Optimization of strain-induced superconductivity in RuO_2 using physics-informed surrogate models

by
Rauan Kaldybayev

Professor Betül Pamuk, Advisor

A thesis submitted in partial fulfillment
of the requirements for the
Degree of Bachelor of Arts with Honors
in Physics

WILLIAMS COLLEGE
Williamstown, Massachusetts
May 26, 2025

Abstract

It was recently discovered that the non-superconducting compound RuO_2 becomes superconducting under the influence of strain. With the help of Density Functional Theory numerical calculations, the onset of superconductivity was attributed to the increase of the electronic density of states at the Fermi level [1]. Searching for other strains that produce superconductivity followed soon after [2]. The search is complicated by the computational cost of DFT simulations, which require supercomputer time. In our work, we introduce physics-informed surrogate models for various electronic properties of RuO_2, which allows us to bypass expensive DFT computations while keeping the accuracy within acceptable levels. The application of these surrogate models makes feasible a systematic and expansive search for strains that produce the greatest density of states at the Fermi level. We perform this expansive search and identify strains that could potentially lead to superconductivity. Biaxial tensile strain in the (001) crystal orientation emerges as one of the best possible strains for RuO_2 and it is, at the same time, experimentally accessible with the help of biaxial epitaxy.

Executive Summary

Desirable properties such as superconductivity can be induced in materials using biaxial strain [1], and strain-induced change in material properties has garnered considerable scientific [3, 4, 5, 6] and engineering [7, 8, 9, 10, 11, 12, 13, 14] interest. Physical experiment remains as the ultimate judge of the usefulness of a particular strain, but Density Functional Theory (DFT) has simplified the search for useful strains by allowing one to computationally estimate the effects of a given strain without having to physically realize it in experiment. However, even DFT is not fast enough for an exhaustive search of the space of all possible strains, since the current approach to DFT requires human attention and supercomputer time for each individual strain.

We propose a methodology where reasonably accurate yet very fast surrogate models are used to perform a systematic and exhaustive search of the space of possible strains. After promising strains have been identified by this method, they can be studied in detail using individual DFT analysis and, if their useful properties are confirmed, proposed to experimentalists for physical realization. (Diagram 1.)

$$\begin{array}{c}\text{Model based}\\ \text{exhaustive search}\end{array} \xrightarrow{\text{Suggests}} \begin{array}{c}\text{DFT based}\\ \text{examination}\end{array} \xrightarrow{\text{Suggests}} \text{Experiment} \tag{1}$$

Our guiding principle is that oftentimes, the response of a "nice" physical quantity to small perturbations is well-approximated by the first few terms of its Taylor series. Strain is a small quantity. Our "nice" physical quantities are the elastic energy density and the electronic bands energies of RuO_2. We expand them as a Taylor series in the strain, up to some finite order. Thanks to rutile symmetry, we are able to restrict the number of independent coefficients in the power series for the elastic energy. The models are linear in their parameters, so upon obtaining training data from DFT, we simply fit the model coefficients using least squares. The phenomena we are ultimately interested in, namely the relaxation of RuO_2 under biaxial strain and the response of the Fermi level density of states (DOS) to strain, appear to depend on the strain in a complicated way. We are able to accurately predict these phenomena by combining our models for the elastic energy and electronic bands of RuO_2 with principles from solid state physics.

Using the model, we survey the space of all strains and find a strain configuration that is predicted by DFT to increase the Fermi level density of states by 77% (Figure 2). For comparison, the strain in Ruf et. al. [1] that led to superconductivity increased the Fermi level DOS by roughly 30%. The model-optimized strain (equation 5.3) is illustrated in Figure

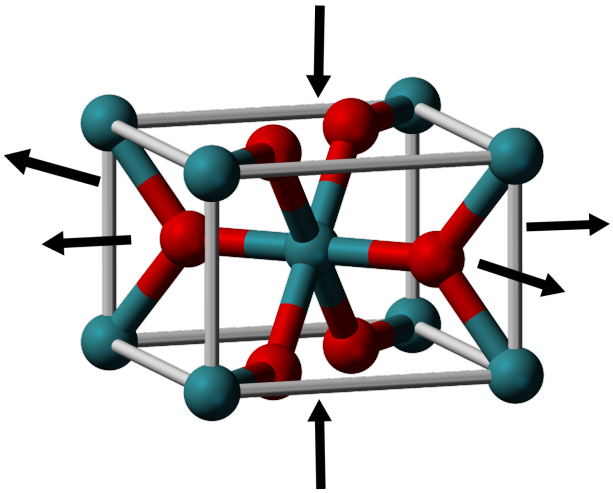


Figure 1: A pictorial illustration of the strain that is predicted to induce the greatest Fermi level DOS increase in RuO_2. The xy plane is uniformly expanded, while the z direction is contracted.

1; it corresponds to expanding the xy plane and contracting the z direction. Its eigenvalues and eigenvectors are

$$\begin{aligned} \lambda_1 &= -4.78\%, \quad & \mathbf{v}_1 &= (-0.09, -0.09, 0.99), \\ \lambda_2 &= 4.57\%, & \mathbf{v}_2 &= (0.93, -0.36, 0.05), \\ \lambda_3 &= 4.84\%, & \mathbf{v}_3 &= (-0.35, -0.93, -0.12). \end{aligned}$$

Luckily, the optimized strain configuration is very close to biaxial tensile strain in the xy plane with the z direction left free. We use DFT to compute the Fermi level DOS of RuO_2 under various values of xy biaxial strain. The Fermi level density of states is predicted to increase with xy strain, with a peak at around 4.5% tensile strain (Figure 3). At the 4.5% peak, DFT predicts a 77%[1] increase in Fermi level DOS. Preliminary DFT calculations also indicate significant softening of phonon modes, although careful DFT analysis must be conducted to corroborate this claim because phonon calculations are difficult in terms of convergence. For comparison, the strain in Ruf et. al. [1] that led to superconductivity is thought to increase the Fermi level DOS by roughly 30%, also with a softening of phonon modes. In the electron-phonon picture of superconductivity [15] [16], a high electron density at the Fermi level and a softening of phonon modes are hallmarks of superconductivity. Therefore, we expect that superconductivity is likely to occur under 4.5% tensile strain in the xy plane. Biaxial strain in the xy plane can be experimentally realized using biaxial epitaxy, for example using the Molecular Beam Epitaxy (MBE) [17] [18] technique.

[1]The 77% Fermi level DOS increase happens to coincide with the 77% increase from the optimized strain, although the two strains are slightly different.

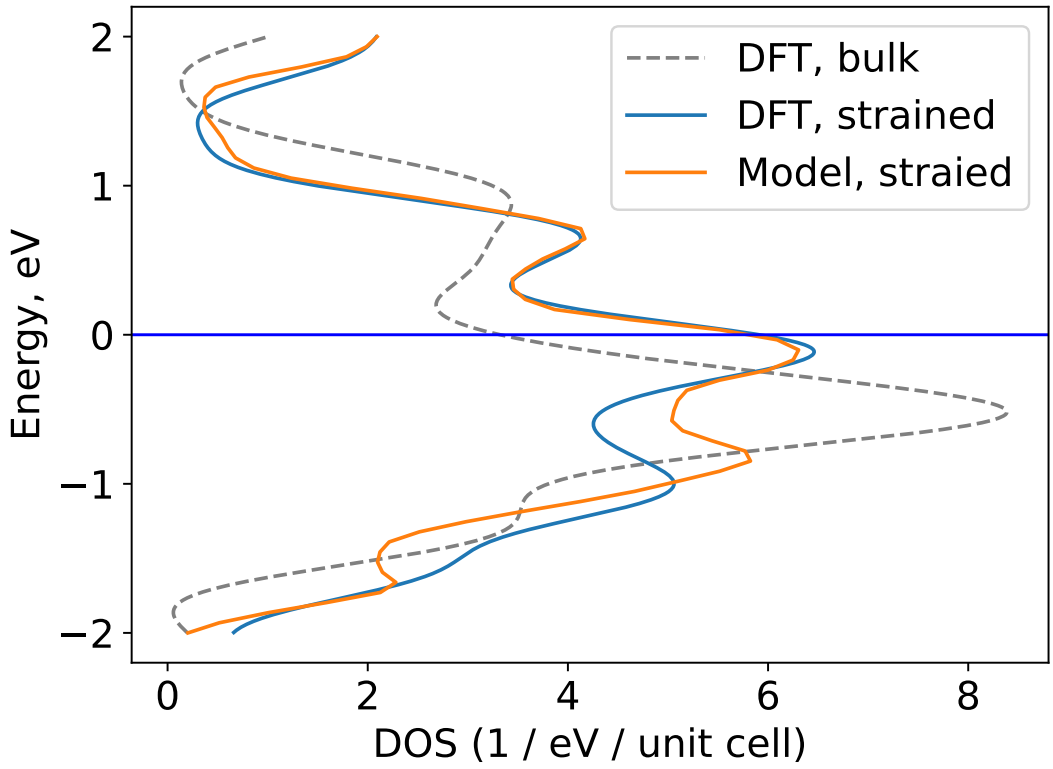


Figure 2: Strain brings the peak in the electronic density of states (DOS) to the Fermi level, which is set to 0 eV. The strain was found using an approximate linear model. The model's prediction is supported by a direct DFT computation.

Experimental realization of tensile biaxial stain in the xy plane stands as a natural direction of future work, since DFT appears to predict a high likelihood of strain-induced superconductivity under this strain.

In summary, we have used DFT data to produce very fast yet reasonably accurate surrogate models. Using these models, we surveyed the space of all possible strains and identified biaxial tensile strain in the xy plane (the (001) crystal plane) as an interesting strain direction. We expect that it increases the electronic density of states at the Fermi level and potentially leads to a considerable softening of phonon modes, which in the electron-phonon picture of superconductivity makes it a likely candidate for strain-induced superconductivity. Biaxial strain in the xy plane can be experimentally realized using Molecular Beam Epitaxy (MBE) in the (001) crystal orientation.

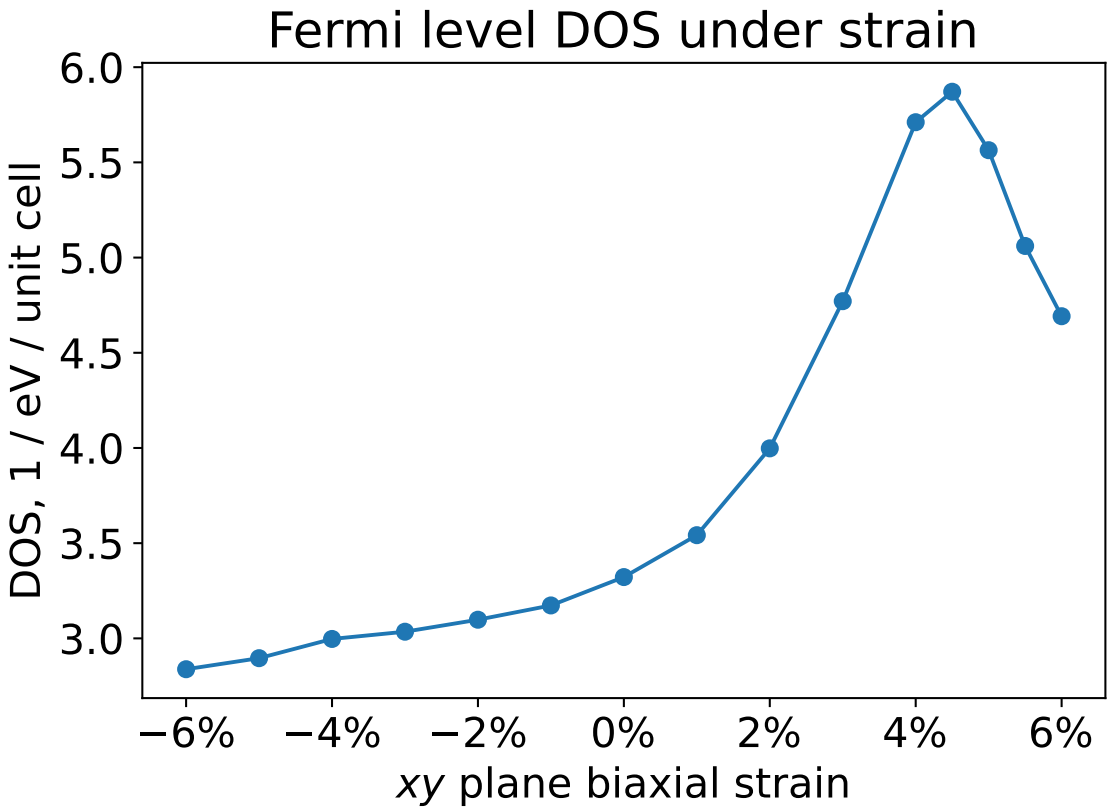


Figure 3: Fermi level density of states under biaxial strain in the xy plane (the (001) direction), predicted using DFT. Under 4.5% tensile strain, the Fermi level DOS is predicted to change from 3.322 to 5.871 states per eV per unit cell, making an increase of 77%.

Acknowledgments

The greatest acknowledgment goes to my advisor, Professor Betul Pamuk, for her kind and constructive support of my explorations in this thesis, as well as to my second reader, Professor David R. Tucker-Smith, for his valuable feedback. I am grateful to Williams College Libraries for providing me with laptop loans, which allowed me to continue working while my personal computer was experiencing technical issues. I thank my advisor and Johns Hopkins University for providing me access to the Rockfish supercomputing cluster, and I thank Williams College for giving me access to their HPCC cluster.

Contents

List of Figures

Chapter 1

Introduction

It was recently discovered [1] [19] that while ruthenium oxide RuO_2 is not superconducting under normal conditions[1], it becomes superconducting when subject to strain. Specifically, when a thin film of RuO_2 is grown on a crystal of rutile TiO_2 in a certain way, the RuO_2 thin film adapts to the crystal lattice parameters of TiO_2 and becomes superconducting. The strain is only about 5% in magnitude, yet it brings about superconductivity. Using first-principles Density Functional Theory (DFT) numerical calculations, Ruf et. al. [1] predicted that the Fermi level electronic density of states (DOS) of RuO_2 is significantly altered by strain, and that softened phonon modes appear in response to strain. In the electron-phonon picture of superconductivity [15] [16], a high Fermi level DOS and the softening of phonon modes are hallmarks of superconductivity.

Strain-induced change in material properties has garnered considerable scientific [3, 4, 5, 6] and engineering [7, 8, 9, 10, 11, 12, 13, 14] interest. Strain allows one to tune materials to exhibit curious behaviors or desirable properties, superconductivity being one of them. Ongoing research presented at a conference meeting [2] illustrates how DFT can guide experiments. They theoretically predicted strain-induced superconductivity in RuO_2 epitaxial thin films on TiO_2 in the (100) orientation, which led them to produce the physical setup in the laboratory. They then performed electronic transport measurements of the RuO_2 thin films and successfully observed superconductivity. Instead of searching through possible strain configurations by realizing and measuring them physically, which is hard, they did the searching using DFT simulations.

But the current DFT approach [1, 14, 5] only allows for a few strains to be examined at a time, with human attention and supercomputer time. How can we search the (rather large) space of all possible strains? We propose a wholesale pre-DFT search based on approximate models. Firstly, we perform a systematic and exhaustive search of the space of all possible strains using very fast yet reasonably accurate approximate models that we develop. This allows us to identify a few interesting strains. Secondly, for strains of interest, one can perform more in-depth DFT analysis. Thirdly, it is possible to propose experimental scenarios

[1] Or, the superconducting transition temperature is below the available measurement threshold.

that might be of interest for realization in the laboratory. This is illustrated in diagram 1.1.

$$\begin{array}{c}\text{Model based}\\ \text{exhaustive search}\end{array} \xRightarrow{\text{Guides}} \begin{array}{c}\text{DFT based}\\ \text{examination}\end{array} \xRightarrow{\text{Guides}} \text{Experiment} \tag{1.1}$$

This is enabled by surrogate models. In engineering and machine learning, a *surrogate model* [20, 21] is a simple mathematical model that one trains to approximate an expensive-to-compute target function. Our target function is the electronic structure of RuO_2, and our input variable is strain. We perform Density Functional Theory [22, 23, 24, 25] numerical calculations to obtain learning and training data for our models. The models we use are linear in their parameters, so the fitting of their coefficients is simply done using least squares.

Our guiding principle is that oftentimes, the response of a "nice" physical quantity to small perturbations is well-approximated by the first few terms of its Taylor series. Strain is a small quantity. Our "nice" physical quantities are the elastic energy density and the electronic bands energies of RuO_2. We construct the surrogate models by replacing the complicated target function with a finite power series. Thanks to rutile symmetry, we are able to restrict the number of independent coefficients in the power series for the elastic energy. For the electronic bands, we find that a first-order power series already suffices. More complicated processes include the relaxation of RuO_2 under biaxial strain, the evolution of the Fermi level in response to strain, and the evolution of the electronic density of states (DOS) in response to strain. We successfully predict these nonlinear phenomena by combining our models for the elastic energy and electronic bands of RuO_2 with principles from solid state physics.

Using the model, we survey the space of all strains and predict a strain configuration that increases the Fermi level density of states by as much as 78%. The optimized lattice parameters are given in equation 1.2. We observe considerable softening of phonon modes. For comparison, the strain in Ruf et. al. [1] that led to superconductivity was predicted, by DFT, to increase the Fermi level DOS by roughly 30%, also with what appears to be a softening of phonon modes. In the electron-phonon picture of superconductivity [15] [16], a high electron density at the Fermi level and the softening of phonon modes are hallmarks of superconductivity. The lattice parameters corresponding to the strain we propose are, in units of Bohr,

$$\mathbf{a}_1 = \begin{pmatrix} 8.485 \\ -0.183 \\ 0.151 \end{pmatrix}, \quad \mathbf{a}_2 = \begin{pmatrix} -0.183 \\ 8.808 \\ 0.201 \end{pmatrix}, \quad \mathbf{a}_3 = \begin{pmatrix} 0.105 \\ 0.139 \\ 5.646 \end{pmatrix} \tag{1.2}$$

Luckily, the optimized strain configuration is very close to biaxial dilation in the xy plane with the z direction left free. This strain can be realized in experiment via biaxial epitaxy in the (001) crystal orientation, using a method such as Molecular Beam Epitaxy [17] [18]. Using direct DFT calculations, we verify the model's prediction of the increased Fermi level density of states. The Fermi level density of states increases in a nonlinear way with xy strain, increasing slowly at first but then quickly accelerating pace. DFT calculations predict that at 4.5% strain in the xy plane, the Fermi level density of states increases by

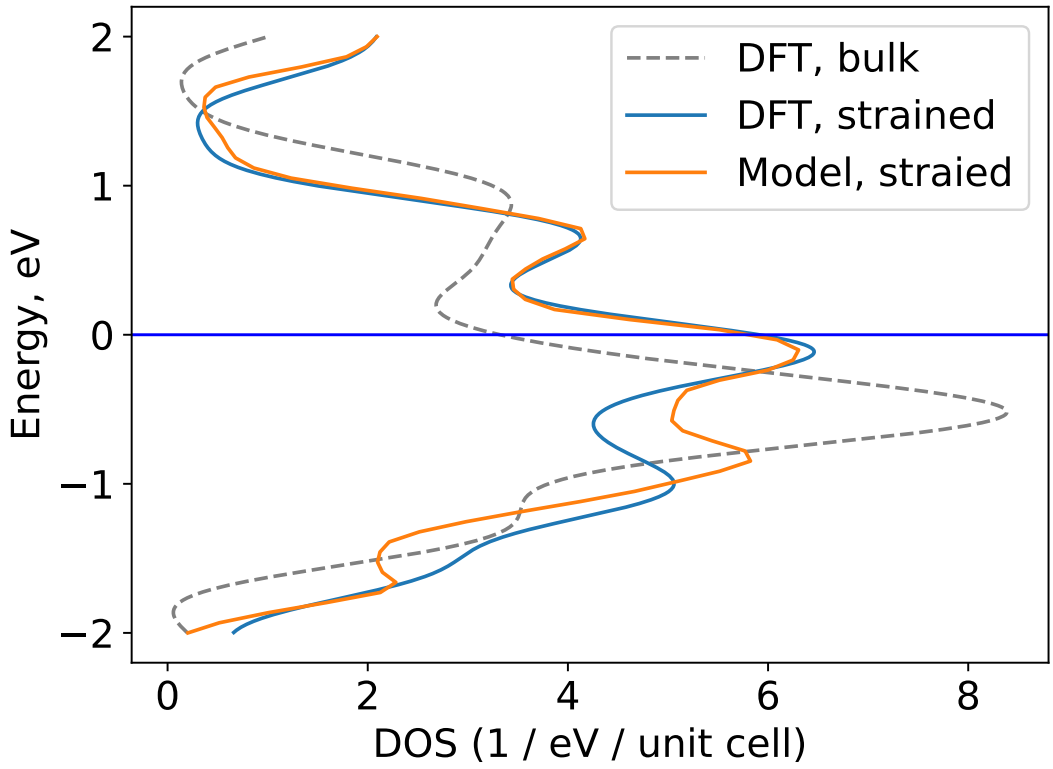


Figure 1.1: Strain brings the peak in the electronic density of states (DOS) to the Fermi level. The strain was found using an approximate linear model. The model's prediction is supported by a direct DFT computation.

as much as 77%, and significant softening of phonon modes occurs. Biaxial strain in the xy plane can be experimentally realized using biaxial epitaxy, for example using the Molecular Beam Epitaxy (MBE) [17] [18] technique.

Experimental implementation of biaxial stain in the xy plane stands as a natural direction of future work, since DFT seems to predict a high likelihood of strain-induced superconductivity in that case.

