

DISS. ETH NO. 25670

# IDENTIFYING TOPOLOGICAL SEMIMETALS

A thesis submitted to attain the degree of  
DOCTOR OF SCIENCES of ETH ZURICH  
(Dr. sc. ETH Zurich)

presented by

DOMINIK ANDRE GRESCH

MSc. in Physics ETH Zurich

born on 24.10.1991

citizen of  
Freienbach SZ

accepted on the recommendation of  
Prof. Dr. Matthias Troyer, examiner  
Prof. Dr. Titus Neupert, co-examiner  
Prof. Dr. Alexey A. Soluyanov, co-examiner

2018

Dominik André Gresch  
*Identifying Topological Semimetals*  
Diss. ETH No. 25670

Digital Object Identifier [doi:10.3929/ethz-b-000308602](https://doi.org/10.3929/ethz-b-000308602)  
E-mail [greschd@phys.ethz.ch](mailto:greschd@phys.ethz.ch)

# Abstract

Geometric properties of electron states in crystalline solids lead to a *topological* classification of materials. A remarkable consequence of this topological viewpoint is that it reveals a deep link between the bulk properties of a material and electronic states which form on its surface. This leads to unique transport properties, the most well-known example being the integer quantum Hall effect. In topological semimetals, the bulk features of interest are nodes in the band structure, where occupied and unoccupied states are not separated by an energy gap. This leads to interesting low-energy excitations, some of which are the condensed matter equivalent of fundamental particles. The Weyl Fermion for example is realized in topological semimetals, which is theoretically postulated but eludes experimental verification in high-energy physics. Crystals however do not have a continuous translational symmetry, and thus do not need to fulfill the so-called Lorentz invariance present in high-energy physics. This allows for Fermions to exist in materials which do not have a fundamental counterpart. The main topic of this thesis is the study and identification of topological semimetals. We propose a mechanism for Weyl Fermions to form under the influence of an external magnetic field. This effect could help explain the anisotropic negative magnetoresistance in transition metal dipnictides. We also study several novel topological material candidates, hosting a plethora of Weyl Fermions and topological nodal lines. In addition to studying specific material examples, we also present several tools and algorithms which enhance the process of identifying topological materials. First, we present an algorithm for evaluating the phase diagram of a system with discrete phases. This is useful in identifying topological phases, but also applicable to other fields of computational physics. Furthermore, we develop tools that simplify the creation of  $\mathbf{k} \cdot \mathbf{p}$  and tight-binding models to study crystalline systems. A particular focus lies on the construction of models which preserve the crystal symmetries, since these play a crucial role in determining the topology of a material. And finally, we develop an algorithm that reliably finds and classifies topological nodal features.



# Zusammenfassung

Geometrische Eigenschaften der Elektronenzustände in Kristallen erlauben eine *topologische* Klassifizierung der Materialien. Eine ausserordentliche Konsequenz dieser topologischen Sichtweise ist, dass sie einen tiefen Zusammenhang zwischen den Eigenschaften im Inneren eines Materials und Oberflächenzuständen der Elektronen offenbart. Dies hat einzigartige Transporteigenschaften zur Folge, von denen der ganzzahlige Quanten-Hall-Effekt das bekannteste Beispiel ist. In topologischen Halbmetallen sind diejenigen Teile der Bandstruktur von besonderem Interesse, an denen sich die besetzten und unbesetzten Zustände berühren, sie also nicht durch eine Energielücke getrennt sind. Dies führt zu niederenergetischen Anregungen, von denen einige das Äquivalent elementarer Teilchen in kondensierter Materie sind. Das Weyl Fermion zum Beispiel existiert in topologischen Halbmetallen, während es in der Hochenergiephysik zwar theoretisch postuliert aber nicht experimentell nachgewiesen ist. Allerdings haben Kristalle keine kontinuierliche Translations-symmetrie, weshalb die sogenannte Lorentz-Invarianz anders als in der Hochenergiephysik nicht erhalten bleiben muss. Dies erlaubt in Materialien die Existenz von Fermionen welche kein elementares Gegenstück haben. Das Hauptthema dieser Dissertation ist das Studium und Identifizieren topologischer Halbmetalle. Wir präsentieren einen Mechanismus, wie sich Weyl Fermionen unter Einfluss eines externen magnetischen Feldes bilden können. Dieser Effekt könnte dazu beitragen, die anisotrope negative Magnetoresistivität von Übergangsmetall-Dipnictiden zu erklären. Ausserdem untersuchen wir mehrere neuartige Materialkandidaten, die eine Fülle von Weyl-Fermionen und topologischen Linienstrukturen beherbergen. Neben der Untersuchung spezifischer Materialbeispiele präsentieren wir verschiedene Werkzeuge und Algorithmen, die der Identifizierung topologischer Materialien dienen. Zunächst stellen wir einen Algorithmus zur Berechnung von Phasendiagrammen mit diskreten Phasen vor. Dies ist nützlich im Zusammenhang mit topologischen Phasen, aber auch in anderen Teilgebieten der Computerphysik anwendbar. Des weiteren entwickeln wir Werkzeuge, welche das Erstellen von  $\mathbf{k} \cdot \mathbf{p}$  - und Tight-Binding-Modellen für Kristalle erleichtern. Speziell im Fokus sind Techniken, durch welche die erstellten Modelle die Kristallsymmetrien erhalten. Diese sind für die topologische Struktur der Materialien von zentraler Bedeutung. Schliess-

---

lich entwickeln wir Algorithmus der zuverlässig die topologischen Merkmale von Halbmetallen findet und klassifiziert.

# Contents

<b>Abstract</b>	<b>iii</b>
<b>Zusammenfassung</b>	<b>v</b>
<b>Contents</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 A Short Introduction to Band Theory</b>	<b>5</b>
2.1 Periodic Lattices . . . . .	5
2.2 Bloch States . . . . .	6
2.3 Numerical Methods . . . . .	7
2.3.1 Density Functional Theory . . . . .	7
2.3.2 Tight-Binding Models . . . . .	9
2.3.3 $\mathbf{k} \cdot \mathbf{p}$ Models . . . . .	10
<b>3 Topological Properties of Electronic Bands</b>	<b>13</b>
3.1 Topology in Non-interacting Materials . . . . .	13
3.1.1 Topological Properties . . . . .	13
3.1.2 Bulk-edge Correspondence . . . . .	14
3.2 The Chern Number . . . . .	16
3.2.1 The Berry Phase and Chern Invariant . . . . .	16
3.2.2 The Chern Number as Change in Berry Phase . . . . .	18
3.2.3 Wilson Loop and Hybrid Wannier Charge Centers . . . . .	19
3.3 Time-reversal Symmetry: $\mathbb{Z}_2$ Classification . . . . .	21
3.3.1 Individual Chern numbers . . . . .	22
<b>4 Type-II Weyl Semimetals</b>	<b>27</b>
4.1 Weyl Nodes . . . . .	27
4.2 Type-II Weyl Semimetals . . . . .	29
4.2.1 Material Examples . . . . .	31

4.3	Other Topological Semimetals . . . . .	32
<b>5</b>	<b>An Algorithm for Calculating Phase Diagrams</b>	<b>33</b>
5.1	Problem Description . . . . .	33
5.2	Description of the Algorithm . . . . .	34
5.2.1	Computational Complexity . . . . .	34
5.2.2	Asynchronous Execution . . . . .	36
5.3	Convergence Properties . . . . .	36
5.3.1	A Note on Terminology . . . . .	36
5.3.2	Phase Containing a Starting Point . . . . .	37
5.3.3	Sufficient Condition for Convergence to Covering a Phase . . . . .	39
5.4	Conclusion . . . . .	42
<b>6</b>	<b>Weyl Fermions Emerging in a Magnetic Field</b>	<b>45</b>
6.1	Atomic and Electronic Structure of AB <sub>2</sub> compounds . . . . .	46
6.1.1	Atomic Structure . . . . .	46
6.1.2	Electronic Structure . . . . .	47
6.1.3	Symmetry Operations and $\mathbf{k} \cdot \mathbf{p}$ Model . . . . .	51
6.2	Band Structure Topology . . . . .	54
6.2.1	Band Structure Topology without Magnetic Fields . . . . .	54
6.2.2	Effect of Zeeman Splitting on the $\mathbf{k} \cdot \mathbf{p}$ Model for TaAs <sub>2</sub> . . . . .	55
6.2.3	Effect of Zeeman Splitting on the Tight-binding Model for NbSb <sub>2</sub> . . . . .	59
6.2.4	Limitations of the Model for Magnetic Field . . . . .	61
6.3	Conclusions . . . . .	65
<b>7</b>	<b>General Form of Symmetric <math>\mathbf{k} \cdot \mathbf{p}</math> Models</b>	<b>67</b>
7.1	Symmetry-Constrained Hamiltonian . . . . .	67
7.1.1	Calculating the Matrix Form of $\hat{F}_g$ . . . . .	69
7.1.2	Calculating the Matrix Form of $\hat{G}_g$ . . . . .	70
7.1.3	Calculating the Invariant Subspace of the Symmetry Group . . . . .	70
7.1.4	Limitations . . . . .	71
7.2	Example Application . . . . .	71
7.3	Conclusion . . . . .	72
<b>8</b>	<b>Automated Construction of Wannier Tight-binding Models</b>	<b>73</b>
8.1	Construction of Wannier Tight-binding Models . . . . .	75
8.1.1	Wannier Tight-binding Construction . . . . .	75
8.1.2	Symmetrization . . . . .	76
8.1.3	Optimization for Band Structure Fit . . . . .	82
8.2	Implementation in AiiDA Workflows . . . . .	83

---

8.2.1	Modular Workflow Design . . . . .	84
8.2.2	Tight-binding Extraction Workflow . . . . .	87
8.3	Strain-dependent Tight-binding Models for Majorana Devices . . . . .	88
8.3.1	Strained Tight-binding Workflow . . . . .	90
8.3.2	First-principles Calculations . . . . .	91
8.3.3	Strain Interpolation . . . . .	92
8.3.4	Results . . . . .	93
8.4	Conclusion and Outlook . . . . .	95
<b>9</b>	<b>Automated Classification of Band Degeneracies</b>	<b>99</b>
9.1	Nodal Search Algorithm . . . . .	99
9.2	Identification of Nodal Features . . . . .	103
9.2.1	Clustering . . . . .	103
9.2.2	Dimensionality Analysis . . . . .	103
9.2.3	Classification . . . . .	106
9.3	Materials Applications . . . . .	112
9.3.1	Validation on Known Topological Semimetals . . . . .	112
9.3.2	Application to Novel Topological Semimetals . . . . .	114
9.4	Conclusion . . . . .	130
<b>10</b>	<b>Conclusion and Outlook</b>	<b>131</b>
	<b>List of Publications</b>	<b>133</b>
	<b>Bibliography</b>	<b>135</b>
	<b>Acknowledgements</b>	<b>155</b>
	<b>Curriculum Vitae</b>	<b>157</b>



# 1

## Introduction

Crystalline states of matter – solids in which atoms form regular, repeating patterns – are ubiquitous in almost any aspect of modern technology. Ferromagnetic materials for example are essential in constructing electric motors, generators or transformers. Semiconductors have powered the computational revolution of the recent decades. Solid-state lasers, light-emitting diodes, solar cells, digital image sensors, and many more technologies fundamentally rely on such crystalline solids. Unsurprisingly, the study of such systems is a major field in condensed matter physics. A particularly active topic is the study of electronic states, since these give rise to many of the aforementioned technological applications. An important factor when studying electrons in a solid is the strength of correlations between individual electrons. When these interactions are strong, quantum effects can become dominant, leading to phenomena such as superconductivity, charge density waves, and Mott insulators to name just a few. A much simpler picture arises when these interactions are weak enough to be ignored. In this case, each electron can be described as being under the influence of an average periodic potential created by the ionic cores and surrounding electrons. The whole system can then be described as a set of independent wave-like states which are either occupied by an electron or unoccupied, depending on their energy. Despite this simplification, surprisingly complex phenomena can still be observed. This was first realized in 1980 by von Klitzing in the integer quantum Hall effect [10]. When applying a strong magnetic field to a two-dimensional electron gas with high mobility, the Hall conductivity forms striking plateaus. The values of these plateaus are integer multiples of a universal constant, with a remarkably low error of less than  $10^{-6}$  even in the early experiment. The reason for this effect is that the Landau levels – electron states induced by the magnetic field – produce chiral edge states. Each

---

occupied Landau level contributes one such edge state, and consequently an integer contribution to the Hall conductivity. Thouless *et al.* [11] realized that this phase can be interpreted in the context of the mathematical theory of topology: The Landau levels can be interpreted as a geometric object which has a nontrivial topological shape. In 1988, Haldane [12] generalized this topological interpretation to arbitrary electronic bands, opening up the possibility for a realization of the quantum Hall effect even in the absence of a magnetic field. These developments form the starting point for the field of topological matter, which studies electronic states from the viewpoint of their topological structure. A common feature of topologically nontrivial phases is that they have a deep connection between this geometric shape of electronic bands in the bulk of the material and conducting edge states.

Another fascinating property of crystalline systems is that they form low-energy excitations which are mathematically similar to actual or hypothetical elementary particles. A well-known example is produced when an electron is missing from its usual position, leaving behind a *hole* of positive charge. These so-called quasiparticles appear as emergent phenomena of the collective behavior of electrons. In many cases, describing the physical properties of solids is more intuitive in this quasiparticle picture. For example, both electrons and holes appear to have a certain *effective mass* which is different from the intrinsic electron mass. Of course the electrons do not actually acquire or lose mass, but describing an electron of a different mass is mathematically more elegant than modeling the complex interactions between electrons and nuclei. In topological semimetals, this notion of low-energy excitations forming quasiparticles is joined with the description of electronic states from the viewpoint of topology. They host quasiparticles which are not only the condensed-matter equivalent of fundamental particles, but also form – and are protected by – a nontrivial topological phase.

The topics covered in this thesis fall into one of two categories: First, the study of specific materials and topological semimetal phases. And second, the creation of techniques, algorithms and tools for simplifying and enhancing this study of topological semimetals. The thesis is structured as follows: In chapter 2, we give a brief description of band theory, which is the general physical context in which topological phases are subsequently described. We discuss numerical approaches to solving this problem, namely density functional theory, tight-binding models, and the  $\mathbf{k} \cdot \mathbf{p}$  approximation. In chapter 3, we dive into the topological aspects of band theory. We motivate the topological nature of electronic bands from its mathematical origin, and describe how a topological invariant called the *Chern number* can be used as a basis for identifying topological phases. Concluding the introductory part of this thesis, chapter 4 describes topological states in the context of semimetals. As an example of such a topological semimetal state, we describe the Weyl phase. We also review recent advances in classifying topological semimetal phases, focusing

mainly on our discovery of the type-II Weyl phase. Departing slightly from the main topic of the thesis, in chapter 5 we introduce an algorithm for computing discrete phase diagrams. We show how this algorithm can reduce the number of phase evaluations required by focusing on the phase boundaries, and describe its convergence properties. Returning to the topic of topological semimetals, in chapter 6 we show how an external magnetic field can produce Weyl nodes in a previously topologically trivial materials. In particular, we study a  $\mathbf{k} \cdot \mathbf{p}$  model of TaAs<sub>2</sub> and a tight-binding model of NbSb<sub>2</sub>. By employing the algorithm introduced in chapter 5, we compute phase diagrams for the number of Weyl nodes as a function of applied magnetic field. Chapters 7 and 8 introduce techniques for simplifying the study of topological materials, as well as the band structure problem in general. In chapter 7, we present a method for automatically generating the general form of a  $\mathbf{k} \cdot \mathbf{p}$  model in the presence of symmetry. By reformulating this problem in the language of linear algebra, we enable solving it with the help of a computer algebra system. The tight-binding approach to solving the electronic structure problem is the topic of chapter 8. We present automated workflows for generating tight-binding models from first-principles calculations. Building upon the technique of Wannier tight-binding models, we introduce a method for symmetrizing these models based on performing a group average. Furthermore, we employ local optimization to automatically choose energy input parameters required by the Wannier tight-binding procedure. Finally, in chapter 9 we describe an algorithm for automatically finding and identifying nodal features. Using this procedure, we classify two known topological semimetal, as well as five novel candidate materials.



# 2

## A Short Introduction to Band Theory

Parts of this chapter were previously published in ref. [8]

Band theory is an approximate way of describing the electronic states in regular, crystalline materials. The topological phenomena discussed in this thesis are founded on this theory, and so we give a brief introduction to band theory in this introductory chapter. This description is far from complete, focusing instead on the parts which are required for the understanding of the topics covered in this thesis. A more thorough description can be found in any solid state physics textbook. In the first section of this chapter, we discuss the definition and symmetries of crystalline systems. Next, we show how the electronic states are defined in the context of weak interactions. And finally, we describe the numerical methods used throughout this thesis to obtain a description of these electronic states.

### 2.1 Periodic Lattices

Atoms in solids often arrange themselves in regular, repeating patterns, forming what is known as a *crystal structure*. The unit of periodicity in such materials is a parallelepiped spanned by vectors  $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ , the *unit cell*. Within each unit cell, the relative positions of atoms are the same. While the choice of unit cell is not unique, there is a unique minimum volume that a unit cell must have. Such a unit cell with minimal volume is called a *reduced* unit cell. A perfect, infinite crystal is symmetric under any translation by a *lattice vector*  $\mathbf{R}$  defined as

$$\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3, \quad (2.1)$$

where  $n_i$  are integers. Due to this symmetry, the Hamiltonian  $\mathcal{H}$  describing the crystalline system and all translation operators  $\hat{T}_{\mathbf{R}}$  commute and are simultane-

ously diagonalizable. In this diagonal basis, the eigenstates of the translation operators are given by

$$\hat{T}_{\mathbf{R}} |\psi_{\mathbf{k}}\rangle = e^{-i\mathbf{k}\cdot\mathbf{R}} |\psi_{\mathbf{k}}\rangle, \quad (2.2)$$

where  $\mathbf{k}$  is a vector of quantum numbers. The physical interpretation of  $\mathbf{k}$  is as a crystal wave-vector. It is unique only up to translations in the so-called *reciprocal* lattice that is spanned by  $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$ , which are defined by

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{i,j}. \quad (2.3)$$

In addition to the purely translational symmetries, crystal structures can have further symmetries such as rotations, reflections, screw axes or glide planes. These symmetries can be written in the form

$$g = \{S|\mathbf{t}\}, \quad (2.4)$$

where  $S$  is a matrix operation (rotation, reflection or rotoreflection), and  $\mathbf{t}$  is a translation vector. Using this notation, the symmetry acts on a general position  $\mathbf{r}$  as

$$g\mathbf{r} = S\mathbf{r} + \mathbf{t}. \quad (2.5)$$

Based on their symmetry group, three-dimensional crystal structures are classified into one of 230 space groups. If the symmetry group of a material (excluding pure translations) can be expressed using only the matrix  $S$  for any choice of origin, it is called a *symmorphic* group. If on the other hand some symmetries always contain a non-zero translation vector, as is the case for screw axes or glide planes, it is called non-symmorphic. The symmetries present in a crystal greatly influence its physical properties, as we shall see throughout this thesis.

Of course, we have considered an idealized crystal structure so far. In real materials, the presence of defects in the crystal structure can influence their physical properties. Another such effect comes from the finite size of samples. Since the typical side length of a unit cell is a few Ångström, this is however usually negligible in macroscopic samples.

## 2.2 Bloch States

In general, the physics of solids is determined by the movement of both nuclei and electrons, as well as their interactions. However, nuclei are orders of magnitude heavier than electrons. When treating the electronic configuration, we can thus use the Born-Oppenheimer approximation, which assumes the nuclei to stay fixed. Instead of considering the full  $n$ -electron problem, we make another approximation to consider the electron-electron interaction as an average potential acting on a

single particle. Surprisingly, this *mean field* approximation of electronic states still gives a reasonable description for many systems. This approximation however can not describe effects which arise from strongly correlated electrons.

In this non-interacting limit, electronic states in crystalline materials can be described by a single-particle Hamiltonian  $\mathcal{H}(\mathbf{k})$ , which is a smooth function of the crystal wave-vector  $\mathbf{k}$ . The possible electronic states are given by the solutions of the time-independent Schrödinger equation

$$\mathcal{H}(\mathbf{k}) |\psi_{n,\mathbf{k}}\rangle = \epsilon_{n,\mathbf{k}} |\psi_{n,\mathbf{k}}\rangle. \quad (2.6)$$

These so-called *Bloch states*  $|\psi_{n,\mathbf{k}}\rangle$  are a superposition of plane waves with wave-vector  $\mathbf{k}$ . As such, they can be written as

$$|\psi_{n,\mathbf{k}}\rangle = e^{i\mathbf{k}\cdot\mathbf{r}} |u_{n,\mathbf{k}}\rangle, \quad (2.7)$$

where  $|u_{n,\mathbf{k}}\rangle$  is *cell-periodic*. This property is known as the Bloch theorem [13]. The energy eigenvalues  $\epsilon_{n,\mathbf{k}}$  are called *energy bands*, with  $n$  being their band index.

The bulk properties of materials are determined largely by their band structure. For example, a material is insulating if there is an energy gap between the eigenstates which are occupied by electrons, and those that are empty, as shown in fig. 2.1(a). Conversely, a material is conducting if there is no such energy gap, such as when an energy band is only occupied for certain values of  $\mathbf{k}$  as shown in fig. 2.1(b).

As we shall see in the next chapter, the band structure does not completely capture the physics of a given material. Instead, taking into account the shape of the Bloch states  $|\psi_{n,\mathbf{k}}\rangle$  leads to a topological classification of materials. A common feature of such topological phases is that they exhibit surface states, which are protected in such a way that they cannot be removed by small perturbations.

## 2.3 Numerical Methods

Before diving into the topological properties of band structures, we discuss some common numerical methods used to treat the electronic structure problem. First, we discuss density functional theory, a method for calculating the electronic structure directly from the atomic configuration. We then discuss two methods for constructing a simplified effective model, the tight-binding and  $\mathbf{k} \cdot \mathbf{p}$  methods.

### 2.3.1 Density Functional Theory

Density functional theory (DFT) [14, 15] is the most commonly used technique for calculating the electronic structure of real materials in the weakly interacting

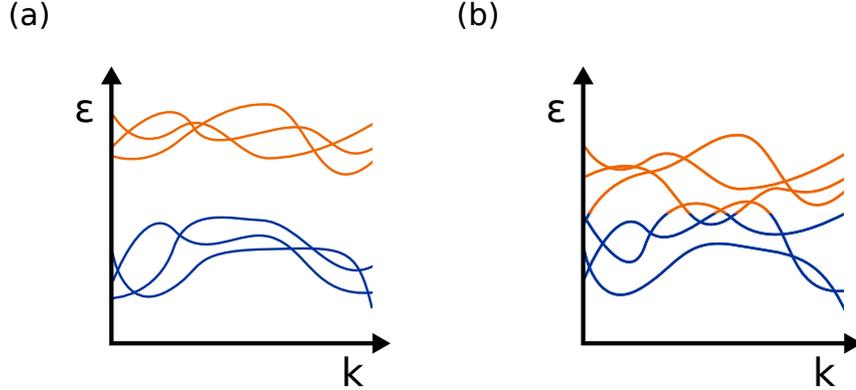


Figure 2.1: Band structures along a one-dimensional line in  $\mathbf{k}$ -space. (a) Band structure of an insulating material. Occupied (blue) and unoccupied (orange) states are separated by an energy gap for all  $\mathbf{k}$ . (b) Band structure of a conducting material. Occupied and unoccupied states touch, and some bands are partially occupied.

limit. It is based on theorems by Hohenberg and Kohn [14], which state that a *universal* density functional  $F[\rho]$  exists such that for any external potential  $V_{\text{ext}}$ , the energy functional

$$E[\rho] = F[\rho] + \int V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) \, d^3\mathbf{r} \quad (2.8)$$

is minimized exactly when  $\rho$  is the ground state density  $\rho_0$ . The minimum energy  $E_0 = E[\rho_0]$  is the ground state energy.

While this description in principle gives a simple way of solving the electronic structure problem by minimizing a functional, the fundamental problem of DFT is that the functional  $F[\rho]$  is not known. A common ansatz [15] used to tackle this problem is to explicitly write the contributions to  $F$  from kinetic energy  $T$  and Coulomb repulsion, and define an “exchange correlation” term  $E_{xc}$  to contain the missing terms:

$$F[\rho] = T[\rho] + \frac{e^2}{2} \iint \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, d^3\mathbf{r} \, d^3\mathbf{r}' + E_{xc}[\rho]. \quad (2.9)$$

This approach obviously just shifts the problem to finding an exact form of  $E_{xc}$ . Different approximate methods such as the local density approximation (LDA) [14], generalized gradient approximation (GGA) [15, 16], and methods including a Hartree-Fock term [17–20] have been developed to treat this issue. The advantage of the formulation in terms of eq. (2.9) is that it can be interpreted as a

single-particle density in an effective potential [15]

$$V_{\text{eff}}(\mathbf{r}) = V_{\text{ext}} + e^2 \int \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' + \frac{\delta E_{xc}[\rho]}{\delta \rho(\mathbf{r})}. \quad (2.10)$$

For a given  $V_{\text{eff}}$ , the single-particle wave function can be obtained by solving the Schrödinger equation

$$\left( -\frac{1}{2}\nabla^2 + V_{\text{eff}}(\mathbf{r}) \right) \psi_i(\mathbf{r}) = \varepsilon_i \psi_i(\mathbf{r}), \quad (2.11)$$

from which the density can be obtained via

$$\rho(\mathbf{r}) = \sum_i |\psi_i(\mathbf{r})|^2, \quad (2.12)$$

where the sum goes only over occupied states. Since the density itself is also part of the effective potential, these equations need to be solved in a self-consistent way. This can be done by choosing an initial density  $\rho$  and then iteratively updating it using eqs. (2.10) to (2.12) until it converges. This is known as the Kohn-Sham scheme [15].

Using DFT, it is possible to solve the electronic structure problem from *first principles*, requiring only the atomic configuration as input. However, the computational cost of this approach is large, since all electronic states need to be calculated. Instead, one can use so-called *pseudopotentials*, which describe the effect of an atomic nucleus and its core electrons. While this reduces the computational cost, the construction of effective pseudopotentials which are transferable (applicable for many different materials) is not an easy task. Often, experimental data is used to improve the accuracy of pseudopotentials. Thanks to various efforts however [21–29], nowadays high-quality pseudopotentials are readily available.

Throughout this thesis, we use DFT as a tool to obtain the electronic structure of real materials, using well-established codes [30–34].

### 2.3.2 Tight-Binding Models

Tight-binding models provide an intuitive, real-space description of crystalline materials. They consist of localized orbitals  $\{|\phi_{1,\mathbf{R}}\rangle, \dots, |\phi_{N,\mathbf{R}}\rangle\}$ , centered at positions  $\mathbf{t}_1, \dots, \mathbf{t}_N$  inside the unit cell given by the lattice vector  $\mathbf{R}$ . The movement of electrons between these orbitals is described by so-called “hopping terms”. For a tight-binding model with Hamiltonian  $\mathcal{H}$ , the transition probability from  $|\phi_{j,\mathbf{R}}\rangle$  to  $|\phi_{i,\mathbf{R}'}\rangle$  is given by

$$\langle \phi_{i,\mathbf{R}'} | \mathcal{H} | \phi_{j,\mathbf{R}} \rangle = \langle \phi_{i,\mathbf{0}} | \mathcal{H} | \phi_{j,\mathbf{R}-\mathbf{R}'} \rangle = H^{ij}[\mathbf{R} - \mathbf{R}']. \quad (2.13)$$

This relation defines the real-space Hamiltonian matrix  $H[\mathbf{R}]$ . The matrix element  $H_{i,i}[\mathbf{0}]$  describes the so-called *on-site* energy of the orbital  $|\phi_i\rangle$ . If we define an extended Bloch-like basis as<sup>1</sup>

$$|\psi_{i,\mathbf{k}}\rangle = \frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot(\mathbf{R}+\mathbf{t}_i)} |\phi_{i,\mathbf{R}}\rangle, \quad (2.14)$$

where  $N$  is the number of lattice sites<sup>2</sup>, we can write the  $\mathbf{k}$ -space Hamiltonian matrix as

$$\begin{aligned} \mathcal{H}^{ij}(\mathbf{k}) &= \langle \psi_{i,\mathbf{k}} | \mathcal{H} | \psi_{j,\mathbf{k}} \rangle = \frac{1}{N} \sum_{\mathbf{R}, \mathbf{R}'} e^{-i\mathbf{k}\cdot(\mathbf{R}'+\mathbf{t}_i)} \langle \phi_{i,\mathbf{R}'} | \mathcal{H} | \phi_{j,\mathbf{R}} \rangle e^{i\mathbf{k}\cdot(\mathbf{R}+\mathbf{t}_j)} \\ &= \frac{1}{N} \sum_{\mathbf{R}, \mathbf{R}'} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}'+\mathbf{t}_j-\mathbf{t}_i)} H^{ij}[\mathbf{R}-\mathbf{R}'] = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot(\mathbf{R}+\mathbf{t}_j-\mathbf{t}_i)} H^{ij}[\mathbf{R}], \end{aligned} \quad (2.15)$$

where we replaced  $\mathbf{R}-\mathbf{R}'$  with  $\mathbf{R}$  in the last step. The tight-binding description can be interpreted as an approximation of the  $\mathbf{k}$ -space Hamiltonian using Fourier coefficients. Including longer-range hopping terms increases the accuracy of this approximation, as does the inclusion of more (possibly unoccupied) orbitals.

In contrast to DFT, tight-binding is *not* a first-principles method. As such, the hopping parameters need to be obtained from a different method. This can be either a theoretical consideration, creating a toy model, or a first-principles calculation to describe realistic materials. The latter is the subject of chapter 8.

### 2.3.3 $\mathbf{k} \cdot \mathbf{p}$ Models

Finally, another way of approximating the  $\mathbf{k}$ -space Hamiltonian is the so-called  $\mathbf{k} \cdot \mathbf{p}$  [37] method. Whereas the tight-binding method approximates  $\mathcal{H}(\mathbf{k})$  with a Fourier expansion, the  $\mathbf{k} \cdot \mathbf{p}$  method does so using a Taylor series around a chosen point  $\mathbf{k}_0$ . The  $\mathbf{k} \cdot \mathbf{p}$  model up to  $n$ -th order can be written as

$$\mathcal{H}(\mathbf{k}) = \sum_{\substack{\alpha, \beta, \gamma \in \mathbb{N}_0 \\ \alpha + \beta + \gamma \leq n}} H^{\alpha\beta\gamma} (k_x^*)^\alpha (k_y^*)^\beta (k_z^*)^\gamma, \quad (2.16)$$

where  $\mathbf{k}^* = \mathbf{k} - \mathbf{k}_0$  is the relative  $\mathbf{k}$ -vector. A notable consequence of this approximation is that the model is no longer periodic in reciprocal space.<sup>3</sup> Since

---

<sup>1</sup>In ref. [35], this is defined as convention I. Unless otherwise specified, this is the convention we use throughout this thesis. Convention II does not explicitly take into account  $\mathbf{t}$ , and is more efficient for computing the energy eigenvalues. As such, it is used as a default in our TBmodels [36] code.

<sup>2</sup>For simplicity we assume the system to be finite here. This allows for creating a normalized Bloch-like basis.

<sup>3</sup>The model can be made periodic, but only at the cost of introducing a discontinuity.

$\mathbf{k} \cdot \mathbf{p}$  models are computationally very cheap, they are often used when only the behavior in the vicinity of  $\mathbf{k}_0$  is important. In chapter 7, we show how  $\mathbf{k} \cdot \mathbf{p}$  models can be constructed in such a way that they respect the crystal symmetries.



# 3

## Topological Properties of Electronic Bands

Parts of this chapter were previously published in ref. [8]

In the previous chapter, we have identified the band structure as a fundamental tool for determining the properties of weakly interacting materials. Here, we extend upon this view by taking into account the shape of the Bloch states themselves. In particular, this gives rise to a *topological* classification of materials. The defining feature of such topological phases is that their Bloch states cannot be smoothly transformed into each other. We start by motivating and defining this notion of topological phases. Next, we give an example of such a phase and show how it can be identified numerically. Finally, we show how topological classification can be enriched in the presence of symmetries, using the example of time-reversal symmetry.

### 3.1 Topology in Non-interacting Materials

In this first section, we will introduce the notion of topological properties in the context of non-interacting materials. From their basic definition, we will see that topological phases must exhibit some interesting physical phenomena.

#### 3.1.1 Topological Properties

To motivate the concept of topological classification, we first show an example from its mathematical origins in geometry: Closed, orientable two-dimensional surfaces can be classified by their number of holes, called *genus*. A sphere for example has

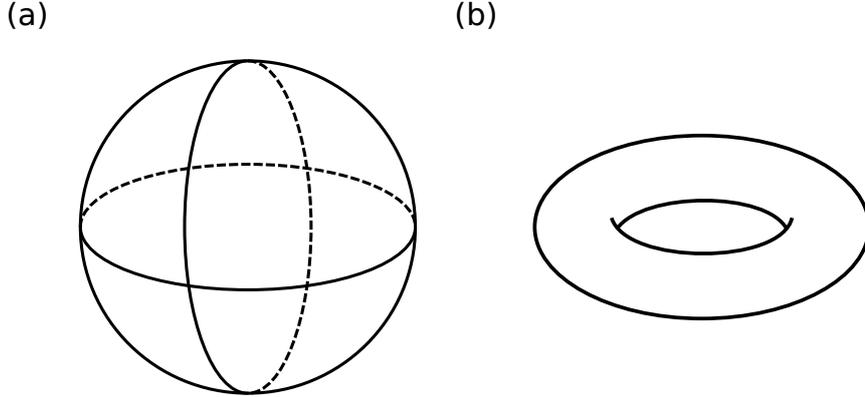


Figure 3.1: Examples of closed orientable surfaces: (a) A sphere has no holes (b) A torus has one hole

no holes, while a torus has exactly one (see Fig. 3.1). This property is conserved under smooth deformations of the surface. The only way to add or remove a hole is by tearing and gluing the surface. The genus is an example for a topological invariant – a quantized property that cannot be changed without changing the topological phase. For this reason, topological invariants are commonly used to *identify* topological phases.

In order to define topological phases for materials, we need a geometric object on which the topological properties can be defined. For this purpose, we pick a set of bands  $B$ . A very common choice for  $B$  is to pick the occupied subspace<sup>1</sup>. The set of states  $\{|u_{n,\mathbf{k}}\rangle\}_{n \in B}$  span a vector space  $V_{\mathbf{k}}$  (over  $\mathbb{C}$ ) for each  $\mathbf{k}$ . If  $V_{\mathbf{k}}$  is a smooth function of  $\mathbf{k}$  and the space where  $\mathbf{k}$  itself is defined is a manifold, this defines a so-called *fiber bundle*.

A simple geometrical example of a fiber bundle is given by a one-dimensional vector space defined on a circle. If the vector space is orthogonal to the plane described by the circle, the resulting object is a cylinder, as shown in Figure 3.2 (a). If however the basis vector rotates by  $\pi$  as it goes around the circle, the resulting object is a Möbius strip. These two objects cannot be smoothly transformed into each other, making them topologically distinct.

### 3.1.2 Bulk-edge Correspondence

In the previous section, the fact that the vector space  $V_{\mathbf{k}}$  needs to be a smooth function of  $\mathbf{k}$  was mentioned. This has a profound impact on the physical properties

---

<sup>1</sup>This is not always possible, as in the case of semimetals where the occupation number changes with  $\mathbf{k}$ . In these cases, one often picks the  $N$  lowest energy bands instead.

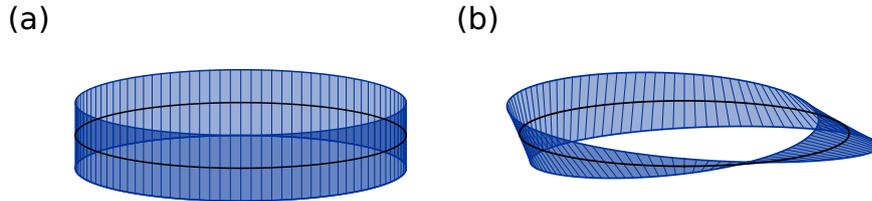


Figure 3.2: (a) A cylinder, spanned by a vector which does not rotate as it goes around a circle. (b) A Möbius strip, spanned by a vector which rotates by  $\pi$  as it goes around a circle.

of topological states, as we shall now see.

Even though the Hamiltonian  $\mathcal{H}(\mathbf{k})$  is a smooth function of  $\mathbf{k}$ , the same is not necessarily true for  $V_{\mathbf{k}}$ . Consider the following one-dimensional example:

$$\mathcal{H}(k) = -\cos(k) |a\rangle\langle a| + \cos(k) |b\rangle\langle b|, \quad (3.1)$$

where  $|a\rangle$  and  $|b\rangle$  are two arbitrary orthogonal states. For  $k = 0$ , the energy eigenvalues of  $|a\rangle$  and  $|b\rangle$  are  $-1$  and  $1$ , respectively. Consequently,  $|a\rangle$  has band index 1 while  $|b\rangle$  has index 2. As  $k$  changes, the energy eigenvalues shift until they are equal at  $k = \pi/2$ . At this point, the vector space  $V_{\mathbf{k}} = \text{span}(\{|u_{n,\mathbf{k}}\}_{n \in \{1\}})$  switches discretely from being spanned by  $|a\rangle$  to being spanned by  $|b\rangle$ . As a result, this space does not meet the criteria for topological categorization.

The smoothness of the vector space  $V_{\mathbf{k}}$  can be broken if the order of energy eigenvalues between the states which are in the set  $B$  and those which are not changes. This can easily be avoided if we restrict our possible choice of bands  $B$ , such that they are always separated from the other bands by a direct energy gap. In other words, topological properties are defined for *isolated* sets of bands, which form smooth fiber bundles.

Another way to frame this is by looking at the possible transformations that can be done to a material without changing its topological properties. In addition to requiring that these transformations smoothly change the Hamiltonian, we impose that the band gap remains open. This definition leads to a remarkable physical property of topological phases: At the boundaries of topologically non-trivial insulating materials, stable conducting edge states must form. In going from the bulk of the topological material to vacuum, the system undergoes a smooth transition from a non-trivial to a trivial (vacuum) state. To allow for this, the aforementioned condition that the bands are separated in energy must be broken. This effect is known as the bulk-boundary correspondence, and variations of this effect govern the interesting transport phenomena to be found in many topological materials [12, 38, 39].

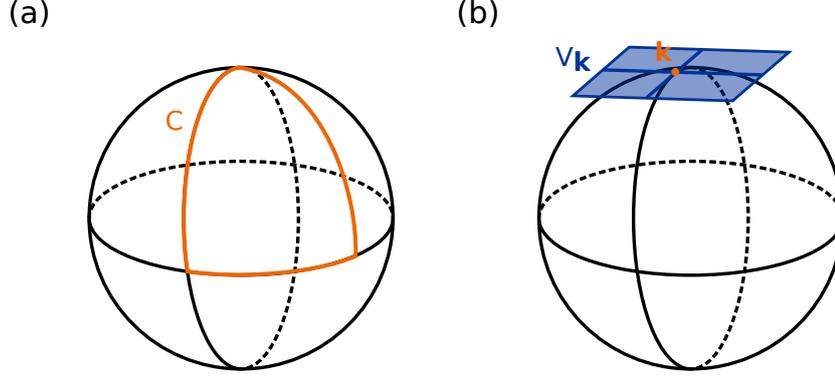


Figure 3.3: (a) A closed path  $\mathcal{C}$  on the surface of a sphere. (b) The tangential vector space  $V_{\mathbf{k}}$  for a given point  $\mathbf{k}$  on a sphere.

## 3.2 The Chern Number

In the previous section, we have seen how topological properties in crystalline materials are defined on a conceptual level. Now, we will show an example for a topological invariant, which can be used to classify many topological phases of matter.

### 3.2.1 The Berry Phase and Chern Invariant

The basis for defining a topological invariant for electronic bands is the notion of a *geometric phase*. To illustrate this phase, imagine a closed loop  $\mathcal{C}$  on a manifold. As an example, we choose a closed loop on a sphere, as shown in Figure 3.3(a). Adding the plane tangential to the sphere at each point gives us a fiber bundle (see Fig. 3.3(b)).

Now we choose a vector in the tangential space, and move it along  $\mathcal{C}$  in such a way that it remains locally parallel to itself, as shown in Figure 3.4. This process is called *parallel transport*. We observe that the vector is rotated by some angle  $\phi$  as it traverses the path  $\mathcal{C}$ . Since this angle depends only on the geometry of the fiber bundle, it is called a geometric phase.

For electronic bands, such a phase, known as *Berry's phase*<sup>2</sup>, can be written as [40]

$$\gamma_{\mathcal{C}} = i \oint_{\mathcal{C}} \sum_{n \in B} \langle u_{n,\mathbf{k}} | \nabla_{\mathbf{k}} | u_{n,\mathbf{k}} \rangle \cdot d\mathbf{k}, \quad (3.2)$$

---

<sup>2</sup>For simplicity, we consider the *total* Berry phase of all bands. The Berry phase can also be defined for a single band, in which case the sum over bands is dropped.

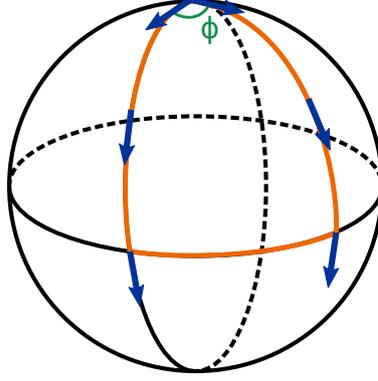


Figure 3.4: Parallel transport of a vector on a closed path on a sphere rotates the vector by an angle  $\phi$ .

where  $\mathcal{C}$  is a closed loop in reciprocal space. Unlike the example above, the Berry's phase represents a rotation in the complex phase of a vector, not its real-space direction.<sup>3</sup> It is gauge-invariant up to multiples of  $2\pi$  [40]. By defining the *Berry potential* (or Berry connection)

$$\mathcal{A}(\mathbf{k}) = i \sum_{n \in B} \langle u_{n,\mathbf{k}} | \nabla_{\mathbf{k}} | u_{n,\mathbf{k}} \rangle, \quad (3.3)$$

the Berry phase can be re-written as

$$\gamma_{\mathcal{C}} = \oint_{\mathcal{C}} \mathcal{A}(\mathbf{k}) \cdot d\mathbf{k}. \quad (3.4)$$

Note that unlike the Berry phase, the Berry potential is not a gauge-invariant quantity. If the Berry potential is a smooth function of  $\mathbf{k}$  (an important prerequisite, as we shall see soon), we can use Stokes' theorem to rewrite the Berry phase as a surface integral

$$\gamma_{\mathcal{C}} = \int_S \nabla_{\mathbf{k}} \wedge \mathcal{A}(\mathbf{k}) \cdot d\mathbf{k}, \quad (3.5)$$

where  $\mathcal{C} = \partial S$ . Introducing the *Berry curvature*

$$\mathcal{F} = \nabla_{\mathbf{k}} \wedge \mathcal{A}(\mathbf{k}), \quad (3.6)$$

which is again gauge-invariant, we can write this as

$$\gamma_{\mathcal{C}} = \int_S \mathcal{F}(\mathbf{k}) \cdot d\mathbf{S}. \quad (3.7)$$

<sup>3</sup>To see this, try calculating the Berry phase for  $|u_k\rangle = e^{ik/2} \begin{pmatrix} \cos(k) \\ \sin(k) \end{pmatrix}$ , for  $k \in [0, 2\pi]$ .

For a closed, orientable two-dimensional surface  $S$  in reciprocal space, we can now define the Chern invariant as [41, 11]

$$C = \frac{1}{2\pi} \int_S \mathcal{F}(\mathbf{k}) \cdot d\mathbf{S}. \quad (3.8)$$

Since the edge of a closed surface is a trivial path, equation 3.4 seems to suggest that the Chern number is always zero. We must now remember however that equations 3.5 and 3.7 are valid only if the Berry potential  $\mathcal{A}(\mathbf{k})$  is smooth. Previously, we discussed that  $V_{\mathbf{k}}$  spanned by  $|u_{n,\mathbf{k}}\rangle$  must be a smooth function of  $\mathbf{k}$  if we wish to define a topological classification. However, there can still be a winding in the phase of  $|u_{n,\mathbf{k}}\rangle$  which makes the Berry potential non-smooth. As a result, the Chern number can take any integer value. In fact, the presence of a non-zero Chern number can be viewed as a topological obstruction to finding a globally smooth gauge [42, 43].

#### 3.2.2 The Chern Number as Change in Berry Phase

Having defined the Chern number in terms of the cell-periodic states  $|u_{n,\mathbf{k}}\rangle$ , we will now show an alternative form that is easier to calculate numerically and is used within the `Z2Pack` code [6]. For simplicity, we will look at the example where  $S$  is the Brillouin zone  $\mathbf{k} \in [0, 1)^2$  of a two-dimensional material, in reduced coordinates<sup>4</sup>. The results are equally applicable to other closed two-dimensional surfaces.

We divide the surface integral (eq. 3.8) for the Chern number into small segments  $S^i$ , as shown in Fig. 3.5(a). The segments should be small enough that

$$C_{\text{part.}}^i = \frac{1}{2\pi} \int_{S^i} \mathcal{F}(\mathbf{k}) \cdot d\mathbf{S} \quad (3.9)$$

is much smaller than one. The Chern number is then given as the sum of all segment integrals,

$$C = \sum_i C_{\text{part.}}^i. \quad (3.10)$$

Since  $\mathcal{A}(\mathbf{k})$  can be made to be *locally* smooth [44, 45], we can use Stokes' theorem to obtain

$$C_{\text{part.}}^i \mod 1 = \frac{1}{2\pi} \int_{\partial S^i} \mathcal{A}(\mathbf{k}) \cdot d\mathbf{k} \mod 1 = \frac{\gamma_{\partial S^i}}{2\pi} \mod 1, \quad (3.11)$$

---

<sup>4</sup>*Reduced* coordinates refer to coordinates in the (real- or reciprocal-space) lattice basis, as opposed to cartesian coordinates.

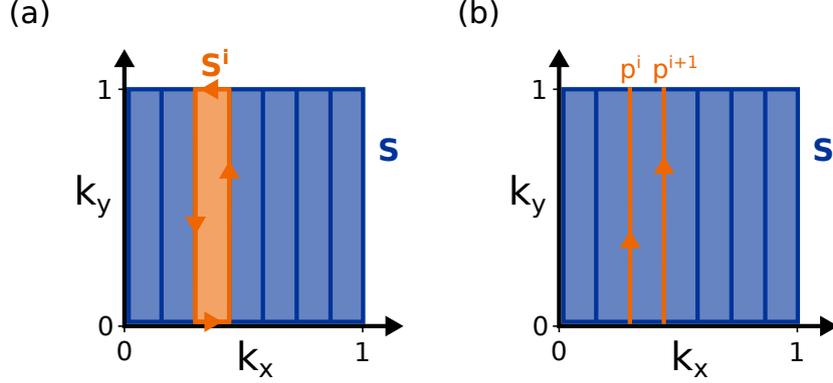


Figure 3.5: (a) The surface  $S$  is divided into segments  $S^i$ . For each segment, the flux of Berry connection can be calculated from the Berry phase around its boundary. (b) The top and bottom paths of each boundary cancel, leaving paths  $p^i$  which cross the Brillouin zone at a constant  $k_x$ .

where the modulus comes from the fact that the Berry phase is defined only modulo  $2\pi$ . Since we imposed that  $C_{\text{part.}}^i$  must be much smaller than one, we can still uniquely determine its value from  $\gamma_{\partial S^i}/2\pi$  by adding an integer that minimizes the absolute value. Since the top and bottom parts of  $\partial S^i$  cancel out due to periodicity, we can write the Berry phase as

$$\gamma_{\partial S^i} = \gamma_{p^{i+1}} - \gamma_{p^i}, \quad (3.12)$$

where  $p^i$  and  $p^{i+1}$  are the paths at either side of the segment  $S^i$ , as shown in Fig. 3.5(b). The Berry phase can also be understood as a function of  $k_x$ , since each path  $p^i$  is given by a fixed  $k_x$ . Because both  $\gamma$  and  $k_x$  are periodic, the Berry phase describes a line on a torus, as shown in Fig. 3.6. The winding number of this line around the torus is exactly the Chern number [46]. In other words, the Chern number can be calculated by *continuously* tracking the Berry phase on lines of constant  $k_x$  as it goes across the Brillouin zone.

