, using Eq. (1). Finally, $\mathbb{L}$ outputs the three generating vectors as column vectors of a 3 x 3 matrix ordered as follows: $[\vec{v}_0, \vec{v}_1, \vec{v}_2]$.

Now we recall symmetry group S_4, which can permute the order of square vonorms but leaves the geometry of the lattice unchanged. In the context of function $\mathbb{L}$, each $g \in S_4$ can lead to a distinct unit cell and orientation of the lattice. In general, $\mathbb{L}(g(\vec{vo})) = A_g \mathbb{L}(\vec{vo})\mu_g^{-1}$, where A_g is a pure rotation or rotoinversion (unitary) and μ_g is a unimodular matrix (Appendix, Proof 5). Considering this change in unit cell, the relative basis coordinates must be left-multiplied by μ_g to account for the change, whereas, with the affine transformation, A, the relative basis coordinates do not change (Appendix, Proof 6) .

For completeness, we list a generating set of S_4 and their unimodular counterparts μ_g. There is an isomorphism between these groups (Appendix, Proof 7); thus, any permutation and corresponding matrix can be constructed by the analogous group multiplication. We restrict μ_g to have a positive determinant so only proper rotations are required to connect neighboring crystal structures (Appendix, Proof 5). Physically, we base this restriction on the idea that when connecting points in crystallographic phase space, basis atoms are never allowed to invert through the origin. Such an inversion is physically prohibited, as atoms would approach infinitesimally close to one another. Further, inversion of the basis typically represents a large, abrupt change in crystal structure. Our goal here is to consider small crystallographic perturbations, composing them together to naturally reconstruct larger structural transformations:

$$0 \leftrightarrow 1 : \mu_g = -\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{11}$$

$$0 \leftrightarrow 2 : \mu_g = -\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \tag{12}$$

$$0 \leftrightarrow 3 : \mu_g = -\begin{pmatrix} -1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \tag{13}$$

$$1 \leftrightarrow 2 : \mu_g = -\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \tag{14}$$

$$1 \leftrightarrow 3 : \mu_g = -\begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix}, \tag{15}$$

$$2 \leftrightarrow 3 : \mu_g = -\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \end{pmatrix}. \tag{16}$$

Previously, we considered the orbit of vonorms under the action of S_4; however, in the context of a lattice + crystal basis, we need to consider the simultaneous orbit of the atomic basis relative coordinates as well. Here, we define $\tilde{\vec{vo}}$ as the maximally ascending representative of the S_4 orbit on the string of lattice lengths. Some elements in S_4 will leave $\tilde{\vec{vo}}$ unchanged. Together these form a subgroup $\mathbb{H}$ of S_4 (orbit-stabilizer theorem). We can orbit the basis coordinates with all the elements of $\mathbb{H}$, forming the set of all possible geometrically equivalent crystal bases for the given unit cell $\mathbb{L}(\tilde{\vec{vo}})$. For each one of these geometrically equivalent bases, we calculate the corresponding maximally ascending basis coordinate string. Then, comparing these strings to one another, we choose a representative that is globally maximally ascending. Together, the maximally ascending lattice vonorms and this maximally ascending basis string constitute a unique representation of the crystal as a string of numbers, termed here crystal normal form (CNF).

We note that the decision to use integers as a representation removes ambiguity from the calculation of $\mathbb{H}$. If floating-point arithmetic were used, the determination as to which vonorms are equal in value would be difficult and ultimately would rely on a tolerance. Further, because integers are used, we can guarantee that high-symmetry lattices, those with many equal vonorms, will be included and well-represented in the discretization of crystallographic phase space.

Taken together, the aforementioned sets of 42 vonorm covariations alongside the $8n$ crystal basis variations create a set of geometrically similar crystals, termed here crystal neighbors (example in the Appendix). By decreasing ξ and increasing δ, these variations can be made arbitrarily small. When a neighbor is calculated, the vonorms or basis relative coordinates are varied, then the neighbor is put into CNF. This ensures that each crystal is canonically represented and calculated only one time, eliminating redundancies. Under this scheme, it can be shown that neighbor relationships between crystal structures are reciprocal and that any crystal structure can be connected to any other via a series of neighbor connections, aka small structural perturbations (Appendix, Proof 8).

V. A CONSIDERATION OF SUPERCELLS

The framework outlined so far is sufficient to relate any two crystal structures with identical number and type of atoms. In other words, all transformation pathways consisting of strains and a superposition of gamma-point phonon modes can be enumerated for a given crystal. By generating supercells of the corresponding starting and ending crystal structures, it is possible to find additional transformation pathways involving generally different strains and other phonon modes with nonzero wave numbers in the first Brillouin zone. By repeating this procedure for all possible larger supercell sizes, one may theoretically derive all possible transformation pathways. In practice, computational resources qualify the max size that may be considered.

In general, an exhaustive set of supercell-generating matrices can be derived as detailed in the Appendix, Proof 9. From these, one may collect all the geometrically unique supercells by calculating the corresponding integer coordinates in CNF and filtering out duplicates. Because there can be many

supercells with differing lattice and basis geometries, an individual crystal structure may correspond to many distinct coordinate strings when represented as a supercell. Further, any path connecting any two supercells of the different structures is a valid transformation connecting the crystals.

VI. SEARCHING CRYSTALLOGRAPHIC PHASE SPACE

As a review, due to the nature of the crystallographic map, each crystal structure has a minimum of 42 neighbors and can be connected to any other via a series of neighbors. This creates an implicit graph data structure that can be systematically searched to discover pathways between different crystals.

If density-functional theory is used to calculate the energy of a series of neighboring crystal structures, it is possible to find minimal-activation-energy-barrier pathways of arbitrary strain and gamma-point phonon-mode complexity. Further, the full exploration of crystallographic phase space clearly shows the energetic landscape, giving insight into finite-temperature behavior as well. Naturally, crystallographic phase space is generally high dimensional, and a fine grid is required to well-characterize the energy variation over the space. As such, running DFT calculations over a significant region of phase space is too costly for most applications. However, using interatomic potentials or analyzing crystal structures by examining lengths and angles between atoms is more computationally tractable.

Naturally, algorithms like the A^* search algorithm [9] are tempting to use. When the overall cost function corresponds to maximum energy along the transformation pathway, it guarantees finding the lowest activation-energy barrier between the phases, as it is an admissible heuristic search algorithm. If the cost function is changed from maximum energy to other geometric considerations, such as closest-atomic-approach, then one can guarantee paths that satisfy certain geometric invariants. With the appropriate heuristics, this framework could provide guesses of transformation pathways based purely on geometric constraints.

We mention too that there may be other cases where the phase space is too high-dimensional to be computationally tractable. In these cases, development of aggressive heuristics and a greedy search algorithm may speed discovery of transformation pathways but at the cost of guaranteeing the optimal path. Further, noting that the energy landscape is highly nonlinear with respect to atomic displacement, there is a balance when setting parameters ξ and δ. Too fine increases computational cost in finding pathways. Too coarse means potentially skipping over important features of the energy landscape. These are concerns to be addressed in future work.

VII. APPLICATION: ZR CRYSTAL SYSTEM

As an example of applying this framework and highlighting its unique strengths, we study zirconium crystal systems.

Many transition metals have been well-researched in *ab initio* studies, and there are known low-energy canonical pathways connecting various phases, namely, fcc, bcc, and hcp. For the Zr system, studied using density-functional theory, the energies and structures are approximately as follows:

Starting with bcc: $E = -9.016\frac{\text{meV}}{\text{atom}}$,

$$a_{\text{bcc}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} = \frac{a_{\text{bcc}}}{2}\begin{pmatrix}-1&1&1\\1&-1&1\\1&1&-1\end{pmatrix}\begin{pmatrix}2&1&1\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}-1&0&0\\1&0&1\\1&1&0\end{pmatrix},$$

$$\mathbb{B}_{\text{bcc}} = \{\vec{0}, \langle 1/2, 1/2, 1/2\rangle\}, \quad \mathbb{B}_{\text{bcc}} = \{\vec{0}\}, \quad a_{\text{bcc}} = 3.52\,\text{Å}.$$

Then fcc: $E = -9.049\frac{\text{meV}}{\text{atom}}$,

$$a_{\text{fcc}}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} = \frac{a_{\text{fcc}}}{2}\begin{pmatrix}1&1&-1\\-1&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}2&1&0\\0&1&0\\0&0&2\end{pmatrix}\begin{pmatrix}0&-1&0\\1&1&1\\0&0&1\end{pmatrix},$$

$$\mathbb{B}_{\text{fcc}} = \{\vec{0}, \langle 1/2, 1/2, 0\rangle, \langle 1/2, 0, 1/2\rangle, \langle 0, 1/2, 1/2\rangle\},$$

$$\mathbb{B}_{\text{fcc}} = \{\vec{0}\}, \quad a_{\text{fcc}} = 4.46\,\text{Å}.$$

Then hcp: $E = -9.081\frac{\text{meV}}{\text{atom}}$,

$$\begin{pmatrix}a&-a/2&0\\0&a\sqrt{3}/2&0\\0&0&c\end{pmatrix} \quad \mathbb{B}_{\text{hcp}} = \{\vec{0}, \langle 2/3, 1/3, 1/2\rangle\},$$

$$a_{\text{hcp}} = 3.19\text{Å}, \quad \left(\frac{c}{a}\right)_{\text{hcp}} = 1.60.$$

All the above structures represent relaxed structures using the PBE-sol functional, and a force cutoff of 0.003 eV/Å.

A low-energy transformation pathway connecting bcc to fcc is known to be a tetragonal strain, with elongation oriented along one of the canonical cubic cell axes and equal contraction in the perpendicular plane. This is visualized in Fig. 1 for the Zirconium crystal system. A low-energy transformation pathway connecting bcc to hcp is known to be a similar tetragonal strain with contraction and elongation in the perpendicular plane located along the face diagonals. This strain is coupled with a minimal-shuffling phonon mode. The combined strain-phonon energy landscape is visualized in Fig. 2 for the Zirconium crystal system.

By using the A^* search algorithm in conjunction with density-functional theory to calculate static energies, we search the crystallographic phase space of two-atom zirconium unit cells, starting with the stable hcp phase. We then search for target structures: fcc and bcc phases. These phases have primitive unit cells of one atom per unit cell; as such, we search for any geometrically unique index-2 supercell of the fcc or bcc structure.