Chapter 2 lays the foundation for subsequent chapters by discussing some basic geometry of crystal lattices. In Chapter 3, we Taylor-expand the elastic energy density of RuO_2 with respect to strain, considering quadratic and higher terms and using rutile symmetry to reduce the number of independent expansion coefficients. The high-order expansion is necessary because the first-order terms are zero, since the bulk (relaxed) state is a local energy minimum. In Chapter 4, we Taylor-expand the electronic bands energies of RuO_2, terminating the expansion at the first order and thus approximating the electronic bands' response to strain as linear. The coefficients of the Taylor expansions are fitted as model parameters based on an initial set of data points obtained using Density Functional Theory

(DFT) first-principles numerical calculations. Once the initial DFT calculations are performed and the models are fitted, no DFT calculation is necessary to make (approximate) calculations concerning the elastic or electronic properties of RuO_2. Chapter 5 discusses the physical predictions made using the methods developed in the previous chapters.

1.1 Notation

- $\mathbb{R}$ – the field of real numbers.
- $\mathbb{Z}$ – the ring of integers.
- $\mathbb{R}^n$, $\mathbb{Z}^n$ – set of tuples of n real numbers (integers). Identified with n-dimensional Euclidean space and an n-dimensional lattice, respectively. In particular, $\mathbb{R}^2$ is the Euclidean plane and $\mathbb{R}^3$ is the Euclidean space.
- $\mathbf{x} = (x_1, \ldots, x_n) = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ – column vector $\mathbf{x} \in \mathbb{R}^n$.
- $\begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix}$ – row vector. (Note the absence of commas between x_i.)
- $\mathbb{R}^{n\times m}$ – the ring of $n \times m$ matrices. If $A \in \mathbb{R}^{n\times m}$ is such a matrix, the transpose $A^T \in \mathbb{R}^{m\times n}$ is the $m \times n$ matrix. The matrices have the form
$$A = \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nm} \end{pmatrix}, \quad A^T = \begin{pmatrix} a_{11} & \ldots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1m} & \ldots & a_{nm} \end{pmatrix}.$$
- A vector $\mathbf{x} \in \mathbb{R}^n$ can be thought of as an $n \times 1$ matrix, $\mathbf{x} \in \mathbb{R}^{n\times 1}$. The transpose $\mathbf{x}^T \in \mathbb{R}^{1\times n}$ is the row vector $\begin{pmatrix} x_1 & \ldots & x_n \end{pmatrix}$.
- When $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ are vectors, we identify their inner product $x_1y_1 + \cdots + x_ny_n$ with the 1×1 matrix $\mathbf{x}^T\mathbf{y}$.
- When $\mathbf{x}^{(1)}, \ldots, \mathbf{x}^{(m)} \in \mathbb{R}^n$ are n-dimensional vectors, with $\mathbf{x}^{(i)} = (x_1^{(i)}, \ldots, x_n^{(i)})$ the matrix
$$A = \begin{pmatrix} \mathbf{x}^{(1)} & \ldots & \mathbf{x}^{(m)} \end{pmatrix} = \begin{pmatrix} x_1^{(1)} & \ldots & x_1^{(m)} \\ \vdots & \ddots & \vdots \\ x_n^{(1)} & \ldots & x_n^{(m)} \end{pmatrix}$$
is defined by taking $\mathbf{x}^{(i)}$ to be the columns of A. Likewise, we can create a matrix whose rows are the transposes of $\mathbf{x}^{(i)}$, $A^T = \begin{pmatrix} (\mathbf{x}^{(1)})^T \\ \vdots \\ (\mathbf{x}^{(m)})^T \end{pmatrix}$.

Chapter 2

Geometry of crystals and their slices

The present chapter is a prelude that sets up language for the future chapters. Sections 2.1 through 2.4 rephrase some familiar notions in a convenient language, while Sections 2.5 and 2.6 develop an uncomplicated but possibly new framework for quantifying biaxial strain.

Section 2.1 introduces *Bravais matrices* as a formulation of the usual notion of crystal lattice basis vectors. Sections 2.2, 2.3, and 2.4 recall Miller indexes from solid state physics and deformation gradients from elasticity theory. A good introduction to crystals is given in Ashcroft and Mermin's classic textbook [26], although the reader need only know the notion of a Bravais lattice, a notion that we briefly explain. And while we talk about the deformation gradient from elasticity theory [27], we are only considering strain that is uniform across all of space, and as a result, we do not get into the full complexities of elasticity theory. Indeed, this chapter is quite simple.

After recalling basic notions, we formulate in Sections 2.5 and 2.6 a systematic description of biaxial strain, specifically the restrictions it places and the freedom it leaves. Biaxial strain is when a crystal is forced to conform to some lattice parameters within some plane but free to relax in the out-of-plane direction. This situation is of practical interest because it can readily be implemented in experiments using biaxial epitaxy. We will see that a crystal lattice has 9 degrees of freedom and that biaxial strain fixes 6, so that 3 are remaining. We parametrize these 3 degrees of freedom. Our matrix based description seems perhaps more versatile, albeit more cumbersome, than the eigenvector-eigenvalue based description one often finds in materials science literature [1]. In Chapter 3, we combine this framework with elastic data about RuO_2 to predict the relaxation of RuO_2 under a given biaxial strain.

2.1 Bravais matrices

Recall 2.1.1 (Bravais lattice). Polonium is a metal that crystallizes in the *simple cubic structure*: the atoms arrange themselves in a perfectly regular rectangular array. In other words, there exist *lattice vectors* $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^3$, orthogonal and all having the same length, such that if one polonium atom is located at position $\mathbf{r}$, then the positions of all other polonium atoms are given by $\mathbf{r} + i\,\mathbf{a}_1 + j\,\mathbf{a}_2 + k\,\mathbf{a}_3$ for integer i, j, k. We call the set of all translations

$i\,\mathbf{a}_1 + j\,\mathbf{a}_2 + k\,\mathbf{a}_3$ with respect to which the crystal is symmetric its *Bravais lattice*. A surprising result from solid state physics is that *every* crystal has a Bravais lattice [26]. Roughly speaking, a crystal is a material that is periodic in space in a discrete way. We can then find lattice vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^3$ such that the crystal is obtained by putting together the copies, translated in space by $i\,\mathbf{a}_1 + j\,\mathbf{a}_2 + k\,\mathbf{a}_3$ for i, j, k integers, of some finite arrangement of atoms. While this description seems to single out crystals with "parallelogram-esque" symmetry, it turns out to apply even to those crystals we would intuitively consider e.g., "triangular" or "hexagonal" in nature.

Consider a crystal with lattice vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^3$. It is often convenient to express positions in reference to the crystal lattice. A relative crystal position $\mathbf{u} = (u_1, u_2, u_3)$ corresponds to the physical position

$$\mathbf{r} = (x, y, z) = u_1\,\mathbf{a}_1 + u_2\,\mathbf{a}_2 + u_3\,\mathbf{a}_3.$$

This can be written compactly as $\mathbf{r} = B\,\mathbf{u}$, where we define the *Bravais matrix* B as the matrix whose columns are $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$,

$$B = \begin{pmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{pmatrix} = \begin{pmatrix} a_1^x & a_2^x & a_3^x \\ a_1^y & a_2^y & a_3^y \\ a_1^z & a_2^z & a_3^z \end{pmatrix}. \tag{2.1}$$

For example, Bravais lattice points are exactly the integer combinations of $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$, and therefore, they are of the form $B\,\mathbf{n}$ for triplets of integers $\mathbf{n} \in \mathbb{Z}^3$. We think of the Bravais matrix as converting from relative crystal positions to physical positions in space.

Just as the lattice basis vectors are not unique, so is the Bravais matrix. Whenever A is an invertible 3×3 matrix with integer coefficients, the Bravais matrix BA defines the same lattice as B. For example, vectors $\mathbf{a}_1 + \mathbf{a}_2, \mathbf{a}_2, \mathbf{a}_3$ define the same lattice as $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$. Another way to phrase the equivalence of $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ and $\mathbf{a}_1 + \mathbf{a}_2, \mathbf{a}_2, \mathbf{a}_3$ is that Bravais matrix BA defines the same lattice as B when

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

There is also rotation symmetry: whenever U is a rotation matrix, the lattice defined by UB is just a rotation of the lattice defined by B. When discussing a crystal, we will implicitly fix a Bravais matrix for it. In particular, "the" Bravais matrix of RuO_2 is [28]

$$B = \begin{pmatrix} 4.492 & 0 & 0 \\ 0 & 4.492 & 0 \\ 0 & 0 & 3.106 \end{pmatrix}, \tag{2.2}$$

where the lattice parameters are expressed in units of angstrom. These values are obtained from Density Functional Theory (DFT) [1].

2.2 Deformation gradient in terms of Bravais matrices

A 3×3 matrix with real coefficients $F \in \mathbb{R}^{3\times 3}$ encodes a linear transformation, namely the transformation

$$\mathbf{r} \;\mapsto\; F\,\mathbf{r}. \tag{2.3}$$

Suppose we take a crystal and move every atom from its initial position $\mathbf{r}$ to a new position $F\,\mathbf{r}$. That is, we apply the transformation defined by F to the crystal. When $F^T F = 1$ is the identity matrix, there are two options: (i) $\det F = 1$ and the transformation is a rotation, (ii) $\det F = -1$ and the transformation is a reflection. In all other cases, F does not preserve distances, and the crystal is strained. While the strain tensor in a medium generally depends on position [29] [30] [27], we are looking at uniform (position-independent) strain. For example, a 5% dilation in the z direction and a 5% shear in the xy plane correspond, respectively, to the matrices

$$F_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1.05 \end{pmatrix}, \quad F_2 = \begin{pmatrix} 1 & 0.05 & 0 \\ 0.05 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Consider a crystal with Bravais matrix B. A point with relative crystal position $\mathbf{u} \in \mathbb{R}^3$ is located at $B\,\mathbf{u}$. Suppose the crystal is now transformed by F, so that the crystal lattice changes. The point with relative position $\mathbf{u}$ moves together with the lattice to $FB\,\mathbf{u}$. Since the relative position and the physical position are related by FB, we see that the Bravais matrix of the transformed crystal is

$$B' = FB. \tag{2.4}$$

Here, again, B is the un-strained (bulk) Bravais matrix and B' is the strained Bravais matrix. In other words, the lattice vectors $\mathbf{a}'_1, \mathbf{a}'_2, \mathbf{a}'_3$ of the transformed crystal are given by $\mathbf{a}'_i = F\,\mathbf{a}_i$. In the context of elasticity theory, F is called the *deformation gradient* [29]. We can write F in terms of Bravais matrices as

$$F = B'\,B^{-1}. \tag{2.5}$$

(We use that the lattice vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ of a crystal do not lie in the same plane, and that therefore the Bravais matrix is always invertible.)

2.3 Polar decomposition

Via the polar decomposition [29] (Appendix A), any deformation gradient F can be factored as

$$F = U\,D, \tag{2.6}$$

where U is a 3×3 orthogonal matrix and D is a 3×3 positive semidefinite matrix. Physically achievable transformation matrices F must be not too far from being a rotation matrix, since substantial compression, dilation, or shear destroys a crystal. In particular, we postulate that $\det F > 0$ because having $\det F = 0$ would squish the crystal to a plane, while $\det F < 0$

corresponds to bending the crystal so hard that it changes handedness. This implies that U is a rotation matrix and D is positive definite.

We call D the *deformation matrix* and U the *post-rotation*. Transforming the crystal by F is the same as first acting on it by D and then rotating it by U (Figure 2.1). Since rotation leaves all physical properties the same, it is really the deformation matrix D that we are interested in. We define the *strain matrix* E as the difference between D and the 3×3 identity matrix,

$$E = D - 1. \tag{2.7}$$

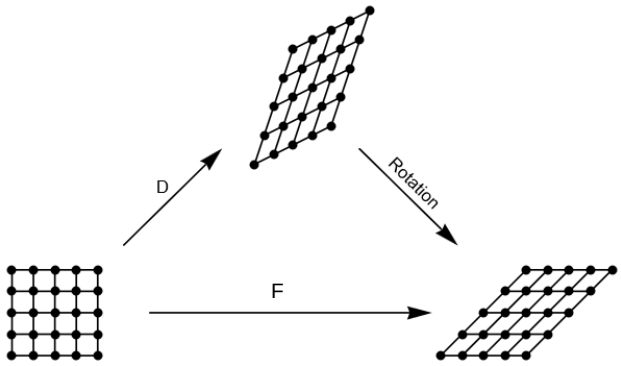


Figure 2.1: A transformation F of the lattice can be decomposed into a positive definite matrix D followed by a rotation U.

While F, an arbitrary 3×3 matrix, has nine degrees of freedom, D and E only have six because they are symmetric. Recall that a rotation is described by three angles. We have removed three inconsequential degrees of freedom by considering F up to a rotation.

2.4 Slices and Miller indexes

Definition 2.4.1 (Integer slice)**.** For a triplet of integers $\mathbf{n} = (n_1, n_2, n_3) \in \mathbb{Z}^3$ where at least one of n_1, n_2, n_3 is nonzero, we define the *integer slice* $\mathrm{Slc}(\mathbf{n})$ as the plane in $\mathbb{Z}^3$ orthogonal to $\mathbf{n}$,

$$\mathrm{Slc}(\mathbf{n}) = \{\mathbf{s} = (s_1, s_2, s_3) \in \mathbb{Z}^3 : \mathbf{n} \cdot \mathbf{s} = n_1\, s_1 + n_2\, s_2 + n_3\, s_3 = 0\}. \tag{2.8}$$

Remark 2.4.2. In mathematics, $\mathrm{Slc}(\mathbf{n})$ would be called a $\mathbb{Z}$-module.

Definition 2.4.3 (Integer basis)**.** If $\mathbf{p}, \mathbf{q} \in \mathrm{Slc}(\mathbf{n})$ are triplets of integers such that every $\mathbf{s} \in \mathrm{Slc}(\mathbf{n})$ can be expressed as a linear combination $\mathbf{s} = a\,\mathbf{p} + b\,\mathbf{q}$ with integer coefficients $a, b \in \mathbb{Z}$, the tuples $\mathbf{p}, \mathbf{q}$ are said to form an *integer basis* for $\mathrm{Slc}(\mathbf{n})$.

Consider a crystal with Bravais matrix B. A *crystal plane* with *Miller indexes* $\mathbf{n} = (n_1, n_2, n_3) \in \mathbb{Z}^3$ is a plane within the Bravais lattice consisting of points $\mathbf{r} = B\,\mathbf{s}$, where $\mathbf{s} \in \mathrm{Slc}(\mathbf{n})$ is a triplet of integers orthogonal to $\mathbf{n}$,

$$B\,\mathrm{Slc}(\mathbf{n}) = \{B\,\mathbf{s} : \mathbf{s} \in Z^3, \mathbf{n} \cdot \mathbf{s} = 0\}.$$

Planes coming from a set of Miller indexes are called *crystal planes*; they are important and natural in the study of crystals [26].

We note that a crystal plane with Miller indexes $\mathbf{n}$ is a two-dimensional Bravais lattice with primitive lattice vectors $B\,\mathbf{p}$, $B\,\mathbf{q}$, where $\mathbf{p}, \mathbf{q}$ are an integer basis for $\mathrm{Slc}(\mathbf{n})$.

2.5 Constraints imposed by biaxial strain

Recall that a transformation during which a crystal is deformed and rotated is represented by a matrix $F \in \mathbb{R}^{3\times 3}$. If strain within a specific plane is fixed, what values can F take? This question is relevant, among other things, to perform calculations about biaxial epitaxy.

2.5.1 Quantifying in-plane strain

The first challenge that we must address is to quantify what we mean by "fixing the in-plane strain." Fix a two-dimensional vector space $\alpha \subseteq \mathbb{R}^3$; this is the plane of strain. Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$ be an orthonormal basis for α. In this orthonormal basis, we can specify the in-plane strain and rotation by a 2×2 matrix

$$\tilde{F} \in \mathbb{R}^{2\times 2}. \tag{2.9}$$

Consider the matrix P whose rows are $\mathbf{u}$ and $\mathbf{v}$,

$$P = \begin{pmatrix} \mathbf{u}^T \\ \mathbf{v}^T \end{pmatrix} \in \mathbb{R}^{2\times 3}, \tag{2.10}$$

which we informally think of as a "projection" from $\mathbb{R}^3$ to the plane of strain. For a vector $\mathbf{r} \in \mathbb{R}^3$, P acts as $P\,\mathbf{r} = (\mathbf{u}\cdot\mathbf{r},\, \mathbf{v}\cdot\mathbf{r})$; the out-of-plane direction is sent to zero by P. The following 3×2 matrix equation (where F is the variable being constrained) expresses the constraint F "acts like $\tilde{F}$ within the plane α,"

$$FP^T = P^T\tilde{F}. \tag{2.11}$$

Indeed, consider a vector $\mathbf{r} \in \alpha$. Since $\mathbf{r} \in \alpha$, we can write it as $\mathbf{r} = P^T\mathbf{x}$, where $\mathbf{x} \in \mathbb{R}^2$ expresses $\mathbf{r}$ in terms of in-plane coordinates. Acting by $\tilde{F}$ on the in-plane coordinates $\mathbf{x}$, we get $\tilde{F}\mathbf{x}$; elevating it to $\mathbb{R}^3$ using P^T yields $P^T\tilde{F}\mathbf{x}$. On the other hand, we can also use F to take $\mathbf{r} = P^T\mathbf{x}$ to $F\mathbf{r} = FP^T\mathbf{x}$. The two end results should agree for all $\mathbf{x} \in \mathbb{R}^2$. Therefore, $FP^T = P^TF$.

If $\tilde{F} = 1$ is the identity, equation 2.11 requires that $F\mathbf{r} = \mathbf{r}$ whenever $\mathbf{r} \in \alpha$. If $\tilde{F} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ exchanges the first and second variables, $F\mathbf{u} = \mathbf{v}$ and $F\mathbf{v} = \mathbf{u}$. If $F = \begin{pmatrix} \lambda & 0 \\ 0 & 1 \end{pmatrix}$, then F multiplies $\mathbf{u}$ by λ and leaves $\mathbf{v}$ unchanged. We see that the action of F within the plane α is determined by $\tilde{F}$ if we interpret the latter as expressed in the basis of $\mathbf{u}, \mathbf{v}$.

2.5.2 Parametrizing compatible strains

What are all the deformation gradients F compatible with biaxial strain $\tilde{F}$ in the plane α, i.e., those matrices F satisfying equation 2.11? The action of F on the perpendicular direction

$$\mathbf{w} = \mathbf{u} \times \mathbf{v} \tag{2.12}$$

is not constrained by equation 2.11. Indeed, whenever F satisfies $FP^T = P^T\tilde{F}$, the modified matrix $F + \mathbf{r}\,\mathbf{w}^T$ also satisfies the equation because $\mathbf{w}^T P^T = 0$, where $\mathbf{r} \in \mathbb{R}^3$ can be any vector. Noticing that $PP^T = 1$ is the 2×2 identity matrix, it is easy to verify that $F = P^T\tilde{F}P$ satisfies equation 2.11. Therefore, the matrices F solving equation 2.11 are precisely those of the form

$$F = P^T\tilde{F}P + (\mathbf{w} + \mathbf{r})\mathbf{w}^T, \tag{2.13}$$

where $\mathbf{r} \in \mathbb{R}^3$ is any vector. Indeed, F has 9 degrees of freedom because it is a 3×3 matrix. Equation 2.11 is a linear equation in F that takes away 6 degrees of freedom because both sides of the equation are 3×2 matrices. There are three degrees of freedom remaining, and these are parametrized by $\mathbf{r} \in \mathbb{R}^3$. Equation 2.11 does not constrain where the out-of-plane direction $\mathbf{w}$ is mapped, and this freedom is accounted for by $\mathbf{r}$.

2.6 Biaxial epitaxy

2.6.1 Overview of molecular beam epitaxy

Molecular Beam Epitaxy (MBE) [18] [17] is a technique for producing thin crystalline films in a certain way. We start with a chunk of substrate material and cut it in such a way that some crystal plane ((100), (221), etc.) is exposed. Another material, called the *epilayer*, is slowly deposited onto the substrate. Under favorable conditions, the epilayer crystallizes as it lands on the substrate, adapting to the crystal structure of the substrate. We end up with a thin crystalline film of the epilayer material fused to the substrate.

The epilayer, in general, undergoes strain to adapt to the substrate, while the substrate is not strained. Indeed, the substrate-epilayer system will organize itself in such a way as to minimize its total potential energy. Elastic energy is the integral of elastic energy density over volume. The system organizes itself in such a way as to minimize the total elastic energy, and since the substrate has a much greater volume than the epilayer, it undergoes basically no strain.