### 3.2.3 Wilson Loop and Hybrid Wannier Charge Centers

The problem of calculating the Chern number is now reduced to calculating the Berry phase for closed loops in the Brillouin zone. This can be done by calculating the so-called Wilson loop [47]  $W(C)$ . The Wilson loop can be understood as a matrix that maps the states at a starting point  $\mathbf{k}_0$  along the loop onto their images after parallel transport along  $C$ . For a discretization  $\{\mathbf{k}_0, \dots, \mathbf{k}_{n-1}, \mathbf{k}_n = \mathbf{k}_0\}$  of the

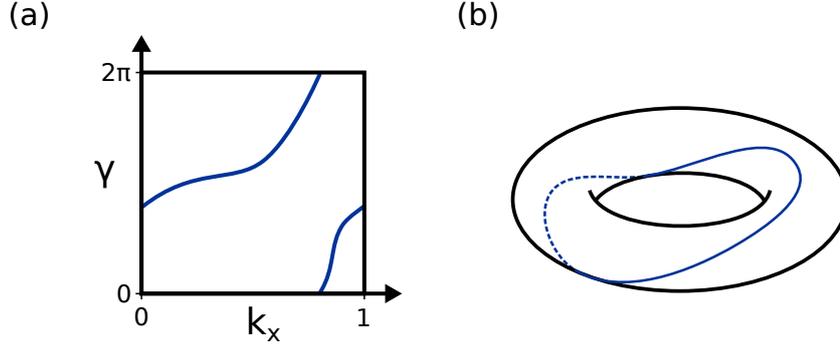


Figure 3.6: (a) The Berry phase  $\gamma$  as a function  $k_x$  for an example system with  $C = 1$ . (b) Because both  $k_x$  and  $\gamma$  are periodic, the Chern number can be seen as the winding number of the Berry phase around a torus.

path  $C$ , the Wilson loop can be approximated as [47][6]

$$W(C) = M^{\mathbf{k}_0, \mathbf{k}_1} \cdot \dots \cdot M^{\mathbf{k}_{n-1}, \mathbf{k}_n}, \quad (3.13)$$

where

$$M_{m,n}^{\mathbf{k}_i, \mathbf{k}_j} = \langle u_{m, \mathbf{k}_i} | u_{n, \mathbf{k}_j} \rangle \quad (3.14)$$

are the overlap matrices between Bloch functions at different  $\mathbf{k}$ . The eigenvalues  $\lambda_i$  of the Wilson loop are connected to the total Berry phase by [48]

$$\gamma_C = \sum_i \arg \lambda_i. \quad (3.15)$$

This reflects the fact that each  $\lambda_i$  is the rotation angle that is acquired by an eigenstate of the Wilson loop as it traverses the path  $C$ . Since the overlap matrices  $M$  can be readily computed, this gives a method for calculating the Chern number numerically. Of course, the convergence of the Wilson loop eigenvalues with respect to the discretization of  $C$  needs to be accounted for.

Another, equivalent, approach to calculating the Berry phase is by computing so-called hybrid Wannier charge centers [49, 50]. This method is based on the notion of Wannier orbitals, which are given by Fourier transforming the Bloch states:

$$|\mathbf{R}n\rangle = \frac{V}{(2\pi)^d} \int_{BZ} e^{-i\mathbf{k}\cdot\mathbf{R}} |\psi_{n, \mathbf{k}}\rangle d\mathbf{k}, \quad (3.16)$$

where  $d$  is the dimensionality of the system, and  $V$  is the unit cell volume. The resulting orbitals are localized, in contrast to the extended nature of the Bloch waves. Since the Bloch states used to construct Wannier orbitals can be changed by

a gauge transformation, the same holds for the Wannier orbitals. Their properties, in particular the localization and position in real space, depend sensitively on this choice of gauge [51]. For the purposes of computing topological invariants, we introduce *hybrid* Wannier orbitals [49, 52], which are Fourier transformed only in one spatial direction and remain extended in the others:

$$|R_x, k_y, k_z; n\rangle = \frac{a_x}{2\pi} \int_{-\pi/a_x}^{\pi/a_x} e^{-ik_x R_x} |\psi_{n,\mathbf{k}}\rangle. \quad (3.17)$$

The average position of such an orbital can be thought of as a function of the remaining reciprocal space variables:

$$\bar{x}_n(k_y, k_z) = \langle 0, k_y, k_z; n | \hat{x} | 0, k_y, k_z; n \rangle. \quad (3.18)$$

This quantity, known as the hybrid Wannier charge center (HWCC) is directly related to the Berry phase:

$$\gamma_C = \frac{2\pi}{a} \sum_n \bar{x}_n, \quad (3.19)$$

where  $C$  is the path along which the hybrid Wannier orbitals were Fourier transformed. Moreover, if the gauge is chosen such that these hybrid Wannier orbitals are maximally localized, the individual HWCC correspond to the eigenvalues of the Wilson loop [6]

$$\bar{x}_i = \frac{2\pi}{a} \arg(\lambda_i), \quad (3.20)$$

up to possible reordering.

This equivalence between hybrid Wannier charge centers and the Berry phase gives rise to a physical interpretation of the Chern number  $C$ . As the momentum ( $k_x$ , in the case of Fig. 3.6) is varied across the Brillouin zone, the average position of the electrons in the orthogonal direction can change. Due to the periodicity of  $k_x$ , it must come back to the same position within the unit cell, but it can change into a different unit cell. This represents a *charge pumping* process, where each cycle of  $k_x$  moves the charge by  $C$  unit cells.

### 3.3 Time-reversal Symmetry: $\mathbb{Z}_2$ Classification

In the previous sections, we have seen how an isolated set of bands can be classified topologically according to their Chern number. Now, we will show how this classification can be enriched in the presence of symmetries. In particular, we will show that time-reversal invariant materials can be classified according to a  $\mathbb{Z}_2$  index. After a theoretical introduction, we describe how the  $\mathbb{Z}_2$  index is computed numerically.

### 3.3.1 Individual Chern numbers

In section 3.1.1, we have seen that topological phases can be defined on manifolds in reciprocal space, if we choose a set of Bloch functions  $\{|u_{n,\mathbf{k}}\rangle\}$  such that they span a smooth vector space  $V_{\mathbf{k}}$ . The most convenient way of achieving this smoothness is to choose an isolated set of bands, which is what we have used so far. This leads to a classification into topological states which can only be adiabatically changed by closing the band gap, and are characterized by the Chern number. However, choosing isolated bands is by no means the only possible way to create a smooth  $V_{\mathbf{k}}$ . For the Hamiltonian of eq. 3.1 for example, we could just pick state  $|a\rangle$  for all  $\mathbf{k}$ .

Here, we aim to find a more complex topological classification by sub-dividing the occupied states into smooth parts. In general, if the Hilbert space  $\mathcal{H}$  of a given problem can be written as a sum of smooth Hilbert spaces,

$$\mathcal{H} = \bigoplus_i \mathcal{H}_i, \quad (3.21)$$

then each of the Hilbert spaces has a well-defined Chern number  $C_i$ . These *individual Chern numbers* [49] sum together to the Chern number of the full Hilbert space:

$$C = \sum_i C_i. \quad (3.22)$$

However, in general these individual Chern numbers do not carry much meaning, since the choice how to split up the Hilbert space is arbitrary. In the presence of a symmetry  $S$  however, the Hilbert space can be split up according to the symmetry eigenvalues. For example, consider a mirror symmetry with eigenvalues  $\pm i$ . On the mirror-symmetric surface,  $S$  and  $H(\mathbf{k})$  commute. Therefore, the Bloch functions  $|u_{n,\mathbf{k}}\rangle$  can be separated into  $+i$  and  $-i$  eigenstates. Both eigenspaces have a well-defined Chern number:

$$C = C_i + C_{-i}. \quad (3.23)$$

This gives rise to a *symmetry-protected* [53–55] topological classification. Materials can have a zero total Chern number, but non-zero individual Chern numbers. Such a topological phase is protected as long as both the band gap remains open and the symmetry is respected. If the symmetry is broken, a mixing of the two eigenspaces can change the topological phase.

Time-reversal symmetry  $\theta$  leads to a particularly interesting and well-known topological classification. Unlike spatial symmetries, it is an anti-unitary symmetry and squares to  $-1$  in the spinful case. As a result, the Bloch functions come in

so-called *Kramers pairs* [38, 39]

$$\begin{aligned}\theta |u_{m,\mathbf{k}}^I\rangle &= |u_{m,\mathbf{k}}^{II}\rangle \\ \theta |u_{m,\mathbf{k}}^{II}\rangle &= -|u_{m,\mathbf{k}}^I\rangle.\end{aligned}\tag{3.24}$$

There is a gauge in which these states are smooth [44, 45], and thus they have well-defined, opposite [50] individual Chern numbers

$$C_m^I = -C_m^{II}.\tag{3.25}$$

Furthermore, the hybrid Wannier charge centers are related by [50]

$$\bar{x}_m^I(k_y) = \bar{x}_m^{II}(-k_y),\tag{3.26}$$

meaning that they are degenerate for the time-reversal invariant lines  $k_y = 0, \pi$ . In order to define a topological invariant, we group the states by their pair indices  $I, II$ . The two groups then have individual Chern numbers

$$C^I = -C^{II}.\tag{3.27}$$

However, these Chern numbers are not gauge invariant. This can be seen by changing the sign of one of the two states:

$$\begin{aligned}|\tilde{u}_m^{II}\rangle &= |u_m^I\rangle \\ |\tilde{u}_m^I\rangle &= -|u_m^{II}\rangle\end{aligned}\tag{3.28}$$

These states still obey eq. 3.24, and the individual Chern number of each state remains the same. Yet the two states have switched their pair indices. As a result, the Chern number  $C^I$  is changed by  $C_m^{II} - C_m^I = 2C_m^{II}$ . Since this re-labeling of Kramers pairs can only ever change the Chern numbers by an even number, a topological invariant can be defined as

$$\mathbb{Z}_2 = C_m^I \pmod{2}.\tag{3.29}$$

In practice, the states do not need to be split by their pair indices to calculate the  $\mathbb{Z}_2$  invariant. Instead, we can use the fact that the hybrid Wannier charge centers must be doubly degenerate at the time-reversal invariant momenta. An arbitrary line between zero and  $\pi$  (dotted green line in Fig. 3.7) will cross an even number of HWCC in the topologically trivial case, and an odd number in the non-trivial case [50]. This principle is used in Z2Pack to calculate the  $\mathbb{Z}_2$  invariant.

When computing the  $\mathbb{Z}_2$  invariant numerically, the challenge using the approach described above is that we cannot uniquely identify hybrid Wannier charge centers.

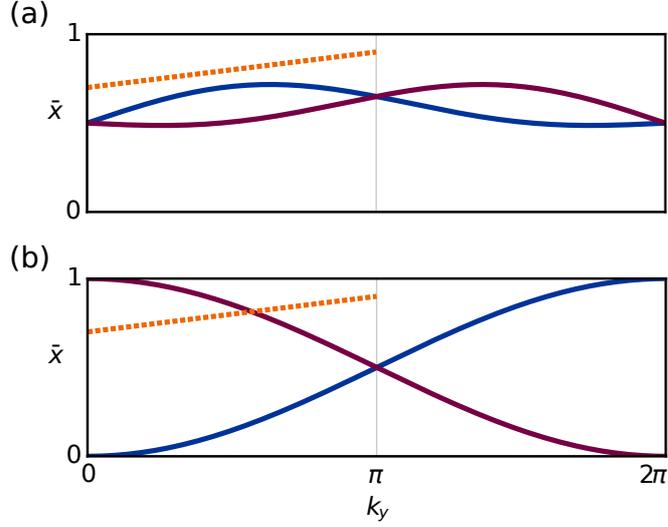


Figure 3.7: Figure previously published in refs. [1, 6]. Hybrid Wannier charge centers for a two-band time-reversal invariant system. (a) Trivial phase. The two bands each have a zero individual Chern number. (b) Non-trivial phase. The two bands have individual Chern numbers  $\pm 1$ .

In other words, we do not know how the HWCC connect between two discrete values of  $k_y$ . We can get around this issue however by choosing the line  $x_{\text{cut}}(k_y)$  for which the number of crossings is counted in a clever way. Since we want a crossing to be as obvious as possible, we choose it to always be in the middle of the largest gap between any two HWCC, as shown in figure 3.8. The number of crossings is then counted by summing up the HWCC which lie between the current and previous value of the largest gap.

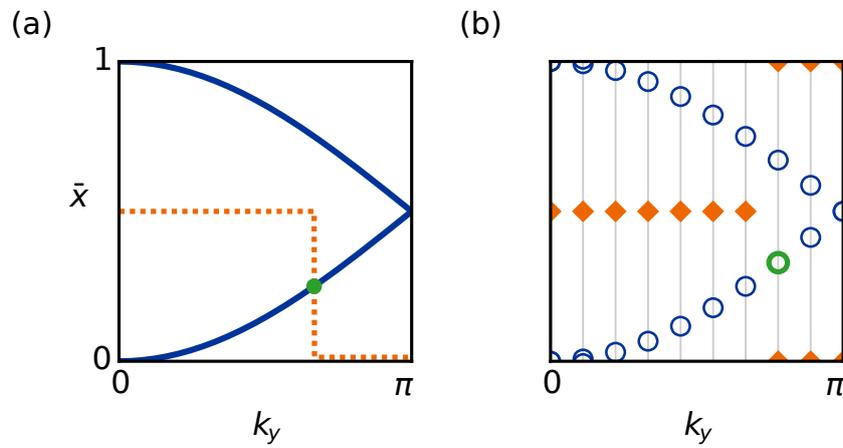


Figure 3.8: Figure previously published in refs. [1, 6]. Sketch showing the  $\mathbb{Z}_2$  calculation. (a) Continuous case. The HWCC (solid blue line) are crossed exactly once by  $x_{\text{cut}}$  (dashed orange line), at the green point. (b) Discrete case. The HWCC (blue circles) and middle of the largest gap (orange diamonds) are known only for discrete  $k_y$ . Crossings are counted when the HWCC value lies between the largest gaps of the current and previous lines (green circle).



# 4

## Type-II Weyl Semimetals

In the previous chapter, we have seen that the Chern number is well-defined on closed surfaces in reciprocal space for an *isolated* set of bands. This definition naturally leads to a topological classification of insulating materials, the Chern [12, 56] and  $\mathbb{Z}_2$  [38, 39] insulators. By no means however is it a necessary requirement for topological classification that a material be insulating. On the contrary, we can use the fact that even for semimetallic materials, a large part of the Brillouin zone still has a direct band gap<sup>1</sup>. In semimetals, then, topological phases usually come in the form of nodal features, such as nodal lines or points. The topological protection then comes in the form of integer invariants on lines or spheres surrounding these nodes. In this chapter, we will first discuss a simple such nodal feature, the Weyl node [57–61]. Next, we will discuss how a change in the Weyl node band structure leads to a second type [2] of Weyl node, and show some material examples for this phase. Finally, we will give a brief overview of other topological semimetal phases.

### 4.1 Weyl Nodes

The Weyl node [57–60] is a linear crossing of two bands in three dimensions. In its simplest form, it can be locally described by the Hamiltonian

$$\mathcal{H}(\mathbf{k}) = \pm \sum_i k_i \sigma_i, \quad (4.1)$$

where  $\sigma_i$  are the Pauli matrices, and the sign the chirality of the node. The two bands which cross at a Weyl node form linear cones, as shown in fig. 4.1. Topo-

---

<sup>1</sup>The bandstructure has a *direct band gap* at a given  $\mathbf{k}$ -point if there is an energy gap between the  $N$ th and  $N + 1$ st bands, where  $N$  is the occupation number.

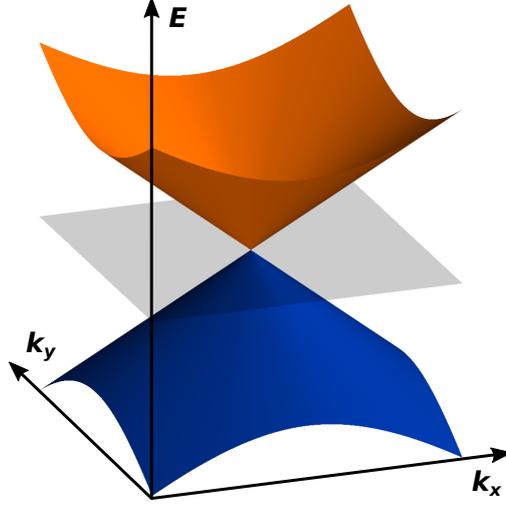


Figure 4.1: The band structure of a rotationally symmetric Weyl node forms two opposing cones (orange, blue) touching at the Weyl node.

logically, a Weyl node of positive chirality acts as a quantized source of Berry curvature, and a negative-chirality node acts as a sink. We can interpret this as a *topological charge*. As a result, Weyl nodes are stable against small perturbations of the system: They can be created or annihilated only in pairs of opposite chirality. This is remarkable insofar as the Weyl node does not require any symmetries to form. Instead, it is a topological feature which forms accidentally but is nevertheless stable.

This topological charge of Weyl nodes can be used to identify them through a topological invariant. Following Gauss' law, the total flux of Berry curvature through a closed surface must be proportional to the total topological charge it encloses. In other words, the Chern number on a sphere surrounding a single Weyl node is plus or minus one, depending on the chirality of the node [2]. We can use this fact to numerically establish the presence of a Weyl node, using the `Z2Pack` code [6].

Since the total topological charge of a material must be zero [62, 63], the sum of Chern numbers for all Weyl nodes in a system must also cancel out. At the very least, Weyl nodes thus come in pairs of two. The potential number and position of Weyl nodes is further influenced by symmetries. Inversion symmetry  $P$  maps a Weyl node  $W$  at position  $\mathbf{k}$  into its partner  $-W$  of opposite chirality, at position  $-\mathbf{k}$ . Time-reversal  $\mathcal{T}$  on the other hand also maps the Weyl node at  $\mathbf{k}$  into one at  $-\mathbf{k}$ , but preserves chirality. As a result, the minimum number of Weyl nodes to exist in a time-reversal symmetric system is four. When the product symmetry

$P*\mathcal{T}$  is present, a Weyl node is mapped back into itself, but with opposite chirality. Consequently, simple Weyl nodes can not exist in systems which observe  $P*\mathcal{T}$ . The topological node which consists of such superimposed Weyl nodes of opposite chirality is called a Dirac node [64–68]. Since the Weyl nodes of opposite chirality could annihilate each other and leave a direct band gap, additional symmetries are required for Dirac nodes to be stable.

The bulk-boundary correspondence introduced in section 3.1.2 can again be used to determine the surface states of a Weyl semimetal. Since the Weyl nodes can annihilate only in pairs, and surrounding vacuum does not contain Weyl nodes, there must be a state on the surface connecting Weyl nodes of opposite chirality. This surface state is called a Fermi arc [69, 70]. Similarly, when a magnetic field is applied in any direction, two Weyl nodes of opposite chirality are connected by a chiral Landau level [71–73]. This so-called chiral anomaly contributes negatively to the magnetoresistance of Weyl semimetals.

## 4.2 Type-II Weyl Semimetals

In the previous section, we have considered only the simplest, rotationally symmetric Weyl node. When considering the general form of linear point crossings between two bands, we find that a second type of Weyl nodes exist [2], as we will show in the following. These *type-II* Weyl nodes exhibit markedly different physics from their type-I counterparts.

The general Hamiltonian for linear crossings of two bands is given by

$$\mathcal{H}(\mathbf{k}) = \varepsilon_0 + \mathbf{k}^T A \boldsymbol{\sigma}, \quad (4.2)$$

where  $\mathbf{k} = (k_x, k_y, k_z)^T$  is position in reciprocal space,  $\boldsymbol{\sigma} = (\sigma_0, \sigma_x, \sigma_y, \sigma_z)^T$  the Pauli vector, and  $A$  a real-valued  $3 \times 4$  coefficient matrix. Of particular importance is the contribution of the  $\sigma_0$  terms. These represent a breaking of continuous rotational symmetry, which is part of the so-called *Lorentz invariance*. They are thus not allowed in high-energy physics. This limitation does not exist for condensed-matter systems, and thus we can create quasiparticles without a high-energy physics equivalent.

Accordingly, we split the coefficient matrix into two parts

$$A = \begin{pmatrix} | & & \\ \mathbf{v} & B^T & \\ | & & \end{pmatrix}, \quad (4.3)$$

where  $\mathbf{v}$  is a vector of length three and contains the  $\sigma_0$  terms, and  $B$  is a  $3 \times 3$  matrix and contains the  $\sigma_{\{x,y,z\}}$  terms. The energy bands of this Hamiltonian are

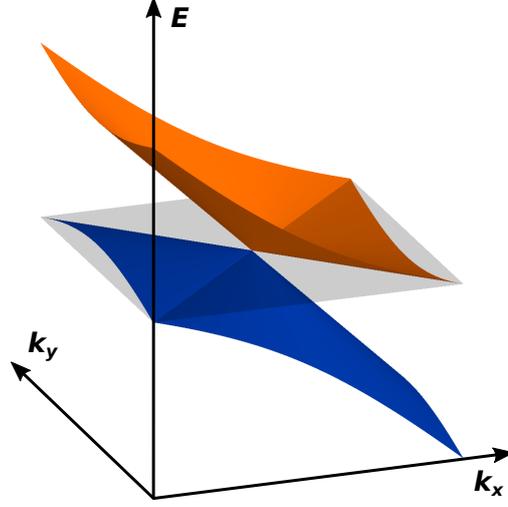


Figure 4.2: In a Weyl node of type II, the upper cone (orange) of the band structure dips below the nodal point energy (gray plane), resulting in an open Fermi surface.

given by a “kinetic” term  $T$  given by  $\mathbf{v}$ , and a “potential” term  $U$  given by  $B$ :

$$\varepsilon_{\pm} = \varepsilon_0 + \mathbf{v}^T \mathbf{k} \pm \sqrt{\mathbf{k}^T B^T B \mathbf{k}} = \varepsilon_0 + T(\mathbf{k}) \pm U(\mathbf{k}) \quad (4.4)$$

While the kinetic term contributes to both bands equally, the potential term is responsible for the splitting between the two bands. By squaring the two terms, we find that the kinetic term dominates when

$$\mathbf{k}^T \mathbf{v} \mathbf{v}^T \mathbf{k} > \mathbf{k}^T B^T B \mathbf{k}. \quad (4.5)$$

This condition can be satisfied – for some  $\mathbf{k}$  – exactly if the symmetric matrix

$$C = B^T B - \mathbf{v} \mathbf{v}^T \quad (4.6)$$

has negative eigenvalues. If this is the case, the Weyl node changes its shape in an interesting way: While the upper band of the regular (or type-I) Weyl node forms an upright cones as shown in fig. 4.1, the “kinetic” term tilts and stretches this cone in such a way that it partially dips below the energy level of the Weyl node itself, as shown in fig. 4.2. Correspondingly, the Fermi surface of this newly discovered [2] type-II Weyl semimetal changes drastically: It goes from being a closed sphere surrounding the Weyl point to being an open surface. Due to this change in the Fermi surface, the chiral Landau level appears only for magnetic field directions where  $T(\mathbf{k}) > U(\mathbf{k})$  [2], and the corresponding negative contribution to magnetoresistance is anisotropic.

It bears noting that, while the type-II Weyl state has markedly different physical properties from its type-I counterpart, the two states are not *topologically* distinct. A type-II Weyl node can smoothly be transformed into a type-I one and vice versa, by tilting the cone.

### 4.2.1 Material Examples

In refs. [2, 3, 5], we have predicted type-II Weyl nodes to exist in WTe<sub>2</sub> [2], MoTe<sub>2</sub> [5], WP<sub>2</sub> and MoP<sub>2</sub> [3] compounds. In this section, we give a brief summary of these results.

The type-II Weyl state was initially discovered in tungsten ditelluride (WTe<sub>2</sub>) [2]. This material crystallizes in an orthorhombic lattice, with space group  $Pmn2_1$  (#31). It has time-reversal symmetry, but not inversion. As a result, the product symmetry  $P * \mathcal{T}$  is broken and Weyl nodes are allowed. Four pairs of Weyl points, related by mirror symmetries, are present in WTe<sub>2</sub>. One of these pairs consists of a negative chirality Weyl node at  $k_1 = (0.1214, 0.0454, 0)$ , 58 meV above the Fermi level, and a positive chirality one at  $k_2 = (0.1218, 0.0382, 0)$ , 52 meV above the Fermi level. The type-II nature of these Weyl nodes comes from the fact that they are formed at a touching point of electron and hole pockets.

It is noteworthy that the Weyl points of opposite chirality are very close to each other in WTe<sub>2</sub>. As a result, they can easily be annihilated by applying strain, for example under tensile strain in  $\hat{x}$  - direction. Compressive strain in the same direction on the other hand separates the Weyl points further, moving the Weyl node initially located at  $k_2$  moves towards the  $k_y = 0$  plane. At roughly 2% compressive strain, this Weyl node meets its mirror partner at  $k_y = 0$  and annihilates, leaving only four Weyl nodes in total – two symmetry-related pairs. As discussed in section 4.1, this is the minimum number of Weyl nodes which is allowed under time-reversal symmetry.

In the structurally and chemically similar compound molybdenum ditelluride (MoTe<sub>2</sub>), this arrangement of four Weyl nodes in the  $k_z = 0$  plane is exactly what we found [5], in a previously unreported orthorhombic  $\gamma$  configuration. However, in MoTe<sub>2</sub> the arrangement of Weyl nodes is also very sensitive to details in the crystal structure. Before the work of ref. [5], a slightly different orthorhombic configuration of MoTe<sub>2</sub> was already predicted to host type-II Weyl nodes [74]. In this structure, it hosts eight in-plane [74] and an additional 16 out-of-plane [5] Weyl nodes. As such, both WTe<sub>2</sub> and MoTe<sub>2</sub> lie on a cusp of topological phase transitions. Nevertheless, spectroscopic evidence [75–78] supports the existence of Weyl nodes in MoTe<sub>2</sub>.

A more robust type-II Weyl state is present in the transition metal diphosphides MoP<sub>2</sub> and WP<sub>2</sub> [3]. Again, eight Weyl points are present in the  $k_z = 0$  plane, in pairs related by mirror symmetry. However, unlike the case of WTe<sub>2</sub> where the nodes in each pair have opposite chirality, these are pairs of Weyl nodes with the

*same* chirality. As a result, they can not easily be annihilated. Spectroscopic and transport evidence for the existence of this type-II Weyl state was also found [79].

## 4.3 Other Topological Semimetals

Spurred by the realization that the limits of Lorentz invariance do not apply to condensed matter systems, the recent years have seen the discovery of a plethora of novel topological phases in semimetals [80–84]. For example, Winkler et al. [80] discovered a threefold degenerate nodal point in  $\text{InAs}_{0.5}\text{Sb}_{0.5}$ . This nodal point occurs along a high-symmetry line, where a single band crosses a two-fold degenerate one. Even higher degeneracies are possible on high-symmetry points, where Bradlyn et al. [81] have discovered three-, six-, and eight-fold degenerate topological nodes.

Apart from the point nodes which exhibit a topological charge, *line* nodes can also be topologically protected [83, 85–88]. In line nodes, the topological nature can be determined by calculating the Berry phase on a closed loop  $\mathcal{C}$  linking the nodal line. While the Berry phase on a closed loop is not quantized in general, the topological phase is revealed when contracting  $\mathcal{C}$  while keeping it linked with the nodal line. If there is no nodal line passing through  $\mathcal{C}$ , the Berry phase must approach zero as  $\mathcal{C}$  approaches a trivial path. If there *is* a nodal line passing through  $\mathcal{C}$  on the other hand, the Berry phase can approach a non-zero value. The reason for this is that  $\mathcal{C}$  cannot be fully contracted without passing a node, where the Berry phase is undefined. As a result, the nodal line can not be gapped locally – instead, it must be fully contracted to be removed.

Recently, Bzúšek et al. [84] have shown that such nodal lines can not only form accidentally, but also be enforced by crystal symmetries. In fact, these nodal lines form chains spanning across the Brillouin zone. In section 9.3.1, we will briefly discuss such a nodal chain semimetal.

# 5

## An Algorithm for Calculating Phase Diagrams

Before continuing to describe how topological semimetal states can arise in real materials, we digress to introduce an algorithm for computing phase diagrams. This is a common task in computational physics, which can naively be achieved by evaluating the phase on a regular mesh of points in the phase space. However, for an  $n$  - dimensional phase diagram this approach scales with  $\mathcal{O}(1/\Delta x^n)$  in the discretization step size  $\Delta x$ . Since the essential features of a discrete phase diagram are the boundaries between phases, the scaling can be reduced to  $\mathcal{O}(1/\Delta x^{n-1})$  by concentrating the evaluation of the phase around these boundaries<sup>1</sup>. In this chapter, we present such an algorithm and its implementation, and discuss its convergence properties. An example application of the algorithm is shown in the subsequent chapter.

### 5.1 Problem Description

In the following, we consider an  $n$  - dimensional phase space  $S$ , consisting of discrete phases  $S = P_1 \cup P_2 \cup \dots \cup P_N$ . We assume there is an evaluation function which can uniquely identify the phase of a given point  $p \in S$ . This condition is sufficient for running the phase diagram algorithm as described in section 5.2. For the discussions of the computational complexity (section 5.2.1) and convergence characteristics (section 5.3) of the algorithm, we will further make the following

---

<sup>1</sup>For the special case of one-dimensional problems, the algorithm reduced to the well-known bisection method and scales with  $\mathcal{O}(\log(1/\Delta x))$ .

assumptions:

1. Each phase has dimension  $n$  everywhere, that is the phase does not have any lower-dimensional parts.
2. For each of the connected components  $Q_i^j$  of  $P_i = Q_i^1 \cup \dots \cup Q_i^M$ , the interior  $Q_i^j - \partial Q_i^j$  is also connected.
3. The phase boundaries  $\partial Q_i^j$  are of dimension  $n - 1$ . Note that this excludes fractals.

## 5.2 Description of the Algorithm

The algorithm for calculating phase diagrams is based on the idea of a quadtree structure: Initially, the phase space is covered by boxes ( $n$  - dimensional hyper-rectangles) of equal size. These are then recursively split into smaller boxes. To decide which boxes should be split, each box is assigned a phase according to the evaluated points it contains. If the points all have the same phase, the box is considered to be of that phase. If there are points of different phase within the box on the other hand, the phase of the box is considered undecided. These undecided boxes are split, if they have not reached a given minimum size.

When splitting a box, the phase is evaluated at its corners and midpoint<sup>2</sup>. Since these points could also be contained in any of the neighboring boxes, the phases of these boxes need to be updated. This step is crucial in creating the convergence properties shown in section 5.3, since it allows boxes to be split even if they were initially found to have a definite phase, as shown in fig. 5.1(c). To make this process efficient, each box needs to maintain a list of neighboring boxes, which also needs to be updated when splitting a box.

Given these rules for splitting boxes, running the algorithm is simple: First, the whole phase space is covered in regular boxes, which are all scheduled to be split<sup>3</sup>. The procedure is then run until no more boxes larger than the minimum size that need splitting remain.

### 5.2.1 Computational Complexity

To estimate the computational complexity of the algorithm, we note that a box can only be split if it contains points of at least two phases. Consequently, the

---

<sup>2</sup>This choice of points to evaluate is the minimum number required to produce the convergence properties shown in section 5.3. Alternatively, one could also choose to evaluate all corners of the split boxes.

<sup>3</sup>Alternatively, one could evaluate the corners of these boxes to decide whether they should be split.

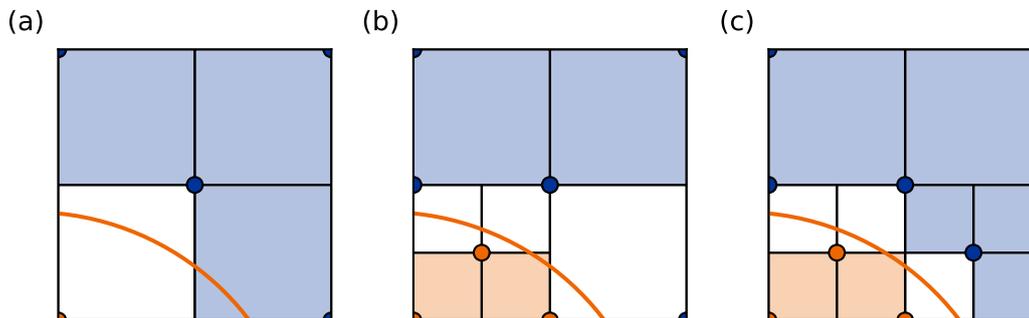


Figure 5.1: Illustration of a single step of the phase diagram algorithm. The exact phase boundary is shown with an orange line. Boxes are colored according to their phase, and boxes of undecided phase are shown in white. (a) Initially, only the lower left box has an undecided phase. (b) The calculation of the corners of the lower left box causes the lower right box to become undecided. (c) The lower right box is split. Only undecided boxes of a smaller size remain.

box must intersect a phase boundary. We can thus find an upper bound for the number of boxes that need to be split by assuming that all boxes which intersect a phase boundary are split.

For simplicity, we re-scale the phase space such that the initial boxes are hypercubes of side length  $d$ . After being split  $s$  times, the boxes will thus have side length  $d_s = d/2^s$ . These boxes can only be split if they are within  $\sqrt{n} \cdot d_s$  of the nearest phase boundary, where  $\sqrt{n}$  is the length of the  $n$ -dimensional diagonal. Consequently, the total volume of boxes that can be split scales as  $\mathcal{O}(d_s)$ . The number of boxes contained in this volume scales as

$$\frac{1}{d_s^n} \cdot \mathcal{O}(d_s) = \mathcal{O}\left(\frac{1}{d_s^{n-1}}\right). \quad (5.1)$$

The computational effort for splitting a box is a constant. Thus, the total complexity for calculating the phase diagram with accuracy  $\Delta x = d/2^m$  is given by

$$\mathcal{O}\left(\sum_{s=1}^m \frac{1}{d_s^{n-1}}\right). \quad (5.2)$$

Here,  $m = \log_2(d/\Delta x)$  is the required number of steps.

First, we consider the special case of  $n = 1$ . Since  $d_s^0 = 1$ , the computational complexity is

$$\mathcal{O}\left(\sum_{s=1}^m 1\right) = \mathcal{O}(m) = \mathcal{O}(\log(d/\Delta x)) = \mathcal{O}(\log(1/\Delta x)). \quad (5.3)$$

### 5.3 Convergence Properties

---

For higher-dimensional problems,

$$\sum_{s=1}^m \frac{1}{d_s^{n-1}} = \sum_{s=1}^m \left( \frac{2^s}{2^m \Delta x} \right)^{n-1} = \frac{1}{\Delta x^{n-1}} \underbrace{\sum_{s=1}^m \frac{1}{2^{(m-s)(n-1)}}}_{\text{converging}} = \mathcal{O}(1/\Delta x^{n-1}), \quad (5.4)$$

and thus the algorithm scales with  $\mathcal{O}(1/\Delta x^{n-1})$ .

#### 5.2.2 Asynchronous Execution

On a technical note, this algorithm lends itself well to be implemented in an asynchronous manner, using coroutines. Contrary to subroutines which run sequentially until they are completed, a coroutine can suspend its execution to wait for some other task to be completed, and then continue executing. Here, the routine which splits a given box can suspend its execution while waiting for the phase calculation to be completed. This allows other boxes to be split in the meantime, which can in turn schedule the phase calculation at a given point. In some cases, this can dramatically reduce the time to solution for calculating the phase diagram.

For example, if many phase calculations can be run concurrently as independent serial processes – a technique known as serial farming – an asynchronous implementation guarantees the maximal utilization of the available computational resources without requiring concurrent processes. If on the other hand there is a significant overhead in starting the process which computes the phase, it can be beneficial to group calculations together into batches that are computed in parallel on a single process. In this scenario, the asynchronous execution again improves throughput, by making sure that the maximum number of points can be grouped together.

## 5.3 Convergence Properties

In this section, we will discuss the convergence properties of the phase calculation algorithm. In particular, we will show some criteria for determining if the algorithm will converge towards the correct phase diagram, and sketch a proof for these criteria.

### 5.3.1 A Note on Terminology

To show the convergence properties of the phase calculation algorithm, we must first define some nomenclature:

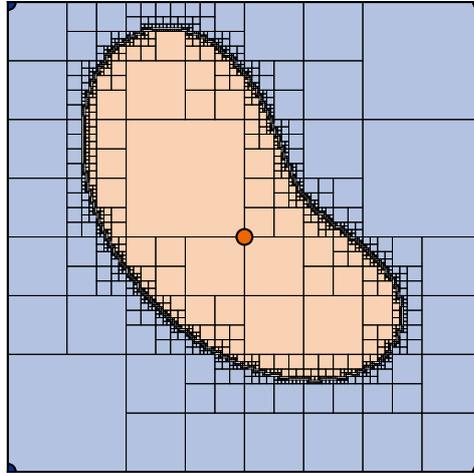


Figure 5.2: Phase diagram of a two-phase system where each phase contains an initial point in its interior. Boxes are shaded according to their phase, and initial points are marked with a dot.

- Two neighboring boxes are called *direct neighbors* if they share an  $(n - 1)$  - dimensional surface.
- For a set  $A$ ,  $\partial_\mu A = \left\{ x \in A \mid \inf_{y \in \partial A} \|x - y\| \leq \mu \right\}$  defines the border of  $A$  with *thickness*  $\mu$ .
- $\text{int}_\mu(A) = A - \partial_\mu A$  defines the interior of  $A$  where the border of thickness  $\mu$  is removed.
- The algorithm is said to *converge towards covering* a connected subset  $Q_i^j$  if for any  $\varepsilon > 0$  a minimum box size  $\Delta x$  can be chosen such that  $\text{int}_\varepsilon(Q_i^j)$  is completely covered by boxes of phase  $P_i$ .

### 5.3.2 Phase Containing a Starting Point

The convergence properties of the algorithm for calculating phase diagrams depend sensitively on the choice of the initial mesh. In this section, we will show that the algorithm converges towards covering a connected phase if it contains an initial point, as shown in fig. 5.2. We will start by showing some basic properties of the algorithm, which will be useful in proving the convergence.

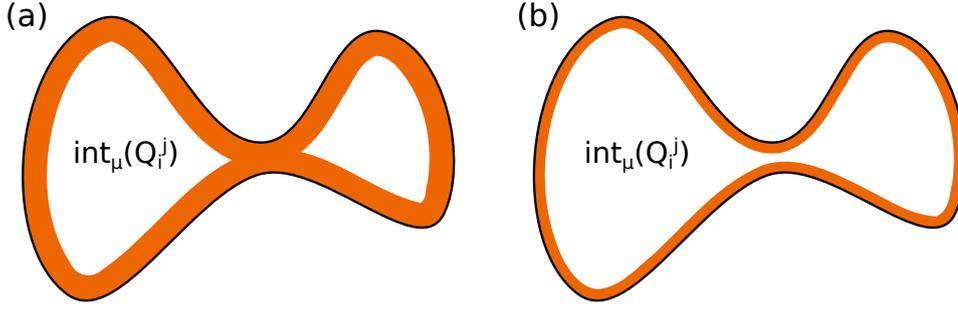


Figure 5.3: Interior  $\text{int}_\mu(Q_i^j)$  of a connected phase, with a border of thickness  $\mu$  (orange line) removed. (a) For large values of  $\mu$ , the interior can be disconnected. (b) Making  $\mu$  small enough re-connects the parts of the interior.

**Lemma 1.** *Direct neighbors cannot have different decided phases.*

*Proof.* By construction, each  $(n - 1)$  - dimensional surface of a box contains at least one evaluated point. Since this is true for the smaller box in a pair of direct neighbors, the two boxes share at least one point. Consequently, the phase of each box is either undecided, or the same phase as the shared point. Thus, the two neighbors cannot have *different* decided phases.  $\square$

**Lemma 2.**  $\text{int}_\mu(Q_i^j)$  is connected for some  $\mu > 0$ .

*Proof.* This follows from the assumption that each phase must have dimension  $n$  everywhere, as illustrated in fig. 5.3.  $\square$

**Lemma 3.** *The boxes which intersect  $\text{int}_{\sqrt{n}\Delta x}(Q_i^j)$  cannot have undecided phase.*

*Proof.* By construction, the boxes which have undecided phase must have the minimal side length  $\Delta x$ , and their diagonal has length  $\sqrt{n}\Delta x$ . In order to have an undecided phase, they must also intersect a phase boundary. Since the distance to the nearest phase boundary is larger than  $\sqrt{n}\Delta x$  for any point within  $\text{int}_{\sqrt{n}\Delta x}(Q_i^j)$ , the boxes which intersect it cannot have undecided phase.  $\square$

**Theorem 1.** *If a connected component  $Q_i^j$  contains an evaluated point  $p$  in its interior, the algorithm converges towards covering it.*

*Proof.* Given  $\varepsilon > 0$ , choose  $\Delta x$  such that:

- $\varepsilon > \sqrt{n}\Delta x$ ,
- $\tilde{Q} := \text{int}_{\sqrt{n}\Delta x}(Q_i^j)$  is connected (following lemma 2),
- $\tilde{Q}$  contains the evaluated point  $p$ .

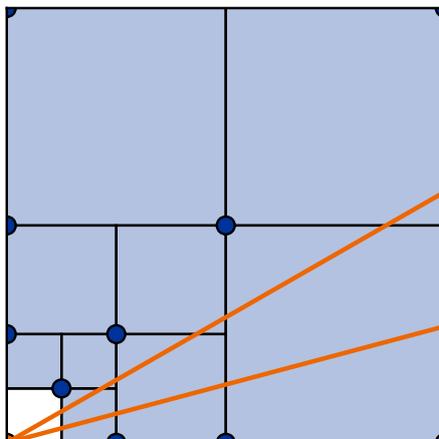


Figure 5.4: Example phase which does not converge even though it contains an initial point (the lower left corner) at the border. The phase boundaries are shown in orange, and boxes are colored according to their phase.

Following lemma 3 we know that the box containing  $p$  cannot have undecided phase, hence it must have the (correct) phase  $P_i$ . The rest of the boxes which intersect  $\tilde{Q}$  can be connected to that box via direct neighbors, hence they must have the same phase. Consequently,  $\text{int}_\varepsilon(Q_i^j)$  is covered by boxes of phase  $P_i$ .  $\square$

Note that we have used the fact that  $p$  must be in the *interior* of the connected phase in choosing  $\Delta x$  such that  $p$  is contained in  $\tilde{Q}$ . In fact, theorem 1 is no longer true if the point  $p$  lies on the border  $\partial Q_i^j$ , as is shown in fig. 5.4.

Also, note that the initial point in the interior of the connected phase is used only to ensure that there is at least one box intersecting  $\tilde{Q}$  which has the correct phase. Hence, we can write a modified version of theorem 1, which will be useful in proving theorem 3:

**Theorem 2.** *If the interior of a connected component  $Q_i^j$  intersects with a box of phase  $P_i$ , the algorithm converges towards covering  $Q_i^j$ .*

### 5.3.3 Sufficient Condition for Convergence to Covering a Phase

In this section, we will show a sufficient condition for convergence to covering a phase, even if the phase does not contain an initial point. We will first show a statement about a phase diagram with three phases, and then generalize this to an arbitrary number of phases.

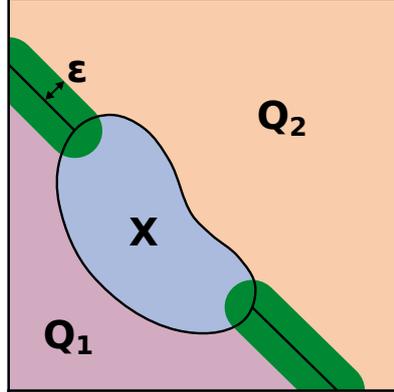


Figure 5.5: Phase diagram with three phases  $P_1$  (purple),  $P_2$  (orange) and  $P_X$  (blue). The points within distance  $\varepsilon$  of both  $P_1$  and  $P_2$  are marked in green. The set  $S_\varepsilon$  defined in eq. (5.5) contains all points except those marked in green.

**Theorem 3.** Let  $Q_1$ ,  $Q_2$  and  $X$  be three connected subsets of different phases  $P_1$ ,  $P_2$  and  $P_X$ . Both  $Q_1$  and  $Q_2$  contain an evaluated point in their interior, and they both touch  $X$  with an  $(n - 1)$  - dimensional boundary.<sup>4</sup> The algorithm then also converges towards covering  $X$ .

*Proof.* First, we define a set  $S_\varepsilon$  containing all points in  $X$ ,  $Q_1$  and  $Q_2$  except those within a distance  $\varepsilon$  from both  $P_1$  and  $P_2$ , as illustrated in fig. 5.5:

$$S_\varepsilon := \left\{ x \in Q_1 \cup Q_2 \cup X \mid \max \left( \inf_{y \in P_1} \|x - y\|, \inf_{y \in P_2} \|x - y\| \right) > \varepsilon \right\} \quad (5.5)$$

Since  $Q_1$  and  $Q_2$  are both connected to  $X$  by an  $(n - 1)$  - dimensional border, we can choose  $\Delta x$  small enough that  $S = S_{2\sqrt{n}\Delta x}$  is connected. Since a box of undecided phase must have border length  $\Delta x$ , it cannot intersect  $S$ ,  $P_1$  and  $P_2$  simultaneously. Also, if an undecided box intersecting  $S$  contains phase  $P_1$ , neighboring undecided boxes cannot contain phase  $P_2$ .

We can construct a path consisting of boxes of side length  $\Delta x$  intersecting  $S$ , connecting the evaluated point in  $Q_1$  to the one in  $Q_2$ . The boxes do not need to exist in the result of the algorithm, but they should be aligned to the evaluation grid, as shown in fig. 5.6(a).

Given this path, we can construct a path consisting of boxes that *are* calculated by the algorithm by choosing all boxes which intersect the previous path, as shown in fig. 5.6(b). Since this path connects boxes of phase  $P_1$  and  $P_2$  via

<sup>4</sup>This implies that  $X$  is *not* an island completely surrounded by only  $Q_1$  or  $Q_2$ .

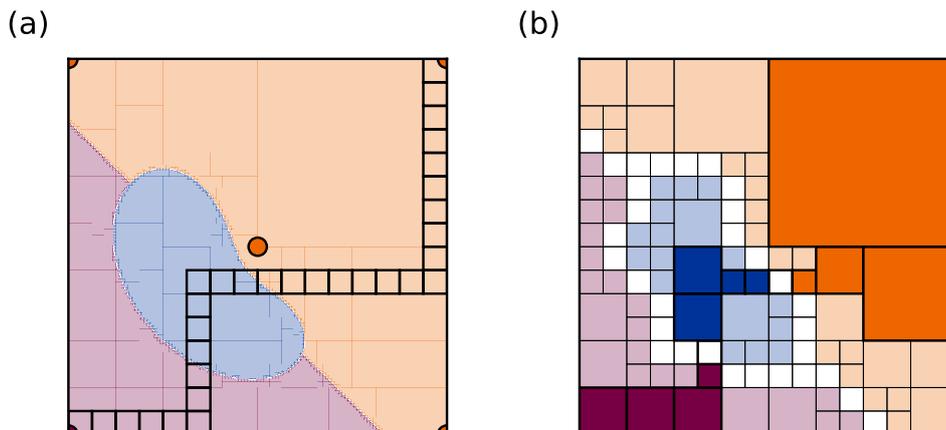


Figure 5.6: Example phase diagram with three phases, where only two contain a starting point. The phases  $P_1$  and  $P_2$  are shown in orange and purple respectively, and  $X$  is shown in blue. (a) A path of boxes constructed as described in the proof of theorem 3. The starting points are shown with dots, and the final phase diagram is shown in light colors. (b) Boxes which are calculated for the same value of  $\Delta x$ . Boxes which intersect the path shown in the left panel are shown in darker color.

direct neighbors, it must contain at least one box of undecided phase. Because undecided boxes cannot contain phases  $P_1$  and  $P_2$  simultaneously, it must contain either phases  $P_1$  and  $X$ , or  $P_2$  and  $X$ . Without loss of generality, we pick the *last* undecided box containing phases  $\{P_1, X\}$  along the path. The next box  $B$  along the path can then have either phase  $P_1$  or  $X$ . We can show by exclusion that the next box  $B$  must have phase  $X$ :

- By construction, it cannot contain phases  $\{P_1, X\}$ .
- It cannot have phase  $P_2$ , since it is the direct neighbor of a box with phases  $\{P_1, X\}$
- It cannot have undecided phase  $\{P_2, X\}$  since all boxes intersect  $S$ , and its direct neighbor has undecided phase  $\{P_1, X\}$ .
- If the box has phase  $P_1$ , the path from  $B$  to the box containing the initial point in  $Q_2$  again connects boxes of phase  $P_1$  and  $P_2$  via direct neighbors. As a result, there must be another undecided box along the path, connecting  $P_1$  to the next phase. Since there cannot be any undecided boxes containing both  $P_1$  and  $P_2$ , it must again contain  $\{P_1, X\}$ . This contradicts the construction that we picked the *last* such box.

