Title

An Algorithm for Exploring Crystallographic Configuration Space

Author

Mrdjenovich, David James

Publication Date

2022

Peer reviewed|Thesis/dissertation

An Algorithm for Exploring Crystallographic Configuration Space

by

David James Mrdjenovich

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Engineering - Materials Science and Engineering

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Kristin Persson, Chair
Professor Daryl Chrzan
Professor Jeffrey Neaton

Fall 2022

An Algorithm for Exploring Crystallographic Configuration Space

Abstract

An Algorithm for Exploring Crystallographic Configuration Space

by

David James Mrdjenovich

Doctor of Philosophy in Engineering - Materials Science and Engineering

University of California, Berkeley

Professor Kristin Persson, Chair

Structure-property relationships comprise the foundation of the field of Materials Science. When characterizing the structure of a crystalline solid, traditional approaches have focused on using three generating vectors to define a lattice. These vectors, when combined with offsets defining normalized positions of atoms within each identical unit cell, form the complete specification of a crystal. This intuitive definition forms the current standard for the field of materials science. However, this choice of definition is highly-redundant and obscures relationships between different crystal structures. For instance, there exists an infinite number unit cells and atomic offsets that can define the same crystal structure. Further, it is computationally and mathematically difficult to determine the perturbations necessary to transform one crystal structure into another.

Here, outlined in this thesis, I propose a new scheme for specifying crystal structure based off the complementary ideas of a lattice-normal-form and a basis-normal-form. Together, these provide a unique fingerprint for each crystal structure that is a function only of physical geometry. Further, I develop a traversal algorithm that connects together every possible crystal structure via a series of neighbors, each of which represents a small perturbation to the original crystal geometry: either via a small strain or a small phonon mode amplitude.

Building on this framework, I outline a new algorithm for exploring crystallographic phase space to find saddle points, and I re-derive known low-energy transformation pathways in the Zirconium crystal system. Together, this study and this framework show the power of a crystallographic map: a means to systematically explore crystallographic phase space. In this application we use the map to find saddle points; however, it is my hope that this new tool will lead to the elucidation and prediction of many more crystal phenomena.

I dedicate this work to my parents

Contents

List of Figures

List of Tables

Acknowledgments

While my thesis has one author, so many others helped me get to the point of publishing this work.

Firstly, I want to thank my parents: Jamie and Charles. Without their influence, I would have never had the strength and drive that has powered me through this thesis and has now granted me so many opportunities. I have found a lot of personal meaning in this work: discovering a deep mathematical connection to a fundamental form of matter, crystals, has been immensely satisfying for me. I credit my parents for fostering my curiosity and endowing me with an appreciation for the esoteric beauty of this world. In the end, it is because of them that I will be able to find joy through science and learning for the rest of my life. The last nine years have been difficult for us as a family. I'm so grateful that, with time, we have been able to come together and rebuild our relationship. I look forward to making new memories together and enjoying the years to come.

Secondly, I would like to thank my husband, Marcelo. A PhD is difficult during the best of times, and, in the face of a global pandemic and multiple periods of societal lockdown, it becomes nearly impossible. It is because of Marcelo that I was able to finish my dissertation and this PhD. He helped me out of a very difficult headspace, and, with his support, I was able to develop my main findings contained herein. He has brought so much love, happiness, growth, and spontaneity to my life, and I'm looking forward to a beautiful future full of joy and adventure together. To Marcelo: thank you for always being there for me, for showing me an entirely different perspective of the world, for bringing me peace, for showing me a new way to enjoy life. I can't wait to see everything that we can build and experience as the years go on.

Thirdly, I would like to thank my advisor, Kristin. For always making time to meet with me, for always being understanding, having empathy, and caring about my well-being alongside my academic progress, I couldn't be more grateful. We started our journey together seven and a half years ago when she first arrived at Berkeley, and I feel we have been together through such tremendous growth. To Kristin: thank you for always believing in me; thank you for always cheering me on and keeping me positive; thank you for choosing this project for me. I know it was a difficult endeavour, and, at times, it seemed an impossible task. But in many ways, this was the perfect project for me that aligned with my interests and played to my strengths. Thank you for your foresight. I certainly couldn't have finished this PhD without your support and guidance over the years. I will always remember our time and meetings together fondly. I hope we can stay in touch as the years go on.

Fourthly, I would like to thank my friends Raj, Rolando, Ryan, and Mandy. Together all of you collectively kept me sane and filled my heart over the years. In many low moments you were always there to build me back up again, to listen to me vent, to offer advice, and to

remind me that I wasn't alone. The four of you have really taught me what friendship is and what it means to be a good person. I wouldn't be who I am today without your influence. Together, we have laughed; we have cried; we have gone to more yoga classes than I can count; we have mused about the nature of life; we have gone vegan. We made a life together during our time at UC Berkeley, and, it was the collective joy we found and built together that kept me going over the years. Thank you from the bottom of my heart.

And to so many others: Ann, Matt, Guy, Shivani, Rachel, Matt, Vish, and more, a huge thank you. From lunches, coffees, and learning crystallography, to going on many trips together, we made countless memories that will stay with me forever. My life is so much better with all of you in it.

Chapter 1

Introduction

A General Perspective

Mankind's ability to manipulate materials has defined fundamental engineering and societal limitations throughout the most recent geological period. In fact, historians often denote distinct windows of human development by the materials that were used at the time: stone age, bronze age, iron age, atomic age, and silicon age respectively[1][2]. As science and industry have advanced, improved understanding of the connection between atomic structure and materials properties has spurred the invention of countless technologies that enable our modern lifestyle of relative abundance. And meeting the challenges of tomorrow will undoubtedly require even more materials innovation.

A canonical example of materials-based invention is the development of the incandescent lightbulb. Many decades of painstaking work were required before a commercial product was realized, the main bottleneck being the filament [3]. This tiny wire, glowing due to resistive heating, would vaporize from the heat, preventing the bulb from working for any reasonable length of time. To solve the issue, painstaking trial and error was required, until, eventually, it was discovered that Tungsten could best resist the high temperatures. This vignette serves as a metaphor for the historical approach to materials engineering: long periods of expensive experimentation. While many solutions have been found to a large number of engineering problems, many others remain unanswered, waiting on the discovery of a more-performant or less-costly material.

To streamline and improve this process, we need a way to predict materials properties without needing to physically test possibly many thousands of candidate materials. And, with the advent of modern supercomputing resources, it has been possible to use quantum mechanics to predict an ever-increasing number of materials properties, just from knowing the atomic composition and geometry of the material. By simulating properties on a computer, one may test millions of materials in a short period of time and then recommend a handful candidates for confirmational laboratory testing. This not only accelerates materials inno-

vation but also saves material resources [4].

While promising, this approach of using computational simulations to predict materials properties is itself limited. The input to the simulation is often a perfect, periodic crystal and the output is a corresponding energy. Real materials always contain defects, which alter the energy and properties. Engineering applications often involve the interaction between different materials, such as at an interface. Further, the precision and accuracy of the energy may be lower than expected due to the physical approximations used in the calculation. An increasing number of these complications can be addressed, but even so, there remain challenges in the interpretation of the results.

Fundamentally, simulations give the energy as a function of crystal geometry, and materials properties need to be inferred from this variation, termed here the "energy landscape". It is well known, based on the principal of thermodynamic energy minimization, that local minima in this landscape are the only possible observable crystal structures for a fixed collection of atoms [5]. By measuring the local curvature of this minima, the change in energy as the material is deformed, it is possible to gauge materials strength, thermal expansion, and more.

However, there remain many other non-local properties that can be inferred from the landscape, and, of particular importance, are allotropic phase transformations and finite-temperature properties. In order to study these, one must understand the full variation of energy as a function of crystal geometry to know the location of saddle points in the landscape or the overall shape and dimensions of the local minimum energy well. While this is potentially calculable using today's techniques, it is an effort ultimately stymied by the complexities of crystal geometry. As such, this thesis will address these complexities, providing a "crystallographic map", an algorithm to systematically and completely explore the geometric configuration space of all crystal structures in a deterministic, discretized, and straightforward fashion.

In the subfield of computational materials science I hope this algorithm may assist in various distinct areas:

1. Increasing the community's understanding of the underlying geometry and mathematics of crystallography.

2. The coding of crystallographic search algorithms: e.g. what are all the different ways that a crystal can distort to transform from one structure to another? Possibly sorted-by a certain value metric or subject to certain constraints.

3. The unambiguous geometric numeric representation of a crystal structure for purposes of materials structure databases and possible machine learning applications.

4. The construction of accurate, interpolatable energetic landscape databases as a function of geometry.

5. The discovery of candidate activation energy barriers separating different crystal structures.

6. Insight into finite temperature properties based on a Monte-Carlo sampling of this energetic landscape.

Relationship to the Field

Exploring atomic mechanisms of allotropic phase changes, including ferroelectric and piezoelectric transformations, has long been an important avenue for designing precision electronics and assessing materials stability. While it is possible to simulate any crystalline configuration of atoms, there is no guarantee that such a configuration is physically realizable in a laboratory. This serves as a fundamental limitation for translating computational results into real-world outcomes.

However, a great step forward would be finding some way to pre-identify which materials certainly cannot be made. If two phases of different energies are separated by a small or zero activation energy barrier then it can be thermodynamically guaranteed that the higher-energy phase will transform into the lower-energy phase; thus, the higher-energy phase cannot exist. These activation energy barriers appear as distant saddle points in the energy landscape, hence, the importance of exploring the larger landscape more completely.

In the past, various approaches have been employed by the materials science community to find saddle points within the larger phase space. These include the minimal-atomic-motion ansatz [6], symmetry group-subgroup relationships [7], and the generalized solid-state nudged-elastic-band algorithm [8]. However these all have a common limitation: they explore a small line or narrow family of one-dimensional lines through this crystallographic phase space. Ultimately, there is no guarantee that any of the techniques will explore the region of phase space that contains the actual lowest activation energy barrier. Hence, a technique that surveys the entire landscape is a significant improvement relative to past approaches as it is conceptually guaranteed to identify the minimum-energy saddle point between two phases.

In general, much work has been done calculating various strain and phonon pathways connecting two crystal structures. There exist, for example, techniques calculating strains [9] connecting two lattices or total atomic displacements connecting two crystallographic motifs. However, this once again only captures a small region of the total crystallographic phase space, generally a one or two dimensional cross section. With a full parameterization of the lattice and crystal motif degrees of freedom, one can explore any combination of phonon modes and micro-strains in any order, building them up incrementally to connect any two

crystal structures. This provides for a much richer picture and potential understanding of the atomic mechanism underlying allotropic phase changes.

As computational materials science has burgeoned, so have databases of materials properties, storing crystal structures and their associated properties. However, the unique representation of a crystal structure is itself a challenging problem due to arbitrary orientation and origin, arbitrary choice of unit cell geometry, arbitrary choice of periodicity, arbitrary ordering of identical atoms, and floating point round-off errors. The approach outlined in this thesis, will be useful for solving all of these ambiguities with the added bonus of a numerically stable discretization that connects all crystals via a series of small strains and phonon mode amplitudes. Ultimately this thesis defines the vector space of all crystal structures and defines an asymmetric unit within this space, a region sufficient to parameterize all crystal structures with a given chemical formula unit. I anticipate this may prove useful for machine learning and energy prediction in the future as well.

Zirconium Crystal System

Transformation pathways between various transition metal phases have been well-studied in the literature, principally due to the fact that an analysis of the high symmetry and simple crystal motif reveals many straightforward strains and atomic displacements. Zirconium is no exception to this trend. As such, many canonical transformation pathways can be constructed to study the relative phase stability of each phase.

In this thesis, Zr will serve as an example system, for assessing relative phase stability with a new approach of exhaustively and methodically searching crystallographic phase space. To perform validation later on in this thesis, canonical transformation pathways will be illustrated and enumerated below as a reference. We begin with an overview of the phases, unit cells, and coordinate systems to be used then enumerate pathways, progressing from simple to complex.

The crystal structures can be represented by three lattice generators arranged as cartesian column vectors, first in the canonical Bravais lattice setting and then in the primitive unit cell defined by 3 Voronoi vectors. They are related via multiplication of a sublattice generating matrix alongside a unimodular matrix. The relaxed volume of the unit cell is given for Zr. The crystallographic motifs are listed below first for the Bravais setting then for the primitive setting where applicable.

Starting with BCC: $E = -9.016\frac{meV}{atom}$

$$a_{BCC}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} = \frac{a_{BCC}}{2}\begin{pmatrix}-1&1&1\\1&-1&1\\1&1&-1\end{pmatrix}\begin{pmatrix}2&1&1\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}-1&0&0\\1&0&1\\1&1&0\end{pmatrix}$$

$$\mathbb{B}_{BCC} = \{\vec{0}, < 1/2, 1/2, 1/2 >\} \qquad \mathbb{B}_{BCC} = \{\vec{0}\} \qquad a_{BCC} = 3.52\mathring{A}$$

Then FCC: $E = -9.049\frac{meV}{atom}$

$$a_{FCC}\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} = \frac{a_{FCC}}{2}\begin{pmatrix}1&1&-1\\-1&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}2&1&0\\0&1&0\\0&0&2\end{pmatrix}\begin{pmatrix}0&-1&0\\1&1&1\\0&0&1\end{pmatrix}$$

$$\mathbb{B}_{FCC} = \{\vec{0}, < 1/2, 1/2, 0 >, < 1/2, 0, 1/2 >, < 0, 1/2, 1/2 >\}$$
$$\mathbb{B}_{FCC} = \{\vec{0}\} \qquad a_{FCC} = 4.46\mathring{A}$$

Then HCP: $E = -9.081\frac{meV}{atom}$

$$\begin{pmatrix}a&-a/2&0\\0&a\sqrt{3}/2&0\\0&0&c\end{pmatrix} \qquad \mathbb{B}_{HCP} = \{\vec{0}, < 2/3, 1/3, 1/2 >\}$$

$$a_{HCP} = 3.19\mathring{A} \qquad \left(\frac{c}{a}\right)_{HCP} = 1.60$$

All above structures represent relaxed structures using the PBE-sol functional, and a force cut-off of $0.003\text{eV}/\mathring{A}$

The known low-energy transformation pathways connecting BCC to the other phases, when represented in cartesian coordinates as a strain followed by rotation, involve two distinct types of strains: tetragonal and trigonal strains. The tetragonal (4) strain involves principal strains oriented along the x, y, and z axes (all 4-fold symmetry axes), one of which is an elongation and two of which are an equal contraction. The trigonal (3) strain involves principal strains oriented along the $< 1, 1, 1 >$ axis (a 3-fold symmetry axis) and two of which are an equal contraction in the orthogonal plane. A general form of these matrices is presented below with the corresponding matrix diagonalization to showcase the principal axes.

$$\epsilon_4 = \begin{pmatrix} \beta_4 & 0 & 0 \\ 0 & \beta_4 & 0 \\ 0 & 0 & \alpha_4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta_4 & 0 & 0 \\ 0 & \beta_4 & 0 \\ 0 & 0 & \alpha_4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}^{-1}$$

$$\epsilon_3 = \frac{1}{3} \begin{pmatrix} \alpha_3 + 2\beta_3 & \alpha_3 - \beta_3 & \alpha_3 - \beta_3 \\ \alpha_3 - \beta_3 & \alpha_3 + 2\beta_3 & \alpha_3 - \beta_3 \\ \alpha_3 - \beta_3 & \alpha_3 - \beta_3 & \alpha_3 + 2\beta_3 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix} \begin{pmatrix} \alpha_3 & 0 & 0 \\ 0 & \beta_3 & 0 \\ 0 & 0 & \beta_3 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{pmatrix}^{-1}$$

For these general strains ϵ_4 and ϵ_3, the parameters, α_4, β_4, α_3, and β_3 can be tuned to yield pure-strain transformation pathways between specific BCC and FCC crystal structures.

Starting with the tetragonal strain, we note that the canonical BCC crystal structure is already in a body-centered tetragonal setting, and the FCC crystal structure can be put into a body-centered tetragonal setting via multiplication below, first a 45^o rotation clockwise about the $\hat{z}$ axis acting on the primitive unit cell multiplied by a sub-lattice generating matrix and an accompanying unimodular matrix:

$$\frac{a_{FCC}}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \times \frac{a_{FCC}}{2} \begin{pmatrix} 1 & 1 & -1 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

In this setting, then, the tetragonal strain is computed as follows:

$$(1 + \epsilon_4)\, a_{BCC} I = \frac{a_{FCC}}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} \qquad \epsilon_4 = \frac{a_{FCC}}{a_{BCC}\sqrt{2}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \sqrt{2} \end{pmatrix} - I$$

Thus,

$$\beta_4 = \frac{a_{FCC}}{a_{BCC}\sqrt{2}} - 1 \qquad \alpha_4 = \frac{a_{FCC}}{a_{BCC}} - 1$$

Then for the trigonal strain, we put both the BCC and FCC crystal structures in an obverse rhombohedral setting. This can be accomplished by multiplying the primitive unit cells by the following matrices:

$$\frac{a_{BCC}}{2}\begin{pmatrix}2&0&1\\-2&2&1\\0&-2&1\end{pmatrix}=\frac{a_{BCC}}{2}\begin{pmatrix}-1&1&1\\1&-1&1\\1&1&-1\end{pmatrix}\begin{pmatrix}3&2&2\\0&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}-1&0&-1\\1&-1&1\\0&1&1\end{pmatrix}$$

$$\frac{a_{FCC}}{2}\begin{pmatrix}1&0&2\\-1&1&2\\0&-1&2\end{pmatrix}=\frac{a_{FCC}}{2}\begin{pmatrix}1&1&-1\\-1&1&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&0&0\\0&3&0\\0&0&1\end{pmatrix}\begin{pmatrix}1&-1&1\\0&0&1\\0&-1&2\end{pmatrix}$$

In these settings, then, the trigonal strain is computed as follows:

$$(1+\epsilon_3)\frac{a_{BCC}}{2}\begin{pmatrix}2&0&1\\-2&2&1\\0&-2&1\end{pmatrix}=\frac{a_{FCC}}{2}\begin{pmatrix}1&0&2\\-1&1&2\\0&-1&2\end{pmatrix}\qquad\epsilon_3=\frac{a_{FCC}}{a_{BCC}}\begin{pmatrix}1&1/2&1/2\\1/2&1&1/2\\1/2&1/2&1\end{pmatrix}-I$$

Thus,

$$\beta_3=\frac{a_{FCC}}{6a_{BCC}}-\frac{1}{3}\qquad\alpha_3=\frac{2a_{FCC}}{3a_{BCC}}-\frac{1}{3}$$

Note that for the obverse rhombohedral setting, both BCC and FCC crystal structures have the following basis:

$$\mathbb{B}_{BCC,Rhom}=\mathbb{B}_{FCC,Rhom}=\{\vec{0},<1/3,2/3,1/3>,<2/3,1/3,2/3>\}$$

Finally, the BCC phase can transform into the HCP phase, via a combined strain + phonon path involving a quasi-tetragonal strain along $\hat{x}$ and a transverse phonon mode with wave vector $\vec{k}$, with coordinates shown in the reciprocal primitive lattice of BCC, cartesian coordinates, and the modified BCC lattice setting:

$$[\vec{k}]_{\mathbb{L}^*_{BCC}}=<1/2,0,0>\qquad\vec{k}\propto<0,1,1>\qquad[\vec{k}]_{\mathbb{L}'_{BCC}}\propto<0,0,1>$$

Here, the transformation A is broken down into a rotation matrix times a stretch matrix. Modified BCC lattice and basis settings are shown to demonstrate strain connection. Finally the strain is diagonalized to gain insight into its essential action.

$$A=\begin{pmatrix}-0.500&0.612&-0.612\\0.866&0.354&-0.354\\0&-0.707&-0.707\end{pmatrix}\times\begin{pmatrix}0.905&0&0\\0&1.07&-0.041\\0&-0.041&1.07\end{pmatrix}$$

$$[\mathbb{L}_{BCC}]' = \frac{a_{BCC}}{2}\begin{pmatrix}-1 & 1 & 1\\ 1 & -1 & 1\\ 1 & 1 & -1\end{pmatrix}\begin{pmatrix}2 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}\begin{pmatrix}0 & 0 & -1\\ -1 & 1 & -1\\ 0 & 1 & -1\end{pmatrix} = \frac{a_{BCC}}{2}\begin{pmatrix}-1 & 2 & 0\\ 1 & 0 & -2\\ -1 & 0 & -2\end{pmatrix}$$

$$A[\mathbb{L}_{BCC}]' = \begin{pmatrix}3.19 & -1.59 & 0\\ 0 & 2.76 & 0\\ 0 & 0 & 5.11\end{pmatrix} = [\mathbb{L}_{HCP}]$$

$$\mathbb{B}'_{BCC} = \{\vec{0}, < 0, 1/2, 1/2 >\} \qquad \mathbb{B}_{HCP} = \{\vec{0}, < 2/3, 1/3, 1/2 >\}$$

$$\epsilon = \begin{pmatrix}0 & 0 & 1\\ \sqrt{2}/2 & \sqrt{2}/2 & 0\\ \sqrt{2}/2 & -\sqrt{2}/2 & 0\end{pmatrix}\begin{pmatrix}0.03 & 0 & 0\\ 0 & 0.11 & 0\\ 0 & 0 & -0.10\end{pmatrix}\begin{pmatrix}0 & 0 & 1\\ \sqrt{2}/2 & \sqrt{2}/2 & 0\\ \sqrt{2}/2 & -\sqrt{2}/2 & 0\end{pmatrix}^{-1}$$

Choosing the perpendicular direction $\vec{v} =< 0, 1, -1 >$ for the polarization of the phonon mode, the bases can be transformed into one another: $[\vec{v}]_{\mathbb{L}'_{BCC}} =< 2, 1, 0 >$. When scaled by $1/6$ and subtracted from basis vector $< 0, 1/2, 1/2 >$, the HCP basis results.