For illustrative purposes, a single pathway is chosen for analysis, out of the many possible pathways all sharing the same activation energy barrier. To make this choice, all shortest paths connecting the starting and ending structures

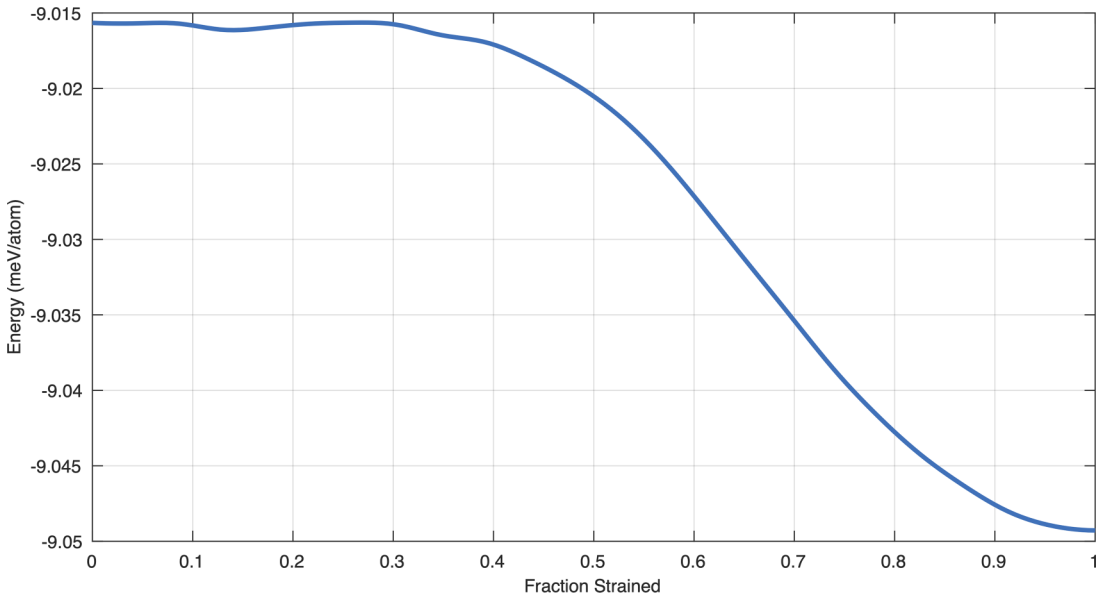

FIG. 1. The canonical strain transformation from bcc to fcc along the tetragonal strain path is visualized, using the calculated minimum-strain pathway. Along the x axis is % transformed in terms of strain. Along the y axis is the energy. This pathway also shows the dynamical instability of the bcc structure at low temperatures.

are considered. Then, of these paths, the unique shortest path whose energy stays the lowest in the first n steps is considered.

We conduct this search using discretization parameters as follows: $\xi = 1.5\,\text{Å}^2$ and $\delta = 30$. With this choice, the hcp, bcc, and fcc structures have coordinates as presented below:

hcp: 7, 7, 17, 24, 7, 24, 24, 10, 20, 15,
bcc: 6, 8, 17, 23, 6, 23, 25, 0, 15, 15,
bcc: 8, 8, 8, 21, 15, 15, 15, 15, 15, 15,
fcc: 6, 7, 13, 25, 13, 18, 20, 15, 15, 15,
fcc: 7, 7, 20, 20, 7, 20, 27, 0, 15, 15.

For the transformation from hcp to bcc, the selected minimum energy transformation pathway proceeds as an interleaved strain + phonon pathway. The bcc crystal structure is higher in energy than the hcp crystal structure, and the transformation pathway from bcc to hcp has a zero activation-energy barrier. The cumulative strains and atomic displacements along the pathway are shown below, starting from the bcc structure, all diagonalized to highlight the principal axes. The coordinate system is the same as that presented in the canonical bcc unit cell setting in the preceding paragraphs. Each strain is shown at 20% transformation pathway

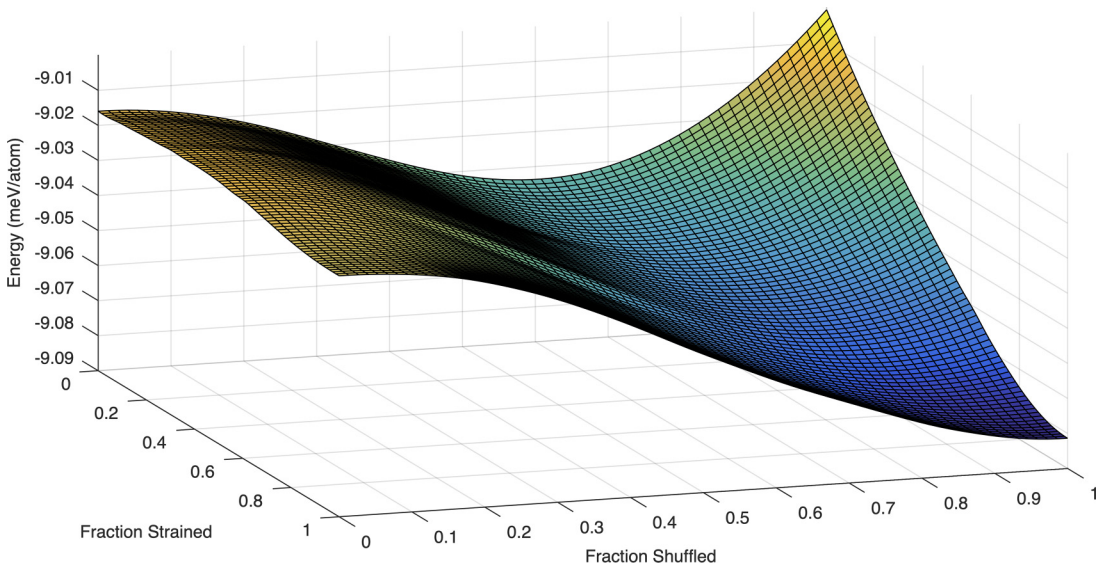

FIG. 2. The canonical strain-phonon transformation from bcc to hcp is visualized, using the calculated minimum-strain, minimal-shuffling pathway. Along the x axis is % transformed in terms of strain. Along the y axis is % transformed in terms of phonon mode. Along the z axis is the energy. This pathway also shows the dynamical instability of the bcc structure at low temperatures.

increments:

Start (bcc): 6, 8, 17, 23, 6, 23, 25, 0, 15, 15.
$E = -9.012 \frac{\text{meV}}{\text{atom}}$.
20%: 6, 6, 20, 25, 6, 25, 26, 0, 15, 15.

$$\epsilon = \begin{pmatrix} -0.057 & 0.533 & 0.845 \\ 0.996 & -0.027 & 0.084 \\ 0.067 & 0.846 & -0.529 \end{pmatrix} \times \begin{pmatrix} -0.150 & 0.000 & 0.000 \\ 0.000 & 0.040 & 0.000 \\ 0.000 & 0.000 & 0.105 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = \vec{0},$$

$E = -9.023 \frac{\text{meV}}{\text{atom}}$.
40%: 6, 7, 19, 25, 6, 25, 26, 2, 16, 15.

$$\epsilon = \begin{pmatrix} -0.000 & -0.707 & -0.707 \\ 1.000 & -0.000 & -0.000 \\ -0.000 & -0.707 & 0.707 \end{pmatrix} \times \begin{pmatrix} -0.080 & 0.000 & 0.000 \\ 0.000 & 0.014 & 0.000 \\ 0.000 & 0.000 & 0.072 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = \langle 2/30, 1/30, 0 \rangle,$$

$E = -9.024 \frac{\text{meV}}{\text{atom}}$.
60%: 6, 7, 19, 25, 6, 25, 26, 6, 18, 15.

$$\epsilon = \begin{pmatrix} -0.000 & -0.707 & -0.707 \\ 1.000 & -0.000 & -0.000 \\ -0.000 & -0.707 & 0.707 \end{pmatrix} \times \begin{pmatrix} -0.080 & 0.000 & 0.000 \\ 0.000 & 0.014 & 0.000 \\ 0.000 & 0.000 & 0.072 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = \langle 6/30, 3/30, 0 \rangle,$$

$E = -9.029 \frac{\text{meV}}{\text{atom}}$.
80%: 6, 7, 18, 25, 7, 24, 25, 9, 17, 15.

$$\epsilon = \begin{pmatrix} -0.180 & -0.707 & 0.684 \\ 0.967 & 0.000 & 0.255 \\ -0.180 & 0.707 & 0.684 \end{pmatrix} \times \begin{pmatrix} -0.092 & 0.000 & 0.000 \\ 0.000 & 0.043 & 0.000 \\ 0.000 & 0.000 & 0.083 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = \langle 9/30, 2/30, 0 \rangle,$$

$E = -9.051 \frac{\text{meV}}{\text{atom}}$.
100% (hcp): 7, 7, 17, 24, 7, 24, 24, 10, 20, 15.

$$\epsilon = \begin{pmatrix} -0.000 & 0.707 & 0.707 \\ 1.000 & 0.000 & 0.000 \\ 0.000 & -0.707 & 0.707 \end{pmatrix} \times \begin{pmatrix} -0.080 & 0.000 & 0.000 \\ 0.000 & 0.014 & 0.000 \\ 0.000 & 0.000 & 0.127 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = \langle 10/30, 5/30, 0 \rangle,$$

$E = -9.075 \frac{\text{meV}}{\text{atom}}$.

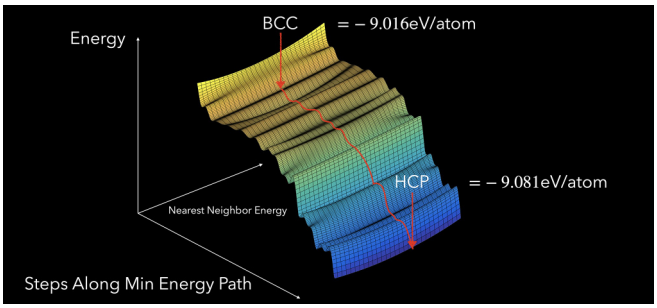

FIG. 3. The search for transformation pathways from bcc to hcp is visualized with the phase space surrounding a path of minimal activation energy barrier highlighted. Along the x axis, the number of steps along the transformation pathway are shown. Along the y axis, two of the lowest energy orthogonal nearest neighbors are shown to the left and right. Along the z axis is the energy. This pathway shows the dynamical instability of the bcc structure at low temperatures.

We note that in contrast to conventional strain-shuffling transformation pathways, characterized by linear interpolation, this pathway proceeds as a superposition of microstrains with distinct principal axes and atomic shuffles in distinct directions. By comparison, it is a more dynamic pathway.

It is difficult to fully visualize the energetic landscape across all degrees of freedom, as it is a ten-dimensional space in this study. However, it is possible to create an energy plot by considering a particular low-energy pathway through the landscape. Along the x axis, we can measure % progression along the pathway and along the z axis we can measure the energy. Then, for each step along the pathway, we can consider the set of nearest neighbors not on the path. Taking the two nearest neighbors of lowest energy and plotting them at ± 1 along the y axis, the plot is then three-dimensional. While not showing the complete information of the landscape, this representation allows one to see whether the minimum energy pathway contains points that are low in energy compared to its neighboring phase space. This plot is shown below for hcp to bcc in Fig. 3 and then hcp to fcc in Fig. 4.