From theorem 2, it follows that the algorithm will converge towards covering  $X$ .  $\square$

Having shown a criterion for phase diagrams of three phases to converge, we can now generalize this to many phases:

**Theorem 4.** *Let  $X$  be a connected subset of a phase  $P_X$ . If there are two connected sets  $S = \{Q_i^j\}_{i,j}$  and  $S' = \{Q_{i'}^{j'}\}_{i',j'}$ , such that*

- *$S$  and  $S'$  are connected to  $X$  via an  $(n - 1)$  - dimensional border*
- *$S$  and  $S'$  contain a starting point in their interior*
- *The phases in  $S$  and  $S'$  are pairwise different, i.e.  $\{P_i\}_i \cap \{P_{i'}\}_{i'} = \emptyset$ ,*

*the algorithm converges towards covering  $X$ .*

*Proof.* First, we consider a phase diagram where all phases with indices  $i$  are assigned a phase  $P_S$ , and all phases with indices  $i'$  are assigned phase  $P_{S'}$ . From theorem 3, it follows that the algorithm will converge towards covering  $X$ . To recreate the original problem, we split the phases  $P_S$  and  $P_{S'}$  back up into their original phases. Because this process only ever *adds* different phases and does not merge them together, it can only add more boxes of undecided phase. Consequently, all boxes which are split in the simplified phase diagram will still be split, and the algorithm will still converge towards covering  $X$ .  $\square$

This property can allow calculating even complex phase diagrams with only few starting points. As an example, fig. 5.7 shows the phase diagram of water, without the critical point. This phase diagram was calculated using only one initial box, with data extracted from the figure [89]. Even though the five initial points only show four phases, the algorithm correctly identifies all fourteen phases from the input figure.

## 5.4 Conclusion

We have introduced an algorithm for calculating diagrams of discrete phases. By constricting the phase evaluations to the vicinity of phase boundaries, this algorithm can significantly reduce the number of such evaluations needed to calculate the phase diagram. While there are some cases where the algorithm will not converge to the true phase diagram, namely when a phase is completely enclosed in another phase, we have shown a relatively simple criterion which guarantees convergence. We expect this algorithm to be applicable to a wide range of physical problems.

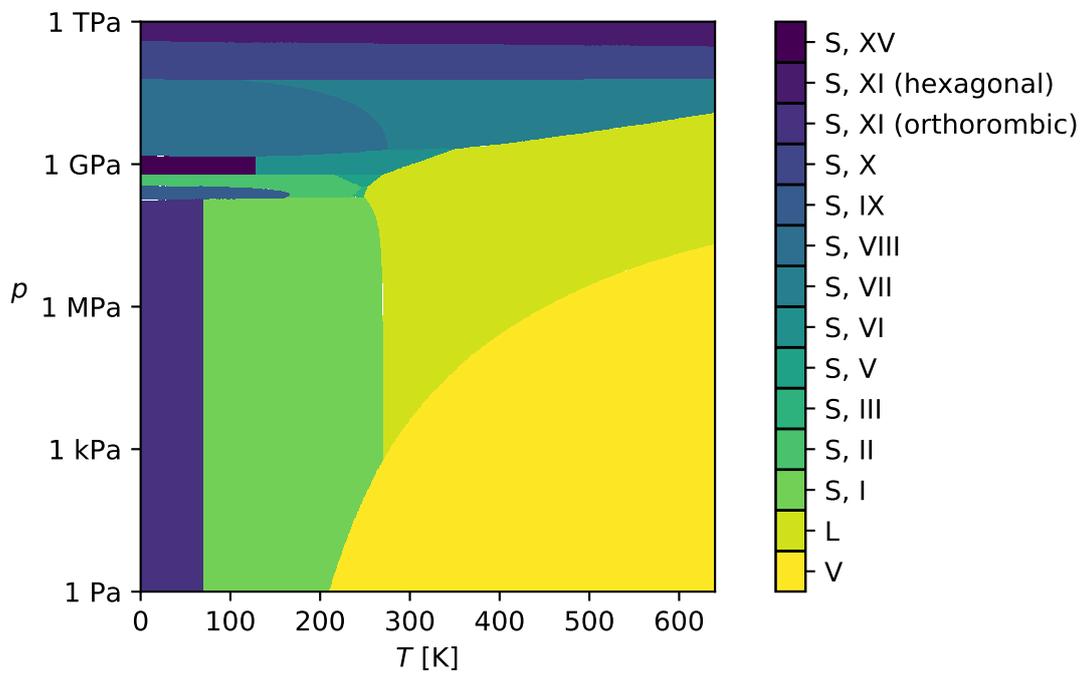


Figure 5.7: Phase diagram of water, calculated with only one initial box. The phase data is extracted from the figure [89].



# 6

## Weyl Fermions Emerging in a Magnetic Field

The content of this chapter was previously published in ref. [7].

Recently, transition metal dipnictides of the type  $AB_2$  ( $A \in \{\text{Ta}, \text{Nb}\}$ ,  $B \in \{\text{As}, \text{Sb}\}$ ) have gained a lot of attention [90–97] for their giant magnetoresistance. These materials are semimetals, but without a direct closure of the band gap. Consequently, they do not host any Weyl or Dirac points.

The exact nature of magnetoresistance in these materials - especially the dependency on the direction of the magnetic field - is still under active investigation. Negative magnetoresistance has been observed experimentally for  $\text{NbAs}_2$  [95, 96],  $\text{TaAs}_2$  [95] and  $\text{TaSb}_2$  [95, 98]. Anomalous - albeit not negative - magnetoresistance has been observed for  $\text{NbSb}_2$ [90] and  $\text{TaAs}_2$  [93]. However, there are also experiments which point to the contrary, which is that there is no negative magnetoresistance in these materials. In Ref. [94], negative magnetoresistance was observed at first but then determined to be an artifact of the measurement setup.

In the following, we propose a mechanism to produce Weyl nodes in these materials under the influence of a magnetic field. The chiral anomaly associated with these Weyl nodes is a possible source of negative magnetoresistance. Such an appearance of Weyl points under magnetic field has recently been proposed in Ref. [99]. The mechanism with which the Weyl points appear, however, is a different one – in this work the Weyl points appear from a previously gapped state, while the Ref. [99] discusses Weyl points arising from the splitting of a four-fold crossing.

The chapter is structured as follows: In the first section, the atomic and electronic structure of the four compounds is described. A four-band Hamiltonian for

TaAs<sub>2</sub> is derived from symmetry considerations and fitted to the band structure. In the second section, the topology of the band structure is studied, first without magnetic field and then by applying a Zeeman term. We find that this leads to the appearance of Weyl points.

## 6.1 Atomic and Electronic Structure of AB<sub>2</sub> compounds

### 6.1.1 Atomic Structure

In the following, the atomic structure of TaAs<sub>2</sub> [100], TaSb<sub>2</sub> [101], NbSb<sub>2</sub> [102] and NbAs<sub>2</sub> [103] is described.

The reduced unit cell of AB<sub>2</sub> compounds has the general form

$$\begin{aligned} a_1 &= (a, b, 0) \\ a_2 &= (-a, b, 0) \\ a_3 &= (-c, 0, d) \end{aligned} \tag{6.1}$$

with parameters as given in table 6.1 [100, 102].

	<b>a</b>	<b>b</b>	<b>c</b>	<b>d</b>
<b>TaAs<sub>2</sub></b>	4.6655	1.6915	3.8420	6.7330
<b>TaSb<sub>2</sub></b>	5.11	1.822	4.1950	7.1502
<b>NbAs<sub>2</sub></b>	4.684	1.698	3.8309	6.7933
<b>NbSb<sub>2</sub></b>	5.1198	1.8159	4.1705	7.2134

Table 6.1: Unit cell dimensions (in Å) for AB<sub>2</sub> compounds.

Each unit cell contains 2 formula units. The atoms are located at general Wyckoff positions  $(x, -x, y), (-x, x, -y)$ , for  $(x, y)$  as shown in table 6.2 [100, 102].

Figure 6.1 shows the reduced unit cell and 1. BZ of TaAs<sub>2</sub>. The k-point path along which band structure calculations are performed is indicated. In the basis reciprocal to that of eq. (6.1), the special k-points are given by

$$\begin{aligned} \Gamma &= (0, 0, 0) \\ A &= (0, 0, 0.5) \\ L &= (0.5, 0, 0.5) \\ M &= (0.5, 0.5, 0.5) \\ V &= (0.5, 0, 0) \\ Y &= (0.5, 0.5, 0). \end{aligned} \tag{6.2}$$

	<b>A</b>	<b>B1</b>	<b>B2</b>
<b>TaAs<sub>2</sub></b>	(0.157, 0.1959)	(0.4054, 0.1082)	(0.1389, 0.5265)
<b>TaSb<sub>2</sub></b>	(0.152, 0.19)	(0.405, 0.113)	(0.147, 0.535)
<b>NbAs<sub>2</sub></b>	(0.1574, 0.1965)	(0.4059, 0.1084)	(0.14, 0.528)
<b>NbSb<sub>2</sub></b>	(0.1521, 0.1903)	(0.4051, 0.1127)	(0.1475, 0.5346)

Table 6.2: Atomic positions  $(x, y)$ .

### 6.1.2 Electronic Structure

Electronic structure calculations were performed in VASP [30], with projector augmented-wave (PAW) [104, 105] pseudopotentials. The PBE approximation [16] was used, and spin-orbit coupling was included in the potentials. The self-consistent field (SCF) calculations were performed on a  $11 \times 11 \times 5$   $\Gamma$  - centered grid for TaAs<sub>2</sub>, and a  $10 \times 10 \times 5$   $\Gamma$  - centered grid for NbSb<sub>2</sub>. The energy cutoff given in the potential files was used, which is 293.2 eV for NbAs<sub>2</sub> and NbSb<sub>2</sub>, and 223.7 eV for TaAs<sub>2</sub> and TaSb<sub>2</sub>.

Additionally, the PBE calculations were tested against the accurate HSE06 hybrid functional [106, 20]. The hybrid SCF calculations for the band structures were performed on a  $\Gamma$  - centered  $6 \times 6 \times 4$  grid for all materials. For the generation of the Wannier tight-binding model of NbSb<sub>2</sub> a  $\Gamma$  - centered  $10 \times 10 \times 5$  grid was used.

The band structure of TaAs<sub>2</sub> and NbSb<sub>2</sub> is shown in figs. 6.2 and 6.3. Both materials exhibit a pair of electron and hole pockets near the  $M$  - point, where the minimum band gap is about 318 meV (120 meV without hybrid functionals) in the case of TaAs<sub>2</sub>, 151 meV (98 meV) for TaSb<sub>2</sub>, 261 meV (22 meV) for NbAs<sub>2</sub>, and 67 meV (18 meV) in the case of NbSb<sub>2</sub>. A more complete calculation of the band structure can be found for example in Ref. [107].

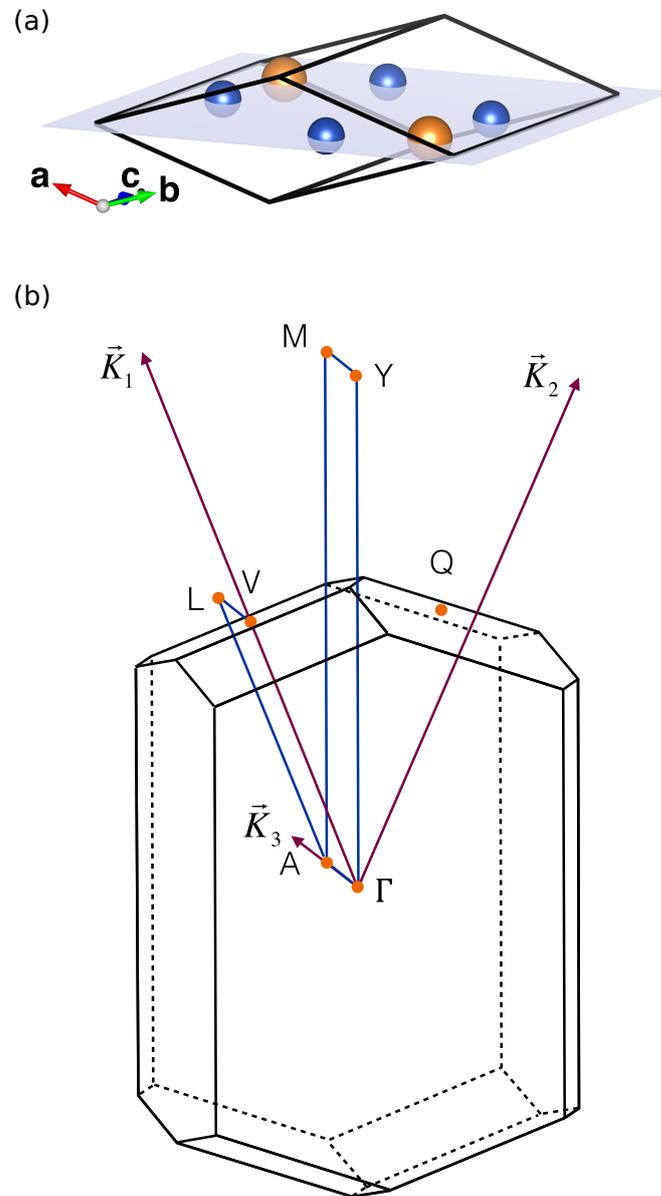


Figure 6.1: (a) Reduced unit cell of  $TaAs_2$ . (b) First BZ of  $TaAs_2$ . The  $\mathbf{k}$ -point path is indicated in blue.

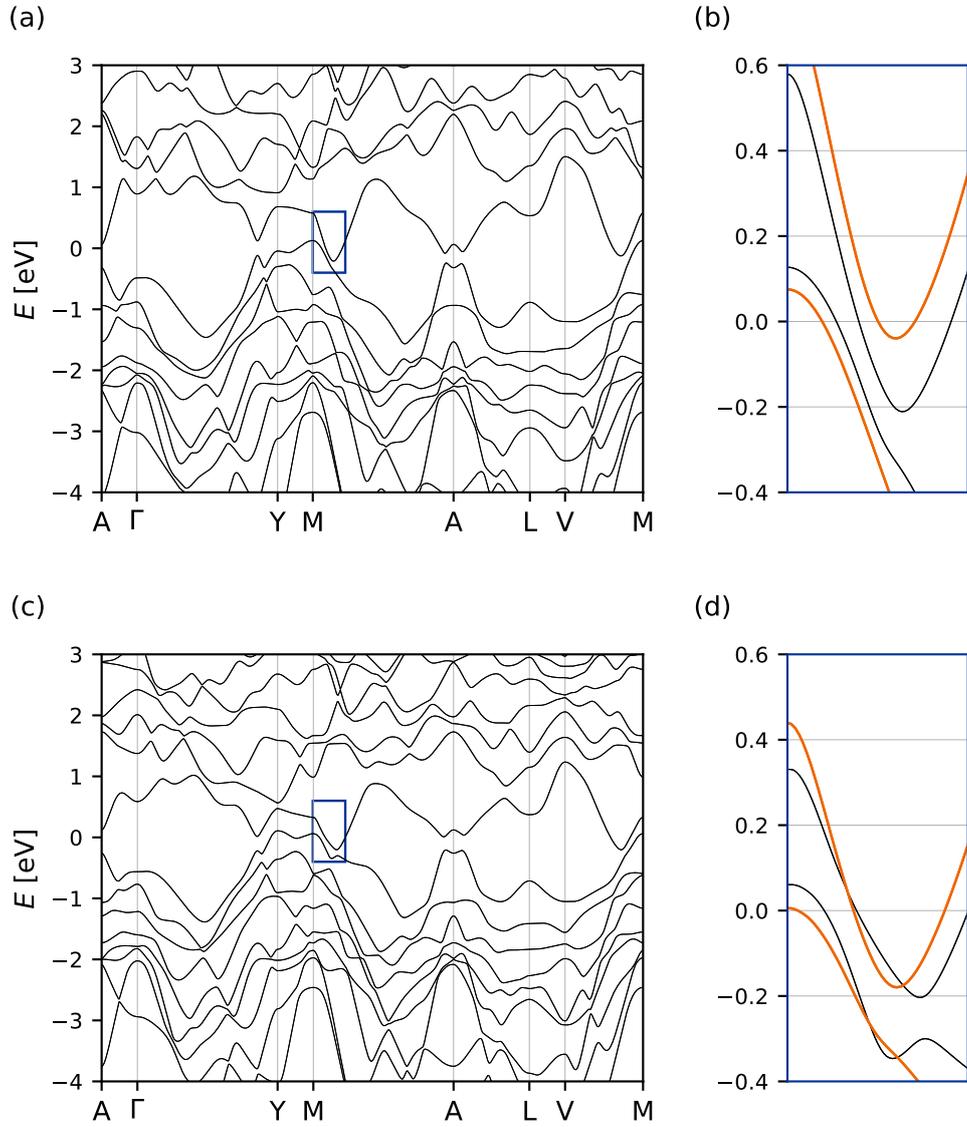


Figure 6.2: Calculated band structures of  $\text{TaX}_2$  compounds. The inset shows electron and hole pockets around  $M$ . The orange line represents calculations using hybrid functionals. (a-b)  $\text{TaAs}_2$  (c-d)  $\text{TaSb}_2$

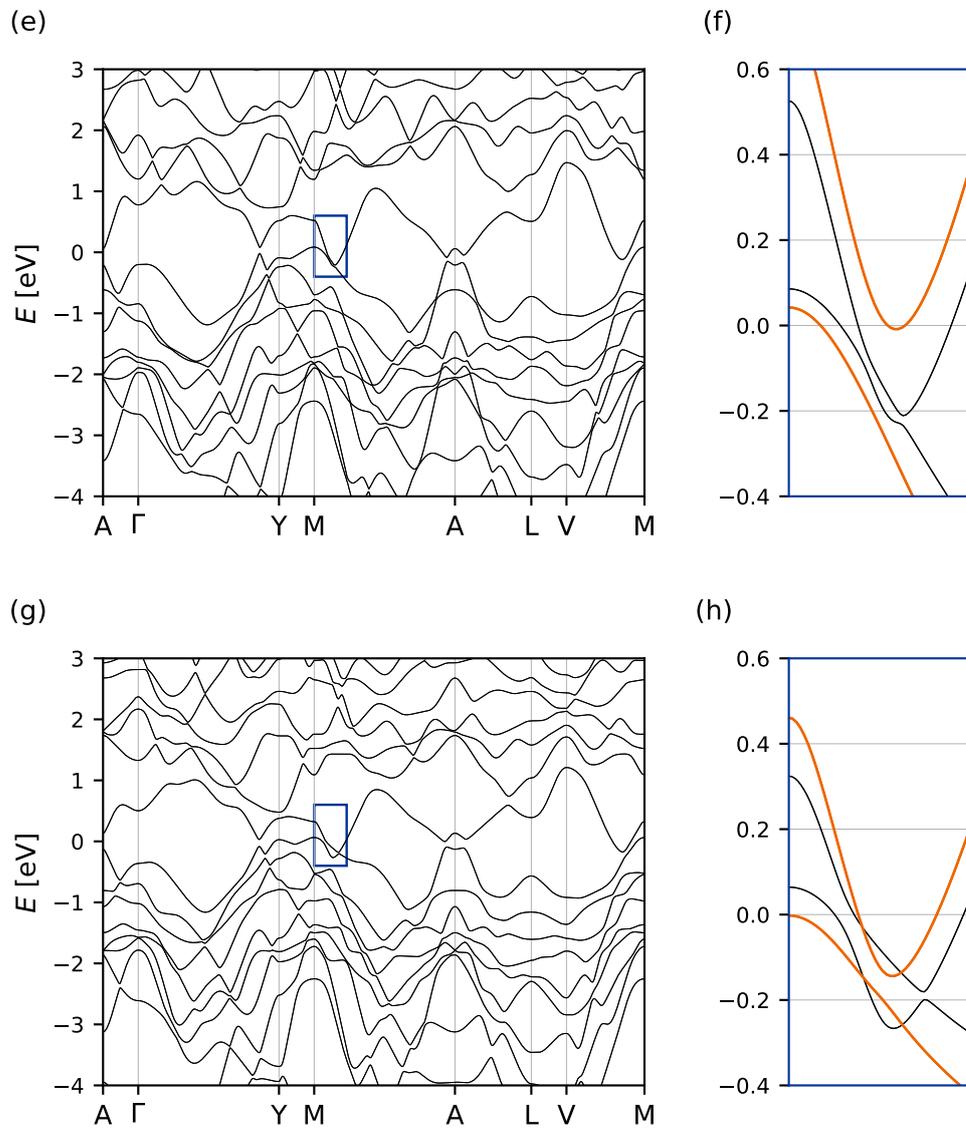


Figure 6.3: Calculated band structures of  $NbX_2$  compounds. The inset shows electron and hole pockets around  $M$ . The orange line represents calculations using hybrid functionals. (e-f)  $NbAs_2$  (g-h)  $NbSb_2$

### 6.1.3 Symmetry Operations and $\mathbf{k} \cdot \mathbf{p}$ Model

The  $\text{AB}_2$  compounds studied here have  $C_{2/m}$  symmetry (space group 12). The rotation axis is along the cartesian  $y$ -axis. In reduced coordinates, the symmetry matrices are as follows:

$$\begin{aligned}
 &\bullet \text{ Identity} && E = \mathbb{I}_{3 \times 3} && (6.3) \\
 &\bullet \text{ Rotation} && C_{2y} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix} \\
 &\bullet \text{ Parity} && P = -\mathbb{I}_{3 \times 3} \\
 &\bullet \text{ Mirror} && M_y = PC_{2y} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}
 \end{aligned}$$

	$E$	$C_{2y}$	$P$	$M_y$
$\Gamma_3^+$	1	$i$	1	$i$
$\Gamma_4^+$	1	$-i$	1	$-i$
$\Gamma_3^-$	1	$i$	-1	$-i$
$\Gamma_4^-$	1	$-i$	-1	$i$

Table 6.3: Character table for the relevant double group representations of  $C_{2m}$  [108].

From the first-principles wave-functions, the representations corresponding to the two highest valence and two lowest conduction bands at the  $M$  - point were determined using the `WIEN2k` code [33, 109]. They were found to be  $\Gamma_3^+, \Gamma_4^+$  and  $\Gamma_3^-, \Gamma_4^-$ , respectively. Their characters are shown in table 6.3, which comes from table 15 on page 35 in Koster et al. [108]. Consequently, the symmetry represen-

tations in these four bands are given by

$$\begin{aligned}
 & \bullet \text{ Identity} & E &= \mathbb{I}_{4 \times 4} & (6.4) \\
 & \bullet \text{ Rotation} & C_{2y} &= \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & i & 0 \\ 0 & 0 & 0 & -i \end{pmatrix} \\
 & \bullet \text{ Parity} & P &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \\
 & \bullet \text{ Mirror} & M_y = PC_{2y} &= \begin{pmatrix} i & 0 & 0 & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & 0 & 0 & i \end{pmatrix} \\
 & \bullet \text{ Time-reversal} & \mathcal{T} &= \begin{pmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \hat{K}
 \end{aligned}$$

For each of the symmetry operations  $g$ , the constraint

$$\mathcal{H}(\mathbf{k}) = D(g)\mathcal{H}(g^{-1}\mathbf{k})D(g^{-1}) \quad (6.5)$$

is imposed on the  $4 \times 4$  Hamiltonian, where  $D(g)$  is the symmetry representation. By applying these constraints on the general form of a four-band Hamiltonian

$$\mathcal{H}(\mathbf{k}) = \sum_{i,j \in \{0,x,y,z\}} C_{ij}(\mathbf{k})(\sigma_i \otimes \sigma_j), \quad (6.6)$$

we find the Hamiltonian to be of the form

$$\begin{aligned}
 \mathcal{H}(\mathbf{k}) &= C_{00}(\mathbf{k})(\sigma_0 \otimes \sigma_0) + C_{xx}(\mathbf{k})(\sigma_x \otimes \sigma_x) + \\
 & C_{xy}(\mathbf{k})(\sigma_x \otimes \sigma_y) + C_{xz}(\mathbf{k})(\sigma_x \otimes \sigma_z) + \\
 & C_{y0}(\mathbf{k})(\sigma_y \otimes \sigma_0) + C_{z0}(\mathbf{k})(\sigma_z \otimes \sigma_0), \quad (6.7)
 \end{aligned}$$

where the  $C_{ij}(\mathbf{k})$  are given up to second order in  $\mathbf{k}^* = \mathbf{k} - M$  (in reduced coordi-

nates) by

$$\begin{aligned}
 C_{00}(\mathbf{k}^*) &= C_{00}^1 + C_{00}^{x^2+y^2}((k_x^*)^2 + (k_y^*)^2) + \\
 &C_{00}^{xy} k_x^* k_y^* + C_{00}^{xz-yz}(k_x^* k_z^* - k_y^* k_z^*) + \\
 &C_{00}^{z^2} (k_z^*)^2
 \end{aligned} \tag{6.8}$$

$$\begin{aligned}
 C_{z0}(\mathbf{k}^*) &= C_{z0}^1 + C_{z0}^{x^2+y^2}((k_x^*)^2 + (k_y^*)^2) + \\
 &C_{z0}^{xy} k_x^* k_y^* + C_{z0}^{xz-yz}(k_x^* k_z^* - k_y^* k_z^*) + \\
 &C_{z0}^{z^2} (k_z^*)^2
 \end{aligned} \tag{6.9}$$

$$C_{xx}(\mathbf{k}^*) = C_{xx}^{x-y}(k_x^* - k_y^*) + C_{xx}^z k_z^* \tag{6.10}$$

$$C_{xy}(\mathbf{k}^*) = C_{xy}^{x-y}(k_x^* - k_y^*) + C_{xy}^z k_z^* \tag{6.11}$$

$$C_{xz}(\mathbf{k}^*) = C_{xz}^{x+y}(k_x^* + k_y^*) \tag{6.12}$$

$$C_{y0}(\mathbf{k}^*) = C_{y0}^{x+y}(k_x^* + k_y^*). \tag{6.13}$$

The general procedure for calculating the terms allowed by symmetry in a  $\mathbf{k} \cdot \mathbf{p}$  model is described in chapter 7.

These 16 parameters were numerically fitted to the band structure of TaAs<sub>2</sub> using the `scipy` [110] package, to obtain the values in table 6.4. The resulting band structure around the  $M$ -point is shown in fig. 6.4. Comparing it to the band structure obtained from first-principles reveals that the approximation is accurate in the immediate vicinity of the  $M$ -point, but breaks down at around 6% of the distance along the line  $M-A$ . Importantly, the minimum band gap is not preserved in this model. Nevertheless, the model can be used to qualitatively study effects in TaAs<sub>2</sub>, owing to the fact that it contains the correct symmetry representations.

[eV]	$C_{00}^1$	= 7.066	$C_{z0}^1$	= -0.224
[eVÅ]	$C_{xz}^{x+y}$	= 1.272	$C_{y0}^{x+y}$	= 1.270
	$C_{xx}^{x-y}$	= -0.061	$C_{xy}^{x-y}$	= -1.999
	$C_{xx}^z$	= -0.554	$C_{xy}^z$	= -0.253
[eVÅ <sup>2</sup> ]	$C_{00}^{x^2+y^2}$	= -71.21	$C_{z0}^{x^2+y^2}$	= 56.30
	$C_{00}^{xy}$	= -137.1	$C_{z0}^{xy}$	= 123.1
	$C_{00}^{xz-yz}$	= 1.52	$C_{z0}^{xz-yz}$	= -1.49
	$C_{00}^{z^2}$	= -0.84	$C_{z0}^{z^2}$	= -1.88

Table 6.4: Parameters of the  $4 \times 4$  Hamiltonian of TaAs<sub>2</sub> around  $M$  up to second order.

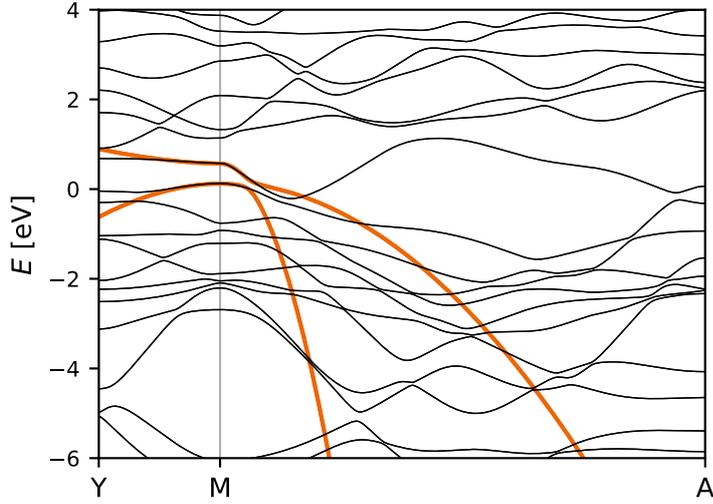


Figure 6.4: TaAs<sub>2</sub> band structure of the  $\mathbf{k} \cdot \mathbf{p}$  model (thick orange line), compared to the first-principles result (black lines).

## 6.2 Band Structure Topology

In this section, we describe the band structure topology and the influence of magnetic field. First, we describe the topology in the absence of magnetic field for all four compounds. Then, we show that Weyl points appear under sufficient magnetic field. This result is shown first for the  $\mathbf{k} \cdot \mathbf{p}$  model of TaAs<sub>2</sub> derived in section 6.1.3, and then for a tight-binding model of NbSb<sub>2</sub> derived from first-principles.

### 6.2.1 Band Structure Topology without Magnetic Fields

In the absence of magnetic field, there is no direct band gap closure in AB<sub>2</sub> compounds. Since the valence bands thus form a well-defined manifold, they can be classified – just like insulators – according to the topology of these valence bands. Because time-reversal symmetry is fulfilled, a  $\mathbb{Z}_2$  classification is possible.

All compounds were found to be weak topological insulators, with  $\mathbb{Z}_2$  indices 0; (111). That is, all time-reversal invariant planes  $k_i = 0, 0.5$  have a non-trivial  $\mathbb{Z}_2$  index  $\Delta = 1$ . This result was derived from first-principles using the Z2Pack code [6], and agrees with previous studies [93, 95, 107]. The corresponding evolution of Wannier charge centers is shown, for the case of TaAs<sub>2</sub>, in fig. 6.5.

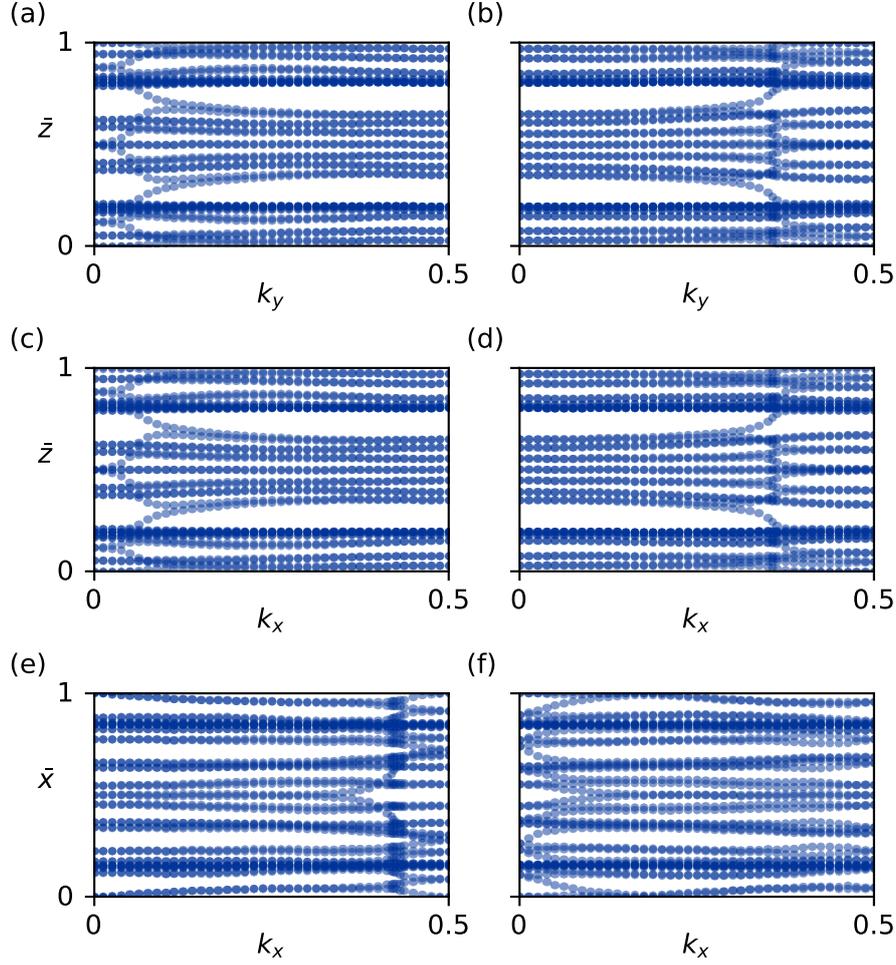


Figure 6.5: Wannier charge center evolution for the time-reversal invariant planes of TaAs<sub>2</sub>. (a)  $k_x = 0$  (b)  $k_x = 0.5$  (c)  $k_y = 0$  (d)  $k_y = 0.5$  (e)  $k_z = 0$  (f)  $k_z = 0.5$

## 6.2.2 Effect of Zeeman Splitting on the $\mathbf{k} \cdot \mathbf{p}$ Model for TaAs<sub>2</sub>

Here we study the effects of magnetic field on TaAs<sub>2</sub> by adding a Zeeman splitting term to the  $\mathbf{k} \cdot \mathbf{p}$  model derived in section 6.1.3 (eq. (6.7)). The splitting term is given by

$$\Delta\mathcal{H} = c_x\sigma_0 \otimes \sigma_y + c_y\sigma_0 \otimes \sigma_z + c_z\sigma_0 \otimes \sigma_x, \quad (6.14)$$

where  $c_i$  is the strength of the Zeeman splitting induced by the magnetic field in that direction, that is

$$c_i = \sum_j g_{ij}\mu_B H_j. \quad (6.15)$$

This assumes that the g-factor is equal for all bands. The limitations of this approximation are discussed in section 6.2.4.

### Magnetic Field Along the Rotation Axis $\hat{y}$

When magnetic field is applied along the rotation axis  $\hat{y}$ , the Zeeman term (eq. (6.14)) takes the form

$$\Delta\mathcal{H} = c_y\sigma_0 \otimes \sigma_z. \quad (6.16)$$

This term preserves all spatial symmetries of the system, breaking only time-reversal.

Along the  $M - A$  line, the  $C_{xx}$  and  $C_{yy}$  contributions to the Hamiltonian vanish since  $k_x^* = k_y^*$  and  $k_z^* = 0$ . Consequently, the energy eigenvalues are given by

$$E(\mathbf{k}) = C_{00}(\mathbf{k}) \pm c_y \mp \sqrt{C_{xz}(\mathbf{k})^2 + C_{y0}(\mathbf{k})^2 + C_{z0}(\mathbf{k})^2} \quad (6.17)$$

The Zeeman term counteracts the original splitting (square root term), such that for sufficient magnetic field there will be a direct band gap closure. Away from the  $M - A$  line, the band gap remains open, giving rise to a Weyl point.

When the Zeeman splitting is gradually switched on, two pairs of Weyl points form at about  $c_y = 0.11$  eV. Increasing the Zeeman splitting leads to a separation between the two nodes in a pair, with one node each moving towards the  $M$ -point. Finally, at  $c_y \approx 0.25$  eV, these two nodes meet at  $M$  and annihilate. This process is shown in fig. 6.6.

The existence of these Weyl points was confirmed by verifying that the nodes are a source or sink of Berry curvature. For this purpose, the Chern number of spheres surrounding the points was calculated by tracking hybrid Wannier charge centers (HWCC) on loops around the sphere [2, 5, 3, 6], using the Z2Pack code [6]. Figure 6.7 shows the evolution of the sum of HWCC for two of four nodes found at  $c_y = 0.12$  eV, demonstrating that the two points are Weyl nodes of opposite chirality.

### General Magnetic Field Direction

Finally, the effects of a magnetic field in a general direction were studied. It turns out that, even though such a field breaks the spatial symmetries of the system, Weyl nodes still appear under a strong enough magnetic field. When magnetic field is applied in  $\hat{x}$  or  $\hat{z}$  - direction, a single pair of Weyl points emerges from the  $M$  point. These Weyl nodes are located on the  $k_x = -k_y$  plane, as shown in table 6.5.

Figure 6.8 shows the number of Weyl points as a function of the Zeeman splitting. To obtain this phase diagram, candidate Weyl points were identi-

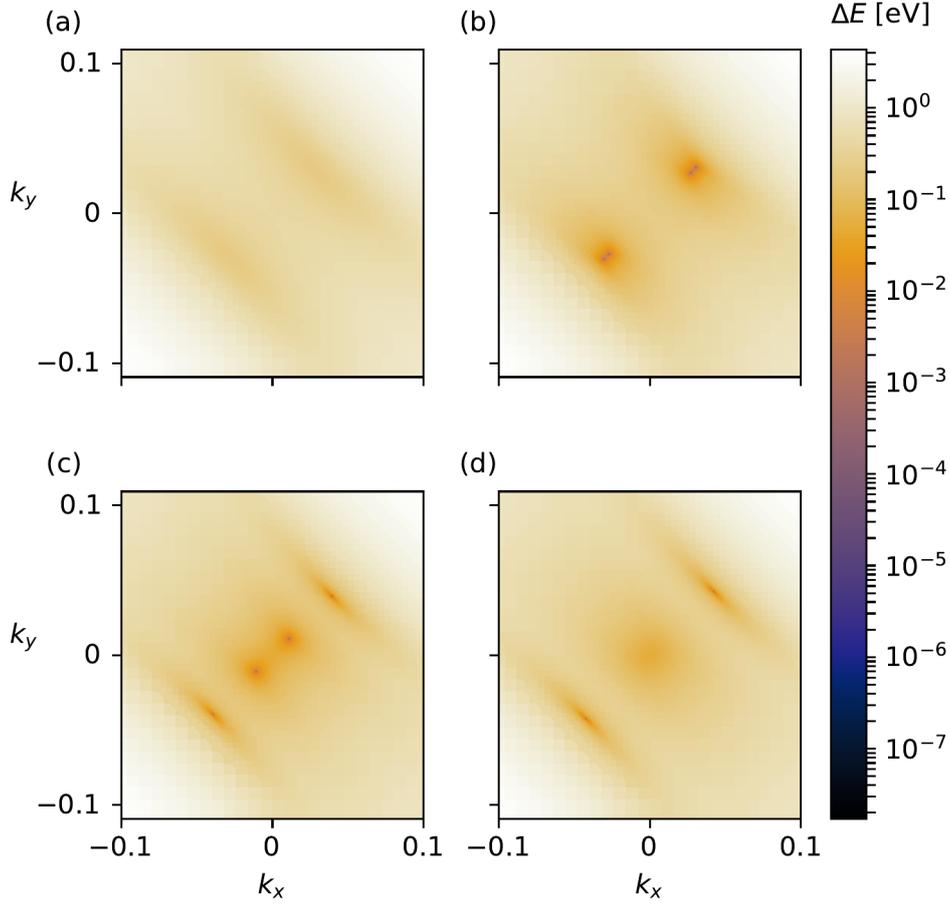


Figure 6.6: Band gap of TaAs<sub>2</sub> in the  $k_z^* = 0$  plane for different values of the magnetic field in  $y$ -direction, calculated from the  $\mathbf{k}\cdot\mathbf{p}$  model. A dark spot indicates the presence of a Weyl point. (a) No magnetic field. There are no Weyl points present (b)  $c_y = 0.11$  eV. Four Weyl points have appeared on the  $k_x = k_y$  line. (c)  $c_y = 0.2$  eV. The pair of Weyl points move further apart. (d)  $c_y = 0.25$  eV. Two Weyl points has annihilated at  $M$ , leaving the other two.

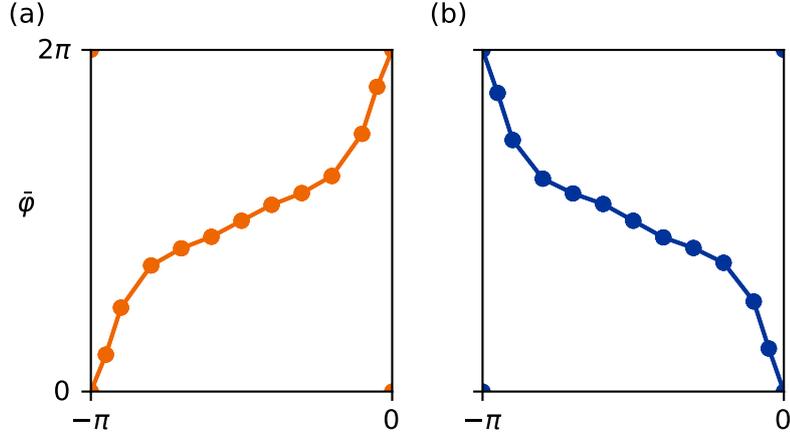


Figure 6.7: Evolution of the sum of HWCC on spheres surrounding the Weyl points at  $c_y = 0.12$  eV splitting. (a) Weyl point at  $\mathbf{k} = (0.5247, 0.5247, 0.5)$ , having positive chirality  $C = +1$  (b) Weyl point at  $\mathbf{k} = (0.53258, 0.53258, 0.5)$  with negative chirality  $C = -1$

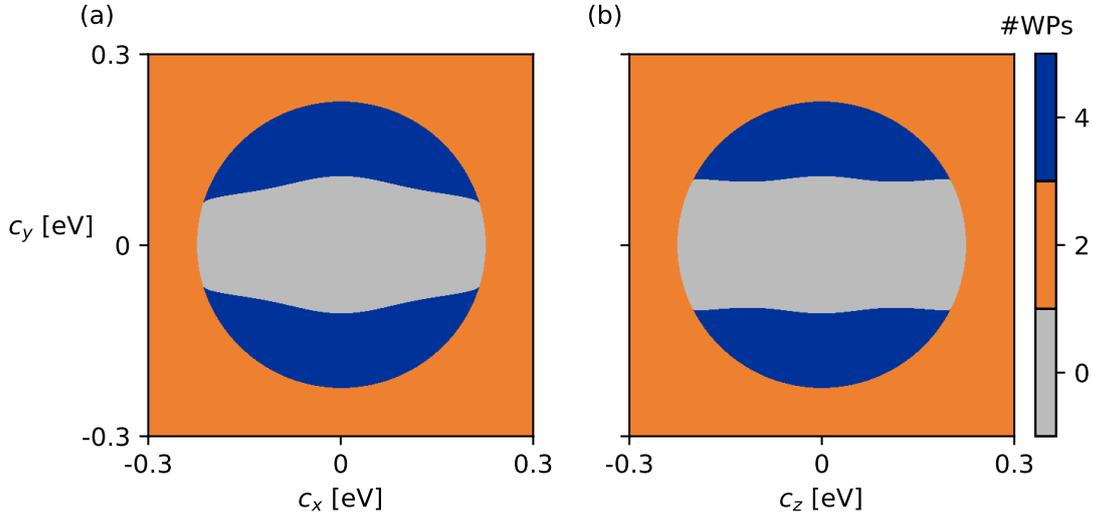


Figure 6.8: Phase diagram showing the number of Weyl points in the  $\mathbf{k} \cdot \mathbf{p}$  model of TaAs<sub>2</sub> as a function of the Zeeman splitting (in eV).

Splitting [eV]	Weyl position $\mathbf{k}^*$	Chirality
$c_x = 0.225$	$(-0.0042, 0.0042, 0.00093)$	-1
	$(0.0042, -0.0042, -0.00093)$	1
$c_x = 0.25$	$(-0.025, 0.025, 0.0054)$	-1
	$(0.025, -0.025, -0.0054)$	1
$c_x = 0.3$	$(-0.044, 0.044, 0.0098)$	-1
	$(0.044, -0.044, -0.0098)$	1
$c_z = 0.225$	$(0.0011, -0.0011, -0.018)$	-1
	$(-0.0011, 0.0011, 0.018)$	1
$c_z = 0.25$	$(0.0066, -0.0066, -0.11)$	-1
	$(-0.0066, 0.0066, 0.11)$	1
$c_z = 0.3$	$(0.012, -0.012, -0.18)$	-1
	$(-0.012, 0.012, 0.18)$	1

Table 6.5: Position  $\mathbf{k}^* = \mathbf{k} - M$  (in reduced coordinates) and chirality of Weyl points for Zeeman splittings in  $\hat{x}$  and  $\hat{z}$ -direction.

fied using a quasi Newton algorithm to find minima in the band gap (using `scipy.optimize.minimize` [110]), for different initial guesses. In a second step, the Chern number on a small sphere (radius  $10^{-4} \text{ \AA}^{-1}$ ) surrounding the candidate points was evaluated (using Z2Pack [6]), keeping only points with a non-zero Chern number. Finally, duplicate points were eliminated by checking whether two points lie within the diameter of the sphere of one another. An improved version of this procedure will be discussed in detail in chapter 9.

### 6.2.3 Effect of Zeeman Splitting on the Tight-binding Model for NbSb<sub>2</sub>

Having studied the effects of Zeeman splitting on the  $\mathbf{k} \cdot \mathbf{p}$  model for TaAs<sub>2</sub>, we now study a more realistic tight-binding model for NbSb<sub>2</sub>, derived from a first-principles calculation with hybrid functionals using the Wannier90 code [111, 112]. NbSb<sub>2</sub> was chosen because it has the smallest direct band gap of the four materials, making it the most promising candidate for hosting Weyl points at realistic magnetic field strength.

The Zeeman splitting for this model can again be expressed by adding the corresponding terms to the Hamiltonian

$$\Delta\mathcal{H} = c_x\sigma_x \otimes \mathbb{I}_{22 \times 22} + c_y\sigma_y \otimes \mathbb{I}_{22 \times 22} + c_z\sigma_z \otimes \mathbb{I}_{22 \times 22}, \quad (6.18)$$

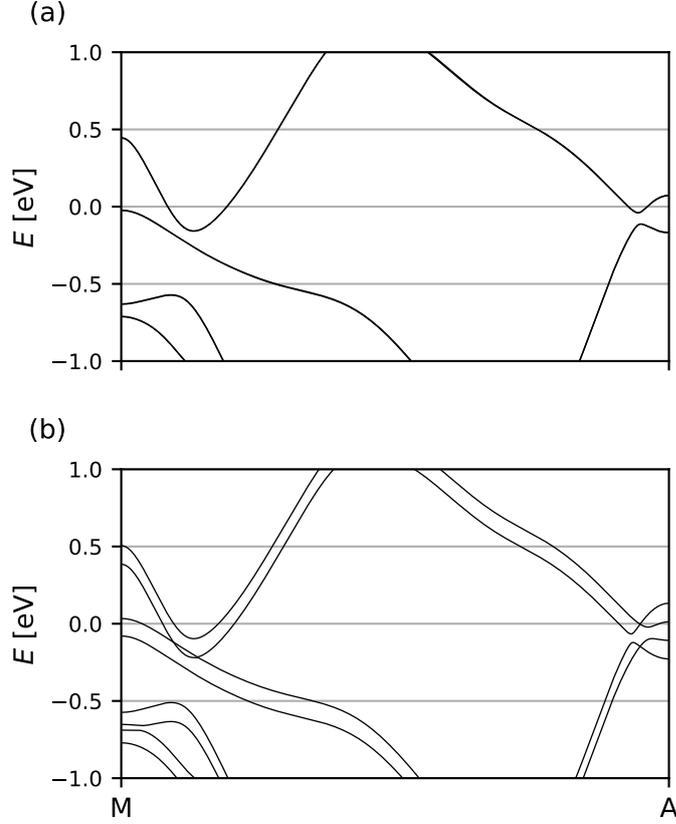


Figure 6.9: Band structure of the tight-binding model for NbSb<sub>2</sub> along the  $M - A$  line. (a) Without Zeeman splitting. (b) With  $c_y = 0.06$  eV Zeeman splitting.

where the change in the splitting terms (compared to eq. (6.14)) is due to the different orbital basis used for the tight-binding model. We search for Weyl points between the last valence band and the first conduction band.

First we study the effect of applying a magnetic field in the  $y$ -direction. Figure 6.9 shows the effect of this splitting along the  $M - A$  line. For  $c_y \approx 0.06$  eV, two pairs of Weyl points appear close to the  $M - A$  line. The reason these points are not exactly on the line is because the crystal symmetry is broken when constructing the Wannier-based tight-binding model [111, 112]. Apart from this numerical difference, this effect is analogous to the case of the  $\mathbf{k} \cdot \mathbf{p}$  model for TaAs<sub>2</sub>, where the two pairs of Weyl points appeared at  $c_y = 0.11$  eV.