The transformation from FCC to HCP can also be modeled here as a combination of two pathways. First performing the tetragonal BCC to FCC transformation in reverse and then performing the BCC to HCP transformation. Note that this hypothetical pathway is not mentioned in the literature, particularly since the FCC phase is not observed; however, it will be shown later in Study 2 that this pathway has a minimum activation energy barrier.

In all of these cases, the strain matrix and phonon mode amplitudes can be scaled by a factor $\lambda \in [0, 1]$ to linearly interpolate the transformation pathway. By sampling the energy for different values of λ, it is possible to derive the activation energy barrier for the path along the corresponding degrees of freedom. Because relative coordinate displacements commute with the action of strains, all intermediate values of partial strain and partial phonon mode action are well-defined.

Experimentally, the HCP phase of Zirconium is observed at standard temperature and pressure. Upon heating to 1135K at standard pressure, Zirconium undergoes a phase transformation to BCC. The FCC phase is not observed as a stable structure under a wide variety of pressure and temperature differences. [10].

Chapter 2

Methods

Mathematical Background

Group and Representation Theories

In this section we briefly review the salient mathematical results pertinent to the techniques used in this thesis. For a full treatment, please refer to [11].

We begin with the definition of a group. A group $\mathbb{G}$ is a set of elements with an associative bijection $\circ : \mathbb{G} \times \mathbb{G} \to \mathbb{G}$. $\mathbb{G}$ must contain an identity element $e : e \circ g = g \circ e = g \ \forall g \in \mathbb{G}$. For every element $g \in \mathbb{G}$ there must exist a corresponding inverse element, denoted g^{-1}, such that: $g \circ g^{-1} = g^{-1} \circ g = e$.

A group may be abelian, meaning that its products are commutative: $g_1 \circ g_2 = g_2 \circ g_1 \ \forall g_1, g_2 \in \mathbb{G}$.

Given two groups, $\mathbb{G}_1$ and $\mathbb{G}_2$, a direct-product group can be derived: $\mathbb{G} := \mathbb{G}_1 \times \mathbb{G}_2$ with associated natural bijection: $\circ : (a_1, a_2) \circ (b_1, b_2) = (a_1 \circ b_1, a_2 \circ b_2)$.

The elements in $\mathbb{G}$ each have a corresponding order. For element g, its order corresponds to the minimum number of times it must be multiplied by itself until the identity element e is reached. This can be written as $g^n = e$, $n = |g|$, for some minimal value of n. For a direct-product group, the order of element (g_1, g_2) is equal to $\text{lcm}(|g_1|, |g_2|)$, where lcm stands for least-common-multiple. If e is never reached $|g| := \infty$.

A group $\mathbb{G}$ can have a subgroup $\mathbb{H}$, meaning that $\mathbb{H}$ is a subset of $\mathbb{G}$, and $\circ$, $\mathbb{G}$'s mapping function, maps all elements in $\mathbb{H}$ onto other elements in $\mathbb{H}$, $\circ : \mathbb{H} \times \mathbb{H} \to \mathbb{H}$. When this occurs, cosets can be defined as either $\mathbb{H}_g := \{h \circ g | h \in \mathbb{H}\}, g \in \mathbb{G}$ or ${}_g\mathbb{H} := \{g \circ h | h \in \mathbb{H}\}, g \in \mathbb{G}$. The distinct cosets of $\mathbb{H}$ naturally form a partition on $\mathbb{G}$'s elements with equally-sized equivalence classes. In this case equivalence corresponds to coset membership: $a \sim b \to a, b \in \mathbb{H}_g$ for

some $g \in \mathbb{G}$.

A subgroup $\mathbb{H}$ of $\mathbb{G}$ can be termed normal if $\forall g \in \mathbb{G}, h \in \mathbb{H}, g \circ h \circ g^{-1} = h' : h' \in \mathbb{H}$. This is often denoted as $\mathbb{H} \trianglelefteq \mathbb{G}$. The cosets of a normal subgroup themselves form a group, known as a factor group, here denoted as $\mathbb{G}/\mathbb{H}$. The products of this factor group can be calculated as follows: $\mathbb{H}_{g_1} \circ \mathbb{H}_{g_2} = \mathbb{H}_{g_1 \circ g_2}$. If a subgroup $\mathbb{H} \leq \mathbb{G}$, contains exactly half of $\mathbb{G}$'s elements, then $\mathbb{H}$ is automatically a normal subgroup, and $\mathbb{H}$'s cosets automatically form a factor group. If $\mathbb{G}$ is abelian, $\mathbb{H}$ is automatically normal.

Two groups $\mathbb{G}_1$ and $\mathbb{G}_2$ can be connected via a homomorphism $f : \mathbb{G}_1 \to \mathbb{G}_2, f(x \circ y) = f(x) \circ f(y)$. Thus, the homomorphism, in some sense, preserves the multiplication action of the groups. In general, one may examine the image of f: $\text{img}f := \{f(g) | g \in \mathbb{G}_1\}$, which is a subgroup of $\mathbb{G}_2$. One may also examine the kernel of f: $\ker f := \{g \in \mathbb{G}_1 | f(g) = e\}$, which is a normal subgroup of $\mathbb{G}_1$. When the image of f equals $\mathbb{G}_2$ then f is surjective. When the kernel of f is trivial, then f is injective. When f is bijective it is called an isomorphism, and when $\mathbb{G}_1 = \mathbb{G}_2$, an isomorphism is called an automorphism.

A group $\mathbb{G}$ may "act" on an arbitrary set Ω through a defined group action $f : \mathbb{G} \times \Omega \to \Omega : f(g_1 \circ g_2, \omega) = f(g_1, f(g_2, \omega)) \ \forall \omega \in \Omega, g_1, g_2 \in \mathbb{G}$. In this sense, the objects of Ω get interchanged by elements of $\mathbb{G}$ in a manner that is reflective of the underlying group structure of $\mathbb{G}$. In general, one may examine the orbit of an element $\omega \in \Omega$, defined as $\{f(g, \omega) | g \in \mathbb{G}\}$. This corresponds to all elements of Ω to which f maps ω. This orbit establishes an equivalence relationship and natural equivalence classes on Ω corresponding to orbit membership. One may also examine the stabilizer of an element $\omega \in \Omega$, defined as $\{g \in \mathbb{G} | f(g, \omega) = \omega\}$. The stabilizer of an element forms a subgroup of $\mathbb{G}$. A coset ${}_g\text{stab}(\omega)$ corresponds to all elements $\tilde{g} \in \mathbb{G} : \tilde{g}(\omega) = g(\omega)$.

Given a group $\mathbb{G}$ and a vector space $\mathbb{V}$, a representation can be defined as a homomorphism $\mathbb{G} \to \hat{A}(\mathbb{V})$, where $\hat{A}$ is a group of linear (matrix) operators defined on $\mathbb{V}$. For each $g \in \mathbb{G}$, we define A_g to be the corresponding operator in $\hat{A}(\mathbb{V})$. In general, one may always define an appropriate inner product on $\mathbb{V}$ so that the corresponding matrix representation of the linear operators is unitary. Alternatively, for a given inner product, the matrices of the representation can all be conjugated by a particular matrix B, effectively choosing a unitary representation.

One may view the collective action of the matrix group as dividing the vector space $\mathbb{V}$ into a collection of invariant subspaces. In general, a representation is termed reducible if there exists a non-zero subspace $\tilde{\mathbb{V}} \leq \mathbb{V}$ such that $\forall \vec{v} \in \tilde{\mathbb{V}}, A_g \vec{v} \in \tilde{\mathbb{V}} \ \forall g \in \mathbb{G}$. The action of any representation can be analyzed on increasingly small subspaces of $\mathbb{V}$ until no such non-zero invariant subspace can be found. These minimal subspaces and the action of the matrices on these spaces are termed irreducible representations.

Irreducible representations have a number of useful properties. Of premier importance in this thesis is the orthogonality relationship among irreducible representation matrix elements (i, j) of non-equivalent irreducible representations (α and β). This can be summarized as follows:

$$< (A_g^{(\alpha)})_{ij}, (A_g^{(\beta)})_{kl} > = \frac{1}{|\mathbb{G}|} \sum_{g \in \mathbb{G}} (A_g^{(\alpha)})_{ij} (A_g^{(\beta)})^*_{kl} = \delta_{i,k} \delta_{j,l} \delta_{\alpha,\beta} \tag{2.1}$$

It can be shown that for abelian groups, the resulting irreducible representations are one-dimensional. In this case, matrices reduce down to trivial scaling action on these one-dimensional subspaces. This will be used several times in the following derivation of the Voronoi representation of lattices.

Lattice Theory

The backbone of this thesis, the numerical representation of the space of 3D lattices, rests on the original work of Conway and Sloane [12]. While an infinite number of generating sets exist for a lattice, each lattice can be represented uniquely by a Wigner-Seitz unit cell and, as will be shown, an obtuse superbasis. This construction is an essential component to any 3D lattice's unambiguous representation. I believe that it is of great pedagogical value to outline the main findings of this work alongside my own extensions in one cogent presentation below. As such, I will distinguish provided theorems and proofs from Conway and Sloane with an explicit citation and leave my own theorems and proofs without any markings.

We begin with the definition of a Wigner-Seitz cell of a lattice. The Wigner-Seitz cell $V(\vec{l})$ of a given lattice point $\vec{l}$ in lattice $\mathbb{L}$ is defined as the set of points in $\mathbb{R}^n$ that are closer to or equally-close to $\vec{l}$ as compared to all other lattice points. We can define this set as follows:

$$V\left(\vec{l}\right) := \left\{ \vec{x} \in \mathbb{R}^n : \left| \vec{x} - \vec{l} \right| \leq |\vec{x} - \vec{v}| \, \forall \vec{v} \in \mathbb{L} \right\} \tag{2.2}$$

Building off this definition, we can then define a Voronoi vector of the lattice. Consider a vector $\vec{v}$ of the lattice $\mathbb{L}$. We can construct a perpendicular plane to this vector that bisects $\vec{v}$. If this plane intersects with the origin-centered Wigner-Seitz cell at any point, then $\vec{v}$ is a Voronoi vector. We can define the set of Voronoi vectors as follows:

$$\left\{ \vec{v} \in \mathbb{L} : \left(V(\vec{0}) \cap \left\{ \vec{x} \in \mathbb{R}^3 : \vec{x} \cdot \vec{v} = \frac{1}{2} \vec{v} \cdot \vec{v} \right\} \right) \neq \emptyset \right\} \tag{2.3}$$

Note that that, in contrast to the crystallographic convention, using this definition leads to neighboring Wigner-Seitz cells sharing points on the boundary. We can subdivide Voronoi vectors into two distinct camps: relevant and irrelevant Voronoi vectors. In the relevant case, the perpendicular bisector of the Voronoi vector and the origin-centered Wigner-Seitz cell

share a common "face" (3D). In the case of n dimensions, a "face" is a locus of points that must be defined by a linear combination of $n-1$ linearly-independent vectors. Otherwise the Voronoi vector is irrelevant.

To analyze the Voronoi vectors of a lattice $\mathbb{L}$ we need to define a related construction, $\mathbb{L}$'s double lattice $2\mathbb{L}$, the collection of all vectors of $\mathbb{L}$ scaled by a factor of two:

$$2\mathbb{L} := \{2\vec{v} : \vec{v} \in \mathbb{L}\} \tag{2.4}$$

By this definition, $2\mathbb{L} \trianglelefteq \mathbb{L}$: $2\mathbb{L}$ is a normal subgroup of $\mathbb{L}$. This implies that there exists a factor group $\mathbb{L}/2\mathbb{L}$ with natural equivalence classes within $\mathbb{L}$.

Proof. We note that $2\mathbb{L}$ is a subset of $\mathbb{L}$ by construction; thus, if it is a valid subgroup then it must also be abelian, as $\mathbb{L}$ is itself abelian.

Consider $\vec{v} \in 2\mathbb{L} \rightarrow \exists \vec{w} \in \mathbb{L} | \vec{v} = 2\vec{w}$
Then $\exists -\vec{w} \in \mathbb{L} \rightarrow \exists -2\vec{w} \in 2\mathbb{L}, -2\vec{w} = -\vec{v} \in 2\mathbb{L}$

Consider $\vec{v_1}, \vec{v_2} \in 2\mathbb{L} \rightarrow \exists \vec{w_1}, \vec{w_2} \in \mathbb{L} : \vec{v_1} = 2\vec{w_1}, \vec{v_2} = 2\vec{w_2}$
Then $\exists \vec{w_1} + \vec{w_2} \in \mathbb{L} \rightarrow \exists 2(\vec{w_1} + \vec{w_2}) \in 2\mathbb{L} \rightarrow \exists \vec{v_1} + \vec{v_2} \in 2\mathbb{L}$ □

From these preliminaries, we prove two theorems about Voronoi vectors building to an important conclusion.

Theorem 2.0.1. *A non-zero vector $\vec{v}$ in lattice $\mathbb{L}$ is a Voronoi vector if and only if $\vec{v}$ is a shortest vector in its $2\mathbb{L}$ equivalence class [12]*

Proof. [12] Assume there exists a non-zero Voronoi vector $\vec{v}$ in lattice $\mathbb{L}$ and there exists another vector $\vec{w} \in \mathbb{L} : |\vec{w}| < |\vec{v}|, \vec{w} \in (2\mathbb{L})_{\vec{v}}$

By definition of the coset, $\vec{v} - \vec{w} \in 2\mathbb{L}$, $\vec{v} - \vec{w} + 2\vec{w} = \vec{v} + \vec{w} \in 2\mathbb{L}$

We then define vectors: $\vec{u} := \frac{1}{2}(\vec{v} - \vec{w}) \in \mathbb{L}$, $\vec{t} := \frac{1}{2}(\vec{v} + \vec{w}) \in \mathbb{L}$

Because $\vec{v}$ is a Voronoi vector, we can define a vector $\vec{p}$ that is both within the origin-centered Wigner-Seitz cell and in $\vec{v}$'s perpendicular bisector: $\exists \vec{p} \in \mathbb{R}^n : \vec{p} \cdot \vec{v} = \frac{1}{2}\vec{v} \cdot \vec{v}$, $\vec{p} \in V(\vec{0})$

Notice that $\vec{p} \in V(\vec{0})$ by definition means that: $\left|\vec{p} - \vec{0}\right| \leq \left|\vec{p} - \vec{l}\right| \quad \forall \vec{l} \in \mathbb{L}$.

Then by some algebra: $\vec{p} \cdot \vec{l} \leq \frac{1}{2}\vec{l} \cdot \vec{l} \quad \forall \vec{l} \in \mathbb{L}$.

Applying the above conclusions, we can state the following:

$\vec{p}\cdot\vec{v}=\frac{1}{2}\vec{v}\cdot\vec{v}$, $\vec{p}\cdot\vec{w}\leq\frac{1}{2}\vec{w}\cdot\vec{w}$, $\vec{p}\cdot\vec{u}\leq\frac{1}{2}\vec{u}\cdot\vec{u}$, $\vec{p}\cdot\vec{t}=\frac{1}{2}\vec{t}\cdot\vec{t}$

Now, using algebra, we can compute the quantity: $\vec{p}\cdot\vec{u}+\vec{p}\cdot\vec{t}=\frac{1}{2}\vec{v}\cdot\vec{v}$.

We can also use the inequalities above to show: $\vec{p}\cdot\vec{u}+\vec{p}\cdot\vec{t}\leq\frac{1}{4}\left(\vec{v}\cdot\vec{v}+\vec{w}\cdot\vec{w}\right)$.
Together these imply: $\vec{v}\cdot\vec{v}\leq\vec{w}\cdot\vec{w}$, a contradiction to the original premise. □

Proof. [12] Assume that $\vec{v}$ is a shortest vector in the coset $(2\mathbb{L})_{\vec{v}}$ but that $\vec{v}$ is not a Voronoi vector.

Applying the definitions above, this implies that for some $\vec{l}\in\mathbb{L}$,

$$\left|\vec{x}-\vec{0}\right|>\left|\vec{x}-\vec{l}\right|\forall\left\{\vec{x}\in\mathbb{R}^n:\vec{x}\cdot\vec{v}=\frac{1}{2}\vec{v}\cdot\vec{v}\right\}$$

Re-arranging the inequality, $\vec{x}\cdot\vec{l}>\frac{1}{2}\vec{l}\cdot\vec{l}$

If we consider the special case where $\vec{x}=\frac{1}{2}\vec{v}$, $\frac{1}{2}\vec{v}\cdot\vec{l}>\frac{1}{2}\vec{l}\cdot\vec{l}$

If we consider then another vector in the $(2\mathbb{L})_{\vec{v}}$ coset, $\vec{v}-2\vec{l}$, we can compute its length: $(\vec{v}-2\vec{l})\cdot(\vec{v}-2\vec{l})$

Using the inequality above and some algebra, $(\vec{v}-2\vec{l})\cdot(\vec{v}-2\vec{l})<\vec{v}\cdot\vec{v}$, a contradiction to the original premise. □

Theorem 2.0.2. *A non-zero vector $\vec{v}$ in lattice $\mathbb{L}$ is a relevant Voronoi vector if and only if $\vec{v}$ and $-\vec{v}$ are the ONLY shortest vectors in the coset $(2\mathbb{L})_{\vec{v}}$ [12]*

Proof. Assume there exists a relevant Voronoi vector $\vec{v}$ with an additional shortest vector $\vec{w}\neq-\vec{v}$ in its $(2\mathbb{L})_{\vec{v}}$ equivalence class.

We define vectors: $\vec{u}:=\frac{1}{2}\left(\vec{v}-\vec{w}\right)\in\mathbb{L}$, $\vec{t}:=\frac{1}{2}\left(\vec{v}+\vec{w}\right)\in\mathbb{L}$

We define a set of vectors, $\vec{x}\in\mathbb{R}^n:\vec{x}\cdot\vec{u}\leq\frac{1}{2}\vec{u}\cdot\vec{u},\vec{x}\cdot\vec{t}\leq\frac{1}{2}\vec{t}\cdot\vec{t}$

Now we can compute the quantity: $\vec{x}\cdot\vec{u}+\vec{x}\cdot\vec{t}=\vec{x}\cdot\vec{v}$

Using the inequalities and definitions above: $\vec{x}\cdot\vec{v}\leq\frac{1}{4}\left(\vec{v}\cdot\vec{v}+\vec{w}\cdot\vec{w}\right)=\frac{1}{2}\vec{v}\cdot\vec{v}=\frac{1}{2}\vec{w}\cdot\vec{w}$, noting that both $\vec{v}$ and $\vec{w}$ have the same length, both being shortest vectors in $(2\mathbb{L})_{\vec{v}}$

Analyzing the results above, we see that all points $\vec{x}$ are automatically inside the the bisector

of $\vec{v}$ simply by being inside the bisectors of both $\vec{u}$ and $\vec{t}$. In fact, as shown by the inequalities, the intersection of the bisectors of $\vec{u}$ and $\vec{t}$ coincides with the bisector of $\vec{v}$. This means that $\vec{v}$ cannot be a relevant Voronoi vector, a contradiction to the original premise. □

Proof. Assume that $\pm\vec{v}$ are the only shortest vectors in the coset $(2\mathbb{L})_{\vec{v}}$, but that $\vec{v}$ is not a relevant Voronoi vector.

This implies that $\frac{1}{2}\vec{v}$ lies on or outside another Voronoi vector's bisecting plane. We'll define this other Voronoi vector as $\vec{w}$.

This condition implies: $\frac{1}{2}\vec{v}\cdot\vec{w} \geq \frac{1}{2}\vec{w}\cdot\vec{w}$

We then compute the dot product: $(\vec{v}-2\vec{w})\cdot(\vec{v}-2\vec{w}) \leq \vec{v}\cdot\vec{v}$, a contradiction to the original premise. □

Combining these theorems together, the conclusion is striking. Noting that for an n-dimensional lattice $|\mathbb{L}/2\mathbb{L}| = 2^n$, there are then at most $2^n - 1$ relevant Voronoi vectors for any lattice. This can be shown by constructing a binary string of n digits and computing the number of distinct possibly binary strings. Each distinct string represents a sum of distinct generators of $\mathbb{L}$ that won't be included in $2\mathbb{L}$ and is in a distinct coset. For the 3-dimensional application, this means that 7 vectors and their negatives are at most are required to define any Wigner-Seitz cell of any lattice, and, often, this number will be reduced when the lattice has a non-trivial holohedry.