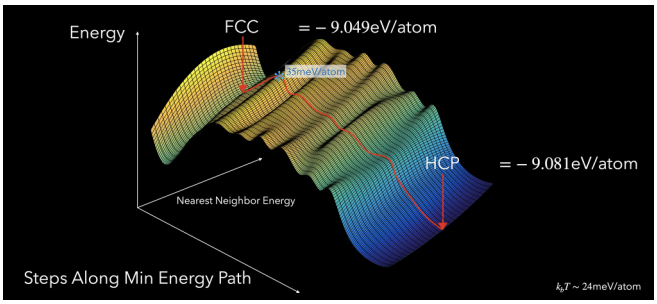

FIG. 4. The search for transformation pathways from fcc to hcp is visualized with the phase space surrounding a path of minimal activation energy barrier highlighted. Along the x axis, the number of steps along the transformation pathway is shown. Along the y axis, two of the lowest energy orthogonal nearest neighbors are shown to the left and right. Along the z axis is the energy. This pathway shows the metastability of the fcc structure at low temperatures.

We contrast this to the calculated energy landscape of the aforementioned canonical bcc–hcp. Along the x axis is % transformed in terms of strain. Along the y axis is % transformed in terms of phonon mode. Along the z axis is the energy.

For the transformation from hcp to fcc, the minimum energy transformation pathway proceeds as an interleaved series of incremental strains and phonon mode steps. The fcc crystal structure is higher in energy than the hcp crystal structure, and the transformation pathway from fcc to hcp has a nonzero activation energy barrier, indicating the metastability of the fcc structure. The energy at this barrier is $-9.015\,\mathrm{meV/atom}$, corresponding to a barrier height of $34\,\mathrm{meV/atom}$ when transforming from fcc to hcp. The crystal structure at this barrier is close to bcc in terms of geometry. The cumulative strains and atomic displacements along the pathway are shown below, starting from the fcc structure. The coordinate system is the same as that presented in the canonical fcc unit cell setting in the preceding paragraphs:

Start (fcc): 7, 7, 20, 20, 7, 20, 27, 0, 15, 15.
$E = -9.029\frac{\mathrm{meV}}{\mathrm{atom}}$.
20%: 6, 9, 17, 19, 6, 20, 25, 0, 15, 15.

$$\epsilon = \begin{pmatrix} 0.144 & 0.759 & -0.635 \\ 0.973 & -0.226 & -0.050 \\ -0.182 & -0.611 & -0.771 \end{pmatrix} \times \begin{pmatrix} -0.151 & 0.000 & 0.000 \\ 0.000 & -0.026 & 0.000 \\ 0.000 & 0.000 & 0.166 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = \vec{0},$$

$E = -9.019\frac{\mathrm{meV}}{\mathrm{atom}}$.
40%: 6, 9, 16, 20, 6, 20, 25, 2, 16, 15.

$$\epsilon = \begin{pmatrix} 0.052 & 0.705 & 0.707 \\ 0.997 & -0.074 & -0.000 \\ -0.052 & -0.705 & 0.707 \end{pmatrix} \times \begin{pmatrix} -0.190 & 0.000 & 0.000 \\ 0.000 & 0.009 & 0.000 \\ 0.000 & 0.000 & 0.164 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = \langle 2/30, 1/30, 0 \rangle,$$

$E = -9.019\frac{\mathrm{meV}}{\mathrm{atom}}$.
60%: 6, 9, 16, 22, 6, 22, 25, 5, 17, 15.

$$\epsilon = \begin{pmatrix} 0.032 & 0.706 & -0.707 \\ 0.999 & -0.045 & 0.000 \\ -0.032 & -0.706 & -0.707 \end{pmatrix} \times \begin{pmatrix} -0.237 & 0.000 & 0.000 \\ 0.000 & 0.081 & 0.000 \\ 0.000 & 0.000 & 0.164 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = \langle 5/30, 2/30, 0 \rangle,$$

$E = -9.019\frac{\mathrm{meV}}{\mathrm{atom}}$.
80%: 7, 7, 17, 23, 7, 23, 24, 6, 18, 15.

$$\epsilon = \begin{pmatrix} -0.120 & 0.707 & -0.697 \\ 0.986 & -0.000 & -0.169 \\ 0.120 & 0.707 & 0.697 \end{pmatrix} \times \begin{pmatrix} -0.143 & 0.000 & 0.000 \\ 0.000 & 0.026 & 0.000 \\ 0.000 & 0.000 & 0.172 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = \langle 6/30, 3/30, 0 \rangle,$$

$E = -9.044\frac{\mathrm{meV}}{\mathrm{atom}}$. 100% (hcp): 7, 7, 17, 24, 7, 24, 24, 10, 20, 15.

$$\epsilon = \begin{pmatrix} -0.102 & 0.707 & 0.700 \\ 0.102 & 0.707 & -0.700 \\ 0.990 & -0.000 & 0.144 \end{pmatrix} \times \begin{pmatrix} -0.164 & 0.000 & 0.000 \\ 0.000 & 0.026 & 0.000 \\ 0.000 & 0.000 & 0.202 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = \langle 10/30, 5/30, 0 \rangle,$$

$E = -9.075\frac{\mathrm{meV}}{\mathrm{atom}}$.

We note that, once again, in contrast to conventional strain transformation pathways, characterized by linear interpolation, this pathway proceeds as a superposition of strains with distinct principal axes and atomic shuffles in multiple directions. By comparison, it is again a more dynamic pathway.

We visualize the local energy environment along the calculated fcc-hcp pathway in the plot shown below.

This identified transformation pathway is comparable to a combined transformation pathway consisting of the canonical tetragonal strain from fcc to bcc and then the canonical strain-shuffle transformation from bcc to hcp. Notably, the activation energy barriers are within 4 meV (measured in a two-atom unit cell), which is insignificant relative to room-temperature thermal vibrations of around $25\frac{\mathrm{meV}}{\mathrm{atom}}$ and limited precision due to large discretization sizes of ξ and δ. The bcc to hcp canonical transformation energetics are shown above. Below, the fcc to bcc transformation is shown in a similar plot. Along the x axis is % transformed in terms of strain. Along the y axis is the energy.

Taken together, these plots illustrate that this formalism is capable of finding additional minimal-energy transformation pathways that differ from known low-energy transformation pathways observed in the literature. More generally, using the delineated algorithm, many other pathways can be enumerated as well. However, all the presently identified pathways and the previously known transformation pathways share the same activation energy barriers. This helps to verify the validity of this formalism for use as a general search algorithm for finding activation energy barriers, as it agrees with previous results.

VIII. APPLICATION: CARBON CRYSTAL SYSTEM

Studying the carbon crystal system, we use this framework to examine the connection between three common crystal systems: diamond, graphite, and lonsdaleite. A choice of primitive unit cells are specified below as well as their CNF coordinate strings, using discretization parameters $\xi = 0.1$ Å and $\delta = 50$.

Diamond:
Lattice: $\mathbb{L} = \{\langle 1.7835\text{Å}, 1.7835\ \text{Å}, 0\rangle, \langle 1.7835\ \text{Å}, 0, 1.7835\ \text{Å}\rangle, \langle 0, 1.7835\ \text{Å}, 1.7835\ \text{Å}\rangle\}$.
Basis: $\mathbb{B} = \{\langle 0, 0, 0\rangle, \langle 1/4, 1/4, 1/4\rangle\}$.
CNF: 63, 64, 64, 64, 64, 64, 127, 25, 12, 37.
Graphite:
Lattice: $\mathbb{L} = \{\langle 1.228\ \text{Å}, 2.127\ \text{Å}, 0\rangle, \langle 1.228\ \text{Å}, -2.127\ \text{Å}, 0\rangle, \langle 0, 0, 6.708\ \text{Å}\rangle\}$.
Basis: $\mathbb{B} = \{\langle 0, 0, 0\rangle, \langle 1/3, 2/3, 0\rangle, \langle 0, 0, 1/2\rangle, \langle 2/3, 1/3, 1/2\rangle\}$.
CNF: 60, 60, 450, 510, 60, 510, 510, 0, 0, 25, 17, 33, 0, 33, 17, 25.
Lonsdaleite:
Lattice: $\mathbb{L} = \{\langle 1.26\ \text{Å}, 2.182\ \text{Å}, 0\rangle, \langle 1.26\ \text{Å}, -2.182\ \text{Å}, 0\rangle, \langle 0, 0, 4.12\ \text{Å}\rangle\}$.
Basis: $\mathbb{B} = \{\langle 1/3, 2/3, 1/16\rangle, \langle 2/3, 1/3, 15/16\rangle, \langle 1/3, 2/3, 7/16\rangle, \langle 2/3, 1/3, 9/16\rangle\}$.
CNF: 63, 64, 170, 233, 64, 233, 233, 0, 0, 19, 33, 16, 25, 33, 17, 44.

Notably, the primitive unit cells of graphite and lonsdaleite have four atoms per unit cell; however, the primitive unit cell of diamond has just two atoms per unit cell. To use this formalism, we must look for connections between the index-2 supercells of diamond and the allotropes of graphite and lonsdaleite so the number of atoms are the same. Referencing Proof 9, there are seven possible index-2 sublattice generating matrices in three dimensions, listed here:

Matrices:
$$\begin{pmatrix}2&0&0\\0&1&0\\0&0&1\end{pmatrix}\ \begin{pmatrix}2&1&0\\0&1&0\\0&0&1\end{pmatrix}\ \begin{pmatrix}2&0&1\\0&1&0\\0&0&1\end{pmatrix}\ \begin{pmatrix}2&1&1\\0&1&0\\0&0&1\end{pmatrix}$$
$$\begin{pmatrix}1&0&0\\0&2&1\\0&0&1\end{pmatrix}\ \begin{pmatrix}1&0&0\\0&2&0\\0&0&1\end{pmatrix}\ \begin{pmatrix}1&0&0\\0&1&0\\0&0&2\end{pmatrix}.$$

When the corresponding sublattice of diamond is created using each matrix, the following crystal structures result, coordinates listed in CNF (same discretization parameters):

Substructures of diamond:
63, 64, 127, 253, 127, 189, 191, 0, 25, 12, 25, 0, 37, 25, 25, 25
63, 64, 127, 253, 127, 189, 191, 0, 25, 13, 25, 0, 38, 25, 25, 25
63, 64, 127, 253, 127, 189, 191, 0, 25, 12, 25, 0, 37, 25, 25, 25
64, 64, 191, 191, 64, 191, 255, 0, 25, 25, 13, 6, 19, 13, 31, 44
64, 64, 191, 191, 64, 191, 255, 0, 25, 25, 13, 6, 19, 13, 31, 44
64, 64, 191, 191, 64, 191, 255, 0, 25, 25, 12, 6, 19, 12, 31, 44
64, 64, 191, 191, 64, 191, 255, 0, 25, 25, 12, 6, 19, 12, 31, 44.