Table 6.6 shows the positions, chirality and type of Weyl points for selected values of the Zeeman splitting. It shows that Weyl points appear even at smaller values of  $c_y$  away from the  $M - A$  line. This is a crucial difference to the  $\mathbf{k} \cdot \mathbf{p}$  model which is valid only near the  $M$  point. Furthermore, all Weyl points found for

these splitting values are of type II [2]. Type - II Weyl points have a tilted energy spectrum, making their Fermi surface open instead of point-like. As a consequence, their chiral anomaly – and their effect on magnetoresistance – is expected to be anisotropic.

Finally, a phase diagram showing the number of Weyl points as a function of magnetic field was calculated (see fig. 6.10). Unlike for the  $\mathbf{k} \cdot \mathbf{p}$  model, the number of Weyl points keeps increasing when the applied Zeeman term grows stronger. Again, the reason for this difference is that Weyl points also form far away from the  $M$  point, where the  $\mathbf{k} \cdot \mathbf{p}$  approximation is no longer applicable.

For some values of the splitting, the phase diagram shows an odd number of Weyl points, which is physically impossible. The reason for this is that the numerical procedure used to identify the number of Weyl points may not find a Weyl point if it is too close to another Weyl point. Since this problem occurs only rarely (see fig. 6.11), the phase diagram is still valid overall. Also, the procedure ensures that no Weyl point can be counted twice, so the phase diagram represents a lower limit for the real number of Weyl points. Thus, the general result that the number of Weyl points increases with stronger Zeeman splitting remains valid.

## 6.2.4 Limitations of the Model for Magnetic Field

In the previous sections the effect of magnetic field was modeled by applying a Zeeman splitting to the model Hamiltonian. The discussion was simplified by assuming that the g-factor is equal for all energy bands, and independent of  $\mathbf{k}$ . Here we discuss how the results might change if this assumption is not made.

If the g-factor is  $k$ -dependent, but still the same for all energy bands, the results above will change quantitatively, but not qualitatively. The reason for this is that a Weyl node that appears at a specific  $k$ -point will still be there, but for a different magnetic field. That is, the order in which the Weyl nodes at different  $k$ -points appear might change, but not the overall picture that there is an increasing number of Weyl points with stronger magnetic field.

The same is true if the g-factor varies for different energy bands, as long as the sign of the g-factor remains the same. Because the appearance of Weyl points is due to the relative Zeeman splitting between the last valence and first electron bands, it does not matter how much the splitting on each band contributes.

If the g-factors in the relevant bands have opposite sign however, there is a qualitative change in the behavior. This is illustrated in the following with the example of the  $\mathbf{k} \cdot \mathbf{p}$  model of TaAs<sub>2</sub> discussed in sections 6.1.3 and 6.2.2. To account for the opposite sign of the g-factor for valence and conduction bands, the Zeeman splitting term (eq. (6.14)) is changed to

$$\Delta\mathcal{H} = c_x\sigma_z \otimes \sigma_y + c_y\sigma_z \otimes \sigma_z + c_z\sigma_z \otimes \sigma_x. \quad (6.19)$$

Split. [eV]	Position $\mathbf{k}$	Chir.	Type
$c_x = 0.045$	(0.4393, 0.4460, 0.5004)	+1	II
	(0.4359, 0.4444, 0.5026)	-1	II
	(0.5641, 0.5556, 0.4974)	+1	II
	(0.5607, 0.5540, 0.4996)	-1	II
$c_y = 0.03$	(0.3670, 0.5141, 0.0977)	+1	II
	(0.3655, 0.5142, 0.1004)	-1	II
	(0.6345, 0.4858, 0.8997)	+1	II
	(0.6330, 0.4858, 0.9023)	-1	II
$c_y = 0.04$	(0.3724, 0.5116, 0.0890)	+1	II
	(0.3627, 0.5135, 0.1055)	-1	II
	(0.6373, 0.4865, 0.8945)	+1	II
	(0.6276, 0.4884, 0.9110)	-1	II
	(0.9028, 0.0340, 0.5451)	+1	II
	(0.9018, 0.0354, 0.5390)	-1	II
	(0.0982, 0.9646, 0.4610)	+1	II
	(0.0974, 0.9658, 0.4545)	-1	II
$c_y = 0.06$	(0.3791, 0.5068, 0.0775)	+1	II
	(0.3592, 0.5131, 0.1108)	-1	II
	(0.6407, 0.4869, 0.8892)	+1	II
	(0.6211, 0.4929, 0.9222)	-1	II
	(0.9033, 0.0328, 0.5532)	+1	II
	(0.9006, 0.0364, 0.5314)	-1	II
	(0.0994, 0.9636, 0.4686)	+1	II
	(0.0968, 0.9671, 0.4467)	-1	II
	(0.4493, 0.4555, 0.5031)	+1	II
	(0.4309, 0.4320, 0.4825)	-1	II
	(0.5691, 0.5680, 0.5175)	+1	II
	(0.5507, 0.5445, 0.4969)	-1	II
$c_z = 0.0475$	(0.4494, 0.4384, 0.4853)	+1	II
	(0.4420, 0.4366, 0.4816)	-1	II
	(0.5580, 0.5634, 0.5184)	+1	II
	(0.5506, 0.5616, 0.5147)	-1	II

Table 6.6: Weyl point positions (in reduced coordinates), chirality and type for different values of the Zeeman splitting in the tight-binding model for NbSb<sub>2</sub>.

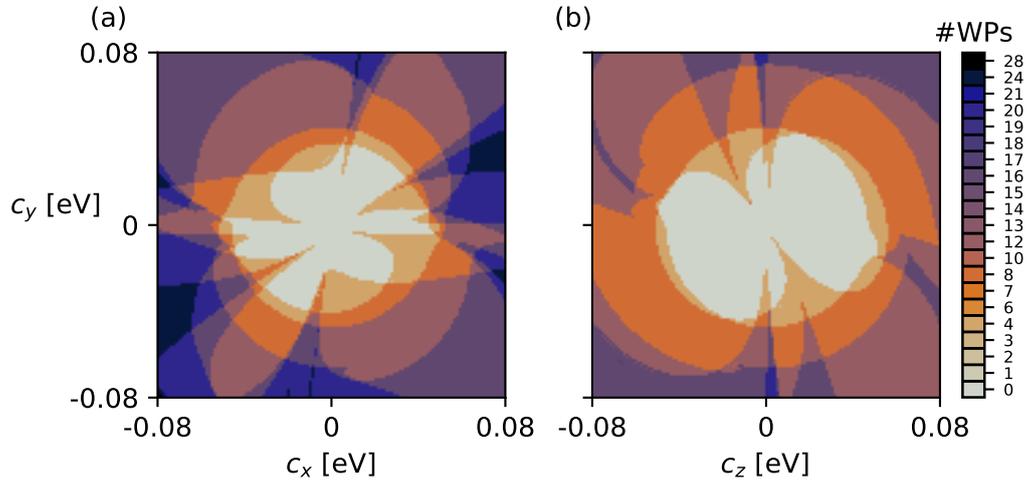


Figure 6.10: Phase diagram showing the number of Weyl points as a function of Zeeman splitting (in eV) for the tight-binding model for NbSb<sub>2</sub>.

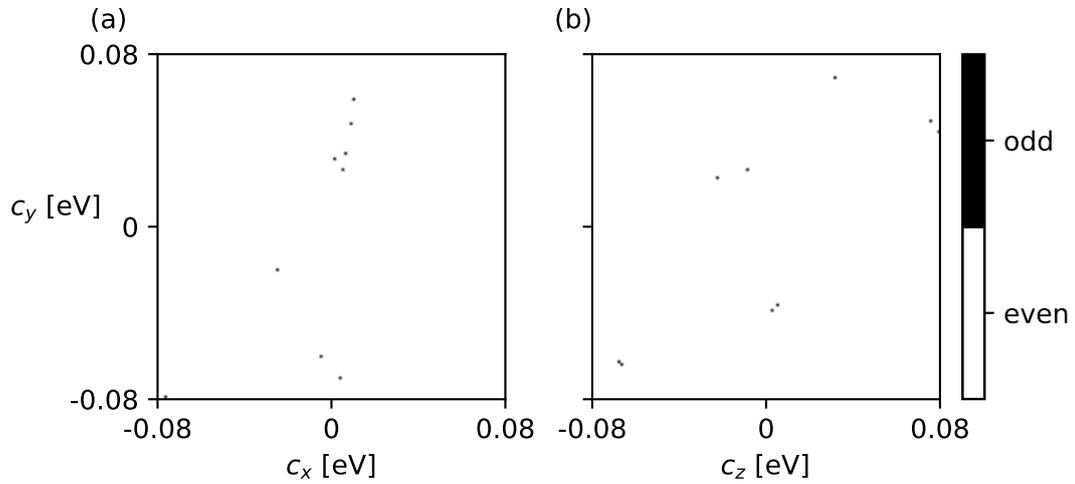


Figure 6.11: Phase diagram showing whether the number of Weyl points shown in fig. 6.10 is even (physically possible) and odd (non-physical). The odd phases are a relict of the numerical evaluation of the phase.

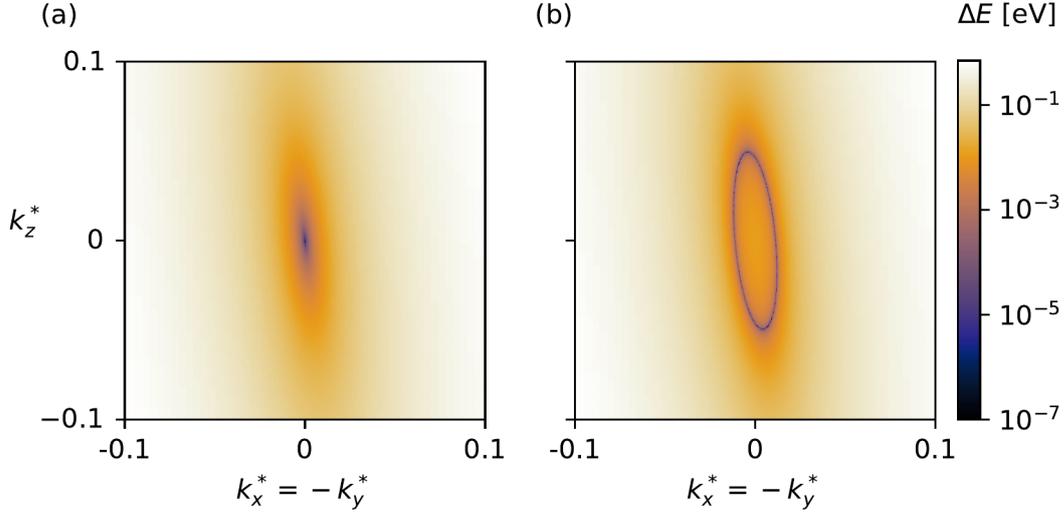


Figure 6.12: Band gap of TaAs<sub>2</sub> on the mirror plane with Zeeman splitting as given in eq. (6.19). (a) At  $c_y \approx 0.2242$  eV, a node appears at the  $M$ -point. (b) The node expands into a line for stronger splitting ( $c_y = 0.23$  eV).

With  $c_y$  splitting, the energy bands on the mirror plane is then given by

$$E(\mathbf{k}) = C_{00}(\mathbf{k}) \pm c_y \mp \sqrt{C_{xx}(\mathbf{k})^2 + C_{xy}(\mathbf{k})^2 + C_{z0}(\mathbf{k})^2}. \quad (6.20)$$

As in eq. (6.17), the Zeeman term counteracts the original splitting. The difference to the previous case is that this equation holds on an entire plane in reciprocal space instead of just a line. As consequence, we can expect the appearance of a nodal line with sufficient Zeeman splitting. Indeed, a nodal line appears for  $c_y \gtrsim 0.2242$  eV, as shown in fig. 6.12. The Berry phase on a closed path around this nodal line was calculated to be  $\pi$ , using the Z2Pack [6] code. This verifies the topological nature of the nodal line.

In conclusion, the qualitative result obtained above remains intact when the  $g$ -factors are assumed to be  $k$ -dependent and different for valence and conduction bands, as long as they keep the same sign. A more adequate model of the magnetic field is needed to establish the exact qualitative and quantitative nature of the topological phases with applied magnetic field. The current results indicate that Weyl nodes will appear at least for some directions of magnetic field.

### 6.3 Conclusions

We studied the topological phase of transition metal dipnictides of the type  $AB_2$  ( $A \in \{\text{Ta}, \text{Nb}\}$ ,  $B \in \{\text{As}, \text{Sb}\}$ ), with and without external magnetic field. In the absence of magnetic field, we found – in accordance with previous results [93, 95] – that these materials can be classified as weak topological insulators despite having an indirect band gap closure.

The effect of a magnetic field was studied by applying a Zeeman splitting. We found that Weyl points can appear. We showed this result first from theoretical considerations on a four-band  $\mathbf{k} \cdot \mathbf{p}$  model, and numerically on a  $\mathbf{k} \cdot \mathbf{p}$  model of  $\text{TaAs}_2$  and a tight-binding model of  $\text{NbSb}_2$ . In the tight-binding model, we found the number of Weyl points to be increasing with growing magnetic field. For specific values of the Zeeman splitting, the type of the Weyl points in the tight-binding model was studied, and they were all found to be of type II.

The appearance of such field-induced Weyl points could help explain the reduced or negative magneto-resistivity in these materials. However, it is unclear whether the Weyl points studied here appear at a magnetic field that is realistic to observe in experiments. Further studies, in particular to obtain a realistic  $g$ -factor and more reliable data for the direct band gap, are required to accurately estimate the required magnetic field. Furthermore, it is known that modeling a strong magnetic field with only Zeeman splitting is not sufficient, and a more accurate model should be considered. Finally, the effect of these Weyl points on the magnetoresistance should be calculated. This is influenced by the orientation of the type-II Weyl points, and their distance from the Fermi level.

Consequently, there are three open questions which require further investigation: First, whether the appearance of field-induced Weyl points is realistic in these  $AB_2$  compounds. Second, if these Weyl points do appear, whether they alone are responsible for the experimentally observed behavior of magneto-resistance or if there are other effects. Finally, whether there are other compounds which contain the same kind of field-induced Weyl points, possibly appearing already at weaker magnetic field.



# 7

## General Form of Symmetric $\mathbf{k}\cdot\mathbf{p}$ Models

In the previous chapter, we utilized a  $\mathbf{k}\cdot\mathbf{p}$  model to locally describe the Hamiltonian for TaAs<sub>2</sub>. An important point in this construction is that the  $\mathbf{k}\cdot\mathbf{p}$  Hamiltonian must be restricted to terms which are allowed by symmetries. Usually, the general form of such a *symmetric*  $\mathbf{k}\cdot\mathbf{p}$  Hamiltonian is calculated by hand. In this chapter, we describe how this process can be automated using a computer algebra system and present open-source implementation of the algorithm, the `kdotp-symmetry` code.

### 7.1 Symmetry-Constrained Hamiltonian

For a general Hamiltonian  $\mathcal{H}(\mathbf{k})$ , the constraint imposed by a symmetry operation  $g$  is given by [113]

$$\mathcal{H}(\mathbf{k}) = D(g)\mathcal{H}(g^{-1}\mathbf{k})D(g^{-1}), \quad (7.1)$$

where  $D(g)$  is the symmetry representation. Consequently, a Hamiltonian that respects a given symmetry group  $G$  must observe eq. (7.1) for all symmetries<sup>1</sup>  $g$  in  $G$ .

In the following, we will reformulate this condition in terms of linear operations on the space of possible Hamiltonian matrices. This allows for implementing an automated system to generate the set of allowed matrices.

---

<sup>1</sup>Note that it is a sufficient condition to observe eq. (7.1) for a generating subset of  $G$ .

First, we will consider the action of the symmetry (eq. (7.1)) on the  $\mathbf{k}$  - vector. We will limit the discussion to some finite set of basis functions, for example monomials of  $\{k_x, k_y, k_z\}$  up to second order. Let  $V \subset \mathcal{F}(\mathbb{R}^3, \mathbb{R})$  be the vector space which contains these functions of  $\mathbf{k}$ . We define

$$\begin{aligned}\hat{F}_g : V &\longrightarrow V \\ f &\longmapsto \tilde{f}_g \\ \tilde{f}_g(\mathbf{k}) &= f(g^{-1}\mathbf{k}),\end{aligned}\tag{7.2}$$

which is a linear operator on  $V$  that captures the action of the symmetry on functions of  $\mathbf{k}$ . We require that the choice of basis functions  $V$  is closed under  $\hat{F}_g$  for all symmetry operations. That is,

$$\forall g \in G, f \in V : \hat{F}_g(f) \in V.\tag{7.3}$$

Second, let us consider the space of hermitian  $N \times N$  matrices, which we will denote  $W$ . Again, we define a linear operator which captures the action of a symmetry as given in eq. (7.1):

$$\begin{aligned}\hat{G}_g : W &\longrightarrow W \\ A &\longmapsto D(g)AD(g^{-1}).\end{aligned}\tag{7.4}$$

This operator is unitary under the Frobenius inner product, which is defined as

$$\langle A, B \rangle_F = \text{tr}(A^\dagger B).\tag{7.5}$$

Since the Hamiltonian is a hermitian matrix, with  $\mathbf{k}$  - dependence as given by  $V$ , it can be expressed as a vector in  $V \otimes W$ . Using this notation, the symmetry constraint eq. (7.1) can be written as

$$\forall g \in G : \left( \hat{F}_g \otimes \hat{G}_g \right) (\mathcal{H}) = \mathcal{H},\tag{7.6}$$

and thus

$$\mathcal{H} \in \bigcap_{g \in G} \text{Eig} \left( \hat{F}_g \otimes \hat{G}_g, 1 \right).\tag{7.7}$$

In conclusion, the problem of finding the general form of the Hamiltonian which respects the symmetry group  $G$  is equivalent to calculating this subspace. In the next section, we will detail how this calculation is implemented in the `kdotp-symmetry` code.

The space of valid Hamiltonian matrices (eq. (7.7)) is calculated as follows: As a first step, the matrix form of the two linear operators  $\hat{F}_g$  (eq. (7.2)) and  $\hat{G}_g$  (eq. (7.6)) is calculated. These describe the action of the symmetry on functions of

$\mathbf{k}$  and hermitian matrices, respectively. The matrix form is then used to compute the cross product which describes the full symmetry operation on the Hamiltonian, and the invariant space  $\text{Eig}(\hat{F}_g \otimes \hat{G}_g)$  is calculated. This procedure is repeated for all symmetry operations  $g$  in a generating set of the full symmetry group  $G$ . Finally, the intersection of the invariant subspaces for all symmetry operations is computed to give the final result. In this section, the implementation of each of these steps in the `kdotp-symmetry` code is described in detail.

The code is a library written in the Python [114] programming language. Most notably, it uses the `sympy` [115] package for symbolic computations. The `numpy` [116] and `scipy` [110] packages are also used, for numerical calculations.

### 7.1.1 Calculating the Matrix Form of $\hat{F}_g$

The operator  $\hat{F}_g$  acts on the space of real-valued functions of  $\mathbf{k}$ . In `kdotp-symmetry`, a set of basis vectors, which are `sympy` expressions of  $k_x$ ,  $k_y$  and  $k_z$ , must be given to define this vector space  $V$ . The action of the symmetry  $g$  on this space is defined by the real-space rotation matrix  $S_g$ . For unitary symmetry operations, the  $\mathbf{k}$ -space matrix  $R_g$  is related to  $S_g$  by<sup>2</sup>

$$R_g = (S_g^T)^{-1}. \quad (7.8)$$

In the anti-unitary case, this relation reads

$$R_g = -(S_g^T)^{-1}. \quad (7.9)$$

Consequently, the expression  $g^{-1}\mathbf{k}$  evaluates to  $S_g^T\mathbf{k}$  and  $-S_g^T\mathbf{k}$  for the unitary and anti-unitary cases, respectively. By simultaneously substituting

$$k_i \mapsto (g^{-1}\mathbf{k})_i = \pm (S_g\mathbf{k})_i, \quad (7.10)$$

for  $i \in \{x, y, z\}$ , the operation  $\hat{F}_g$  can be applied on the basis vectors  $\{v_1, \dots, v_n\}$ . This is done using the `subs` method of `sympy` expressions.

The final step to obtain the matrix form of the operator  $\hat{F}_g$  is to express these transformed basis vectors  $\{\hat{F}_g v_1, \dots, \hat{F}_g v_n\}$  in terms of the original basis. This means solving for the coefficients  $\lambda_i$  in

$$\hat{F}_g v_i = \lambda_1^i v_1 + \dots + \lambda_n^i v_n. \quad (7.11)$$

Algebraically solving this equation is a challenging task, especially if complicated functions  $v_i$  are involved. The following procedure however has shown to work well

---

<sup>2</sup>Note that  $R_g$  and  $S_g$  are not necessarily unitary, since the coordinates can be given w.r.t. a non-orthogonal basis.

in practice: Equation (7.11) is evaluated (algebraically) for a set of random values of  $\mathbf{k}$ . Assuming that the initial basis vectors were indeed linearly independent, this will produce a system of linear equations that can be exactly solved for  $\lambda_j^i$ . To ensure that the result is not accidental, which might happen if the  $v_j$  are linearly dependent, or  $\hat{F}_g v_i$  is not in the space spanned by  $\{v_j\}_j$ , eq. (7.11) is then verified using the obtained values for  $\lambda_j^i$ .

Repeating this procedure for all transformed basis vectors  $\hat{F}_g v_i$ , the matrix form of  $\hat{F}_g$  can be obtained as

$$\mathcal{M}(\hat{F}_g) = \begin{pmatrix} \lambda_1^1 & \cdots & \lambda_1^n \\ \vdots & \ddots & \vdots \\ \lambda_n^1 & \cdots & \lambda_n^n \end{pmatrix}. \quad (7.12)$$

### 7.1.2 Calculating the Matrix Form of $\hat{G}_g$

The general procedure for calculating the matrix form of  $\hat{G}_g$  is the same as for  $\hat{F}_g$ . First the transformed basis vectors  $\hat{G}_g w_i$  are calculated, where  $\{w_1, \dots, w_m\}$  is the basis of  $W$ . This is simply a matter of applying eq. (7.6), with the given symmetry representation  $D(g)$ :

$$\hat{G}_g w_i = D(g) w_i D(g)^{-1} \quad (7.13)$$

Second, the transformed basis vectors need to be expressed in terms of the original basis

$$\hat{G}_g w_i = \mu_1^i w_1 + \dots + \mu_m^i w_m. \quad (7.14)$$

We require that the basis  $\{w_i\}_i$  of  $W$  is orthogonal with respect to the Frobenius inner product. The coefficients  $\mu_j^i$  can then be simply computed as

$$\mu_j^i = \frac{\langle \hat{G}_g w_i, w_j \rangle_F}{\langle w_j, w_j \rangle_F}. \quad (7.15)$$

The matrix form is then given as

$$\mathcal{M}(\hat{G}_g) = \begin{pmatrix} \mu_1^1 & \cdots & \mu_1^n \\ \vdots & \ddots & \vdots \\ \mu_n^1 & \cdots & \mu_n^n \end{pmatrix}. \quad (7.16)$$

### 7.1.3 Calculating the Invariant Subspace of the Symmetry Group

Given the matrix forms of both  $\hat{F}_g$  and  $\hat{G}_g$ , the total action of the symmetry on the Hamiltonian can be calculated as  $\mathcal{M}(\hat{F}_g) \otimes \mathcal{M}(\hat{G}_g)$ . The space of Hamiltonians

which are invariant under  $g$  can then be calculated as the nullspace of

$$\mathcal{M}(\hat{F}_g) \otimes \mathcal{M}(\hat{G}_g) - \mathbb{I}_{n \cdot m \times n \cdot m}, \quad (7.17)$$

using the `nullspace` method of `sympy` matrices. To validate this result, the dimension of the nullspace is compared against the number of eigenvalues close to zero calculated numerically using `numpy` and `scipy`.

Finally, the intersection of the invariant spaces for each symmetry operation  $g$  must be calculated. This is done using the Zassenhaus algorithm [117], which is implemented using `sympy`'s method of generating the reduced row echelon form, `sympy.Matrix.rref`.

### 7.1.4 Limitations

The current implementation of the `kdotp-symmetry` code is limited to symmorphic symmetry groups (see section 2.1). However, it should be possible to implement the non-symmorphic case without substantial changes to the architecture of the code.

## 7.2 Example Application

An example application of the `kdotp-symmetry` code is the  $\mathbf{k} \cdot \mathbf{p}$  model for  $\text{TaAs}_2$  used in section 6.1.3. In this section, we provide some additional details of how the basis for this model is computed.

As a basis of the space of hermitian  $4 \times 4$  matrices, we use the tensor products of Pauli matrices:

$$\mathcal{B}_W = \{\sigma_i \otimes \sigma_j | i, j \in \{0, x, y, z\}\} \quad (7.18)$$

Because the symmetry operators cannot change the order of a function of  $\mathbf{k}$ , we can consider only the subspace of  $V$  of a given order  $n$ , and use the monomial basis

$$\mathcal{B}_{V^n} = \{k_x^i k_y^j k_z^l | i, j, l \in \mathbb{N}_0; i + j + l = n\}. \quad (7.19)$$

This separation into subspaces significantly reduces the computational complexity of the task, since the basis size of the full space  $V \otimes W$  is reduced.

Since only a generating subset of the full symmetry group is required to produce the correct Hamiltonian form, we choose the  $C_{2y}$ , parity, and time-reversal symmetries, as given in eq. (6.3). Their corresponding representation matrices are given in eq. (6.4). Using these symmetries with the given bases for  $W$  and  $V$ , the `kdotp-symmetry` code produces the following basis for the Hamiltonian matrix, up to second order:

- Zeroth order:

$$\sigma_0 \otimes \sigma_0, \sigma_z \otimes \sigma_0 \quad (7.20)$$

- First order:

$$\begin{aligned} (k_x - k_y)\sigma_x \otimes \sigma_x, (k_x - k_y)\sigma_x \otimes \sigma_y, (k_x + k_y)\sigma_x \otimes \sigma_z, \\ (k_x + k_y)\sigma_y \otimes \sigma_z, k_z\sigma_x \otimes \sigma_x, k_z\sigma_x \otimes \sigma_y \end{aligned} \quad (7.21)$$

- Second order:

$$\begin{aligned} (k_x^2 + k_y^2)\sigma_0 \otimes \sigma_0, (k_x^2 + k_y^2)\sigma_z \otimes \sigma_0, k_x k_y \sigma_0 \otimes \sigma_0, k_x k_y \sigma_z \otimes \sigma_0, \\ (k_x k_z - k_y k_z)\sigma_0 \otimes \sigma_0, (k_x k_z - k_y k_z)\sigma_z \otimes \sigma_0, k_z^2 \sigma_0 \otimes \sigma_0, k_z^2 \sigma_z \otimes \sigma_0 \end{aligned} \quad (7.22)$$

This is exactly the basis used in section 6.1.3, up to some relabeling.

## 7.3 Conclusion

In this chapter, we have presented `kdotp-symmetry`, a tool for computing the general form of a Hamiltonian matrix for  $\mathbf{k} \cdot \mathbf{p}$  models under a given symmetry group. Since the calculation of such models by hand is cumbersome and error-prone, this should be a useful tool to simplify such calculations. The automation of the process is made possible by reformulating the problem as a linear algebra problem, which can then be solved using a computer algebra system.

# 8

## Automated Construction of Wannier Tight-binding Models

This chapter was previously published in reference [155].

A significant part of materials science is devoted to the problem of finding the electronic structure of a given material. As a result, numerous computational techniques have been developed to study this problem. These techniques can roughly be classified into two kinds: *First-principles* methods solve the problem using the fundamental physical principles and properties of atoms comprising the material. For weakly-interacting systems, density functional theory (DFT) [14] is the dominant (mean field) technique for solving the electronic structure problem from first principles.

In contrast, *empirical* methods aim to capture the relevant physical properties using a simplified model. Such models are usually matched to known properties of the material, which can be obtained from either experiments or first-principles calculations. An example of such an empirical method is given by the tight-binding approximation, which describes a material as a set of localized orbitals and pre-defined electron hopping terms between them. While the first-principles methods typically have superior accuracy, empirical methods are often used due to their lower computational cost. In particular, calculations of complex device geometries are often inaccessible to a direct first-principles study. As such, the construction of reliable empirical models is of significant importance. And the technique of creating Wannier tight-binding models [51, 118] from first-principles calculations is arguably one of the most popular tools in nowadays computational materials science. The use of Wannier tight-binding models allows one to combine the simplicity of empirical methods with the correct wave function properties obtained

---

from first-principles.

In recent years, *high-throughput* techniques made a profound impact in various fields of materials science [119–122]. While the domain eludes a strict definition, a common feature of such techniques is that computational tools are applied to a wide range of candidate materials, or variations of a given material, in search of some beneficial property. Existing codes and techniques are combined and applied on a scale that was not previously possible. A range of automated frameworks [123, 124] support this by facilitating the combination of separate calculations into logical workflows. The challenge in designing such a high-throughput workflow is to make it resilient to varying input parameters. Since the number of calculations performed is too large to be human-controlled, many decisions – for example which calculation to perform based on the output of a previous calculation – need to be encoded into the automated workflow.

In this chapter, we introduce steps for addressing two standardly known problems of using Wannier90 [111, 112] in combination with any *ab initio* software to construct tight-binding models: the absence of symmetries present in the original compound in the obtained tight-binding model, and the necessity to search for optimal inner and outer energy windows for projection of the first-principles energy bands. We do not, however, treat the issue of selecting the initial projections used by Wannier90. As such, we create automated workflows which are applicable to large classes of materials with similar orbital character of the bands of interest. However, these workflows are not yet applicable to high-throughput scenarios in the sense that they can trivially be applied to arbitrary compounds. Nevertheless, the presented workflows are written in a way that they could be combined with efforts to address the problem of selecting initial projections [125].

In section 8.1, we review the general process of calculating the Wannier tight-binding models by means of Wannier90 and explain the proposed and implemented symmetrization and automatic energy window choice procedures. Sec. 8.2 describes how these procedures are used for the development of an automated workflow using the AiiDA [123] framework. While this workflow automates the tight-binding calculation itself, there are still some tunable parameters which might be eliminated by a more sophisticated system. By using a modular design approach, we provide an extensible framework for implementing such improvements. In the final section, we illustrate the application of this workflow to calculate tight-binding models for strained III-V semiconductor materials. These are useful in the pursuit of Majorana devices [126–128], enabling the study of transport properties for different topological devices with III-V semiconductor quantum wells, where strains play an important role in the topological transition.

## 8.1 Construction of Wannier Tight-binding Models

In this section, we describe the process of generating *symmetrized* Wannier-like tight-binding (SWTB) models. First, we give a short description of the method as introduced in the works of Refs. [51, 118] and implemented in Wannier90 [111, 112] software package. Next, we describe a method for symmetrizing these WTBs in a post-processing step. Finally, we describe a scheme to enhance the band-structure accuracy by *optimizing the energy windows* used by Wannier90.

### 8.1.1 Wannier Tight-binding Construction

Tight-binding models represent a common way to describe crystalline systems in a computationally cheap way. The material is described as a system of localized orbitals with positions  $\mathbf{t}_i$  in the unit cell, and hopping terms  $H^{ij}[\mathbf{R}]$  between the  $j$ -th orbital in the unit cell at location  $\mathbf{R}$  and the  $i$ -th orbital in the home unit cell  $\mathbf{R} = \mathbf{0}$ . From these parameters, the matrix Hamiltonian can be written as <sup>1</sup>

$$\mathcal{H}^{ij}(\mathbf{k}) = \sum_{\mathbf{R}} H^{ij}[\mathbf{R}] e^{i\mathbf{k} \cdot (\mathbf{R} + \mathbf{t}_j - \mathbf{t}_i)}. \quad (8.1)$$

For the case of spinful systems, we choose the indices  $i, j$  to include the spin index for simplicity.

The Wannier tight-binding (WTB) method utilizes localized Wannier functions as basis orbitals to capture the compound’s physics. These basis Wannier functions are obtained from first-principles simulations. This procedure is based on the work of Refs. [51, 118] and implemented in the Wannier90 [111, 112] code. After obtaining the necessary Wannier90 input files from a first-principles calculation, two steps are performed to construct these Wannier functions:

In a first step, the Bloch wave-functions  $|\psi_{n,\mathbf{k}}\rangle$  calculated by the first-principles code are *disentangled* to obtain  $M$  wave-functions, where  $M$  is the target number of basis Wannier functions in WTB. For selecting the Bloch wave-functions which are involved in this procedure, one needs to choose an *outer* energy window. Optionally, an *inner* energy window can be chosen. States inside this inner window will be preserved by the disentanglement. An optimization routine is performed to select the  $M$  states such that the “change of character”  $\Omega_I$  (defined in Ref. [118]) is minimized. As an initial guess for this optimization procedure,  $M$  localized trial orbitals  $|g_m\rangle$  are used. Because the disentanglement procedure needs to discard some states, it usually changes both the symmetry and the energy bands of the model in comparison with first-principles results. Consequently, choosing good

<sup>1</sup>In this work, we use the tight-binding convention I of Ref. [35].

values for both the energy windows and the trial orbitals has a strong effect on the quality of the resulting model.

As a second (optional) step, another optimization is performed to find a unitary transformation such that the resulting Wannier functions are maximally localized [51]. Again, the trial orbitals  $|g_m\rangle$  are used to create an initial guess for this optimization. Typically, these orbitals are chosen to be those chemical atomic orbitals that contribute most to the bands of interest. A method for constructing Wannier orbitals without the need for such a guess is described in Ref. [125].

## 8.1.2 Symmetrization

An important feature of tight-binding models, especially for studying topological effects, is that they preserve certain crystal symmetries. For a given symmetry group  $G$ , the symmetry constraint on the Hamiltonian matrix is given by [113]

$$\forall g \in G : \mathcal{H}(\mathbf{k}) = D^{\mathbf{k}}(g)\mathcal{H}(g^{-1}\mathbf{k})D^{\mathbf{k}}(g^{-1}), \quad (8.2)$$

where  $D^{\mathbf{k}}(g)$  is the  $\mathbf{k}$ -dependent representation of the symmetry  $g$  from the group  $G$ . We define the  $\mathbf{k}$  - *independent* part  $D(g)$  of the representation as

$$D^{\mathbf{k}}(g) = e^{i\alpha_g \cdot \mathbf{k}} D(g), \quad (8.3)$$

where  $\alpha_g$  is the translation vector of the symmetry. In the following, we will assume that the representation matrix of *unitary* operations is given by

$$D^{\mathbf{k}}(g) = e^{i\alpha_g \cdot \mathbf{k}} D(g) = e^{i\alpha_g \cdot \mathbf{k}} U_g, \quad (8.4)$$

where  $\alpha_g$  is the translation vector and  $U_g$  is a unitary matrix. For anti-unitary operations, we assume they are of the form

$$D^{\mathbf{k}}(g) = e^{i\alpha_g \cdot \mathbf{k}} D(g) = e^{i\alpha_g \cdot \mathbf{k}} U_g \hat{K}, \quad (8.5)$$

where  $\hat{K}$  represents complex conjugation. A convenient property of the representation matrices is that

$$D^{\mathbf{k}}(g)AD^{\mathbf{k}}(g^{-1}) = D(g)AD(g^{-1}) \quad (8.6)$$

for any matrix  $A$ . In the unitary case, this is shown by

$$D^{\mathbf{k}}(g)AD^{\mathbf{k}}(g^{-1}) = e^{i\alpha_g \cdot \mathbf{k}} U_g A (e^{i\alpha_g \cdot \mathbf{k}} U_g)^{-1} \quad (8.7)$$

$$= e^{i\alpha_g \cdot \mathbf{k}} U_g A U_g^\dagger e^{-i\alpha_g \cdot \mathbf{k}} = U_g A U_g^\dagger = D(g)AD(g^{-1}), \quad (8.8)$$

and in the anti-unitary case by

$$D^{\mathbf{k}}(g)AD^{\mathbf{k}}(g^{-1}) = e^{i\alpha_g \cdot \mathbf{k}} U_g \hat{K} A (e^{i\alpha_g \cdot \mathbf{k}} U_g \hat{K})^{-1} \quad (8.9)$$

$$= e^{i\alpha_g \cdot \mathbf{k}} U_g \hat{K} A \hat{K} U_g^\dagger e^{-i\alpha_g \cdot \mathbf{k}} = U_g \hat{K} A \hat{K} U_g^\dagger = D(g)AD(g^{-1}). \quad (8.10)$$

### Symmetrized Hamiltonian as Group Average

For a Hamiltonian which does not fulfill these symmetry constraints, we define the *symmetrized* Hamiltonian as the group average

$$\tilde{\mathcal{H}}(\mathbf{k}) = \frac{1}{|G|} \sum_{g \in G} D^{\mathbf{k}}(g) \mathcal{H}(g^{-1}\mathbf{k}) D^{\mathbf{k}}(g^{-1}). \quad (8.11)$$

Following eq. (8.6), we can equivalently write the symmetrized Hamiltonian as

$$\tilde{\mathcal{H}}(\mathbf{k}) = \frac{1}{|G|} \sum_{g \in G} D(g) \mathcal{H}(g^{-1}\mathbf{k}) D(g^{-1}). \quad (8.12)$$

This procedure projects the Hamiltonian onto the symmetric subspace, meaning that the modified Hamiltonian respects eq. (8.2):

$$\begin{aligned} D^{\mathbf{k}}(g') \tilde{\mathcal{H}}([g']^{-1}\mathbf{k}) D^{\mathbf{k}}([g']^{-1}) &\stackrel{\text{eq. (8.6)}}{=} D(g') \tilde{\mathcal{H}}([g']^{-1}\mathbf{k}) D([g']^{-1}) \\ &= \frac{1}{|G|} \sum_{g \in G} D(g') D(g) \mathcal{H}(g^{-1}[g']^{-1}\mathbf{k}) D(g^{-1}) D([g']^{-1}) \\ &= \frac{1}{|G|} \sum_{g \in G} D(g'g) \mathcal{H}([g'g]^{-1}\mathbf{k}) D([g'g]^{-1}) \\ &\stackrel{g''=g'g}{=} \frac{1}{|G|} \sum_{g'' \in G} D(g'') \mathcal{H}([g'']^{-1}\mathbf{k}) D([g'']^{-1}) = \tilde{\mathcal{H}}(\mathbf{k}) \end{aligned} \quad (8.13)$$

Furthermore, if the original Hamiltonian is already symmetric, the original and symmetrized Hamiltonians are identical.

$$\begin{aligned} \tilde{\mathcal{H}}^{\text{symm.}}(\mathbf{k}) &= \frac{1}{|G|} \sum_{g \in G} D(g) \mathcal{H}^{\text{symm.}}(g^{-1}\mathbf{k}) D(g^{-1}) \\ &\stackrel{\text{eq. (7.1)}}{=} \frac{1}{|G|} \sum_{g \in G} \mathcal{H}^{\text{symm.}}(\mathbf{k}) = \mathcal{H}^{\text{symm.}}(\mathbf{k}) \end{aligned} \quad (8.14)$$

Since this construction does not explicitly construct the corresponding Wannier functions, we term these models *symmetrized Wannier-like* tight-binding models (SWTB).

It is important to note that the eigenstates and eigenvalues of the symmetrized Hamiltonian may differ significantly from those of the non-symmetrized Hamiltonian. In fact, for an anti-symmetric initial Hamiltonian, meaning that

$$D^{\mathbf{k}}(g) \mathcal{H}(g^{-1}\mathbf{k}) D^{\mathbf{k}}(g^{-1}) = -\mathcal{H}(\mathbf{k}) \quad (8.15)$$

for some symmetry  $g$ , the symmetrized result vanishes completely. However, given a Hamiltonian which *almost* respects the symmetry, this technique can effectively eliminate small symmetry-breaking terms.

In the context of tight-binding models, this symmetrization technique can only straightforwardly be applied when the underlying basis set is symmetric. If the tight-binding basis contains an orbital  $|\alpha\rangle$  centered around the position  $\mathbf{r}$ , it must also contain  $g|\alpha\rangle$  centered around  $g\mathbf{r}$  for all symmetries  $g \in G$ . For example, if the model for a material which has  $C_4^x$  symmetry contains a  $p_x$  orbital at the origin, it must also contain a  $p_y$  orbital at the origin.

For Wannier tight-binding models, this means that the technique can generally only be applied when the step of maximally localizing the Wannier functions is omitted, and pre-defined atomic orbitals are used. When this condition is met however, the method can be applied for both unitary and anti-unitary symmetries, as well as non-symmorphic symmetry groups.

### Hopping Matrices of the Symmetrized Hamiltonian

To apply the group average to tight-binding models, it is convenient to rewrite Eq. 8.11 directly in terms of the hopping matrices  $H[\mathbf{R}]$ . To obtain these, we first notice that  $D_{il}(g) \neq 0$  only if  $g\mathbf{t}_l - \mathbf{t}_i \in \mathbb{Z}^d$ , meaning that orbitals centered at  $\mathbf{t}_l$  are mapped onto  $\mathbf{t}_i$ , up to a possible lattice translation. Using eqs. (8.1) and (8.12), we can write the symmetrized Hamiltonian as

$$\tilde{\mathcal{H}}^{ij}(\mathbf{k}) = \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m \\ \mathbf{R}}} D_{il}(g) H^{lm}[\mathbf{R}] e^{i(g^{-1}\mathbf{k}) \cdot (\mathbf{R} + \mathbf{t}_m - \mathbf{t}_l)} D_{mj}(g^{-1}), \quad (8.16)$$

where the indices  $l, m$  only go over non-zero  $D_{il}(g)$  and  $D_{mj}(g^{-1})$ . Writing the real-space operator for  $g$  in Seitz notation [129]

$$g_{\mathbf{r}} = \{S_g | \boldsymbol{\alpha}_g\}, \quad g_{\mathbf{r}}^{-1} = \{S_g^{-1} | -S_g^{-1}\boldsymbol{\alpha}_g\}, \quad (8.17)$$

where  $S_g$  is the rotational part, and  $\boldsymbol{\alpha}_g$  is the translation vector of the symmetry, this means that

$$g\mathbf{t}_l - \mathbf{t}_i = S_g\mathbf{t}_l + \boldsymbol{\alpha}_g - \mathbf{t}_i \in \mathbb{Z}^d \quad (8.18)$$

$$g^{-1}\mathbf{t}_j - \mathbf{t}_m = S_g^{-1}\mathbf{t}_j - S_g^{-1}\boldsymbol{\alpha}_g - \mathbf{t}_m \in \mathbb{Z}^d \Rightarrow \mathbf{t}_j - \boldsymbol{\alpha}_g - S_g\mathbf{t}_m \in \mathbb{Z}^d \quad (8.19)$$

$$\stackrel{-(8.18)-(8.19)}{\Rightarrow} \mathbf{T}_{ij}^{ml} = S_g(\mathbf{t}_m - \mathbf{t}_l) - (\mathbf{t}_j - \mathbf{t}_i) \in \mathbb{Z}^d. \quad (8.20)$$

Next, we must consider how  $g$  acts on the reciprocal-space vector  $\mathbf{k}$ . For an (anti-)unitary operator, we know that

$$\langle \mathbf{k}, \mathbf{r} \rangle = \pm \langle g\mathbf{k}, g\mathbf{r} \rangle, \quad (8.21)$$

where the positive (negative) sign corresponds to the unitary (anti-unitary) case. Since  $g$  acts on  $\mathbf{r}$  with  $S_g$ , it follows that

$$\begin{aligned}\langle \mathbf{k}, \mathbf{r} \rangle &= \pm \langle A\mathbf{k}, S_g\mathbf{r} \rangle \\ \mathbf{k}^T \mathbf{r} &= \pm \mathbf{k}^T A^T S_g \mathbf{r},\end{aligned}\tag{8.22}$$

where  $A$  is the operation which acts upon  $\mathbf{k}$  when  $g$  is applied. Since this is true for all  $\mathbf{k}$  and  $\mathbf{r}$ ,

$$A = \pm (S_g^T)^{-1},\tag{8.23}$$

and thus

$$g^{-1}\mathbf{k} = \pm S_g^T \mathbf{k}.\tag{8.24}$$

For the next step, we treat the unitary and anti-unitary cases separately for clarity.

### 1. Unitary case

By applying eq. (8.24) to eq. (8.16), we get

$$\begin{aligned}\tilde{\mathcal{H}}^{ij}(\mathbf{k}) &= \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m \\ \mathbf{R}}} D_{il}(g) H^{lm}[\mathbf{R}] e^{i(S_g^T \mathbf{k}) \cdot (\mathbf{R} + \mathbf{t}_m - \mathbf{t}_l)} D_{mj}(g^{-1}) \\ &= \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m \\ \mathbf{R}}} D_{il}(g) H^{lm}[\mathbf{R}] e^{i\mathbf{k} \cdot [S_g(\mathbf{R} + \mathbf{t}_m - \mathbf{t}_l)]} D_{mj}(g^{-1}).\end{aligned}\tag{8.25}$$

Applying eq. (8.4), we obtain

$$\begin{aligned}\tilde{\mathcal{H}}^{ij}(\mathbf{k}) &= \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m \\ \mathbf{R}}} (U_g)_{il} H^{lm}[\mathbf{R}] e^{i\mathbf{k} \cdot [S_g(\mathbf{R} + \mathbf{t}_m - \mathbf{t}_l)]} (U_g^\dagger)_{mj} \\ &= \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m \\ \mathbf{R}}} D_{il}(g) H^{lm}[\mathbf{R}] D_{mj}(g^{-1}) e^{i\mathbf{k} \cdot [S_g(\mathbf{R} + \mathbf{t}_m - \mathbf{t}_l)]}.\end{aligned}\tag{8.26}$$

## 2. Anti-unitary case

In the anti-unitary case, we get

$$\begin{aligned}\tilde{\mathcal{H}}^{ij}(\mathbf{k}) &= \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m \\ \mathbf{R}}} D_{il}(g) H^{lm}[\mathbf{R}] e^{i(-S_g^T \mathbf{k}) \cdot (\mathbf{R} + \mathbf{t}_m - \mathbf{t}_l)} D_{mj}(g^{-1}) \\ &= \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m \\ \mathbf{R}}} D_{il}(g) H^{lm}[\mathbf{R}] e^{-i\mathbf{k} \cdot [S_g(\mathbf{R} + \mathbf{t}_m - \mathbf{t}_l)]} D_{mj}(g^{-1}).\end{aligned}\quad (8.27)$$

Applying eq. (8.5) to eq. (8.27), we get

$$\begin{aligned}\tilde{\mathcal{H}}^{ij}(\mathbf{k}) &= \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m \\ \mathbf{R}}} (U_g)_{il} \hat{K} H^{lm}[\mathbf{R}] e^{i\mathbf{k} \cdot [-S_g(\mathbf{R} + \mathbf{t}_m - \mathbf{t}_l)]} \hat{K} (U_g^\dagger)_{mj} \\ &= \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m \\ \mathbf{R}}} (U_g)_{il} \hat{K} H^{lm}[\mathbf{R}] \hat{K} (U_g^\dagger)_{mj} e^{i\mathbf{k} \cdot [S_g(\mathbf{R} + \mathbf{t}_m - \mathbf{t}_l)]} \\ &= \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m \\ \mathbf{R}}} D_{il}(g) H^{lm}[\mathbf{R}] D_{mj}(g^{-1}) e^{i\mathbf{k} \cdot [S_g(\mathbf{R} + \mathbf{t}_m - \mathbf{t}_l)]}.\end{aligned}\quad (8.28)$$

We observe that the result is the same for the unitary and anti-unitary cases, and treat them together in the following. Next, we substitute  $\mathbf{t}_m - \mathbf{t}_l$  using  $\mathbf{T}_{ij}^{ml}$  defined above, and define  $\mathbf{R}' = S_g \mathbf{R} + \mathbf{T}_{ij}^{ml}$ . Since  $\mathbf{R}'$  is again a lattice vector, we can change the summation from  $\mathbf{R}$  to  $\mathbf{R}'$ :

$$\tilde{\mathcal{H}}^{ij}(\mathbf{k}) = \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m \\ \mathbf{R}}} D_{il}(g) H^{lm}[\mathbf{R}] D_{mj}(g^{-1}) e^{i\mathbf{k} \cdot [S_g \mathbf{R} + \mathbf{T}_{ij}^{ml} + \mathbf{t}_j - \mathbf{t}_l]} \quad (8.29)$$

$$= \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m \\ \mathbf{R}'}} D_{il}(g) H^{lm}[S_g^{-1}(\mathbf{R}' - \mathbf{T}_{ij}^{ml})] D_{mj}(g^{-1}) e^{i\mathbf{k} \cdot (\mathbf{R}' + \mathbf{t}_j - \mathbf{t}_l)}. \quad (8.30)$$

Finally, we again use eq. (8.1) to obtain the symmetrized real-space hopping matrices

$$\tilde{H}^{ij}[\mathbf{R}'] = \frac{1}{|G|} \sum_{\substack{g \in G \\ l, m}} D_{il}(g) H^{lm}[S_g^{-1}(\mathbf{R}' - \mathbf{T}_{ij}^{ml})] D_{mj}(g^{-1}). \quad (8.31)$$

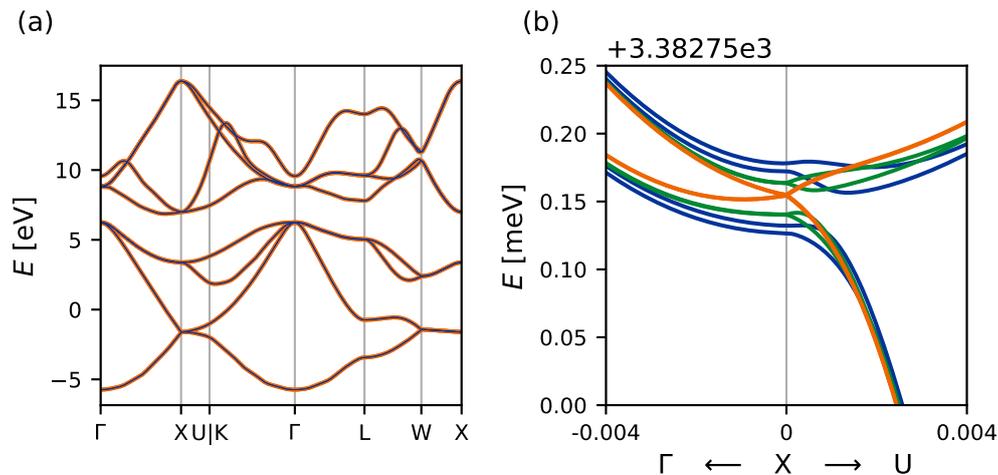


Figure 8.1: Comparison of the initial (blue) and symmetrized (orange) band structure for a tight-binding model of silicon with atom-centered  $sp^3$  orbitals. (a) In the eV scale, there are no visible differences between the two models. (b) A zoom in around the  $X$  point on the meV scale reveals a slight lifting of the band degeneracies in the initial model. This incorrectness is resolved in the symmetrized model. For comparison, a symmetrized band structure taking into account only symmorphic symmetries (green) is also shown.