To gain more insight into the Voronoi vectors of a lattice, it is necessary to introduce another related construction. Any lattice of n dimensions is defined by n linearly-independent generating vectors, forming a generating set: $\{\vec{v_1}, ..., \vec{v_n}\}$. To this set we append an additional vector: $\vec{v_0} := -\sum_i \vec{v_i}$. Together these $n+1$ vectors form a "superbasis" for the lattice. Just as a lattice has an infinite number of generating sets, it also has an infinite number of superbases.

A superbasis has a number of nice properties worth mentioning. By construction, a permutational symmetry exists: any of the original generating vectors and the augmentation vector $\vec{v_0}$ can be interchanged to form another valid generating set for the lattice and another valid superbasis. Further, the lengths and angles of the superbasis are related, a relationship that can be shown algebraically. The specific relationship for 3D space is shown below:

Proof. Consider the transposition of $\vec{v_0}$ with any other generating set vector $\vec{v_i}$.

Formatting the generating set of $\mathbb{L}$ as an $n \times n$ matrix, this transposition is accomplished via a right matrix multiplication, a matrix A that is the identity matrix with one column

replaced by all -1 entries.

A is a unimodular matrix, as all entries are integers and the determinant is necessarily ± 1; thus, the set is still a generating set after the transposition. □

Proof. Consider the transposition of $\vec{v_0}$ with any other generating set vector $\vec{v_i}$.
Re-calculating the quantity: $-\sum_x \vec{v_x} = \vec{v_i}$; thus, the set is still a valid superbasis. □

Proof. Consider the quantities below:

$$\begin{aligned}
\vec{v_0}\cdot\vec{v_0} = \vec{v_0}\cdot(-\vec{v_1}-\vec{v_2}-\vec{v_3}) &= -\vec{v_0}\cdot\vec{v_1}-\vec{v_0}\cdot\vec{v_2}-\vec{v_0}\cdot\vec{v_3}\\
\vec{v_1}\cdot\vec{v_1} = \vec{v_1}\cdot(-\vec{v_0}-\vec{v_2}-\vec{v_3}) &= -\vec{v_0}\cdot\vec{v_1}-\vec{v_1}\cdot\vec{v_2}-\vec{v_1}\cdot\vec{v_3}\\
\vec{v_2}\cdot\vec{v_2} = \vec{v_2}\cdot(-\vec{v_0}-\vec{v_1}-\vec{v_3}) &= -\vec{v_0}\cdot\vec{v_2}-\vec{v_1}\cdot\vec{v_2}-\vec{v_2}\cdot\vec{v_3}\\
\vec{v_3}\cdot\vec{v_3} = \vec{v_3}\cdot(-\vec{v_0}-\vec{v_1}-\vec{v_2}) &= -\vec{v_0}\cdot\vec{v_3}-\vec{v_1}\cdot\vec{v_3}-\vec{v_2}\cdot\vec{v_3}\\
(\vec{v_0}+\vec{v_1})\cdot(\vec{v_0}+\vec{v_1}) = (\vec{v_0}+\vec{v_1})\cdot(-\vec{v_2}-\vec{v_3}) &= -\vec{v_0}\cdot\vec{v_2}-\vec{v_0}\cdot\vec{v_3}-\vec{v_1}\cdot\vec{v_2}-\vec{v_1}\cdot\vec{v_3}\\
(\vec{v_0}+\vec{v_2})\cdot(\vec{v_0}+\vec{v_2}) = (\vec{v_0}+\vec{v_2})\cdot(-\vec{v_1}-\vec{v_3}) &= -\vec{v_0}\cdot\vec{v_1}-\vec{v_0}\cdot\vec{v_3}-\vec{v_1}\cdot\vec{v_2}-\vec{v_2}\cdot\vec{v_3}
\end{aligned}$$

These can be summarized in the following matrix equation:

$$\begin{pmatrix}
-1 & -1 & -1 & 0 & 0 & 0\\
-1 & 0 & 0 & -1 & -1 & 0\\
0 & -1 & 0 & -1 & 0 & -1\\
0 & 0 & -1 & 0 & -1 & -1\\
0 & -1 & -1 & -1 & -1 & 0\\
-1 & 0 & -1 & -1 & 0 & -1
\end{pmatrix}
\begin{pmatrix}
\vec{v_0}\cdot\vec{v_1}\\
\vec{v_0}\cdot\vec{v_2}\\
\vec{v_0}\cdot\vec{v_3}\\
\vec{v_1}\cdot\vec{v_2}\\
\vec{v_1}\cdot\vec{v_3}\\
\vec{v_2}\cdot\vec{v_3}
\end{pmatrix}
=
\begin{pmatrix}
v_0^2\\
v_1^2\\
v_2^2\\
v_3^2\\
(\vec{v_0}+\vec{v_1})^2\\
(\vec{v_0}+\vec{v_2})^2
\end{pmatrix}
\tag{2.5}$$

□

With these preliminaries, we now show an important connection between a non-strictly-obtuse superbasis of a lattice and its Voronoi vectors. A non-strictly-obtuse superbasis is one in which all dot products between the vectors are less than or equal to zero. And this special construction directly reveals the Voronoi vectors of the lattice. In particular, each vector of the non-strictly-obtuse superbasis is a Voronoi vector, alongside all non-zero sub-sums of these vectors.

Theorem 2.0.3. *Given lattice $\mathbb{L}$ with a non-strictly-obtuse superbasis $\tilde{\mathbb{B}} := \{\vec{v_i} : i \in \mathbb{Z} \cap [0, n]\}$, certain sums of these vectors correspond to Voronoi vectors. Specifically, the set $\left\{\vec{v_s} : s \subseteq \{0, ..., n\}, |s| \in (0, n), \vec{v_s} := \sum_{i \in s} \vec{v_i}\right\}$ gives the $2(2^n - 1)$ Voronoi vectors of $\mathbb{L}$, a superset of its relevant Voronoi vectors [12]*

Proof. [12] Consider $\vec{v} = \sum_i m_i \vec{v_i}, m_i \in \mathbb{Z}$.

We compute the length of $\vec{v}$ and rearrange the result:

$$\vec{v} \cdot \vec{v} = \sum_i \sum_j m_i m_j \vec{v_i} \cdot \vec{v_j} = \sum_i m_i^2 \vec{v_i} \cdot \vec{v_i} + \sum_{i \neq j} m_i m_j \vec{v_i} \cdot \vec{v_j}$$

Replacing the self-dot-product:

$$\vec{v} \cdot \vec{v} = \sum_i \sum_{j \neq i} m_i m_j \vec{v_i} \cdot \vec{v_j} - m_i^2 \vec{v_i} \cdot \vec{v_j}$$

Re-indexing the summation:

$$\vec{v} \cdot \vec{v} = \sum_{i<j} 2 m_i m_j \vec{v_i} \cdot \vec{v_j} - m_i^2 \vec{v_i} \cdot \vec{v_j} - m_j^2 \vec{v_i} \cdot \vec{v_j}$$

Combining terms:

$$\vec{v} \cdot \vec{v} = \sum_{i<j} -\vec{v_i} \cdot \vec{v_j} (m_i - m_j)^2$$

Analyzing this expression, we note that because $\tilde{\mathbb{B}}$ is non-strictly-obtuse, each $\vec{v_i} \cdot \vec{v_j} \leq 0$.

Consider shifting any combination of m_i coefficients by even integers $(2z_i)$:

$$\vec{v} \cdot \vec{v} = \sum_{i<j} -\vec{v_i} \cdot \vec{v_j} \left((m_i - m_j)^2 + (2z_i - 2z_j)^2 + 2(m_i - m_j)(2z_i - 2z_j)\right)$$

Without loss of generality, assume that each m_i coefficient is either 0 or 1. Then the terms in the above expression are as follows:

$$\vec{v} \cdot \vec{v} = \sum_{i<j} -\vec{v_i} \cdot \vec{v_j} \left(\{0, 1\} + \Delta^2 + 2\{-1, 0, 1\}\Delta\right), \ \Delta \in 2\mathbb{N}_0$$

When presented in this form, the sum is clearly minimized when Δ is 0 or possibly 2 (equally-minimized). All other values of Δ increase the length of $\vec{v}$.

Because every vector in $2\mathbb{L}$ must have even coordinates m_i, shifting m_i values by all multiples

of 2 will orbit through all members of the $(2\mathbb{L})_{\vec{v}}$ coset.

Thus, when Δ is 0, $\vec{v}$ is a shortest vector in $(2\mathbb{L})_{\vec{v}}$ and therefore a Voronoi vector.
When Δ is 0, then each m_i is 0 or 1, with $\vec{v}$ corresponding to a subsum as described in the theorem.

Consider changing subset S for its compliment S^c. In this case every m_i equal to 0 changes to 1, and every m_i equal to 1 changes to 0. Noting that $\sum \vec{v_i} = 0$, this change transforms $\vec{v}$ to $-\vec{v}$, but the length is unchanged. Thus, $\vec{v}_{S^c} = -\vec{v}_S$, and they are part of the same $2\mathbb{L}$ coset.

Consider $S_1 \neq S_2, S_2^c$ and the corresponding coordinates m_i for $\vec{v}_{S_1}$ and $\vec{v}_{S_2}$. These coordinates will be unequal and no possible modification of $m_i \pm 2$ can get the coordinates to coincide; thus, $\vec{v_{S_1}}$ and $\vec{v_{S_2}}$ are in different $2\mathbb{L}$ cosets. □

We can comment further on the properties of an obtuse superbasis of a lattice by noticing whether the dot products are strictly non-zero. In general, at most three dot products of a superbasis can be zero in 3D space, otherwise some lengths of the superbasis would have to be zero. This is revealed via the matrix equation [eq. 2.5], relating the lengths and angles of a superbasis in 3D space.

Theorem 2.0.4. *Given a lattice $\mathbb{L}$ with an obtuse superbasis $\tilde{\mathbb{B}}$, the Voronoi vectors in each coset are relevant if and only if $\vec{v_i} \cdot \vec{v_j} < 0 \quad \forall i, j$ [12]*

Proof. [12] Referencing the above proof, the length of a vector $\vec{v_S}$ is given by the following:

$$\vec{v_S} \cdot \vec{v_S} = \sum_{i<j} -\vec{v_i} \cdot \vec{v_j} \left((m_i - m_j)^2 + (2z_i - 2z_j)^2 + 2(m_i - m_j)(2z_i - 2z_j)\right)$$

If all $\vec{v_i} \cdot \vec{v_j} < 0$, there exist 2 minimizing solutions: $z_i = 0$ and $z_i - z_j = -(m_i - m_j)$, this corresponds to $\vec{v}_S$ and $-\vec{v}_S$ respectively.

All other values for z_i are not minimal; thus, $\vec{v_S}$ must be a relevant Voronoi vector. □

Now that the relationship between a non-strictly-obtuse superbasis and the Voronoi vectors of a lattice has been established, it's important to know how to construct an obtuse superbasis starting from any generating set of the lattice. The Selling algorithm can be used to accomplish this, and, in so doing, proves that every 3D lattice possesses a non-strictly-obtuse superbasis.

Theorem 2.0.5. *Any 3D lattice possesses a non-strictly-obtuse superbasis [12]*

Proof. [12] Assume lattice $\mathbb{L}$ has a superbasis $\tilde{\mathbb{B}} := \{\vec{v_0}, \vec{v_1}, \vec{v_2}, \vec{v_3}\}$.

Referencing the proof above, the $\mathbb{L}/2\mathbb{L}$ factor group can be enumerated as follows, noting that each $\vec{v_S}$ falls into a distinct coset:

$$\mathbb{L}/2\mathbb{L} = \left\{(2\mathbb{L})_{\vec{v_0}}, (2\mathbb{L})_{\vec{v_1}}, (2\mathbb{L})_{\vec{v_2}}, (2\mathbb{L})_{\vec{v_3}}, (2\mathbb{L})_{\vec{v_0}+\vec{v_1}}, (2\mathbb{L})_{\vec{v_0}+\vec{v_2}}, (2\mathbb{L})_{\vec{v_0}+\vec{v_3}}, (2\mathbb{L})_{\vec{0}}\right\}$$

Consider the following change of basis:

$$\vec{v_1}' = -\vec{v_1}$$
$$\vec{v_2}' = \vec{v_1} + \vec{v_2}$$
$$\vec{v_3}' = \vec{v_3}$$
$$\vec{v_0}' = \vec{v_0} + \vec{v_1}$$

This is accomplished by right multiplication of the following unimodular matrix of the generators $[\vec{v_1}, \vec{v_2}, \vec{v_3}]$, arranged as column vectors of a matrix:

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

This transformation leads to the following new coset representatives:

$$\mathbb{L}/2\mathbb{L} = \left\{(2\mathbb{L})_{\vec{v_0}}, (2\mathbb{L})_{-\vec{v_1}}, (2\mathbb{L})_{-\vec{v_2}}, (2\mathbb{L})_{\vec{v_3}}, (2\mathbb{L})_{\vec{v_0}+\vec{v_1}}, (2\mathbb{L})_{\vec{v_1}-\vec{v_3}}, (2\mathbb{L})_{-\vec{v_0}-\vec{v_3}}, (2\mathbb{L})_{\vec{0}}\right\}$$

We notice that all representatives are simply inverted or the same length, except for the change: $(2\mathbb{L})_{\vec{v_0}+\vec{v_2}} = (2\mathbb{L})_{-\vec{v_1}-\vec{v_3}} \to (2\mathbb{L})_{\vec{v_1}-\vec{v_3}}$

Comparing the coset representative lengths:

$$(\vec{v_1} + \vec{v_3})^2 = v_1^2 + v_3^2 + 2\vec{v_1} \cdot \vec{v_3}$$
$$(\vec{v_1} - \vec{v_3})^2 = v_1^2 + v_3^2 - 2\vec{v_1} \cdot \vec{v_3}$$

The square norm changed by $-4\vec{v_1} \cdot \vec{v_3}$. If this dot product were positive, then the transformation preserved the norm of each coset representative, but it reduced the norm of the coset representative $\vec{v_1} + \vec{v_3}$, by using the new representative $\vec{v_1} - \vec{v_3}$.

By swapping labels when necessary and repeating this process, we can continue to find shorter representatives until all $\vec{v_i} \cdot \vec{v_j} \leq 0 \quad \forall i, j$. This means that the superbasis will be

non-strictly obtuse and the coset representatives will be as small as possible, hence, Voronoi vectors. □

Now, for the purposes of this thesis, it is necessary to study the possible lengths and angles of this superbasis further, revealing additional features of a non-strictly-obtuse superbasis. To begin this study, we define additional terms.

We define a vonorm of a lattice vector, $vo(\vec{l}), \vec{l} \in \mathbb{L}$ as the smallest vector length in $\vec{l}$'s $2\mathbb{L}$ equivalence class, aka. the coset $(2\mathbb{L})_{\vec{l}}$. This is denoted as follows:

$$vo(\vec{l}) := min(\vec{x} \cdot \vec{x} : \vec{x} \in (2\mathbb{L})_{\vec{v}}) \tag{2.6}$$

An n-dimensional lattice will then have up to 2^n distinct vonorms, one of which is zero, in correspondence with its factor group $\mathbb{L}/2\mathbb{L}$, which has 2^n members.

We define a character, F as a bounded-norm, 1-dimensional representation of the lattice group $\mathbb{L}$ on the vector space $\mathbb{R}$. This can be denoted as follows:

$$F : \mathbb{L} \to \{\pm 1\}, F(\vec{u} + \vec{v}) = F(\vec{u}) \times F(\vec{v})$$

Given a character, F, there exists a derived character $\tilde{F}$ as a bounded-norm, 1-dimensional representation of the factor group $\mathbb{L}/2\mathbb{L}$ on the vector space $\mathbb{R}$. This can be denoted as follows:

$$\tilde{F} : \mathbb{L}/2\mathbb{L} \to \{\pm 1\}, \tilde{F}((2\mathbb{L})_{\vec{v}}) = F(\vec{v})$$

Proof. Consider $\vec{v_1}, \vec{v_2} \in (2\mathbb{L})_{\vec{v_1}} \to \vec{v_2} = \vec{v_1} + 2\vec{l}, \vec{l} \in \mathbb{L}$

Then,

$$\tilde{F}\left((2\mathbb{L})_{\vec{v_2}}\right) = F(\vec{v_2}) = F\left(\vec{v_1} + \vec{l} + \vec{l}\right) = F(\vec{v_1}) F\left(\vec{l}\right)^2 = F(\vec{v_1}) = \tilde{F}\left((2\mathbb{L})_{\vec{v_1}}\right)$$

Consider $\vec{v_1}, \vec{v_2} \in \mathbb{L}$:

Then,

$$\tilde{F}\left((2\mathbb{L})_{\vec{v_1}} \cdot (2\mathbb{L})_{\vec{v_2}}\right) = \tilde{F}\left((2\mathbb{L})_{\vec{v_1}+\vec{v_2}}\right) = F(\vec{v_1} + \vec{v_2}) = F(\vec{v_1}) F(\vec{v_2}) = \tilde{F}\left((2\mathbb{L})_{\vec{v_1}}\right) \tilde{F}\left((2\mathbb{L})_{\vec{v_2}}\right)$$

□

A one-dimensional representation is irreducible trivially; thus, distinct representations $\tilde{F}_1$ and $\tilde{F}_2$ are orthogonal under the inner product:

$$\left\langle \tilde{F}_1, \tilde{F}_2 \right\rangle := \sum_{g \in \mathbb{L}/2\mathbb{L}} \tilde{F}_1(g)\tilde{F}_2(g) = \begin{cases} 0 & \text{if } \tilde{F}_1 \neq \tilde{F}_2 \\ |\mathbb{L}/2\mathbb{L}| & \text{else} \end{cases}$$

Recalling that there are 2^n distinct cosets in the factor group $\mathbb{L}/2\mathbb{L}$, there are up to 2^n orthogonal derived characters for the lattice and 2^n orthogonal characters.

In analogy to vonorms, the $\mathbb{L}/2\mathbb{L}$ factor group, and derived characters, we now define conorms, the derived character group, and meta-characters.

We define a conorm of a derived character, $\tilde{F}$ as a specific summation of vonorms. It is denoted as follows:

$$co\left(\tilde{F}\right) := -\frac{1}{2^{n-1}} \sum_{g \in \mathbb{L}/2\mathbb{L}} \tilde{F}(g) vo(g) \tag{2.7}$$

As such, the conorms of the lattice are completely specified by the vonorms of the lattice.

The derived characters of the lattice form a group under functional composition, here termed the derived character group. This group action can be summarized as:

$$\left(\tilde{F}_1 \cdot \tilde{F}_2\right)\left((2\mathbb{L})_{\vec{v}}\right) := \tilde{F}_1\left((2\mathbb{L})_{\vec{v}}\right) \tilde{F}_2\left((2\mathbb{L})_{\vec{v}}\right) \tag{2.8}$$

Proof. By construction, this composition always forms a valid derived character.

The identity element is the derived character: $\tilde{F}\left(g\right) = 1 \quad \forall g \in \mathbb{L}/2\mathbb{L}$

For all $\tilde{F}$, the inverse $\tilde{F}^{-1} = \tilde{F}$. □

We then define the meta-character $\tilde{G}_{\vec{v}}$ representation of the derived character group on $\mathbb{R}$. This can be written as:

$$\tilde{G}_{\vec{v}} : \left\{\tilde{F}\right\} \to \{\pm 1\} : \tilde{G}_{\vec{v}}\left(\tilde{F}\right) = \tilde{F}\left((2\mathbb{L})_{\vec{v}}\right)$$

We see that $\tilde{G}_{\vec{v}} \simeq \mathbb{L}/2\mathbb{L}$, as a derived character $\tilde{F}$ has only one value per coset.