2.6.2 Quantifying biaxial epitaxy

Recall that the epilayer material is a crystal. Denote by B_1 the Bravais matrix for this material under normal conditions, without epitaxy. Denote by F_1 the deformation gradient

that the epilayer undergoes during epitaxy, so that the new Bravais matrix is $F_1 B_1$.[1] Likewise, denote by B_2 and F_2 the Bravais matrix and deformation gradient for the substrate. As argued in the preceding subsection, the substrate undergoes no strain, so by the polar decomposition discussed in Section 2.3, F_2 is a rotation matrix.

$$\begin{aligned} \text{Epilayer: } & B_1, F_1, \\ \text{Substrate: } & B_2, F_2, \qquad F_2 \text{ a rotation.} \end{aligned}$$

The exposed plane of the epilayer is a two-dimensional lattice. As mentioned in Section 2.4, one can compute triplets of integers $\mathbf{n}_1, \mathbf{m}_1 \in \mathbb{Z}^3$ so that this plane consists of points[2]

$$i\, F_1\, B_1\, \mathbf{n}_1 + j\, F_1\, B_1\, \mathbf{m}_1, \tag{2.14}$$

where $i, j \in \mathbb{Z}$ are integers. That is, if the substrate was magically evaporated and only the epilayer would be left, one would see exposed a two-dimensional sublattice of the epilayer's Bravais lattice given by equation 2.14. Likewise, the points of the exposed plane of the substrate are

$$i'\, F_2\, B_2\, \mathbf{n}_2 + j'\, F_2\, B_2\, \mathbf{m}_2, \tag{2.15}$$

where $i', j' \in \mathbb{Z}$ are integers. For the sublattices from equations 2.15 and 2.14 to match is to have a one-to-one mapping between pairs (i, j) and (i', j'). Clearly, the mapping is also linear. Thus, there has to exist a 2×2 invertible matrix A with integer entries so that

$$A \begin{pmatrix} i \\ j \end{pmatrix} = \begin{pmatrix} i' \\ j' \end{pmatrix}. \tag{2.16}$$

That is, a point $i\, B_1\, \mathbf{n}_1 + j\, B_1\, \mathbf{m}_1$ of the epilayer is matched to $i'\, B_2\, \mathbf{n}_2 + j'\, B_2\, \mathbf{m}_2$ of the substrate. Let us introduce the *sublattice generating matrices* $K_1, K_2 \in \mathbb{Z}^{3\times 2}$ as

$$K_1 \coloneqq \begin{pmatrix} \mathbf{n}_1 & \mathbf{m}_1 \end{pmatrix} \in \mathbb{Z}^{3\times 2}, \qquad K_2 \coloneqq \begin{pmatrix} \mathbf{n}_2 & \mathbf{m}_2 \end{pmatrix} \in \mathbb{Z}^{3\times 2}. \tag{2.17}$$

The entries of these 3×2 matrices are integers, and they quantify the crystal planes being "glued" together. The 2×2 matrix A quantifies in what way the crystal planes are glued. Write $\mathbf{x}$ for (i, j) and $\mathbf{x}'$ for $(i', j') = A\,\mathbf{x}$. With this notation, we see that a point $B_1\, K_1\, \mathbf{x}$ of the epilayer is matched to $B_2\, K_2\, \mathbf{x}' = B_2\, K_2\, A\, \mathbf{x}$ of the substrate. This holds for all pairs of integers $\mathbf{x} = (i, j)$. Therefore, we have the equality of matrices

$$\boxed{F_1 B_1 K_1 = F_2 B_2 K_2 A.} \tag{2.18}$$

[1] The epilayer thin film becomes strained not by starting out unstrained and being continuously deformed, but rather, by crystallizing on the substrate already in a strained configuration. In this sense, the deformation gradient as discussed in Section 2.2 doesn't really make sense. A more proper way to think of F_1 would be as follows. If we look at the epilayer thin film, the crystal lattice should be only a small deformation of the original crystal lattice of the epilayer material. We can therefore find a matrix F_1 that is close to being a rotation matrix and relates the two lattices.

[2] We will be sloppy and identify points with vectors, assuming additionally that the origin of coordinates lies within the exposed plane of the epilayer. Speaking more properly, the exposed plane of the epilayer is a set of points, and $i\, B_1\, \mathbf{n}_1 + j\, B_1\, \mathbf{m}_1$ are those translations with respect to which the set is invariant.

We call it the *sublattice matching equation*. It is a necessary condition for the epilayer and the substrate to match perfectly. When the sublattice matching equation is not satisfied, the surface of crystal 1 exposed by the first cut has a different periodicity from the surface of crystal 2 exposed by the second cut, and the two surfaces cannot be matched perfectly. When the equation *is* satisfied, depending on the atomic structure of crystals 1 and 2, it might be possible for the two cut-crystals to fit together perfectly.

2.6.3 Parameterizing compatible strains

Letting $K_2' = K_2 A$ and using it instead of K_2, we can assume without loss of generality that $A = 1$ is the identity, so that the sublattice matching equation (eq. 2.18) becomes

$$F_1 B_1 K_1 = F_2 B_2 K_2, \qquad F_2 \text{ a rotation matrix.} \tag{2.19}$$

It is to be thought of as an equation in F_1 and F_2. Indeed, B_1 and B_2 are fixed by our choice of the substrate and epilayer materials, while K_1, K_2 are fixed by which crystal planes we are gluing together in both ways; the "unknowns" are F_1 and F_2.

The sublattice matching equation equation does not have a unique solution for three reasons. The first two are relatively unsurprising, but the third reason requires attention.

1. (K_1 and K_2 are parameters.) The sublattice generating index matrices $K_1, K_2 \in \mathbb{Z}^{3\times 2}$ (equation 2.17) are parameters that must be chosen. The choice of $K_2 = \begin{pmatrix}\mathbf{n}_2 & \mathbf{m}_2\end{pmatrix}$ specifies the plane $\{i\, B_2\, \mathbf{n}_2 + j\, B_2\, \mathbf{m}_2 : i, j \in \mathbb{Z}\}$ of the substrate crystal on which the epilayer crystal will be grown, while the choice of $K_1 = \begin{pmatrix}\mathbf{n}_1 & \mathbf{m}_1\end{pmatrix}$ specifies the orientation in which the epilayer crystal will grow on top of the substrate crystal. Thus, equation 2.18 is really not one equation but a family of equations, one for each choice of crystal orientations.

2. (Rotation symmetry.) If U is any 3×3 rotation matrix and F_1 and F_2 solve the sublattice matching equation, then UF_1 and UF_2 solve the sublattice matching equation because $(UF_1)B_1K_1 = (UF_2)B_2K_2$. This corresponds to the fact that when two crystals are glued together, one can rotate the entire system, producing a setup that differs only trivially from the first one. Therefore, we may choose F_2 to be any rotation matrix we desire.

3. (Freedom of the perpendicular direction.) Fix $K_1, K_2 \in \mathbb{Z}^{3\times 2}$ and a rotation matrix F_2. Then, the sublattice matching equation $F_1 B_1 K_1 = F_2 B_2 K_2$ becomes a linear equation for $F_1 \in \mathbb{R}^{3\times 3}$ that takes away 6 degrees of freedom. Since an arbitrary 3×3 matrix F_1 has 9 degrees of freedom, there are 3 degrees of freedom remaining. That is, kinematically, if we consider the space of all possible lattice parameters (equivalently, Bravais matrices) that would allow the epilayer to match with the substrate along the plane of growth, it is three-dimensional. In the actual physical experiment, of course, the epilayer just has some specific value of lattice parameters. How does the three-dimensional space of possible lattice parameters get reduced to one point? Within the

constraints imposed by epitaxy, the epilayer crystal will grow in such a way that its elastic energy is minimized – this is discussed in Section 3.7.

Leaving elastic energy minimization to Chapter 3, let us parametrize the space of kinematically allowed lattice parameters, e.g., the space of lattice parameters that allow the epilayer to match with the substrate along the plane of growth. An instead of parameterizing lattice parameters (or Bravais matrices B_{strained}), we take the equivalent problem of parameterizing deformation gradients $F_1 = B_{\text{strained}} B_{\text{bulk}}^{-1}$. Firstly, we pick the parameters K_1 and K_2. Secondly, we find one solution (F_1, F_2) to the sublattice matching equation using Algorithm 1, provided in Appendix B. The substrate's rotational freedom is fixed, so that it is now locked in place. Equation 2.19 then becomes a linear equation in F_1. Thirdly, in a similar way to Subsection 2.5.2, we parametrize all solutions up to a rotation. The process is somewhat technical, but it is described in Appendix B. The end result is a parametrization

$$F_1 = D + \mathbf{r}\,\mathbf{w}^T \tag{2.20}$$

of epilayer deformation gradients that are allowed by the sublattice matching equation. Here, $\mathbf{r}$ is a vector that is free to take any value (it parametrizes the freedom in biaxial epitaxy), D is a symmetric matrix representing one possible value of the deformation gradient, and $\mathbf{w}$ is a normal vector to the plane of straining. The value of $\mathbf{r}$ that corresponds to the actual lattice parameters assumed by the epilayer upon crystallization is one that makes the elastic energy the smallest – this is discussed in detail in Section 3.7. The parametrization is up to a rotation, meaning that to recover all solutions to equation 2.19, we would need consider not only those solutions described by equation 2.20, but also their rotations, namely $(U\,F_1, U\,F_2)$ for rotation matrices U.

The parametrization (eq. 2.20) has the property that D is a "reasonable" positive definite matrix, and $\mathbf{r} = 0$ corresponds to a reasonable deformation of the epilayer. Picking large $\mathbf{r}$ would impose a large strain on the epilayer, causing it to break. Therefore, in equation 2.20, we need to only consider small values of $\mathbf{r} \in \mathbb{R}^3$.

So what cell parameters does the epilayer actually take during biaxial epitaxy? What value of $\mathbf{r}$ is chosen? Answering this question requires more than knowing the geometry of the epilayer and substrate crystal – namely, we need to know the elastic energy of the epilayer as a function of its lattice parameters. The lattice parameters taken will be those that minimize elastic energy under the given biaxial strain. We proceed to develop in Chapter 3 a model for the elastic energy density of RuO_2 as a function of strain, and in Section 3.7 we are able to fully answer the question of what lattice parameters are taken under a given biaxial strain.

Chapter 3

Elastic energy of rutile materials

3.1 Strain and stress

We emphasize the distinction between strain and stress. Strain is a purely geometric notion, while stress quantifies the pressures and tensions inside the material induced by strain [29, 30].

$$\begin{aligned} \text{Strain} &= (\text{deformation of the crystal lattice}), \\ \text{Stress} &= (\text{elastic forces in response to strain}). \end{aligned} \tag{3.1}$$

Consider a macroscopic rectangular box filled with RuO_2, and imagine that we have at our disposal a powerful machine with the following functionality. For whatever box dimensions we desire, the machine is able to apply such forces to the box that the box, and the RuO_2 inside it, conform to the desired dimensions. The desired deformation corresponds to the strain, whereas the forces the machine has to apply correspond to the stress.

Considering this hypothetical setup is a useful middle step towards describing realistic physical experiments. An important scenario is biaxial strain, a situation where a crystal is forced to conform to specified lattice parameters in a given plane but is left free to relax the out-of-plane direction. Recall that in Subsection 2.5.2, we characterized the out-of-plane freedom that the crystal has under biaxial strain. In Section 3.6, we develop a model for the elastic energy of RuO_2 as a function of strain. Under the constraints imposed by biaxial strain, RuO_2 settles in those lattice parameters that minimize its elastic energy. Combining the results of Subsection 2.5.2 and Section 3.6, we develop in Section 3.7 a way to calculate, for a given biaxial strain, the relaxed lattice parameters of RuO2.

3.2 Elastic energy density

Consider a crystal with a Bravais matrix B. For any deformation gradient F, a DFT calculation allows us to compute $W(F)$, the *energy density*[1] of the crystal when it is deformed to

[1]We compute it by running a "relax" calculation in Quantum ESPRESSO [22] [23] [24], so that the cell parameters are fixed to a given value but the atoms are allowed to move within the cell to find a least-energy

have cell parameters $B' = FB$,

$$W(F) = (\text{energy density with Bravais matrix } B' = FB). \tag{3.2}$$

If the material is subject to small strain, $F = 1 + E$, the energy density increases. Indeed, if there was a small strain E such that $W(1 + E) < W(1)$, the crystal would spontaneously deform from Bravais matrix B to $(1 + E)B$. This cannot happen, however, because the material is stable in maintaining its Bravais matrix and resists strain. We recognize that the difference in the energy densities of the strained crystal and the unstrained one,

$$W(F) - W(1), \tag{3.3}$$

as the elastic energy density. The elastic energy density is zero when no strain is applied and is positive for any small strain.

For large strain, the DFT-predicted elastic energy density might potentially be negative, either due to a computational artifact of DFT or a true physical phase transition that occurs under large strain. However, we need not worry about this fact because we will only consider small strain. The distinction between "small" and "large" strain is not something we make precise, but very roughly speaking, strain up to 5% is small, while 20% or greater is large. Strains as large as 20% are certainly not in the elastic regime, and we will not need to worry about them.

3.3 $P136$ symmetry of the elastic energy function

As described in Part 7 of the International Tables for Crystallography, Volume A [31], the space group P136, with respect to which RuO_2 is symmetric, is generated by 90-degree rotations around the z axis, reflections around the yz plane (e.g., changing x to $-x$), and reflection around the xy plane (e.g., changing z to $-z$). That is, whenever a matrix M is equal to either of the matrices

$$C_4 = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad m_x = \begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad m_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

then the RuO_2 crystal is symmetric with respect to $M \in \{C_4, m_x, m_z\}$ in the sense that the energy is invariant when the deformation gradient is right-multiplied by M,

$$W(FM) = W(F). \tag{3.4}$$

Indeed, recall that $W(F)$ is the energy density of the crystal when it is deformed to have Bravais matrix FB. Since transforming the crystal by $M \in \{C_4, m_x, m_z\}$ just yields the same crystal again, $W(FM)$ and $W(F)$ just both express the energy density of the same

configuration. The resulting energy density is expressed in Rydberg per unit cell.

crystal and are therefore equal. A more detailed account is given in Section 3.4 of [27]. Let G be the set of matrices that can be generated as products of C_4, m_x, and m_z. We note that the elastic energy density is invariant under any element of G. Indeed, when $M_1, M_2 \in G$ are two matrices in G, for any deformation gradient F,

$$W(FM_2M_1) = W((FM_2)M_1) = W(FM_2) = W(F),$$

so $M_2M_1 \in G$. For example, the matrix m_y that sends y to $-y$ and leaves x and z the same is in G because it can be realized as $m_y = C^2 m_x m_z$.

Moreover, whenever M is a rotation matrix, the elastic energy is invariant when the deformation gradient F is left-multiplied by M,

$$W(MF) = W(F), \tag{3.5}$$

since the lattice defined by MF is simply a rotation of the lattice defined by F and rotation leaves physical quantities invariant. It is curious that symmetry with respect to C_4, m_x, m_z is with right multiplication, while symmetry with respect to rotation matrices is with left multiplication.

The matrices C_4 and m_xm_y, m_ym_z, m_zm_x are simultaneously rotation matrices and elements of G. These four matrices, together with their inverses, are elements of G and also rotation matrices. Indeed, the matrix $m_xm_y = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ can be realized as a 180-degree rotation around the z axis, and likewise for m_ym_z and m_zm_x. The matrix C_4 is a 90-degree rotation around the z-axis. The matrices m_xm_y, m_ym_z, m_zm_x are their own inverses, while the inverse of C_4 is C_4^3. Let $M \in \{C_4, m_xm_y, m_ym_z, m_zm_x\}$, and consider a deformation gradient $F = 1 + E$ for some strain matrix E. By equations 3.5 and 3.4,

$$W(1 + E) = W(1 + M\,E\,M^{-1}). \tag{3.6}$$

Indeed,

$$W(1 + M\,E\,M^{-1}) = W(M(1 + E)M^{-1}) = W(M(1 + E)) = W(1 + E).$$

Equation 3.6 states that conjugating the strain matrix by $M \in \{C_4, m_xm_y, m_ym_z, m_zm_x\}$ leaves the elastic energy unchanged. In Section 3.4, we define a convenient parametrization of strain to leverage this symmetry.

3.4 A highly symmetric parametrization of strain

In the Bravais lattice of a rutile material, the xy plane is comprised of squares, and the z axis is perpendicular to this plane. The xy plane is symmetric with respect to 90 degree rotations. The xy plane and the z axis should be, in some sense, independent, except for the fact that compressing the crystal in the z direction should make it "want" to expand

	τ_1	τ_2	τ_3	τ_4	τ_5	τ_6
C	$+$	$+$	$-$	$-$	τ_6	$-\tau_5$
$m_x m_y$	$+$	$+$	$+$	$+$	$-$	$-$
$m_y m_z$	$+$	$+$	$+$	$-$	$-$	$+$
$m_z m_x$	$+$	$+$	$+$	$-$	$+$	$-$

Table 3.1: Conjugation of $\tau_1, \ldots, \tau_6$ by some elements of the $P136$ symmetry group. Here, "$+$" means that $M\,\tau_i\,M^{-1} = \tau_i$, while "$-$" means that $M\,\tau_i\,M^{-1} = -\tau_i$.

in the xy plane. Driven by this physical intuition, we write down a collection of matrices $\tau_1, \ldots, \tau_6$ that happen to transform nicely under conjugation by C_4, $m_x m_y$, $m_y m_z$, and $m_z m_x$ discussed in Section 3.3. The matrices are,

$$\tau_1 = \begin{pmatrix} 1 & & \\ & 1 & \\ & & 0 \end{pmatrix}, \qquad \tau_2 = \begin{pmatrix} 0 & & \\ & 0 & \\ & & \sqrt{2} \end{pmatrix}, \qquad \tau_3 = \begin{pmatrix} 1 & & \\ & -1 & \\ & & 0 \end{pmatrix},$$
$$\tau_4 = \begin{pmatrix} 0 & 1 & \\ 1 & 0 & \\ & & 0 \end{pmatrix}, \qquad \tau_5 = \begin{pmatrix} 0 & & 1 \\ & 0 & \\ 1 & & 0 \end{pmatrix}, \qquad \tau_6 = \begin{pmatrix} 0 & & \\ & 0 & 1 \\ & 1 & 0 \end{pmatrix}, \tag{3.7}$$

where missing matrix entries are zero. The normalization for τ_2 is chosen so that the matrices are orthogonal under the trace inner product (Appendix A.3) and all have norm-squared 2,

$$\operatorname{tr}\left(\tau_i^T \tau_j\right) = 2\,\delta_{ij}. \tag{3.8}$$

We think of τ_1 as a dilation in the xy plane; τ_2 as a dilation in the z direction; τ_3 and τ_4 as hyperbolic rotations in the xy plane; and τ_5, τ_6 as hyperbolic rotations that mix the x, y directions with the z direction. These represent "reasonable" physical strains that should be largely decoupled from one another, except possibly that compression in the xy plane should be coupled with compression along the z axis.[2] To make this intuition precise, we compute the conjugations $M\,\tau_i\,M^{-1}$ for $M \in \{C_4, m_x m_y, m_y m_z, m_z m_x\}$ and $i = 1, \ldots, 6$. These are presented in Table 3.1. In Section 3.5, we explore the physical implications of these conjugation relations in the context of the energy density $W(F)$.

3.4.1 Orthogonality of the basis

The tau matrices form a basis for the vector space of symmetric 3×3 matrices. Therefore, every strain matrix E can be expressed as

$$E = c_1\,\tau_1 + \cdots + c_6\,\tau_6, \tag{3.9}$$

[2]For τ_1, for example, we imagine expanding the material in the xy plane while forcing the z direction to remain unchanged. There might be stress forces in the z direction; in this sense, to produce a τ_1 strain would requite applying forces not only in the xy plane but in the z direction too. One can imagine filling a box with RuO_2 and changing the dimensions of the box in the desired way.

where the *strain numbers* can be computed, using equation 3.8, as

$$c_i = \frac{1}{2}\,\mathrm{tr}\big(\tau_i^T\, E\big). \tag{3.10}$$

Since the strain matrix E is symmetric, its eigenvalues $\lambda_1, \ldots, \lambda_3$ are real. Per equation 3.8, the norm-squared of the strain is twice the Euclidean norm-squared of $(c_1, \ldots, c_6) \in \mathbb{R}^6$,

$$\begin{aligned} ||E||^2 = \sum_{i,j=1}^{3} E_{ij}^2 = \lambda_1^2 + \cdots + \lambda_3^2 = \mathrm{tr}(E^T\, E) \\ = 2\left(c_1^2 + \cdots + c_6^2\right). \end{aligned} \tag{3.11}$$

3.5 Enumeration of Taylor series terms compatible with symmetry

Recall that a strain matrix E can be decomposed in terms of $\tau_1, \ldots, \tau_6$ as $c_1\tau_1 + \cdots + c_6\tau_6$ where $c_1, \ldots, c_6$ are the strain numbers we introduced in Subsection 3.4.1. The energy density $W(1+E)$ at this strain can be expanded as a Taylor series

$$W(1+E) \;=\; \sum_{n_1=0}^{\infty} \cdots \sum_{n_6=0}^{\infty} K_{n_1,\ldots,n_6}\, c_1^{n_1} \ldots c_6^{n_6} \tag{3.12}$$

with some coefficients $K_{n_1,\ldots,n_6} \in \mathbb{R}$. The term $K_{0,\ldots,0} = W(1)$ is the energy density with no strain; it is the reference value so that the difference $W(1+E) - K_{0,\ldots,0}$ is the elastic energy density. The six linear terms $K_{1,0,\ldots,0}$, $K_{0,1,0,\ldots,0}$, $\ldots$, $K_{0,\ldots,0,1}$ are equal to zero, since for all i, the derivative $\partial W/\partial c_i$ evaluated at zero strain $c_1 = \cdots = c_6 = 0$ is zero; the unstrained configuration is a local minimum of energy. (If the derivative were nonzero, the crystal would spontaneously deform from its unstrained configuration, which does not happen.) The second-order terms, such as $K_{2,0,\ldots,0}$, encode the stiffness tensor of the crystal. Third and higher-order terms encode small corrections studied in nonlinear elasticity theory. For $n \geq 0$, there are

$$\binom{n+5}{5} \;\sim\; n^5/120 \tag{3.13}$$

terms of total degree $n_1 + \cdots + n_6 = n$, as can be seen by imagining that we are counting the number of permutations of n balls and 5 spacers. To leading order, this number grows as $n^5/120$. There are

$$\binom{7}{5} = 21$$

quadratic terms – or is there really? The reader familiar with elasticity theory [30] [29] [27] might recall that the stiffness tensor of a rutile material can be specified by 6 numbers; this economy is due to rutile symmetry. Equation 3.6 states that the energy at strain E is the

same as that at strain $M\,E\,M^{-1}$, where M is one of C_4, $m_x m_y$, $m_y m_z$, and $m_z m_x$. Combined with equation 3.12, this means that

$$\sum_{n_1=0}^{\infty} \cdots \sum_{n_6=0}^{\infty} K_{n_1,\ldots,n_6}\, c_1^{n_1} \ldots c_6^{n_6} \;=\; \sum_{n_1=0}^{\infty} \cdots \sum_{n_6=0}^{\infty} K_{n_1,\ldots,n_6}\, (c_1')^{n_1} \ldots (c_6')^{n_6}, \tag{3.14}$$

where $c_1', \ldots, c_6'$ are determined from $c_1, \ldots, c_6$ according to Table 3.1 and by requiring that $M\,E\,M^{-1}$ is $c_1'\tau_1 + \cdots + c_6'\tau_6$. In order for equation 3.14 to hold, the coefficients before equal monomials on the right-hand side and left-hand side have to be equal. Let us first consider the transformation $m_x m_y$, which negates the signs of c_5 and c_6 and leaves $c_1, \ldots, c_4$ unchanged. Equation 3.14 then becomes

$$\sum_{n_1=0}^{\infty} \cdots \sum_{n_6=0}^{\infty} K_{n_1,\ldots,n_6}\, c_1^{n_1} \ldots c_6^{n_6} \;=\; \sum_{n_1=0}^{\infty} \cdots \sum_{n_6=0}^{\infty} K_{n_1,\ldots,n_6}\, c_1^{n_1} \ldots c_6^{n_6}\, (-1)^{n_5+n_6},$$

or equivalently,

$$\sum_{n_1=0}^{\infty} \cdots \sum_{n_6=0}^{\infty} \left(1 + (-1)^{1+n_5+n_6}\right) K_{n_1,\ldots,n_6}\, c_1^{n_1} \ldots c_6^{n_6} \;=\; 0,$$

from which it follows that $K_{n_1,\ldots,n_6} = 0$ whenever $n_5 + n_6$ is odd. Likewise, by considering the conjugation by $c_1, \ldots, c_6$ by $m_y m_z$ and $m_z m_x$, we see from Table 3.1 that $K_{n_1,\ldots,n_6} = 0$ whenever $n_4 + n_5$ or $n_4 + n_6$ is odd. It follows from here that n_4, n_5, n_6 must all have the same parity for $K_{n_1,\ldots,n_6}$ to be nonzero.