## Example Results

Fig. 8.1 shows the results of this symmetrization procedure on a tight-binding model for bulk silicon, in the diamond cubic crystal structure, with atom-centered  $sp^3$  orbitals. The initial model already approximately fulfills the symmetry condition, which is reflected in the fact that the band structure does not change in the electronvolt scale. However, at the sub-millielectronvolt scale the band degeneracies are lifted in the original model, but restored after the symmetrization procedure. Since the symmetry group of the diamond cubic structure  $Fd\bar{3}m$  (no. 227) is non-symmorphic, this example demonstrates that the symmetrization technique is capable also of enforcing such symmetries. In panel b of fig. 8.1, we compare the symmetrization using the full symmetry group to a partial symmetrization enforcing only the symmorphic subgroup. Adding non-symmorphic symmetries enforces the four-fold degeneracy at the  $X$  point and two-fold degeneracy on the  $X-U$  line, whereas symmorphic symmetries only enforce a two-fold degeneracy on the  $\Gamma-X$  line.

## Construction of Matrix Representations

To determine the matrix representations  $D(g)$ , we use the fact that Wannier90 allows one to manually choose the trial orbitals  $|g_m\rangle$ . As a result, the basis after the disentanglement procedure corresponds to the chosen orbitals, up to some numerical error. Since the behavior of the basis orbitals under symmetries is known,  $D(g)$  can be determined in this way. For the treatment of spin, we use the rotation matrices as given in ref. [130]. The action of time-reversal on the spin basis  $\{|\uparrow\rangle, |\downarrow\rangle\}$  is given by  $\sigma_y \hat{K}$ , where  $\hat{K}$  represents complex conjugation. An automated method for generating the representation matrices for given atomic orbitals is available in the `symmetry-representation` package. Importantly, we used Wannier90 *without* performing the maximal localization step. It is the case in the illustrated application of Sec. 8.3, where this allows us to preserve the orbital basis. Alternatively, one could use the basis transformation matrices  $U^{(\mathbf{k})}$  provided by Wannier90 [111] to transform  $D(g)$  into the maximally-localized basis. While this approach produces computationally cheaper localized models, the drawback is that the basis is different for each produced tight-binding model. As a result, comparing models is more difficult. Also, linear interpolation between models, as described in section 8.3.3, would require a change of basis.

## Alternative Methods

Another approach to obtaining symmetric tight-binding models is to use the site-symmetry mode implemented in Wannier90 [131]. However, this method is limited to symmetries which leave a given real-space coordinate invariant (site symmetries), and does not include time-reversal. The method presented here has no such limitation, but is instead limited to models which have a symmetric set of basis functions as described above. The site-symmetry mode also relies on obtaining the symmetry information from the first-principles code, which is currently implemented only for Quantum Espresso [31, 32]. The workflow described in Sec. 8.2 could be adapted to allow using this approach with only minimal changes.

### 8.1.3 Optimization for Band Structure Fit

As described above, an important parameter in running Wannier90 is the choice of the so-called energy windows [111]. There are two such windows: The *outer* window determines which states are taken into account for the disentanglement procedure. At every  $\mathbf{k}$ -point, it must contain at least  $M$  bands, where  $M$  is the desired number of bands in the tight-binding model. The *inner* (or frozen) window on the other hand determines which states should not be modified during disentanglement. It can contain at most  $M$  bands at any given  $\mathbf{k}$ .

Since the quality of the resulting tight-binding model depends sensitively on the choice of energy windows, a strategy for reliably choosing good windows is required. A straightforward way of achieving this is by iteratively optimizing the window values. Having constructed and symmetrized a tight-binding model, its quality can be determined by comparing its band structure to a reference computed directly from first-principles <sup>2</sup>. As a measure of their mismatch, we choose the average difference between the energy eigenvalues

$$\Delta = \frac{1}{M} \frac{1}{N_{\mathbf{k}}} \sum_{i=1}^M \sum_{\mathbf{k}} |\varepsilon_{i,\mathbf{k}}^{\text{DFT}} - \varepsilon_{i,\mathbf{k}}^{\text{TB}}|. \quad (8.32)$$

Some values of the energy windows cannot produce a tight-binding model, for example if the outer window contains less than  $M$  bands. As a result, finding appropriate energy windows is a constrained, four-dimensional optimization problem. The Nelder-Mead (downhill simplex) algorithm [132] can be used to solve this problem <sup>3</sup>.

Fig. 8.2 shows the result of such an optimization procedure, for unstrained InSb as described in section 8.3. A clear improvement is visible between the tight-binding model obtained with the initial windows chosen by hand (panel a), and the optimized window values (panel b). In particular, the conduction bands at the  $X$  and  $Z$  points are represented more accurately in the optimized model. Since the given bands for InSb are not entangled, it is also possible to skip the disentanglement step completely by using the `exclude_bands` parameter of Wannier90 to ignore all other energy bands. The resulting band structure is shown in fig. 8.2(c). Nevertheless, we find that the band structure using optimized disentanglement is slightly better ( $\Delta = 0.0327$ ) than the one without disentanglement ( $\Delta = 0.0375$ ), especially for the four lowest conduction bands on the  $Z - \Gamma - X$  line. Hence, it can be useful to apply the disentanglement procedure and energy window optimization even in cases where the bands are not inherently entangled, especially when the time required to run the tight-binding calculation is short compared to the initial first-principles calculation.

## 8.2 Implementation in AiiDA Workflows

The AiiDA [123] platform is a Python [114] framework for performing high-throughput calculations, focused on the field of materials physics. It enables reproducible research by keeping track of inputs, outputs and settings for each calculation. On top

---

<sup>2</sup>Because the first-principles calculation usually contains more than  $M$  bands, we need to choose which bands should be represented by the tight-binding model.

<sup>3</sup>The constraint is implemented by assigning an infinite value of  $\Delta$  to invalid energy windows.

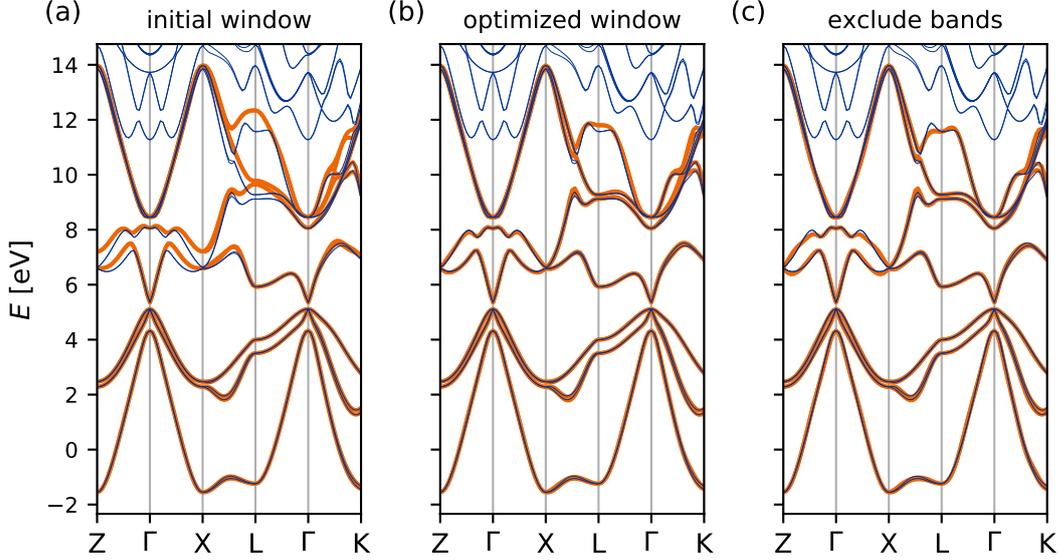


Figure 8.2: Comparison between the reference first-principles band structure (blue) and band structures calculated from tight-binding models (orange) for InSb. The tight-binding model in (a) was calculated with the initial energy window, whereas (b) shows the model using the optimized energy window as detailed in table 8.2. The model in (c) was calculated without the disentanglement procedure, using the `exclude_bands` parameter of Wannier90.

of this so-called *provenance* layer, it provides a toolset for automatically chaining calculations into user-defined workflows.

In this section, we describe the implementation of the Wannier tight-binding extraction scheme as an AiiDA workflow. This automation enables the high-throughput application to the study of strain effects (described in section 8.3). Special care has been taken to design the workflow in a modular way, which enables re-using parts of the workflow for purposes other than tight-binding extraction. We first discuss these design principles, before showing how they are applied in the tight-binding workflows.

The code for the AiiDA workflows is available in the open-source `aiida-tbextraction` package.

### 8.2.1 Modular Workflow Design

The basic principle of modular workflow design is to split up a single monolithic workflow into minimal sub-workflows or calculations that perform exactly one task. For example, the tight-binding model created by Wannier90 is post-processed by

parsing it to an HDF5 format, followed by optionally changing the order of the basis and symmetrizing the model. While this could easily be implemented in a single script, splitting these three steps up into separate calculations allows separately re-using each of the steps.

More complex workflows are created by combining multiple sub-workflows into a logical unit at a higher abstraction level. Inputs to the sub-workflow are either forwarded directly from the input to the parent workflow or created within the parent workflow. Similarly, outputs from the sub-workflow can either be forwarded to be an output of the parent workflow or consumed directly to guide the further execution of the parent workflow.

Since a complex workflow can consist of multiple layers of wrapped sub-workflows, this modular approach is maintainable only if the overhead of forwarding input and output is minimal. Following the single responsibility principle, a parent workflow should not have to change if an input or output parameter of a sub-workflow changes, unless it directly interacts with this parameter. To achieve this, a syntax is needed to specify that a parent workflow will *inherit* inputs or outputs of a sub-workflow, without explicitly listing each parameter. In AiiDA, such a feature is available in the newly-introduced *expose* functionality. It allows implicitly forwarding input and output values of a sub-workflow instead of having to explicitly specify each value. Listing 8.1 shows a simple workflow with two inputs **a** and **b**, and one output **c**. A parent workflow that only wraps this workflow is shown in listings 8.2 and 8.3 with and without using the *expose* functionality, respectively. Import statements are omitted in all listings for brevity.

---

```

1 class SubWF(WorkChain):
2     @classmethod
3     def define(cls, spec):
4         spec.input('a', valid_type=Int)
5         spec.input('b', valid_type=Int)
6         spec.output('c', valid_type=Int)
7         ...

```

---

Listing 8.1: A simple workflow with inputs **a** and **b**, and output **c**. The steps executing the workflow are omitted.

---

```

1 class ParentWF(WorkChain):
2     @classmethod
3     def define(cls, spec):
4         spec.expose_inputs(SubWF)
5         spec.expose_outputs(SubWF)
6
7         spec.outline(cls.invoke_subwf, cls.write_outputs)
8
9

```

---

---

```

10     def invoke_subwf(self):
11         return ToContext(
12             sub_wf=self.submit(SubWF, **self.exposed_inputs(SubWF))
13         )
14
15     def write_outputs(self):
16         self.out_many(self.exposed_outputs(self.ctx.sub_wf))

```

---

Listing 8.2: A workflow that wraps SubWF by using the *expose* functionality.

---

```

1 class ParentWF(WorkChain):
2     @classmethod
3     def define(cls, spec):
4         spec.input('a', valid_type=Int)
5         spec.input('b', valid_type=Int)
6         spec.output('c', valid_type=Int)
7
8         spec.outline(cls.invoke_subwf, cls.write_outputs)
9
10    def invoke_subwf(self):
11        return ToContext(
12            sub_wf=self.submit(
13                SubWF,
14                a=self.inputs.a,
15                b=self.inputs.b
16            )
17        )
18
19    def write_outputs(self):
20        self.out('c', self.ctx.sub_wf.out.c)

```

---

Listing 8.3: A workflow that wraps SubWF without using the *expose* functionality.

The modular architecture improves not only the re-usability, but also the flexibility of workflows. Often, a given part of a workflow could be performed in different ways. For example, many different codes can perform the first-principles calculations in the tight-binding extraction workflows. Additionally, one might want to add steps such as relaxation or cut-off energy convergence.

To allow for this, the parent workflow can allow for dynamically selecting a workflow for performing a given task by passing it as an input <sup>4</sup>. An abstract workflow class defines the interface that a workflow must fulfill so that it can be used to perform the task. If needed, the parent workflow can allow for dynamic inputs, which are just forwarded to the specific workflow implementing the in-

---

<sup>4</sup>For storing the workflow in the AiiDA database, it needs to be converted into an AiiDA data type. We chose to convert it into a string containing the fully qualified class name, from which we import the workflow when needed.

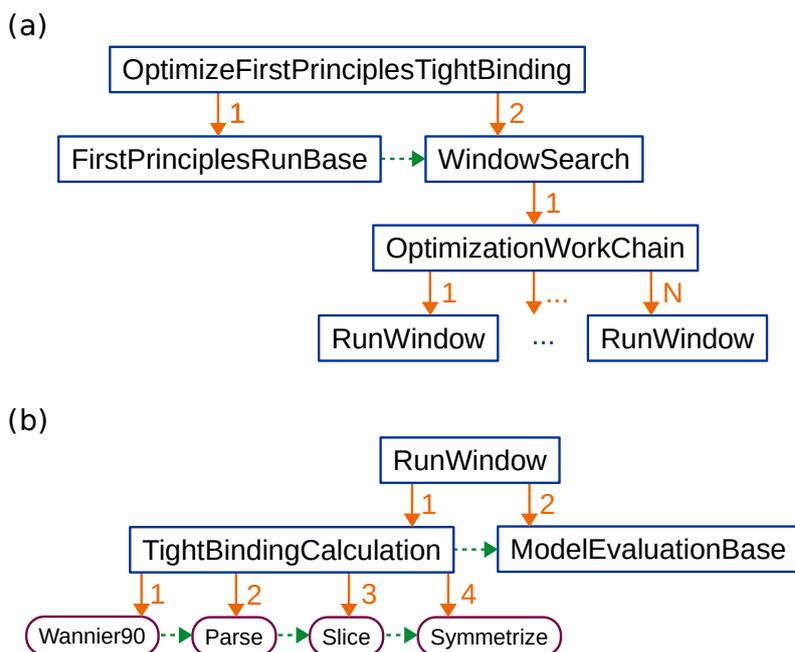


Figure 8.3: Schematic of the AiiDA workflow for creating tight-binding models with energy window optimization. Workflows are shown in blue, and calculations in purple. Orange arrows show calls from parent- to child-workflows (or calculations). Dashed green arrows show the implicit data dependency between workflows of the same level. In calculation names, the suffix `Calculation` is omitted for brevity.

terface. In this way, the parent workflow can act as a template that defines an abstract series of steps, without knowledge of the detailed input flags available on each step.

## 8.2.2 Tight-binding Extraction Workflow

Having discussed the design principles for modular workflows, we now show how these are applied to create a workflow for the construction of tight-binding models. This workflow is implemented in the `OptimizeFirstPrinciplesTightBinding` class as sketched in fig. 8.3. At the uppermost level, the workflow has two parts: `FirstPrinciplesRunBase`, which executes the first-principles calculations, and `WindowSearch` which calculates the tight-binding model with energy window optimization.

Since different first-principles codes can produce the input files required by `Wannier90`, `FirstPrinciplesRunBase` defines only the minimum interface needed to perform this task. As described in the previous section, a workflow that implements this interface for a specific first-principles code can then be chosen dynami-

cally. As a result, the subsequent parts of the workflow are independent of which first-principles code is used.

The `WindowSearch` workflow performs the Nelder-Mead algorithm [132] for finding the optimal energy window. Because optimization schemes are useful outside of this specific application, we implemented the Nelder-Mead method in a general way. The `OptimizationWorkChain`, defined in the `aiida-optimize` module, can be used to solve generic optimization problems in the context of AiiDA workflows. It requires two inputs: A workflow which defines the function to be optimized, and an engine that implements the optimization method. Consequently, changing the whole workflow to use a different optimization method would be a simple matter of using a different engine.

Because AiiDA workflows need to be able to stop and re-start after any given step, the engine is written in an object-oriented instead of a procedural way. While this complicates implementing the Nelder-Mead method, it allows for serializing and storing the state of the engine.

The function which is optimized by the `OptimizationWorkChain` is implemented in the `RunWindow` workflow. It again consists of two parts: `TightBindingCalculation` creates the tight-binding model itself, and `ModelEvaluationBase` evaluates the quality of the model. The first step in the `TightBindingCalculation` workflow is to run Wannier90 on the given input parameters. In a second step, the Wannier90 output is parsed and converted into the TBmodels [36] HDF5 format. A third, optional, “slicing” step is used to either permute the basis orbitals or discard some orbitals. Finally, the (also optional) symmetrization procedure is performed. Both the `Slice` and the `Symmetrize` calculation have a TBmodels HDF5 file as both input and output, meaning that they could be chained arbitrarily with other such post-processing steps.

For the evaluation of the tight-binding model, we again use an abstract interface class, `ModelEvaluationBase`. While for the purposes of this thesis we used the average difference of band energies (eq. (8.32)) as a measure of model quality, other quantities might be more appropriate for different applications.

## 8.3 Strain-dependent Tight-binding Models for Majorana Devices

The quest for Majorana zero modes (MZMs) in condensed matter systems has recently attracted a lot of interest [126–128, 133–137]. The non-abelian exchange statistics of Majorana Fermions makes these zero modes promising candidates for the realization of topological quantum computation devices [126, 138]. Experimental investigations of possible MZMs focus on the proposal by Lutchyn et. al. and

Oreg et. al. [127, 128] in which MZMs appear on the boundaries of proximitized spin-orbit coupled quantum wires. Current experimental setups include semiconducting InAs nanowires with epitaxial superconducting Al [139], and InAs/GaSb heterostructures in which the quantum spin Hall effect [38, 140] can be realized providing the possibility to proximity couple the helical edge state [134, 136]. While there is a good deal of evidence suggesting that MZMs exist in the wire-based setups [141, 142], a conclusive proof requires directly showing the braiding statistics of MZMs. An important step in realizing braiding with the systems based on the helical edge state is the search for optimized device and material properties. For optimizing the topological gap, a better theoretical understanding of the electronic structure in such devices is required. In this section, we show how the workflows can be used to generate tight-binding models which form the basis for accurate device simulations. While these device simulations themselves are outside the scope of this work, this shows the potential use of the method for a topic of active research in current condensed matter physics.

Highly accurate first-principles methods, using hybrid functionals [18], or the *GW* approximation [143], are computationally too demanding for the simulation of realistic device geometries and heterostructures. State of the art simulations of such structures use the  $\mathbf{k}\cdot\mathbf{p}$  method [37], or empirical tight-binding (ETB) methods [144]. In both of these methods the Hamiltonian is parametrized by a small number of parameters which are obtained empirically, for example via fitting to the first-principles band structure. For both of these methods the choice of parameters is ambiguous and one can obtain a good fit of the band structure while at the same time the electronic wavefunction might be wrongly represented. This might lead to unphysical solutions in confined geometries [145, 146], and low transferability of the bulk models to the heterostructure in general. Recently, it was shown that better matching the ETB with the first-principles calculations can improve their transferability [146, 147].

Realistic simulations of heterostructures require a correct treatment of strains at interfaces. In the  $\mathbf{k}\cdot\mathbf{p}$  and the ETB method this is usually done by strain-dependent parameter sets. However, often the symmetries are not broken correctly. In this context, the Wannier or Wannier-like tight-binding models can offer a significant improvement by accurately representing the first-principles wavefunction and correctly capturing the effect of strain. As a demonstration of the AiiDA workflows, we construct SWTB models for the III-V semiconductors InSb, InAs and GaSb.

Including spin-orbit coupling (SOC), we require only 14 basis functions, namely  $s$  and  $p$  orbitals centered on the In/Ga atom, and  $p$  orbitals centered on the As/Sb atom. The popular  $sp^3d^5s^*$  ETB models on the other hand require 40 [148] basis functions. The reason for this is that WTB models generally include longer-range neighbor interactions, whereas ETB is typically limited to nearest-neighbor

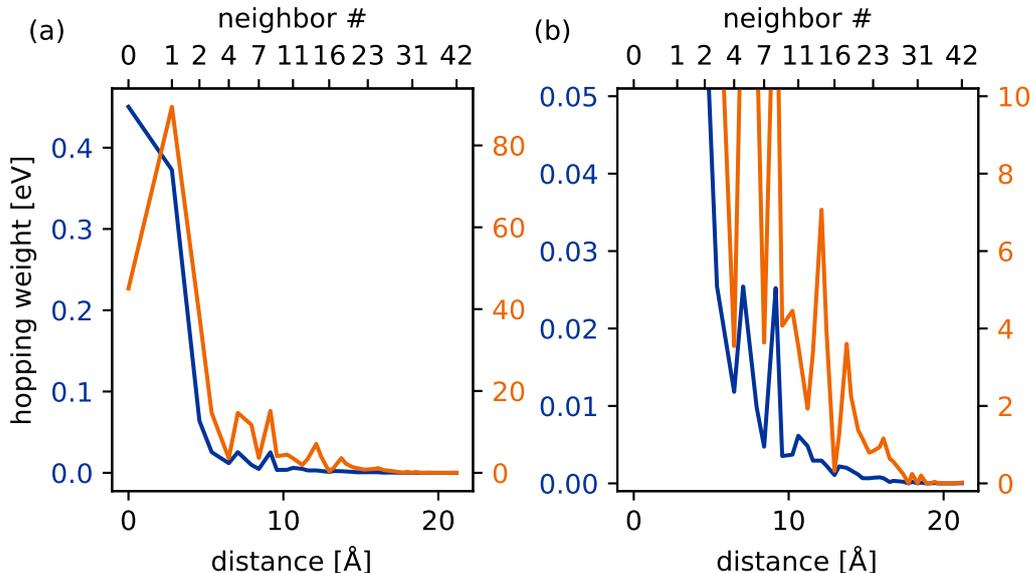


Figure 8.4: Average (blue, left axis) and total (orange, right axis) weights of the hopping parameters for the unstrained InSb tight-binding model, as a function of distance.

(or next-nearest-neighbor in some cases [149]) interactions to keep the number of parameters manageable. As illustrated in fig. 8.4, the produced tight-binding models include long-range hopping parameters, with amplitudes quickly decaying with distance.

To account for strain, we construct tight-binding models with biaxial (001), (110) and (111) strains, and the uniaxial [110] strain. For each material and strain direction, we calculated 16 models in the range of  $\pm 4\%$  strain. Including the unstrained models, we constructed a total of 195 tight-binding models, showing the applicability of the AiiDA workflow to a large number of chemically and structurally similar compounds.

### 8.3.1 Strained Tight-binding Workflow

To automatically extract tight-binding models for different strain directions and strengths, we define an additional workflow, `OptimizeStrainedFirstPrinciplesTightBinding`, as shown in fig. 8.5. The first step in this workflow, `ApplyStrainsWithSymmetry`, creates the strained structures from the initial structure and strain parameters. Since strain can break crystal symmetries, the symmetries of the unstrained system are tested against the strained structure. With the strained structures and the remaining symmetries, we then use the `OptimizeFirstPrinciples`

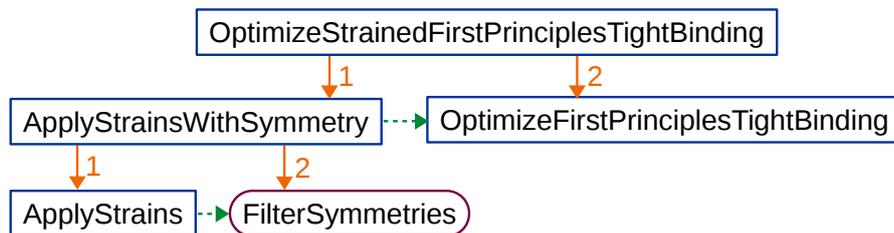


Figure 8.5: Sketch of the workflow for constructing strained tight-binding models. The color scheme is the same as in fig. 8.3.

TightBinding workflow to create a tight-binding model for each strain value.

### 8.3.2 First-principles Calculations

In the first step of generating the SWTB we need to carry out a first-principles calculation of the bulk semiconductor structure. We performed all first-principles calculations using the Vienna Ab-initio Simulation Package (VASP) utilizing projector augmented-wave (PAW) basis sets [30]. To obtain an accurate prediction of the band gap we employed hybrid functionals [106]. The HSE03/HSE06 hybrid functionals proved to be successful in computing band structures of III-V semiconductors [19]. These hybrid functionals are constructed by replacing a quarter of the density functional short-range exchange (which is the Perdew-Burke-Enzerhof functional in our case [16]) with its Hartree-Fock counterpart. The screening parameter  $\mu$  defines the separation into long- and short-range parts. In the popular HSE06 scheme, it is set to  $\mu = 0.2 \text{ \AA}^{-1}$ . We treated  $\mu$  as an empirical parameter such that the calculated band gap is fitted to the experimental value. In this work, we used  $\mu_{\text{InAs}} = 0.20 \text{ \AA}^{-1}$ ,  $\mu_{\text{GaSb}} = 0.15 \text{ \AA}^{-1}$  and  $\mu_{\text{InSb}} = 0.23 \text{ \AA}^{-1}$ , following the prescriptions of Ref. [150]. Since the SOC of III-V semiconductors is significant, we accounted for it by using scalar-relativistic PAW potentials.

InAs, GaSb and InSb crystallize in the zincblende structure with space group  $T_d^2$  (no. 216). For the unstrained structures we perform the first-principles calculation with the experimental lattice constant  $a$  at 300K, that is  $a_{\text{InAs}} = 6.058 \text{ \AA}$ ,  $a_{\text{GaSb}} = 6.096 \text{ \AA}$ ,  $a_{\text{InSb}} = 6.479 \text{ \AA}$ , from ref. [151]. A plane-wave energy cutoff of 380 eV was used for all calculations. The Brillouin-zone integrations were sampled by a  $6 \times 6 \times 6$   $\Gamma$ -centered  $k$ -points mesh.

To get optimal results from the Wannier90 code in conjunction with VASP we found that it is necessary to turn symmetries off in VASP, that is setting the ISYM-tag to 0. Since the states are obtained by a numerical diagonalization routine, they obtain a random phase at each  $\mathbf{k}$ -point. When symmetries are enabled however, the phases are the same for all vectors forming the star of  $\mathbf{k}$ . Since the convergence

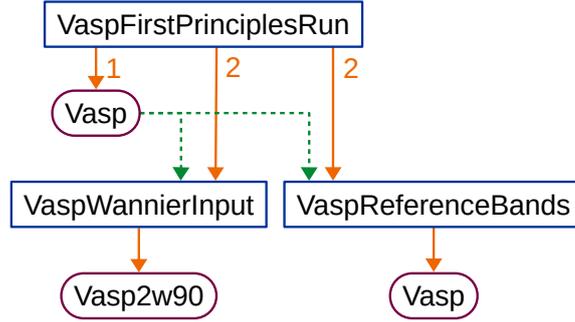


Figure 8.6: Sketch of the `FirstPrinciplesRunBase` subclass used for calculating the Wannier90 input and reference bands with VASP and hybrid functionals.

of Wannier90 is better if the numerical phases are random, turning symmetries off generally results in more localized Wannier functions after the projection step.

The interface for running first-principles calculations in the tight-binding extraction workflow is defined in the `FirstPrinciplesRunBase` class (see section 8.2.2). Here, we describe the specific sub-class used to implement these calculations with VASP [30], `VaspFirstPrinciplesRun` (see fig. 8.6). In a first step, this workflow performs a self-consistent calculation. The resulting wave-function is then passed to calculations for the reference band-structure and the input files for Wannier90. Two workflows `VaspReferenceBands` and `VaspWannierInput` are used to perform these calculations. The workflows are thin wrappers around the corresponding calculations from the `aiida-vasp` plugin [152], providing additional input and output validation. For the band-structure calculation, the workflow also adds the  $\mathbf{k}$ -point grid needed for hybrid functional calculations.

### 8.3.3 Strain Interpolation

Using the AiiDA workflow, we obtained tight-binding models for strains in the range of  $\pm 4\%$ , in steps of  $0.5\%$ . However, it is sometimes useful to have a finer control over the strain value without having to run additional first-principles calculations. A common way of obtaining this is by linear interpolation of the hopping parameters. Given two strain values  $s_1$  and  $s_2$ , for which the hopping parameter  $H^{s_i}[\mathbf{R}]$  are known, the hopping parameters for an unknown  $s^*$  can be calculated as

$$H^{s^*}[\mathbf{R}] = \alpha H^{s_1}[\mathbf{R}] + (1 - \alpha) H^{s_2}[\mathbf{R}], \quad (8.33)$$

where

$$\alpha = \frac{s^* - s_2}{s_1 - s_2}. \quad (8.34)$$

Since this method assumes that the hopping parameters are a linear function

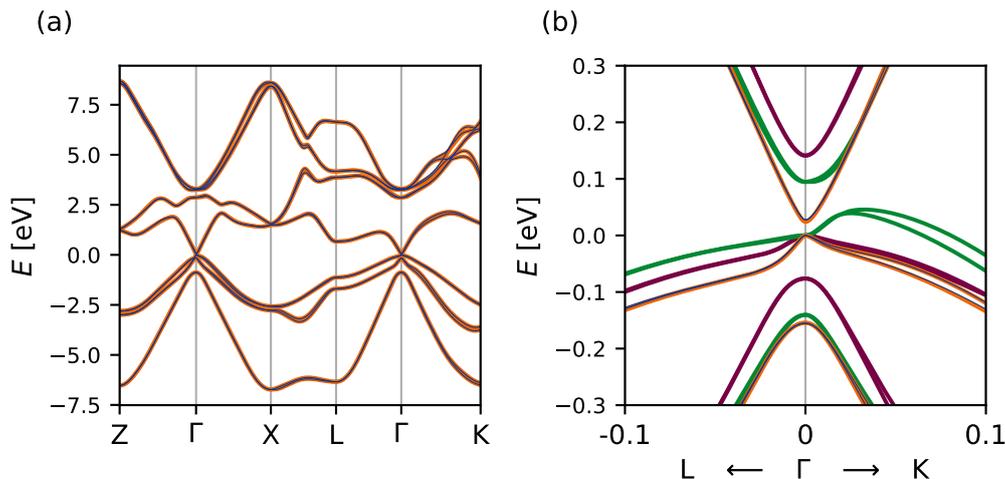


Figure 8.7: Comparison between the InSb band structure obtained directly from the tight-binding model with 2% biaxial (001) strain (blue), and from the linear interpolation (orange) between models with 1% and 3% strain. The energy scale is fixed by setting the top of the valence bands at  $\Gamma$  to zero. (a) At the electron-volt scale, the only visible difference is in the upper bands along the  $\Gamma$  - K line. (b) Close-up of the bands around  $\Gamma$ . The bands for 1% (purple) and 3% strain (green) are also shown.

of strain value, it becomes unreliable when  $s^*$  is too far away from  $s_1$  and  $s_2$ . For this reason, we compared a tight-binding model for InSb with 2% biaxial (001) strain obtained from linear interpolation of 1% and 3% strain models with one calculated directly from first-principles. Figure 8.7 shows a comparison of the two band-structures, which we find to be almost identical.

Important to note is that while linear interpolation works well for strains of the same kind, this is not necessarily the case when combining two models with different strain directions. The reason for this is that the symmetries of a particular structure depend on the direction of the applied strain, but (unless it is zero) not on its strength. As a result, a tight-binding model resulting from linear interpolation between two models of a different strain direction would not have the correct symmetries.

### 8.3.4 Results

To validate the tight-binding models obtained using the `aidda-tbextraction` workflows, several material parameters were calculated. Table 8.1 shows effective masses and g-factors for the unstrained models, in comparison to first-principles [150] and experimental [150, 153] values. Effective masses for the tight-binding models

Material	Method	$ m_{\text{SO}}^* $	$ m_{\text{LH}}^* $	$ m_{\text{HH}}^* $	$ m_e^* $	g-factor
<b>InSb</b>	HSE <sub>b<sub>fit</sub></sub>	0.129	0.018	0.245	0.017	
	SWTB	0.118	0.016	0.219	0.015	-49.8
	Expt.	0.110	0.015	0.263	0.014	-50.6
<b>InAs</b>	HSE <sub>b<sub>fit</sub></sub>	0.112	0.033	0.343	0.027	
	SWTB	0.118	0.036	0.340	0.029	-15.3
	Expt.	0.140	0.027	0.333	0.026	-15
<b>GaSb</b>	HSE <sub>b<sub>fit</sub></sub>	0.143	0.047	0.235	0.042	
	SWTB	0.124	0.039	0.20	0.036	-15.1
	Expt.	0.120	0.044	0.250	0.039	-7.8

Table 8.1: Effective masses of light hole (LH), heavy hole (HH), split-off hole and electron at  $\Gamma$  point along [100] direction in the unstrained case. Values for symmetrized Wannier-like tight-binding models (SWTB) are compared to first-principles (HSE<sub>b<sub>fit</sub></sub>) [150] and experimental results [150, 153].

Material		Energy Windows [eV]	$\Delta$
InSb	initial	(-4.5, [-4, 6.5], 16)	0.107
	optimized	(-4.44, [-3.24, 8.67], 14.01)	0.033
InAs	initial	(-4.5, [-4, 6.5], 16)	0.113
	optimized	(-4.44, [-3.59, 7.34], 15.04)	0.046
GaSb	initial	(-4.5, [-4.5, 7], 16)	0.082
	optimized	(-5.35, [-3.34, 7.90], 14.27)	0.043

Table 8.2: Initial and optimized energy windows used for calculating unstrained tight-binding models, and the corresponding band-structure mismatch as defined in eq. (8.32).

were calculated using second-order polynomial fit with range  $0.001 \text{ \AA}^{-1}$ . The g-factor calculations were performed using both perturbation theory and a Landau level calculation [154], with good agreement ( $< 0.5\%$  difference) between the two methods.

The effect of the energy window optimization is shown in table 8.2, which lists the initial and optimized windows, as well as the corresponding band-structure mismatch. As shown in fig. 8.2, it can be seen that the mismatch is substantially reduced after optimization.

Finally, the effect of strain on the energy levels at high-symmetry points is

shown in figs. 8.8 and 8.9. The numerical data is listed in the supplementary files of ref. [155].

In the supplementary materials of ref. [155], an export of the AiiDA database is given. This database contains the full provenance of each calculation performed to create the tight-binding models. For ease of accessibility, a separate data set containing only the 195 strained tight-binding models is also given.

## 8.4 Conclusion and Outlook

We have implemented a workflow for the *automatic* construction of Wannier tight-binding models from first-principles calculations. Building on the known procedure for calculating these models, we introduced a post-processing step to symmetrize the models, and an optimization of the energy windows used for disentanglement. These workflows are implemented in the `aiida-tbextraction` package, which is a free and open-source plugin for the AiiDA framework. As a test case, tight-binding models for strained III-V semiconductor materials were calculated. These results should enable device simulations for Majorana designs and other quantum devices.

The workflows have been implemented in a modular and extensible way. As a result, they can be used as building blocks for further improvements in automating the process of generating Wannier tight-binding models. Possible directions include extending the number of first-principles codes which are compatible with the plugin, adding different fitness criteria for the energy window optimization, and further minimizing the number of tunable parameters. For example, the need for choosing initial trial orbitals could be eliminated either by using another optimization step, or by utilizing the method of Ref. [125].

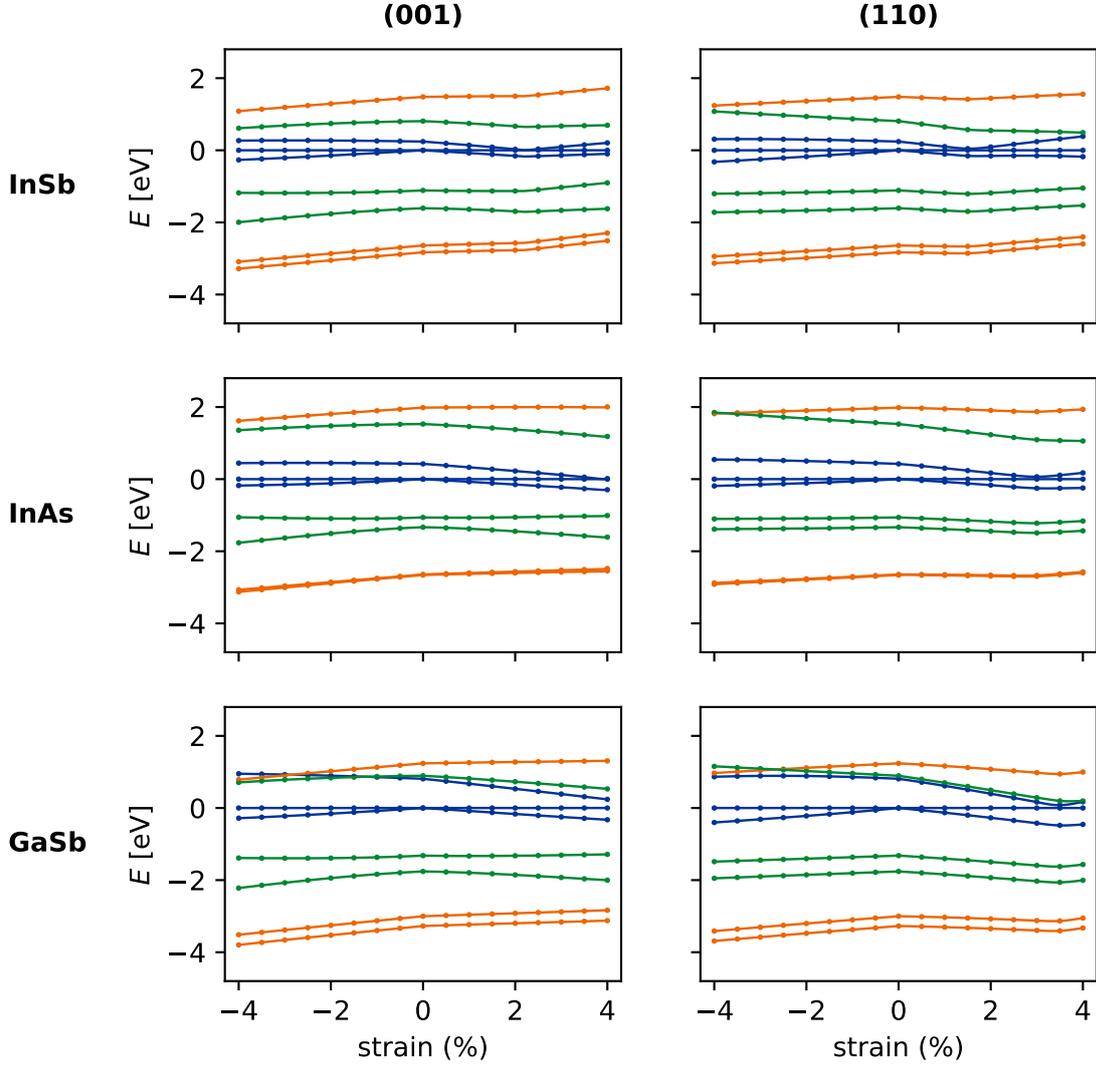


Figure 8.8: Strain dependence of band energies for biaxial (001) and (110) strains. The two highest valence bands and the lowest conduction band are shown at the  $\Gamma$  (blue),  $X$  (orange) and  $L$  (green) points, where each band is doubly degenerate. Energy values are shifted such that the valence band maximum at  $\Gamma$  is zero. The line represents values calculated from the tight-binding models with linear interpolation (eq. (8.33)) in steps of 0.1%. For comparison, the points show values calculated from first-principles. We find a good agreement between the tight-binding and first-principles values.

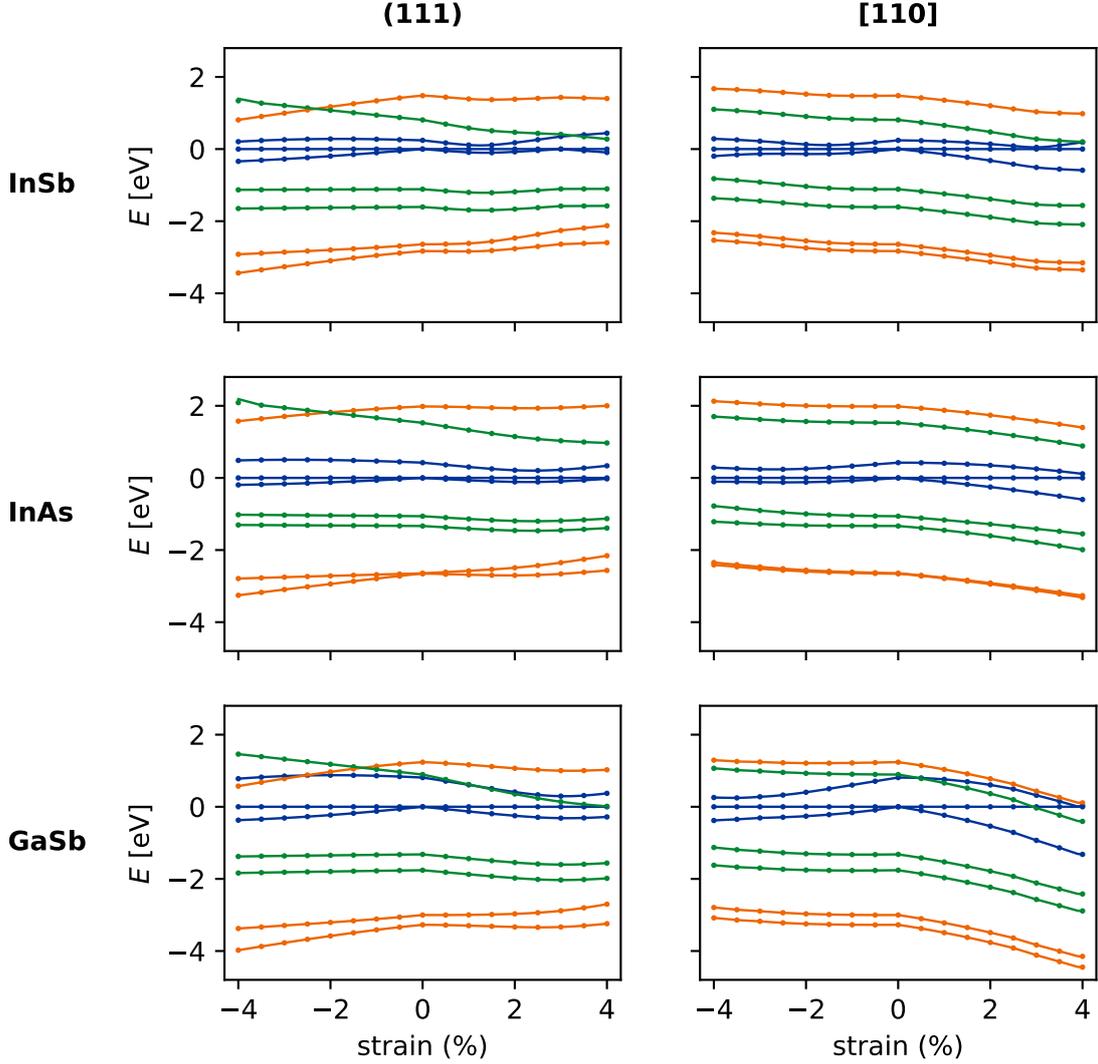


Figure 8.9: Strain dependence of band energies for biaxial (111) and uniaxial [110] strains. The two highest valence bands and the lowest conduction band are shown at the  $\Gamma$  (blue),  $X$  (orange) and  $L$  (green) points, where each band is doubly degenerate. Energy values are shifted such that the valence band maximum at  $\Gamma$  is zero. The line represents values calculated from the tight-binding models with linear interpolation (eq. (8.33)) in steps of 0.1%. For comparison, the points show values calculated from first-principles. We find a good agreement between the tight-binding and first-principles values, except for the conduction band value at the  $L$  - point at  $-4\%$  biaxial (111) strain.



# 9

## Automated Classification of Band Degeneracies

In this final chapter, we discuss the problem of finding and classifying all band degeneracies present in a given system. We split this problem into two steps: First, we use an optimization scheme to search for possible nodal features. The goal of this step is to produce a point cloud which lies densely in the nodal features of the material. In the second step, we aim to identify the nodal features from this point cloud. This procedure is an improvement and generalization of the method used in section 6.2.2.

Following this separation into two steps, the first two sections of this chapter describe the searching and identification steps, respectively. In the final section, we show some applications of the procedure. We first validate the method using known topological semimetals, and then apply it to classify novel topological semimetals.

The methods described here are implemented in the open-source `nodefinder` library.

### 9.1 Nodal Search Algorithm

In order to find the nodal features present in a given material, we need a function that evaluates the band gap, as a function of position  $\mathbf{k}$  in reciprocal space. This is typically done using a tight-binding or  $\mathbf{k} \cdot \mathbf{p}$  model. Calculating the gap directly from first principles is also possible, but significantly more computationally demanding.

The goal of this first step in classifying band degeneracies is to solve the root-finding problem for this band gap function. Since the band gap is a relatively

well-behaved function – it is continuous, and smooth outside of a set of measure zero – finding *a* nodal point is not a difficult task (if it exists), and one of many existing algorithms can be used for this purpose [132, 156]. However, in order to classify all nodal features, we need to find *all* roots of the given function. While techniques for finding all roots (or minima) such as simulated annealing [157] or minima hopping [158] exist, these perform well only if each minimum to be found is a single point. In contrast, the nodal features in materials can also have higher dimension, such as nodal lines or planes.

For such higher-dimensional nodal features, the goal is to create a point cloud of nodal points which lies densely in the nodal feature. Specifically, for a nodal feature  $F$ , the point cloud  $C$  should be such that no point in  $F$  is further than a given cutoff distance  $d_c$  away from the nearest nodal point:

$$\sup_{\mathbf{k} \in F} \left[ \min_{\mathbf{k}_i \in C} \|\mathbf{k}_i - \mathbf{k}\| \right] \leq d_c. \quad (9.1)$$

As a consequence, two points  $\mathbf{k}_i, \mathbf{k}_j \in C$  should be no further than twice the cutoff distance apart. We call this distance *feature size*

$$d_f = 2 d_c. \quad (9.2)$$

To achieve this, we first perform a Nelder-Mead [132] optimization from a mesh of starting positions, as illustrated in fig. 9.1 (panels a and b). The mesh should be sufficiently dense that at least one resulting point lies on each nodal feature. By default, we use a mesh of side-length 10. Since the search space needs to be covered roughly equally, the Nelder-Mead method is a good choice of optimization technique. Unlike other techniques such as steepest descent, it does not perform large jumps, and instead favors minima which are close to the starting point.