Proof.

$$\tilde{G}_{\vec{v}}\left(\tilde{F}_1 \cdot \tilde{F}_2\right) = \left(\tilde{F}_1 \cdot \tilde{F}_2\right)\left((2\mathbb{L})_{\vec{v}}\right) = \tilde{F}_1\left((2\mathbb{L})_{\vec{v}}\right)\tilde{F}_2\left((2\mathbb{L})_{\vec{v}}\right) = \tilde{G}_{\vec{v}}\left(\tilde{F}_1\right)\tilde{G}_{\vec{v}}\left(\tilde{F}_2\right)$$

□

Noting that the $\tilde{G}$ meta character is one-dimensional, it is irreducible trivially, and distinct representations will be orthogonal under the following inner product:

$$\left\langle \tilde{G}_{\vec{v_1}}, \tilde{G}_{\vec{v_2}} \right\rangle := \sum_{\tilde{F} \in \left\{\tilde{F}\right\}} \tilde{G}_{\vec{v_1}}\left(\tilde{F}\right)\tilde{G}_{\vec{v_2}}\left(\tilde{F}\right) = \sum_{\tilde{F} \in \left\{\tilde{F}\right\}} \tilde{F}\left((2\mathbb{L})_{\vec{v_1}}\right)\tilde{F}\left((2\mathbb{L})_{\vec{v_2}}\right) = \begin{cases} 0 \\ |\mathbb{L}/2\mathbb{L}| \end{cases}$$

This inner product vanishes if $(2\mathbb{L})_{\vec{v}_1} \neq (2\mathbb{L})_{\vec{v}_2}$, otherwise it equals $|\mathbb{L}/2\mathbb{L}|$

Theorem 2.0.6. *The vonorms are completely specified by the conorms of the lattice: [12]*

$$vo\left(\vec{v}\right) = \frac{-1}{2^{n-1}} \sum_{\tilde{F}} \tilde{F}\left((2\mathbb{L})_{\vec{v}}\right) co\left(\tilde{F}\right)$$

Proof. We show this is true by simplifying the right hand side with some algebra:

$$vo\left(\vec{v}\right) = \frac{1}{2^n} \sum_{\tilde{F}} \tilde{F}\left((2\mathbb{L})_{\vec{v}}\right) \sum_{g \in \mathbb{L}/2\mathbb{L}} \tilde{F}(g) vo(g)$$

Changing the order of the summation:

$$vo\left(\vec{v}\right) = \frac{1}{2^n} \sum_{g \in \mathbb{L}/2\mathbb{L}} vo(g) \sum_{\tilde{F}} \tilde{F}(g)\tilde{F}\left((2\mathbb{L})_{\vec{v}}\right)$$

Applying the orthogonality relationship between derived characters:

$$vo\left(\vec{v}\right) = \frac{1}{2^n} \sum_{g \in \mathbb{L}/2\mathbb{L}} vo(g) \times 2^n \delta_{g,(2\mathbb{L})_{\vec{v}}} = vo\left(\vec{v}\right)$$

□

Putting these conceptual pieces together, it's clear that the vonorms and the conorms of the lattice fundamentally encode the same information: each can be derived from the other. The cosets, $\mathbb{L}/2\mathbb{L}$, and the derived characters of the lattice also exist analogously, and a

non-strictly-obtuse superbasis of the lattice can be used to enumerate both sets. This is shown in a similar theorem below.

Theorem 2.0.7. *Given a lattice* $\mathbb{L}$ *with a non-strictly-obtuse superbasis:* $\tilde{B} := \{\vec{v_0}, ..., \vec{v_n}\}$*, The set:*

$$\tilde{F}_S\left((2\mathbb{L})_{\vec{v_i}}\right) = \begin{cases} -1 & i \in S \\ +1 & i \notin S \end{cases} \qquad s \subseteq \{0, ..., n\} : |s| \ \% \ 2 = 0$$

is the set of all possible derived characters on $\mathbb{L}$ *[12]*

Proof. Consider a vector $\vec{v} := \sum m_i \vec{v_i}$ and the fact that $\sum \vec{v_i} = \vec{0}$. Then,

$$\vec{v} = \sum (m_i - 1)\, \vec{v_i}$$

Consider now the derived character $\tilde{F}_S$:

$$\tilde{F}_S\left((2\mathbb{L})_{\sum m_i \vec{v_i}}\right) = \prod \tilde{F}_S\left((2\mathbb{L})_{\vec{v_i}}\right)^{m_i} = \tilde{F}_S\left((2\mathbb{L})_{\sum (m_i - 1)\vec{v_i}}\right) = \prod \tilde{F}_S\left((2\mathbb{L})_{\vec{v_i}}\right)^{m_i - 1}$$

Thus,

$$\prod \tilde{F}_S\left((2\mathbb{L})_{\vec{v_i}}\right) = 1$$

Then, $\tilde{F}_S$ must have an even number of negative values and S must have an even number of elements.

Assume $S_1 \neq S_2$ then $\exists \vec{v} : \tilde{F}_{S_1}\left((2\mathbb{L})_{\vec{v}}\right) \neq \tilde{F}_{S_2}\left((2\mathbb{L})_{\vec{v}}\right)$ and $\tilde{F}_{S_1} \neq \tilde{F}_{S_2}$.

There are 2^n even subsets of $\{0, ..., n\}$, each of which enumerates the characters.

Verifying that $\tilde{F}$ is a valid representation:

$$\tilde{F}_S\left((2\mathbb{L})_{(\sum m_i \vec{v_i} + \sum n_i \vec{v_i})}\right) = \tilde{F}_S\left((2\mathbb{L})_{\left(\sum_{i \in S} (m_i + n_i)\vec{v_i} + \sum_{i \notin S} (m_i + n_i)\vec{v_i}\right)}\right) = \prod_{i \in S} (-1)^{m_i + n_i}$$

$$\tilde{F}_S\left((2\mathbb{L})_{(\sum m_i \vec{v_i})}\right) \tilde{F}_S\left((2\mathbb{L})_{(\sum n_i \vec{v_i})}\right) = \prod_{i \in S} (-1)^{m_i} \prod_{i \in S} (-1)^{n_i}$$

□

Now as a final yet salient proof, we establish the geometric interpretation of the conorms of the lattice, binding together the previous concepts.

Theorem 2.0.8. *Given lattice $\mathbb{L}$ with a non-strictly-obtuse superbasis $\tilde{\mathbb{B}} := \{\vec{v_0}, ..., \vec{v_n}\}$, the non-identity conorms of $\mathbb{L}$ can be written as follows: [12]*

$$co\left(\tilde{F}_S\right) = \begin{cases} -\vec{v_i} \cdot \vec{v_j} & \text{if } S = \{i, j\} \\ 0 & \text{else} \end{cases} \qquad S \subseteq \{0, ..., n\},\ |S| \in 2\mathbb{Z}$$

Proof. For this sub-proof, we define a normal subgroup, $\mathbb{H}_{ij}$, of $\mathbb{L}/2\mathbb{L}$ and show that it's valid. Further we define a derived orthogonality relationship between the members of $\mathbb{H}_{ij}$

$$\mathbb{H}_{ij} := \left\{(2\mathbb{L})_{\vec{v}_S} : i, j \notin S\right\} \trianglelefteq \mathbb{L}/2\mathbb{L}, \qquad \vec{v_S} := \sum_{i \in S} \vec{v_i}$$

Consider $h \in \mathbb{H}_{ij}$, by definition, $h \cdot h = (2\mathbb{L})_{2\vec{v}_S} = (2\mathbb{L})_{\vec{0}} = e$; thus, $h = h^{-1}$, $h^{-1} \in \mathbb{H}$.

Consider $h_1, h_2 \in \mathbb{H}_{ij}$. Both h_1 and h_2 individually can be expressed as $(2\mathbb{L})_{\sum m_x \vec{v_x}}$ where $m_i = m_j = 0$. The composition $h_1 \cdot h_2$ can then be expressed analogously as $(2\mathbb{L})_{\sum m_x \vec{v_x}}$, and $m_i = m_j = 0$. Thus, $h_1 \cdot h_2 \in \mathbb{H}_{ij}$.

Consider $g_1, g_2 \notin \mathbb{H}_{ij}$. We can find some element w such that $g_2 = g_1 \cdot w \rightarrow w = g_1^{-1} \cdot g_2$.

By definition of $\mathbb{H}_{ij}$, g_1 and g_2, when expressed as $(2\mathbb{L})_{\sum m_x \vec{v_x}}$ must have either $m_i = 1$ or $m_j = 1$, though not both.

Consider: if $m_i = m_j = 0$ then g must be in $\mathbb{H}_{ij}$. If $m_i = m_j = 1$ then we can add $-\sum \vec{v_i} = \vec{0}$ to g, preserving the vector, but changing $m_i = m_j = 0$, meaning $g \in \mathbb{H}_{ij}$.

Thus, when g_1 and g_2 are added, the result w must have $m_i = m_j = 1$, or it must have $m_i = 2$ and $m_j = 0$, or $m_i = 0$ and $m_j = 2$. In the latter two cases we can subtract $2\vec{v_i}$ or $2\vec{v_j}$, preserving the $\mathbb{L}/2\mathbb{L}$ coset. The result will have $m_i = 0$ and $m_j = 0$, meaning that $w \in \mathbb{H}_{ij}$. In the first case, we can once more add $-\sum \vec{v_i} = \vec{0}$ to w. The result will have $m_i = 0$ and $m_j = 0$, meaning that $w \in \mathbb{H}_{ij}$.

Because $w \in \mathbb{H}_{ij}$, it follows that the group $\mathbb{L}/2\mathbb{L}$ must contain 2 cosets of $\mathbb{H}_{ij}$. $\mathbb{H}_{ij}$ is therefore halving group and therefore a normal subgroup of $\mathbb{L}/2\mathbb{L}$. Because $\mathbb{H}_{ij}$ is a subgroup of $\mathbb{L}/2\mathbb{L}$, all derived characters $\tilde{F}$ have defined values on the elements of $\mathbb{H}_{ij}$ and they form another one-dimensional representation with an analogous orthogonality relationship:

$$\left\langle \tilde{F}_1, \tilde{F}_2 \right\rangle := \sum_{h \in \mathbb{H}_{ij}} \tilde{F}_1(h)\, \tilde{F}_2(h) = \begin{cases} 0 & \text{if } \tilde{F}_1 \neq \tilde{F}_2 \\ |\mathbb{H}_{ij}| & \text{if } \tilde{F}_1 = \tilde{F}_2 \end{cases}$$

□

Proof. Given lattice $\mathbb{L}$ with obtuse superbasis $\tilde{\mathbb{B}}$, the non-zero vonorms of each $(2\mathbb{L})_{\vec{v}_S}$ coset defined by $\left\{\vec{v}_S | S \subseteq \{0, ..., n\}, |S| \in (0, n), \vec{v}_S := \sum_{i \in S} \vec{v}_i\right\}$ are given by the following:

$$\vec{v}_S \cdot \vec{v}_S = \sum_{i \in S, j \notin S} -\vec{v}_i \cdot \vec{v}_j$$

We can show this via some algebra, beginning with the definition of $\vec{v}_S$

$$\vec{v}_S \cdot \vec{v}_S = \sum_{i \in S} \sum_{j \in S} \vec{v}_i \cdot \vec{v}_j = \sum_{i \in S} \vec{v}_i \cdot \vec{v}_i + \sum_{i \in S} \sum_{j \in S, j \neq i} \vec{v}_i \cdot \vec{v}_j$$

Using the construction of a superbasis, we can re-express $\vec{v}_i$ as a sum over all other vectors in the superbasis:

$$\vec{v}_S \cdot \vec{v}_S = \sum_{i \in S} \sum_{j \neq i} -\vec{v}_i \cdot \vec{v}_j + \sum_{i \in S} \sum_{j \in S, j \neq i} \vec{v}_i \cdot \vec{v}_j$$

Analyzing the bounds of summation, we can re-express the sums again:

$$\vec{v}_S \cdot \vec{v}_S = \sum_{i \in S} \sum_{j \in S^c} -\vec{v}_i \cdot \vec{v}_j$$

□

Proof. Now we apply the constructions above to demonstrate the proof, starting with the definition of a conorm:

$$co\left(\tilde{F}_S\right) = \frac{-1}{2^{n-1}} \sum_{g \in \mathbb{L}/2\mathbb{L}} \tilde{F}_S(g) vo(g) = \frac{1}{2^n} \sum_{\tilde{S} \subseteq \{0,...,n\}} \tilde{F}_S\left(\vec{v}_{\tilde{S}}\right) \sum_{i \in \tilde{S}, j \notin \tilde{S}} \vec{v}_i \cdot \vec{v}_j$$

We re-order the summations:

$$co\left(\tilde{F}_S\right) = \frac{1}{2^n} \sum_{i<j} \vec{v}_i \cdot \vec{v}_j \sum_{\tilde{S} \subseteq \{0,...,n\} : i \in \tilde{S}, j \notin \tilde{S}} \tilde{F}_S\left(\vec{v}_{\tilde{S}}\right)$$

We can re-express the inner summand using the definition of $\mathbb{H}_{ij}$:

$$co\left(\tilde{F}_S\right) = \frac{1}{2^{n-1}} \sum_{i<j} \vec{v}_i \cdot \vec{v}_j \sum_{h \in \mathbb{H}_{ij}} \tilde{F}_S\left(h \cdot (2\mathbb{L})_{\vec{v}_i}\right)$$

We can take the term $\tilde{F}_S\left((2\mathbb{L})_{\vec{v}_i}\right)$ out of the inner summand. We note that the identity conorm $\tilde{F}_I$ is always equal to 1. Thus, multiplying the inner summand by 1 we can add this

additional term as well.

$$co\left(\tilde{F}_S\right) = \frac{1}{2^{n-1}} \sum_{i<j} \vec{v_i} \cdot \vec{v_j}\ \tilde{F}_S\left((2\mathbb{L})_{\vec{v_i}}\right) \sum_{h \in \mathbb{H}_{ij}} \tilde{F}_S(h)\, \tilde{F}_I(h)$$

Recalling the orthogonality relation above, we see that $\tilde{F}_S(h)$ must equal 1 for all $h \in \mathbb{H}_{ij}$, otherwise it will be orthogonal to $\tilde{F}_I$. This means that if $S \neq \{i, j\}$ then the sum will be zero. Representing this mathematically:

$$co\left(\tilde{F}_S\right) = \frac{1}{2^{n-1}} \sum_{i<j} \vec{v_i} \cdot \vec{v_j}\ \tilde{F}_S\left((2\mathbb{L})_{\vec{v_i}}\right) 2^{n-1} \delta_{\{i,j\},\{a,b\}=S}$$

$$co\left(\tilde{F}_S\right) = \vec{v_a} \cdot \vec{v_b}\ \tilde{F}_{\{a,b\}}\left((2\mathbb{L})_{\vec{v_a}}\right)$$

$$co\left(\tilde{F}_S\right) = -\vec{v_a} \cdot \vec{v_b}\ \delta_{\{a,b\},S}$$

□

This proof establishes an important connection between conorms and the dot products of a lattice's non-strictly-obtuse superbasis. As such, the lengths of the non-strictly-obtuse superbasis (vonorms), uniquely define the conorms, which uniquely define the dot products of the lattice's non-strictly-obtuse superbasis. The vonorm lengths and conorm values, both inherent geometric properties of the lattice, uniquely determine the geometry of the non-strictly-obtuse superbasis. By extension, then, every non-strictly-obtuse superbasis of a lattice must have the same geometry and is thus a unique geometric construction for a lattice.

We can now enumerate the cosets / vonorms of the lattice alongside the derived lattice characters / conorms of a 3D lattice, in a table and in a compact matrix relationship. Vectors $\vec{v_i}$ represent the vectors of the lattice's non-strictly-obtuse superbasis.

$$\begin{pmatrix} -1 & -1 & 1 & 1 & 1 & -1 & -1 \\ -1 & 1 & -1 & 1 & -1 & 1 & -1 \\ -1 & 1 & 1 & -1 & -1 & -1 & 1 \\ 1 & -1 & -1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 & 1 & -1 & -1 \\ -1 & -1 & -1 & -1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} v_0^2 \\ v_1^2 \\ v_2^2 \\ v_3^2 \\ (\vec{v_0} + \vec{v_1})^2 \\ (\vec{v_0} + \vec{v_2})^2 \\ (\vec{v_0} + \vec{v_3})^2 \end{pmatrix} = 4 \begin{pmatrix} \vec{v_0} \cdot \vec{v_1} \\ \vec{v_0} \cdot \vec{v_2} \\ \vec{v_0} \cdot \vec{v_3} \\ \vec{v_1} \cdot \vec{v_2} \\ \vec{v_1} \cdot \vec{v_3} \\ \vec{v_2} \cdot \vec{v_3} \\ 0 \end{pmatrix} \tag{2.9}$$

As a final note, the ultimate row of the matrix shows the relationship:

$$v_0^2 + v_1^2 + v_2^2 + v_3^2 = (\vec{v_0} + \vec{v_1})^2 + (\vec{v_0} + \vec{v_2})^2 + (\vec{v_0} + \vec{v_3})^2 \tag{2.10}$$

	$(2\mathbb{L})_{\vec{0}}$	$(2\mathbb{L})_{\vec{v}_0}$	$(2\mathbb{L})_{\vec{v}_1}$	$(2\mathbb{L})_{\vec{v}_2}$	$(2\mathbb{L})_{\vec{v}_3}$	$(2\mathbb{L})_{\vec{v}_0+\vec{v}_1}$	$(2\mathbb{L})_{\vec{v}_0+\vec{v}_2}$	$(2\mathbb{L})_{\vec{v}_0+\vec{v}_3}$
$\tilde{F}_{\emptyset}$	1	1	1	1	1	1	1	1
$\tilde{F}_{\{0,1\}}$	1	-1	-1	1	1	1	-1	-1
$\tilde{F}_{\{0,2\}}$	1	-1	1	-1	1	-1	1	-1
$\tilde{F}_{\{0,3\}}$	1	-1	1	1	-1	-1	-1	1
$\tilde{F}_{\{1,2\}}$	1	1	-1	-1	1	-1	-1	1
$\tilde{F}_{\{1,3\}}$	1	1	-1	1	-1	-1	1	-1
$\tilde{F}_{\{2,3\}}$	1	1	1	-1	-1	1	-1	-1
$\tilde{F}_{\{0,1,2,3\}}$	1	-1	-1	-1	-1	1	1	1

Table 2.1: An enumeration of the 8 distinct derived characters ($\tilde{F}$) of a 3D lattice ($\mathbb{L}$) with superbasis ($\{\vec{v}_i\}$) and their values on the 8 distinct cosets comprising the double-lattice factor group ($\mathbb{L}/2\mathbb{L}$)

While there are 7 distinct vonorms, this condition establishes a fundamental mathematical constraint on their values. Thus, a 3D lattice has 6 geometric degrees of freedom, in agreement with the larger corpus of 3D lattice theories.

Here we have developed a description of the space of all 3D lattices, a 6-dimensional surface embedded within $\mathbb{R}^7$, corresponding to the region of space where the canonical invariant: $v_0^2+v_1^2+v_2^2+v_3^2 = (\vec{v_0}+\vec{v_1})^2+(\vec{v_0}+\vec{v_2})^2+(\vec{v_0}+\vec{v_3})^2$ holds true. This surface is not a vector space, as certain combinations of vector lengths don't correspond to a valid non-strictly-obtuse superbasis: it must be explicitly verified that each dot product $\vec{v}_i \cdot \vec{v}_j$ is less-than or equal-to zero. However, the regions corresponding to valid superbasis lengths must lie within this hyperplane with normal vector $< 1, 1, 1, 1, -1, -1, -1 >$ to be consistent with the aforementioned canonical constraint. The plane has an orthonormal basis given as follows:

$$\left\{ \frac{1}{2}\begin{pmatrix}1\\1\\-1\\-1\\0\\0\\0\end{pmatrix}, \frac{1}{2}\begin{pmatrix}1\\-1\\-1\\1\\0\\0\\0\end{pmatrix}, \frac{1}{2}\begin{pmatrix}-1\\1\\-1\\1\\0\\0\\0\end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix}0\\0\\0\\0\\2\\-1\\-1\end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix}0\\0\\0\\0\\0\\1\\-1\end{pmatrix}, \frac{1}{\sqrt{84}}\begin{pmatrix}3\\3\\3\\3\\4\\4\\4\end{pmatrix} \right\}$$

Crystallographic Motif Theory

Having developed a mathematical description of the geometry of all 3D lattices, it is useful to also explicitly mathematically describe all possible motifs for a crystal. For n atoms in a unit cell, there exist $3n$ degrees of freedom: each atom can be perturbed in any direction in 3D space. However, out of the $3n$ total degrees of freedom, 3 are redundant, as globally translating all atoms identically and simultaneously does not change the geometry of the crystal.

Given a basis for $\mathbb{R}^3$, the three coordinates of each atom can be varied independently, and the $3n$ degrees of freedom can be built up from the tensor product of n degrees of freedom 3 times: once per coordinate. Then, analyzing one coordinate, we have n degrees of freedom, one per atom, and global translation along that coordinate can be represented as $< 1, ..., 1 >$, a homogeneous vector of 1's repeated n times. Coming up with $n-1$ linearly-independent vectors will provide a basis for the space of all atomic translations along the chosen coordinate. Further, repeating this for each of the 3 coordinates, provides a basis for all possible atomic translations within the crystal when the appropriate tensor-product is calculated, with zeros along the other dimensions.

To calculate $n-1$ linearly-independent basis vectors, we will turn to the theory of a discrete Fourier series. In general the vector $< 1, ..., 1 >$ can be interpreted as a zero-frequency discrete cosine wave defined on n points. As such, there naturally arise $n-1$ orthogonal vectors.