$$K_{n_1,\ldots,n_6} \neq 0 \quad\Longrightarrow\quad n_4 \equiv n_5 \equiv n_6 \pmod 2. \tag{3.15}$$

Under C_4, the strain numbers c_1 and c_2 remain the same, c_3 and c_4 pick up a minus sign, while $c_5' = -c_6$ and $c_6' = c_5$. Equation 3.12 transforms to

$$\begin{aligned} W(1 + C_4\,E\,C_4^3) \;&=\; \sum_{n_1,\ldots,n_6} K_{n_1,\ldots,n_5,n_6}\, c_1^{n_1} c_2^{n_2} (-c_3)^{n_3} (-c_4)^{n_4} (-c_6)^{n_5} c_5^{n_6} \\ &=\; \sum_{n_1,\ldots,n_6} K_{n_1,\ldots,n_6,n_5}\, (-1)^{n_3+n_4+n_6}\, c_1^{n_1} \ldots c_6^{n_6}. \end{aligned}$$

By equation 3.15, if $K_{n_1,\ldots,n_6}$ is nonzero, then $n_4 + n_6 \equiv 0 \pmod 2$, so $(-1)^{n_4+n_6} = 1$; if $K_{n_1,\ldots,n_6}$ is zero, then it doesn't matter if we multiply it by -1. Therefore, we may replace $(-1)^{n_3+n_4+n_6}$ with $(-1)^{n_3}$. Since $W(1+E) = W(1 + C_4\,E\,C_4^3)$, we therefore have

$$K_{n_1,\ldots,n_5,n_6} = (-1)^{n_3} K_{n_1,\ldots,n_6,n_5}, \tag{3.16}$$

that is, the K coefficients with n_5 and n_6 exchanged are related by a factor of $(-1)^{n_3}$. An immediate consequence is that $K_{n_1,\ldots,n_6} = 0$ whenever $n_5 = n_6$ but n_3 is odd.

Let us examine which quadratic terms can be nonzero. Since n_4, n_5, n_6 all have the same parity, and since $n_1 + \cdots + n_6 = 2$, there are four options: $n_4 = 2$, $n_5 = 2$, $n_6 = 2$, or $n_4 = n_5 = n_6 = 0$. In the latter case, since $n_5 = n_6$, by equation 3.16, it has to be that n_3 is even – so either $n_3 = 2$ or $n_3 = 0$. If $n_3 = 0$, then $n_1 + n_2 = 0$ leaves three options: $(2, 0)$, $(1, 1)$, and $(0, 2)$. Thus, there are seven tuples $(n_1, \ldots, n_6)$ satisfying equation 3.15 and 3.16, namely

$$\begin{array}{llll} (2,0,0,0,0,0), & (0,2,0,0,0,0), & (1,1,0,0,0,0), & (0,0,2,0,0,0), \\ (0,0,0,2,0,0), & (0,0,0,0,2,0), & (0,0,0,0,0,2). & \end{array}$$

Moreover, by equation 3.16, the last two tuples have equal coefficients. These correspond to the 6 free components of the stiffness tensor of a rutile material! All except one of these tuples yield squares of the strain numbers $c_1^2, \ldots, c_6^2$. The one exception, $(1, 1, 0, 0, 0, 0)$, corresponds to the term $c_1 c_2$. Recall from equation 3.7 that c_1 represents dilation in the xy plane, while c_2 represents dilation in the z direction. Thus, the only two coupled strain directions are dilations/contractions in the xy plane and the z direction. It makes sense that these two strains should be coupled, since compressing RuO_2 in the z direction should make it "want" to expand in the xy plane, and vice versa. Yay!

Let us now proceed to try and enumerate higher terms compatible with rutile symmetry. These higher terms will allow us to account for the small nonlinear corrections on top of the dominant terms considered in the linear theory of elasticity. In light of equations 3.15 and 3.16, the Taylor series for the energy density $W(1 + E)$ is

$$\begin{aligned} W(1+E) &= \sum_{n_1,n_2,n_3,n_4=0}^{\infty} c_1^{n_1} \ldots c_4^{n_4}\, \tilde{W}_{n_1,\ldots,n_4}, \\ \tilde{W}_{n_1,\ldots,n_4} &= \sum_{\substack{a<b \\ a,b\equiv n_4}} K_{n_1,\ldots,a,b}\, \left(c_5^a c_6^b + (-1)^{n_3} c_5^b c_6^a\right) + \begin{cases} \sum_{a\equiv n_4} K_{n_1,\ldots,a,a}\, c_5^a c_6^a & n_3 \text{ even} \\ 0 & n_3 \text{ odd} \end{cases}, \end{aligned} \tag{3.17}$$

where the sum over a, b goes over those pairs (a, b) of nonnegative integers that $a < b$ and a, b both have the same parity as n_4. The number of independent coefficients $K_{n_1,\ldots,n_6}$ of total order 2 through 8, as well as the number one would have without symmetry (eq. 3.13), is given in Table 3.2. For example, there are $1+6 = 7$ terms up to order 2 when we use rutile symmetry because there is one term of degree zero $K_{0,\ldots,0}$, six independent terms of degree two (as illustrated a few paragraphs above), and zero terms of degree one. The degree-one terms are always zero because the unstrained configuration is a local minimum of energy.

3.5.1 Compact notation for the series

Recall equation 3.12 decomposing the elastic energy density into a Taylor series,

$$W(1+E) = \sum_{\mathbf{n}\in\mathbb{Z}^6_{\geq 0}} K_{\mathbf{n}}\, c^{\mathbf{n}},$$

	Rutile symmetry		**General case**	
n	$\deg = n$	$\deg \leq n$	$\deg = n$	$\deg \leq n$
2	6	7	21	22
3	12	19	56	78
4	25	44	126	204
5	44	88	252	456
6	77	165	462	918
7	124	289	792	1710
8	196	485	1287	2997

Table 3.2: The number of coefficients $K_{n_1,\ldots,n_6}$ with total degree $\deg = n_1 + \cdots + n_6$ either equal or at most n, for n from 2 to 8. The first column shows the counts with rutile symmetry (eq. 3.17), while the second column shows the counts in the general case (eq. 3.12).

where $\mathbb{Z}^6_{\geq 0}$ is the set of all tuples $\mathbf{n} = (n_1, \ldots, n_6)$ of six nonnegative integers, $K_{\mathbf{n}}$ is a shorthand for $K_{n_1,\ldots,n_6}$, and $c^{\mathbf{n}}$ is a shorthand for $c_1^{n_1} \ldots c_6^{n_6}$. Let us also recall equation 3.17, which uses rutile symmetry to reduce the number of independent coefficients $K_{\mathbf{n}}$. With the notation we are about to define below, one may write the equation as

$$W(1 + E) = \sum_{\mathbf{n} \in S} K_{\mathbf{n}}\, c[\mathbf{n}]. \tag{3.18}$$

In equation 3.18, $S \subseteq \mathbb{Z}^6_{\geq 0}$ is the set consisting of all tuples $(n_1, \ldots, n_4, a, b)$ considered in equation 3.17. More precisely, a tuple of nonnegative integers $(n_1, \ldots, n_4, a, b)$ lies in S if all of the following conditions are satisfied.

(i) a and b have the same parity as n_4, $a \equiv n_4 \pmod 2$ and $b \equiv n_4 \pmod 2$.

(ii) The total degree is not one, $n_1 + \cdots + n_4 + a + b \neq 1$.

(iii) Either $a < b$, or $a = b$ and $n_3 \equiv 0 \pmod 2$.

We require in point (ii) that the total degree is not 1 because any nonzero linear terms in equation 3.17 would contradict the fact that the unstrained configuration is a local minimum of energy density. Also, $c[\mathbf{n}]$ for $\mathbf{n} = (n_1, \ldots, n_4, a, b) \in S$ is defined as

$$c[\mathbf{n}] = \begin{cases} c_1^{n_1} \ldots c_4^{n_4} c_5^a c_6^b + (-1)^{n_3} c_1^{n_1} \ldots c_4^{n_4} c_5^b c_6^a & a < b, \\ c_1^{n_1} \ldots c_4^{n_4} c_5^a c_6^a & a = b. \end{cases} \tag{3.19}$$

3.6 A model for the elastic energy density

Equation 3.18 expresses the energy density of a rutile material as a Taylor series $\sum_{\mathbf{n} \in S} K_{\mathbf{n}}\, c[\mathbf{n}]$, summing over terms compatible with rutile symmetry. We now want to consider finite approximations to the full Taylor series. Let S_n be the set of all $\mathbf{n} \in S$ of total degree at most

n,

$$S_n = \{\mathbf{n} = (n_1, \ldots, n_6) \in S : n_1 + \cdots + n_6 \leq n\}. \tag{3.20}$$

Note that since first-order terms are removed from S to make sure that the unstrained configuration is a local minimum of energy (Subsection 3.5.1). The "deg $\leq n$" side of the left column of Table 3.2 gives the numbers $|S_n|$ of elements in S_n, so that, for example, S_5 has 88 elements.

Taking equation 3.17 and leaving terms of total degree up to n, we arrive at the approximate equality

$$W(1+E) \approx \sum_{\mathbf{n}\in S} K_{\mathbf{n}}\, c[\mathbf{n}]. \tag{3.21}$$

We randomly choose some number p of strain matrices $E^{(1)}, \ldots, E^{(p)}$ and use Density Functional Theory (DFT) to compute, for each strain matrix E_i, the energy density $W^{(i)} = W(1 + E^{(i)})$ (the computation is described later in this section). The right-hand side of equation 3.21 is linear in the model parameters $K_{\mathbf{n}}$, and it is a standard exercise to fit the model parameters to data using the method of least squares. Namely, for each strain $E^{(i)}$, let us compute the strain numbers $c_1^{(p)}, \ldots, c_6^{(p)}$ according to equation 3.10. Let l be the number $|S_n|$ of elements in S_n; enumerate the elements of S_n in some order,

$$S_n = \{\mathbf{n}_1, \ldots, \mathbf{n}_l\}.$$

Then, equation 3.21 can be written in matrix form as

$$\begin{pmatrix} W^{(1)} \\ \vdots \\ W^{(p)} \end{pmatrix} \approx \begin{pmatrix} c^{(1)}[\mathbf{n}_1] & \ldots & c^{(p)}[\mathbf{n}_l] \\ \vdots & \ddots & \vdots \\ c^{(p)}[\mathbf{n}_1] & \ldots & c^{(p)}[\mathbf{n}_l] \end{pmatrix} \begin{pmatrix} K_{\mathbf{n}_1} \\ \vdots \\ K_{\mathbf{n}_l} \end{pmatrix}. \tag{3.22}$$

Provided $p \geq l$ – that is, we have not more parameters than data points, which we enforce by making n not too large – we can use the method of least-squares to find a value of $(K_{\mathbf{n}_1}, \ldots, K_{\mathbf{n}_l})$ that makes the norm-squared of the residual in equation 3.22 as small as possible. This is how we fit our model parameters $K[\mathbf{n}_i]$.

Upon fitting the model, we predict the energy density at Bravais matrix B' as follows. Firstly, compute the deformation gradient $F = B'\, B^{-1}$. Secondly, use the polar decomposition to write $F = UD$, where U is a rotation and D is positive definite. Thirdly, compute the strain $E = F - 1$. Fourthly, compute the strain numbers $c_1, \ldots, c_6$ as $c_i = (1/2)\,\mathrm{tr}(\tau_i\, E)$. Fifthly, substitute the values of $c_1, \ldots, c_6$ into equation 3.19 to compute $\mathbf{c}[\mathbf{n}_1], \ldots, \mathbf{c}[\mathbf{n}_l]$. And sixthly, compute the predicted energy density as

$$\widehat{W}(F) = \begin{pmatrix} \mathbf{c}[\mathbf{n}_1] & \ldots & \mathbf{c}[\mathbf{n}_l] \end{pmatrix} \begin{pmatrix} K_{\mathbf{n}_1} \\ \vdots \\ K_{[\mathbf{n}_l} \end{pmatrix}. \tag{3.23}$$

3.6.1 Small-strain dataset

We acquire data by running Density Functional Theory (DFT) calculations. We randomly select some number N of matrices ϵ,

$$\begin{pmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} \\ \epsilon_{21} & \epsilon_{22} & \epsilon_{23} \\ \epsilon_{31} & \epsilon_{32} & \epsilon_{33} \end{pmatrix},$$

where each ϵ_{ij} is selected from a Gaussian distribution with some standard deviation σ and mean zero. Then $||\epsilon||^2 = 9 \cdot \sigma^2$, so $||\epsilon|| = 3\sigma$. Since ϵ is not necessarily symmetric, it is generally not the strain matrix. The strain matrix will be approximately $E \approx (\epsilon + \epsilon^T)/2$, so it will also have a magnitude of around 3% on average. We compute the strained Bravais matrices as

$$B_1' = (1 + \epsilon_1)\, B,\ \ldots,\ B_N' = (1 + \epsilon_N)\, B,$$

where B is the Bravais matrix of bulk (unstrained) RuO_2. For each B_i', we run a `relax` calculation on Quantum ESPRESSO [22] [23] [24] to calculate the energy density at the given lattice parameters, where the ruthenium and oxygen atoms are allowed to move within the crystal cell to find a configuration of least energy.

We have $N = 223$ data points, generated with $\sigma = 0.012$. Fix a maximum degree n, so that terms of degree up to n are included in the polynomial model. We allocate $p = 180$ data points for fitting the model and $q = 43$ for testing, using the `random.sample` function in Python to randomly distribute data between the fitting and testing datasets, fixing the random number generator seed at 0 for consistency between subsequent runs. Denote by C the matrix on the right-hand side of equation 3.22, and by $\mathbf{w}$ the vector on the left-hand side. We compute the model parameters $(K_{\mathbf{n}_1}, \ldots, K_{\mathbf{n}_l})$ by running `numpy.linalg.lstsq`, which outputs a value of $(K_{\mathbf{n}_1}, \ldots, K_{\mathbf{n}_l})$ that makes the norm-squared of the left-hand side and the right-hand side in 3.22 as small as possible. Having fit the model, we proceed to testing. For each of the $q = 43$ testing data points (which were not used for fitting), we use equation 3.23 to compute the predicted energy density $\widehat{W}_i$. We compute the root-mean-squared residual as

$$\Delta W = \sqrt{\frac{1}{q} \sum_{i=1}^{q} \left(\widehat{W}_i - W_i\right)^2},$$

where W_i is the DFT-calculated energy density at the i-th testing data point and $\widehat{W}_i$ is the predicted energy density, for $1 \le i \le q = 43$. The standard deviation of the DFT-calculated energy densities is

$$\sigma_W = \sqrt{\frac{1}{q} \sum_{i=1}^{q} \left(\mu(W) - W_i\right)^2} = 2.432 \times 10^{-3}\ \text{Ry} = 33.09\ \text{meV},$$

where $\mu(W) = (W_1 + \cdots + W_q)/q$ is the mean energy density among the testing data points. The standard deviation σ_W indicates the scale on which the energy density varies. We also

n	# params.	$\log_{10}$(rel. err.)
0	1	0.00
2	7	−0.99
3	19	−2.07
4	44	−1.97
5	88	−1.65
6	165	−1.39

Table 3.3: A polynomial model described in Section 3.6 is fitted to $p = 178$ data points and tested on $q = 45$ data points with small strain. As the degree n of the model is increased, the number of model parameters increases. The $n = 3$ model with 19 parameters is found to have the greatest accuracy (smallest relative error).

compute the root-mean-square magnitude of the strain from the testing data to be 3.0%. We compute the base-10 logarithm of the relative error, $\log_{10}(\Delta W/\sigma_W)$, which measures how accurate the polynomial model is compared to using a constant value for the predicted energy density. This single number gives a summary of the model's accuracy on the testing data points, none of which were used for training. Recall that for each degree n, there is a different model – namely, for a degree n, we include in our polynomial model all Taylor series terms of degree up to n that are compatible with symmetry. We try $n = 2, 3, 4, 5, 6$. The results are given in Table 3.3.

As seen from Table 3.3, the $n = 3$ model with 19 parameters is the most accurate, with a relative error of about $10^{-2} = 1\%$. This high degree of accuracy suggests that the model successfully predicts the energy density of RuO_2 as a function of small strain.

3.6.2 Large-strain dataset

As will be discussed later, one of our motivations for fitting the model for the elastic energy of RuO_2 is to be able to perform relaxation calculations for biaxial strain. To achieve a high degree of accuracy with high biaxial strain, we wish to include high-order Taylor series terms in our model. Now consider the polynomial model fitted from small strain (Subsection 3.6.1). We note that for $n = 4$, even though the model has more data points to learn from than parameters to fit (180 data points for 44 parameters), the model is not very successful at generalizing from training data to testing data, as evidenced by the fact that the $n = 4$ model is slightly less accurate than the $n = 3$ model. Our proposed explanation is that the training and testing datasets consist of strains so close to zero that the contribution of higher-order Taylor series terms is overshadowed by numerical errors of DFT, and that therefore, these higher-order terms cannot be effectively learned. We now prepare a dataset with greater values of strain.

Using the same procedure as in Subsection 3.6.1, we prepare 80 data points with ϵ chosen randomly with $\sigma = 0.018$, and another 160 data points with ϵ chosen randomly with $\sigma = 0.028$. We now have 463 data points. Performing the same testing steps as

n	# params.	$\log_{10}$(rel. err.)
0	1	0.00
2	7	−0.82
3	19	−1.24
4	44	−1.37
5	88	−1.39
6	165	−1.44
7	289	−0.89

Table 3.4: A polynomial model described in Section 3.6 is fitted to $p = 370$ data points and tested on $q = 93$ data points with large strain. As the degree n of the model is increased, the number of model parameters increases. The $n = 6$ model with 165 parameters is found to have the greatest accuracy (smallest relative error).

in the previous subsection, we catalog in Table 3.4 the relative error for model degrees $n = 0, 2, 3, 4, 5, 6, 7$. The standard deviation of the testing energies is now 42.85 meV, and the root-mean-square strain is 4.8%. The largest strain in the testing strains is 8.8%. While strains as large as 8.8% would probably cause the RuO_2 crystal to crack and therefore are not of experimental interest, it is perhaps computationally useful to push the boundaries of the strain region we are considering. By looking at large strains, we can fit even the high-order K coefficients and, hopefully, achieving a high prediction accuracy in the moderate strain region that we are ultimately interested in. Since we are working with large strain, the prediction task is harder, and the accuracy is lower than in the small-strain case. The $n = 6$ model is found to be the most accurate, with a relative error of about $10^{-1.44} \approx 3.6\%$.

It is very curious to note that for the large-strain dataset, as the model degree changes from 0 to 4, the relative error decreases in an approximately exponential way, as can be seen from Figure 3.1. The trend then slows down and reverses as n becomes too large, around $n = 5$. The exponential scaling matches our expectation. Indeed, if terms of degree up to n are included in the truncated Taylor series, we expect the discrepancy between the truncated series and the true function to scale as the $(n+1)$-st power of the strain, since it is the terms of degree $n + 1$ that are the lowest-degree omitted terms. The discrepancy between the model and the true function can be expected to be roughly the n-th power of the average magnitude of strain. At the same time, we must note that while the average magnitude of strain for the large-strain dataset was 4.8%, the improvement in relative error between the $n = 2$ and $n = 3$ models is about $10^{-0.4} \approx 0.4$, significantly less than what we would have expected based on the suggestion outlined in this paragraph.