Next we perform a *refinement* step, where the surrounding area for each nodal point is explored. On a small box around a given nodal point, we again place starting positions on a mesh, as shown in fig. 9.1 (panels c and d).<sup>1</sup> This refinement step is repeated for each nodal point which is at least  $\frac{2}{3}d_c$  away from the nearest neighbor for which a refinement was already performed. A smaller value than the cutoff distance is used to improve the resilience of the procedure, so that eq. (9.1) will be fulfilled. We call this the *reduced* cutoff distance

$$\tilde{d}_c = \frac{2}{3}d_c. \quad (9.3)$$

When performing the refinement step, a possible issue is that the minimization will keep finding the same nodal points instead of exploring new ones. This is

---

<sup>1</sup>The default size of this box is  $\frac{10}{3}d_c = 5\tilde{d}_c$ , and we use a mesh of side-length 3.

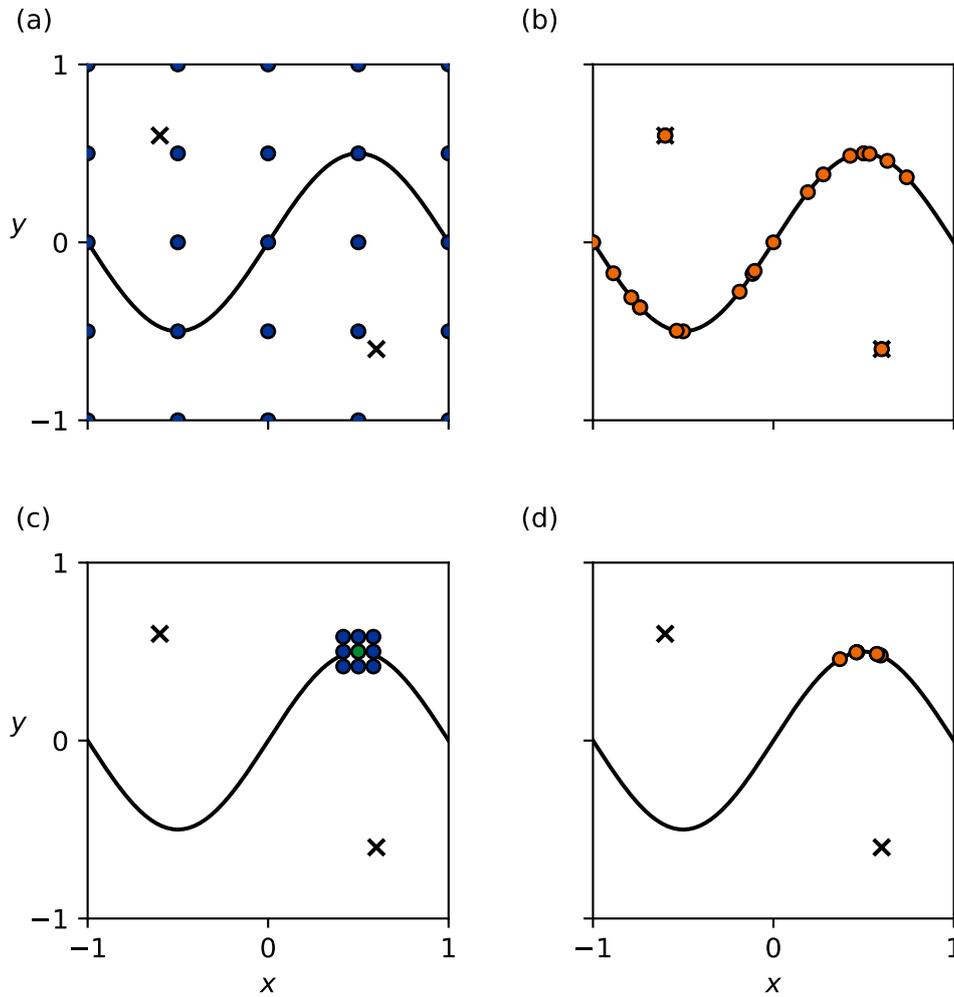


Figure 9.1: Illustration of the nodal search algorithm for a system with two nodal points (marked with crosses) and one nodal line. (a) Initial starting points (blue), on a  $5 \times 5$  mesh. (b) Nodal points (orange) found after the initial minimization step. (c) Refinement starting points (blue) around a given nodal point (green), on a  $3 \times 3$  mesh. (d) Nodal points (orange) found with the refinement starting points of panel c.

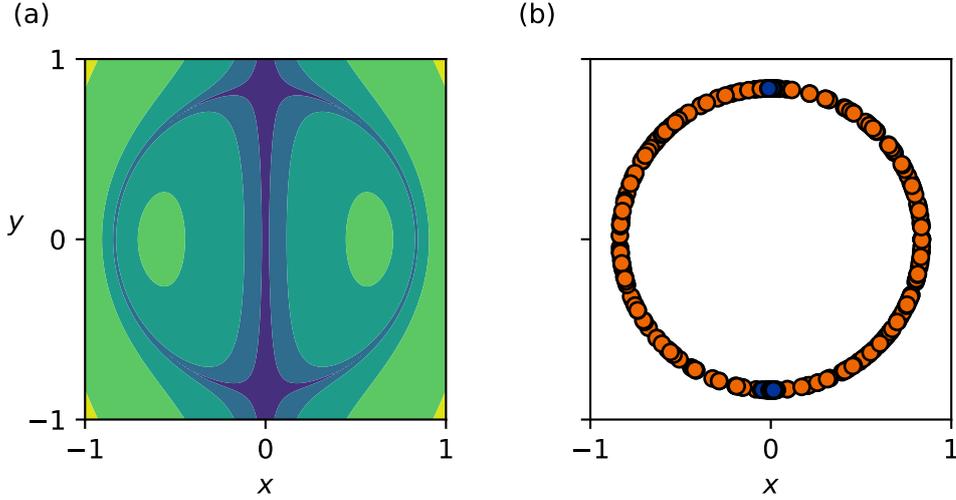


Figure 9.2: Example of a biased potential,  $f(x, y) = |x^2 + y^2 - 0.7| \cdot (10x^2 + 0.01)$ . (a) Contour plot of the potential, with logarithmic height levels. (b) Nodal search results with (orange) and without (blue) artificial repulsive potential, using a  $2 \times 2$  initial mesh,  $3 \times 3$  refinement mesh, and  $d_f = 0.2$ .

especially true when the potential is biased such that certain parts of the nodal feature are more attractive, as shown in fig. 9.2 (panel a). To avoid this problem, we add an artificial repulsive potential

$$V_{\text{art.}}(\mathbf{k}) = \begin{cases} \infty & \text{if } \min_{\mathbf{k}_i \in C} \|\mathbf{k} - \mathbf{k}_i\| \leq \tilde{d}_c \\ 0 & \text{else} \end{cases}, \quad (9.4)$$

where  $C$  is the set of nodes which were already found. The effect of this potential is that nodal points which were already found are avoided, and exploration of new areas is improved. The main advantage of using an *infinite* repulsive potential is that it is scale-invariant, meaning that its behavior does not depend on the scale of the gap function. It is made possible because the Nelder-Mead algorithm only compares function values, and does not compute derivatives.

Crucially, after the minimization using the artificial potential has converged, we perform a second minimization starting from the previous result *without* artificial potential. If the result of the first minimization is a true minimum, it will not move any further. However, if it is a local minimum introduced only due to the artificial potential, it will then converge back to the real minimum.

By combining the Nelder-Mead algorithm from multiple starting points with a refinement step and an artificial repulsive potential, we find that we can reliably search for nodal features in real materials, as will be shown in section 9.3.

## 9.2 Identification of Nodal Features

Having generated a point cloud describing the nodal features, we need to identify the individual features. This is done in three steps: First, the point cloud is separated into clusters, each corresponding to a distinct feature. Next, the dimensionality of each component is determined. Finally, the shape and topological properties of the nodal feature can be determined.

In doing this analysis, we assume that the nodal points lie densely in the features to be identified, as defined in eq. (9.1).

### 9.2.1 Clustering

The first step in the identification of nodal features is to separate the full point cloud into parts which belong to the same feature. For this purpose, we consider two nodal points to be “neighbors” if they are less than the feature size  $d_f$  apart. Points which are related by a (possibly indirect) neighbor relationship are then grouped together. Since this is an equivalence relation, it defines a unique separation of the points into groups. In order for this procedure to correctly assign points to nodal features, two conditions must be met: First, the nodal points must lie densely in the nodal feature, as defined in eq. (9.1). If this condition is not met, one nodal feature could be split into multiple groups. Second, the distance between two nodal features must be at least  $d_f$ . If this is not the case, two features will incorrectly be classified as the same.

To compute this clustering of nodal points, we first generate a list of neighbors for each nodal point. This can be done efficiently because the nodal points are stored in a cell list. We also keep track of the distance to each neighbor, which will be used in the subsequent steps. Using this graph representation of the nodal points the connected components can easily be determined, for example using breadth-first search. We use the `networkx` library [159] for this purpose.

### 9.2.2 Dimensionality Analysis

Given a cloud of nodal points, we would like to determine the dimension of the feature that they describe. This is a problem known as finding the *intrinsic dimension* of a dataset, and is commonly studied in the context of dimensionality reduction. As such, a number of techniques have been developed for estimating this dimension [160–167]. Ref. [160] for example estimates the dimension by statistically evaluating either distances or angles between a given point and its neighbors. Here, we use a very similar technique, but using instead the  $n$  - dimensional volume spanned by neighbor vectors as a measure. In the following, we describe this process in some detail.

First, we consider the idealized case of an  $n$ -dimensional *flat*  $F$  (a point, line, plane, etc.) in  $l$ -dimensional space. If we pick two random points  $p_0$  and  $p_1$  on  $F$ , the Euclidean distance  $\|p_1 - p_0\|$  is nonzero in the generic case, unless  $F$  is a point. When picking three points  $p_0, p_1$  and  $p_2$ , the area spanned by their connecting vectors  $\mathbf{k}_1 := p_1 - p_0$  and  $\mathbf{k}_2 := p_2 - p_0$  is nonzero in the generic case, unless  $F$  is a line or a point. This can be generalized to arbitrary dimensions as follows: To determine if  $F$  has at least dimension  $m \leq l$ , we pick  $m + 1$  random points, and construct the vectors  $\mathbf{k}_1 := p_1 - p_0, \dots, \mathbf{k}_m := p_m - p_0$ . From the singular value decomposition of the matrix

$$M = \begin{pmatrix} | & & | \\ \mathbf{r}_1 & \cdots & \mathbf{k}_m \\ | & & | \end{pmatrix} = U\Sigma V^\dagger, \quad (9.5)$$

we can compute the  $m$ -dimensional volume

$$V^m = |\det \Sigma| \quad (9.6)$$

spanned by these vectors. Again, it is zero in the generic case only if the dimension of  $F$  is smaller than  $m$ . By finding the first  $m$  for which  $V^m$  is zero, we can thus determine the dimension of  $F$  as  $n = m - 1$ .

Of course, the nodal features of interest are in general not flat. Nevertheless, the dimension can be determined if they are differentiable outside of a subset of measure zero. In that case, we can make use of the fact that  $F$  is almost flat in a small region around a generic point. Instead of picking any random points on  $F$  to compute the  $m$ -dimensional volume, we thus pick points within a certain radius  $r$  around the initial point  $p_0$ . A convenient choice of  $r$  is the feature size  $r = d_f$ , since we have already obtained a list of neighbors with maximum distance  $d_f$  for each point. The average  $m$ -dimensional volume should then either be close to zero if the dimension of  $F$  is less than  $m$ , or close to the exact value for a flat  $F$  which can be calculated (choosing  $r = 1$  for simplicity) as

$$\bar{V}_{\text{exact}}^m = \frac{1}{(V_{S^m})^m} \int_{S^m} \cdots \int_{S^m} \left| \det \begin{pmatrix} | & & | \\ \mathbf{k}_1 & \cdots & \mathbf{k}_m \\ | & & | \end{pmatrix} \right| d\mathbf{k}_1 \cdots d\mathbf{k}_m, \quad (9.7)$$

where  $S^m$  is the  $m$ -dimensional unit sphere, and  $V_{S^m}$  is its volume. By choosing to calculate the volume directly in  $m$ -dimensional space instead of a higher dimension we can avoid the singular-value decomposition and directly use the determinant. For  $m = 1, 2, 3$ , this can easily be evaluated algebraically:

**Case  $m = 1$**

The unit sphere in one dimension is simply the interval  $[-1, 1]$ , leading to

$$\bar{V}_{\text{exact}}^1 = \frac{1}{2} \int_{-1}^1 |k| dk = \frac{1}{2}. \quad (9.8)$$

**Case  $m = 2$**

In two dimensions, we can simplify the calculation by noticing that the problem is rotationally invariant. By an appropriate change of coordinates, we can thus keep the direction of  $\mathbf{k}_1$  fixed and only rotate  $\mathbf{k}_2$ . In polar coordinates, we get

$$\begin{aligned} \bar{V}_{\text{exact}}^2 &= \frac{1}{\pi^2} \int_0^1 dr_1 \int_0^1 dr_2 \int_0^{2\pi} d\varphi_1 \int_0^{2\pi} d\varphi_2 r_1 r_2 \left| \det \begin{pmatrix} 0 & r_2 \cos \varphi_2 \\ r_1 & r_2 \sin \varphi_2 \end{pmatrix} \right| \\ &= \frac{2\pi}{\pi^2} \left( \int_0^1 r^2 dr \right)^2 \int_0^{2\pi} |\sin \varphi_2| d\varphi_2 = \frac{8}{9\pi} \approx 0.283. \end{aligned} \quad (9.9)$$

**Case  $m = 3$**

The same rotational symmetry trick can be applied in three dimensions, to obtain

$$\begin{aligned} \bar{V}_{\text{exact}}^3 &= \left( \frac{3}{4\pi} \right)^3 4\pi \left( \int_0^1 r^3 dr \right)^3 \int_0^{2\pi} d\varphi_2 \int_0^{2\pi} d\varphi_3 \int_0^{\pi} d\vartheta_2 \int_0^{\pi} d\vartheta_3 \times \\ &\quad \times \sin \vartheta_2 \sin \vartheta_3 \left| \det \begin{pmatrix} 0 & \cos \varphi_2 \sin \vartheta_2 & \cos \varphi_3 \sin \vartheta_3 \\ 0 & \sin \varphi_2 \sin \vartheta_2 & \sin \varphi_3 \sin \vartheta_3 \\ 1 & \cos \vartheta_2 & \cos \vartheta_3 \end{pmatrix} \right| \\ &= \frac{27}{4^5 \pi^2} \left( \int_0^{\pi} (\sin \vartheta)^2 d\vartheta \right)^2 \int_0^{2\pi} d\varphi_2 \int_0^{2\pi} d\varphi_3 |\sin(\varphi_2 - \varphi_3)| \\ &= \frac{27}{4^5 \pi^2} \left( \frac{\pi}{2} \right)^2 8\pi = \frac{27\pi}{512} \approx 0.166. \end{aligned} \quad (9.10)$$

The actual values computed however can differ significantly from this theoretical value, due mostly to three factors:

1. The feature  $F$  is not exactly flat within the sphere of radius  $r$ , as mentioned previously. Figure 9.3 shows the distribution of distance (one-dimensional volume) and area (two-dimensional volume) for a straight and a curved line. While the area is exactly zero for the straight line, the curvature adds a non-zero contribution.
2. Since there is only a finite number of neighboring points within the given radius, the value for the  $m$ -dimensional volume is only approximate, as illustrated in fig. 9.4.
3. The positions of the nodal points are not exact. As illustrated in fig. 9.4 (panels g and h), this again adds a non-zero contribution even if  $m > n$ .

Despite these sources of error, the estimate for the  $m$ -dimensional volume is still roughly zero when the dimension  $n$  of the feature is less than  $m$ , and close to the exact value when  $n \geq m$ . We consider a point  $p_0$  to have a “local dimension” of at least  $m$  if the  $m$ -dimensional volume  $\bar{V}_{\text{num.}}^m$  computed numerically is at least half the exact value, i.e.

$$\bar{V}_{\text{num.}}^m \geq \bar{V}_{\text{exact}}^m/2. \quad (9.11)$$

In this way, we can assign a local dimension to each of the nodal points. The dimension of the nodal feature is then chosen to be the most common local dimension, under the condition that at least two thirds of all points have that dimension. In this way, a stable result can be found even in the presence of significant curvature and error in the nodal point positions.

Unfortunately, the number of volume calculations that are needed to calculate the local dimension in this way scales linearly with the number of points, and exponentially with the number of neighbors for each point. To circumvent this, we can choose to calculate the local dimension only for a subset of points. After a minimum number of points are calculated, we keep adding points only until the quorum of two thirds having the same local dimension is reached. Similarly, we can calculate the local dimension for a given point using only a subset of all neighbor combinations. As a stopping criterion, we require that the estimated error for  $\bar{V}_{\text{num.}}^m$  is less than half its distance to the cutoff value  $\bar{V}_{\text{exact}}^m/2$ :

$$\sqrt{\frac{\text{Var}[V_{\text{num.}}^m]}{N-1}} < \frac{1}{2} |\bar{V}_{\text{num.}}^m - \bar{V}_{\text{exact}}^m/2|, \quad (9.12)$$

where  $N$  is the number of neighbor combinations that have been evaluated.

### 9.2.3 Classification

The final step in identifying nodal features is to find a parametrization for its shape, and where applicable determine the topological properties of the node. Since this

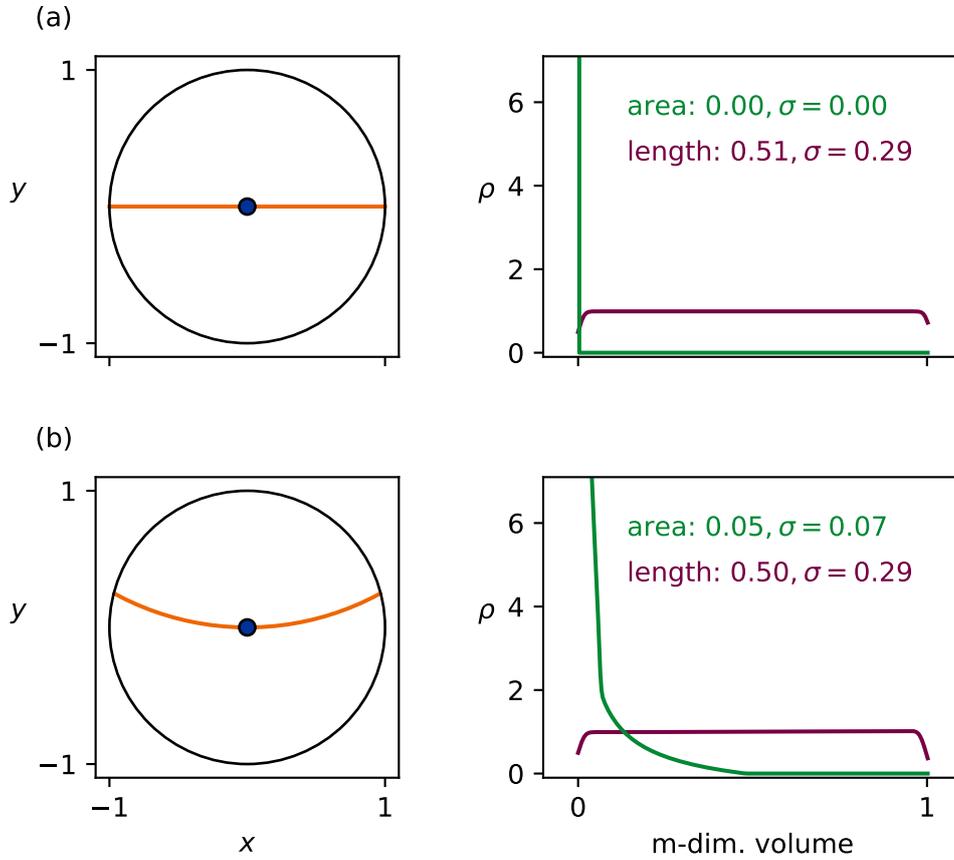


Figure 9.3: Exact distribution of length and area with neighbors (orange) on a line through the origin (blue). The distributions  $\rho$  (right panels) are sampled using a kernel density estimator with bandwidth factor 0.05. (a) Distribution for a straight line. (b) Distribution for a curved line. The curvature adds a non-zero contribution to the area distribution.

is done differently for each node dimension, we limit the discussion to point and line nodes, which are the most common nodal features in topological metals.

### Weyl Node Identification

For nodal points, the classification step is quite straightforward. First, the average position is computed from the given nodal points. In doing so, one needs to take into account that the coordinate system used can be periodic. If that is the case,

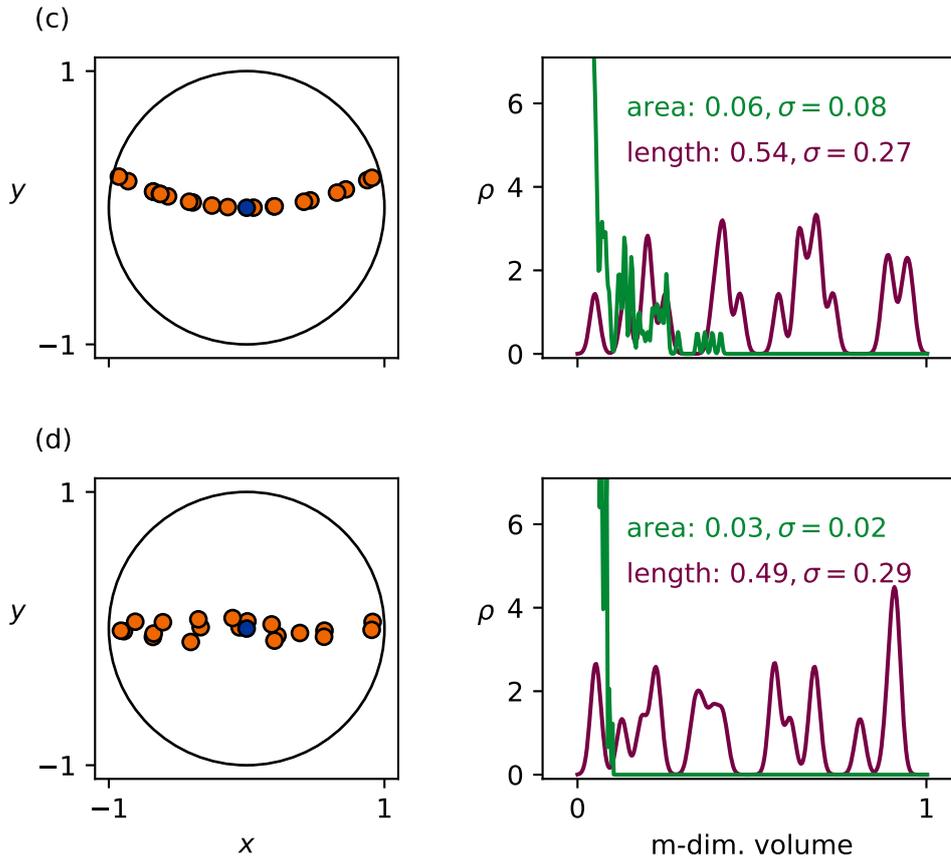


Figure 9.4: Distribution of length and area with neighbors (orange) sampled from a line through the origin (blue). The distributions  $\rho$  (right panels) are sampled using a kernel density estimator with bandwidth factor 0.05. (c) Curved line, evaluated with 20 random neighboring points. The random sampling causes fluctuations compared to the exact distribution (fig. 9.3 (b)). (d) Straight line, with Gaussian noise of width  $\sigma = 0.05$  in the neighbor positions. The error in the nodal point position adds a nonzero contribution to the area compared to the exact result (fig. 9.3 (a)).

the average position cannot simply be calculated as an average

$$\bar{\mathbf{k}} = \frac{1}{N} \sum_{i=1}^N \mathbf{k}_i \quad (9.13)$$

of position vectors, since the different  $\mathbf{r}_i$  might be on opposite sides of the periodic boundary. For example, the average between vectors  $\mathbf{k}_1 = (0.99)$  and  $\mathbf{k}_2 = (0.01)$  would incorrectly yield  $\bar{\mathbf{k}} = (0.5)$  instead of  $\bar{\mathbf{k}} = (0.)$ . For this reason, we need to first calculate the shortest connecting vector  $\boldsymbol{\tau}_i$  between a given position  $k_i$  and the (arbitrarily chosen) first position  $k_1$

$$\mathbf{k}_i = \mathbf{k}_1 + \boldsymbol{\tau}_i + \mathbf{T}, \quad (9.14)$$

where  $\mathbf{T}$  is a reciprocal lattice vector chosen such that the length of  $\boldsymbol{\tau}_i$  is minimized. The average position can then be computed as

$$\bar{\mathbf{k}} = \mathbf{k}_1 + \frac{1}{N} \sum_{i=1}^N \boldsymbol{\tau}_i. \quad (9.15)$$

Having calculated the average position, the topological properties of the nodal point can be evaluated by calculating the Chern number on a sphere surrounding the point, as described in chapter 3. For higher-order topological nodes, such as for example Dirac points, the topological invariant to be calculated depends on the symmetry that protects them. As a consequence, this is currently not done automatically.

### Line Identification

In the case of nodal lines, the goal is to find a minimal representation that accurately represents the nodal line. For simple open or closed paths, we would like to obtain a parametrization of the line. To obtain such a representation, we aim to start at a given nodal point, and then iteratively *follow* the nodal line by going in the direction where its neighbors are most dense. In each step, we add the current node and the edge connecting it to the previous node to a result graph  $g_{\text{res}}$ . In the following, we describe in detail how these nodes are chosen.

Given a starting point  $\mathbf{p}_0$  with neighbors  $\{\mathbf{n}_i\}_i$ , we first define the unit vector in the direction of the  $i$ -th neighbor

$$\mathbf{e}_i := \frac{\mathbf{n}_i - \mathbf{p}_0}{\|\mathbf{n}_i - \mathbf{p}_0\|}. \quad (9.16)$$

To determine how many *other* neighbors lie in the same direction, we assign a weight

$$w_i = \sum_{j \neq i} \max(\mathbf{e}_i \cdot \mathbf{e}_j, 0) \quad (9.17)$$

to each neighbor. In this weight, negative values are ignored because neighbors typically come in two groups of roughly opposite direction, as shown in fig. 9.5 (panel b). A possible choice for the next point would be to simply pick the neighbor

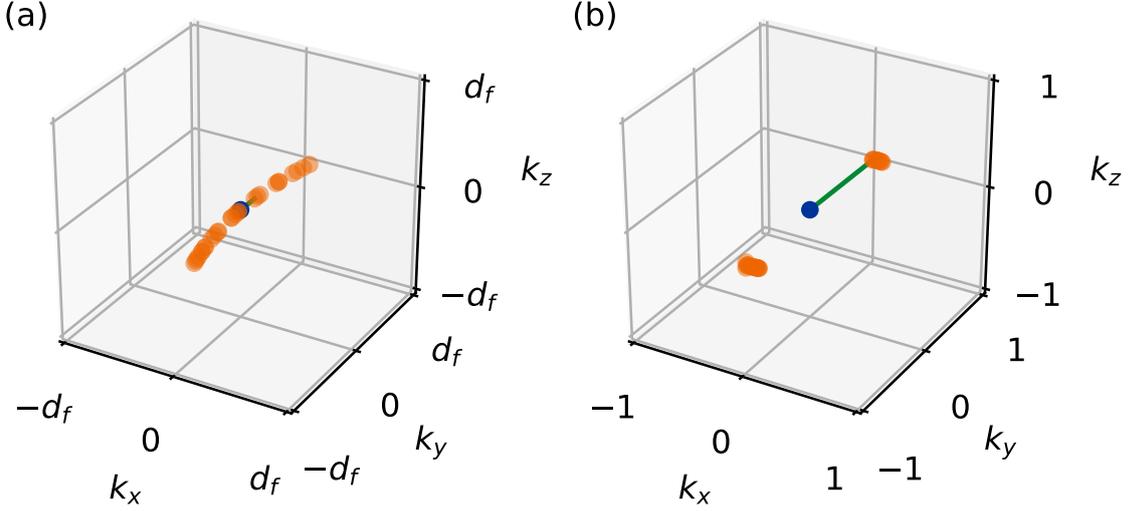


Figure 9.5: Illustration of the process for choosing the next point along the nodal line. The starting point  $\mathbf{p}_0$  is shown in blue, and all possible neighbors in orange. The point which was chosen is indicated with a green line. In panel (b), the neighbor positions are projected onto the unit sphere.

with the maximum weight and continue from there. However, it is also desirable that the next point is not too far from the existing one, to create a smoother representation of the nodal line. To account for this we multiply the weight by an empirical factor

$$(1 + 5 [d_i/d_f])^{-1}, \quad (9.18)$$

where  $d_i = \|\mathbf{n}_i - \mathbf{p}_0\|$  is the distance to each neighbor. When choosing subsequent steps, we also need to take into account that the representation of the nodal line should continue in the same direction along the line. This can be done by multiplying the weight by a factor

$$\mathbf{e}_i \cdot (\mathbf{p}_k - \mathbf{p}_{k-1}), \quad (9.19)$$

where  $\mathbf{p}_k$  and  $\mathbf{p}_{k-1}$  are the current and previous node respectively, and  $\mathbf{n}_i$  and  $\mathbf{e}_i$  are calculated with respect to  $\mathbf{p}_k$ . In this way, only neighbors which are in the positive half-plane with respect to the current line direction are considered. An additional issue coming from this factor is that  $\mathbf{p}_k - \mathbf{p}_{k-1}$  is numerically unstable if  $\mathbf{p}_k$  and  $\mathbf{p}_{k-1}$  are too close. This can be avoided by another factor in the weight

$$\alpha_i = \begin{cases} 1 & \text{if } d_i/d_f > 10^{-3} \\ 0 & \text{else} \end{cases}, \quad (9.20)$$

which prohibits selecting neighbors which are less than a thousandth of the feature size away. Furthermore, we want to avoid passing through the line twice. That is, if a node already exists in the result graph  $g_{\text{res}}$ , it is prioritized using the factor

$$\beta_i = \begin{cases} 2 & \text{if } n_i \in g_{\text{res}} \\ 1 & \text{else} \end{cases}. \quad (9.21)$$

Nodes which represent a crossing of two lines, having a degree of more than two in the result graph, are also prioritized with the factor

$$\delta_i = \begin{cases} 2 & \text{if } \text{deg}_{\text{res}}(n_i) > 2 \\ 1 & \text{else} \end{cases}. \quad (9.22)$$

Finally, the total weight by which the next point  $\mathbf{p}_{k+1}$  is selected is given by

$$\tilde{w}_i = \alpha_i \beta_i \delta_i w_i \frac{\mathbf{e}_i \cdot (\mathbf{p}_k - \mathbf{p}_{k-1})}{1 + 5 [d_i/d_f]}. \quad (9.23)$$

Using this heuristic, we trace the shape of the line until there are no more neighboring points with positive weights. This means that the end of the line is reached, since there are no further points in the half-plane given by the direction  $\mathbf{p}_k - \mathbf{p}_{k-1}$ . In the case of a closed loop, we stop if at least two consecutive nodal points were already part of the result graph  $g_{\text{res}}$ . Since a nodal line feature can be more complex and is not necessarily completely traced in the first run, this procedure is repeated until all nodes in the initial graph have at least one neighbor in the result graph. The new starting point is chosen to be (in order of precedence):

1. A node on the result graph with degree one. In this case, the direction is forced to be opposite the outgoing edge which already connects to the node. To see why this is necessary, consider an open line: when tracing the line from  $\mathbf{p}_0$ , one of two possible directions is chosen arbitrarily. To complete the line, the other direction must also be considered.<sup>2</sup>
2. Any node which does *not* have a neighbor on the result graph  $g_{\text{res}}$ .

Once this simplified graph  $g_{\text{res}}$  is constructed, the two cases of an open line and a closed loop can be identified by computing the degree of all vertices in the graph. Since there can be an arbitrary number of vertices of degree two, we ignore them for simplicity. If there are no other vertices, the shape of the nodal line must be a closed loop. If there are only two vertices of degree one, it must be an open line.

---

<sup>2</sup>Nodes at which a trace *ended*, and those for which this procedure was already done, are excluded.

In these cases, it is simple to create a parametrization of the line: For an open line, we can start at either one of the vertices with degree one and traverse the graph to create an ordered list of positions. For a closed loop, we can start at any vertex and traverse the graph in one of the two possible directions to again create an ordered list of positions. For more complex shapes, the degree of the vertices is still a useful index, but it no longer uniquely defines the shape of the line.

## 9.3 Materials Applications

Having described the two-step process of classifying topological semimetals, we now demonstrate its use on real materials. To reduce the computational cost required to find band degeneracies, we first create tight-binding models for each of the materials described here, using the method shown in chapter 8. First-principles calculations were performed with the VASP package, using projector augmented-wave (PAW) basis sets [30].

### 9.3.1 Validation on Known Topological Semimetals

As a first step, we apply the method of classifying band degeneracies to two known topological semimetals, molybdenum ditelluride ( $\text{MoTe}_2$ ) [5] and iridium tetrafluoride ( $\text{IrF}_4$ ) [84]. This serves as a test and validation of the method, by checking the results obtained here against previous calculations.

#### Type-II Weyl Semimetal $\text{MoTe}_2$

Molybdenum ditelluride ( $\text{MoTe}_2$ ) is a promising material candidate to host the type-II Weyl phase, as indicated both by calculations [74][5] and spectroscopic measurements [75–78]. Here, we study its orthorhombic  $\gamma$  phase (space group no. 31), as reported in ref. [5]. In this phase,  $\text{MoTe}_2$  hosts only four type-II Weyl points [5], related by symmetry.

To construct a tight-binding model for  $\text{MoTe}_2$ , we use  $d$  orbitals centered on the molybdenum atoms, and  $p$  orbitals centered on the tellurium atoms as a basis. First-principles calculations were calculated using a  $8 \times 6 \times 4$   $\mathbf{k}$ -point mesh, and a cutoff energy of 300 eV. For the subsequent search of band degeneracies, a feature size of  $d_f = 0.01$  and a  $5 \times 5 \times 5$  mesh of initial starting points was used. As expected, we find four type-II Weyl nodes in the  $k_z = 0$  plane, as listed in

---

<sup>3</sup>Positions are given in reduced reciprocal coordinates.

<sup>4</sup>Note that the chirality is inverted compared to ref. [5], due to using a different convention in the definition of chirality.

<sup>5</sup>Minimum eigenvalue of the matrix  $C$ , as defined in eq. (4.6).

No.	Position <sup>3</sup>	$E - E_F$ [eV]	Chirality <sup>4</sup>	Type	$\lambda_{\min}(C)$ <sup>5</sup>
1	(0.1018, 0.0498, 0.0000)	0.069	-1	II	-5.002
2	(0.1018, 0.9502, 0.0000)	0.069	1	II	-5.002

Table 9.1: List of Weyl nodes in MoTe<sub>2</sub>, showing only one Weyl node in each pair of nodes related by time-reversal symmetry. The type is calculated from the minimum eigenvalue  $\lambda_{\min}$  of the matrix  $C$ , as defined in eq. (4.6). Negative minimum eigenvalues correspond to type-II Weyl nodes, whereas positive minimum eigenvalues correspond to type-I Weyl nodes.

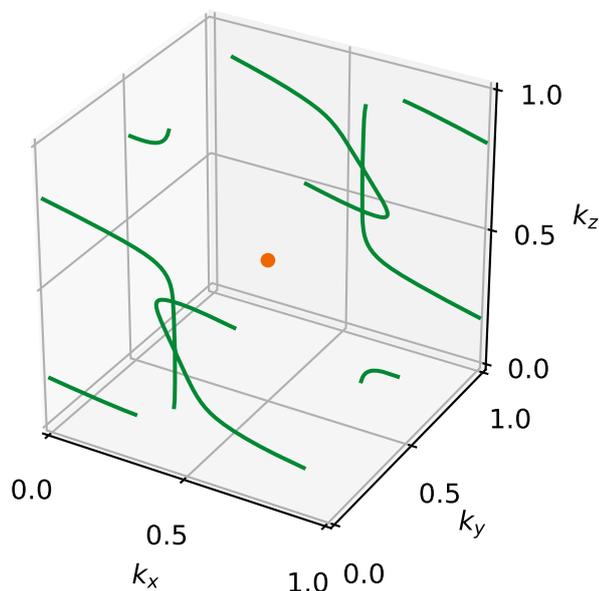


Figure 9.6: Nodal chain (green) and Dirac node (orange) of IrF<sub>4</sub>, in reduced coordinates.

table 9.1. The exact positions of the Weyl nodes are very close to the values of ref. [5], which places the Weyl node no. 1 at  $\mathbf{k} = (0.1011, 0.0503, 0)$ .

#### Nodal-Chain Semimetal IrF<sub>4</sub>

As an example of a nodal line semimetal, we study iridium tetrafluoride (IrF<sub>4</sub>), which belongs to the non-symmorphic space group  $Fdd2$  (no. 43) [168]. Ref. [84] predicts this material to host a nodal chain, and a single Dirac point.

We use a  $6 \times 6 \times 6$   $\mathbf{k}$ -point mesh for first-principles calculations of  $\text{IrF}_4$ , with an energy cutoff of 300 eV. To construct a tight-binding model, we use a basis of  $s$  and  $d$  orbitals centered on the iridium atoms, and  $p$  orbitals centered on the fluorine atoms. For the search of nodal features, we use a feature size of  $d_f = 0.01$  and a  $5 \times 5 \times 5$  initial mesh. The result of this calculation is shown in fig. 9.6, clearly showing both the nodal chain and the Dirac node.

### 9.3.2 Application to Novel Topological Semimetals

Having tested the method for classifying band degeneracies on known topological semimetals, we now study materials of unknown topological nature. To select appropriate candidate materials, we used the materials project database [169], which utilizes the `pymatgen` [170] and `fireworks` [124] libraries. For the reference band structure, reciprocal space paths were selected with the SeeK-path library [171], utilizing the `spglib` library [172] and ASE parser [173]. The Bilbao crystallographic server [174–176] was used to transform some input files.

The materials presented here were chosen without taking into account their growth, or whether topological features might be hidden by other bulk bands. As such, we do not expect that they are good material candidates for any specific application. Nevertheless, we demonstrate that this method can reliably identify topological features of unknown materials. This opens up the possibility of a more targeted materials search.

#### Weyl Semimetal $\text{CaCuO}_2$

In a high-pressure phase, calcium copper dioxide ( $\text{CaCuO}_2$ ) crystallizes in an orthorhombic structure [177] with space group  $P2_12_12$  (no. 18). Its unit cell is given by (in Ångström)

$$\begin{aligned} \mathbf{a}_1 &= (5.4603, & 0, & 0) \\ \mathbf{a}_2 &= (0, & 11.1351 & 0) \\ \mathbf{a}_3 &= (0, & 0, & 3.2029). \end{aligned}$$

In reduced coordinates, there are four equivalent symmetry-related positions

$$\begin{aligned} & (x, & y, & z) \\ & (-x, & -y, & z) \\ & (x + 1/2, & -y + 1/2, & -z) \\ & (-x + 1/2, & y + 1/2, & -z), \end{aligned}$$

and the atoms are positioned – omitting the symmetry-related positions – at

<b>Ca</b>	(0.4756,	0.3627,	0.5046)
<b>Cu</b>	(0.9850,	0.3781,	0.0071)
<b>O 1</b>	(0.2733,	0.7457,	0.9990)
<b>O 2</b>	(0.2565,	0.4939,	0.0155).

For computing a tight-binding model of  $\text{CaCuO}_2$ , we performed first-principles calculations using a  $5 \times 2 \times 10$   $\mathbf{k}$ -point mesh, with a cutoff energy of 520 eV. We chose  $s$  and  $d$  orbitals centered on the calcium atoms,  $s$ ,  $p$ , and  $d$  orbitals centered on the copper atoms, and  $s$  and  $p$  orbitals centered on the oxygen atoms as a basis for the tight-binding model. The band structure of this model is shown in fig. 9.7, compared to the first-principles reference. Note that there is a significant difference between the two band structures along the  $Y - \Gamma$  line, for the higher conduction bands. The difference is smaller however in the bands which form the nodal features discussed here.

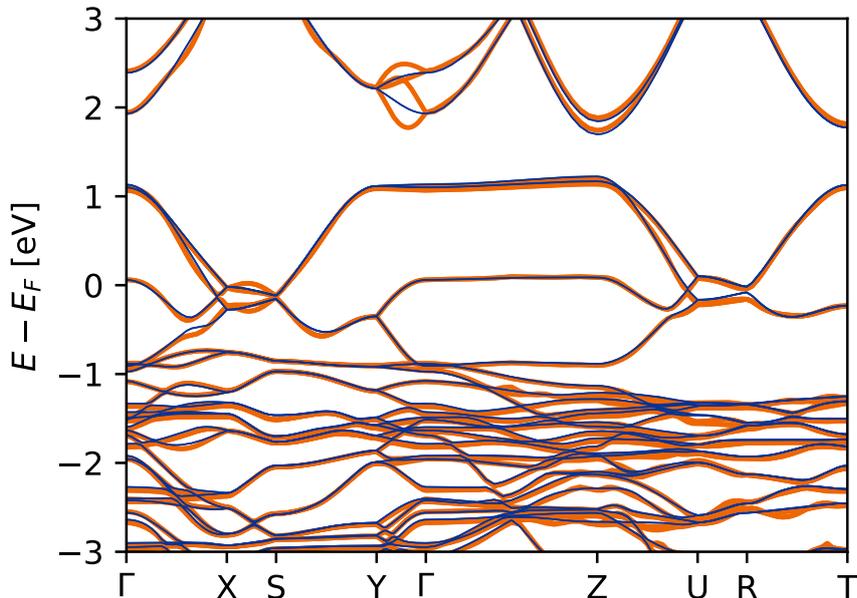


Figure 9.7: Band structure of  $\text{CaCuO}_2$  computed from first-principles (blue), compared to the tight-binding values (orange). The tight-binding energy values are shifted to have the same valence band maximum at  $\Gamma$  as the first-principles results.

Using the tight-binding model, we run the node classification algorithm with a feature size of  $d_f = 0.002$  and an initial mesh of size  $15 \times 15 \times 15$ . We find that  $\text{CaCuO}_2$  hosts 64 Weyl nodes, as listed in table 9.2 and shown in table 9.2. All Weyl nodes are located in the vicinity of the  $k_x = 0.5$  plane, between  $k_x = 0.448$  and  $k_x = 0.552$ . We find that Weyl nodes of both type-I and type-II are present in

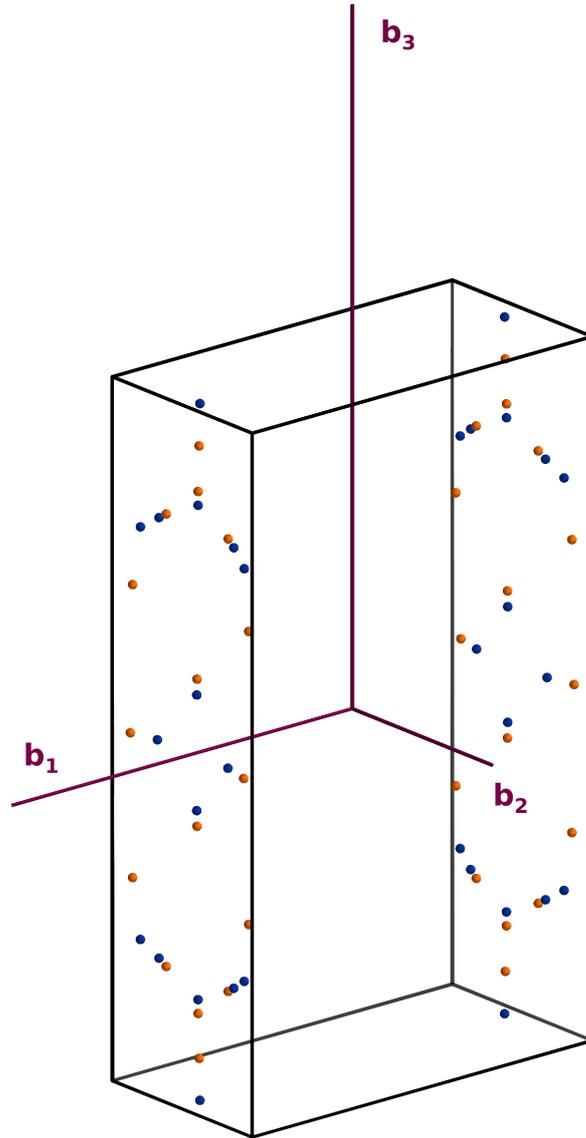


Figure 9.8: Weyl nodes in the Brillouin zone of  $\text{CaCuO}_2$ . Positive chirality nodes are colored orange, and negative chirality ones are colored blue.

the system. For many Weyl nodes the minimum eigenvalues of the matrix  $C$  (as defined in eq. (4.6)) is close to zero, meaning that they are close to a type transition. In these cases, the type of the Weyl node should be considered undecided, due to the possible dependence on numerical details.

No.	Position <sup>3</sup>	$E - E_F$ [eV]	Chirality	Type	$\lambda_{\min}(C)^5$
1	(0.4480, 1.0000, 0.4949)	-0.123	-1	I	0.010
2	(0.4480, 0.0000, 0.5051)	-0.123	-1	I	0.010
3	(0.4503, 0.0000, 0.4350)	-0.122	1	I	0.012
4	(0.4503, 1.0000, 0.5650)	-0.122	1	I	0.012
5	(0.4534, 1.0000, 0.3707)	-0.125	1	II	-0.046
6	(0.4534, 1.0000, 0.6293)	-0.125	1	II	-0.046
7	(0.4535, 1.0000, 0.6490)	-0.132	-1	II	-0.239
8	(0.4535, 0.0000, 0.3510)	-0.132	-1	II	-0.239
9	(0.4559, 0.2215, 0.3214)	-0.154	1	II	-0.142
10	(0.4559, 0.7785, 0.6786)	-0.154	1	II	-0.142
11	(0.4559, 0.7785, 0.3214)	-0.154	1	II	-0.142
12	(0.4559, 0.2215, 0.6786)	-0.154	1	II	-0.142
13	(0.4565, 1.0000, 0.1045)	-0.172	1	I	0.057
14	(0.4565, 0.0000, 0.8955)	-0.172	1	I	0.057
15	(0.4572, 0.0000, 0.0821)	-0.174	-1	I	0.004
16	(0.4572, 0.0000, 0.9179)	-0.174	-1	I	0.004
17	(0.4583, 0.2673, 0.6871)	-0.163	-1	II	-0.301
18	(0.4583, 0.2673, 0.3129)	-0.163	-1	II	-0.301
19	(0.4583, 0.7327, 0.3129)	-0.163	-1	II	-0.301
20	(0.4583, 0.7327, 0.6871)	-0.163	-1	II	-0.301
21	(0.4694, 0.2520, 0.0000)	-0.180	-1	I	0.000
22	(0.4694, 0.7480, 0.0000)	-0.180	-1	I	0.000
23	(0.4703, 0.3711, 0.2930)	-0.181	-1	II	-0.016
24	(0.4703, 0.6289, 0.2930)	-0.181	-1	II	-0.016
25	(0.4703, 0.6289, 0.7070)	-0.181	-1	II	-0.016
26	(0.4703, 0.3711, 0.7070)	-0.181	-1	II	-0.016
27	(0.4753, 0.4150, 0.7919)	-0.185	1	I	0.006
28	(0.4753, 0.5850, 0.2081)	-0.185	1	I	0.006
29	(0.4753, 0.4150, 0.2081)	-0.185	1	I	0.006
30	(0.4753, 0.5850, 0.7919)	-0.185	1	I	0.006
31	(0.4860, 0.5952, 0.0000)	-0.180	1	I	0.000
32	(0.4860, 0.4048, 1.0000)	-0.180	1	I	0.000

Table 9.2: List of Weyl nodes in  $\text{CaCuO}_2$ , showing only one Weyl node in each pair of nodes related by time-reversal symmetry.