Immediately it is clear that there are only two geometrically distinct diamond substructures, as there are two sets of nearly identical coordinate strings. The small variation is negligible, as they are all equal to one or two neighbor steps from one another. We will term these two substructures diamond A and diamond B, respectively.

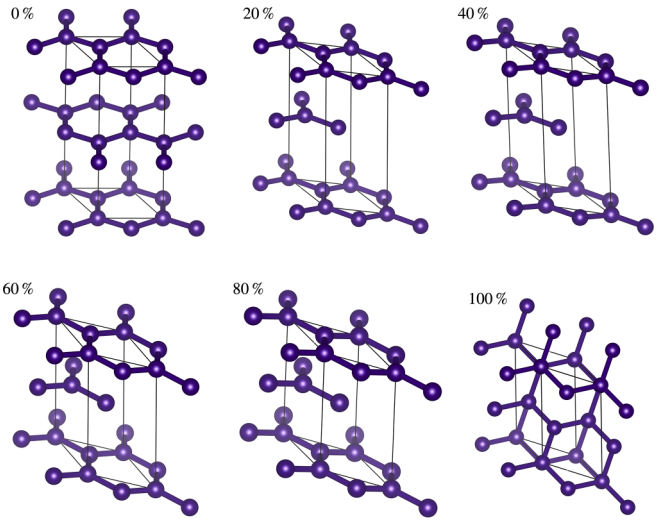

FIG. 5. A transformation pathway from graphite to diamond (B) is visualized with intermediate phases highlighted at 0%, 20%, 40%, 60%, 80%, and 100% transformed.

From each one of these index-2 substructures, it is possible to find pathways connecting diamond to both graphite and lonsdaleite. In this example, we use the square difference in each CNF coordinate as the A^* search heuristic weighting function as opposed to the energy (used in the previous study). Pathways generated in this way will tend to have the fewest number of intermediate neighbors. This corresponds to a minimal amount of overall strain and atomic shuffling, which is a common metric used for conceptualizing possible transformations. We add the additional physical constraint that any pathway in which the carbon atoms approach closer than 1.4 Å is disqualified. This is 90% the equilibrium bond length in diamond. We note that there are many other possible geometric criteria as well as possible interatomic potentials that could be utilized in lieu of this particular choice. Such choices might generate better possible guesses for the suite of minimal-activation-energy-barrier pathways. We use this method here for simplicity, as the pathways can be found quickly, less than one second on modern desktop computers.

We summarize the pathways from graphite and lonsdaleite to diamond (B) below (full pathways in the Appendix). Figure 5 shows the intermediate structures in the transformation from graphite to diamond. Figure 6 shows the intermediate structures from lonsdaleite to diamond.

Start (graphite): 60, 60, 450, 510, 60, 510, 510, 0, 0, 25, 17, 33, 0, 33, 17, 25.

20%: 60, 60, 411, 413, 60, 413, 471, 8, 42, 25, 17, 33, 0, 25, 25, 25.

$$\epsilon = \begin{pmatrix}-0.00&-0.00&1.00\\0.63&-0.78&0.00\\0.78&0.63&0.00\end{pmatrix} \times \begin{pmatrix}-0.13&0.00&0.00\\0.00&0.08&0.00\\0.00&0.00&0.00\end{pmatrix} P^{-1},$$

$$\vec{\delta} = [0.024, -0.11, 0.029][-0.0072, 0.12, -0.015] \times [0.013, -0.12, 0.020][-0.029, 0.10, -0.033].$$

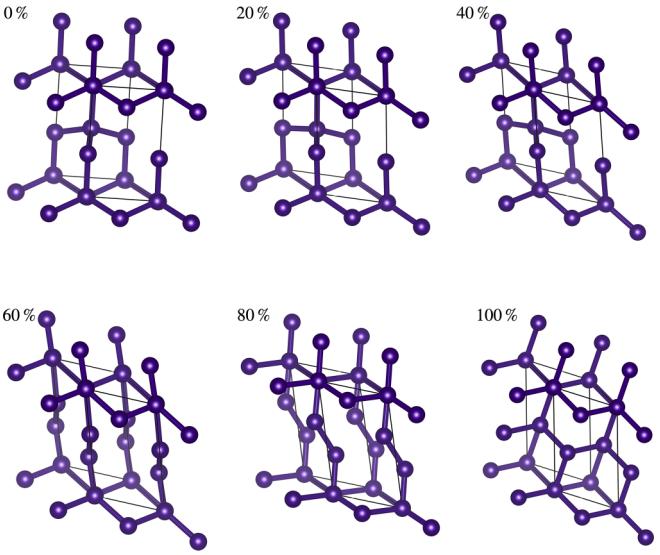

FIG. 6. A transformation pathway from lonsdaleite to diamond (B) is visualized with intermediate phases highlighted at 0%, 20%, 40%, 60%, 80%, and 100% transformed.

40%: 60, 60, 343, 345, 60, 345, 403, 8, 42, 25, 17, 33, 0, 25, 25, 25.

$$\epsilon = \begin{pmatrix} -0.00 & 0.00 & 1.00 \\ 0.00 & 1.00 & 0.00 \\ 1.00 & 0.00 & 0.00 \end{pmatrix} \begin{pmatrix} -0.09 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = [-0.023, -0.0081, -0.015][0.023, 0.0081, 0.015] \times [-0.023, -0.0081, -0.015][0.023, 0.0081, 0.015].$$

60%: 60, 60, 276, 277, 60, 277, 336, 8, 42, 25, 17, 33, 0, 25, 25, 25.

$$\epsilon = \begin{pmatrix} -0.50 & -0.02 & -0.87 \\ 0.87 & -0.01 & -0.50 \\ -0.00 & -1.00 & 0.02 \end{pmatrix} \times \begin{pmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & -0.11 & 0.00 \\ 0.00 & 0.00 & 0.00 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = [-0.028, -0.0084, -0.018][0.027, 0.0083, 0.018] \times [-0.027, -0.0083, -0.018][0.028, 0.0084, 0.018].$$

80%: 60, 60, 208, 209, 60, 209, 268, 8, 42, 25, 17, 33, 0, 25, 25, 25.

$$\epsilon = \begin{pmatrix} 0.00 & 0.00 & 1.00 \\ 0.00 & 1.00 & 0.00 \\ 1.00 & 0.00 & 0.00 \end{pmatrix} \begin{pmatrix} -0.14 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = [-0.036, -0.012, -0.024][0.036, 0.012, 0.024] \times [-0.036, -0.012, -0.024][0.036, 0.012, 0.024].$$

100% (diamond B): 64, 64, 191, 191, 64, 191, 255, 0, 25, 25, 13, 6, 19, 13, 31, 44.

$$\epsilon = \begin{pmatrix} -0.50 & -0.03 & -0.87 \\ 0.87 & -0.02 & -0.50 \\ -0.00 & -1.00 & 0.03 \end{pmatrix} \times \begin{pmatrix} 0.03 & 0.00 & 0.00 \\ 0.00 & -0.05 & 0.00 \\ 0.00 & 0.00 & 0.03 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = [-0.083, 0.060, 0.11][-0.057, -0.10, -0.20] \times [0.083, 0.10, 0.23][0.057, -0.055, -0.14].$$

Start (lonsdaleite): 63, 64, 170, 233, 64, 233, 233, 0, 0, 19, 33, 16, 25, 33, 17, 44.

20%: 63, 64, 175, 208, 64, 207, 239, 0, 1, 19, 16, 33, 24, 17, 34, 44.

$$\epsilon = \begin{pmatrix} -0.04 & 0.04 & 1.00 \\ -0.66 & -0.75 & 0.00 \\ 0.75 & -0.66 & 0.05 \end{pmatrix} \times \begin{pmatrix} 0.08 & 0.00 & 0.00 \\ 0.00 & -0.08 & 0.00 \\ 0.00 & 0.00 & 0.01 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = [0.021, 0.023, 0.037][-0.0019, -0.011, -0.029] \times [-0.013, -0.038, -0.043][-0.0064, 0.026, 0.035].$$

40%: 63, 64, 190, 192, 64, 192, 253, 0, 0, 19, 35, 17, 26, 35, 18, 45.

$$\epsilon = \begin{pmatrix} -0.52 & -0.58 & 0.63 \\ 0.85 & -0.35 & 0.38 \\ 0.00 & 0.74 & 0.68 \end{pmatrix} \times \begin{pmatrix} -0.00 & 0.00 & 0.00 \\ 0.00 & 0.09 & 0.00 \\ 0.00 & 0.00 & -0.08 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = [0.0088, 0.0070, 0.0052][-0.00095, -0.060, -0.034] \times [-0.0090, 0.060, 0.034][0.0012, -0.0070, -0.0052].$$

60%: 63, 64, 191, 191, 64, 191, 254, 4, 8, 27, 8, 22, 13, 15, 33, 43.

$$\epsilon = \begin{pmatrix} -0.50 & -0.64 & -0.59 \\ 0.87 & -0.36 & -0.34 \\ 0.00 & 0.68 & -0.73 \end{pmatrix} \times \begin{pmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & 0.01 & 0.00 \\ 0.00 & 0.00 & -0.01 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = [0.11, 0.12, 0.12][0.085, -0.038, -0.11] \times [-0.055, 0.039, 0.12][-0.13, -0.12, -0.13].$$

80%: 64, 65, 191, 191, 64, 192, 255, 0, 17, 29, 12, 16, 16, 13, 33, 46.

$$\epsilon = \begin{pmatrix} -0.15 & 0.52 & 0.86 \\ 0.09 & 0.81 & -0.49 \\ 0.98 & -0.29 & 0.17 \end{pmatrix} \times \begin{pmatrix} 0.00 & 0.00 & 0.00 \\ 0.00 & -0.00 & 0.00 \\ 0.00 & 0.00 & 0.01 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = [0.040, 0.051, 0.024][0.00013, -0.071, 0.17] \times [-0.020, 0.071, -0.16][-0.020, -0.051, -0.034].$$

100% (diamond B): 64, 64, 191, 191, 64, 191, 255, 0, 25, 25, 13, 6, 19, 13, 31, 44.

$$\epsilon = \begin{pmatrix} 0.00 & -0.49 & 0.87 \\ 0.33 & 0.82 & 0.46 \\ 0.94 & -0.29 & -0.16 \end{pmatrix} \times \begin{pmatrix} -0.00 & 0.00 & 0.00 \\ 0.00 & -0.01 & 0.00 \\ 0.00 & 0.00 & 0.00 \end{pmatrix} P^{-1},$$

$$\vec{\delta} = [-0.015, -0.019, 0.0057][0.065, 0.059, 0.24] \times [-0.075, -0.059, -0.25][0.025, 0.019, 0.0043].$$

IX. SUMMARY

As presented, any crystal structure can be connected to any other crystal structure with the same formula unit via a series of similar neighbors: small incremental strains and phonon mode amplitudes. These neighbor relationships are reciprocal, forming a well-defined grid. Further, each crystal structure can be uniquely represented as a series of integers. As such, the entire 3D crystallographic configuration space, for a fixed number of atoms per unit cell, can be discretized and explored using the parameters ξ and δ.