To show this, we define the following family of functions within the vector space of all complex-valued function on a set of n points, alongside a canonical inner product:

$$f_\omega : \mathbb{Z} \cap [1, n] \to \mathbb{C} \mid f(t) = e^{i\omega t} \qquad < f_{\omega_1}, f_{\omega_2} >:= \sum_{t=1}^{n} f_{\omega_1}(t) f_{\omega_2}(t)^*$$

We mandate that each function f_ω be orthogonal to the homogeneous translation vector $< 1, ..., 1 >$. Thus:

$$< f_\omega, 1 >= \sum_{t=1}^{n} e^{i\omega t} = e^{i\omega}\left(\frac{e^{i\omega n}-1}{e^{i\omega}-1}\right) = 0 \ \rightarrow \ e^{i\omega n} = 1 \ \rightarrow \ \omega = \frac{a2\pi}{n}, a \in \mathbb{Z}$$

It can clearly be seen that $a \in \mathbb{Z} \cap [0, n)$ provides all possible unique orthogonal ω values. Taking $a \rightarrow a + n$ will shift $\omega \rightarrow \omega + 2\pi$, and f_ω and $f_{\omega+2\pi}$ are identically-valued. Where $a = 0$, f_ω is the homogeneous translation vector. We can show that each of the n distinct values for a correspond to orthogonal f_ω functions, forming an orthogonal basis for $\mathbb{C}^n$. Consider:

$$< f_{\frac{a_1 2\pi}{n}}, f_{\frac{a_2 2\pi}{n}} >= \sum_{t=1}^{n} e^{i\frac{2\pi}{n}(a_1-a_2)} = e^{i\frac{2\pi}{n}(a_1-a_2)}\left(\frac{e^{i2\pi(a_1-a_2)}-1}{e^{i\frac{2\pi}{n}(a_1-a_2)}-1}\right) = \begin{cases} 0 & \text{if } a_1 \neq a_2 \\ n & \text{else} \end{cases}$$

Of course, our field of interest is the set of real numbers $\mathbb{R}$; thus, we can consider real combinations of the functions f_ω in the form of sine and cosine waves. These combinations will be real-valued and orthogonal under the action of the previously-defined inner product:

$$\cos(\omega t) = \frac{1}{2}\left(f_\omega + f_{-\omega}\right) \qquad \sin(\omega t) = \frac{1}{2i}\left(f_\omega - f_{-\omega}\right)$$

$$< \cos(\omega_1 t), \cos(\omega_2 t) >= \frac{1}{4}\left(< f_{\omega_1}, f_{\omega_2} > + < f_{\omega_1}, f_{-\omega_2} > + < f_{-\omega_1}, f_{\omega_2} > + < f_{-\omega_1}, f_{-\omega_2} >\right)$$

$$< \cos(\omega_1 t), \cos(\omega_2 t) >= \frac{n}{4}\left(\delta_{\omega_1,\omega_2} + \delta_{\omega_1+\omega_2,2\pi m} + \delta_{-\omega_1-\omega_2,2\pi m} + \delta_{\omega_1,\omega_2}\right)$$

$$< \sin(\omega_1 t), \sin(\omega_2 t) >= \frac{1}{4}\left(< f_{\omega_1}, f_{\omega_2} > - < f_{\omega_1}, f_{-\omega_2} > - < f_{-\omega_1}, f_{\omega_2} > + < f_{-\omega_1}, f_{-\omega_2} >\right)$$

$$< \sin(\omega_1 t), \sin(\omega_2 t) >= \frac{n}{4}\left(\delta_{\omega_1,\omega_2} - \delta_{\omega_1+\omega_2,2\pi m} - \delta_{-\omega_1-\omega_2,2\pi m} + \delta_{\omega_1,\omega_2}\right)$$

$$< \cos(\omega_1 t), \sin(\omega_2 t) >= -\frac{1}{4i}\left(< f_{\omega_1}, f_{\omega_2} > - < f_{\omega_1}, f_{-\omega_2} > + < f_{-\omega_1}, f_{\omega_2} > - < f_{-\omega_1}, f_{-\omega_2} >\right)$$

$$< \cos(\omega_1 t), \sin(\omega_2 t) >= -\frac{n}{4i}\left(\delta_{\omega_1,\omega_2} - \delta_{\omega_1+\omega_2,2\pi m} + \delta_{-\omega_1-\omega_2,2\pi m} - \delta_{\omega_1,\omega_2}\right) = 0$$

With this information we can define the following functions to serve as a basis for all atomic translations along a chosen coordinate, omitting the redundant $< 1, ..., 1 >$ homogeneous

translation vector:

$$\begin{cases} n \text{ odd} & \begin{cases} \omega = \frac{2\pi}{n} a & \left\{\sqrt{\frac{2}{n}}\cos(\omega t)\right\} \\ a \in \mathbb{Z} \cap [1, n/2) & \left\{\sqrt{\frac{2}{n}}\sin(\omega t)\right\} \end{cases} \\ n \text{ even} & \begin{cases} \omega = \frac{2\pi}{n} a & \left\{\sqrt{\frac{2}{n}}\cos(\omega t)\right\} \\ a \in \mathbb{Z} \cap [1, n/2) & \left\{\sqrt{\frac{2}{n}}\sin(\omega t)\right\} \\ & \left\{\sqrt{\frac{1}{n}}\cos(\pi t)\right\} \end{cases} \end{cases}$$

Each individual coordinate of each atom has this basis. With the appropriate tensor products, then, a basis for the $3(n-1)$ degrees of freedom is defined.

Take for example a 3-atom unit cell. We can define a basis (6 vectors) for the atomic displacements, listing the coordinates in order as follows: (atom1 coord1), (atom1 coord2), (atom1 coord3), (atom2 coord1), (atom2 coord2), (atom2 coord3), (atom3 coord1), (atom3 coord2), (atom3 coord3):

$$\mathbb{B} = \left\{ \begin{pmatrix} -0.408 \\ 0 \\ 0 \\ -0.408 \\ 0 \\ 0 \\ 0.816 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0.707 \\ 0 \\ 0 \\ -0.707 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ -0.408 \\ 0 \\ 0 \\ -0.408 \\ 0 \\ 0 \\ 0.816 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0.707 \\ 0 \\ 0 \\ -0.707 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ -0.408 \\ 0 \\ 0 \\ -0.408 \\ 0 \\ 0 \\ 0.816 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0.707 \\ 0 \\ 0 \\ -0.707 \\ 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

We note that this basis is valid for any choice of coordinate system spanning $\mathbb{R}^3$, whether Cartesian (x, y, z) or direct (a, b, c).

Crystallographic Representations

In general, the lattice and basis degrees of freedom are independent; however, when direct coordinates are used, the choice of unit cell is important. Consider a lattice $\mathbb{L}$ with its generating vectors arranged as column vectors of the matrix $[\mathbb{L}]$. Any generating set of the lattice can be chosen via right multiplication of a unimodular matrix (μ): $[\mathbb{L}] \to [\mathbb{L}]\mu$.

We can see this is true considering the following: take lattice $\mathbb{L}_1$ generated by $\{\vec{a}, \vec{b}, \vec{c}\}$ and lattice $\mathbb{L}_2$ generated by $\{\vec{x}, \vec{y}, \vec{z}\}$. By definition of a lattice, $\mathbb{L}_1 \subseteq \mathbb{L}_2 \leftrightarrow [\vec{v}_1]_{\mathbb{L}_2} \in \mathbb{Z}^3 \ \forall \vec{v}_1 \in \mathbb{L}_1$.

Then considering a mapping $A : A[\vec{v}_1]_{\mathbb{L}_1} = [\vec{v}_1]_{\mathbb{L}_2}$, A maps every possible set of three integer coordinates to another set of integer coordinates. This means that A must act as a matrix with all integer coordinates: $A \in GL_3(\mathbb{Z})$. If we consider the special case where $\mathbb{L}_1 = \mathbb{L}_2$, it follows that $\mathbb{L}_2 \subseteq \mathbb{L}_1$. We can analyze the inverse mapping $A^{-1} : A^{-1}[\vec{v}_1]_{\mathbb{L}_2} = [\vec{v}_1]_{\mathbb{L}_1}$, drawing the same conclusion: $A^{-1} \in GL_3(\mathbb{Z})$. Calculating the determinant of each, $|A|, |A^{-1}| \in \mathbb{Z}$, as each is a matrix composed solely of integers. We note that $|A| = |A^{-1}|^{-1} \in \mathbb{Z}$; thus, $|A| = \pm 1$, as only these integers are equal to their multiplicative inverses. With this additional constraint, $A \in SL_3(\mathbb{Z})$, a unimodular matrix.

Given a vector $\vec{v}$ in lattice $\mathbb{L}$, we can say that: $\vec{v} = [\mathbb{L}][\vec{v}]_{\mathbb{L}}$. Here $[\vec{v}]_{\mathbb{L}} \in \mathbb{Z}^3$, a coordinate vector of the lattice. We take a unimodular matrix A as defined before. Then, multiplying by unity, $\vec{v} = [\mathbb{L}]AA^{-1}[\vec{v}]_{\mathbb{L}}$. We can interpret this equation as a set of new lattice generators $[\mathbb{L}]A$ and a new integer coordinate string $A^{-1}[\vec{v}]_{\mathbb{L}} \in \mathbb{Z}^3$.

Applying this to the relative coordinates of the crystallographic motif, when a new unit cell is chosen given by generators $[\mathbb{L}]\mu$, the cartesian coordinates of the motif remain unchanged, yet the relative coordinates change. In an exact analogy to the previous equation $\vec{b} = [\mathbb{L}][\vec{b}]_{\mathbb{L}} = [\mathbb{L}]\mu\mu^{-1}[\vec{b}]_{\mathbb{L}}$. Thus, where a unimodular matrix changes the choice of unit cell, the coordinates of the motif in that new unit cell are related by a left multiplication of the inverse of that particular unimodular matrix.

In general, every crystal structure has an underlying fundamental periodicity, represented by the primitive unit cell of its underlying lattice. There is only one primitive lattice for a given crystal. However, in the mathematical specification of the crystal's geometry, it is possible to choose larger unit cells, defining a sublattice of this fundamental lattice. With this choice of sublattice, it is possible to expand the crystallographic motif to include additional atoms and recreate the same original crystal structure, in terms of a set of displacements from the origin. In these cases, there are often many distinct ways to define possible sublattices and their associated bases so that the original crystal structure is preserved.

Here, using a series of proofs and results from representation theory, we will delineate an approach for deriving geometrically-unique sublattices of index N for a given crystal.

Consider a vector $\vec{k} \in \mathbb{R}^n$ and a lattice $\mathbb{L}$. Then we can define a $\vec{k}$-derived sublattice $\mathbb{L}_{\vec{k}} := \left\{ \vec{v} \in \mathbb{L} : e^{i\vec{k}\cdot\vec{v}} = 1 \right\} \leq \mathbb{L}$. From this definition, a couple of facts also follow: $\mathbb{L}_{\vec{k}} = \mathbb{L}_{-\vec{k}}$, $\mathbb{L}_{\vec{k}} \leq \mathbb{L}_{z\vec{k}}, z \in \mathbb{Z}$.

Proof. Take $\vec{v}_1, \vec{v}_2 \in \mathbb{L}_{\vec{k}}$

Consider:

$$e^{i\vec{k}\cdot(\vec{v}_1+\vec{v}_2)} = e^{i\vec{k}\cdot\vec{v}_1}e^{i\vec{k}\cdot\vec{v}_2} = 1 \qquad \therefore\ \vec{v}_1+\vec{v}_2 \in \mathbb{L}_{\vec{k}}$$

Consider:

$$e^{-i\vec{k}\cdot\vec{v}_1} = \left(e^{i\vec{k}\cdot\vec{v}_1}\right)^{-1} = 1 \qquad \therefore\ -\vec{v}_1 \in \mathbb{L}_{\vec{k}}$$

Consider:

$$e^{i\vec{k}\cdot\vec{v}} = 1 = e^{-i\vec{k}\cdot\vec{v}} \qquad \therefore\ \vec{v} \in \mathbb{L}_{-\vec{k}} \quad \forall\vec{v} \in \mathbb{L}_{\vec{k}} \qquad \therefore\ \vec{v} \in \mathbb{L}_{\vec{k}} \quad \forall\vec{v} \in \mathbb{L}_{-\vec{k}}$$

Consider:

$$\vec{v} \in \mathbb{L}_{\vec{k}} \ \rightarrow\ e^{iz\vec{k}\cdot\vec{v}} = \left(e^{i\vec{k}\cdot\vec{v}}\right)^{z} = 1 \qquad \therefore\ \vec{v} \in \mathbb{L}_{z\vec{k}}$$

□

It can also be shown that $\vec{k}$, lies in the reciprocal lattice of $\mathbb{L}_{\vec{k}}$, denoted here $\mathbb{L}^*_{\vec{k}}$.

Proof. Consider:

$$\left\{e^{i\vec{k}\cdot\vec{v}} : \vec{v} \in \mathbb{L}_{\vec{k}}\right\} = \{1\} \qquad \therefore\ \vec{k}\cdot\vec{v} = 2\pi n, n \in \mathbb{Z} \quad \forall\vec{v} \in \mathbb{L}_{\vec{k}}$$

Thus, $\vec{k} \in \mathbb{L}^*_{\vec{k}}$, $\vec{k}$ is, by definition, a reciprocal lattice vector of $\mathbb{L}_{\vec{k}}$ □

Now we investigate further the connection between $\vec{k}$ and the nature of the sublattice $\mathbb{L}_{\vec{k}}$.

Consider the coset decomposition of $\mathbb{L}$ in terms of cosets of the subgroup $\mathbb{L}_{\vec{k}}$. As such, we express:

$$\mathbb{L} = \bigcup_{\vec{x}_i} \left(\mathbb{L}_{\vec{k}}\right)_{\vec{x}_i} \qquad \vec{x}_i \in \mathbb{L}$$

Take $e^{i\vec{k}\cdot\vec{x}_1} = e^{i\vec{k}\cdot\vec{x}_2}$

$$\therefore\ \vec{k}\cdot\vec{x}_1 = \vec{k}\cdot\vec{x}_2 + 2\pi n, n \in \mathbb{Z} \ \rightarrow\ \vec{k}\cdot(\vec{x}_1-\vec{x}_2) = 2\pi n \ \rightarrow\ \vec{x}_1-\vec{x}_2 \in \mathbb{L}_{\vec{k}} \ \rightarrow\ \vec{x}_2 \in \left(\mathbb{L}_{\vec{k}}\right)_{\vec{x}_1}$$

Take $\vec{x}_1, \vec{x}_2 \in \left(\mathbb{L}_{\vec{k}}\right)_{\vec{x}_1}$

$$\therefore\ \vec{x}_1 = \vec{x}_2 + \vec{t} : \vec{t} \in \mathbb{L}_{\vec{k}}\ \rightarrow\ e^{i\vec{k}\cdot\vec{x}_1} = e^{i\vec{k}\cdot\left(\vec{x}_2+\vec{t}\right)} = e^{i\vec{k}\cdot\vec{x}_2}$$

Thus, the values $\left\{e^{i\vec{k}\cdot\vec{v}} : \vec{v} \in \mathbb{L}\right\}$ enumerate the distinct cosets of $\mathbb{L}_{\vec{k}}$; further, if the subgroup $\mathbb{L}_{\vec{k}}$ is to be an index N subgroup, the aforementioned set must have N distinct values.

If we consider, $\mathbb{L}$ having generators $\{\vec{t}_i, ...\}$ and $\mathbb{L}^*$ having generators $\{\vec{t}^*_i, ...\}$, we can expand the dot product $\vec{k}\cdot\vec{v}, \vec{v} \in \mathbb{L}$ in terms of these coordinates:

$$\vec{k}\cdot\vec{v} = \left(\sum_i c_i\vec{t}^*_i\right)\cdot\left(\sum_i z_i\vec{t}_i\right) = 2\pi\sum_i c_i z_i, z_i \in \mathbb{Z}$$

Further we can show that $\mathbb{L}_{\vec{k}} = \mathbb{L}_{\vec{k}+\vec{g}}, \vec{g} \in \mathbb{L}^*$:

$$\vec{g} \in \mathbb{L}^*\ \rightarrow\ e^{i\vec{g}\cdot\vec{v}} = 1\ \forall\vec{v} \in \mathbb{L}\ \therefore\ \vec{v} \in \mathbb{L}_{\vec{k}} \leftrightarrow e^{i\vec{k}\cdot\vec{v}} = 1 = e^{i\left(\vec{k}+\vec{g}\right)\cdot\vec{v}} \leftrightarrow \vec{v} \in \mathbb{L}_{\vec{k}+\vec{g}}$$

Thus, all unique sublattices $\mathbb{L}_{\vec{k}}$ can be generated using $\vec{k}$ within one primitive unit cell of $\mathbb{L}^*$. This confines c_i coordinate values to the interval $[0, 1)$. Further, to form an index N subgroup, we need $\{\vec{c} : |\{\vec{c}\cdot\vec{z}\ \%\ 1\}| = N\}$. That is, the integer sum of all components c_i must lead to N distinct values in the interval $[0, 1)$. Values outside this interval, alias back into the interval due to the periodicity of $e^{i\vec{k}\cdot\vec{v}}$, hence the $\%$ modulus operator.

When explicitly stated in this form, it's clear that all coordinates c_i of $\vec{k}$ must be rational numbers otherwise $\{\vec{c}\cdot\vec{z}\ \%\ 1\}$ would have an infinite number of elements. As such we can express each:

$$c_i = m_i/n_i, \qquad m_i, n_i \in \mathbb{Z}, \qquad m_i, n_i \text{ coprime}$$

We can interpret each coordinate $c_i\ \%\ 1$ as its own cyclic group of order n_i. Then, together, $<\vec{k}>$ serves as a direct product group of each of its coordinates' cyclic groups. Referring to the mathematical preliminaries: $|<\vec{k}>| = \text{lcm}(\{n_i\})$. Further, $<\vec{k}> \leq \{\vec{c}\cdot\vec{z}\ \%\ 1\}$, as $\vec{k}$ self-addition is equivalent to multiplication by homogeneously-valued $\vec{z}$.

From these conclusions, we can show that $|\{\vec{c}\cdot\vec{z}\ \%\ 1\}| = \text{lcm}(\{n_i\})$

Proof. Any integer sum of c_i values can be put under a common denominator $\eta := \text{lcm}(\{n_i\})$, whose numerator μ may or may not be coprime with η.

This means that at most $\{\vec{c}\cdot\vec{z}\ \%\ 1\}$ contains at most η elements, potentially less.

However, because $<\vec{k}>\leq\{\vec{c}\cdot\vec{z}\ \%\ 1\}$, we can say that $\{\vec{c}\cdot\vec{z}\ \%\ 1\}$ contains at least η elements, potentially more.

Taken together, these statements imply that $\{\vec{c}\cdot\vec{z}\ \%\ 1\}$ contains exactly η elements. □

We can perform the same analysis as before; however, this time analyzing the reciprocal lattices $\mathbb{L}^*$ and $\mathbb{L}^*_{\vec{k}}$. Referencing the above math proofs, $\mathbb{L}^*\leq\mathbb{L}^*_{\vec{k}}$, as $e^{i\vec{g}\cdot\vec{v}}=1\ \forall\vec{v}\in\mathbb{L}_{\vec{k}},\vec{g}\in\mathbb{L}^*$. As before we can express $\mathbb{L}^*_{\vec{k}}$ as a union of cosets of $\mathbb{L}^*$:

$$\mathbb{L}^*_{\vec{k}}=\bigcup_{\vec{k}_i}(\mathbb{L}^*)_{\vec{k}_i}$$

We can perform the same analysis on these cosets as before.

Take $\vec{v}\in\mathbb{L}$, $e^{i\vec{k}_1\cdot\vec{v}}=e^{i\vec{k}_2\cdot\vec{v}}$

$$\therefore\ \vec{k}_1\cdot\vec{v}=\vec{k}_2\cdot\vec{v}+2\pi n, n\in\mathbb{Z}\ \rightarrow\ \left(\vec{k}_1-\vec{k}_2\right)\cdot\vec{v}=2\pi n\ \rightarrow\ \vec{k}_1-\vec{k}_2\in\mathbb{L}^*\ \rightarrow\ \vec{k}_2\in(\mathbb{L}^*)_{\vec{k}_1}$$

Take $\vec{k}_1,\vec{k}_2\in(\mathbb{L}^*)_{\vec{k}_1}$, $\vec{v}\in\mathbb{L}$

$$\therefore\ \vec{k}_1=\vec{k}_2+\vec{t}^*:\vec{t}^*\in\mathbb{L}^*\ \rightarrow\ e^{i\vec{k}_1\cdot\vec{v}}=e^{i\left(\vec{k}_2+\vec{t}^*\right)\cdot\vec{v}}=e^{i\vec{k}_2\cdot\vec{v}}$$

Thus, as before, the values $\left\{e^{i\vec{k}\cdot\vec{v}}:\vec{k}\in\mathbb{L}^*_{\vec{k}}\right\}$ enumerate the distinct cosets of $\mathbb{L}^*$. Where we are constructing a sublattice of index N, this set must have N distinct values.

With this in mind, for a given sublattice $\mathbb{L}_{\vec{k}}$, we can show that the set of all possible unique generators of $\mathbb{L}_{\vec{k}}$ is a subset of $<\vec{k}>$.

Proof. Each of the N $\vec{k}\in<\vec{k}>$ falls in a distinct $\mathbb{L}^*$ coset.

Considering the c_i coordinates of $\vec{k}$, being $\mathbb{L}^*$ equivalent corresponds to $c_i\ \%\ 1$ equivalence. By construction of $<\vec{k}>$ as a direct product group, each $\vec{k}\in<\vec{k}>$ must be un-equivalent, as $<\vec{k}>$ contains N elements.

From before, given $\vec{k}$ generates $\mathbb{L}_{\vec{k}}$ it follows that $\vec{k}\in\mathbb{L}^*_{\vec{k}}$.