### Type-II Weyl Semimetal $\text{Sc}_6\text{N}_2\text{O}_5$

Because Weyl nodes can not be adiabatically gapped out, they can be stable even in the absence of any crystal symmetries. Here, we show a material which exemplifies this:  $\text{Sc}_6\text{N}_2\text{O}_5$  is predicted computationally [169] to crystallize in the trivial space group  $P1$  (no. 1), having only time-reversal symmetry. This predicted structure has a primitive cell with unit vectors (in Ångström)

$$\begin{aligned}\mathbf{a}_1 &= (0.9209, -5.3916, 0) \\ \mathbf{a}_2 &= (-5.4656, 0, 0) \\ \mathbf{a}_3 &= (0.9354, 1.1047, -5.2781).\end{aligned}$$

Since the space group is  $P1$ , there are no symmetry-related equivalent positions in the unit cell. The atomic positions in reduced coordinates are as follows:

$$\begin{aligned}\text{Sc 1} & (-0.5740, -0.1542, -0.2762) \\ \text{Sc 2} & (-0.8525, -0.7181, -0.4316) \\ \text{Sc 3} & (-0.7260, -0.4338, -0.8599) \\ \text{Sc 4} & (-0.2706, -0.5631, -0.1488) \\ \text{Sc 5} & (-0.1419, -0.2785, -0.5563) \\ \text{Sc 6} & (-0.4357, -0.8523, -0.7234) \\ \text{N 1} & (-0.9222, -0.3751, -0.2236) \\ \text{N 2} & (-0.4994, -0.5016, -0.5009) \\ \text{O 1} & (-0.7722, -0.0761, -0.6241) \\ \text{O 2} & (-0.6221, -0.7764, -0.0758) \\ \text{O 3} & (-0.3770, -0.2223, -0.9205) \\ \text{O 4} & (-0.2260, -0.9220, -0.3805) \\ \text{O 5} & (-0.0804, -0.6265, -0.7784)\end{aligned}$$

We perform first-principles calculations using a  $4 \times 4 \times 4$   $\mathbf{k}$ -point mesh, with an energy cutoff of 520 eV. For constructing the tight-binding model, a basis of  $s$  and  $d$  orbitals centered on the scandium atoms,  $s$  and  $p$  orbitals centered on the oxygen atoms, and  $s$  and  $p$  orbitals centered on the nitrogen atoms is used. Figure 9.9 shows the band structure of both the tight-binding model and the first-principles calculated.

By running the node finding algorithm with  $d_f = 0.004$  and an initial mesh of size  $8 \times 8 \times 8$ , we find 24 Weyl nodes, as listed in table 9.3 and shown in fig. 9.10. Most of these Weyl nodes are of type-II, except four which are very close to the type transition. These four Weyl nodes – number 5 and 6, and their time-reversal images – form two pairs which are close to each other and both have positive chirality. While most other Weyl points are close to a node of the opposite chirality and thus prone to being annihilated, another two pairs of Weyl nodes with *negative* chirality – number 11 and 12, and their time-reversal images – also exist.

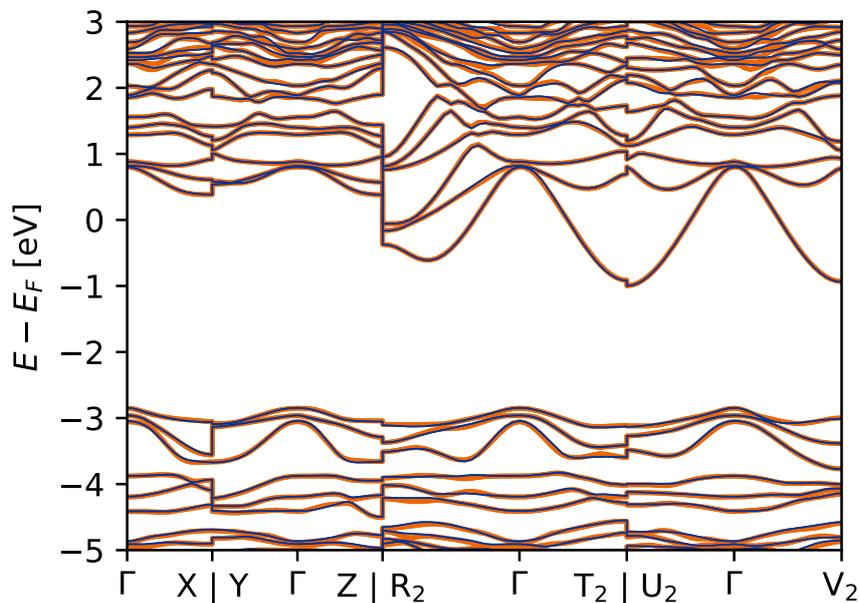


Figure 9.9: Band structure of  $\text{Sc}_6\text{N}_2\text{O}_5$  computed from first-principles (blue), compared to the tight-binding values (orange). The tight-binding energy values are shifted to have the same valence band maximum at  $\Gamma$  as the first-principles results.

Interestingly, the negative chirality and positive chirality pairs are quite far apart in energy. Unfortunately, all Weyl nodes found here more than 0.2 eV away from the Fermi energy<sup>6</sup>. As such, they very probably have negligible influence on the transport properties of the material.

<sup>6</sup>It bears mentioning that the Fermi energy estimate obtained from the first-principles calculation might well be inaccurate.

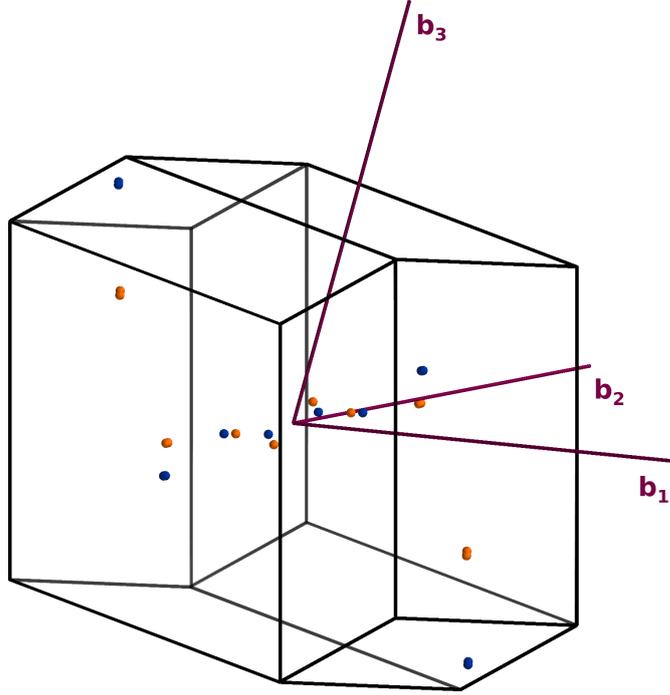


Figure 9.10: Weyl nodes in the Brillouin zone of  $\text{Sc}_6\text{N}_2\text{O}_5$ . Positive chirality nodes are colored orange, and negative chirality ones are colored blue.

No.	Position <sup>3</sup>	$E - E_F$ [eV]	Chirality	Type	$\lambda_{\min}(\mathbf{C})^5$
1	(0.0001, 0.9218, 0.9843)	0.772	-1	II	-0.660
2	(0.0060, 0.0401, 0.0458)	0.795	1	II	-0.697
3	(0.0086, 0.1828, 0.0008)	0.686	1	II	-1.148
4	(0.0088, 0.2242, 0.9953)	0.647	-1	II	-1.070
5	(0.0118, 0.5559, 0.0152)	0.534	1	II	-0.005
6	(0.0149, 0.5449, 0.0157)	0.534	1	I	0.002
7	(0.0678, 0.3085, 0.0884)	0.685	-1	II	-2.150
8	(0.0728, 0.3078, 0.0894)	0.691	-1	II	-2.011
9	(0.2090, 0.4483, 0.6556)	-0.269	1	II	-2.030
10	(0.2144, 0.4439, 0.6459)	-0.286	1	II	-1.690
11	(0.3393, 0.3775, 0.4156)	-0.257	-1	II	-2.541
12	(0.3421, 0.3759, 0.4090)	-0.245	-1	II	-2.842

Table 9.3: List of Weyl nodes in  $\text{Sc}_6\text{N}_2\text{O}_5$ , showing only one Weyl node in each pair of nodes related by time-reversal symmetry.

### Weyl and Nodal Line Semimetal RbPdBr<sub>3</sub>

Next, we discuss another structure which is only predicted computationally, rubidium palladium tribromide (RbPdBr<sub>3</sub>) in the monoclinic  $P1m1$  space group (no. 6) [178]. Its unit cell is given (in Ångström) by

$$\begin{aligned}\mathbf{a}_1 &= (5.2891, 0, 0) \\ \mathbf{a}_2 &= (0, 5.2962, 0) \\ \mathbf{a}_3 &= (-0.0040, 0, 5.2922),\end{aligned}$$

with atoms placed as follows (in reduced coordinates):

$$\begin{aligned}\mathbf{Rb} & (0.9937, 0.0000, 0.0010) \\ \mathbf{Pd} & (0.4849, 0.5000, 0.4959) \\ \mathbf{Br\ 1} & (0.4873, 0.0000, 0.4960) \\ \mathbf{Br\ 2} & (0.9853, 0.5000, 0.5010) \\ \mathbf{Br\ 3} & (0.4888, 0.5000, 0.9960)\end{aligned}$$

Notably, this structure is only slightly distorted from having a cubic lattice, with inversion symmetry. In that case, of course the existence of simple Weyl points would be prohibited by the product symmetry of inversion and time-reversal. If this structure is indeed stable, it would be interesting to see whether its structure is cubic or monoclinic.

To create a tight-binding model for RbPdBr<sub>3</sub>, we performed first-principles calculations with an  $8 \times 8 \times 8$   $\mathbf{k}$ -point mesh and an energy cutoff of 520 eV. As a basis for the tight-binding model, we choose  $s$  and  $p$  orbitals centered on the rubidium atom,  $s$  and  $d$  orbitals centered on the palladium atom, and  $s$  and  $p$  orbitals centered on the bromine atoms. Figure 9.11 shows the band structure of tight-binding model, in comparison to the first-principles calculation. As with CaCuO<sub>2</sub>, there is a significant difference in the higher conduction bands, which should not influence the nodal structure.

We run the node finding algorithm on this tight-binding model of RbPdBr<sub>3</sub> with a feature size of  $d_f = 0.002$ , and an initial mesh of size  $8 \times 8 \times 8$ . As listed in table 9.4 and shown in fig. 9.12, we find twelve type-II Weyl nodes and two small nodal loops. The nodal features are concentrated around the  $\mathbf{k} = (0.5, 0.5, 0.5)$  point, and all lie more than 1.5 eV above the Fermi level. The two nodal lines lie on the  $k_y = 0.5$  plane. Of the twelve Weyl nodes, eight form pairs of opposite chirality, while the remaining four are further away from the nearest neighboring Weyl point. An interesting feature are the two nodal lines, especially because they are very small with less than 0.01 distance (in reduced reciprocal coordinates) across. Nevertheless, they are topologically nontrivial, as shown by the Berry phase calculation presented in fig. 9.13. Due to their small size, the nodal lines could very easily be contracted and consequently. However, it should be noted

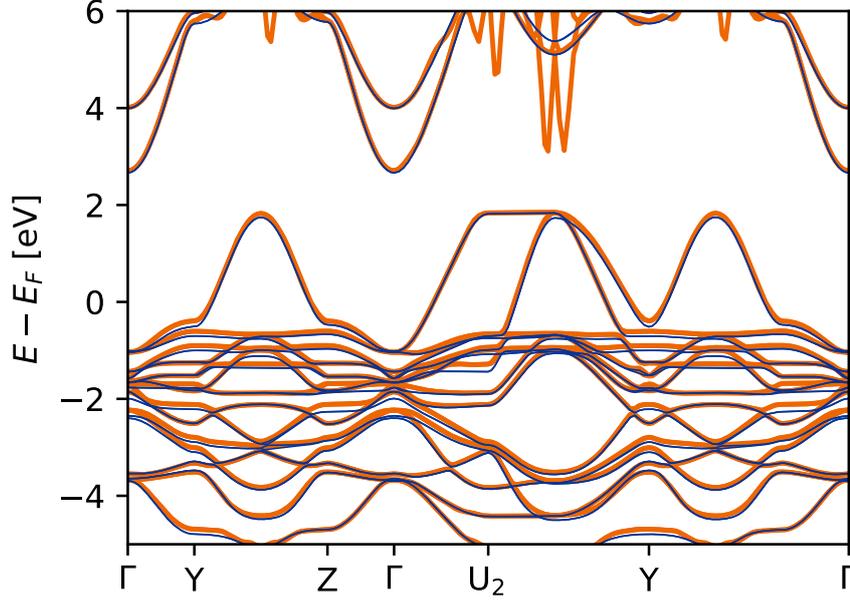


Figure 9.11: Band structure of RbPdBr<sub>3</sub> computed from first-principles (blue), compared to the tight-binding values (orange). The tight-binding energy values are shifted to have the same valence band maximum at  $\Gamma$  as the first-principles results.

No.	Position <sup>3</sup>	$E - E_F$ [eV]	Chirality	Type	$\lambda_{\min}(\mathbf{C})^5$
1	(0.4175, 0.4222, 0.5824)	1.545	-1	II	-12.115
2	(0.4175, 0.5778, 0.5824)	1.545	1	II	-12.115
3	(0.4410, 0.5553, 0.4408)	1.663	1	II	-7.283
4	(0.4410, 0.4447, 0.4408)	1.663	-1	II	-7.283
5	(0.4464, 0.5482, 0.4488)	1.694	-1	II	-6.565
6	(0.4464, 0.4518, 0.4488)	1.694	1	II	-6.565

Table 9.4: List of Weyl nodes in RbPdBr<sub>3</sub>, showing only one Weyl node in each pair of nodes related by time-reversal symmetry.

that it is possible for the nodal lines to be stabilized and protected by additional symmetries, for example with the mechanism described in ref. [179].

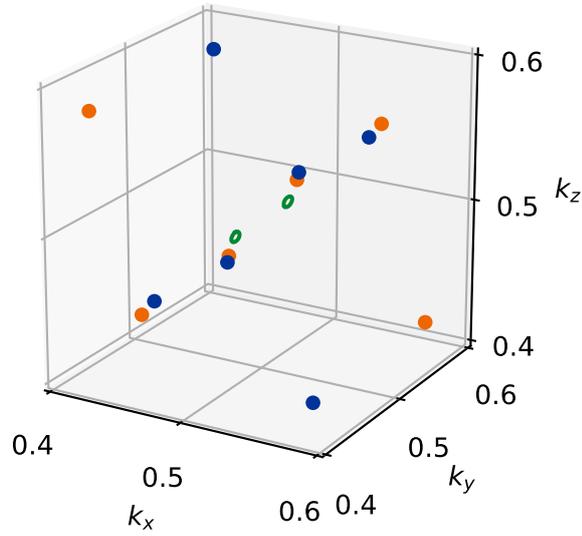


Figure 9.12: Weyl nodes and nodal loops in  $\text{RbPdBr}_3$ . Positive chirality Weyl nodes are colored orange, and negative chirality ones blue. The nodal loops are colored in green. No nodal features are present outside the box  $\mathbf{k} \in [0.4, 0.6]^3$  shown here.

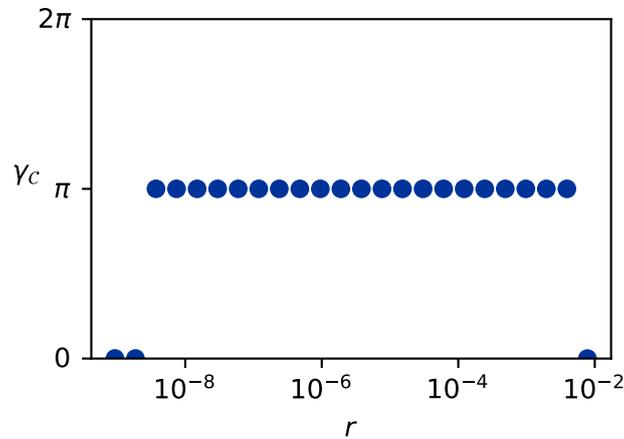


Figure 9.13: Berry phase on a circle around one of the two nodal loops found in  $\text{RbPdBr}_3$ , as a function of the circle radius. For large radii, the circle encloses the nodal loop instead of linking with it, so the Berry phase is zero. For very small radii ( $r < 2 \cdot 10^{-9}$ ), the circle is not guaranteed to link the nodal loop due to numerical inaccuracies. Since the Berry phase close to  $\pi$  for a large range of intermediate radii, we can conclude that the nodal loop is topologically nontrivial.

**Weyl and Nodal Line Semimetal AsPd<sub>2</sub>**

As a further material example we discuss AsPd<sub>2</sub>, in an orthorhombic structure with space group  $Cmc2_1$  (no. 36) [180]. The reduced unit cell is given by (in Ångström)

$$\begin{aligned}\mathbf{a}_1 &= (1.6656, -8.5965, 0.0000) \\ \mathbf{a}_2 &= (1.6656, 8.5965, 0.0000) \\ \mathbf{a}_3 &= (0.0000, 0.0000, 6.7049).\end{aligned}$$

In reduced coordinates, the atoms take positions  $(x, -x, z)$  and  $(-x, x, z + 1/2)$  for the following values of  $(x, z)$ :

$$\begin{aligned}\mathbf{As\ 1} & (0.2039, 0.7961, 0.2682) \\ \mathbf{As\ 2} & (0.4416, 0.5584, 0.3027) \\ \mathbf{Pd\ 1} & (0.3254, 0.6746, 0.0392) \\ \mathbf{Pd\ 2} & (0.4150, 0.5850, 0.6731) \\ \mathbf{Pd\ 3} & (0.1709, 0.8291, 0.8923) \\ \mathbf{Pd\ 4} & (0.0383, 0.9617, 0.5102)\end{aligned}$$

To create a tight-binding model for AsPd<sub>2</sub>, we run a first-principles calculation using a  $6 \times 6 \times 3$   $\mathbf{k}$ -point mesh and a cutoff energy of 520 eV. The basis orbitals we use are  $s$  and  $p$  orbitals centered on the arsenic atoms and  $d$  orbitals centered on the palladium atom. The resulting first-principles and tight-binding band structures are shown in fig. 9.14. While there is some visible difference between the first-principles and tight-binding energy values around the Fermi level, they appear to be the same qualitatively.

We run the node finding algorithm for a starting mesh of size  $15 \times 15 \times 15$ , and feature size  $d_f = 0.01$  to find 20 Weyl points and two  $C$ -shaped nodal line as shown in fig. 9.15 and listed in table 9.5. Due to the complex shape of the nodal lines with paths in close proximity and acute angles, the refinement as described in section 9.1 does not produce an accurate nodal point cloud describing this nodal line. This can be corrected by using an improved set of starting points for each refinement step. Instead of using a regular mesh, we use 30 points regularly arranged on a sphere of radius  $1.1 d_f$  around the point to be refined.

Of the 20 Weyl points, eight are of type I and twelve are of type II. They are all slightly below the Fermi level, between  $-11$  meV and  $-57$  meV. The nodal lines even cross the Fermi level, ranging from  $-0.264$  to  $0.055$  eV. They lie at the Brillouin zone boundary, on the  $k_x = k_y$  plane in reciprocal coordinates. Among the material candidates presented here AsPd<sub>2</sub> seems the most promising, both because it has a relatively clean Fermi surface, and because the nodal features lie close to the Fermi level.

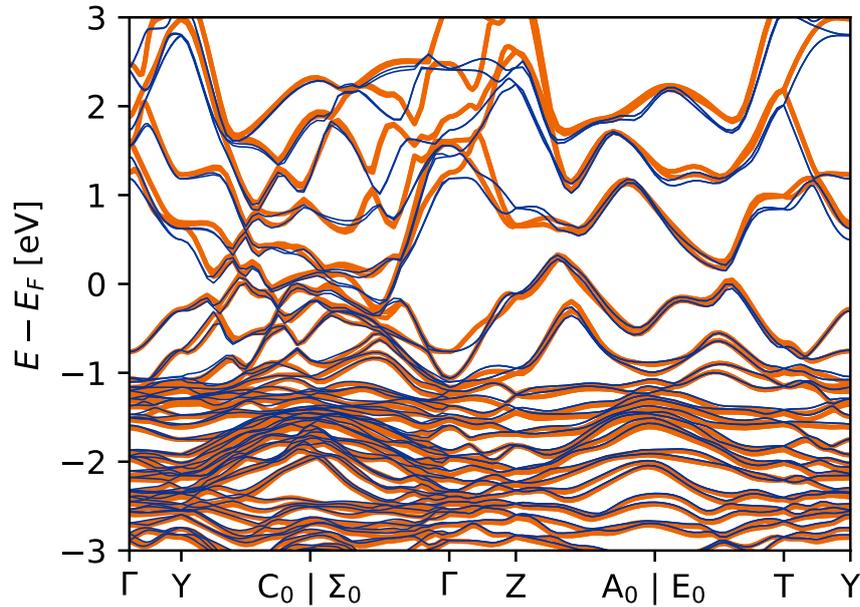


Figure 9.14: Band structure of  $\text{AsPd}_2$  computed from first-principles (blue), compared to the tight-binding values (orange). The tight-binding energy values are shifted to have the same valence band maximum at  $\Gamma$  as the first-principles results.

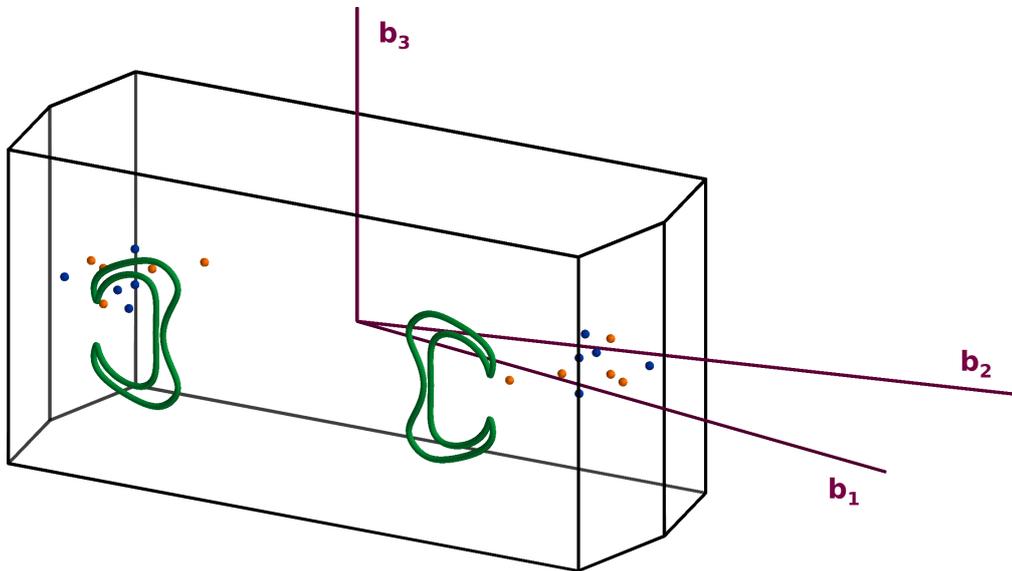


Figure 9.15: Weyl nodes and nodal loops in  $\text{AsPd}_2$ . Positive chirality Weyl nodes are colored orange, and negative chirality ones blue. The nodal loops are colored in green.

No.	Position <sup>3</sup>	$E - E_F$ [eV]	Chirality	Type	$\lambda_{\min}(\mathbf{C})^5$
1	(0.0521, 0.3230, 1.0000)	-0.011	-1	II	-0.765
2	(0.0769, 0.3247, 0.0568)	-0.039	1	II	-0.018
3	(0.0769, 0.3247, 0.9432)	-0.039	1	II	-0.018
4	(0.1312, 0.3403, 0.0000)	-0.057	-1	I	0.196
5	(0.1360, 0.5425, 0.0000)	-0.047	-1	I	0.048
6	(0.3230, 0.0521, 0.0000)	-0.011	1	II	-0.765
7	(0.3247, 0.0769, 0.9432)	-0.039	-1	II	-0.018
8	(0.3247, 0.0769, 0.0568)	-0.039	-1	II	-0.018
9	(0.3403, 0.1312, 0.0000)	-0.057	1	I	0.196
10	(0.4575, 0.8640, 0.0000)	-0.047	1	I	0.048

Table 9.5: List of Weyl nodes in AsPd<sub>2</sub>, showing only one Weyl node in each pair of nodes related by time-reversal symmetry.

### Weyl and Nodal Line Semimetal GaAu<sub>2</sub>

Finally, we discuss GaAu<sub>2</sub> in the structure reported in ref. [181] which has the same orthorhombic space group as AsPd<sub>2</sub>, *Cmc2<sub>1</sub>* (no. 36). The primitive unit cell is given by

$$\begin{aligned}\mathbf{a}_1 &= (1.6950, -9.0543, 0) \\ \mathbf{a}_2 &= (1.6950, 9.0543, 0) \\ \mathbf{a}_3 &= (0, 0, 6.9947).\end{aligned}$$

In reduced coordinates, the atoms take positions  $(x, -x, z)$  and  $(-x, x, z + 1/2)$  for the following values of  $(x, z)$ :

$$\begin{aligned}\mathbf{Ga\ 1} & (0.2908, 0.2276) \\ \mathbf{Ga\ 2} & (0.0594, 0.1935) \\ \mathbf{Au\ 1} & (0.0863, 0.8209) \\ \mathbf{Au\ 2} & (0.1696, 0.4507) \\ \mathbf{Au\ 3} & (0.3227, 0.5967) \\ \mathbf{Au\ 4} & (0.4647, 0.9985)\end{aligned}$$

To create a tight-binding model of GaAu<sub>2</sub>, we perform first-principles calculations using a  $6 \times 6 \times 3$  mesh of  $\mathbf{k}$ -points, and an energy cutoff of 520 eV. As a basis of the tight-binding model, we use *s*, *p* and *d* orbitals centered on the gallium atoms, and *s* and *d* orbitals centered on the gold atoms. The resulting tight-binding band structure is shown in fig. 9.16, in comparison to the first-principles reference. It should be noted here that the tight-binding model shows significant differences from the first-principles calculations, although the correct connectivity of the bands seems to be preserved.

We run the node finding algorithm with an initial mesh size of size  $20 \times 20 \times 20$  and feature size  $d_f = 0.005$ . As for AsPd<sub>2</sub>, we use a spherical refinement mesh with 30 starting points. We identify 68 type-II Weyl nodes and four nodal lines in GaAu<sub>2</sub>, as listed in table 9.6 and shown in fig. 9.17. Of the Weyl nodes, the closest to the Fermi level are eight nodes (nos. 1, 2, 22, 23 and their time-reversal images), which lie 85 meV above the Fermi energy. The four nodal lines are again located at the Brillouin zone boundary, on the  $k_x = k_y$  plane. Their energy ranges between 356 meV and 83 meV below the Fermi level.

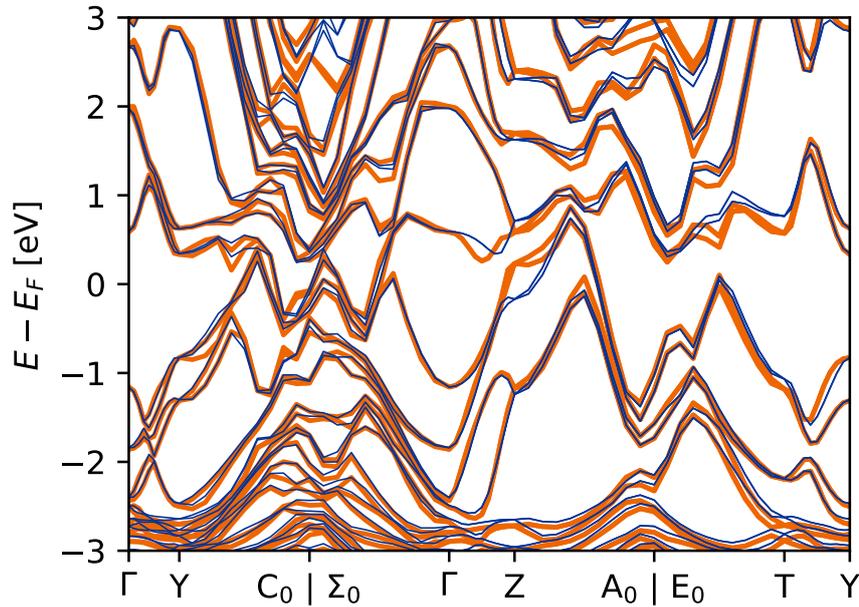


Figure 9.16: Band structure of  $\text{GaAu}_2$  computed from first-principles (blue), compared to the tight-binding values (orange). The tight-binding energy values are shifted to have the same valence band maximum at  $\Gamma$  as the first-principles results.

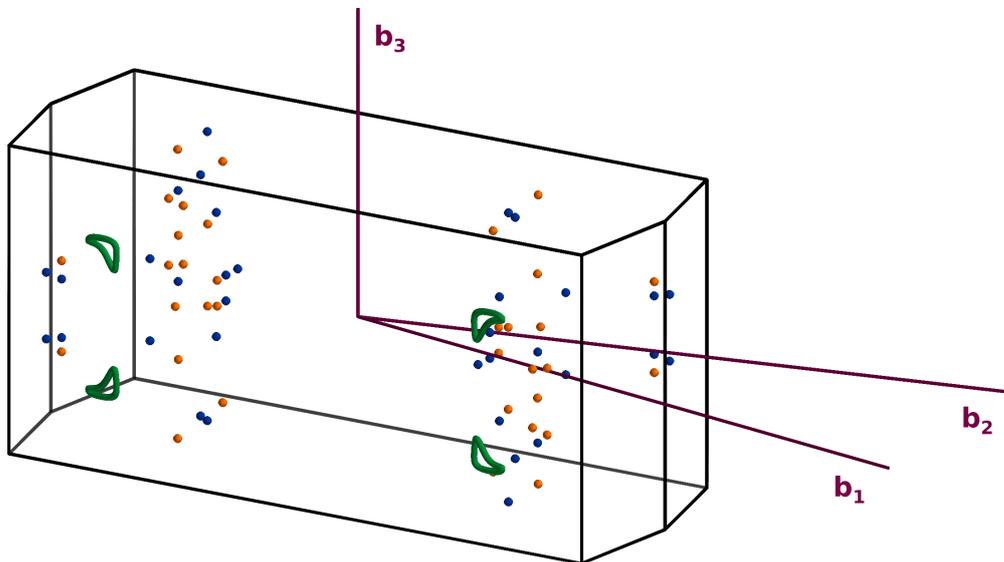


Figure 9.17: Weyl nodes and nodal loops in  $\text{GaAu}_2$ . Positive chirality Weyl nodes are colored orange, and negative chirality ones blue. The nodal loops are colored in green.

No.	Position <sup>3</sup>	$E - E_F$ [eV]	Chirality	Type	$\lambda_{\min}(C)^5$
1	(0.0166, 0.7129, 0.7986)	0.085	1	II	-6.717
2	(0.0166, 0.7129, 0.2014)	0.085	1	II	-6.717
3	(0.0210, 0.2578, 0.5311)	0.582	1	II	-49.164
4	(0.0210, 0.2578, 0.4689)	0.582	1	II	-49.164
5	(0.0349, 0.2115, 0.6087)	0.700	-1	II	-0.398
6	(0.0349, 0.2115, 0.3913)	0.700	-1	II	-0.398
7	(0.0790, 0.6191, 0.1331)	-0.133	-1	II	-2.479
8	(0.0790, 0.6191, 0.8669)	-0.133	-1	II	-2.479
9	(0.0798, 0.1500, 0.9582)	0.252	1	II	-4.381
10	(0.0798, 0.1500, 0.0418)	0.252	1	II	-4.381
11	(0.1220, 0.6229, 0.0000)	-0.387	1	II	-1.808
12	(0.1500, 0.0798, 0.0418)	0.252	-1	II	-4.381
13	(0.1500, 0.0798, 0.9582)	0.252	-1	II	-4.381
14	(0.2115, 0.0349, 0.3913)	0.700	1	II	-0.398
15	(0.2115, 0.0349, 0.6087)	0.700	1	II	-0.398
16	(0.2578, 0.0210, 0.5311)	0.582	-1	II	-49.164
17	(0.2578, 0.0210, 0.4689)	0.582	-1	II	-49.164
18	(0.2649, 0.3339, 0.1475)	-0.361	1	II	-5.826
19	(0.2649, 0.3339, 0.8525)	-0.361	1	II	-5.826
20	(0.2777, 0.3002, 0.8918)	-0.360	-1	II	-61.596
21	(0.2777, 0.3002, 0.1082)	-0.360	-1	II	-61.596
22	(0.2871, 0.9834, 0.2014)	0.085	-1	II	-6.717
23	(0.2871, 0.9834, 0.7986)	0.085	-1	II	-6.717
24	(0.2874, 0.3157, 0.9050)	-0.463	-1	II	-1.110
25	(0.2874, 0.3157, 0.0950)	-0.463	-1	II	-1.110
26	(0.3002, 0.2777, 0.1082)	-0.360	1	II	-61.596
27	(0.3002, 0.2777, 0.8918)	-0.360	1	II	-61.596
28	(0.3157, 0.2874, 0.0950)	-0.463	1	II	-1.110
29	(0.3157, 0.2874, 0.9050)	-0.463	1	II	-1.110
30	(0.3339, 0.2649, 0.8525)	-0.361	-1	II	-5.826
31	(0.3339, 0.2649, 0.1475)	-0.361	-1	II	-5.826
32	(0.3771, 0.8780, 0.0000)	-0.387	-1	II	-1.808
33	(0.3809, 0.9210, 0.8669)	-0.133	1	II	-2.479
34	(0.3809, 0.9210, 0.1331)	-0.133	1	II	-2.479

Table 9.6: List of Weyl nodes in GaAu<sub>2</sub>, showing only one Weyl node in each pair of nodes related by time-reversal symmetry.

## 9.4 Conclusion

We have created a two-step algorithm for classifying nodal features in materials. In the first step, a root-finding method is used to create a cloud of nodal points. In the second step, this point cloud is separated into individual features, and their dimension and shape is analyzed. By combining this procedure with the workflows for creating tight-binding models discussed in chapter 8, we were able to classify several topological semimetals, both known and unknown. While the novel topological semimetals presented here may not be particularly useful candidate materials themselves, this demonstrates a significant simplification in the process of identifying topological semimetals. By making the implementation of the methods described here publicly available, we hope to enable both the study of topological semimetal phases by non-specialists, and a targeted high-throughput search for specific material properties. A possible improvement of the node finding algorithm would be to adaptively increase the size of the initial mesh. This could be based on whether the current mesh still finds new nodal features.

# 10

## Conclusion and Outlook

The topics we have studied in this thesis fall broadly in one of two categories: First, the study of topological semimetal phases in real materials. And second, the creation of algorithms enabling and enhancing this study.

On the first subject, we start in chapter 4, where we review our discovery of the type-II Weyl state. This is continued in chapter 6, where we studied the effect of magnetic field on transition metal dipnictides. By approximating this effect as a Zeeman term, we find that magnetic field induces the presence of Weyl nodes. An open question in this research is whether these Weyl nodes persist when a more accurate model of the magnetic field effect is used. Also, it is not known whether these Weyl points are responsible for the reduced or negative magnetoresistance observed in these compounds. This could be investigated with transport calculations. And finally, in chapter 9 we study several different materials and classify their nodal structure topologically. We find an abundance of Weyl nodes, which indicates that this state might be quite common in both semimetals and metals when allowed by symmetry.

The creation of algorithms simplifying the study of topological semimetals – which arguably make out the more significant part of this thesis – starts in chapter 5, where we introduce a general algorithm for computing phase diagrams. The algorithm is based on a quadtree approach, and can be seen as a multidimensional generalization of binary search. We thoroughly analyze the convergence properties of the algorithm, defining a criterion for convergence based on the initial starting grid. Since this algorithm is not specific to topological phase diagrams, we expect it to be suitable for a wide variety of applications. An example application is shown in chapter 6, where we use it to calculate phase diagrams showing the number of Weyl points as a function of magnetic field strength. A tool simplifying

---

the study of the electronic structure problem in general is presented in chapter 7. We study the problem of generating  $\mathbf{k}\cdot\mathbf{p}$  models which conserve a given set of crystal symmetries. By expressing this problem in terms of vector spaces, we enable automatically solving it with a computer algebra system. The subsequent chapter 8 treats a similar problem, the construction of symmetric tight-binding models from first-principles calculations. We introduce an approach to symmetrizing existing tight-binding models based on performing a group average. Based on the Wannier90 code and the AiiDA framework, we implement automated workflows that perform this construction of tight-binding models. To eliminate the need for manually choosing energy windows used by Wannier90, we implement an optimization routine to automatically select the appropriate values by comparing the resulting band structure to a first-principles reference. And finally, in chapter 9 we implement a two-step algorithm for classifying nodal features in band structures. In the first step, we use a root-finding approach to create a cloud of nodal points. In the second step, the features present in these point clouds are extracted and analyzed. This is applied to the study of topological semimetals, calculating topological invariants to establish the nontrivial nature of these nodes.

By making all these tools publicly available and publishing comprehensive documentation, we aim to make the task of studying topological semimetal states approachable to non-specialists.

# List of Publications

- [1] **Dominik Gresch**. *Identifying Topological States in Matter*. Master Thesis, ETH Zurich, May 2015.
- [2] Alexey A. Soluyanov, **Dominik Gresch**, Zhijun Wang, QuanSheng Wu, Matthias Troyer, Xi Dai, and B. Andrei Bernevig. Type-II Weyl semimetals. *Nature*, 527(7579):495–498, November 2015.
- [3] Gabriel Autès, **Dominik Gresch**, Matthias Troyer, Alexey A. Soluyanov, and Oleg V. Yazyev. Robust Type-II Weyl Semimetal Phase in Transition Metal Diphosphides  $XP_2$  ( $X=Mo, W$ ). *Phys. Rev. Lett.*, 117(6):066402, August 2016.
- [4] Alexey A. Soluyanov, **Dominik Gresch**, Matthias Troyer, Roman M. Lutchyn, Bela Bauer, and Chetan Nayak. Optimizing spin-orbit splittings in InSb Majorana nanowires. *Phys. Rev. B*, 93(11):115317, March 2016.
- [5] Zhijun Wang, **Dominik Gresch**, Alexey A. Soluyanov, Weiwei Xie, S. Kushwaha, Xi Dai, Matthias Troyer, Robert J. Cava, and B. Andrei Bernevig.  $MoTe_2$ : A Type-II Weyl Topological Metal. *Phys. Rev. Lett.*, 117(5):056805, July 2016.
- [6] **Dominik Gresch**, Gabriel Autès, Oleg V. Yazyev, Matthias Troyer, David Vanderbilt, B. Andrei Bernevig, and Alexey A. Soluyanov. Z2Pack: Numerical implementation of hybrid Wannier centers for identifying topological materials. *Phys. Rev. B*, 95(7):075146, February 2017.
- [7] **Dominik Gresch**, QuanSheng Wu, Georg W. Winkler, and Alexey A. Soluyanov. Hidden Weyl points in centrosymmetric paramagnetic metals. *New J. Phys.*, 19(3):035001, 2017.
- [8] **Dominik Gresch** and Alexey Soluyanov. Calculating Topological Invariants with Z2Pack. In Dario Bercioux, Jérôme Cayssol, Maia G. Vergniory, and M. Reyes Calvo, editors, *Topological Matter: Lectures from the Topological Matter School 2017*, Springer Series in Solid-State Sciences, pages 63–92. Springer International Publishing, Cham, 2018.

- [9] **Dominik Gresch**, QuanSheng Wu, Georg W. Winkler, Rico Häuselmann, Matthias Troyer, and Alexey A. Soluyanov. Automated construction of symmetrized Wannier-like tight-binding models from ab initio calculations. *Phys. Rev. Materials*, 2(10):103805, October 2018.

# Bibliography

- [10] K. v. Klitzing, G. Dorda, and M. Pepper. New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance. *Phys. Rev. Lett.*, 45(6):494–497, August 1980. doi:10.1103/PhysRevLett.45.494.
- [11] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs. Quantized Hall Conductance in a Two-Dimensional Periodic Potential. *Phys. Rev. Lett.*, 49(6):405–408, August 1982. doi:10.1103/PhysRevLett.49.405.
- [12] F. D. M. Haldane. Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the “Parity Anomaly”. *Phys. Rev. Lett.*, 61(18):2015–2018, October 1988. doi:10.1103/PhysRevLett.61.2015.
- [13] Felix Bloch. Über die Quantenmechanik der Elektronen in Kristallgittern. *Z. Physik*, 52(7-8):555–600, July 1929. doi:10.1007/BF01339455.
- [14] P. Hohenberg and W. Kohn. Inhomogeneous Electron Gas. *Phys. Rev.*, 136(3B):B864–B871, November 1964. doi:10.1103/PhysRev.136.B864.
- [15] W. Kohn and L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.*, 140(4A):A1133–A1138, November 1965. doi:10.1103/PhysRev.140.A1133.
- [16] John P. Perdew, Kieron Burke, and Matthias Ernzerhof. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.*, 77(18):3865–3868, October 1996. doi:10.1103/PhysRevLett.77.3865.
- [17] J. C. Slater. A Simplification of the Hartree-Fock Method. *Phys. Rev.*, 81(3):385–390, February 1951. doi:10.1103/PhysRev.81.385.
- [18] Axel D. Becke. Density-functional thermochemistry. III. The role of exact exchange. *The Journal of Chemical Physics*, 98(7):5648–5652, April 1993. doi:10.1063/1.464913.

- [19] Jochen Heyd and Gustavo E. Scuseria. Efficient hybrid density functional calculations in solids: Assessment of the Heyd–Scuseria–Ernzerhof screened Coulomb hybrid functional. *The Journal of Chemical Physics*, 121(3):1187–1192, July 2004. doi:[10.1063/1.1760074](https://doi.org/10.1063/1.1760074).
- [20] Aliaksandr V. Krukau, Oleg A. Vydrov, Artur F. Izmaylov, and Gustavo E. Scuseria. Influence of the exchange screening parameter on the performance of screened hybrid functionals. *The Journal of Chemical Physics*, 125(22):224106, December 2006. doi:[10.1063/1.2404663](https://doi.org/10.1063/1.2404663).
- [21] Alex Willand, Yaroslav O. Kvashnin, Luigi Genovese, Álvaro Vázquez-Mayagoitia, Arpan Krishna Deb, Ali Sadeghi, Thierry Deutsch, and Stefan Goedecker. Norm-conserving pseudopotentials with chemical accuracy compared to all-electron calculations. *The Journal of Chemical Physics*, 138(10):104109, March 2013. doi:[10.1063/1.4793260](https://doi.org/10.1063/1.4793260).
- [22] E. Kucukbenli, M. Monni, B. I. Adetunji, X. Ge, G. A. Adebayo, N. Marzari, S. de Gironcoli, and A. Dal Corso. Projector augmented-wave and all-electron calculations across the periodic table: A comparison of structural and energetic properties. *ArXiv14043015 Cond-Mat*, April 2014. arXiv:[1404.3015](https://arxiv.org/abs/1404.3015).
- [23] Andrea Dal Corso. Pseudopotentials periodic table: From H to Pu. *Computational Materials Science*, 95:337–350, December 2014. doi:[10.1016/j.commatsci.2014.07.043](https://doi.org/10.1016/j.commatsci.2014.07.043).
- [24] M. Topsakal and R. M. Wentzcovitch. Accurate projected augmented wave (PAW) datasets for rare-earth elements (RE=La–Lu). *Computational Materials Science*, 95:263–270, December 2014. doi:[10.1016/j.commatsci.2014.07.030](https://doi.org/10.1016/j.commatsci.2014.07.030).
- [25] Kevin F. Garrity, Joseph W. Bennett, Karin M. Rabe, and David Vanderbilt. Pseudopotentials for high-throughput DFT calculations. *Computational Materials Science*, 81:446–452, January 2014. doi:[10.1016/j.commatsci.2013.08.053](https://doi.org/10.1016/j.commatsci.2013.08.053).
- [26] Martin Schlipf and François Gygi. Optimization algorithm for the generation of ONCV pseudopotentials. *Computer Physics Communications*, 196:36–44, November 2015. doi:[10.1016/j.cpc.2015.05.011](https://doi.org/10.1016/j.cpc.2015.05.011).
- [27] Kurt Lejaeghere, Gustav Bihlmayer, Torbjörn Björkman, Peter Blaha, Stefan Blügel, Volker Blum, Damien Caliste, Ivano E. Castelli, Stewart J. Clark, Andrea Dal Corso, Stefano de Gironcoli, Thierry Deutsch, John Kay Dewhurst, Igor Di Marco, Claudia Draxl, Marcin Dułak, Olle Eriksson, José A.

- Flores-Livas, Kevin F. Garrity, Luigi Genovese, Paolo Giannozzi, Matteo Giantomassi, Stefan Goedecker, Xavier Gonze, Oscar Grånäs, E. K. U. Gross, Andris Gulans, François Gygi, D. R. Hamann, Phil J. Hasnip, N. a. W. Holzwarth, Diana Iuşan, Dominik B. Jochym, François Jollet, Daniel Jones, Georg Kresse, Klaus Koepernik, Emine Küçükbenli, Yaroslav O. Kvashnin, Inka L. M. Loch, Sven Lubeck, Martijn Marsman, Nicola Marzari, Ulrike Nitzsche, Lars Nordström, Taisuke Ozaki, Lorenzo Paulatto, Chris J. Pickard, Ward Poelmans, Matt I. J. Probert, Keith Refson, Manuel Richter, Gian-Marco Rignanese, Santanu Saha, Matthias Scheffler, Martin Schlipf, Karlheinz Schwarz, Sangeeta Sharma, Francesca Tavazza, Patrik Thunström, Alexandre Tkatchenko, Marc Torrent, David Vanderbilt, Michiel J. van Setten, Veronique Van Speybroeck, John M. Wills, Jonathan R. Yates, Guo-Xu Zhang, and Stefaan Cottenier. Reproducibility in density functional theory calculations of solids. *Science*, 351(6280):aad3000, March 2016. doi:[10.1126/science.aad3000](https://doi.org/10.1126/science.aad3000).
- [28] Gianluca Prandini, Antimo Marrazzo, Ivano E. Castelli, Nicolas Mounet, and Nicola Marzari. Precision and efficiency in solid-state pseudopotential calculations. *ArXiv180605609 Cond-Mat Physicsphysics*, June 2018. arXiv:[1806.05609](https://arxiv.org/abs/1806.05609).
- [29] M. J. van Setten, M. Giantomassi, E. Bousquet, M. J. Verstraete, D. R. Hamann, X. Gonze, and G. M. Rignanese. The PseudoDojo: Training and grading a 85 element optimized norm-conserving pseudopotential table. *Computer Physics Communications*, 226:39–54, May 2018. doi:[10.1016/j.cpc.2018.01.012](https://doi.org/10.1016/j.cpc.2018.01.012).
- [30] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54(16):11169–11186, October 1996. doi:[10.1103/PhysRevB.54.11169](https://doi.org/10.1103/PhysRevB.54.11169).
- [31] Paolo Giannozzi, Stefano Baroni, Nicola Bonini, Matteo Calandra, Roberto Car, Carlo Cavazzoni, Davide Ceresoli, Guido L. Chiarotti, Matteo Cococcioni, Ismaila Dabo, Andrea Dal Corso, Stefano de Gironcoli, Stefano Fabris, Guido Fratesi, Ralph Gebauer, Uwe Gerstmann, Christos Gougousis, Anton Kokalj, Michele Lazzeri, Layla Martin-Samos, Nicola Marzari, Francesco Mauri, Riccardo Mazzarello, Stefano Paolini, Alfredo Pasquarello, Lorenzo Paulatto, Carlo Sbraccia, Sandro Scandolo, Gabriele Sclauzero, Ari P. Seitsonen, Alexander Smogunov, Paolo Umari, and Renata M. Wentzcovitch. QUANTUM ESPRESSO: A modular and open-source software project for quantum simulations of materials. *J. Phys.: Condens. Matter*, 21(39):395502, 2009. doi:[10.1088/0953-8984/21/39/395502](https://doi.org/10.1088/0953-8984/21/39/395502).