Employing this approach, we anticipate many possible applications to the field of solid-state physics and materials science. By studying a general collection of crystal structures and evaluating their energies using first-principles techniques, we can gain greater insight into the phase space of materials and assess materials stability, among other functional properties. The larger energetic landscape reveals activation energy barriers separating distinct phases of a material and may elucidate general transformation pathways induced by external fields, such as magnetic, electric, and mechanical energy gradients. The curvature around energetic minima yields second-order response functions, such as elastic tensors and various phonon frequencies. Ultimately, an energy landscape generated using this technique may also enable another way to sample phase space via Monte Carlo approaches. While the orientation of the crystal system has been explicitly discounted in this treatment, with additional effort, the strain and rotation connecting neighboring crystal structures can be calculated using techniques such as Ref. [10]. With these parameters, it is possible to keep track of the overall crystal orientation relative to an external applied field or similar. In this way, coupled properties, such as structural variation in response to external fields or changes in spin state as a response to structural perturbation can be assessed under a consistent global coordinate system.

From a materials engineering perspective, by using this representation in combination with a data set of first-principles materials energies, we anticipate possible utility in fitting interatomic potentials and training machine-learning algorithms. Using an entire energy landscape as a training input would add rich structure-energy information outside of the well-known crystal structures and their perturbations, likely improving the generality of fitted models. Additionally, this technique offers another way to assess crystallographic similarity between materials: the fewer nearest-neighbor hops required, the more similar the crystals. This may prove useful in large databases of crystal structures to quickly screen for duplicates in the database.

ACKNOWLEDGMENTS

This work gratefully acknowledges support from the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Materials Sciences and Engineering Division, for their support under Contract No. DE-AC02-05-CH11231 within the Data Science for Data-Driven Synthesis Science grant (No. KCD2S2).

APPENDIX

1. Proof 1: Nearest-neighbor lattices are connected by small strains

Assume there exist two lattices, $\mathbb{L}_1$ and $\mathbb{L}_2$, with similar square vonorms, differing by a small amount $\sim\xi$.

Assume there exists a transformation A such that $[\mathbb{L}_2] = A[\mathbb{L}_1]$, where $[\mathbb{L}]$ is a 3 x 3 matrix with three generating vectors of $\mathbb{L}$ arranged as column vectors.

Then, it follows that $\langle A\vec{v}_i, A\vec{v}_i\rangle = \langle \vec{v}_i, \vec{v}_i\rangle + \xi_i$, where $\vec{v}_i$ corresponds to one of the seven voronoi vectors, and ξ_i is the difference in the square vonorm value between lattices $\mathbb{L}_1$ and $\mathbb{L}_2$ for the ith square vonorm.

Rearranging $\langle \vec{v}_i, A^TA\vec{v}_i\rangle - \langle \vec{v}_i, \vec{v}_i\rangle = \xi_i$, using the standard definitions of matrix multiplication and conventional cartesian inner product on $\mathbb{R}^n$.

Rearranging $\xi_i = v_i^2\langle \hat{v}_i, (A^TA - I)\hat{v}_i\rangle$, using linearity and the distributive property of the inner product, where $\hat{v}_i$ represents a unit vector in the direction of $\vec{v}_i$.

Assume $A^TA - I$ has an orthonormal eigenbasis $\{\hat{b}_i\}$ with eigenvalues $\{\lambda_i\}$, noting that $A^TA - I$ is real and symmetric, thereby invoking the spectral decomposition theorem.

We can express $\hat{v}_i$ as a linear combination of eigenbasis vectors: $\hat{v}_i = \sum_j c_j\hat{b}_j$.

Rearranging $\frac{\xi_i}{v_i^2} = \sum_{j,k}\langle c_j\hat{b}_j, \lambda_k c_k\hat{b}_k\rangle = \sum_j c_j^2\lambda_j$.

Further, note that $\sum_i c_i^2 = 1$, as $\hat{v}_i$ and $\hat{b}_i$ are all unit vectors.

Thus, ξ_i/v_i^2 is a non-negative weighted average of all λ_i values. When the set of ξ_i/v_i^2 values are near zero for all i, it must follow that for many different weightings, the weighted average of all λ_i values is small. Thus, each λ_i value must be near zero.

Invoking the polar decomposition theorem, $A = RT$, where R is unitary and T is symmetric. Physically, R corresponds to a rotoinversion and T corresponds to a stretch. Thus, the quantity $A^TA - I = T^2 - I$, and the eigenvalues, $\lambda_i = s^2 - 1$, correspond to a square principal stretch value minus 1.

Thus, where λ_i values are small, it follows that the principal stretch values must be near 1, meaning a small strain connects the lattices.

2. Proof 2: Every lattice can be connected to every other lattice via a series of nearest neighbors

We start by referencing Proof 3: every lattice neighbor relationship is reciprocal. Thus, to show lattices $\mathbb{L}_1$ and $\mathbb{L}_2$ are connected by a series of neighbors, it suffices to show that they are both connected to an intermediate lattice $\tilde{\mathbb{L}}$.

With a bit of algebra, it can be shown that the following matrix equation applies for any superbasis, whether or not obtuse:

$$\frac{1}{2}\begin{pmatrix} -1 & -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & -1 & -1 \\ 1 & 0 & 0 & 1 & -1 & -1 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & -1 & 1 & 0 \end{pmatrix}\begin{pmatrix} v_0^2 \\ v_1^2 \\ v_2^2 \\ v_3^2 \\ (\vec{v}_0+\vec{v}_1)^2 \\ (\vec{v}_0+\vec{v}_2)^2 \end{pmatrix} = \begin{pmatrix} \vec{v}_0\cdot\vec{v}_1 \\ \vec{v}_0\cdot\vec{v}_2 \\ \vec{v}_0\cdot\vec{v}_3 \\ \vec{v}_1\cdot\vec{v}_2 \\ \vec{v}_1\cdot\vec{v}_3 \\ \vec{v}_2\cdot\vec{v}_3 \end{pmatrix}.$$

Considering a subset of vonorm modifications $\{\vec{\chi}_i\}$, listed completely in Sec. 10 of the Appendix, we can calculate the image of each modification when multiplied by the above matrix. Notice only the first six square vonorms are listed: the seventh is uniquely implied by Eq. (2):

$$\{\vec{\chi}_i\} = \left\{\begin{pmatrix}0\\0\\1\\0\\1\\0\end{pmatrix},\begin{pmatrix}0\\0\\0\\1\\1\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\\0\\1\end{pmatrix},\begin{pmatrix}0\\0\\0\\1\\0\\1\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\1\\0\\0\\0\end{pmatrix},\begin{pmatrix}1\\0\\0\\0\\0\\0\end{pmatrix},\begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix},\begin{pmatrix}1\\0\\0\\0\\0\\1\end{pmatrix},\begin{pmatrix}0\\0\\1\\0\\0\\1\end{pmatrix},\begin{pmatrix}1\\0\\0\\0\\1\\0\end{pmatrix},\begin{pmatrix}0\\1\\0\\0\\1\\0\end{pmatrix}\right\},$$

$$A\{\vec{\chi}_i\} = \frac{1}{2}\left\{\begin{pmatrix}1\\-1\\0\\-1\\0\\0\end{pmatrix},\begin{pmatrix}1\\0\\-1\\0\\-1\\0\end{pmatrix},\begin{pmatrix}-1\\1\\0\\-1\\0\\0\end{pmatrix},\begin{pmatrix}0\\1\\-1\\0\\0\\-1\end{pmatrix},\begin{pmatrix}-1\\0\\1\\0\\-1\\0\end{pmatrix},\begin{pmatrix}0\\-1\\1\\0\\0\\-1\end{pmatrix},\begin{pmatrix}-1\\-1\\0\\1\\0\\0\end{pmatrix},\ldots\right.$$

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

Noting that the matrix above is invertible, it can be seen that if all dot products $\vec{v}_i \cdot \vec{v}_j$ are zero, then all vonorms must be zero. As such, every 3D lattice must have at least one dot product strictly less than zero.

Examining the images of $\{\vec{\chi}_i\}$ above, every 3D lattice has at least one neighbor that keeps all dot products nonpositive. Further, by combining multiple of these neighbors together in a series, the dot products $\vec{v}_i \cdot \vec{v}_j$ can be made arbitrarily negative. Here we call this specific summation and series of neighbors a canonical obtuse step. Further, when derived in this order, the neighbors, relative to their full S_4 orbit, remain maximally ascending at each step. This is because larger vonorms are incremented first before smaller vonorms are later incremented by the same amount. Thus, the ordering is always unchanged:

$$\vec{\chi}_i + \vec{\chi}_j \ldots = \begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix} + \begin{pmatrix}0\\0\\1\\0\\0\\1\end{pmatrix} + \begin{pmatrix}0\\1\\0\\0\\1\\0\end{pmatrix} + \begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix} + \begin{pmatrix}0\\0\\1\\0\\0\\1\end{pmatrix} + \begin{pmatrix}1\\0\\0\\0\\1\\0\end{pmatrix} + \begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix} + \begin{pmatrix}0\\1\\0\\0\\0\\1\end{pmatrix} + \ldots$$

$$\begin{pmatrix}1\\0\\0\\0\\1\\0\end{pmatrix}+\begin{pmatrix}0\\0\\1\\0\\0\\0\end{pmatrix}+\begin{pmatrix}0\\1\\0\\0\\0\\1\end{pmatrix}+\begin{pmatrix}1\\0\\0\\0\\1\\0\end{pmatrix},$$

$$A(\vec{\chi}_i+\vec{\chi}_j...)=\frac{1}{2}\left(\begin{pmatrix}0\\0\\0\\1\\-1\\-1\end{pmatrix}+\begin{pmatrix}0\\0\\0\\-1\\1\\-1\end{pmatrix}+\begin{pmatrix}0\\0\\0\\-1\\-1\\1\end{pmatrix}+\begin{pmatrix}0\\0\\0\\1\\-1\\-1\end{pmatrix}+\begin{pmatrix}0\\0\\0\\-1\\1\\-1\end{pmatrix}+\begin{pmatrix}0\\-1\\-1\\0\\0\\1\end{pmatrix}+\begin{pmatrix}0\\0\\0\\1\\-1\\-1\end{pmatrix}+...\right.$$

$$\left.\begin{pmatrix}-1\\1\\0\\-1\\0\\0\end{pmatrix}+\begin{pmatrix}0\\-1\\-1\\0\\0\\1\end{pmatrix}+\begin{pmatrix}0\\-1\\1\\0\\0\\-1\end{pmatrix}+\begin{pmatrix}-1\\1\\0\\-1\\0\\0\end{pmatrix}+\begin{pmatrix}0\\-1\\-1\\0\\0\\1\end{pmatrix}\right)=\begin{pmatrix}-1\\-1\\-1\\-1\\-1\\-1\end{pmatrix}.$$

With these preliminaries, we begin the proof.