Noting that $\mathbb{L}^*_{\vec{k}}=\bigcup_{\vec{k}_i\in<\vec{k}>}(\mathbb{L}^*)_{\vec{k}_i}$, it follows that $\vec{k}\in(\mathbb{L}^*)_{\vec{k}_i}$

In conclusion then, $\vec{k}=\vec{k}_i+\vec{z}:\vec{k}_i\in<\vec{k}>,\vec{z}\in\mathbb{Z}^3$ □

Summarizing these findings, $\mathbb{L}_{\vec{k}}$ can only be generated by $\vec{k} = \vec{k}_i \in < \vec{k} > + \vec{z} \in \mathbb{Z}^3$. Thus, to enumerate all $\vec{k}$ sub-lattices of index N it is sufficient to construct all possible $\vec{k}$ vectors with coordinates $c_i = m_i/n_i \in [0, 1)$ satisfying $\text{lcm}(\{n_i\}) = N$. From this set of $\vec{k}$, we can successively filter out all other $\vec{k}_i \in < \vec{k} >$ so that only one representative from each cyclic group remains. This final set of $\vec{k}$ represents all distinct $\mathbb{L}_{\vec{k}}$ sublattices of index N.

While the sublattices $\mathbb{L}_{\vec{k}}$ can be enumerated in this form, it is useful to also calculate the set of generating vectors for $\mathbb{L}_{\vec{k}}$. For this discussion we will turn to the following theorem:

Theorem 2.0.9. *Given $\tilde{\mathbb{L}} \leq \mathbb{L}, \tilde{\mathbb{L}}$ 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 [13].*

Proof. 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 that the the term $\left| \frac{\xi_k}{|z_k|} - s \right|$ lies in the interval $[0, 1)$; thus, we can say $\left| \frac{\xi_k}{|z_k|} - s \right| < 1$, 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. □

Building off this theorem, we can design an algorithm to find the upper-triangular matrix A_k, that will derive the generators of $\mathbb{L}_{\vec{k}}$ via right multiplication of the generators of $\mathbb{L}$, arranged as column vectors of a matrix. The algorithm can be summarized as follows:

1. Fill each column of A_k one-by-one
2. For column (i), find the smallest integer linear combination of $c_j : j \in [1, i)$ that forms the largest denominator (keeping the numerator and denominator coprime), store this sum as $c_s = m_s/n_s$
3. Find the greatest common denominator of n_i and n_s and divide each to determine the unique factors in n_s and n_i. Then multiply both c_i and c_s by the others' unique factors so that both share a common denominator.
4. Take the new c_i value and add the new c_s value repeatedly until an integer results.
5. Keeping track of coefficients, arrange them in the matrix's i^{th} column vector.

Take as an example, the unique $\vec{k}$ sub-lattices of index 4 for a lattice $\mathbb{L}$. The unique $\vec{k}$ can be enumerated as follows (listed in $\mathbb{L}^*$ relative coordinates):

$$\begin{pmatrix} 1/4 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1/4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/4 \\ 1/4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/4 \\ 0 \end{pmatrix}, \begin{pmatrix} 3/4 \\ 1/4 \\ 0 \end{pmatrix}, \begin{pmatrix} 1/4 \\ 1/2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1/4 \end{pmatrix}$$
$$\begin{pmatrix} 1/4 \\ 0 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 0 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 3/4 \\ 0 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 0 \\ 1/4 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/4 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 3/4 \\ 1/4 \\ 1/4 \end{pmatrix}$$
$$\begin{pmatrix} 0 \\ 1/2 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 1/4 \\ 1/2 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 1/2 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 3/4 \\ 1/2 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 0 \\ 3/4 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 1/4 \\ 3/4 \\ 1/4 \end{pmatrix}, \begin{pmatrix} 1/2 \\ 3/4 \\ 1/4 \end{pmatrix}$$

$$\begin{pmatrix}3/4\\3/4\\1/4\end{pmatrix},\begin{pmatrix}1/4\\0\\1/2\end{pmatrix},\begin{pmatrix}0\\1/4\\1/2\end{pmatrix},\begin{pmatrix}1/4\\1/4\\1/2\end{pmatrix},\begin{pmatrix}1/2\\1/4\\1/2\end{pmatrix},\begin{pmatrix}3/4\\1/4\\1/2\end{pmatrix},\begin{pmatrix}1/4\\1/2\\1/2\end{pmatrix}$$

Then constructing a corresponding matrix $A_{\vec{k}}$ for each unique $\vec{k}$:

$$\begin{pmatrix}4&0&0\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}1&0&0\\0&4&0\\0&0&1\end{pmatrix},\begin{pmatrix}4&3&0\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}2&1&0\\0&2&0\\0&0&1\end{pmatrix},\begin{pmatrix}4&1&0\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}4&2&0\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}1&0&0\\0&1&0\\0&0&4\end{pmatrix}$$
$$\begin{pmatrix}4&0&3\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}2&0&1\\0&1&0\\0&0&2\end{pmatrix},\begin{pmatrix}4&0&1\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}1&0&0\\0&4&3\\0&0&1\end{pmatrix},\begin{pmatrix}4&3&3\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}2&1&0\\0&2&3\\0&0&1\end{pmatrix},\begin{pmatrix}4&1&1\\0&1&0\\0&0&1\end{pmatrix}$$
$$\begin{pmatrix}1&0&0\\0&2&1\\0&0&2\end{pmatrix},\begin{pmatrix}4&2&3\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}2&1&1\\0&1&0\\0&0&2\end{pmatrix},\begin{pmatrix}4&2&1\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}1&0&0\\0&4&1\\0&0&1\end{pmatrix},\begin{pmatrix}4&1&3\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}2&1&0\\0&2&1\\0&0&1\end{pmatrix}$$
$$\begin{pmatrix}4&3&1\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}4&0&2\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}1&0&0\\0&4&2\\0&0&1\end{pmatrix},\begin{pmatrix}4&3&2\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}2&1&0\\0&2&2\\0&0&1\end{pmatrix},\begin{pmatrix}4&1&2\\0&1&0\\0&0&1\end{pmatrix},\begin{pmatrix}4&2&2\\0&1&0\\0&0&1\end{pmatrix}$$

These sublattices are unique in the sense that they contain unique subsets of lattice points of $\mathbb{L}$. In the geometric sense, some may be equivalent (aka. being related via an affine transformation). Given a lattice $\mathbb{L}$ has a symmetry operation represented by $A_g = [\mathbb{L}]\mu[\mathbb{L}]^{-1}$, where μ is a unimodular matrix, and $[\mathbb{L}]$ is a matrix with $\mathbb{L}$'s generators arranged as column vectors, the reciprocal lattice $\mathbb{L}^*$ will also have the same symmetry operation given by $A_g = [\mathbb{L}^*]\left(\mu^t\right)^{-1}[\mathbb{L}^*]^{-1}$. We can see this is true following the algebra below.

Proof. By definition of the reciprocal lattice, $[\mathbb{L}]^t[\mathbb{L}^*] = 2\pi I$

$$A_g = [\mathbb{L}]\mu[\mathbb{L}]^{-1} \;\rightarrow\; A_g^t = \left([\mathbb{L}]^{-1}\right)^t \mu^t[\mathbb{L}]^t$$
$$A_g^t[\mathbb{L}^*] = \left([\mathbb{L}]^{-1}\right)^t \mu^t\,(2\pi) \;\rightarrow\; [\mathbb{L}^*]^t A_g = \mu[\mathbb{L}]^{-1}\,(2\pi)$$
$$[\mathbb{L}^*]^t A_g \left([\mathbb{L}^*]^t\right)^{-1} = (2\pi)\,\mu\left(\frac{1}{2\pi}\right) \;\rightarrow\; [\mathbb{L}^*]^t A_g = \mu[\mathbb{L}^*]^t$$
$$A_g^t[\mathbb{L}^*] = [\mathbb{L}^*]\mu^t \;\rightarrow\; [\mathbb{L}^*]\mu^t[\mathbb{L}^*]^{-1}$$

Noting that $A_g^t = A_g^{-1}$, as affine transformations can be represented by unitary matrices.

$$A_g = [\mathbb{L}^*]\left(\mu^t\right)^{-1}[\mathbb{L}^*]^{-1}$$

□

Note that to be a lattice isometry, A_g must coincide with a unimodular matrix when put in the coordinate system of the lattice. Here we see that A_g does correspond to a unimodular matrix in both the $\mathbb{L}$ and $\mathbb{L}^*$ coordinate systems.

Thus given $\vec{k}$ in the coordinate system of $\mathbb{L}^*$, multiplying $\vec{k}$ by $(\mu^t)^{-1}$, the unimodular matrix from above, performs the action of the isometry A_g. Because $\vec{k}$ and $A_g\vec{k}$ are symmetrically equivalent, they will both generate the same sublattice $\mathbb{L}_{\vec{k}}$. With this knowledge, it is possible to filter out all geometrically equivalent lattices $\mathbb{L}_{\vec{k}}$.

Of course, a crystal is composed of a lattice and a basis. Depending on the geometry of the basis, even when two $\mathbb{L}_{\vec{k}}$ sublattices are geometrically equal, the actual crystals may still not coincide due to a difference in basis. In general, this must be checked computationally.

When $\mathbb{L}_{\vec{k}}$ is generated, multiplication by $A_{\vec{k}}$ can be used to generate the new unit cell; however, the basis must also be adjusted to preserve the overall crystal structure. In general the original basis vectors remain unchanged in terms of Cartesian coordinates; however, for every non-identity coset in $\mathbb{L}/\mathbb{L}_{\vec{k}}$, a "missing" translation must be added to form a new basis vector for each original basis atom. This can be accomplished via the following algorithm for a given $\mathbb{L}_{\vec{k}}$ sublattice:

1. Find the smallest integer linear combination of c_i that forms the largest denominator (keeping the numerator and denominator coprime), store this sum as $c_s = m_s/n_s$, keeping track of the coefficients of the sum.

2. Then, adding c_s to itself over and over, each new term falls in a distinct $\mathbb{L}/\mathbb{L}_{\vec{k}}$ coset. The coefficients of each sum correspond to integer linear combinations of the generating set of $\mathbb{L}$ that must be added together to form a "missing" translation.

In general, a lattice may have additional sub-lattices that aren't generated by a particular $\vec{k}$. 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 sub-lattice 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 sub-lattice geometrically. Further, if two γ matrices γ_1 and γ_2 generate the same lattice geometrically then the

following is true:

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

We consider now the following transformations for 3-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 sub-lattice.

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 sub-lattices.

Revisiting the sublattices of index 4, the following additional matrices are revealed using the above γ matrix generation technique:

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$$

Together with the previously-listed matrices, these provide a complete enumeration of all unique index-4 sub-lattices of a given lattice.

For each of these matrices, the "missing" translations can again be found by enumerating the cosets of the factor group $\mathbb{L}/\tilde{\mathbb{L}}$, where $\tilde{\mathbb{L}}$ is a sublattice of $\mathbb{L}$. In general, these coset representatives will have fractional coordinates in the unit cell setting of $\tilde{\mathbb{L}}$, as $[\vec{v}]_{\tilde{\mathbb{L}}} = \gamma^{-1}[\vec{v}]_{\mathbb{L}}$.

Because γ is an upper-triangular matrix, it can be shown that its inverse has diagonal elements: $(\gamma^{-1})_{ii} = \frac{1}{\gamma_{ii}}$. Then, the coset representatives can be enumerated, in coordinates of $[\tilde{\mathbb{L}}]$ with the following algorithm:

1. For non-zero representatives of the form $< \frac{z_1}{\gamma_{11}}, 0, 0 >$, there are $\gamma_{11} - 1$ inequivalent representatives.

2. For non-zero representatives of the form $< \frac{z_1}{\gamma_{11}} + \frac{z_2 m_2}{\gamma_{11}\gamma_{22}}, \frac{z_2}{\gamma_{22}}, 0 >$, there are $\gamma_{11}(\gamma_{22} - 1)$ additional inequivalent representatives.

3. For non-zero representatives of the form $< \frac{z_1}{\gamma_{11}} + \frac{z_2 m_2}{\gamma_{11}\gamma_{22}} + \frac{z_3 m_3'}{\gamma_{11}\gamma_{22}\gamma_{33}}, \frac{z_2}{\gamma_{22}} + \frac{z_3 m_3}{\gamma_{22}\gamma_{33}}, \frac{z_3}{\gamma_{33}} >$, there are $\gamma_{11}\gamma_{22}(\gamma_{33} - 1)$ additional inequivalent representatives.

4. Together these form the $\gamma_{11}\gamma_{22}\gamma_{33} - 1$ non-zero coset representatives of the factor group.

In summary, a given crystal structure has a unique primitive representation in terms of a primitive lattice; however, it has a finite number of larger sublattices (N times as large), and it has an infinite number of possible sublattice sizes $N \in \mathbb{N}_1$. For each choice of sublattice, we can apply the results in the preceding section, representing the crystal structure in terms of 7 lattice vonorms and $3(n-1)$ basis degrees of freedom. For different choices of N, the coordinate system changes, resulting in $7 + 3(Nn - 1)$ degrees of freedom. The increased degrees of freedom conceptually result as a lattice with fewer translations affords fewer geometric constraints. Physically, the $7 + 3(n-1)$ primitive lattice and motif degrees of freedom correspond to all possible strain and gamma-point phonon perturbations of the crystal. As N increases, the larger supercells' degrees of freedom correspond to all possible strains and additional phonon modes with non-zero wavelengths in coordinates of the primitive reciprocal lattice. In general, given a sublattice of index N, these new phonon modes will have wave-vectors $[\vec{k}]_{\mathbb{L}^*} \sim < a/N, b/N, c/N >, a, b, c \in \mathbb{Z}$.

As a final note, in calculating larger sublattices of index N, the crystal structure will have several geometrically equivalent crystal structures, each with a different lattice geometry and corresponding basis. This implies that when represented in the $7 + 3(Nn - 1)$ dimensional space, the same crystal structure will correspond to many different vectors, all derivable using the techniques above. This redundancy inherently arises from the fundamental representation of a crystal structure in terms of a lattice and a basis.

Strains

Any two 3D lattices can be brought into geometric coincidence by means of an invertible mapping function $f : \mathbb{R}^3 \to \mathbb{R}^3$. If we are to consider diffusion-less transformations, we are forced to consider smooth transformations with no shuffling of lattice points along any linear trajectory. In this sense we immediately have a restriction on f. Given $a\vec{v}$, $b\vec{v}$, such that $|a\vec{v}| > |b\vec{v}|$, then it must follow that $|f(a\vec{v})| > |f(b\vec{v})|$ $\forall a, b \in \mathbb{R}, \vec{v} \in \mathbb{R}^3$. Further, noting

that a lattice is a linear integer combination of three generating vectors, the function f must itself be a linear mapping function, otherwise the result of the mapping could not be a lattice.

Any linear transformation can be represented as a 3x3 matrix multiplication on $\mathbb{R}^3$. We note that to avoid inversion through the origin, this transformation must have a positive determinant. As such, we note that proper rotations and positive definite stretches satisfy all aforementioned properties of a valid lattice mapping function.

It can be shown that for any invertible matrix, A, of positive determinant, A can be represented as a product of a unitary and a symmetric matrix: $A = R_1T_1 = T_2R_2$. The precise order is unimportant; though, the matrices will differ depending on the order chosen. Here R_1, R_2 are proper rotation matrices, and T_1, T_2 are stretch matrices. By computing the quantities A^tA and AA^t, we immediately see connections between these quantities:

$$A^tA = T_1^2 = R_2^tT_2^2R_2$$

$$AA^t = R_1T_1^2R_1^t = T_2^2$$

$$\therefore T_1^2 = R_1^tT_2^2R_1 \quad \rightarrow \quad R_1 = R_2,\ T_1 \sim T_2$$

Given that A maps vectors from lattice $\mathbb{L}_1$ to lattice $\mathbb{L}_2$, it must follow that there exists a corresponding matrix $[A]$ that maps all integer coordinates of lattice points in $\mathbb{L}_1$ onto other integer coordinates of lattice points in $\mathbb{L}_2$. These two matrices are related intuitively below with matrices $[\mathbb{L}]$ corresponding to a matrix whose column vectors correspond to the generators of lattice $\mathbb{L}$:

$$A\vec{v} = [\mathbb{L}_2][A][\mathbb{L}_1]^{-1}\vec{v}$$

As $[A]$ links integer coordinates to other integer coordinates for all possible inputs in $\mathbb{Z}^3$, each of its column vectors must consist of only integers, aka. $[A] \in GL_3(\mathbb{Z})$. Given that A is invertible, then $[A]$ must also be invertible. As such:

$$A^{-1} = [\mathbb{L}_1][A]^{-1}[\mathbb{L}_2]^{-1}$$

Following the same logic, it must follow that $[A]^{-1}$ consists only of integers, aka. $[A]^{-1} \in GL_3(\mathbb{Z})$.

Examining the determinant, it follows that $|[A]| \in \mathbb{Z}$ and $|[A]^{-1}| \in \mathbb{Z}$. Further, $|[A]^{-1}| = |[A]|^{-1}$. The only integers satisfying this constraint are ± 1. Where $|\mathbb{L}_1| > 0$ and $|\mathbb{L}_2| > 0$, it follows that the sign of $|A|$ equals the sign of $|[A]|$. Since we are only interested in linear transformations with positive determinant, aka. no inversion, this means that $|[A]| = 1$. Thus, the only valid possibilities for $[A]$ are matrices of all integers with determinant $+1$,

aka. $[A] \in SL_3(\mathbb{Z})+$, a positive unimodular matrix.

Summarizing these results in compact set notation:

$$A \in \left\{ [\mathbb{L}_2][A][\mathbb{L}_1]^{-1} : [A] \in SL_3(\mathbb{Z})+ \right\}$$

By calculating various possible values for A and performing a polar decomposition, the set of all strains and rotations linking two lattices can be enumerated. Care must be taken as this is an infinite set; however, the search for small strains connecting two lattices can be efficient and finitely truncated. Given a maximum strain cutoff defined by the Pythagorean sum of the strain eigenvalues: $pss := \sum \sqrt{\epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2}$, a finite subset of positive unimodular matrices can be considered. This finite subset is composed of all unimodular matrices whose maximum column vector norm is below a certain threshold. The equation below relates this threshold to the strain cutoff [9]:

$$|[A]|_C \leq \frac{|[\mathbb{L}_1]|_C}{\nu_{min}([\mathbb{L}_2])} (pss + 1)$$

Here we denote the maximum column vector norm of a matrix M to be $|M|_C$ and the minimum singular value of a matrix M to be $\nu_{min}(M)$.

DFT Overview

In order to study the variation of crystal energy with respect to crystal geometry, this dissertation uses density functional theory (DFT) to estimate the energy. While the postulates of quantum mechanics and the corresponding Dirac equation are technically sufficient to predict the behavior and properties of any atomic system, the mathematical complexity proves to be too much for most practical systems. Particularly, for the field of solid-state-physics, the non-relativistic Kohn-Sham equations are a useful approximation for the ground-state energy and electronic structure of the system. This approach is detailed below following the text [14].

The Kohn-Sham equations and DFT are built up from two main theorems:

1. For any system, the potential energy of a single electron $v(\vec{r})$ is uniquely determined by the ground state electronic density $\rho(\vec{r})$ up to a non-physical constant, and visa-versa [14]

2. A universal functional $E[\rho, v_{ext}]$ exists, for fixed v_{ext}. The ρ minimizing E corresponds to to the exact ground state charge density, with E giving the exact ground state energy [14]

This represents an enormous simplification relative to the traditional quantum treatment, as one no-longer needs to deal with a many-body wavefunction, rather, one may deal solely with ρ, a function of 3 dimensions. Further, the conclusions apply quite generally, for any given externally applied potential v_{ext} and any set of self-interacting particles. However, the downside to this approach is that the theorems yield no insight into the form of this energy functional $E[\rho, v_{ext}]$, nor any techniques to derive it.

With time, approaches have been developed to approximate this energy functional, which is usually broken up into the following terms: [14]

$$E_{Kohn-Sham} = T[\rho] + \int v_{ext}(\vec{r})\rho(\vec{r})d^3\vec{r} + \frac{1}{2}e^2 \int \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{|\vec{r}_1 - \vec{r}_2|} d^3\vec{r}_1 d^3\vec{r}_2 + E_{XC}$$

The first term is the kinetic energy of the system. The second term is the energy due to the imposed external potential. The third term is called the Hartree term, representing electron-electron self-repulsion. And the final term is the exchange-correlation term, left purposefully undefined.

Using the exact solution from a homogeneous electron gas, the local-density-approximation functional was developed for E_{XC} in which the exchange energy depends linearly on the charge density. This was a useful first step; however, it has several systematic errors, including underestimating lattice parameters.

Other solutions have included GGA functionals, which depend on both the charge density and its gradient. This thesis, in particular, uses the GGA functional developed by Perdew, Burke, and Ernzerhof (PBE-sol), as it provides a better energy estimate for most solids.

To represent the charge density, a set of basis functions is used, here corresponding to planewaves, as they are aptly suited to the periodic boundary conditions present in crystal lattices. These planewaves will include reciprocal lattice vectors, as well as other $\vec{k}$ values in between. These $\vec{k}$ points form a grid that extends outward until a maximum kinetic cutoff energy is reached for practical calculation purposes. Typically this is a safe approximation as ρ is a function with finite curvature and derivatives. By choosing the smallest three vectors of reciprocal lattice's obtuse superbasis to define the primitive unit cell, the $\vec{k}$ values can be sampled in a nearly-uniform grid with angles between 90 and 120 degrees, leading to reliable performance.