3.7 Relaxation under biaxial strain

Biaxial strain is a situation where a crystal is forced to conform to specified lattice parameters within a given plane but left free to relax the out-of-plane direction. In Subsection 2.5.2, we

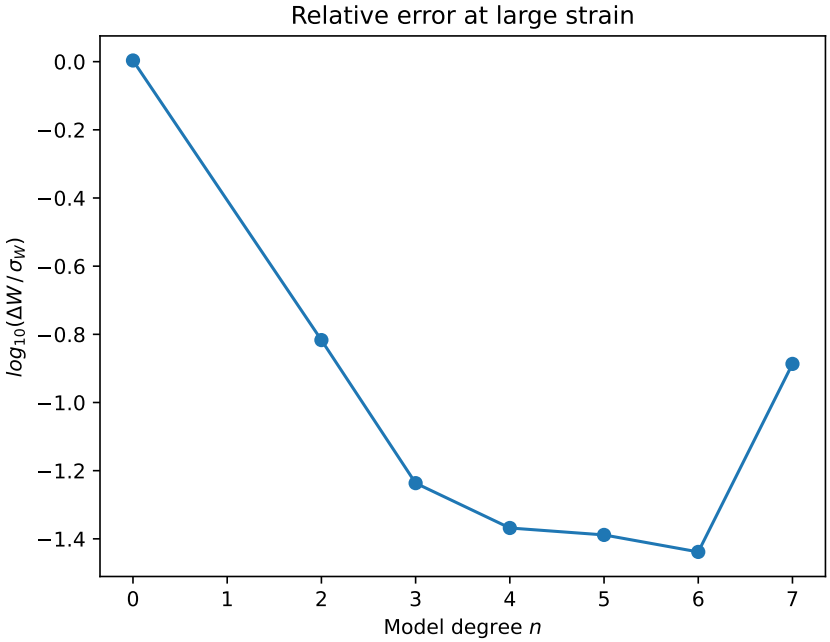


Figure 3.1: The relative error of the polynomial model for the energy density fitted to the large-strain dataset decreases with n approximately exponentially as n changes from 0 to 4. The exponential trend slows down around $n = 4$ and reverses as n goes from 6 to 7.

saw that the possible values of the deformation gradient F can be parametrized using three degrees of freedom $\mathbf{r} = (r_x, r_y, r_z) \in \mathbb{R}^3$ as

$$F = P^T \tilde{F} P + (\mathbf{w} + \mathbf{r})\mathbf{w}^T,$$

where $\tilde{F}$ is a 2×2 matrix that specifies the in-plane deformation, $\mathbf{w}$ is a normal vector to the plane of straining, P is a 2×3 matrix that "projects to the plane of straining," and the free parameter $\mathbf{r} \in \mathbb{R}^3$ decides in what way the out-of-plane direction is deformed, with $\mathbf{r} = 0$ corresponding to no deformation of the out-of-plane direction. Letting $F_0 = P^T \tilde{F} P + \mathbf{w}\,\mathbf{w}^T$ be the deformation gradient when the normal direction $\mathbf{w}$ is left unchanged, we have

$$F = F_0 + \mathbf{r}\,\mathbf{w}^T. \tag{3.24}$$

Recall also that in Section 3.6, we constructed a model for the energy density of RuO_2 as a function of the deformation gradient F. Combining these two results, we can predict the energy density of RuO_2 at a given value of $\mathbf{r}$,

$$\widehat{W}_{\text{biax}}(\mathbf{r}). \tag{3.25}$$

To determine what value of $\mathbf{r}$ corresponds to the relaxed lattice parameters of RuO_2 under biaxial strain, it remains to minimize $\widehat{W}_{\text{biax}}(\mathbf{r})$. Indeed, the lattice parameters that RuO_2

settles into are those that minimize elastic energy under the given biaxial strain. The value of $\mathbf{r}$ corresponding to the crystal's relaxed lattice parameters under biaxial strain is one that minimizes the energy density,

$$\mathbf{r}_{\text{biax-relaxed}} = \text{argmin}_{\mathbf{r}}\, \widehat{W}_{\text{biax}}(\mathbf{r}), \tag{3.26}$$

and the deformation gradient is given by $F_{\text{biax-relaxed}} = F_0 + \mathbf{r}_{\text{biax-relaxed}}\, \mathbf{w}^T$. The relaxed Bravais matrix $B_{\text{biax-relaxed}}$ is obtained from the bulk (unstrained) Bravais matrix B as

$$B_{\text{biax-relaxed}} = F_{\text{biax-relaxed}}\, B. \tag{3.27}$$

Recall from Subsection 2.5.2 that $\mathbf{r} = 0$ corresponds to no deformation of the out-of-plane direction. When minimizing $\widehat{W}_{\text{biax}}(\mathbf{r})$, we must take care to keep $\mathbf{r}$ small. Indeed, the elastic energy model of Section 3.6 contains cubic and higher terms, and these may lead to unphysical minima of energy for large strain. A quadratic model would be much easier to minimize, since such a model would have a unique minimum, and this minimum, if not exactly equal to the true physical minimum, would probably be pretty close to it. Therefore, our procedure for minimizing $\widehat{W}_{\text{biax}}(\mathbf{r})$ is as follows. Firstly, we minimize $\widehat{W}_{\text{biax}}(\mathbf{r})$ as calculated from equation 3.23 with only the quadratic terms included in the Taylor series. This gives us an approximate estimate $\mathbf{r}_{\text{quad}}$ for the location of the energy minimum, the value one would obtain using the linear theory of elasticity [29]. Secondly, using a standard numerical minimization algorithm from the Optuna Python library [32], we search for the energy minimum of $\widehat{W}_{\text{biax}}(\mathbf{r})$, this time with all terms included in equation 3.23, near the approximate minimum $\mathbf{r}_{\text{quad}}$. The second step provides small corrections on top of the approximate estimate of step 1.

3.8 Example: biaxial strain in the xy plane

Can we test the model-based relaxation algorithm against accurate DFT predictions? In this section, we consider a range of biaxial strains in the xy plane.

Recall that a unit cell of RuO_2 is rectangular with side lengths a, a, c for some a, c. The coordinate axes are chosen to coincide with the directions of the unit cell sides. We now pick a small real number ε and apply strain ε in the xy plane, letting the z direction relax. One way to phrase it is that we force a to change by a factor of $1+\varepsilon$ while leaving the z direction free,

$$a \longmapsto (1+\varepsilon)\, a, \quad c \longmapsto c' = ?. \tag{3.28}$$

(By symmetry, we know that the z direction will not get tilted into the x or y directions, so the strain will be described by a rescaling of a and c.) Another way to quantify the biaxial strain is that equation 2.13 becomes

$$F = \begin{pmatrix} 1+\varepsilon & 0 & 0 \\ 0 & 1+\varepsilon & 0 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 0 & r_x \\ 0 & 0 & r_y \\ 0 & 0 & r_z \end{pmatrix}. \tag{3.29}$$

The images under F of $(1,0,0)$ and $(0,1,0)$ are fixed, while the image of the out-of-plane vector $(0,0,1)$ is allowed to be anything because the parameter $\mathbf{r} = (r_x, r_y, r_z)$ can take any value. By symmetry, it is clear that $r_x = r_y = 0$; but let us not postulate that explicitly! We will run the procedure of Section 3.7 in all generality and see that the expected symmetry is, indeed, observed. We introduce a number η,

$$\eta = c'/c - 1 = F_{3,3} - 1$$

to quantify how much the crystal "chooses" to expand in the z direction as a result of xy biaxial strain ε. We wish to predict η as a function of ε.

Equation 3.28 puts the biaxial problem in a way amenable to direct DFT solution via Quantum ESPRESSO, while equation 3.29 puts it in a way amenable to solution via equation 3.26, e.g., using the elastic energy model we develop in Chapter 2. We can thus benchmark our model-based relaxation algorithm against accurate DFT results.

DFT calculation. For 13 data points as ε ranges from -6% to 6% in steps of 1%, we run the `vc-relax` calculation in Quantum ESPRESSO [22] [23] [24] with specifications `nat=6` and `cell_dofree='z'`. The first tells Quantum ESPRESSO to look for a rutile lattice with some parameters a', c', while the second instructs it to find a lattice parameter c' that minimizes the energy density while a' is fixed at $a' = (1+\varepsilon)\,a$. The atoms are allowed to move within the crystal cell to find a configuration of least energy. This way, for each ε, we obtain the value of c' that minimizes the elastic energy density.

Model calculation. As ε ranges from -6% to 6%, we calculate the relaxed deformation gradient F according to the procedure described in Section 3.7. Since the z direction is not tilted into the x or y directions, we expect F to be diagonal with $1+\varepsilon$, $1+\varepsilon$, and $1+\eta$ on the diagonal, and we extract η as $\eta = F_{33} - 1$. For example, for $\varepsilon = 3\%$, the deformation gradient is computed to be

$$F = \begin{pmatrix} 1.030 & 0.000 & -2\times10^{-4} \\ 0.000 & 1.030 & -6\times10^{-4} \\ 0.000 & 0.000 & 0.973 \end{pmatrix},$$

and we calculate $\eta = -2.7\%$. Small off-diagonal terms on the order of 0.05% are discarded. The values of ε for which to perform this computation are chosen from -6% to 6% finely to approximate a continuous range.

The comparison of the model to DFT is shown in Figure 3.2. The DFT values of η (black crosses) are plotted together with the model's prediction (blue solid line). We also plot the values of η one gets from linear elasticity theory, e.g., using $\mathbf{r}_{\text{quad}}$ instead of $\mathbf{r}_{\text{biax-relaxed}}$ when calculating F (dashed gray line). The DFT-calculated η behaves very linearly for ε between -6% and 3%, but then it suddenly changes course and starts going down at a considerably faster rate. The linear approximation for η lags behind the DFT-calculated value for large positive ε, but the model successfully predicts the nonlinearity and agrees with the DFT result even at large ε.

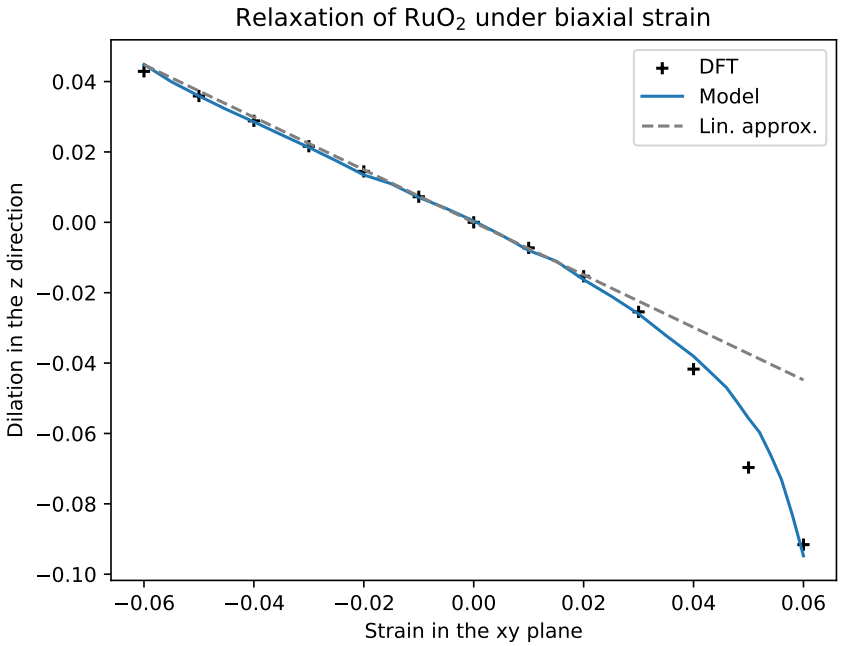


Figure 3.2: The model successfully predicts the z-direction relaxation of RuO_2 under xy plane biaxial strain. The model stays accurate even as the process starts displaying nonlinear behavior at large strain, $\varepsilon \approx 5\,\%$, possibly indicative of a phase transition.

Chapter 4

Electronic density of states

4.1 Overview

We are interested in understanding the electronic structure of materials under strain. We are considering small strain, around 5% or smaller in magnitude. And a rather common maxim for working with small quantities is,

Taylor-expand in small quantities.

We will use an especially simple variant of this approach, expanding to first order. That is, suppose $\epsilon_i(E;\kappa)$ is the energy of an electron with relative wavenumber κ (to be introduced in Section 4.2) in the i-th electronic band in a given crystal (say, RuO_2) subject to strain E. As discussed in Chapter 3, since E is a 3×3 symmetric matrix, it can be parametrized by six real numbers $c_1, \ldots, c_6$ (eq. 3.9). For small strains, the parameters $c_1, \ldots, c_6$ are also small, and we write the strained electronic energy as the unstrained one plus a first-order correction,

$$\epsilon_i(E;\kappa) \approx \epsilon_i(0;\kappa) + \sum_{a=1}^{6} \frac{\partial \epsilon_i(\kappa)}{\partial c_a} c_a. \tag{4.1}$$

This kind of first-order expansion is the subject of Deformation Potential Theory [33]. The coefficients $\frac{\partial \epsilon_i(\kappa)}{\partial c_a}$ are called *deformation potentials*, and they determine the response of the electronic structure to strain.

The bulk (relaxed) electronic structure and the deformation potentials are easy to compute using Density Functional Theory (DFT). To do that, we calculate the electronic energy bands for some number of strain parameters and fit a linear model through the data points. While the response of the energy bands to strain is not exactly linear, we find that the linear model is accurate to within about 10%, and oftentimes more accurate. This makes intuitive sense – straining the crystal introduces a small perturbation the effective Hamiltonian experienced by the electrons, and first-order perturbation theory [34] suggests that small perturbations in the Hamiltonian lead to small perturbations in the energy spectrum. The quadratic terms in the Taylor expansion seem to be small compared to first-order terms.

The dependence of the Fermi level on the strain is less linear. In Figure 4.1, one can see that as RuO_2 is strained in a certain way with increasing magnitude, the Fermi level changes in a rather nonlinear way. This is to be expected, since the Fermi level is determined by trying to fill in some specified number of electronic states. While the energies of the electronic bands are approximately linear in the strain, the Fermi level depends in a complicated way on the energy bands.

The density of states at the Fermi level depends in a complicated and nonlinear way on the strain. Firstly, the Fermi level depends on the strain nonlinearly. Secondly, the density of states is calculated from the energy bands by counting how many electronic states have an energy near the Fermi level – this is a very nonlinear definition.

4.2 Relative wavenumber

Consider a RuO_2 crystal subject to strain E. Since the lattice basis vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ change in response to strain, so do the reciprocal lattice basis vectors $\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3$. The wavenumber $\mathbf{k}$ is then not a convenient number to study because the Brillouin zone changes in response to strain. To hide the strain dependence, we define the *relative wavevector* κ, which is a 3-component vector $\kappa = (\kappa^1, \kappa^2, \kappa^3)$, in such a way that relative wavevector κ represents the "usual" wavevector

$$\mathbf{k} = \kappa^1 \, \mathbf{b}_1 + \kappa^2 \, \mathbf{b}_2 + \kappa^3 \, \mathbf{b}_3. \tag{4.2}$$

For example, the M and Z points in the Brillouin zone always correspond to $\kappa = (1/2, 1/2, 0)$ and $\kappa = (0, 0, 1/2)$, respectively. As we fix κ and vary strain, we can think of it as "studying the same place in the Brillouin zone" as the strain varies and changes the Brillouin zone with it. A big practical reason for the introduction of the relative wavenumber is to facilitate the analysis of data obtained from DFT calculations – we discuss it in the following section.

4.3 A model for the electronic bands as a function of strain

As cell parameters change under small strain, the effective Hamiltonian experienced by the electrons in RuO_2 changes by a small perturbation. In the spirit of first-order perturbation theory in quantum mechanics, we expect that $\varepsilon_i(E; \kappa)$ responds to this perturbation in an approximately linear way. In other words, we expect that

$$\boxed{\epsilon_i(E; \kappa) \approx \epsilon_i(0; \kappa) + \sum_{a=1}^{6} \frac{\partial \epsilon_i(\kappa)}{\partial c_a} \, c_a} \tag{4.3}$$

for some response coefficients $\frac{\partial \epsilon_i(\kappa)}{\partial c_a}$, and where the *strain numbers* $c_a = \frac{1}{2} \operatorname{tr}\big(\tau_a^T E\big)$, introduced in Section 3.4, provide a convenient way to parametrize strain matrices. (Denoting the response coefficients by $\frac{\partial \epsilon_i(\kappa)}{\partial c_a}$ is a bit of an abuse of notation, since we are not considering a

function $\epsilon_i(\kappa)$.) It is interesting to note that the response coefficients $\frac{\partial \epsilon_i(\kappa)}{\partial c_a}$ have the proper name of *deformation potentials*, and the study of various properties of such potentials is referred to as Deformation Potential Theory [33]. The simple equation 4.3 is boxed because, in a sense, it is the driving idea behind the present chapter and perhaps even the whole thesis.

Let us now return to computational matters. Using Density Functional Theory, we calculate the electronic bands energies for $p = 69$ values of strain $E_1, \dots, E_p$. The Brillouin zone is covered by a $16 \times 16 \times 24$ mesh, and the relative wavenumber κ runs over $16^2 \cdot 24 = 6144$ values that are independent of strain. In our calculations, there are $n = 34$ electronic bands. It turns out that for the j-th strain strain E_j, the calculation yields the electronic energies as a list

$$(\kappa_1, (\epsilon_1(E_j; \kappa_1), \dots, \epsilon_n(E_j; \kappa_1))), \dots, (\kappa_{6144}, (\epsilon_1(E_j; \kappa_{6144}), \dots, \epsilon_n(E_j; \kappa_{6144}))),$$

where $\epsilon_i(E; \kappa_l)$ is the energy of an electron in the i-th band with wavevector κ_l when the crystal is strained by the amount E. Importantly, $\epsilon_i(E_j; \kappa_l)$ is the energy with respect to some *arbitrary zero of energy* used by Quantum ESPRESSO [22] [23] [24]; we will later discuss the Fermi level and how to measure energies with respect to it. We rearrange the electronic energies data into a list

$$\begin{aligned} &(\text{band}_1, \dots, \text{band}_n), \\ &\text{band}_i = (\dots, (E_j, \epsilon_i(E_j)), \dots) \end{aligned}$$

so that all the information for the i-th band is collected into one variable band_n. Here, by $\epsilon_i(E_j; *)$ we denote the 6144-component array

$$\epsilon_i(E_j; *) = (\epsilon_i(E_j; \kappa_1), \dots, \epsilon_i(E_j; \kappa_{6144})),$$

that gives the energy value of the i-th band at the j-th strain for the 6144 grid values of κ. Note that the entries of $\epsilon_i(E_j; *)$ are ordered, so that the l-th entry gives the energy at the l-th relative wavenumber κ_l.

Regarding $\epsilon_i(E_j; *)$ as a 6144-dimensional column vector, we write equation 4.3 as

$$\epsilon_i(E; *) \approx C_i \, x(E), \tag{4.4}$$

where C_i is a 6144×7 matrix encoding a linear model for the i-th band's response to strain, and where $x(E) = (1, c_1, \dots, c_6)$ is a 7-dimensional column vector, encoding the strain E, that consists of the strain numbers $c_1, \dots, c_6$ with an extra 1 on the left. The first column of C corresponds to $\varepsilon_i(E; \kappa)$ in equation 4.3, while the other six columns encode the deformation potentials $\frac{\partial \epsilon_i(\kappa)}{\partial c_a}$.

For a given band, the approximate equality (eq. 4.4) should hold for all strain values – we wish to pick C_i such that this is true. It is a standard procedure to use the method of least-squares to find a value for C_i that minimizes the total error-squared

$$\sum_{j=1}^{p} ||C_i x_j - \epsilon_i(E_j; *)||^2.$$

The matrix C_i encodes our linear model of the i-th band's response to strain. The first column of C_i encodes the electronic energies at zero strain, while the other six columns quantify the band's response to strain.

For a given strain E, evaluating equation 4.4 is almost instantaneous because it just requires applying the matrix C_i to the vector $x(E)$. The linear model outlined in this section is also not demanding in terms of memory. Indeed, there are $n = 34$ matrices C_i (one matrix for one electronic band), and each matrix contains $7 \cdot 6144 = 43008$ floating-point numbers. The entire data fits within a few megabytes of space, well within the dynamic memory capabilities of a personal computer. The absolute and relative error of the model for different electronic bands across the training data is given in Table 4.1.