Take two lattices $\mathbb{L}_1$ and $\mathbb{L}_2$. Begin by adding first any necessary $\vec{\chi}_i$ values from above so all $\vec{v}_i \cdot \vec{v}_j$ values are at most $-3/2$. We will then start with these derived neighbors $\tilde{\mathbb{L}}_1$ and $\tilde{\mathbb{L}}_2$, ensuring that the vonorms are in maximally ascending order. At any time during the course of the steps below, if any dot products of either lattice become -1 or larger, both lattices must be modified by a canonical obtuse step to reduce the dot products. This ensures that specific neighbors can continue to be derived while keeping the superbasis obtuse. Applying a canonical obtuse step to both lattices concurrently will not change the difference between the vonorms of both lattices.

Using $\vec{\chi}_i = \langle 0, 0, 0, 1, 0, 0, 1\rangle$, we construct consecutive neighbors from the lattice with the smaller v_3^2 square vonorm until both lattices have equal values for v_3^2.

Using $\vec{\chi}_i = \langle 0, 0, 1, 0, 0, 0, 1\rangle$, we construct consecutive neighbors from the lattice with the smaller v_2^2 square vonorm until both lattices have equal values for v_2^2.

Using $\vec{\chi}_i = \langle 0, 1, 0, 0, 0, 0, 1\rangle$, we construct consecutive neighbors from the lattice with the smaller v_1^2 square vonorm until both lattices have equal values for v_1^2.

Using $\vec{\chi}_i = \langle 1, 0, 0, 0, 0, 0, 1\rangle$, we construct consecutive neighbors from the lattice with the smaller v_0^2 square vonorm until both lattices have equal values for v_0^2.

At this point, both original lattices $\mathbb{L}_1$ and $\mathbb{L}_2$ have been connected via nearest neighbors to new lattices $\tilde{\mathbb{L}}_1'$ and $\tilde{\mathbb{L}}_2'$. Both of these new lattices have identical primary square vonorm values. Further, the progression above ensures that the primary square vonorms are always sorted in ascending order at each intermediate step. Thus, no permutation changing the order of primary square vonorms occurs.

Using $\vec{\chi}_i = \langle 0, 0, 0, 0, \bar{1}, 0, 1\rangle$, we construct consecutive neighbors from the lattice with the smaller $(\vec{v}_0 + \vec{v}_3)^2$ square vonorm until both lattices have equal values for $(\vec{v}_0 + \vec{v}_3)^2$.

Using $\vec{\chi}_i = \langle 0, 0, 0, 0, \bar{1}, 1, 0\rangle$, we construct consecutive neighbors from the lattice with the smaller $(\vec{v}_0 + \vec{v}_2)^2$ square vonorm until both lattices have equal values for $(\vec{v}_0 + \vec{v}_2)^2$.

At this point, both original lattices $\mathbb{L}_1$ and $\mathbb{L}_2$ have been connected via nearest neighbors to new lattices $\tilde{\mathbb{L}}_1''$ and $\tilde{\mathbb{L}}_2''$. Both of these lattices have identical primary square vonorm values and identical values for $(\vec{v}_0 + \vec{v}_3)^2$ and $(\vec{v}_0 + \vec{v}_2)^2$. Further, the progression above ensures that the secondary square vonorms are always sorted in maximally ascending order at each intermediate step. Thus, no permutation changing the order of secondary square vonorms occurs.

Finally, referring to Eq. (2), the value of six square vonorms determines the value of the remaining vonorm. In this case, because both lattices $\tilde{\mathbb{L}}_1''$ and $\tilde{\mathbb{L}}_2''$ have the same six square vonorms, they must have the same seven square vonorms. Thus, both lattices must be equal. As such, any two lattices $\mathbb{L}_1$ and $\mathbb{L}_2$ can be connected by a series of neighbors to a common lattice. Thus, any two lattices can be connected to one another via nearest neighbors.

3. Proof 3: Lattice neighbor relationships are reciprocal

Assume lattice $\mathbb{L}_A$ has neighbor $\mathbb{L}_B$. Mathematically, this means that when a displacement $\vec{\chi}_i$ is added to the vonorms of $\mathbb{L}_A$ and the vonorms are sorted to be maximally ascending, the vonorms of $\mathbb{L}_B$ result:

$$g(\vec{vo}_A + \vec{\chi}_i) = \vec{vo}_B : g \in S_4,$$

We can rearrange this equation algebraically:

$$g^{-1}(\vec{vo}_B - g(\vec{\chi}_i)) = \vec{vo}_A.$$

Referencing Sec. 10 of the Appendix, where all$\{\vec{\chi}_i\}$ are listed, it can be seen that every permutation $g \in S_4$ is a bijection on this set. That is, by construction, the set is closed under this group action. Further, for each $\vec{\chi}_i$ in the set, its negative is included.

Thus, the action $-g$ maps $\vec{\chi}_i$ onto another element of the set: $\vec{\chi}_j$,

$$-g(\vec{\chi}_i) = \vec{\chi}_j,$$

Thus, it can be said that there exists a displacement such that

$$g(\vec{v}o_B + \vec{\chi}_j) = \vec{v}o_A : g \in S_4.$$

As such, lattice $\mathbb{L}_B$ has neighbor lattice $\mathbb{L}_A$ by definition.

4. Proof 4: The lattice mapping function $\mathbb{L}(\vec{v}\vec{o})$ is injective

For this proof, we reference the following truncated relationship between vector norms and dot products for any superbasis:

$$\frac{1}{2}\begin{pmatrix} -1 & -1 & 0 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & -1 & -1 \\ 1 & 0 & 0 & 1 & -1 & -1 \\ 0 & -1 & 0 & -1 & 0 & 1 \\ 0 & 0 & -1 & -1 & 1 & 0 \end{pmatrix}\begin{pmatrix} v_0^2 \\ v_1^2 \\ v_2^2 \\ v_3^2 \\ (\vec{v}_0+\vec{v}_1)^2 \\ (\vec{v}_0+\vec{v}_2)^2 \end{pmatrix} = \begin{pmatrix} \vec{v}_0\cdot\vec{v}_1 \\ \vec{v}_0\cdot\vec{v}_2 \\ \vec{v}_0\cdot\vec{v}_3 \\ \vec{v}_1\cdot\vec{v}_2 \\ \vec{v}_1\cdot\vec{v}_3 \\ \vec{v}_2\cdot\vec{v}_3 \end{pmatrix}.$$

Assume that $\mathbb{L}(\vec{v}o_1) = \mathbb{L}(\vec{v}o_2)$. Then, the generating vectors output must be identical from the function, and they must have equal lengths and angles.

Noting equal generator lengths, it follows that v_0^2, v_1^2, and v_2^2 must be equal for both input square vonorm vectors $\vec{v}o_1$ and $\vec{v}o_2$.

Noting equal angles, the dot products $\vec{v}_0\cdot\vec{v}_1$, $\vec{v}_0\cdot\vec{v}_2$, and $\vec{v}_1\cdot\vec{v}_2$ must be equal for both input square vonorm vectors $\vec{v}o_1$ and $\vec{v}o_2$.

From above, $\vec{v}_0\cdot\vec{v}_1 = -v_0^2 - v_1^2 + (\vec{v}_0+\vec{v}_1)^2$. Because v_0^2 and v_1^2 are equal for both $\vec{v}o_1$ and $\vec{v}o_2$, it must follow that $(\vec{v}_0+\vec{v}_1)^2$ is equal for both $\vec{v}o_1$ and $\vec{v}o_2$.

From above, $\vec{v}_0\cdot\vec{v}_2 = -v_0^2 - v_2^2 + (\vec{v}_0+\vec{v}_2)^2$. Because v_0^2 and v_2^2 are equal for both $\vec{v}o_1$ and $\vec{v}o_2$, it must follow that $(\vec{v}_0+\vec{v}_2)^2$ is equal for both $\vec{v}o_1$ and $\vec{v}o_2$.

From above, $\vec{v}_1\cdot\vec{v}_2 = v_0^2 + v_3^2 - (\vec{v}_0+\vec{v}_1)^2 - (\vec{v}_0+\vec{v}_2)^2$. Because v_0^2, $(\vec{v}_0+\vec{v}_1)^2$, and $(\vec{v}_0+\vec{v}_2)^2$ are equal for both $\vec{v}o_1$ and $\vec{v}o_2$, it must follow that v_3^2 is equal for both $\vec{v}o_1$ and $\vec{v}o_2$.

Combining these results with Eq. (2), because six square vonorms are equal, all seven square vonorms must be equal between $\vec{v}o_1$ and $\vec{v}o_2$. Thus, $\vec{v}o_1 = \vec{v}o_2$.

5. Proof 5: The group S_4 has an action on set of lattice generators $\mathbb{L}(\vec{v}\vec{o})$ arranged as column vectors. This action is as follows: $\mathbb{L}(g(\vec{v}\vec{o})) = A_g\mathbb{L}(\vec{v}\vec{o})\mu_g^{-1} : g \in S_4$, where A_g is a unitary rotoinversion and μ_g is a unimodular matrix

To show a valid group action, we algebraically note the following, referencing that unitary and unimodular matrices form groups under composition:

$$\begin{aligned}\mathbb{L}(g_1\cdot g_2(\vec{v}o)) &= A_{g_1\cdot g_2}\mathbb{L}(\vec{v}o)\mu_{g_1\cdot g_2}^{-1} = A_{g_1}A_{g_2}\mathbb{L}(\vec{v}o)\mu_{g_2}^{-1}\mu_{g_1}^{-1} \\ &= \mathbb{L}(g_1(g_2(\vec{v}o))), \\ \mathbb{L}(e(\vec{v}o)) &= A_e^{-1}\mathbb{L}(\vec{v}o)\mu_e = \mathbb{L}(\vec{v}o).\end{aligned}$$

As a justification for this definition of group action, consider the action of S_4 on the set of seven square vonorms. The lengths and angles of the superbasis are preserved at all times, just reordered. The geometry of the lattice is preserved; although the mapping function $\mathbb{L}(\vec{v}o)$ may change the orientation of the lattice. This is represented by a left multiplication by a unitary matrix, which preserves lengths and angles. Again, considering the preservation of lattice geometry, the unit cell $\mathbb{L}(\vec{v}o)$ must geometrically be a unit cell of the original lattice. Thus, right multiplication by a unimodular matrix represents this geometric constraint.