In general, it is expensive and infeasible to solve the Kohn-Sham equations with all the electrons included for every atom within the crystal. Firstly, solving the equations for one unit cell will provide a solution for the entire crystal by lattice symmetry. Further, it is possible to ignore the behavior of core electrons in most cases, as they remain tightly localized to the nucleus and do not significantly delocalize throughout the crystal. These atoms

can be treated as fixed and provide a fixed pseudo-potential surrounding the atomic cores. Then, we can confine our attention to just the valence electrons and solve the Kohn-Sham equations expeditiously [14].

Practically, the Kohn-Sham equations are solved using a self-consistent approach [14]. An initial guess of the charge density then generates v_{ext} and allows for one to solve for a series of non-interacting pseudo-wavefunctions. These then allow one to re-calculate the charge density. The process repeats until the ground state energy no longer changes significantly from iteration to iteration. This process is implemented using the Vienna ab-initio Simulation Package in this thesis.

Chapter 3

Study 1: A Crystallographic Map

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 behavior such as piezoelectricity, mechanical stability, ferroelectricity, etc. However, techniques for exploring this energetic landscape, particularly to find low-energy transformation pathways, have remained limited.

In this study, we delineate an algorithm for fully exploring crystallographic phase space in a non-redundant, incremental, and rigorous fashion. Further, the approach is discretized, robust to numerical precision errors, and proceeds intuitively. Through the algorithm, we provide new avenues for identifying low-energy transformation pathways, exploring arbitrary paths through phase space, identifying and generating similar crystal structures, and representing crystals numerically. With this work, 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.

Representation of 3D Lattices

Building off the background presented in the previous methodology, we start with the seven squared vonorms of a lattice and their accompanying dot products. The space of all 3D lattices can be represented as integers, and geometrically-similar lattices can be easily found using this representation.

In general, 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. 2.5] 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 transforma-

tions or a re-ordering of the vonorms and dot products.

For the purposes of this and future study, we are concerned only with the effect of crystal geometry on the energy of the system; thus, affine transformations of the crystal can be considered equal in energy. In the presence of external fields, for example, this would not be the case, and orientation degrees of freedom would need to be included in a description of the system.

Studying the re-ordering 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 $\left\{(\vec{v_0}+\vec{v_1})^2, (\vec{v_0}+\vec{v_2})^2, (\vec{v_0}+\vec{v_3})^2\right\}$ "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. 2.9] 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 re-ordered primary and secondary squared vonorms. We choose a unique representative (L) from this orbit by choosing the permutation for which the seven ordered lengths, $\left\{v_0^2, v_1^2, v_2^2, v_3^2, (\vec{v_0}+\vec{v_1})^2, (\vec{v_0}+\vec{v_2})^2, (\vec{v_0}+\vec{v_3})^2\right\}$, 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 $\leq$ the first (j - 1) elements of (M) AND the jth element of (L) is strictly less than the jth element of (M). Thus, for any lattice, we can select a unique ordered string of lengths to represent its geometry unambiguously.

As en 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) [15]:

$$\left\{v_0^2 = 19.2,\ v_1^2 = 21.3,\ v_2^2 = 19.2,\ v_3^2 = 21.3,\right.$$
$$\left.(\vec{v_0}+\vec{v_1})^2 = 40.5,\ (\vec{v_0}+\vec{v_2})^2 = 19.2,\ (\vec{v_0}+\vec{v_3})^2 = 21.3\right\}$$

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.

$$\left\{v_0^2 = 19.2,\ v_1^2 = 19.2,\ v_2^2 = 21.3,\ v_3^2 = 21.3,\right.$$
$$\left.(\vec{v_0}+\vec{v_1})^2 = 19.2,\ (\vec{v_0}+\vec{v_2})^2 = 40.5,\ (\vec{v_0}+\vec{v_3})^2 = 21.3\right\}$$

Seeing that square vonorms v_0^2 and v_1^2 are 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 re-ordered. The result is maximally-ascending:

$$\left\{v_0^2 = 19.2,\ v_1^2 = 19.2,\ v_2^2 = 21.3,\ v_3^2 = 21.3,\right.$$

$$\left.(\vec{v_0} + \vec{v_1})^2 = 19.2,\ (\vec{v_0} + \vec{v_2})^2 = 21.3,\ (\vec{v_0} + \vec{v_3})^2 = 40.5\right\}$$

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 the equation below remains true.

$$-\left(v_0^2 + v_1^2 + v_2^2 + v_3^2\right) + (\vec{v_0} + \vec{v_1})^2 + (\vec{v_0} + \vec{v_2})^2 + (\vec{v_0} + \vec{v_3})^2 = 0 \tag{3.1}$$

This is a necessary invariant of all 3D lattices [12]. Where ξ is small, the new lengths correspond to a lattice, one that is related to the original by a small strain. 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. We can see this in the proof below:

Theorem 3.0.1. *1: Nearest neighbor lattices are connected by small strains*

Proof. Assume there exist 2 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 3x3 matrix with 3 generating vectors of $\mathbb{L}$ arranged as column vectors.

Then, it follows: $\langle A\vec{v_i}, A\vec{v_i}\rangle = \langle \vec{v_i}, \vec{v_i}\rangle + \xi_i$, 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 i^{th} square vonorm.

Re-arranging: $\left\langle \vec{v_i}, A^T A\vec{v_i}\right\rangle - \langle \vec{v_i}, \vec{v_i}\rangle = \xi_i$, using the standard definitions of matrix multiplication and conventional cartesian inner product on $\mathbb{R}^n$.

Re-arranging: $\xi_i = v_i^2 \left\langle \hat{v_i}, \left(A^T A - I\right) \hat{v_i}\right\rangle$, 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 $A^T A - I$ 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: $\hat{v_i} = \sum_j c_j \hat{b_j}$

Re-arranging: $\frac{\xi_i}{v_i^2} = \sum_{j,k} \left\langle c_j \hat{b_j}, \lambda_k c_k \hat{b_k} \right\rangle = \sum_j c_j^2 \lambda_j$

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. □

It is possible to use a measuring unit of small size ξ. Then, by rounding each vonorm to the nearest ξ, while preserving [eq. 2.9], 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 (explicit enumeration in the appendix). Referencing [eq. 2.9], 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 non-strictly 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 non-positive dot products. In floating-point arithmetic, due to limited numerical precision, it can be difficult to determine whether a dot product is positive [16]. 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. Further, by design, if lattice A is a neighbor of lattice B, it follows that lattice B is a neighbor of lattice A. These proofs are shown below:

Theorem 3.0.2. *Every lattice can be connected to every other lattice via a series of nearest neighbors*

Proof. We start by referencing [Theorem 3.0.3]: 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 the appendix, we can calculate the image of each modification when multiplied by the above matrix. Notice only the first 6 square vonorms are listed: the 7th is uniquely implied by [eq. 2.10].

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

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

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 non-positive. 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.

$$\vec{\chi_i}+\vec{\chi_j}... = \begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix}+\begin{pmatrix}0\\0\\1\\0\\0\\1\end{pmatrix}+\begin{pmatrix}0\\1\\0\\0\\1\\0\end{pmatrix}+\begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix}+\begin{pmatrix}0\\0\\1\\0\\0\\1\end{pmatrix}+\begin{pmatrix}1\\0\\0\\0\\1\\0\end{pmatrix}+\begin{pmatrix}0\\0\\0\\1\\0\\0\end{pmatrix}+\begin{pmatrix}0\\1\\0\\0\\0\\1\end{pmatrix}+\begin{pmatrix}1\\0\\0\\0\\1\\0\end{pmatrix}+\begin{pmatrix}0\\0\\1\\0\\0\\0\end{pmatrix}+\begin{pmatrix}0\\1\\0\\0\\0\\1\end{pmatrix}+\begin{pmatrix}1\\0\\0\\0\\1\\0\end{pmatrix}$$

$A(\vec{\chi_i} + \vec{\chi_j}...) =$

$$\frac{1}{2}\left(\begin{pmatrix}0\\0\\0\\1\\-1\\-1\end{pmatrix}+\begin{pmatrix}0\\0\\0\\-1\\1\\-1\end{pmatrix}+\begin{pmatrix}0\\0\\0\\-1\\-1\\1\end{pmatrix}+\begin{pmatrix}0\\0\\0\\1\\-1\\-1\end{pmatrix}+\begin{pmatrix}0\\0\\0\\-1\\1\\-1\end{pmatrix}+\begin{pmatrix}0\\-1\\-1\\0\\0\\1\end{pmatrix}+\begin{pmatrix}0\\0\\0\\1\\-1\\-1\end{pmatrix}+\begin{pmatrix}-1\\1\\0\\-1\\0\\0\end{pmatrix}+\begin{pmatrix}0\\-1\\-1\\0\\0\\1\end{pmatrix}+\right.$$

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

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 that 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 $\vec{\chi_i} = \langle 0,0,0,0,-1,0,1\rangle$, 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 $\vec{\chi_i} = \langle 0,0,0,0,-1,1,0\rangle$, 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. 2.10], the value of 6 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 6 square vonorms, they must have the same 7 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. □

Theorem 3.0.3. *Lattice neighbor relationships are reciprocal.*

Proof. 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 re-arrange this equation algebraically:

$$g^{-1}(\vec{vo}_B - g(\vec{\chi_i})) = \vec{vo}_A$$

Referencing the appendix where $\{\vec{\chi_i}\}$ 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:

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

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

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.

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 $[0, 1)$, arbitrary lattice translations are accounted for. By confining a particular atom to the origin, arbitrary origin is accounted for, and only $n-1$ atomic positions need to 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 non-origin atoms will not change the crystal structure; however, the listing will be re-ordered.

Using the same technique employed with the lattice length strings above, we can calculate the full orbit of strings of basis vectors 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 coordinate, 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 [17]. 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:

2 atoms Sn @0, 0, 0 and @0.5, 0.5, 0.5
4 atoms O @0, 0.7, 0.7 and @0, 0.3, 0.3 and @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 4 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, 5,2,2 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. Together these generate $8n$ neighbors for any crystal basis. If δ is sufficiently large, these neighbors are similar geometrically to the original basis.

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 7 integers that are in an unambiguous order, the following formulas [eq. 3.2-3.9] will always define 3 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. Since the crystal's orientation is arbitrary in this study, a variable orientation arises naturally when using these formulae.

$$\vec{v_0} = \xi v_0 \langle 1, 0, 0 \rangle \tag{3.2}$$

$$\vec{v_1} = \xi v_1 \langle x, y, 0 \rangle \tag{3.3}$$

$$x = \frac{\vec{v_0} \cdot \vec{v_1}}{\xi^2 v_0 v_1} \tag{3.4}$$

$$y = \left(1 - x^2\right)^{1/2} \tag{3.5}$$

$$\vec{v_2} = \xi v_2 \langle a, b, c \rangle \tag{3.6}$$

$$a = \frac{\vec{v_0} \cdot \vec{v_2}}{\xi^2 v_0 v_2} \tag{3.7}$$

$$b = \frac{1}{y}\left(\frac{\vec{v_1} \cdot \vec{v_2}}{\xi^2 v_1 v_2} - xa\right) \tag{3.8}$$

$$c = \left(1 - a^2 - b^2\right)^{1/2} \tag{3.9}$$

We can show this unique connection of vonorms to generators via the following proof:

Theorem 3.0.4. *The lattice mapping function* $\mathbb{L}(\vec{vo})$ *is injective.*

Proof. 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 $\mathbb{L}(\vec{vo}_1) = \mathbb{L}(\vec{vo}_2)$. 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 $\vec{vo}_1$ and $\vec{vo}_2$.

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 $\vec{vo}_1$ and $\vec{vo}_2$.

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 $\vec{vo}_1$ and $\vec{vo}_2$, it must follow that $(\vec{v_0} + \vec{v_1})^2$ is equal for both $\vec{vo}_1$ and $\vec{vo}_2$.

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 $\vec{vo}_1$ and $\vec{vo}_2$, it must follow that $(\vec{v_0} + \vec{v_2})^2$ is equal for both $\vec{vo}_1$ and $\vec{vo}_2$.

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 $\vec{vo}_1$ and $\vec{vo}_2$, it must follow that v_3^2 is equal for both $\vec{vo}_1$ and $\vec{vo}_2$.

Combining these results with [eq. 2.10], because 6 square vonorms are equal, all 7 square vonorms must be equal between $\vec{vo}_1$ and $\vec{vo}_2$. Thus, $\vec{vo}_1 = \vec{vo}_2$. □

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 7 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 7 values to calculate the corresponding dot products, using [eq. 2.9]. Finally, $\mathbb{L}$ outputs the three generating vectors as column vectors of a 3x3 matrix ordered as follows: $[\vec{v_0}, \vec{v_1}, \vec{v_2}]$

Now we recall the symmetry group S_4, which can permute the order of square vonorms but leaves the geometry of the lattice unchanged. In the context of the function $\mathbb{L}$, each $g \in S_4$ can lead to a distinct unit cell and orientation. 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 (shown below). 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 (shown below).

Theorem 3.0.5. *The group S_4 has an action on set of lattice generators $\mathbb{L}(\vec{vo})$ arranged as column vectors. This action is as follows: $\mathbb{L}(g(\vec{vo})) = A_g\mathbb{L}(\vec{vo})\mu_g^{-1} : g \in S_4$, where A_g is a unitary rotoinversion and μ_g is a unimodular matrix.*

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

$$\mathbb{L}(g_1 \cdot g_2(\vec{vo})) = A_{g_1 \cdot g_2}\mathbb{L}(\vec{vo})\mu_{g_1 \cdot g_2}^{-1} = A_{g_1}A_{g_2}\mathbb{L}(\vec{vo})\mu_{g_2}^{-1}\mu_{g_1}^{-1} = \mathbb{L}(g_1(g_2(\vec{vo})))$$

$$\mathbb{L}(e(\vec{vo})) = A_e^{-1}\mathbb{L}(\vec{vo})\mu_e = \mathbb{L}(\vec{vo})$$

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 re-ordered. The geometry of the lattice is preserved; although, the mapping function $\mathbb{L}(\vec{vo})$ 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 $\mathbb{L}(\vec{vo})$ 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 $\mathbb{L}(\vec{vo})$, 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.

Re-arranging 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 occur signifies that $|A_g| > 0$, thus, $|\mu_g| > 0$.

As such, we pick the negative representatives above so that 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. □

Theorem 3.0.6. *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.*

Proof. 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 3 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 and cartesian 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:

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

□

For completeness, we list a generating set of S_4 and their unimodular counterparts μ_g. There is an isomorphism between these groups; thus, any permutation and corresponding matrix can be constructed by the analogous group multiplication (shown below). We restrict μ_g to have a positive determinant so that only proper rotations are required to connect neighboring crystal structures [Theorem 3.0.5]. 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.

$$0 \leftrightarrow 1 : \mu_g = -\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.10}$$

$$0 \leftrightarrow 2 : \mu_g = -\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \tag{3.11}$$

$$0 \leftrightarrow 3 : \mu_g = -\begin{pmatrix} -1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \tag{3.12}$$

$$1 \leftrightarrow 2 : \mu_g = -\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \tag{3.13}$$

$$1 \leftrightarrow 3 : \mu_g = -\begin{pmatrix} 1 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \tag{3.14}$$

$$2 \leftrightarrow 3 : \mu_g = -\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & -1 \end{pmatrix} \tag{3.15}$$

Theorem 3.0.7. *There exists an isomorphism between S_4 and the group of unimodular matrices accomplishing the change of unit cell μ_g*

Proof. Referencing the group action from [Theorem 3.0.5], 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 [Theorem 3.0.5]. 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 $\{\mu_g\}$. □

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 $\tilde{\vec{v}o}$ as the maximally-ascending representative of the S_4 orbit on the string of lattice lengths. Some elements in S_4 will leave $\tilde{\vec{v}o}$ 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 $\mathbb{L}(\tilde{\vec{v}o})$. 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".

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 co-variations alongside the $8n$ crystal basis variations create a set of geometrically-similar crystals, termed here as 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 crystal normal form. 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.

Theorem 3.0.8. *The structure of crystal neighbors is reciprocal and complete.*

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

a) If B is a lattice neighbor:

$$\vec{vo}_B = g(\vec{vo}_A + \vec{\chi_i}) \qquad \mathbb{B}_B = \mu_g \mathbb{B}_A \qquad g \in S_4$$

b) If B is a crystal basis neighbor:

$$\vec{vo}_B = \vec{vo}_A \qquad \mathbb{B}_B = \mu_g(\mathbb{B}_A + \vec{\delta_a}) \qquad g \in S_4$$

We can re-arrange the equations as before:

a) If B is a lattice neighbor:

$$\vec{vo}_A = g^{-1}(\vec{vo}_B - g(\vec{\chi_i})) \qquad \mathbb{B}_A = \mu_g^{-1} \mathbb{B}_B \qquad g \in S_4$$

b) If B is a crystal basis neighbor:

$$\vec{vo}_B = \vec{vo}_A \qquad \mathbb{B}_A = \mu_g^{-1}(\mathbb{B}_B - \mu_g \vec{\delta_a}) \qquad g \in S_4$$

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 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 [Theorem 3.0.2] 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 cell $\mathbb{B}_1$ 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 crystal normal form 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 be repeated also for any different unit cell, and the same results hold, regardless of the unit cell chosen. □

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

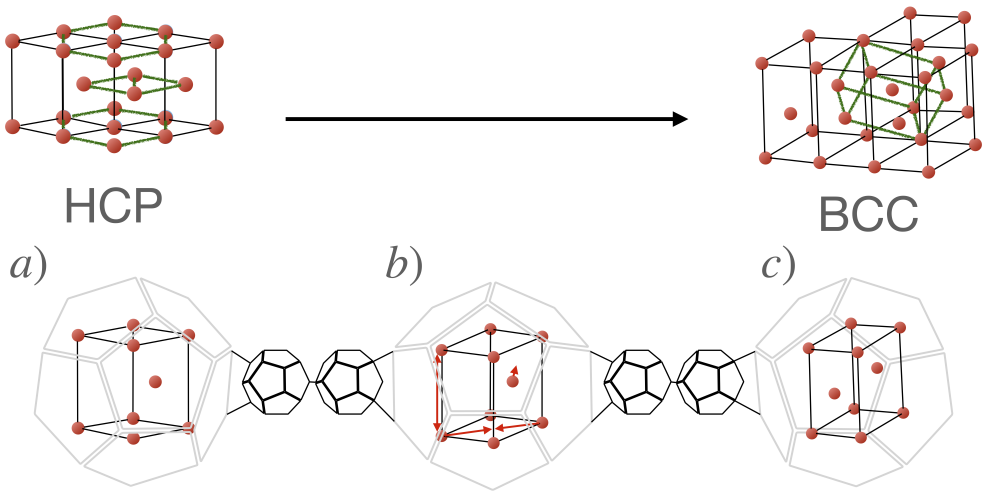

Figure 3.1: The search for transformation pathways from HCP to BCC is visualized below. The hexagonal closest-packed planes are highlighted in green for the HCP structure. The canonical BCC unit cell is highlighted in green for the BCC structure. The unique unit cells defined by canonical mapping function $\mathbb{L}(\vec{vo})$ are shown for HCP (a) and BCC (c). An intermediate structure connects the two (b) with unit cell strain and atomic shuffling visualized via red arrows. Some neighbors are omitted for conceptual clarity. The path highlighted is a shortest path connecting the HCP and BCC crystals. This path closely corresponds to the well-known C' strain, TA1(N) phonon-mode transformation pathway connecting BCC and HCP structures in many metallic systems [18]

.

cell, can be discretized and explored using the parameters ξ and δ. These connections, by design, are insensitive to choice of unit cell, either in the beginning or ending structure, as the vonorms, conorms, and basis normal form of the crystal are a geometrical invariant of the crystal.

As an example, we present a canonical transformation pathway from HCP to BCC found using these techniques in Figure 1. Here two Zirconium atoms per unit cell are searched in the configuration space. We pick a shortest path connecting the structures, signifying a path with the smallest number of nearest-neighbor connections.

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 can 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 a new way to sample phase space using Monte-Carlo approaches to estimate finite temperature properties using ab-initio techniques. 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 [9]. 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 dataset of 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.

Acknowledgements

I would like to sincerely thank Dr. Kristin Persson for her unwavering support and guidance through the duration of this project. 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 (KCD2S2).

Chapter 4

Study 2: The Zr Crystal System

Having created a means of exploring crystallographic phase space, it is now possible to now apply these techniques to find activation energy barriers. As a proof of concept, we will focus on the Zirconium phase space, validating here that this new general approach to finding transformation pathways re-derives known patterns from the literature.

In order to conduct a generalized search for minimal-activation-energy pathways, the following "water-filling" algorithm is employed and can be summarized as follows:

1. Start with a stable crystal structure in canonical integer coordinates.
2. Generate the neighbors of the crystal structure and calculate the energies of those neighbors.
3. Keep track of a maximum energy. For every neighbor under this maximum energy, calculate its neighbors and their associated energies.
4. Repeat the process until all previously-found neighbors are greater in energy than this maximum.
5. Raise the maximum energy and repeat the algorithm until the target crystal structures are found.
6. The maximum energy found upon algorithm termination is the activation energy barrier and all neighbors below this energy form a phase space of physical transformation paths sharing that same activation energy.