4.4 A model for the Fermi level

For the j-th strain E_j, our DFT calculation tells us the Fermi level and the electronic energies. It is a straightforward exercise to count the number $N(E_j)$ of electronic states at or below the Fermi level. For the $p = 69$ calculations we use for our analysis, the numbers $N(E_j)$ are given below.

$$
\begin{gathered}
172009, 172091, 172075, 172071, 172053, 172005, 172069, 172027, 172080, 172094, \\
172078, 172045, 172043, 172021, 172102, 172072, 172015, 172077, 172027, 172111, \\
172029, 172061, 172015, 172060, 172066, 172027, 172055, 172073, 172101, 172045, \\
172057, 172097, 172081, 172090, 172056, 172065, 172071, 172061, 172083, 172048, \\
172055, 172075, 172037, 172037, 172023, 172044, 172058, 172113, 172058, 172063, \\
172117, 172022, 172077, 172119, 172079, 172035, 172095, 171999, 172145, 172109, \\
172063, 172051, 172053, 172097, 172079, 172032, 172043, 172031, 172045.
\end{gathered}
$$

While the strains E_j go up to about 5% in magnitude, the numbers $N(E_j)$ are all within 0.1% of 172062. Recall that the ground state of RuO_2 is one that has the smallest energy. By the Pauli exclusion principle, two electrons cannot occupy the same state. Therefore, if there are N electrons that must be accommodated, the N lowest electronic states are occupied. The energy of the highest occupied state is the Fermi level. Since the number of electrons in RuO_2 does not change under strain, the Fermi level varies with the strain such that the number of electronic states with energy below it is fixed. We conclude that regardless of strain, the first

$$N = 172062 \approx 6144 \times 28$$

electronic states are filled. The slight variation of N from one calculation to another is likely due to a smearing procedure used by Quantum ESPRESSO, wherein the electronic density of states induced by a single electronic state is said to be not a Dirac delta function but a smoothed Gaussian-like distribution. The fact that $N \approx 6144 \times 28$ matches what one would expect theoretically. Indeed, the pseudopotential files we use include 6 electrons per an oxygen atom and 12 electrons per a ruthenium atom. The unit cell of RuO_2 that we

Band num.	$\log_{10}$(rel. err)	Absolute error, eV
1	−1.684	0.0089
2	−1.629	0.0101
3	−1.344	0.0187
4	−1.272	0.0219
5	−1.529	0.0122
6	−1.472	0.0140
7	−1.597	0.0107
8	−1.555	0.0117
9	−0.943	0.0358
10	−0.996	0.0322
11	−1.164	0.0236
12	−1.036	0.0321
13	−0.926	0.0339
14	−0.922	0.0338
15	−0.885	0.0373
16	−0.860	0.0399
17	−1.103	0.0246
18	−1.244	0.0187
19	−1.255	0.0189
20	−1.193	0.0226
21	−1.149	0.0252
22	−1.121	0.0272
23	−1.270	0.0200
24	−1.311	0.0183
25	−1.373	0.0180
26	−1.333	0.0203
27	−1.368	0.0192
28	−1.321	0.0218
29	−1.458	0.0164
30	−1.458	0.0170
31	−1.383	0.0241
32	−1.276	0.0306
33	−1.296	0.0299
34	−1.362	0.0263

Table 4.1: The relative and absolute error of the electronic bands model for different bands.

are using for our calculation has two formula units of the molecule. Therefore, there are $2 \times (12 + 6 + 6) = 56$ electrons to accommodate. With two electrons per band (which we can see from the Pauli exclusion principle and the fact that electrons have two spin states), this means that 28 bands will be filled.

For an arbitrary strain E, the Fermi level can be predicted by calculating the energies of all available electronic states and filling the first N of them. Specifically, the procedure is as follows. For a given value of strain, and for each of the 34 bands, equation 4.4 yields a 6144-component array representing the electronic energies $\epsilon_i(E; *)$ of the band at each of the 6144 sampled values of the relative wavenumber κ. Let us combine the energies of the bands into one $6144 \cdot 34$-component array, and let us sort this array in increasing order. This is the list of all electronic energies from all bands. Taking the N-th element of this array to be the Fermi level, we ensure that N electronic energies are below the Fermi level.

4.5 A model for the electronic density of states

Suppose we have fit a linear model to the electronic energy bands as in Section 4.3. By the results of Section 4.4, we can also predict the Fermi level as a function of strain. For an arbitrary strain matrix E (one that does not have to lie in our training data), the electronic density of states (DOS), with energies measured with respect to the Fermi level, can be predicted as follows.

1. Use equation 4.4 to compute the electronic energies $\epsilon_i(E; *)$.
2. Compute the Fermi level $\varepsilon^{(F)}$ by filling in the lowest N states, as in Section 4.4.
3. Subtract the Fermi level from the electronic energies, so that energy is measured with respect to the Fermi level. Count the number of electronic energies near different energy values (make a histogram of the electronic energies) to obtain the electronic density of states.

None of the steps outlined above involve Density Functional Theory (DFT) calculations. Step 1 involves the evaluation of a linear model, while steps 2 and 3 involve simple counting. As a result, the procedure outlined above for the approximate computation of the electronic density of states (DOS) takes minimal computational resources to execute.

4.6 Benchmarking the model

We benchmark to compare our approximate model to a full DFT calculation by considering the following scenario. Suppose RuO_2 is strained with the following strain matrix,

$$E = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \alpha \end{pmatrix},$$

where α goes over some range of values. That is, while RuO_2 normally has lattice parameters a, c with $a = 4.492\mathring{A}$ and $b = 3.106\mathring{A}$, we keep a the same while changing c to $(1 + \alpha)c$. We let α go from -7% to 7% in steps of 1%. What will be the electronic energy bands, the Fermi level, and the electronic density of states for the different strains? We compare our model's prediction against full DFT calculations.

Recall from Section 4.4 that we predict the Fermi level by firstly predicting the electronic energy bands with respect to the arbitrary zero of energy used by Quantum ESPRESSO, and secondly filling up the first $N = 6144 \times 28$ lowest energy levels. In Figure 4.1, the model's prediction for the Fermi level is compared with the Fermi level as provided directly by the DFT calculation. One can see that the model correctly follows the variation of the Fermi level with strain.

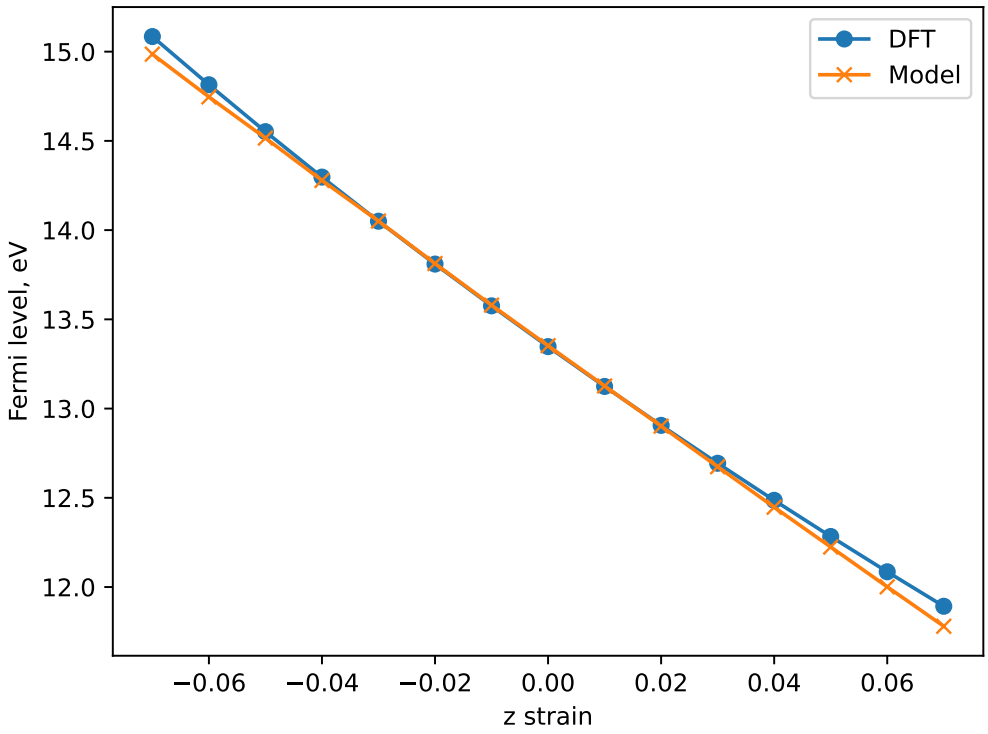


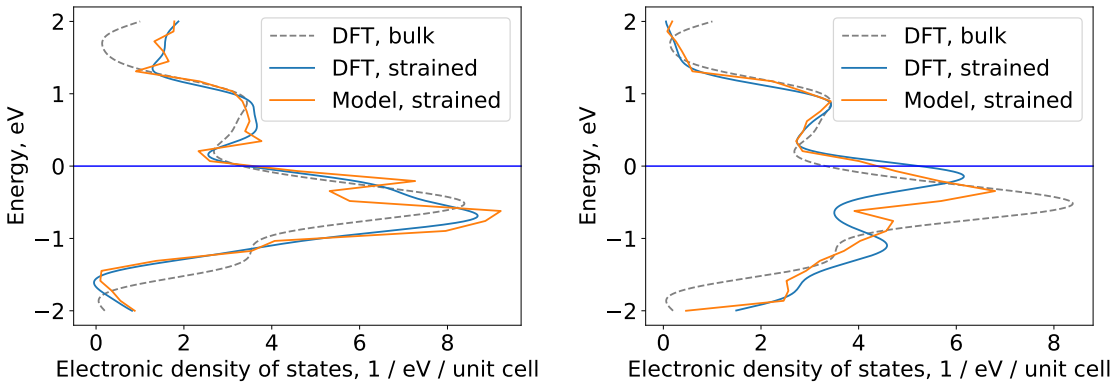
Figure 4.1: The Fermi level as a function of strain in the z direction, computed using the approximate model and using DFT. The approximate model follows the DFT result.

Figure 4.2(a) compares the model's prediction for the electronic density of states (DOS) to the result of a DFT calculation, for $\alpha = 0.07$. The density of states is in units of states per eV per unit cell. While the DOS is not changed much near the Fermi level, the function does change around 1.7 and -1.5 eV, and the model's prediction follows the DFT result in these changes. Our prediction for the DOS comes out to be less smooth than the DFT result.

Figure 4.2(b) makes the comparison when $\alpha = -0.07$; the model's prediction follows the decrease of the peak DOS observed in the DFT result, although the model underestimates the drop in DOS to some extent.

It is curious to note how different Figures 4.2(a) and 4.2(b) are, despite the fact that the strain on the first is negative the strain on the second. Somehow $\alpha = 0.07$ keeps the density of states at the peak around $-0.5\,\text{eV}$ largely intact while changing the DOS at the fringes $\pm 1.5\,\text{eV}$, while $\alpha = -0.07$ keeps the DOS relatively stable at the fringes while decreasing the density of states at the peak at $-0.5\,\text{eV}$. This suggests that the density of states is not well-approximated as a linear function of the strain.

Our model for the density of states, while based on a linear model for the electronic energy bands, captures the qualitative nonlinearity feature that positive α keeps the DOS at the peak relatively unchanged while negative α decreases the peak DOS value. The nonlinearity in our model for the DOS comes from the fact that while the bands are predicted linearly, the counting of the number of electronic states near various energy values is a nonlinear process.



(a) Strain with $\alpha = 7\%$.

(b) Strain with $\alpha = -7\%$.

Figure 4.2: The model's prediction largely follows the DFT-calculated density of states as it changes in response to strain, even for relatively large strains and even when the DOS response to strain becomes nonlinear.

Chapter 5

Optimizing biaxial strain for superconductivity

What strain should we impose on RuO_2 to achieve the greatest extent of strain-stabilized superconductivity? This is the question we look at in this chapter, using the machinery developed in previous chapters to search the space of possible strains. In Section 5.1, we find a strain that optimizes the Fermi level DOS, as predicted by our model. The search is over generic small strains, without taking into account experimental accessibility. Luckily, the optimized strain seems to be close to a biaxial strain one can potentially implement using biaxial epitaxy. We investigate this in Section 5.2 and find that biaxial tensile strain in the xy plane (the (001) crystal plane) seems to be a promising candidate for strain-induced superconductivity.

5.1 Optimizing Fermi level DOS

If we could enact any small strain in RuO_2, what value would we choose to produce the greatest Fermi level DOS?

Upon creating the approximate linear model for the electronic bands energies (Section 4.3), we can evaluate the electronic energies for a given strain at virtually no cost. From the electronic energies, the density of states is computed by counting the number of electronic states at various energy levels – this can also be done quickly. This way, we obtain a model for the density of states as a function of strain. In particular, we can compute the density of states at the Fermi level. Using standard numerical optimization techniques, we can find strain values that maximize the density of states at the Fermi level.

Firstly, we read the file `Cmat.json` produced by our code. As described in Section 4.4, for a given value of strain, this gives us the ability to approximately calculate the Fermi level density of states (DOS) at any given strain. Using the Optuna [32] library in Python, we run a standard maximization procedure to maximize the Fermi level density of states. In the electron-phonon picture of superconductivity [15] [16], a high Fermi level DOS is favorable to superconductivity. We will consider phonons later.

An important consideration in optimizing Fermi level DOS is that strains should be kept small, within a magnitude of about 5%. We enforce this condition as follows. We take as our target function not the density of states at the Fermi level, but the density of states at the Fermi level divided by the *penalty function*

$$f(E) = \exp\left(50 \sum_{i=1}^{3} \max\{0, |\lambda_i| - 0.05\}\right), \tag{5.1}$$

where $\lambda_1, \lambda_2, \lambda_3$ are the eigenvalues of the strain matrix E. The value of $0.05 = 5\%$ was chosen to make sure the strains don't exceed 5%, while the constant of 50 in the exponential, which controls how steep the penalty function is, was chosen arbitrarily. When all the eigenvalues are less than 0.05 in magnitude, $f(E) = 1$, meaning that no penalty is imposed. However, if the condition is not satisfied, the function $f(E)$ quickly grows. Using this penalty function, we prevent strains being greater than about 5% in any given direction.

Given that the Fermi level density of states is a rather nonlinear function of the lattice parameters, one might wonder whether there are several local maxima. We take the approach of maximizing the Fermi level density of states using a standard Optuna optimizer, searching the space of strains that do not exceed the maximum allowed magnitude. The optimizer first samples the space of strains randomly and then proceeds to fine-tune starting from the randomly obtained points. Our assumption is that the Fermi level DOS is a nonlinear function, sampling sufficiently many points would allow us to find the global maximum. We run the optimizer with 10000 samples. The fact that the density of states at the Fermi level can be quickly estimated as a function of the strain allows numerical optimization to run within a few minutes.

The result of the model-based search are as follows. We expect that the Fermi level DOS can be raised to 5.513 states per eV per unit cell, a 66% increase from the bulk value of 3.322. The strained Bravais matrix is

$$B_{\text{optimized}} = \begin{pmatrix} 8.818 & 0.001 & 0.063 \\ 0.001 & 8.834 & 0.068 \\ 0.044 & 0.047 & 5.575 \end{pmatrix}, \tag{5.2}$$

in units of Bohr. The corresponding strain numbers are

$$(c_1, \ldots, c_6) = (4.63\%,\ -3.27\%,\ -0.09\%,\ 0.01\%,\ 0.81\%,\ 0.87\%).$$

The strain matrix is

$$E_{\text{optimized}} = 0.01\ \times \begin{pmatrix} 4.54 & 0.01 & 0.81 \\ 0.01 & 4.72 & 0.87 \\ 0.81 & 0.87 & -4.62 \end{pmatrix} \tag{5.3}$$

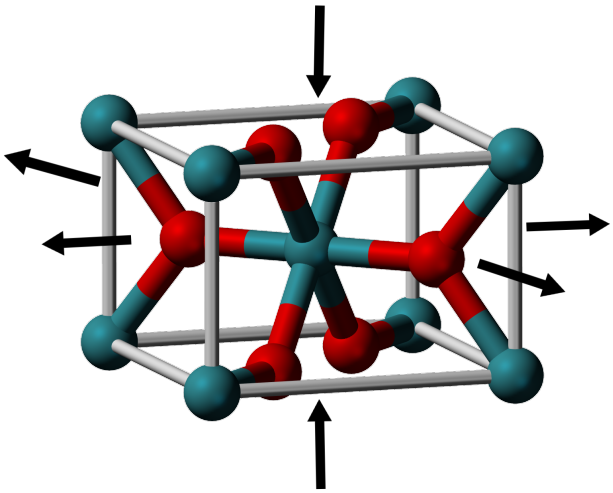


Figure 5.1: A pictorial illustration of the strain that is predicted to induce the greatest Fermi level DOS increase in RuO_2. The xy plane is uniformly expanded, while the z direction is contracted.

and has eigenvalues and eigenvectors

$$\lambda_1 = -4.78\%, \quad \mathbf{v}_1 = (-0.09, -0.09, 0.99), \tag{5.4}$$
$$\lambda_2 = 4.57\%, \quad \mathbf{v}_2 = (0.93, -0.36, 0.05), \tag{5.5}$$
$$\lambda_3 = 4.84\%, \quad \mathbf{v}_3 = (-0.35, -0.93, -0.12). \tag{5.6}$$
$$\tag{5.7}$$

A pictorial illustration of the optimized strain is given in Figure 5.1. It is curious to note that the strain matrix E is almost diagonal, and that the E_{11} and E_{22} components are almost equal. Moreover, the signs of E_{11} and E_{33} tell us that we are dealing with xy expansion and z contraction. We will revisit these properties in Section 5.2.

We now investigate the model-predicted optimized strain using DFT. With the lattice parameters specified by equation 5.2, we run a DFT computation of the electronic density of states. Figure 5.2 provides the DFT-calculated DOS, as well as the model-calculated DOS. A high degree of agreement can be seen. DFT predicts that the Fermi level DOS increases to 5.883 states per eV per unit cell under the optimized strain (eq. 5.3), an increase of 77% from the bulk value of 3.322. This supports the model-based prediction that the Fermi level DOS is significantly increased under the strain of equation 5.3.

5.2 Biaxial tensile strain in the xy plane

Maximizing the Fermi level DOS led us to a strain matrix E that is almost diagonal (eq. 5.3), meaning that the x, y, z directions are almost not mixed between each other. The xy plane is almost uniformly expanded by about 4.6%, while the z direction is compressed by

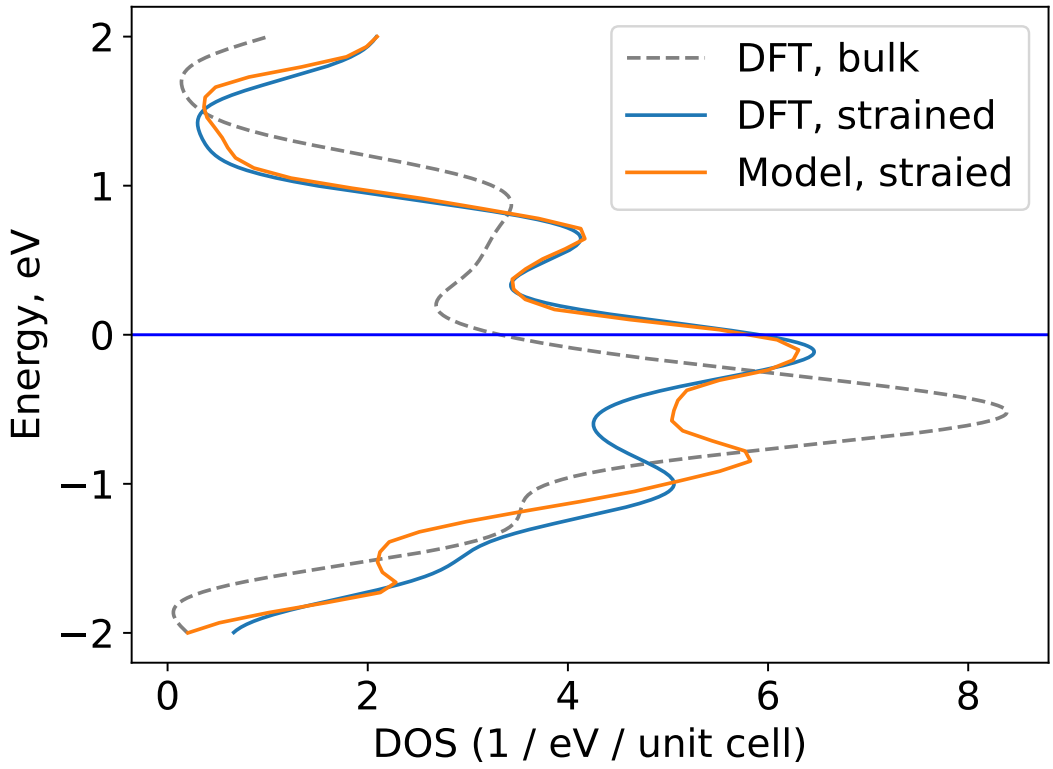


Figure 5.2: Strain brings the peak in the electronic density of states (DOS) to the Fermi level, which is set to 0 eV. The strain was found using an approximate linear model. The model's prediction is supported by a direct DFT computation.

about -4.6%. The absence of shear is a nice simplification. The fact that xy expansion is combined with z contraction is also nice. If RuO_2 were forced to expand in the xy plane, the z direction would naturally "want" to contract. This prompts us to consider biaxial strain in the xy plane, a situation when RuO_2 is forced to conform to certain lattice parameters in the xy plane while the z direction is left free. Specifically, we consider a scenario where RuO_2 is forced to expand or contract uniformly in the xy plane with the z direction left free, the situation described in Section 3.8.

Using the Quantum ESPRESSO [22] [23] [24] DFT package, specifically by running the `vc-relax` and `nscf` calculations, we predict the Fermi level DOS for a range of xy strains. The results are demonstrated in Figure 5.3. At around 4.5%, the Fermi level density of states increases from its bulk value of 3.32 states per eV per unit cell to 5.87, an increase of as much as 77%[1]. These numbers are obtained not using our model but using Density Functional Theory.

[1]The 77% Fermi level DOS increase happens to coincide with the 77% increase from the optimized strain, although the two strains are slightly different.

Phonon modes appear to be softened by biaxial tensile strain in the xy plane. (Figure 5.3.) This is only a tentative possibility at this point, since phonon calculations are complex and require careful convergence analysis to validate. If correct, however, this calculation would position tensile biaxial strain in the xy plane as a strong candidate for strain-induced superconductivity, since an increased in the Fermi level DOS a softening of phonon modes are a positive factor for superconductivity in the electron-phonon picture of superconductivity [15] [16].