- [32] P. Giannozzi, O. Andreussi, T. Brumme, O. Bunau, M. Buongiorno Nardelli, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, M. Cococcioni, N. Colonna, I. Carnimeo, A. Dal Corso, S. de Gironcoli, P. Delugas, R. A. DiStasio Jr, A. Ferretti, A. Floris, G. Fratesi, G. Fugallo, R. Gebauer, U. Gerstmann, F. Giustino, T. Gorni, J. Jia, M. Kawamura, H-Y Ko, A. Kokalj, E. Küçükbenli, M. Lazzeri, M. Marsili, N. Marzari, F. Mauri, N. L. Nguyen, H.-V. Nguyen, A. Otero-de-la-Roza, L. Paulatto, S. Poncé, D. Rocca, R. Sabatini, B. Santra, M. Schlipf, A. P. Seitsonen, A. Smogunov, I Timrov, T. Thonhauser, P. Umari, N. Vast, X. Wu, and S. Baroni. Advanced capabilities for materials modelling with Quantum ESPRESSO. *J. Phys.: Condens. Matter*, 29(46):465901, 2017. doi:10.1088/1361-648X/aa8f79.
- [33] Peter Blaha, Karlheinz Schwarz, GKH Madsen, Dieter Kvasnicka, and Joachim Luitz. *WIEN2k*. Technical Universität Wien Austria, 2001.
- [34] Dr Karlheinz Schwarz. WIEN2k: An Augmented Plane Wave + Local Orbitals Program for Calculating Crystal Properties. page 281.
- [35] Tahir Yusufaly, David Vanderbilt, and Sinisa Coh. Tight-Binding Formalism in the Context of the PythTB Package. [http://www.physics.rutgers.edu/pythtb/\\_downloads/pythtb-formalism.pdf](http://www.physics.rutgers.edu/pythtb/_downloads/pythtb-formalism.pdf), February 2013.
- [36] Dominik Gresch. TBmodels documentation. <http://z2pack.ethz.ch/tbmodels>.
- [37] E. O. Kane. Chapter 3 The  $k \cdot p$  Method. In R. K. Willardson and Albert C. Beer, editors, *Semiconductors and Semimetals*, volume 1 of *Semiconductors and Semimetals*, pages 75–100. Elsevier, January 1966. doi:10.1016/S0080-8784(08)62376-5.
- [38] C. L. Kane and E. J. Mele. Quantum Spin Hall Effect in Graphene. *Phys. Rev. Lett.*, 95(22):226801, November 2005. doi:10.1103/PhysRevLett.95.226801.
- [39] C. L. Kane and E. J. Mele.  $Z_2$  topological Order and the Quantum Spin Hall Effect. *Phys. Rev. Lett.*, 95(14):146802, September 2005. doi:10.1103/PhysRevLett.95.146802.
- [40] J. Zak. Berry’s phase for energy bands in solids. *Phys. Rev. Lett.*, 62(23):2747–2750, June 1989. doi:10.1103/PhysRevLett.62.2747.
- [41] Shiing-Shen Chern. Characteristic classes of Hermitian manifolds. *Ann. Math.*, pages 85–121, 1946.

- 
- [42] D. J. Thouless. Wannier functions for magnetic sub-bands. *J. Phys. C: Solid State Phys.*, 17(12):L325, 1984. doi:[10.1088/0022-3719/17/12/003](https://doi.org/10.1088/0022-3719/17/12/003).
- [43] T. Thonhauser and David Vanderbilt. Insulator/Chern-insulator transition in the Haldane model. *Phys. Rev. B*, 74(23):235111, December 2006. doi:[10.1103/PhysRevB.74.235111](https://doi.org/10.1103/PhysRevB.74.235111).
- [44] Alexey A. Soluyanov and David Vanderbilt. Smooth gauge for topological insulators. *Phys. Rev. B*, 85(11):115415, March 2012. doi:[10.1103/PhysRevB.85.115415](https://doi.org/10.1103/PhysRevB.85.115415).
- [45] Georg W. Winkler, Alexey A. Soluyanov, and Matthias Troyer. Smooth gauge and Wannier functions for topological band structures in arbitrary dimensions. *Phys. Rev. B*, 93(3):035453, January 2016. doi:[10.1103/PhysRevB.93.035453](https://doi.org/10.1103/PhysRevB.93.035453).
- [46] Alexey A. Soluyanov. *Topological Aspects of Band Theory*. PhD thesis, Rutgers University - Graduate School - New Brunswick, 2012. doi:[10.7282/T3N0159G](https://doi.org/10.7282/T3N0159G).
- [47] A. Alexandradinata, Xi Dai, and B. Andrei Bernevig. Wilson-loop characterization of inversion-symmetric topological insulators. *Phys. Rev. B*, 89(15):155114, April 2014. doi:[10.1103/PhysRevB.89.155114](https://doi.org/10.1103/PhysRevB.89.155114).
- [48] Raphaël Leone. The geometry of (non)-Abelian adiabatic pumping. *J. Phys. A: Math. Theor.*, 44(29):295301, 2011. doi:[10.1088/1751-8113/44/29/295301](https://doi.org/10.1088/1751-8113/44/29/295301).
- [49] Alexey A. Soluyanov and David Vanderbilt. Wannier representation of  $\mathbb{Z}_2$  topological insulators. *Phys. Rev. B*, 83(3):035108, January 2011. doi:[10.1103/PhysRevB.83.035108](https://doi.org/10.1103/PhysRevB.83.035108).
- [50] Alexey A. Soluyanov and David Vanderbilt. Computing topological invariants without inversion symmetry. *Phys. Rev. B*, 83(23):235401, June 2011. doi:[10.1103/PhysRevB.83.235401](https://doi.org/10.1103/PhysRevB.83.235401).
- [51] Nicola Marzari and David Vanderbilt. Maximally localized generalized Wannier functions for composite energy bands. *Phys. Rev. B*, 56(20):12847–12865, November 1997. doi:[10.1103/PhysRevB.56.12847](https://doi.org/10.1103/PhysRevB.56.12847).
- [52] Claudia Sgierovello, Maria Peressi, and Raffaele Resta. Electron localization in the insulating state: Application to crystalline semiconductors. *Phys. Rev. B*, 64(11):115202, August 2001. doi:[10.1103/PhysRevB.64.115202](https://doi.org/10.1103/PhysRevB.64.115202).

## BIBLIOGRAPHY

---

- [53] Jeffrey C. Y. Teo, Liang Fu, and C. L. Kane. Surface states and topological invariants in three-dimensional topological insulators: Application to  $\text{Bi}_{1-x}\text{Sb}_x$ . *Phys. Rev. B*, 78(4):045426, July 2008. doi:10.1103/PhysRevB.78.045426.
- [54] Liang Fu. Topological Crystalline Insulators. *Phys. Rev. Lett.*, 106(10):106802, March 2011. doi:10.1103/PhysRevLett.106.106802.
- [55] A. Alexandradinata and B. Andrei Bernevig. Berry-phase description of topological crystalline insulators. *Phys. Rev. B*, 93(20):205104, May 2016. doi:10.1103/PhysRevB.93.205104.
- [56] GE Volovik. An analog of the quantum Hall effect in a superfluid  $^3\text{He}$  film. *Sov. Phys.-JETP Engl. Transl.*, 67(9):1804–1811, 1988.
- [57] Hermann Weyl. Elektron und gravitation. I. *Z. Für Phys.*, 56(5-6):330–352, 1929.
- [58] G. E. Volovik. Momentum space topology of Fermion zero modes brane. *Jetp Lett.*, 75(2):55–58, January 2002. doi:10.1134/1.1466475.
- [59] Grigory E Volovik. *The Universe in a Helium Droplet*, volume 117. Oxford University Press on Demand, 2003.
- [60] Zhong Fang, Naoto Nagaosa, Kei S. Takahashi, Atsushi Asamitsu, Roland Mathieu, Takeshi Ogasawara, Hiroyuki Yamada, Masashi Kawasaki, Yoshinori Tokura, and Kiyoyuki Terakura. The Anomalous Hall Effect and Magnetic Monopoles in Momentum Space. *Science*, 302(5642):92–95, October 2003. doi:10.1126/science.1089408.
- [61] B. Q. Lv, H. M. Weng, B. B. Fu, X. P. Wang, H. Miao, J. Ma, P. Richard, X. C. Huang, L. X. Zhao, G. F. Chen, Z. Fang, X. Dai, T. Qian, and H. Ding. Experimental Discovery of Weyl Semimetal TaAs. *Phys. Rev. X*, 5(3):031013, July 2015. doi:10.1103/PhysRevX.5.031013.
- [62] H. B. Nielsen and M. Ninomiya. Absence of neutrinos on a lattice: (I). Proof by homotopy theory. *Nuclear Physics B*, 185(1):20–40, July 1981. doi:10.1016/0550-3213(81)90361-8.
- [63] H. B. Nielsen and M. Ninomiya. A no-go theorem for regularizing chiral fermions. *Physics Letters B*, 105(2):219–223, October 1981. doi:10.1016/0370-2693(81)91026-1.

- [64] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim. The electronic properties of graphene. *Rev. Mod. Phys.*, 81(1):109–162, January 2009. doi:[10.1103/RevModPhys.81.109](https://doi.org/10.1103/RevModPhys.81.109).
- [65] S. M. Young, S. Zaheer, J. C. Y. Teo, C. L. Kane, E. J. Mele, and A. M. Rappe. Dirac Semimetal in Three Dimensions. *Phys. Rev. Lett.*, 108(14):140405, April 2012. doi:[10.1103/PhysRevLett.108.140405](https://doi.org/10.1103/PhysRevLett.108.140405).
- [66] Zhijun Wang, Yan Sun, Xing-Qiu Chen, Cesare Franchini, Gang Xu, Hongming Weng, Xi Dai, and Zhong Fang. Dirac semimetal and topological phase transitions in  $A_3\text{Bi}$  ( $A = \text{Na}, \text{K}, \text{Rb}$ ). *Phys. Rev. B*, 85(19):195320, May 2012. doi:[10.1103/PhysRevB.85.195320](https://doi.org/10.1103/PhysRevB.85.195320).
- [67] Zhijun Wang, Hongming Weng, Quansheng Wu, Xi Dai, and Zhong Fang. Three-dimensional Dirac semimetal and quantum transport in  $\text{Cd}_3\text{As}_2$ . *Phys. Rev. B*, 88(12):125427, September 2013. doi:[10.1103/PhysRevB.88.125427](https://doi.org/10.1103/PhysRevB.88.125427).
- [68] Sangjun Jeon, Brian B. Zhou, Andras Gyenis, Benjamin E. Feldman, Itamar Kimchi, Andrew C. Potter, Quinn D. Gibson, Robert J. Cava, Ashvin Vishwanath, and Ali Yazdani. Landau quantization and quasiparticle interference in the three-dimensional Dirac semimetal  $\text{Cd}_3\text{As}_2$ . *Nat. Mater.*, 13(9):851–856, September 2014. doi:[10.1038/nmat4023](https://doi.org/10.1038/nmat4023).
- [69] Xiangang Wan, Ari M. Turner, Ashvin Vishwanath, and Sergey Y. Savrasov. Topological semimetal and Fermi-arc surface states in the electronic structure of pyrochlore iridates. *Phys. Rev. B*, 83(20):205101, May 2011. doi:[10.1103/PhysRevB.83.205101](https://doi.org/10.1103/PhysRevB.83.205101).
- [70] M. A. Silaev and G. E. Volovik. Topological Fermi arcs in superfluid  $^3\text{He}$ . *Phys. Rev. B*, 86(21):214511, December 2012. doi:[10.1103/PhysRevB.86.214511](https://doi.org/10.1103/PhysRevB.86.214511).
- [71] H. B. Nielsen and Masao Ninomiya. The Adler-Bell-Jackiw anomaly and Weyl fermions in a crystal. *Physics Letters B*, 130(6):389–396, November 1983. doi:[10.1016/0370-2693\(83\)91529-0](https://doi.org/10.1016/0370-2693(83)91529-0).
- [72] Stephen L. Adler. Axial-Vector Vertex in Spinor Electrodynamics. *Phys. Rev.*, 177(5):2426–2438, January 1969. doi:[10.1103/PhysRev.177.2426](https://doi.org/10.1103/PhysRev.177.2426).
- [73] J. S. Bell and R. Jackiw. A PCAC puzzle:  $\pi^0 \rightarrow \gamma\gamma$  in the  $\sigma$ -model. *Nuovo Cimento A (1965-1970)*, 60(1):47–61, March 1969. doi:[10.1007/BF02823296](https://doi.org/10.1007/BF02823296).

- [74] Yan Sun, Shu-Chun Wu, Mazhar N. Ali, Claudia Felser, and Binghai Yan. Prediction of Weyl semimetal in orthorhombic MoTe<sub>2</sub>. *Phys. Rev. B*, 92(16):161107, October 2015. doi:[10.1103/PhysRevB.92.161107](https://doi.org/10.1103/PhysRevB.92.161107).
- [75] A. Tamai, Q. S. Wu, I. Cucchi, F. Y. Bruno, S. Riccò, T. K. Kim, M. Hoesch, C. Barreateau, E. Giannini, C. Besnard, A. A. Soluyanov, and F. Baumberger. Fermi Arcs and Their Topological Character in the Candidate Type-II Weyl Semimetal MoTe<sub>2</sub>. *Phys. Rev. X*, 6(3):031021, August 2016. doi:[10.1103/PhysRevX.6.031021](https://doi.org/10.1103/PhysRevX.6.031021).
- [76] Yun Wu, Daixiang Mou, Na Hyun Jo, Kewei Sun, Lunan Huang, S. L. Bud'ko, P. C. Canfield, and Adam Kaminski. Observation of Fermi arcs in the type-II Weyl semimetal candidate WTe<sub>2</sub>. *Phys. Rev. B*, 94(12):121113, September 2016. doi:[10.1103/PhysRevB.94.121113](https://doi.org/10.1103/PhysRevB.94.121113).
- [77] Lunan Huang, Timothy M. McCormick, Masayuki Ochi, Zhiying Zhao, Michi-To Suzuki, Ryotaro Arita, Yun Wu, Daixiang Mou, Huibo Cao, Jiaqiang Yan, Nandini Trivedi, and Adam Kaminski. Spectroscopic evidence for a type II Weyl semimetallic state in MoTe<sub>2</sub>. *Nat. Mater.*, 15(11):1155–1160, November 2016. doi:[10.1038/nmat4685](https://doi.org/10.1038/nmat4685).
- [78] Ke Deng, Guoliang Wan, Peng Deng, Kenan Zhang, Shijie Ding, Eryin Wang, Mingzhe Yan, Huaqing Huang, Hongyun Zhang, Zhilin Xu, Jonathan Denlinger, Alexei Fedorov, Haitao Yang, Wenhui Duan, Hong Yao, Yang Wu, Shoushan Fan, Haijun Zhang, Xi Chen, and Shuyun Zhou. Experimental observation of topological Fermi arcs in type-II Weyl semimetal MoTe<sub>2</sub>. *Nat. Phys.*, 12(12):1105–1110, December 2016. doi:[10.1038/nphys3871](https://doi.org/10.1038/nphys3871).
- [79] Nitesh Kumar, Yan Sun, Nan Xu, Kaustuv Manna, Mengyu Yao, Vicky Süss, Inge Leermakers, Olga Young, Tobias Förster, Marcus Schmidt, Horst Borrmann, Binghai Yan, Uli Zeitler, Ming Shi, Claudia Felser, and Chandra Shekhar. Extremely high magnetoresistance and conductivity in the type-II Weyl semimetals WP<sub>2</sub> and MoP<sub>2</sub>. *Nat. Commun.*, 8(1):1642, November 2017. doi:[10.1038/s41467-017-01758-z](https://doi.org/10.1038/s41467-017-01758-z).
- [80] Georg W. Winkler, QuanSheng Wu, Matthias Troyer, Peter Krogstrup, and Alexey A. Soluyanov. Topological Phases in InAs<sub>1-x</sub>Sb<sub>x</sub>: From Novel Topological Semimetal to Majorana Wire. *Phys. Rev. Lett.*, 117(7):076403, August 2016. doi:[10.1103/PhysRevLett.117.076403](https://doi.org/10.1103/PhysRevLett.117.076403).
- [81] Barry Bradlyn, Jennifer Cano, Zhijun Wang, M. G. Vergniory, C. Felser, R. J. Cava, and B. Andrei Bernevig. Beyond Dirac and Weyl

- fermions: Unconventional quasiparticles in conventional crystals. *Science*, 353(6299):aaf5037, August 2016. doi:10.1126/science.aaf5037.
- [82] Zhijun Wang, A. Alexandradinata, R. J. Cava, and B. Andrei Bernevig. Hourglass fermions. *Nature*, 532(7598):189–194, April 2016. doi:10.1038/nature17410.
- [83] Hongming Weng, Yunye Liang, Qiunan Xu, Rui Yu, Zhong Fang, Xi Dai, and Yoshiyuki Kawazoe. Topological node-line semimetal in three-dimensional graphene networks. *Phys. Rev. B*, 92(4):045108, July 2015. doi:10.1103/PhysRevB.92.045108.
- [84] Tomáš Bzdušek, QuanSheng Wu, Andreas Rüegg, Manfred Sigrist, and Alexey A. Soluyanov. Nodal-chain metals. *Nature*, 538(7623):75–78, October 2016. doi:10.1038/nature19099.
- [85] Guang Bian, Tay-Rong Chang, Raman Sankar, Su-Yang Xu, Hao Zheng, Titus Neupert, Ching-Kai Chiu, Shin-Ming Huang, Guoqing Chang, Ilya Belopolski, Daniel S. Sanchez, Madhab Neupane, Nasser Alidoust, Chang Liu, BaoKai Wang, Chi-Cheng Lee, Horng-Tay Jeng, Chenglong Zhang, Zhujun Yuan, Shuang Jia, Arun Bansil, Fangcheng Chou, Hsin Lin, and M. Zahid Hasan. Topological nodal-line fermions in spin-orbit metal PbTaSe<sub>2</sub>. *Nat. Commun.*, 7:10556, February 2016. doi:10.1038/ncomms10556.
- [86] Lilia S. Xie, Leslie M. Schoop, Elizabeth M. Seibel, Quinn D. Gibson, Weiwei Xie, and Robert J. Cava. A new form of Ca<sub>3</sub>P<sub>2</sub> with a ring of Dirac nodes. *APL Materials*, 3(8):083602, July 2015. doi:10.1063/1.4926545.
- [87] Rui Yu, Hongming Weng, Zhong Fang, Xi Dai, and Xiao Hu. Topological Node-Line Semimetal and Dirac Semimetal State in Antiperovskite Cu<sub>3</sub>PdN. *Phys. Rev. Lett.*, 115(3):036807, July 2015. doi:10.1103/PhysRevLett.115.036807.
- [88] Youngkuk Kim, Benjamin J. Wieder, C. L. Kane, and Andrew M. Rappe. Dirac Line Nodes in Inversion-Symmetric Crystals. *Phys. Rev. Lett.*, 115(3):036806, July 2015. doi:10.1103/PhysRevLett.115.036806.
- [89] Cmglee. File:Phase diagram of water.svg. [https://en.wikipedia.org/wiki/File:Phase\\_diagram\\_of\\_water.svg](https://en.wikipedia.org/wiki/File:Phase_diagram_of_water.svg).
- [90] Kefeng Wang, D. Graf, Lijun Li, Limin Wang, and C. Petrovic. Anisotropic giant magnetoresistance in NbSb<sub>2</sub>. *Sci. Rep.*, 4:7328, December 2014. doi:10.1038/srep07328.

- [91] Yi-Yan Wang, Qiao-He Yu, Peng-Jie Guo, Kai Liu, and Tian-Long Xia. Resistivity plateau and extremely large magnetoresistance in NbAs<sub>2</sub> and TaAs<sub>2</sub>. *Phys. Rev. B*, 94(4):041103, July 2016. doi:10.1103/PhysRevB.94.041103.
- [92] Desheng Wu, Jian Liao, Wei Yi, Xia Wang, Peigang Li, Hongming Weng, Youguo Shi, Yongqing Li, Jianlin Luo, Xi Dai, and Zhong Fang. Giant semiclassical magnetoresistance in high mobility TaAs<sub>2</sub> semimetal. *Appl. Phys. Lett.*, 108(4):042105, January 2016. doi:10.1063/1.4940924.
- [93] Yongkang Luo, R. D. McDonald, P. F. S. Rosa, B. Scott, N. Wakeham, N. J. Ghimire, E. D. Bauer, J. D. Thompson, and F. Ronning. Anomalous electronic structure and magnetoresistance in TaAs<sub>2</sub>. *Sci. Rep.*, 6:27294, June 2016. doi:10.1038/srep27294.
- [94] Zhujun Yuan, Hong Lu, Yongjie Liu, Junfeng Wang, and Shuang Jia. Large magnetoresistance in compensated semimetals TaAs<sub>2</sub> and NbAs<sub>2</sub>. *Phys. Rev. B*, 93(18):184405, May 2016. doi:10.1103/PhysRevB.93.184405.
- [95] Yupeng Li, Zhen Wang, Yunhao Lu, Xiaojun Yang, Zhixuan Shen, Feng Sheng, Chunmu Feng, Yi Zheng, and Zhu-An Xu. Negative Magnetoresistance in Topological Semimetals of Transition-Metal Dipnictides with Nontrivial  $\mathbb{Z}_2$  Indices. *ArXiv160304056 Cond-Mat*, March 2016. arXiv:1603.04056.
- [96] Bing Shen, Xiaoyu Deng, Gabriel Kotliar, and Ni Ni. Fermi surface topology and negative longitudinal magnetoresistance observed in the semimetal NbAs<sub>2</sub>. *Phys. Rev. B*, 93(19):195119, May 2016. doi:10.1103/PhysRevB.93.195119.
- [97] Zheng Wang, Yupeng Li, Yunhao Lu, Zhixuan Shen, Feng Sheng, Chunmu Feng, Yi Zheng, and Zhuan Xu. Topological phase transition induced extreme magnetoresistance in TaSb<sub>2</sub>. *ArXiv160301717 Cond-Mat*, March 2016. arXiv:1603.01717.
- [98] Yuke Li, Lin Li, Jialu Wang, Tingting Wang, Xiaofeng Xu, Chuanying Xi, Chao Cao, and Jianhui Dai. Resistivity plateau and negative magnetoresistance in the topological semimetal TaSb<sub>2</sub>. *Phys. Rev. B*, 94(12):121115, September 2016. doi:10.1103/PhysRevB.94.121115.
- [99] Jennifer Cano, Barry Bradlyn, Zhijun Wang, Max Hirschberger, N. P. Ong, and B. A. Bernevig. Chiral anomaly factory: Creating Weyl fermions with a magnetic field. *Phys. Rev. B*, 95(16):161306, April 2017. doi:10.1103/PhysRevB.95.161306.

- [100] RG Ling and C Belin. Affinement de la structure cristalline du diarseniure de tantale. *CR Acad Sci Paris*, 292:891–893, 1981.
- [101] F. Hulliger. New Representatives of the NbAs<sub>2</sub> and ZrAs<sub>2</sub> Structures. *Nature*, 204(4960):775, November 1964. doi:10.1038/204775a0.
- [102] Ya F. Lomnytska and V. V. Berezovets. Phase Relations in the Nb-Ni-Sb System. *Inorg Mater*, 41(11):1166–1171, November 2005. doi:10.1007/s10789-005-0281-z.
- [103] Wolfgang Bensch and Wolfram Heid. NbAs<sub>2</sub>. *Acta Crystallogr. C*, 51(11):2205–2207, 1995.
- [104] P. E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50(24):17953–17979, December 1994. doi:10.1103/PhysRevB.50.17953.
- [105] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59(3):1758–1775, January 1999. doi:10.1103/PhysRevB.59.1758.
- [106] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof. Hybrid functionals based on a screened Coulomb potential. *The Journal of Chemical Physics*, 118(18):8207–8215, April 2003. doi:10.1063/1.1564060.
- [107] Chenchao Xu, Jia Chen, Guo-Xiang Zhi, Yuke Li, Jianhui Dai, and Chao Cao. Electronic structures of transition metal dipnictides  $X\text{Pn}_2$  ( $X = \text{Ta}, \text{Nb}$ ;  $\text{Pn} = \text{P}, \text{As}, \text{Sb}$ ). *Phys. Rev. B*, 93(19):195106, May 2016. doi:10.1103/PhysRevB.93.195106.
- [108] G. F. Koster, J. O. Dimmock, R. G. Wheeler, and H. Statz. *Properties of the Thirty-Two Point Groups*, volume 24. The MIT Press, 1963.
- [109] K. Schwarz, P. Blaha, and G. K. H. Madsen. Electronic structure calculations of solids using the WIEN2k package for material sciences. *Computer Physics Communications*, 147(1):71–76, August 2002. doi:10.1016/S0010-4655(02)00206-0.
- [110] Eric Jones, Travis Oliphant, Pearu Peterson, and others. SciPy: Open source scientific tools for Python. 2001. [Online; accessed 2018-05-31].
- [111] Arash A. Mostofi, Jonathan R. Yates, Young-Su Lee, Ivo Souza, David Vanderbilt, and Nicola Marzari. Wannier90: A tool for obtaining maximally-localised Wannier functions. *Computer Physics Communications*, 178(9):685–699, May 2008. doi:10.1016/j.cpc.2007.11.016.

- [112] Arash A. Mostofi, Jonathan R. Yates, Giovanni Pizzi, Young-Su Lee, Ivo Souza, David Vanderbilt, and Nicola Marzari. An updated version of wannier90: A tool for obtaining maximally-localised Wannier functions. *Computer Physics Communications*, 185(8):2309–2310, August 2014. doi:[10.1016/j.cpc.2014.05.003](https://doi.org/10.1016/j.cpc.2014.05.003).
- [113] Mildred S. Dresselhaus, Gene Dresselhaus, and Ado Jorio. *Group Theory: Application to the Physics of Condensed Matter*. Springer Science & Business Media, December 2007.
- [114] Guido Van Rossum and others. Python Programming Language. In *USENIX Annual Technical Conference*, volume 41, 2007.
- [115] Aaron Meurer, Christopher P. Smith, Mateusz Paprocki, Ondřej Čertík, Sergey B. Kirpichev, Matthew Rocklin, AMiT Kumar, Sergiu Ivanov, Jason K. Moore, Sartaj Singh, Thilina Rathnayake, Sean Vig, Brian E. Granger, Richard P. Muller, Francesco Bonazzi, Harsh Gupta, Shivam Vats, Fredrik Johansson, Fabian Pedregosa, Matthew J. Curry, Andy R. Terrel, Štěpán Roučka, Ashutosh Saboo, Isuru Fernando, Sumith Kulal, Robert Cimrman, and Anthony Scopatz. SymPy: Symbolic computing in Python. *PeerJ Comput. Sci.*, 3:e103, January 2017. doi:[10.7717/peerj-cs.103](https://doi.org/10.7717/peerj-cs.103).
- [116] Travis E Oliphant. *A Guide to NumPy*, volume 1. Trelgol Publishing USA, 2006.
- [117] Eugene M. Luks, Ferenc Rákóczi, and Charles R. B. Wright. Some Algorithms for Nilpotent Permutation Groups. *Journal of Symbolic Computation*, 23(4):335–354, April 1997. doi:[10.1006/jsco.1996.0092](https://doi.org/10.1006/jsco.1996.0092).
- [118] Ivo Souza, Nicola Marzari, and David Vanderbilt. Maximally localized Wannier functions for entangled energy bands. *Phys. Rev. B*, 65(3):035109, December 2001. doi:[10.1103/PhysRevB.65.035109](https://doi.org/10.1103/PhysRevB.65.035109).
- [119] Alberto Franceschetti and Alex Zunger. The inverse band-structure problem of finding an atomic configuration with given electronic properties. *Nature*, 402(6757):60–63, November 1999. doi:[10.1038/46995](https://doi.org/10.1038/46995).
- [120] G. H. Jóhannesson, T. Bligaard, A. V. Ruban, H. L. Skriver, K. W. Jacobsen, and J. K. Nørskov. Combined Electronic Structure and Evolutionary Search Approach to Materials Design. *Phys. Rev. Lett.*, 88(25):255506, June 2002. doi:[10.1103/PhysRevLett.88.255506](https://doi.org/10.1103/PhysRevLett.88.255506).

- [121] Stefano Curtarolo, Dane Morgan, Kristin Persson, John Rodgers, and Gerbrand Ceder. Predicting Crystal Structures with Data Mining of Quantum Calculations. *Phys. Rev. Lett.*, 91(13):135503, September 2003. doi:[10.1103/PhysRevLett.91.135503](https://doi.org/10.1103/PhysRevLett.91.135503).
- [122] Stefano Curtarolo, Gus L. W. Hart, Marco Buongiorno Nardelli, Natalio Mingo, Stefano Sanvito, and Ohad Levy. The high-throughput highway to computational materials design. *Nat. Mater.*, 12(3):191–201, March 2013. doi:[10.1038/nmat3568](https://doi.org/10.1038/nmat3568).
- [123] Giovanni Pizzi, Andrea Cepellotti, Riccardo Sabatini, Nicola Marzari, and Boris Kozinsky. AiiDA: Automated interactive infrastructure and database for computational science. *Computational Materials Science*, 111:218–230, January 2016. doi:[10.1016/j.commatsci.2015.09.013](https://doi.org/10.1016/j.commatsci.2015.09.013).
- [124] Anubhav Jain, Shyue Ping Ong, Wei Chen, Bharat Medasani, Xiaohui Qu, Michael Kocher, Miriam Brafman, Guido Petretto, Gian-Marco Rignanese, Geoffroy Hautier, Daniel Gunter, and Kristin A. Persson. FireWorks: A dynamic workflow system designed for high-throughput applications. *Concurr. Comput. Pract. Exp.*, 27(17):5037–5059, December 2015. doi:[10.1002/cpe.3505](https://doi.org/10.1002/cpe.3505).
- [125] Jamal I. Mustafa, Sinisa Coh, Marvin L. Cohen, and Steven G. Louie. Automated construction of maximally localized Wannier functions: Optimized projection functions method. *Phys. Rev. B*, 92(16):165134, October 2015. doi:[10.1103/PhysRevB.92.165134](https://doi.org/10.1103/PhysRevB.92.165134).
- [126] A. Yu Kitaev. Unpaired Majorana fermions in quantum wires. *Phys.-Usp.*, 44(10S):131, 2001. doi:[10.1070/1063-7869/44/10S/S29](https://doi.org/10.1070/1063-7869/44/10S/S29).
- [127] Roman M. Lutchyn, Jay D. Sau, and S. Das Sarma. Majorana Fermions and a Topological Phase Transition in Semiconductor-Superconductor Heterostructures. *Phys. Rev. Lett.*, 105(7):077001, August 2010. doi:[10.1103/PhysRevLett.105.077001](https://doi.org/10.1103/PhysRevLett.105.077001).
- [128] Yuval Oreg, Gil Refael, and Felix von Oppen. Helical Liquids and Majorana Bound States in Quantum Wires. *Phys. Rev. Lett.*, 105(17):177002, October 2010. doi:[10.1103/PhysRevLett.105.177002](https://doi.org/10.1103/PhysRevLett.105.177002).
- [129] D. B. Litvin and V. Kopský. Seitz notation for symmetry operations of space groups. *Acta Cryst A*, 67(4):415–418, July 2011. doi:[10.1107/S010876731101378X](https://doi.org/10.1107/S010876731101378X).

- [130] Howard E. Haber. Three-Dimensional Proper and Improper Rotation Matrices, 2011.
- [131] R. Sakuma. Symmetry-adapted Wannier functions in the maximal localization procedure. *Phys. Rev. B*, 87(23):235109, June 2013. doi:10.1103/PhysRevB.87.235109.
- [132] J. A. Nelder and R. Mead. A Simplex Method for Function Minimization. *Comput J*, 7(4):308–313, January 1965. doi:10.1093/comjnl/7.4.308.
- [133] Jason Alicea. Majorana fermions in a tunable semiconductor device. *Phys. Rev. B*, 81(12):125318, March 2010. doi:10.1103/PhysRevB.81.125318.
- [134] D. I. Pikulin, J. P. Dahlhaus, M. Wimmer, H. Schomerus, and C. W. J. Beenakker. A zero-voltage conductance peak from weak antilocalization in a Majorana nanowire. *New J. Phys.*, 14(12):125011, 2012. doi:10.1088/1367-2630/14/12/125011.
- [135] Jason Alicea. New directions in the pursuit of Majorana fermions in solid state systems. *Rep. Prog. Phys.*, 75(7):076501, 2012. doi:10.1088/0034-4885/75/7/076501.
- [136] Shuo Mi, D. I. Pikulin, M. Wimmer, and C. W. J. Beenakker. Proposal for the detection and braiding of Majorana fermions in a quantum spin Hall insulator. *Phys. Rev. B*, 87(24):241405, June 2013. doi:10.1103/PhysRevB.87.241405.
- [137] C.W.J. Beenakker. Search for Majorana Fermions in Superconductors. *Annu. Rev. Condens. Matter Phys.*, 4(1):113–136, 2013. doi:10.1146/annurev-conmatphys-030212-184337.
- [138] A. Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2–30, January 2003. doi:10.1016/S0003-4916(02)00018-0.
- [139] M. T. Deng, S. Vaitiekėnas, E. B. Hansen, J. Danon, M. Leijnse, K. Flensberg, J. Nygård, P. Krogstrup, and C. M. Marcus. Majorana bound state in a coupled quantum-dot hybrid-nanowire system. *Science*, 354(6319):1557–1562, December 2016. doi:10.1126/science.aaf3961.
- [140] Chaoxing Liu, Taylor L. Hughes, Xiao-Liang Qi, Kang Wang, and Shou-Cheng Zhang. Quantum Spin Hall Effect in Inverted Type-II Semiconductors. *Phys. Rev. Lett.*, 100(23):236601, June 2008. doi:10.1103/PhysRevLett.100.236601.

- 
- [141] V. Mourik, K. Zuo, S. M. Frolov, S. R. Plissard, E. P. a. M. Bakkers, and L. P. Kouwenhoven. Signatures of Majorana Fermions in Hybrid Superconductor-Semiconductor Nanowire Devices. *Science*, 336(6084):1003–1007, May 2012. doi:[10.1126/science.1222360](https://doi.org/10.1126/science.1222360).
- [142] H. O. H. Churchill, V. Fatemi, K. Grove-Rasmussen, M. T. Deng, P. Caroff, H. Q. Xu, and C. M. Marcus. Superconductor-nanowire devices from tunneling to the multichannel regime: Zero-bias oscillations and magnetoconductance crossover. *Phys. Rev. B*, 87(24):241401, June 2013. doi:[10.1103/PhysRevB.87.241401](https://doi.org/10.1103/PhysRevB.87.241401).
- [143] Lars Hedin. New Method for Calculating the One-Particle Green’s Function with Application to the Electron-Gas Problem. *Phys. Rev.*, 139(3A):A796–A823, August 1965. doi:[10.1103/PhysRev.139.A796](https://doi.org/10.1103/PhysRev.139.A796).
- [144] J. C. Slater and G. F. Koster. Simplified LCAO Method for the Periodic Potential Problem. *Phys. Rev.*, 94(6):1498–1524, June 1954. doi:[10.1103/PhysRev.94.1498](https://doi.org/10.1103/PhysRev.94.1498).
- [145] Yaohua Tan, Michael Povolotskyi, Tillmann Kubis, Yu He, Zhengping Jiang, Gerhard Klimeck, and Timothy B. Boykin. Empirical tight binding parameters for GaAs and MgO with explicit basis through DFT mapping. *J Comput Electron*, 12(1):56–60, March 2013. doi:[10.1007/s10825-013-0436-0](https://doi.org/10.1007/s10825-013-0436-0).
- [146] Yaohua P. Tan, Michael Povolotskyi, Tillmann Kubis, Timothy B. Boykin, and Gerhard Klimeck. Tight-binding analysis of Si and GaAs ultrathin bodies with subatomic wave-function resolution. *Phys. Rev. B*, 92(8):085301, August 2015. doi:[10.1103/PhysRevB.92.085301](https://doi.org/10.1103/PhysRevB.92.085301).
- [147] Yaohua Tan, Michael Povolotskyi, Tillmann Kubis, Timothy B. Boykin, and Gerhard Klimeck. Transferable tight-binding model for strained group IV and III-V materials and heterostructures. *Phys. Rev. B*, 94(4):045311, July 2016. doi:[10.1103/PhysRevB.94.045311](https://doi.org/10.1103/PhysRevB.94.045311).
- [148] Jean-Marc Jancu, Reinhard Scholz, Fabio Beltram, and Franco Bassani. Empirical spds\* tight-binding calculation for cubic semiconductors: General method and material parameters. *Phys. Rev. B*, 57(11):6493–6507, March 1998. doi:[10.1103/PhysRevB.57.6493](https://doi.org/10.1103/PhysRevB.57.6493).
- [149] Timothy B. Boykin. Improved fits of the effective masses at  $\Gamma$  in the spin-orbit, second-nearest-neighbor sp<sup>3</sup>s\* model: Results from analytic expressions. *Phys. Rev. B*, 56(15):9613–9618, October 1997. doi:[10.1103/PhysRevB.56.9613](https://doi.org/10.1103/PhysRevB.56.9613).

## BIBLIOGRAPHY

---

- [150] Yoon-Suk Kim, Martijn Marsman, Georg Kresse, Fabien Tran, and Peter Blaha. Towards efficient band structure and effective mass calculations for III-V direct band-gap semiconductors. *Phys. Rev. B*, 82(20):205212, November 2010. doi:10.1103/PhysRevB.82.205212.
- [151] O. Madelung, U. Rössler, and M. Schulz, editors. *Condensed Matter: Group IV Elements, IV-IV and III-V Compounds*. Landolt-Börnstein - Group III Condensed Matter. Springer, Berlin Heidelberg, 2001. doi:10.1007/B60136.
- [152] Rico Häuselmann. AiiDA-VASP documentation. <https://aiida-vasp.readthedocs.io/>.
- [153] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan. Band parameters for III-V compound semiconductors and their alloys. *Journal of Applied Physics*, 89(11):5815–5875, June 2001. doi:10.1063/1.1368156.
- [154] M. Graf and P. Vogl. Electromagnetic fields and dielectric response in empirical tight-binding theory. *Phys. Rev. B*, 51(8):4940–4949, February 1995. doi:10.1103/PhysRevB.51.4940.
- [155] Dominik Gresch, QuanSheng Wu, Georg W. Winkler, Rico Häuselmann, Matthias Troyer, and Alexey A. Soluyanov. Automated construction of symmetrized Wannier-like tight-binding models from ab initio calculations. *Phys. Rev. Materials*, 2(10):103805, October 2018. doi:10.1103/PhysRevMaterials.2.103805.
- [156] C. G. Broyden. A class of methods for solving nonlinear simultaneous equations. *Math. Comp.*, 19(92):577–593, 1965. doi:10.1090/S0025-5718-1965-0198670-6.
- [157] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by Simulated Annealing. *Science*, 220(4598):671–680, May 1983. doi:10.1126/science.220.4598.671.
- [158] Stefan Goedecker. Minima hopping: An efficient search method for the global minimum of the potential energy surface of complex molecular systems. *The Journal of Chemical Physics*, 120(21):9911–9917, May 2004. doi:10.1063/1.1724816.
- [159] Aric Hagberg, Pieter Swart, and Daniel S Chult. Exploring network structure, dynamics, and function using NetworkX. Technical report, Los Alamos National Lab.(LANL), Los Alamos, NM (United States), 2008.

- [160] G. V. Trunk. Statistical estimation of the intrinsic dimensionality of data collections. *Information and Control*, 12(5):508–525, May 1968. doi:10.1016/S0019-9958(68)90591-3.
- [161] K. Fukunaga and D. R. Olsen. An Algorithm for Finding Intrinsic Dimensionality of Data. *IEEE Trans. Comput.*, C-20(2):176–183, February 1971. doi:10.1109/T-C.1971.223208.
- [162] J. Bruske and G. Sommer. Intrinsic dimensionality estimation with optimally topology preserving maps. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20(5):572–575, May 1998. doi:10.1109/34.682189.
- [163] F. Camastra and A. Vinciarelli. Estimating the intrinsic dimension of data with a fractal-based method. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(10):1404–1407, October 2002. doi:10.1109/TPAMI.2002.1039212.
- [164] Balázs Kégl. Intrinsic dimension estimation using packing numbers. In *Advances in Neural Information Processing Systems*, pages 697–704, 2003.
- [165] Elizaveta Levina and Peter J Bickel. Maximum likelihood estimation of intrinsic dimension. In *Advances in Neural Information Processing Systems*, pages 777–784, 2005.
- [166] Mingyu Fan, Hong Qiao, and Bo Zhang. Intrinsic dimension estimation of manifolds by incising balls. *Pattern Recognition*, 42(5):780–787, May 2009. doi:10.1016/j.patcog.2008.09.016.
- [167] K. M. Carter, R. Raich, and A. O. Hero III. On Local Intrinsic Dimension Estimation and Its Applications. *IEEE Trans. Signal Process.*, 58(2):650–663, February 2010. doi:10.1109/TSP.2009.2031722.
- [168] M. Bartlett and A. Tressaud. Sur un nouveau type structurale de tetrafluorures d’elements de transition: Synthese, etude cristallographique et magnetique de  $\text{IrF}_4$ . *Comptes Rendus Hebd. Seances Acad. Sci. Ser. C Sci. Chim.* 1966-, 278:1501–1504, 1974.
- [169] Anubhav Jain, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, Dan Gunter, David Skinner, Gerbrand Ceder, and Kristin A. Persson. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Materials*, 1(1):011002, July 2013. doi:10.1063/1.4812323.
- [170] Shyue Ping Ong, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L. Chevrier,

- Kristin A. Persson, and Gerbrand Ceder. Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis. *Computational Materials Science*, 68:314–319, February 2013. doi: [10.1016/j.commatsci.2012.10.028](https://doi.org/10.1016/j.commatsci.2012.10.028).
- [171] Yoyo Hinuma, Giovanni Pizzi, Yu Kumagai, Fumiyasu Oba, and Isao Tanaka. Band structure diagram paths based on crystallography. *Computational Materials Science*, 128:140–184, February 2017. doi: [10.1016/j.commatsci.2016.10.015](https://doi.org/10.1016/j.commatsci.2016.10.015).
- [172] Atsushi Togo and Isao Tanaka. Spglib: A software library for crystal symmetry search. *ArXiv180801590 Cond-Mat*, August 2018. arXiv: [1808.01590](https://arxiv.org/abs/1808.01590).
- [173] Ask Hjorth Larsen, Jens Jørgen Mortensen, Jakob Blomqvist, Ivano E. Castelli, Rune Christensen, Marcin Dułak, Jesper Friis, Michael N. Groves, Bjørk Hammer, Cory Hargus, Eric D. Hermes, Paul C. Jennings, Peter Bjerre Jensen, James Kermode, John R. Kitchin, Esben Leonhard Kolsbjerg, Joseph Kubal, Kristen Kaasbjerg, Steen Lysgaard, Jón Bergmann Maronsson, Tristan Maxson, Thomas Olsen, Lars Pastewka, Andrew Peterson, Carsten Rostgaard, Jakob Schiøtz, Ole Schütt, Mikkel Strange, Kristian S. Thygesen, Tejs Vegge, Lasse Vilhelmsen, Michael Walter, Zhenhua Zeng, and Karsten W. Jacobsen. The atomic simulation environment—a Python library for working with atoms. *J. Phys.: Condens. Matter*, 29(27):273002, 2017. doi: [10.1088/1361-648X/aa680e](https://doi.org/10.1088/1361-648X/aa680e).
- [174] M. I. Aroyo, A. Kirov, C. Capillas, J. M. Perez-Mato, and H. Wondratschek. Bilbao Crystallographic Server. II. Representations of crystallographic point groups and space groups. *Acta Cryst A*, 62(2):115–128, March 2006. doi: [10.1107/S0108767305040286](https://doi.org/10.1107/S0108767305040286).
- [175] Mois Ilia Aroyo, Juan Manuel Perez-Mato, Cesar Capillas, Eli Kroumova, Svetoslav Ivantchev, Gotzon Madariaga, Asen Kirov, and Hans Wondratschek. Bilbao Crystallographic Server: I. Databases and crystallographic computing programs. *Z. Für Krist. - Cryst. Mater.*, 221(1):15–27, 2009. doi: [10.1524/zkri.2006.221.1.15](https://doi.org/10.1524/zkri.2006.221.1.15).
- [176] Mois I Aroyo, JM Perez-Mato, D Orobengoa, E Tasci, G De La Flor, and A Kirov. Crystallography online: Bilbao crystallographic server. *Bulg Chem Commun*, 43(2):183–197, 2011.
- [177] N. Kobayashi, Z. Hiroi, and M. Takano. Compounds and phase relations in the (Sr O) - (Ca O) - (Cu O) system under high pressure. *J. Solid State Chem.*, 132:274–283, 1997.

- [178] Sabine Körbel, Miguel A. L. Marques, and Silvana Botti. Stability and electronic properties of new inorganic perovskites from high-throughput ab initio calculations. *J. Mater. Chem. C*, 4(15):3157–3167, April 2016. doi: [10.1039/C5TC04172D](https://doi.org/10.1039/C5TC04172D).
- [179] Xiao-Qi Sun, Tomáš Bzdušek, and Shou-Cheng Zhang. Conversion rules for Weyl points and nodal lines in topological media. *ArXiv180306364 Cond-Mat*, March 2018. [arXiv:1803.06364](https://arxiv.org/abs/1803.06364).
- [180] U. Baelz and K. Schubert. Kristallstruktur von Pd<sub>2</sub>As(r) und Pd<sub>2</sub>Sb. *J. -Common Met.*, 19:300–304, 1969.
- [181] M. Puselj and K. Schubert. Kristallstruktur von Au<sub>2</sub>Ga. *J. -Common Met.*, 38:83–90, 1974.



# Acknowledgements

First I would like to thank Alexey Soluyanov, who supervised me since my semester thesis writing parts of what would later become Z2Pack. He introduced me to the field of topological materials, and his passion and kind-hearted supervision in no small parts motivated me to continue research in this area. Throughout my Ph.D Alexey has been a great mentor, guiding me to become an independent scientist.

Of course I am also indebted to my supervisor Matthias Troyer, who allowed me to pursue this research. Even before I started working in his group, Matthias helped spark my interest in coding through his inspiring lectures. I would also like to thank Matthias for fostering an environment that combines excellent research with a dedication to creating clean code. I have no doubt that the principles I learned from him will continue to guide me well throughout my career.

I would also like to thank Titus Neupert, for reviewing this thesis and acting as a co-examiner.

A special thank you goes to my long-time office mates, Mario Könz and Donjan Rodic. Most importantly, they are great friends who made working in our office a fun experience. They also never hesitated to discuss some obscure part of my code or algorithms in great detail. I can not count the number of times my work has benefited from their inputs.

With QuanSheng Wu and Georg Winkler I shared many insightful discussions, since we formed the topological materials subgroup under the supervision of Alexey. I feel privileged to have worked with such talented and enthusiastic colleagues. Georg was also my “conference buddy” on many occasions, and I have fond memories of exploring the surroundings of each venue with him.

While working with the AiiDA framework, I visited the group of Nicola Marzari for several coding sprints. I would like to thank them for a very productive collaboration and welcoming atmosphere. In particular I would like to thank all the people who worked with me on the AiiDA code, namely Sebastiaan, Giovanni, Martin, Leonid, Antimo, Tiziano, Snehal, Spyros, Leopold, Fernando, and Rico.

Further I would like to thank all members of the computational physics group who I had the pleasure of interacting with. In no particular order, these are Thomas, Damian, Michele, Juan, Andreas, Guglielmo, Giuseppe, Kiryl, Ethan,

---

Lei, Bettina, Jana, Mauro, Hiroshi, Tao, Ilia, Alex, Slava, Vojta, Daniel, and Natalie.

Zu guter Letzt möchte ich mich bei meiner Familie bedanken, bei meinen Eltern Klara und Markus, bei Valentin, Linda und Matteo, und bei Martina, Andi und Étienne. Ihnen verdanke ich die Perspektive für die wichtigen Dinge ausserhalb der Physik. Spezieller Dank gebührt Laura, die mich zu jeder Zeit unterstützt hat, und mich in Höhen wie auch in Tiefen begleitet.

# Curriculum Vitae

## Personal data

NAME	Dominik André Gresch
DATE OF BIRTH	24.10.1991
CITIZEN OF	Freienbach (SZ)
NATIONALITY	Switzerland

## Education

AUG 2015 – <i>Current</i>	DOCTOR OF SCIENCES Computational Physics ETH Zurich, Switzerland
SEP 2013 – AUG 2015	MASTER OF SCIENCE MSc. in Physics, <i>summa cum laude</i> ETH Zurich, Switzerland
SEP 2009 – AUG 2012	BACHELOR OF SCIENCE BSc. in Physics ETH Zurich, Switzerland
SEP 2003 – JUN 2009	MATURA High School Diploma Stiftsschule Einsiedeln, Switzerland

---

## Employment

AUG 2015 – JAN 2019	RESEARCH AND TEACHING ASSISTANT Institute for Theoretical Physics ETH Zurich, Switzerland
SEP 2013 – MAR 2014	TEACHING ASSISTANT IN MATHEMATICS Numerical Methods and Linear Algebra ETH Zurich, Switzerland
OCT 2012 – AUG 2013	SOLDIER Mandatory Military Service Swiss Air Force
SEP 2012	SUBSTITUTE MATHEMATICS TEACHER Kantonsschule Olten
SEP 2011 – MAR 2012	TEACHING ASSISTANT IN MATHEMATICS Numerical Methods, Methods in Mathematical Physics, and Linear Algebra ETH Zurich, Switzerland

## Awards

2016	ETH MEDAL For the Master Thesis <i>Identifying Topological States in Matter</i>
JUL 2009	HONOURABLE MENTION International Physics Olympiad
MAR 2009	GOLD MEDAL Swiss Physics Olympiad

## Memberships

2009 – <i>Current</i>	PHYSICS OLYMPIAD (SWITZERLAND) Secretary, since 2015 Head Organizer Zurich, 2014 – 2018
-----------------------	---