Further, S_4 is simply a permutation of the superbasis vectors labels, and it's the closure of the following simple transpositions: $0\leftrightarrow 1$, $0\leftrightarrow 2$, $0\leftrightarrow 3$, $1\leftrightarrow 2$, $1\leftrightarrow 3$, $2\leftrightarrow 3$. If any transposition occurs, the unit cell $\mathbb{L}(\vec{v}o)$, defined geometrically by the generators $\vec{v}_0$, $\vec{v}_1$, and $\vec{v}_2$, must change to reflect the transposition.

In the order, $0\leftrightarrow 1$, $0\leftrightarrow 2$, $0\leftrightarrow 3$, $1\leftrightarrow 2$, $1\leftrightarrow 3$, $2\leftrightarrow 3$, the generating set $\{\vec{v}_0, \vec{v}_1, \vec{v}_2\}$ gets orbited to the following: $\{\vec{v}_1, \vec{v}_0, \vec{v}_2\}$, $\{\vec{v}_2, \vec{v}_1, \vec{v}_0\}$, $\{-\vec{v}_0-\vec{v}_1-\vec{v}_2, \vec{v}_1, \vec{v}_2\}$, $\{\vec{v}_0, \vec{v}_2, \vec{v}_1\}$, $\{\vec{v}_0, -\vec{v}_0-\vec{v}_1-\vec{v}_2, \vec{v}_2\}$, $\{\vec{v}_0, \vec{v}_1, -\vec{v}_0-\vec{v}_1-\vec{v}_2\}$, noting that $\vec{v}_3 = -(\vec{v}_0+\vec{v}_1+\vec{v}_2)$.

By inspection, these transpositions can be accomplished by the right multiplication of the following unimodular matrices:

$$0\leftrightarrow 1 : \mu_g = \pm\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

$$0\leftrightarrow 2 : \mu_g = \pm\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

$$0\leftrightarrow 3 : \mu_g = \pm\begin{pmatrix} -1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix},$$

$$1\leftrightarrow 2 : \mu_g = \pm\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

$$1\leftrightarrow 3 : \mu_g = \pm\begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix},$$

$$2\leftrightarrow 3 : \mu_g = \pm\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \end{pmatrix}.$$

In listing the unimodular matrices, $\pm$ is included to represent a sign ambiguity, as all lattices have inversion symmetry in their group structure. Therefore, a unit cell or its

inversion will have the same geometry and order of square vonorms. However, when we consider the presence of a crystal basis, we need to pick the appropriate sign for which no inversion occurs. Physically, this ensures that each neighbor involves only a small change in basis. Recall that inversion is a large discontinuous change in crystal basis in the general case: when the basis does not also share inversion symmetry.

Rearranging the definition of the group action,

$$A_g = \mathbb{L}(g(\vec{vo}))\mu_g\mathbb{L}(\vec{vo})^{-1},$$

we can take the determinant of both sides:

$$|A_g| = |\mathbb{L}(g(\vec{vo}))||\mu_g||\mathbb{L}(\vec{vo})^{-1}|.$$

We note that, as defined, $\mathbb{L}(\vec{vo})$ always outputs a right-handed set of generators and has a positive determinant.

Thus, it follows that

$$\text{sign}|A_g| = \text{sign}|\mu_g|.$$

The constraint that no inversion occurs signifies that $|A_g| > 0$, thus, $|\mu_g| > 0$.

As such, we pick the negative representatives above so each μ_g has a positive determinant. The set of unimodular matrices with positive determinant likewise forms a group, so the conclusions above are unchanged.

6. Proof 6: A unitary rotoinversion of the crystal does not change the relative coordinates of the crystal basis; however, a unimodular change of unit cell does change the relative coordinates of the crystal basis

Consider a crystal with basis $\mathbb{B}$, represented as a matrix with each basis atom's Cartesian coordinates corresponding to a column vector. Assume $\mathbb{L}$ to be a matrix representing the three generators of a lattice, arranged as column vectors of the matrix. Then we define $[\mathbb{B}]_\mathbb{L} := \mathbb{L}^{-1}\mathbb{B}$, where $[\mathbb{B}]_\mathbb{L}$ is a matrix with each basis atom's relative coordinates corresponding to a column vector.

When a unitary rotoinversion A of the crystal occurs, $\mathbb{L} \to A\mathbb{L}$, $\mathbb{B} \to A\mathbb{B}$. That is, the lattice generators andCartesian basis coordinates are all modified correspondingly. However, calculating the relative basis coordinates, we see they are unchanged:

$$[\mathbb{B}]_\mathbb{L} \to \mathbb{L}^{-1}A^{-1}A\mathbb{B} = \mathbb{L}^{-1}\mathbb{B} = [\mathbb{B}]_\mathbb{L}.$$

When a unimodular matrix μ_g changes the unit cell of the crystal, no geometric change has occurred; thus, the Cartesian basis coordinates remain the same. However, the geometry of the lattice generators has changed: $\mathbb{L} \to \mathbb{L}\mu_g$. Calculating the relative basis coordinates, we see they are changed:

$$[\mathbb{B}]_\mathbb{L} \to \mu_g^{-1}\mathbb{L}^{-1}\mathbb{B} = \mu_g^{-1}[\mathbb{B}]_\mathbb{L}.$$

7. Proof 7: There exists an isomorphism between S_4 and the group of unimodular matrices accomplishing the change of unit cell μ_g

Referencing the group action from Proof 5, the homomorphism from S_4 to $\{\mu_g\}$ is already defined through the generator relations.

Assume that $\mu_g = I$. This implies that $g \in S_4$ does not change the order of $\vec{v}_0$, $\vec{v}_1$, and $\vec{v}_2$. Thus, g can only orbit $\vec{v}_3$ with itself and must leave the other primary voronoi vectors unchanged. Thus, g can only be the identity permutation, and the homomorphism has a trivial kernel. This means the homomorphism is injective.

We take $\{\mu_g\}$ to be the set of all products of the generators defined in Proof 5. As such, for every μ_g there exists $g \in S_4$ that maps to μ_g by the homomorphism. This means the homomorphism is surjective.

Taken together, this means that the homomorphism is an isomorphism (bijection) between S_4 and the set $\{\mu_g\}$.

8. Proof 8: The structure of crystal neighbors is reciprocal and complete

If crystal B is a neighbor of crystal A, one of the following must be true:

(a) If B is a lattice neighbor:

$\vec{vo}_B = g(\vec{vo}_A + \vec{\chi}_i)$, $\mathbb{B}_B = \mu_g\mathbb{B}_A$, $g \in S_4$.

(b) If B is a crystal basis neighbor:

$\vec{vo}_B = \vec{vo}_A$, $\mathbb{B}_B = \mu_g(\mathbb{B}_A + \vec{\delta}_a)$, $g \in S_4$.

We can rearrange the equations as before:

(a) If B is a lattice neighbor:

$\vec{vo}_A = g^{-1}(\vec{vo}_B - g(\vec{\chi}_i))$, $\mathbb{B}_A = \mu_g^{-1}\mathbb{B}_B$, $g \in S_4$.

(b) If B is a crystal basis neighbor:

$\vec{vo}_B = \vec{vo}_A$, $\mathbb{B}_A = \mu_g^{-1}(\mathbb{B}_B - \mu_g\vec{\delta}_a)$, $g \in S_4$.

Inspecting both cases:

(a) $\{\vec{\chi}_i\}$ is closed under S_4 permutation and negation: $-g(\vec{\chi}_i) = \vec{\chi}_j$. Thus, if B is a lattice neighbor of A, A will be a lattice neighbor of B.

(b) It can be shown that the set $\{\vec{\delta}_i\}$ in Sec. 10 of the Appendix is invariant under the action of the generators of $\{\mu_g\}$ and negation: $-\mu_g\vec{\delta}_i = \vec{\delta}_j$. Thus, if B is a crystal basis neighbor of A, A will be a crystal basis neighbor of B.

Therefore, neighbor relationships are reciprocal.

Continuing the proof, we reference Proof 2 to demonstrate that any crystal lattice can be reached as a series of lattice neighbors from any other crystal lattice. Thus, we need only show that for a fixed unit cell reference frame (implying a fixed lattice), all sets of integer basis coordinates can be reached as a series of crystal neighbors. Together this would imply that any combination of lattice and basis can be reached via a series of neighbors.

To demonstrate this, we take two bases within the same unit cells $\mathbb{B}_1$ and $\mathbb{B}_2$. Due to the integer specification and the fact that the displacements $\{\vec{\delta}_i\}$ span the vector space of all atomic displacements, there exists a finite series of displacements (aka basis coordinate neighbors) connecting the two bases: $\mathbb{B}_2 = \mathbb{B}_1 + \sum \vec{\delta}_i$

Now, throughout the algorithm, as displacements occur, the unit cell may shift (though the lattice is unchanged) $\mathbb{L} \to \mathbb{L}\mu_g^{-1}$. Putting the coordinates into CNF dictates these changes. With each shift, the basis coordinates will change: $\mu_g\mathbb{B}_2 = \mu_g\mathbb{B}_1 + \sum \mu_g\vec{\delta}_i$. Because $\{\vec{\delta}_i\}$ is closed with respect to all possible unimodular multiplication, even in the new cell setting a finite path between the starting and ending bases exists, and the next step of the path is always accessible.

Eventually, the path terminates and the bases coincide. This process can be repeated for any $\mathbb{B}_2$ of the original unit cell and, from the perspective of this unit cell, all possible basis relative

coordinates are accessible. This process can also be repeated for any different unit cell and the same results hold, regardless of the unit cell chosen.

9. Proof 9: All geometrically unique supercells of a given index, *N*, can be enumerated by a finite set of upper triangular matrices

Summarizing previous work, given $\tilde{\mathbb{L}} \leqslant \mathbb{L}$, $\tilde{\mathbb{L}}$ generated by $\{\vec{t}_i, ...\}$, $\mathbb{L}$ generated by $\{\vec{b}_i, ...\}$, then $\{\vec{t}_i, ...\}$ can be chosen such that $\{\vec{t}_i, ...\} = \{\vec{b}_i, ...\}A$, where A is upper triangular [11].

Define vectors $\vec{\gamma}_i \in \tilde{\mathbb{L}} : \vec{\gamma}_i = z_1\vec{b}_1 + ... + z_i\vec{b}_i, z_i \neq 0$, $|z_i|$ chosen to be as small as possible.