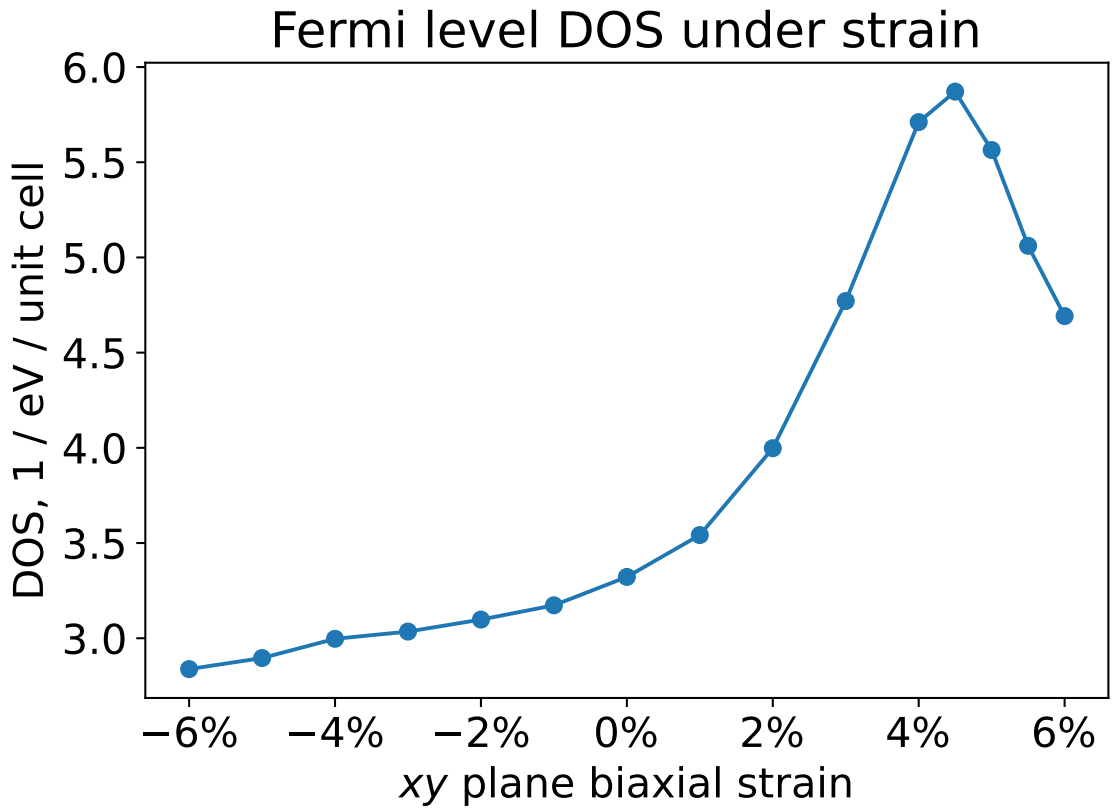


Figure 5.3: Fermi level density of states under biaxial strain in the xy plane (the (001) direction), predicted using DFT. Under 4.5% tensile strain, the Fermi level DOS is predicted to change from 3.322 to 5.871 states per eV per unit cell, making an increase of 77%.

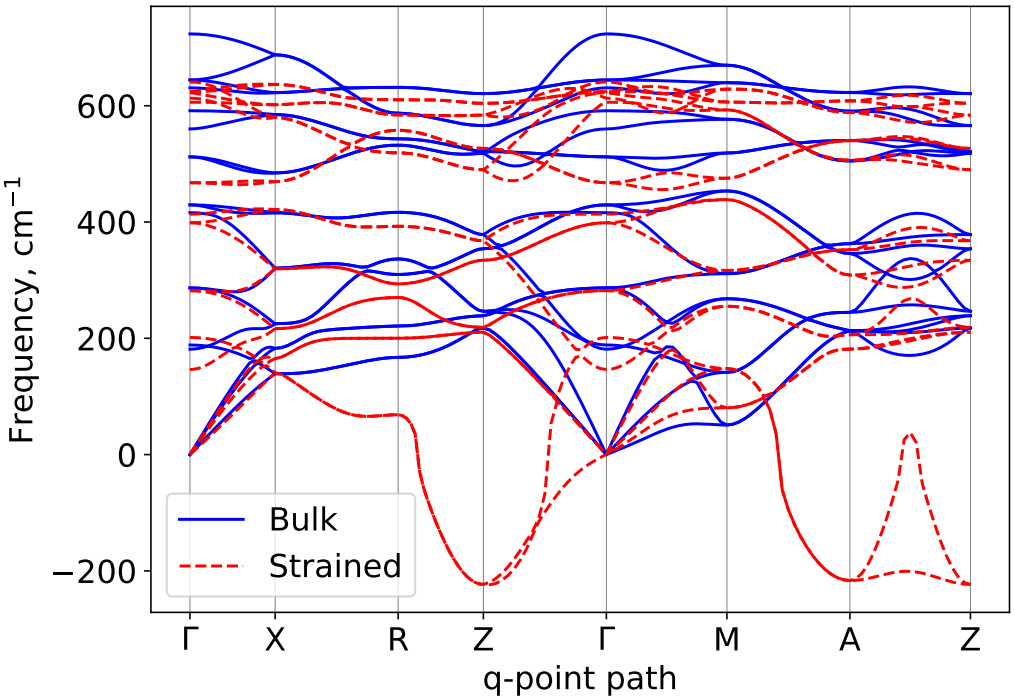


Figure 5.4: Tentatively, Density Functional Theory seems to suggest that 5% tensile strain in the xy plane leads to a softening of phonon modes. Careful convergence analysis must be performed to assess this possibility.

Chapter 6

Conclusion

Using surrogate models, we have systematically examined the space of possible strains and found biaxial strain in the xy plane a promising candidate for strain-induced superconductivity. This strain can be experimentally realized using biaxial epitaxy (for example, Molecular Beam Epitaxy [17] [18]) in the (001) crystal orientation.

We have focused on RuO_2, but the models can readily be trained on any other rutile materials (see Future Work below). The procedures would be the same for another rutile material, and one would need to run DFT calculations to obtain training data. The models could then be used to study these other rutile materials.

Our hope is that the methods developed in the present work can help accelerate materials discovery and shed more light on the understanding of the effects of strain.

6.1 Future work

- **Physical interpretation of model coefficients.** What physics can we extract from the surrogate models for the elastic energy and electronic band structure of RuO_2? What is the meaning of the magnitudes and signs of the coefficients $K_{n_1,\dots,n_6}$ (Section 3.6)? What can we say about the deformation potentials $\frac{\partial \epsilon_i(\kappa)}{\partial c_a}$ (Chapter 5)?

- **Materials beyond RuO_2.** Extending this work from RuO_2 to other rutile materials is a matter of calculations. For a given material, one would need to run the DFT and Python calculations in Section 3.6 to get a model for the elastic energy and those in Section 4.3 to get a model for the electronic bands. The setup of the calculations is the same, and what one needs to do is update the chemical data from RuO_2 to the new material in all Quantum ESPRESSO input files. With the models for the elastic and electronic properties, one can investigate superconductivity and other properties.

  Extending in non-rutile materials is slightly harder but certainly within reach. We would be working with a space group other than $P136$, and one would need to change the symmetry setup in the elastic model, Section 3.6. The matrices $\tau_1, \dots, \tau_6$ would generally not transform nicely under this new group, and one would need to find

another parametrization of strain to leverage symmetry. However, the general approach of Sections 3.5 and 3.6 seems to still apply – to decompose the energy density into a Taylor series and reduce the number of terms using symmetry, and to fit the coefficients from DFT data using least-squares. The analysis of Chapter 5 remains unchanged when one passes from materials with rutile symmetry to those with another symmetry, or even no symmetry at all. Combining the elastic and electronic bands models, we would be able to study superconductivity and other properties.

- **Application of machine learning techniques.** Linear models require only a small amount of data to fit and are, in many cases, very accurate. Combined with appropriate nonlinear functions, such as the counting of electronic energy levels we perform in Section 4.5, they can even be used to simulate nonlinear phenomena like the dependence of the electronic density of states on strain. And while linear models are linear in model parameters, they may be nonlinear in the input variables, such as the polynomial model for the elastic energy we develop in Chapter 3. With all the applicability of linear models, it is natural to ask if the simple learning techniques we have applied in this work would benefit from recent advances [35] [36] [37]. Human physicists speak of, e.g., how different electronic bands correspond to different orbitals. Electronic bands are then understood not simply as quantities, but also in terms of chemical insight about RuO_2. Perhaps language models would be able to pick out some structure in the electronic data that we have not incorporated into our linear models?

  Neural radiance fields (NeRF) [38] have recently been shown to perform luminosity calculations in computer graphics faster and with higher accuracy than "hard-coded" physics-based algorithms. Perhaps a similar result could be possible for electronic and phonon structure calculations? The success of neural radiance fields is quite impressive given how non-linear illumination problems are. Physics-informed neural networks (PINN) [39] are another inspiring example.

- **Modeling the displacement of atoms within the crystal cell.** As the cell parameters are changed in response to strain, ruthenium and oxygen atoms move within the cell to find a configuration of least energy. In which direction and how far would the atoms move in response to different kinds of strain? This question seems to be intimately connected to the physics of the material. The two ruthenium atoms, for example, always remain at relative positions $(0, 0, 0)$ and $(1/2, 1/2, 1/2)$ within the cell – due to symmetry, they remain at these positions under strain. What can one say about the ruthenium atoms? Is there a special way to decompose strain where every component of the decomposition leads to motions of the atoms that have a clear physical interpretations? The atoms only move within the crystal cell by a small amount, and it is a reasonable guess that their displacement is well-approximated as being linear in the strain. By looking at the coefficients of such a model, perhaps, interesting physics could be extracted.

- **A surrogate model for phonons.** The structure of vibrational modes plays a big role

in the theory of electron-phonon superconductors. We have created surrogate models for the elastic and electronic properties of RuO_2 but did not do so for phonons, opting instead to directly calculate the phonon structure using DFT when necessary. Part of the reason is that the electronic density of states varies strongly with strain and provides a pretty good indication of whether superconductivity is likely to occur. But another reason is technical, namely that phonon calculations using DFT are significantly more expensive than electronic calculations. To create a surrogate model for the phonon structure of RuO_2, one would need to find a way to produce a large dataset of phonon calculations. This calculation could be made feasible, for example, by the use of rutile symmetry. If successful, a surrogate model for phonons, together with the existing models for the elastic energy and electronic structure, could be used to make more complete predictions of the superconducting properties of RuO_2.

- **Using symmetry to speed up calculations.** For simplicity of DFT calculations and post-processing, we used a relatively crude approach where symmetry is disregarded. When computing the electronic band energies, for example, we instruct Quantum ESPRESSO to compute the band energies at every point of a mesh in the Brillouin zone; rutile symmetry could be used to reduce the number of k-points to be computed by up to 16 times, but to reap this benefit we would have to both design the calculations in a clever way and undertake some careful post-processing. Elastic and electronic calculations are relatively inexpensive, and we are able to get away with our simple approach to calculations – but it could be that for more demanding computations, like for phonons, taking advantage of symmetry would become necessary. It is a natural direction of future work to rework the DFT calculations in Section 3.6 in a way that takes advantage of rutile symmetry.

- **Efficiently learning the elastic energy using strain data provided by Quantum ESPRESSO.** When fitting a model to the elastic energy of RuO_2 as a function of strain in Chapter 3, we used as our initial data the elastic energy values at various strains, as calculated by Quantum ESPRESSO. But for each such calculation, Quantum ESPRESSO outputs not only the energy but also the stress tensor. While the energy is a single number, the stress tensor has 9 components – so the latter contains more information. Perhaps more vitally, the energy is approximately quadratic in the strain near zero, while the stress is approximately linear in the strain. Since we are considering small strain, this makes the stress tensor a more convenient quantity to fit a model for because the DFT data has more "contrast." Differentiating the energy density of a rutile material (equation 3.17) with respect to strain, one obtains a formula for the stress as a function of strain where the number of parameters $K_{n_1,\dots,n_6}$ is reduced using rutile symmetry. That is, the arguments of Chapter 3 that allowed us to reduce the number of unknown parameters using rutile symmetry still apply if we are interested in stress instead of the energy density. With this in mind, we can fit the coefficients $K_{n_1,\dots,n_6}$ using the stress tensor data provided by Quantum ESPRESSO. Perhaps such an approach would be more accurate and effective than the one we cur-

rently take.

Appendix A

Operations with matrices

A.1 Functions of matrices

Let D be a subset of the complex numbers, and suppose $f : D \to \mathbb{C}$ is a function. Suppose M is an $n \times n$ diagonalizable matrix, and write M as

$$M = P\Lambda P^{-1},$$

where P is the matrix of eigenvectors and $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ is the matrix of eigenvalues. If all the eigenvalues are in the domain of f, $\lambda_i \in D$, we define $f(M)$ to be the matrix

$$f(M) \coloneqq Pf(\Lambda)P^{-1},$$

where $f(\Lambda) = \operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_n))$ is a diagonal matrix consisting of $f(\lambda_i)$. One can check that this is well-defined, in the sense that different choices of diagonalization produce the same $f(M)$. Furthermore, when f is given by a convergent Taylor series $f(x) = \sum_{n=0}^{\infty} c_n x^n/n!$, the matrix $f(M)$ can be evaluated in a very nice way by replacing x with M in the power series of $f(x)$,

$$\begin{aligned} f(M) &= P \operatorname{diag}\left(f(\lambda_i)\right) P^{-1} = P \operatorname{diag}\left(\sum_{n=0}^{\infty} \frac{c_n}{n!}\lambda_i^n\right) P^{-1} \\ &= P \sum_{n=0}^{\infty} \frac{c_n}{n!}\Lambda^n P^{-1} = \sum_{n=0}^{\infty} \frac{c_n}{n!}\left(P\Lambda P^{-1}\right)^n = \sum_{n=0}^{\infty} \frac{c_n}{n!}M^n. \end{aligned}$$

A.2 Polar decomposition of matrices

The polar decomposition is described in Horn and Johnson's textbook [40]. Any $n \times n$ matrix M with real coefficients can be decomposed as the product

$$M = UD,$$

where U is orthogonal and D is positive definite. Indeed, M^TM is positive semidefinite, and we may define D to be the matrix square root of M^TM,

$$D \coloneqq \sqrt{M^TM},$$

as defined in Subsection A.1. Clearly, D is a symmetric matrix with nonnegative eigenvalues. We will now assume, for simplicity, that M is invertible – a general treatment may be found in [40]. If M is invertible, so is D, and we may define

$$U \coloneqq MD^{-1},$$

so that M is decomposed as $M = UD$. Note that

$$U^TU = (D^{-1})^TM^TMD = D^{-1}D^2D = 1,$$

so U is a unitary matrix.

A.3 The trace inner product

The vector space of $n \times n$ matrices is equipped with the inner product

$$\langle A, B\rangle = \mathrm{tr}\left(A^TB\right), \tag{A.1}$$

where A^T is the transpose of A and tr is the trace. One can verify that the usual properties of the dot product hold. That is, when A, B, C are matrices and λ is a number,

- (Bilinearity.) $\langle A, \lambda B\rangle = \lambda\langle A, B\rangle$ and $\langle A, B + C\rangle = \langle A, B\rangle + \langle A, C\rangle$,
- (Symmetry.) $\langle A, B\rangle = \langle B, A\rangle$,
- (Positive-definiteness.) $\langle A, A\rangle = \sum_{i,j=1}^{3} A_{ij}^2 \geq 0$ is the sum of squares of the entries of A, and it is zero if and only if A is zero.

A.4 Matrix norms

Definition A.4.1. For an $n\times n$ matrix A with real entries, define the *spectral radius* $\rho_{\max}(A)$ and *smallness* $\rho_{\min}(A)$ as the largest and smallest possible magnitudes of $A\mathbf{x}$ among all unit vectors $\mathbf{x} \in \mathbb{R}^n$,

$$\rho_{\max}(A) \coloneqq \sup\{||A\mathbf{x}|| : \mathbf{x} \in \mathbb{R}^n, ||\mathbf{x}|| = 1\},$$
$$\rho_{\min}(A) \coloneqq \inf\{||A\mathbf{x}|| : \mathbf{x} \in \mathbb{R}^n, ||\mathbf{x}|| = 1\}.$$

Remark A.4.2. For any vector $\mathbf{x} \in \mathbb{R}^n$, if $\mathbf{x} \neq 0$, we may define the unit vector $\mathbf{y} \coloneqq \mathbf{x}/||\mathbf{x}||$. Then the absolute value of $||A\mathbf{x}||$ is $||\mathbf{x}||$ times $||A\mathbf{y}||$, and the latter is between $\rho_{\min}(A)$ and $\rho_{\max}(A)$. If $\mathbf{x} = 0$, then $||A\mathbf{x}||$ is zero. Thus, in either case,

$$\rho_{\min}(A)\,||\mathbf{x}|| \leq ||A\mathbf{x}|| \leq \rho_{\max}(A)\,||\mathbf{x}||.$$

Proposition A.4.3. *The spectral radius is sub-multiplicative,* $\rho_{\max}(AB) \leq \rho_{\max}(A)\rho_{\max}(B)$.

Proof. Let $\mathbf{x} \in S^{n-1}$ be a unit vector. Then

$$||AB\mathbf{x}|| \leq \rho_{\max}(A)||B\mathbf{x}|| \leq \rho_{\max}(A)\rho_{\max}(B).$$

□

Let S^{n-1} be the $(n-1)$-dimensional unit sphere in $\mathbb{R}^n$. Since the function $f : S^{n-1} \to \mathbb{R}$, $\mathbf{x} \mapsto ||A\mathbf{x}||$ is continuous, and since S^{n-1} is topologically compact, we arrive at the following conclusion.

Remark A.4.4. There exist unit vectors $\mathbf{x}_1, \mathbf{x}_2 \in S^{n-1}$ saturating the inequalities in the definitions of $\rho_{\min}(A)$ and $\rho_{\max}(A)$,

$$\rho_{\max}(A) = ||A\mathbf{x}_1||, \qquad \rho_{\min}(A) = ||A\mathbf{x}_2||.$$

Proposition A.4.5. *For an invertible matrix A, the spectral radius and smallness are related by* $\rho_{\max}(A)\rho_{\min}(A^{-1}) = 1$.

Proof. Let $\mathbf{x}$ be a unit vector such that $\rho_{\min}(A^{-1}) = ||A\mathbf{x}||$. Then

$$1 = ||\mathbf{x}|| = ||AA^{-1}\mathbf{x}|| \leq \rho_{\max}(A)||A^{-1}\mathbf{x}|| = \rho_{\max}(A)\rho_{\min}(A^{-1}).$$

Let $\mathbf{x}$ be a unit vector such that $\rho_{\max}(A) = ||A\mathbf{x}||$. Then

$$1 = ||\mathbf{x}|| = ||A^{-1}A\mathbf{x}|| \geq \rho_{\min}(A^{-1})||A\mathbf{x}|| = \rho_{\min}(A^{-1})\rho_{\max}(A).$$

□

Appendix B

Matching sublattices for biaxial epitaxy

Recall that Section 2.6 quantifies biaxial epitaxy, a process when a crystalline material (the epilayer) is slowly deposited onto a substrate crystal cut in some crystal orientation, resulting (ideally) in a thin crystalline film of the epilayer that matches the lattice parameters of the substrate crystal within the plane of growth. A necessary condition for the two crystals to match is the sublattice matching equation (eq. 2.19),

$$F_1 B_1 K_1 = F_2 B_2 K_2.$$

Here, B_1 and B_2 are Bravais matrices determining the lattice parameters of the substrate and epilayer crystal (these are 3×3 matrices with real coefficients), K_1 and K_2 are 3×2 matrices with integer coefficients determining the plane of growth, and F_1, F_2 are the deformation gradients of the crystals (3×3 matrices with real coefficients). Variables B_1, K_1, F_1 correspond to the epilayer, while B_2, K_2, F_2 correspond to the substrate.

We now present an algorithm to parametrize solutions to the sublattice matching equation. The solutions are pairs (F_1, F_2) of deformation gradients, and B_1, B_2, K_1, K_2 are fixed parameters. Recall from Section 2.6 that we are interested in solutions up to a rotation, meaning that $(U\,F_1, U\,F_2)$ is the same for us as (F_1, F_2) whenever U is a rotation matrix. In Section B.1, we find one solution to the sublattice matching equation, and it has the effect of fixing the rotational freedom. In Section B.2, with F_2 fixed from before, we parametrize the possible values of F_1; since $F_1 B_1 K_1 = F_2 B_2 K_2$ becomes a linear equation, the parametrization is done in the usual way for linear systems [41].

B.1 Finding a particular solution

Algorithm 1 produces one solution (F_1, F_2) to the sublattice matching equation $F_1 B_1 K_1 = F_2 B_2 K_2$. Moreover, it produces a "nice" solution: by Proposition B.1.2, when the substrate and epilayer sublattices are similar, the epilayer deformation gradient F_1 is close to the

identity matrix. This corresponds to the reasonable expectation that when the substrate and epilayer crystals are very compatible, they can fit together with little strain.

Proposition B.1.1. *With the definitions from Algorithm 1, the following identities hold.*

1. $P_1 P_1^T = P_2 P_2^T = 1$ *is the* 2×2 *identity matrix.*
2. $P_1^T P_1 + \mathbf{w}_1 \mathbf{w}_1^T = P_2^T P_2 + \mathbf{w}_2 \mathbf{w}_2^T = 1$ *is the* 3×3 *identity matrix.*
3. $\mathbf{w}_1^T B_1 K_1 = \mathbf{w}_2^T B_2 K_2 = 0$.
4. D *is positive definite.*
5. U *is a rotation matrix.*
6. $D B_1 K_1 = U B_2 K_2$.