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

In order to verify the results of this analysis, we search the 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 index 2 supercell of the FCC or BCC structure. There are two geometrically-unique FCC unit cells of index 2, and there are two geometrically-unique BCC unit cells of index 2. These are found by applying the aforementioned sublattice-generating algorithms and filtering out any structures connected by zero-strain and with equivalent bases. As such, there are four target structures in the search.

We conduct this search using discretization parameters as follows: $\xi = 1.5\mathring{A}^2$ 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: $7, 7, 20, 20, 7, 20, 27, 0, 15, 15$

Recall that the first 7 coordinates correspond to square voronoi vector norms and the final 3 coordinates correspond to the position of the second atom in the unit cell. The first atom is fixed at the origin.

For the transformation from HCP to BCC, the 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 introduction. Each strain is shown at 20% transformation pathway increments:

Start (BCC): 6, 8, 17, 23, 6, 23, 25, 0, 15, 15
$E = -9.012\frac{meV}{atom}$

20%: 6, 6, 20, 25, 6, 25, 26, 0, 15, 15

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

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

$E = -9.023 \frac{meV}{atom}$

40%: 6, 7, 19, 25, 6, 25, 26, 2, 16, 15

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

$$\vec{\delta} =< 2/30, 1/30, 0 >$$

$E = -9.024 \frac{meV}{atom}$

60%: 6, 7, 19, 25, 6, 25, 26, 6, 18, 15

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

$$\vec{\delta} =< 6/30, 3/30, 0 >$$

$E = -9.029 \frac{meV}{atom}$

80%: 6, 7, 18, 25, 7, 24, 25, 9, 17, 15

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

$$\vec{\delta} =< 9/30, 2/30, 0 >$$

$E = -9.051 \frac{meV}{atom}$

100% (HCP): 7, 7, 17, 24, 7, 24, 24, 10, 20, 15

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

$$\vec{\delta} =< 10/30, 5/30, 0 >$$

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

We note that in contrast to conventional strain-shuffling transformation pathways, characterized by linear interpolation, this pathway proceeds as a superposition of micro-strains 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 10-dimensional space in this study. However, it is possible to create a similar plot to those above 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 2 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 first HCP to BCC and then HCP to FCC later on.

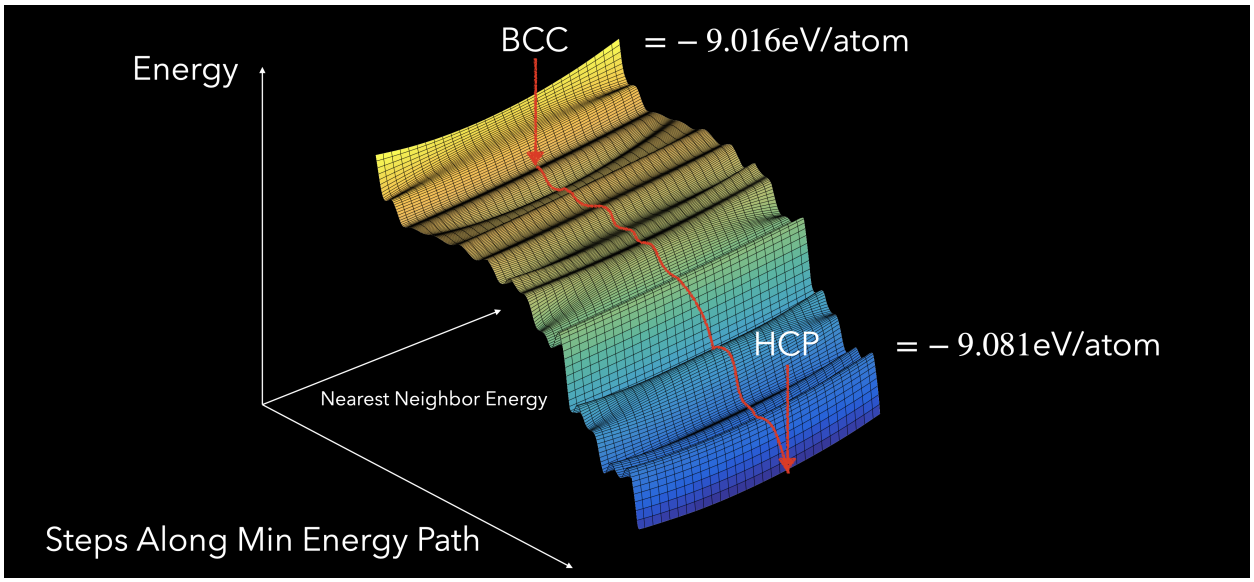

Figure 4.1: The search for transformation pathways from BCC to HCP is visualized with the phase space surrounding a path of minimal activation energy barrier highlighted.Along the x-axis, the number of steps along the transformation pathway are shown. Along the y-axis two of the lowest energy orthogonal 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 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.

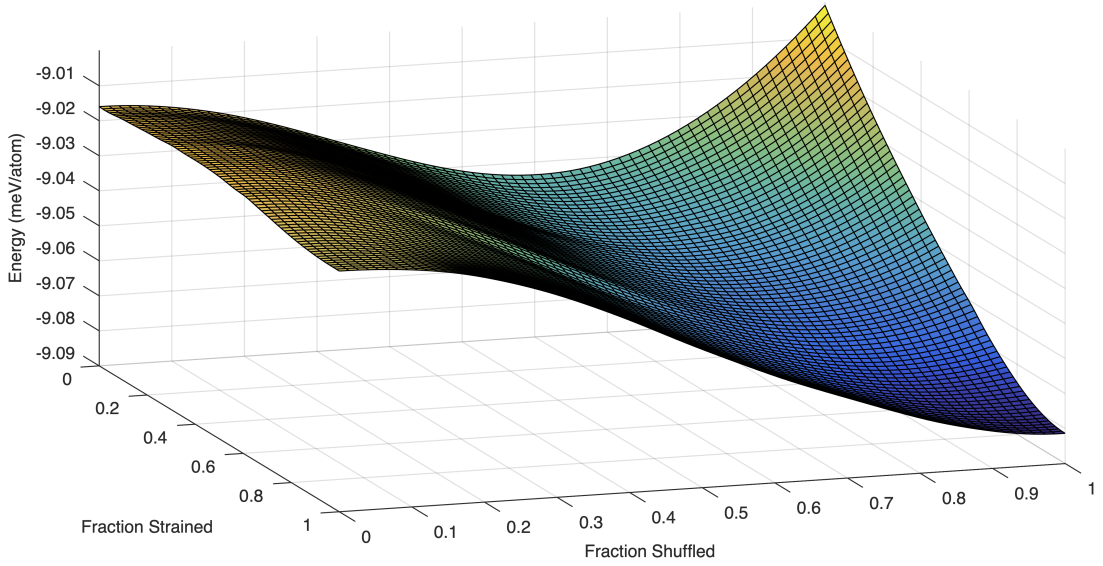

Figure 4.2: The canonical strain-phonon transformation from BCC to HCP is visualized, using the calculated minimum-strain, minimal-shuffling pathway. Along the x-axis is % transformed in terms of strain. Along the y-axis is % transformed in terms of phonon mode. Along the z-axis is the energy. This pathway also shows the dynamical instability of the BCC structure at low temperatures.

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 non-zero activation energy barrier, indicating the metastability of the FCC structure. The energy at this barrier is $-9.015 meV/atom$, corresponding to a barrier height of $34 meV/atom$ 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 introduction:

Start (FCC): 7, 7, 20, 20, 7, 20, 27, 0, 15, 15
$E = -9.029 \frac{meV}{atom}$

20%: 6, 9, 17, 19, 6, 20, 25, 0, 15, 15

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

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

$E = -9.019 \frac{meV}{atom}$

40%: 6, 9, 16, 20, 6, 20, 25, 2, 16, 15

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

$$\vec{\delta} =< 2/30, 1/30, 0 >$$

$E = -9.019 \frac{meV}{atom}$

60%: 6, 9, 16, 22, 6, 22, 25, 5, 17, 15

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

$$\vec{\delta} =< 5/30, 2/30, 0 >$$

$E = -9.019 \frac{meV}{atom}$

80%: 7, 7, 17, 23, 7, 23, 24, 6, 18, 15

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

$$\vec{\delta} =< 6/30, 3/30, 0 >$$

$E = -9.044 \frac{meV}{atom}$

100% (HCP): 7, 7, 17, 24, 7, 24, 24, 10, 20, 15

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

$$\vec{\delta} =< 10/30, 5/30, 0 >$$

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

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 distinct directions. By comparison, it is a more dynamic pathway.

We visualize the local energy environment along the calculated FCC - HCP pathway, in the plot shown below.

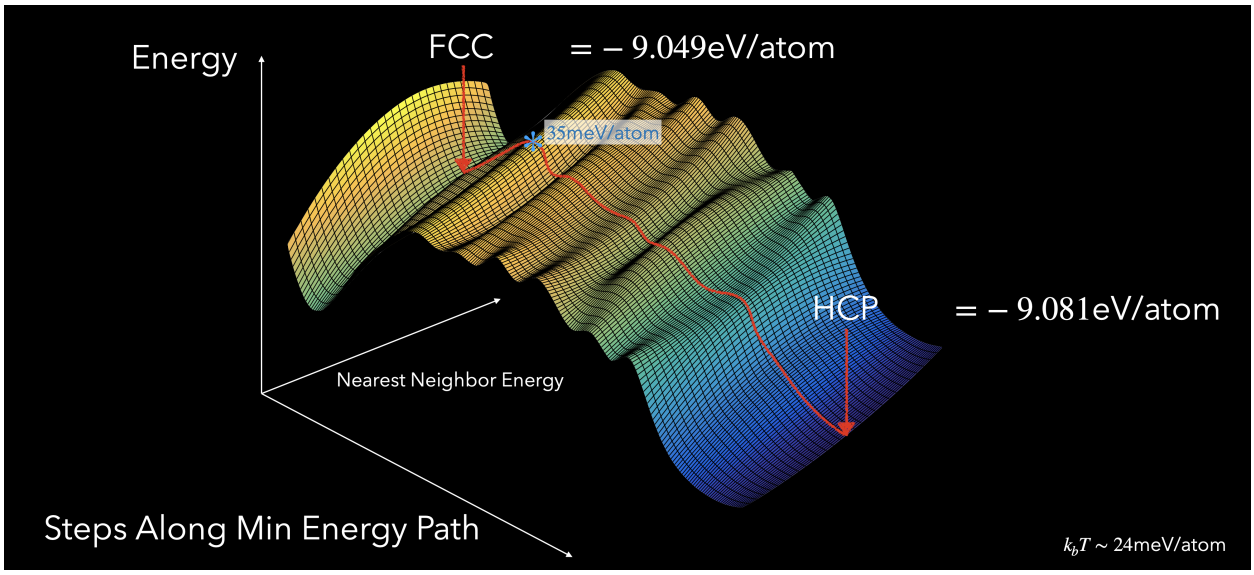

Figure 4.3: The search for transformation pathways from FCC to HCP is visualized with the phase space surrounding a path of minimal activation energy barrier highlighted. Along the x-axis, the number of steps along the transformation pathway are shown. Along the y-axis two of the lowest energy orthogonal nearest neighbors are shown to the left and right. Along the z-axis is the energy. This pathway shows the metastability of the FCC structure at low temperatures.

This new 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 $4meV$ (measured in a 2-atom unit cell), which is insignificant relative to room-temperature thermal vibrations of around $25\frac{meV}{atom}$. 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.

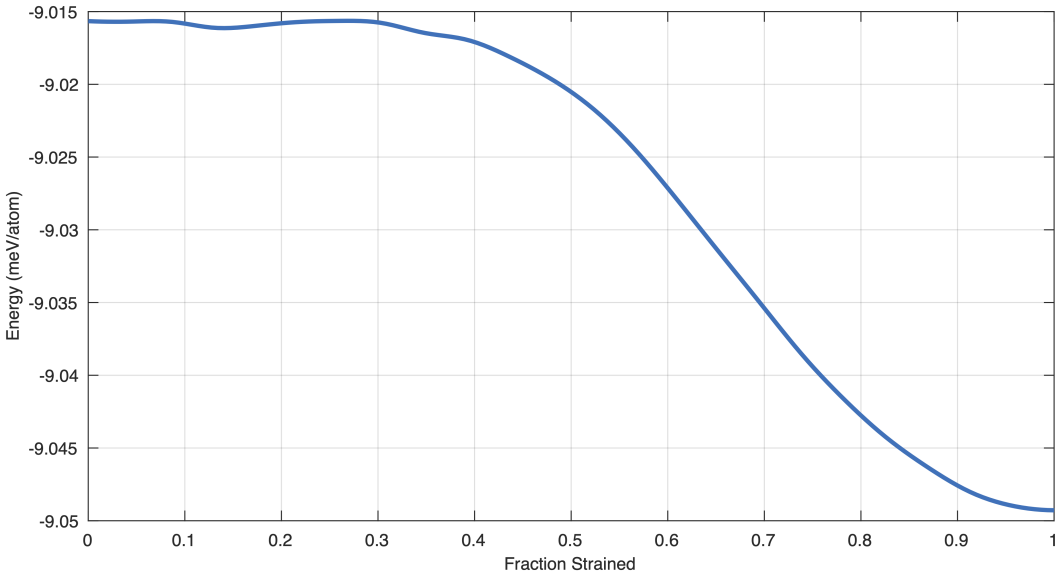

Figure 4.4: The canonical strain transformation from BCC to FCC along the tetragonal strain path is visualized, using the calculated minimum-strain pathway. Along the x-axis is % transformed in terms of strain. Along the y-axis is the energy. This pathway also shows the dynamical instability of the BCC structure at low temperatures.

Taken together, these plots illustrate that this new 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 newly-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.

In conventional approaches to studying martensitic phase transformations, considerable effort goes into managing multiple crystallographic orientations and choices of unit cell. Further, in cases where there are many atoms in a unit cell, it is computationally-expensive to find a minimal-shuffling solution, and there is a separate solution for each unit cell setting. Alternatively, this formalism, focused exclusively on crystal geometry, alleviates this significant overhead and easily generalizes to complicated systems. In terms of finding minimal-strain, minimal-shuffling pathways, a simple breadth-first-search of the discretized parameter space will yield such a path. Further, by employing the described "water-filling" algorithm, we get not only an ensemble of lowest-barrier transformation pathways between two crystal structures, but also an energy landscape: significantly more useful information. We anticipate

that a more thorough understanding of the energy as a function of geometry can yield significant additional insight into finite-temperature properties.

We can now use these activation energy barriers to gain some insight into the phase stability of Zr at atmospheric pressure. At high temperatures the BCC phase, having higher specific entropy, is stabilized relative to the HCP and FCC phases. As the temperature falls, HCP becomes the only observed stable phase (1135K). These calculations support the idea that the BCC phase becomes dynamically unstable with falling temperature. Further, noting that $k_bT \sim 97.8meV/atom$ at this temperature, the crystal has ample thermal energy to overcome the calculated activation energy barrier between FCC and HCP, even cooling significantly from 1135K. Thus, the FCC is never observed as the crystal can easily transform to the more-stable HCP phase.

While this search algorithm considered only the energy of the atomic system, for constant-pressure and/or constant-temperature systems, the corresponding thermodynamic potentials, H (enthalpy), F (Helmholtz free energy), G (Gibbs free energy) can be estimated. Using the principal of minimization of free energy, this same algorithm can be employed, calculating the appropriate free energy to gain insight into relative phase stability at a variety of environmental conditions.

Chapter 5

Outlook

Altogether, this thesis has outlined the salient mathematical framework for understanding crystal geometry and the relationship between different crystal structures in terms of strains and phonon modes. This allows us to define the space of all crystal geometry in new ways and explore it in a comprehensive fashion not previously attempted. When applied to the Zirconium crystal system, this exploration re-derives several known transformation pathways, serving as an initial proof-of-concept.

This framework opens up many exciting new approaches to solving difficult problems in computational materials science. From finding low-energy transformation pathways and gauging relative stability, to relating crystal structures geometrically, filtering out duplicate crystal structures, even gauging finite-temperature properties. As an analogy, this work fundamentally serves as a crystallographic map: a means to uniquely represent any crystal geometry as a string of coordinates. While the complexities involved prevent the representation of this geometry as a vector space, nevertheless, we have been able to create a grid-like structure in which similar crystals are close together and the grid coarseness can be set arbitrarily. As such, with appropriate modification, this tool can be used to provide insight into any desired property as a function of crystal geometry. Simulation tools calculate the property, and this framework keeps track of the crystal geometry.

Going forward, there remain many exciting avenues for further development. Firstly, the phase space grows rapidly more complex as the number of atoms per unit cell is increased. For even 6-atom systems, the aforementioned "water-filling" algorithm becomes impractical as the configuration space is simply too large to exhaustively explore. Even for the two-atom Zr calculations, slightly over 50,000 static calculations were required when using a fine structural search grid coarseness. There certainly is a trade-off between rigorously finding the lowest activation energy pathway vs. a more-targeted heuristic search algorithm. For many systems, more-targeted approaches will be necessary.

I believe there is ample room for new and efficient transformation pathway search algo-

rithms, either engineered or machine-learned, built off of this new formalism. While a given energy calculation takes on the order of minutes for small systems, geometric calculations and comparisons can be calculated on the order of milliseconds. As such, this formalism can enable a wide range of transformation-finding algorithms based on arbitrary geometric constraints, with reasonable performance for larger systems. Notably, this includes the canonical paradigm of minimal-atomic-displacement, represented here as minimum number of hops in the crystallographic parameter space. Further, the ability to set the coarseness of the discretization parameters, allows one to quickly compute a pathway at a coarse setting and then refine the pathway as needed by progressively increasing the density of the grid.

As a final note, I hope that this tool can serve the larger materials science community, both in finding new crystal structures and geometrically connecting them together. Structure-property relationships form the main foundation of the field. I am hopeful these new techniques, combined with ever-growing ab-initio toolkits, will help to accelerate discovery and, ultimately, positively impact the world.

Bibliography

[1] Barbara Ann Kipfer. *Encyclopedic dictionary of archaeology*. Springer, 2021.

[2] Kim Ann Zimmermann. "History of computers: A brief timeline". In: *Live science* (2017).

[3] Marc Tyler Nobleman. *The Light Bulb*. Capstone, 2003.

[4] Anubhav Jain et al. "Commentary: The Materials Project: A materials genome approach to accelerating materials innovation". In: *APL materials* 1.1 (2013), p. 011002.

[5] Lev Davidovich Landau and Evgenii Mikhailovich Lifshitz. *Statistical Physics: Volume 5*. Vol. 5. Elsevier, 2013.

[6] Edgar C Bain and NY Dunkirk. "The nature of martensite". In: *trans. AIME* 70.1 (1924), pp. 25–47.

[7] Harold T Stokes and Dorian M Hatch. "Procedure for obtaining microscopic mechanisms of reconstructive phase transitions in crystalline solids". In: *Physical Review B* 65.14 (2002), p. 144114.

[8] Daniel Sheppard et al. "A generalized solid-state nudged elastic band method". In: *The Journal of chemical physics* 136.7 (2012), p. 074103.

[9] Konstantinos Koumatos and Anton Muehlemann. "Optimality of general lattice transformations with applications to the Bain strain in steel". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 472.2188 (2016), p. 20150865.

[10] SK Sikka, YK Vohra, and R Chidambaram. "Omega phase in materials". In: *Progress in Materials Science* 27.3-4 (1982), pp. 245–310.

[11] David Steven Dummit and Richard M Foote. *Abstract algebra*. Vol. 3. Wiley Hoboken, 2004.

[12] John Horton Conway and Neil JA Sloane. "Low-dimensional lattices. VI. Voronoi reduction of three-dimensional lattices". In: *Proceedings of the Royal Society of London. Series A: Mathematical and Physical Sciences* 436.1896 (1992), pp. 55–68.

[13] John William Scott Cassels. *An introduction to the geometry of numbers*. Springer Science & Business Media, 2012.

[14] Marvin L Cohen and Steven G Louie. *Fundamentals of condensed matter physics*. Cambridge University Press, 2016.

[15] CS Barrett, P Cucka, and KJAC Haefner. "The crystal structure of antimony at 4.2, 78 and 298 K". In: *Acta Crystallographica* 16.6 (1963), pp. 451–453.

[16] James Demmel and Yozo Hida. "Accurate and Efficient Floating Point Summation". In: *SIAM Journal on Scientific Computing* 25.4 (2004), pp. 1214–1248. DOI: 10.1137/S1064827502407627. eprint: https://doi.org/10.1137/S1064827502407627.

[17] T Yamanaka, R Kurashima, and J Mimaki. "X-ray diffraction study of bond character of rutile-type SiO2, GeO2 and SnO2". In: *Zeitschrift für Kristallographie-Crystalline Materials* 215.7 (2000), pp. 424–428.

[18] Mathias Ekman et al. "Ab initio study of the martensitic bcc-hcp transformation in iron". In: *Physical Review B* 58.9 (1998), p. 5296.