$\vec{\gamma}_i$ is well-defined, as the vector $\vec{\gamma}_i = N\vec{b}_i$ is guaranteed to fall in $\tilde{\mathbb{L}}$ by construction; however, there may be other combinations of $\vec{b}_i$ that result in a smaller z_i value.

Assume $\exists \vec{c} \in \tilde{\mathbb{L}} : \vec{c} \notin \text{ispan}\{\vec{\gamma}_i\}, \vec{c} = \xi_1\vec{b}_1 + ... + \xi_k\vec{b}_k, k \in \mathbb{Z} \cap [1, n], \xi_k \neq 0$.

Examine the vector $\vec{c} - s\vec{\gamma}_k \in \tilde{\mathbb{L}}, s \in \mathbb{Z}$: $\vec{c} - s\vec{\gamma}_k = (\xi_1 - sz_1)\vec{b}_1 + ... + (\xi_k - sz_k)\vec{b}_k$.

We can choose s to minimize the integer $|\xi_k - sz_k|$; further, we can show that this minimized value is less than $|z_k|$:

$$|\xi_k - sz_k| = |z_k|\left|\frac{\xi_k}{|z_k|} - s\right|.$$

It is possible to choose s so the the term $|\frac{\xi_k}{|z_k|} - s|$ lies in the interval $[0, 1)$; thus, we can say $|\frac{\xi_k}{|z_k|} - s| < 1$, for an appropriate choice of s.

From this, we see that $|\xi_k - sz_k| < |z_k|$.

If $|\xi_k - sz_k| = 0$, then we repeat the above step using $\vec{c} \to \vec{c} - s\vec{\gamma}_k$ and thus a smaller value of k for this next iteration. If this continues until $\vec{c} = \vec{0}$, then $c \in \text{ispan}\{\vec{\gamma}_i\}$, otherwise we will reach the conclusion below.

If $|\xi_k - sz_k| \neq 0$, then $\vec{\gamma}_k$ could not have been chosen to minimize k, a contradiction.

Thus, it must follow that $\vec{c} \in \text{ispan}\{\vec{\gamma}_i\}$.

This shows that the set $\{\vec{\gamma}_i\}$ serves as a generating set for the lattice $\tilde{\mathbb{L}}$. Then, by construction, the $\vec{\gamma}_i$ vectors can be derived from the set $\{\vec{b}_i, ...\}$ via multiplication of the following upper-triangular matrix by construction:

$$[\vec{\gamma}_1, \vec{\gamma}_2, \vec{\gamma}_3] = [\vec{b}_1, \vec{b}_2, \vec{b}_3]\begin{pmatrix} a & b & d \\ 0 & c & e \\ 0 & 0 & f \end{pmatrix}.$$

Here, matrices $[\vec{\gamma}_i]$ and $[\vec{b}_i]$ contain generators $\{\vec{\gamma}_i\}$ and $\{\vec{b}_i\}$, respectively, as column vectors. These can be derived by referencing the previously proven fact that every sublattice can be generated by multiplying the appropriate upper triangular matrix, γ. For a sublattice of index N, the generating matrix must have a determinant equal to N. This places the following constraint on γ's elements:

$$\prod \gamma_{ii} = N.$$

If N has a prime factorization $N = \prod p_i$, where p_i are primes, then all possible γ matrices involve some distribution of these prime factors along γ's diagonal. Thus, to enumerate all possible matrices $\{\gamma\}$, we must first collect all unique ways to parcel these prime factors along the diagonal.

Note that for every γ matrix, $\gamma\mu : \mu \in SL_3(\mathbb{Z})$, generates the same sublattice geometrically. Further, if two γ matrices γ_1 and γ_2 generate the same lattice geometrically then the following is true:

$$[\mathbb{L}]\gamma_1 = [\tilde{\mathbb{L}}],\ [\mathbb{L}]\gamma_2 = [\tilde{\mathbb{L}}]\mu \quad \to \quad \gamma_1^{-1}\gamma_2 = \mu \in SL_3(\mathbb{Z}).$$

We now consider the following transformations for three dimensions:

$$\begin{pmatrix} a & b & d \\ 0 & c & e \\ 0 & 0 & f \end{pmatrix}\begin{pmatrix} 1 & \pm 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b \pm a & d \\ 0 & c & e \\ 0 & 0 & f \end{pmatrix},$$

$$\begin{pmatrix} a & b & d \\ 0 & c & e \\ 0 & 0 & f \end{pmatrix}\begin{pmatrix} 1 & 0 & \pm 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b & d \pm a \\ 0 & c & e \\ 0 & 0 & f \end{pmatrix},$$

$$\begin{pmatrix} a & b & d \\ 0 & c & e \\ 0 & 0 & f \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \pm 1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} a & b & d \\ 0 & c & e \pm c \\ 0 & 0 & f \end{pmatrix}.$$

Clearly, the integer upper triangular elements of the gamma matrices can be put in the interval including zero but excluding the diagonal element of the corresponding row. If the elements fall outside this range, there always exists a unimodular matrix that will transform the elements back into the appropriate range and preserve the geometry of the generated sublattice.

Building on the prime factorization above, for each parcelling of prime factors, it is possible to exhaustively enumerate values of all off-diagonal elements. This collection of matrices may contain geometric duplicates; however, the products $\gamma_a^{-1}\gamma_b$ must be calculated. If any products produce a unimodular matrix, then the matrices γ_a and γ_b produce geometrically duplicate sublattices.

10. Miscellaneous

Here, we write a complete enumeration of lattice neighbor displacements, $\{\vec{\chi}_i\}$ (42 elements):

$$\pm\left\{\begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}, ...\right.$$

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

Here we write a complete enumeration of crystal basis neighbor displacements: $\{\vec{\delta}_i\}$. These displacements can be applied to any basis atom of the crystal:

$$\pm\left\{\begin{pmatrix}1\\0\\0\end{pmatrix}\begin{pmatrix}0\\1\\0\end{pmatrix}\begin{pmatrix}0\\0\\1\end{pmatrix}\begin{pmatrix}1\\1\\1\end{pmatrix}\right\}.$$

Here we calculate all the neighbors of an example crystal in CNF:
The original crystal in CNF:
6 6 15 16 4 19 20 2 10 10.
The original crystal in CNF:
6 6 15 16 4 19 20 2 10 10.

The lattice neighbors of the original crystal:
6 6 15 16 4 18 21 2 10 10
6 6 15 16 4 19 20 10 2 10
6 6 15 16 3 19 21 2 10 10
6 6 15 16 5 19 19 2 10 10
6 6 15 16 3 20 20 2 10 10
6 6 15 16 5 18 20 2 10 10
6 6 15 15 4 19 19 0 8 10
6 6 15 17 4 19 21 2 10 10
6 6 15 15 4 18 20 0 8 10
6 6 15 17 4 20 20 2 10 10
6 6 15 15 3 19 20 0 8 10
6 6 15 17 5 19 20 2 10 10
6 6 14 16 4 19 19 2 10 10
6 6 16 16 4 19 21 0 8 10
6 6 14 16 4 18 20 2 10 10
6 6 16 16 4 20 20 0 8 10
6 6 14 16 3 19 20 2 10 10
6 6 16 16 5 19 20 0 8 10
6 6 14 17 4 19 20 2 10 10
6 6 15 16 4 19 20 0 8 10
5 6 15 16 4 19 19 10 2 10
6 7 15 16 4 19 21 2 10 10
5 6 15 16 4 20 18 10 2 10
6 7 15 16 4 20 20 2 10 10
5 6 15 16 3 20 19 10 2 10
6 7 15 16 5 19 20 2 10 10
5 6 15 17 4 20 19 10 2 10
6 7 15 15 4 19 20 2 10 10
5 6 16 16 4 19 20 0 8 10
6 7 14 16 4 19 20 2 10 10
5 6 15 16 4 19 19 2 10 10
6 7 15 16 4 21 19 10 2 10
5 6 15 16 4 18 20 2 10 10
6 7 15 16 4 20 20 10 2 10
5 6 15 16 3 19 20 2 10 10
6 7 15 16 5 20 19 10 2 10
5 6 15 17 4 19 20 2 10 10
6 7 15 15 4 19 20 0 8 10
5 6 16 16 4 19 20 2 10 10

6 7 14 16 4 20 19 10 2 10
5 7 15 16 4 19 20 2 10 10
5 7 15 16 4 20 19 10 2 10.

The basis neighbors of the original crystal:
6 6 15 16 4 19 20 1 9 9
6 6 15 16 4 19 20 3 11 11
6 6 15 16 4 19 20 2 10 9
6 6 15 16 4 19 20 2 10 11
6 6 15 16 4 19 20 2 9 10
6 6 15 16 4 19 20 2 11 10
6 6 15 16 4 19 20 1 10 10
6 6 15 16 4 19 20 3 10 10.

[1] H. T. Stokes and D. M. Hatch, Procedure for obtaining microscopic mechanisms of reconstructive phase transitions in crystalline solids, Phys. Rev. B **65**, 144114 (2002).

[2] D. Sheppard, P. Xiao, W. Chemelewski, D. D. Johnson, and G. Henkelman, A generalized solid-state nudged elastic band method, J. Chem. Phys. **136**, 074103 (2012).

[3] D. R. Trinkle, D. M. Hatch, H. T. Stokes, R. G. Hennig, and R. C. Albers, Systematic pathway generation and sorting in martensitic transformations: Titanium α to ω, Phys. Rev. B **72**, 014105 (2005).

[4] M. Valle and A. R. Oganov, Crystal fingerprint space–a novel paradigm for studying crystal-structure sets, Acta Crystallogr. Sect. A **66**, 507 (2010).

[5] J. H. Conway and N. J. A. Sloane, Low-dimensional lattices. VI. Voronoi reduction of three-dimensional lattices, Proc. R. Soc. London A **436**, 55 (1992).

[6] C. S. Barrett, P. Cucka, and K. J. A. C. Haefner, The crystal structure of antimony at 4.2, 78, and 298 k, Acta Crystallogr. **16**, 451 (1963).

[7] J. Demmel and Y. Hida, Accurate and efficient floating point summation, SIAM J. Sci. Comput. **25**, 1214 (2004).

[8] T. Yamanaka, R. Kurashima, and J. Mimaki, X-ray diffraction study of bond character of rutile-type SiO_2, GeO_2 and SnO_2, Z. Kristallogr. **215**, 424 (2000).

[9] P. E. Hart, N. J. Nilsson, and B. Raphael, A formal basis for the heuristic determination of minimum cost paths, IEEE Trans. Syst. Sci. Cybern. **4**, 100 (1968).

[10] K. Koumatos and A. Muehlemann, Optimality of general lattice transformations with applications to the Bain strain in steel, Proc. R. Soc. A **472**, 20150865 (2016).

[11] J. William Scott Cassels, *An Introduction to the Geometry of Numbers* (Springer Science & Business Media, 2012), Chap. 1, p. 11.