Proof. We leave points 1, 2, 3 to the reader. To prove points 4 and 5, define 3×3 matrices S_1, S_2 to transform from the usual basis for $\mathbb{R}^3$ to the right-handed orthonormal bases $\mathbf{u}_1, \mathbf{v}_1, \mathbf{w}_1$ and $\mathbf{u}_2, \mathbf{v}_2, \mathbf{w}_2$,

$$S_1 \coloneqq \begin{pmatrix} \mathbf{u}_1 & \mathbf{v}_1 & \mathbf{w}_1 \end{pmatrix}, \quad S_2 \coloneqq \begin{pmatrix} \mathbf{u}_2 & \mathbf{v}_2 & \mathbf{w}_2 \end{pmatrix}.$$

Since $\tilde{D}$ is positive definite and $D = S_1 \begin{pmatrix} \tilde{D} & \\ & 1 \end{pmatrix} S_1^T$, we see that D is positive definite. To show that U is a rotation matrix, we examine V and W individually. Note that

$$\det S_1 = (\mathbf{u}_1 \times \mathbf{v}_1) \cdot \mathbf{w}_1 = ||\mathbf{w}_1||^2 = 1,$$

and likewise, $\det S_2 = 1$. One can verify that $V S_2 = S_1$. By the multiplicativity of the determinant, it has to be that $\det V = 1$, so V is a rotation matrix. Since $W = S_1 \begin{pmatrix} \tilde{W} & \\ & 1 \end{pmatrix} S_1^T$, we see that $\det W = \det \tilde{W}$. By construction, $\tilde{W}^T \tilde{W} = 1$, so $\det \tilde{W}$ is either 1 or -1. But since $\det c_1$, $\det c_2$, and $\det D$ are all positive, the equation

$$\det c_2 / \det c_1 = \det \tilde{F} = \det \tilde{D} \ \det \tilde{W}$$

requires that $\det \tilde{W} > 0$. Therefore, W is a rotation matrix. To prove point 6, we provide the following calculation,

$$\begin{aligned} D B_1 K_1 &= P_1^T \tilde{D} P_1 B_1 K_1 \\ &= P_1^T \tilde{D} c_1 = P_1^T \tilde{W} \tilde{F} c_1 \\ &= P_1^T \tilde{W} c_2 = U P_2^T c_2 = U B_2 K_2. \end{aligned}$$

In the last line, we use that $V P_2^T = P_1^T$ and that $W P_1^T = P_1^T \tilde{W}$. □

[3]For example, we can use the Gram-Schmidt process. For a nonzero vector $\mathbf{x}$, let $\mathrm{nz}(\mathbf{x}) \coloneqq \mathbf{x}/||\mathbf{x}||$ be the unit vector in the direction of $\mathbf{x}$. Then, we may let $\mathbf{u}_1 = \mathrm{nz}(B_1 \mathbf{n}_1)$ and $\mathbf{v}_1 = \mathrm{nz}(B_1 \mathbf{m}_1 - (\mathbf{u}_1^T B_1 \mathbf{m}_1) \mathbf{u}_1)$.

[4]Strictly speaking, a projection operator on $\mathbb{R}^3$ must be a 3×3 matrix π with $\pi^2 = \pi$. Under this definition, the projection operator onto $B_1 K_1$ is $P_1^T P_1 : \mathbb{R}^3 \to \mathbb{R}^3$, not $P_1 : \mathbb{R}^3 \to \mathbb{R}^2$.

Algorithm 1 Sublattice matching

Require: $B_1, B_2 \in \mathbb{R}^{3\times 3}$ are nonsingular; $K_1 = \begin{pmatrix}\mathbf{n}_1 & \mathbf{m}_1\end{pmatrix} \in \mathbb{Z}^{3\times 2}$ is such that $\mathbf{n}_1$ and $\mathbf{m}_1$ are non-collinear, and likewise for $K_2 = \begin{pmatrix}\mathbf{n}_2 & \mathbf{m}_2\end{pmatrix} \in \mathbb{Z}^{3\times 2}$.

Ensure: Find matrices $F_1 = D$ and $F_2 = U$ such that $DB_1K_1 = UB_2K_2$, where D is positive definite and U a rotation matrix (Proposition B.1.1). Moreover, $D \approx 1$ is close to the identity matrix, as specified in Proposition B.1.2.

Let $\mathbf{u}_1, \mathbf{v}_1 \in \mathbb{R}^3$ be an orthonormal basis for the plane formed by $B_1\,\mathbf{n}_1$ and $B_1\,\mathbf{m}_1$, and likewise, let $\mathbf{u}_2, \mathbf{v}_2 \in \mathbb{R}^3$ be an orthonormal basis for the plane formed by $B_2\,\mathbf{n}_2$ and $B_2\,\mathbf{m}_2$.[1] Define the matrices $P_1, P_2 \in \mathbb{R}^{2\times 3}$ as the "projection matrices"[2]onto the planes $B_1\,K_1$ and $B_2\,K_2$,

$$P_1 \coloneqq \begin{pmatrix}\mathbf{u}_1^T \\ \mathbf{v}_1^T\end{pmatrix}, \quad P_2 \coloneqq \begin{pmatrix}\mathbf{u}_2^T \\ \mathbf{v}_2^T\end{pmatrix}, \tag{B.1}$$

and define the 2×2 matrices c_1 and c_2,

$$c_1 \coloneqq P_1B_1K_1, \quad c_2 \coloneqq P_2B_2K_2, \tag{B.2}$$

so that the columns of c_1 express $B_1\,\mathbf{n}_1$ and $B_2\,\mathbf{m}_1$ in the orthonormal basis defined by $\mathbf{u}_1$ and $\mathbf{v}_1$, and likewise for c_2. One can verify that c_1 and c_2 are nonsingular. We may require the positive orientation condition

$$\det c_1 > 0, \quad \det c_2 > 0, \tag{B.3}$$

for if $\det c_1 < 0$, instead of the basis $(\mathbf{u}_1, \mathbf{v}_1)$ we can equally well use the basis $(\mathbf{v}_1, \mathbf{u}_1)$ – that is, we switch the order of the basis vectors.

Define $\mathbf{w}_1, \mathbf{w}_2 \in \mathbb{R}^3$ as the normal vectors to the planes B_1K_1 and B_2K_2,

$$\mathbf{w}_1 \coloneqq \mathbf{u}_1 \times \mathbf{v}_1, \quad \mathbf{w}_2 \coloneqq \mathbf{u}_2 \times \mathbf{v}_2. \tag{B.4}$$

Let $\tilde{F} \coloneqq c_2\,c_1^{-1}$, and let $\tilde{D} \coloneqq \sqrt{\tilde{F}^T\tilde{F}}$ and $\tilde{W} \coloneqq \left(\tilde{F}\tilde{D}^{-1}\right)^{-1}$, so that $\tilde{F} = \tilde{W}^T\tilde{D}$ is written as a rotation matrix times a positive definite matrix (Subsection A.2). We now "lift" $\tilde{D}$ and $\tilde{W}$ to 3×3 matrices D and W by applying $\tilde{D}$ and $\tilde{W}$ in the B_1K_1 plane while leaving the $\mathbf{w}_1$ direction invariant,

$$D \coloneqq P_1^T\tilde{D}P_1 + \mathbf{w}_1\mathbf{w}_1^T, \quad W \coloneqq P_1^T\tilde{W}P_1 + \mathbf{w}_1\mathbf{w}_1^T. \tag{B.5}$$

Define a 3×3 matrix V by $V \coloneqq \mathbf{u}_1\mathbf{u}_2^T + \mathbf{v}_1\mathbf{v}_2^T + \mathbf{w}_1\mathbf{w}_2^T$. Lastly, let $U = WV$.

Define the inner product matrices $g_1, g_2 \in \mathbb{R}^{2\times 2}$ by

$$g_1 \coloneqq (B_1K_1)^T(B_1K_1), \quad g_2 \coloneqq (B_2K_2)^T(B_2K_2) \tag{B.6}$$

to encode the cell parameters of the sublattices B_1K_1 and B_2K_2 in a basis-independent way.

Proposition B.1.2. *When the substrate and epilayer sublattices have similar parameters, the matrix D as defined in Algorithm 1 is close to the identity matrix. That is, the trace norm (Appendix A.3) of $D^2 - 1$ is bounded by the trace norm of $g_2\, g_1^{-1} - 1$,*

$$||D^2 - 1||_2 = \sqrt{\operatorname{tr}\left((g_2 g_1^{-1} - 1)^2\right)} \leq ||g_2\, g_1^{-1} - 1||_2.$$

Since $E = D - 1$, we see that $||D^2 - 1||_2$ is approximately equal to twice the absolute value of the strain, $2||E||_2$. Thus, $g_1 \approx g_2$ implies small strain in the solution to the sublattice matching equation produced by Algorithm 1.

Proof. By Proposition B.1.1,

$$\begin{aligned}\operatorname{tr}\left((D^2-1)^2\right) &= \operatorname{tr}\left((D^2 - P_1^T P_1 - \mathbf{w}_1\mathbf{w}_1^T)^2\right) \\ &= \operatorname{tr}\left(P_1^T(\tilde{D}^2-1)^2 P_1\right) = \operatorname{tr}\left((\tilde{D}^2-1)^2\right).\end{aligned}$$

Recalling that $\tilde{F} = c_2\, c_1^{-1}$ and that $\tilde{D}^2 = \tilde{F}^T\tilde{F}$, we find

$$\operatorname{tr}\left((\tilde{D}^2-1)^2\right) = \operatorname{tr}\left((c_1^{-T} g_2\, c_1^{-1} - 1)^2\right) = \operatorname{tr}\left((g_2 g_1^{-1} - 1)^2\right),$$

completing the proof. □

B.2 Parameterizing all solutions

Having used Algorithm 1 to find one pair of matrices that solves the sublattice matching equation $F_1B_1K_1 = F_2B_2K_2$ (eq. 2.19), call it (F_{1*}, F_{2*}), how do we parametrize all solutions up to rotation? Recall that we are requiring F_2 to be a rotation matrix because the substrate is not strained. Therefore, if (F_1, F_2) is any solution, then (UF_1, F_{2*}) is an equivalent solution, where the rotation matrix U is given by $U = F_{2*}F_2^{-1}$. We may fix $F_2 = F_{2*}$ and look for matrices $F_1 \in \mathbb{R}^{3\times 3}$ solving the linear equation

$$F_1B_1K_1 = F_{2*}B_2K_2. \tag{B.7}$$

We know that F_{1*}, the value computed by Algorithm 1, is a particular solution to this equation. To find all solutions, let $\mathbf{w} \in \mathbb{R}^3$ be a normal vector to the plane made by $B_1\,\mathbf{n}_1$ and $B_1\,\mathbf{m}_1$. Note that if $\mathbf{r} \in \mathbb{R}^3$ is any vector, then

$$F_1 = F_{1*} + \mathbf{r}\,\mathbf{w}_1^T \tag{B.8}$$

solves equation B.7 because $\mathbf{w}_1^T B_1 K_1 = 0$. Conversely, suppose F_1 is a solution. Since the vectors $\mathbf{w}_1, B_1\mathbf{n}_1, B_1\mathbf{m}_1$ form a basis for $\mathbb{R}^3$, we can write $F_1 - F_{1*} = \mathbf{r}\,\mathbf{w}_1^T + Q\,(B_1K_1)^T$ for some $\mathbf{r} \in \mathbb{R}^3$ and $Q \in \mathbb{R}^{3\times 2}$. Then,

$$0 = (F_1 - F_{1*})B_1K_1 = Q\,(B_1K_1)^T(B_1K_1),$$

so $Q = 0$ and $F_1 = F_{1*} + \mathbf{r}\,\mathbf{w}_1^T$ is of the form of equation B.8. Thus, if (F_{1*}, F_{2*}) solves the sublattice matching equation, the set of all solutions, up to rotation, is parametrized as $(F_{1*} + \mathbf{r}\,\mathbf{w}_1^T, F_{2*})$. The parameter $\mathbf{r} \in \mathbb{R}^3$ specifies how the direction perpendicular to the plane gets transformed by F_1.

We are interested in solutions with small strain – that is, F_1 is "almost a rotation matrix." When we set $\mathbf{r} = 0$, we recover the particular solution from Section B.1 which does have small strain. If we set $\mathbf{r}$ to a large value, F_1 will be far from the identity. Thus, we need to consider small perturbations $\mathbf{r} \approx 0$ from the particular solution produced by Algorithm 1.

Bibliography

[1] J. P. Ruf, H. Paik, N. J. Schreiber, H. P. Nair, L. Miao, J. K. Kawasaki, J. N. Nelson, B. D. Faeth, Y. Lee, B. H. Goodge, et al., Nature Communications **12**, 59 (2021), ISSN 2041-1723, URL https://doi.org/10.1038/s41467-020-20252-7.

[2] B. Pamuk, N. Wadehra, R. Kaldybayev, B. Gregory, S. Zhang, N. Schnitzer, Y. Iguchi, E. Li, L. Kourkoutis, A. Singer, et al., aPS Global Summit Abstracts, Session MAR-W23, (2025).

[3] M. Tuniz, A. Consiglio, G. Pokharel, F. Parmigiani, T. Neupert, R. Thomale, S. K. Chaluvadi, P. Orgiani, G. Sangiovanni, S. D. Wilson, et al., Phys. Rev. Lett. **134**, 066501 (2025), URL https://link.aps.org/doi/10.1103/PhysRevLett.134.066501.

[4] P. Settembri, F. Mazzola, I. Vobornik, J. Fujii, M. Kögler, C.-N. Kuo, C. S. Lue, A. Politano, and G. Profeta, Phys. Rev. B **110**, L201401 (2024), URL https://link.aps.org/doi/10.1103/PhysRevB.110.L201401.

[5] D. Sibanda, S. T. Oyinbo, T.-C. Jen, and A. I. Ibitoye, Processes **10** (2022), ISSN 2227-9717, URL https://www.mdpi.com/2227-9717/10/6/1184.

[6] B. Z. Gregory, J. Strempfer, D. Weinstock, J. P. Ruf, Y. Sun, H. Nair, N. J. Schreiber, D. G. Schlom, K. M. Shen, and A. Singer, Phys. Rev. B **106**, 195135 (2022), URL https://link.aps.org/doi/10.1103/PhysRevB.106.195135.

[7] D. G. Schlom, L.-Q. Chen, X. Pan, A. Schmehl, and M. A. Zurbuchen, Journal of the American Ceramic Society **91**, 2429 (2008), https://ceramics.onlinelibrary.wiley.com/doi/pdf/10.1111/j.1551-2916.2008.02556.x, URL https://ceramics.onlinelibrary.wiley.com/doi/abs/10.1111/j.1551-2916.2008.02556.

[8] Y. Qi, M. A. Sadi, D. Hu, M. Zheng, Z. Wu, Y. Jiang, and Y. P. Chen, Advanced Materials **35**, 2205714 (2023), https://advanced.onlinelibrary.wiley.com/doi/pdf/10.1002/adma.202205714, URL https://advanced.onlinelibrary.wiley.com/doi/abs/10.1002/adma.202205714.

[9] J. Hwang, Z. Feng, N. Charles, X. R. Wang, D. Lee, K. A. Stoerzinger, S. Muy, R. R. Rao, D. Lee, R. Jacobs, et al., Materials Today **31**, 100 (2019), ISSN 1369-7021, URL https://www.sciencedirect.com/science/article/pii/S1369702119300185.

[10] H. S. Inbar, Ph.D. thesis, University of California, Santa Barbara, United States – California (2023), copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works., URL https://ezproxy2.williams.edu/login?url=https://www.proquest.com/dissertations-these

[11] L. Yang, H. Wang, S. Yang, and Y. Lu, Materials for Quantum Technology **4**, 023001 (2024), https://iopscience.iop.org/article/10.1088/2633-4356/ad4e8d/pdf, URL https://app.dimensions.ai/details/publication/pub.1172314351.

[12] T. Li, S. Deng, H. Liu, and J. Chen, Chemical Reviews **124**, 7045 (2024), ISSN 0009-2665, URL https://doi.org/10.1021/acs.chemrev.3c00767.

[13] V. K. Kozin, E. Thingstad, D. Loss, and J. Klinovaja, Phys. Rev. B **111**, 035410 (2025), URL https://link.aps.org/doi/10.1103/PhysRevB.111.035410.

[14] T. Zhang, Y. Liu, J. Yu, Q. Ye, L. Yang, Y. Li, and H. J. Fan, Advanced Materials **34**, 2202195 (2022), https://advanced.onlinelibrary.wiley.com/doi/pdf/10.1002/adma.202202195, URL https://advanced.onlinelibrary.wiley.com/doi/abs/10.1002/adma.202202195.

[15] J. Bardeen, L. N. Cooper, and J. R. Schrieffer, Phys. Rev. **108**, 1175 (1957), URL https://link.aps.org/doi/10.1103/PhysRev.108.1175.

[16] F. Marsiglio, Annals of Physics **417**, 168102 (2020), ISSN 0003-4916, eliashberg theory at 60: Strong-coupling superconductivity and beyond, URL https://www.sciencedirect.com/science/article/pii/S000349162030035X.

[17] W. Nunn, S. Nair, H. Yun, A. Kamath Manjeshwar, A. Rajapitamahuni, D. Lee, K. A. Mkhoyan, and B. Jalan, APL Materials **9**, 091112 (2021), ISSN 2166-532X, https://pubs.aip.org/aip/apm/article-pdf/doi/10.1063/5.0062726/14566450/091112_1_onli URL https://doi.org/10.1063/5.0062726.

[18] P. Keßler, T. Waldsauer, V. Jovic, M. Kamp, M. Schmitt, M. Sing, R. Claessen, and S. Moser, APL Materials **12**, 101110 (2024), ISSN 2166-532X, https://pubs.aip.org/aip/apm/article-pdf/doi/10.1063/5.0217312/20211920/101110_1_5.02 URL https://doi.org/10.1063/5.0217312.

[19] M. Uchida, T. Nomoto, M. Musashi, R. Arita, and M. Kawasaki, Phys. Rev. Lett. **125**, 147001 (2020), URL https://link.aps.org/doi/10.1103/PhysRevLett.125.147001.

[20] A. Aloisio, A. Contento, R. Alaggio, and G. Quaranta, Mechanical Systems and Signal Processing **194**, 110276 (2023), ISSN 0888-3270, URL https://www.sciencedirect.com/science/article/pii/S0888327023001838.

[21] T. Würth, C. Krauß, C. Zimmerling, and L. Kärger, Materials & Design **231**, 112034 (2023), ISSN 0264-1275, URL https://www.sciencedirect.com/science/article/pii/S0264127523004495.

[22] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, et al., Journal of Physics: Condensed Matter **21**, 395502 (19pp) (2009), URL http://www.quantum-espresso.org.

[23] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. B. Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, et al., Journal of Physics: Condensed Matter **29**, 465901 (2017), URL http://stacks.iop.org/0953-8984/29/i=46/a=465901.

[24] P. Giannozzi, O. Baseggio, P. Bonfà, D. Brunato, R. Car, I. Carnimeo, C. Cavazzoni, S. de Gironcoli, P. Delugas, F. Ferrari Ruffino, et al., The Journal of Chemical Physics **152**, 154105 (2020), https://doi.org/10.1063/5.0005082, URL https://doi.org/10.1063/5.0005082.

[25] K. Burke, *The abc of dft* (2007), URL https://dft.uci.edu/doc/g1.pdf.

[26] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt-Saunders, 1976).

[27] R. W. Ogden, *Nonlinear elasticity with application to material modelling*, online (2003).

[28] T. Berlijn, P. C. Snijders, O. Delaire, H.-D. Zhou, T. A. Maier, H.-B. Cao, S.-X. Chi, M. Matsuda, Y. Wang, M. R. Koehler, et al., Phys. Rev. Lett. **118**, 077201 (2017), URL https://link.aps.org/doi/10.1103/PhysRevLett.118.077201.

[29] J. Marsden and T. J. R. Hughes, *Mathematical foundations of elasticity* (Dover Publications, Inc., 1983), ISBN 0-486-67865-2.

[30] D. M. A. Leggett, The Journal of the Royal Aeronautical Society **64**, 176–177 (1960).

[31] T. Hahn, *International Tables for Crystallography, Volume A: Space Group Symmetry* (Springer, 2002).

[32] T. Akiba, S. Sano, T. Yanase, T. Ohta, and M. Koyama, in *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (Association for Computing Machinery, New York, NY, USA, 2019), KDD '19, p. 2623–2631, ISBN 9781450362016, URL https://doi.org/10.1145/3292500.3330701.

[33] J. Bardeen and W. Shockley, Phys. Rev. **80**, 72 (1950), URL https://link.aps.org/doi/10.1103/PhysRev.80.72.

[34] D. J. Griffiths and D. F. Schroeter, *Introduction to Quantum Mechanics* (Cambridge University Press, 2018), 3rd ed.

[35] T. Konno, H. Kurokawa, F. Nabeshima, Y. Sakishita, R. Ogawa, I. Hosako, and A. Maeda, Phys. Rev. B **103**, 014509 (2021), URL https://link.aps.org/doi/10.1103/PhysRevB.103.014509.

[36] Z. Wang, S. Zhang, J. Wang, M. Zhang, Y. Chen, B. Li, T. Zhao, M. Zhu, F. Hu, B. Shen, et al., Acta Materialia **292**, 121031 (2025), ISSN 1359-6454, URL https://www.sciencedirect.com/science/article/pii/S1359645425003210.

[37] J. B. Gibson, A. C. Hire, P. M. Dee, O. Barrera, B. Geisler, P. J. Hirschfeld, and R. G. Hennig, npj Computational Materials **11**, 7 (2025), ISSN 2057-3960, URL https://doi.org/10.1038/s41524-024-01475-4.

[38] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, Commun. ACM **65**, 99–106 (2021), ISSN 0001-0782, URL https://doi.org/10.1145/3503250.

[39] M. Raissi, P. Perdikaris, and G. Karniadakis, Journal of Computational Physics **378**, 686 (2019), ISSN 0021-9991, URL https://www.sciencedirect.com/science/article/pii/S0021999118307125.

[40] R. A. Horn and C. R. Johnson, *Matrix Analysis* (Cambridge University Press, 1985).

[41] G. Strang, *Introduction to Linear Algebra* (CUP, 2023), 6th ed.