Electronic Structure and Transport in Solids from First Principles

by

Jamal Ibrahim Mustafa

A dissertation submitted in partial satisfaction of the requirements for the degree of Doctor of Philosophy in Physics in the Graduate Division of the University of California, Berkeley

Committee in charge:
Professor Steven G. Louie, Chair
Professor Michael F. Crommie
Professor Mark D. Asta

Summer 2016
Electronic Structure and Transport in Solids from First Principles

Copyright 2016
by
Jamal Ibrahim Mustafa
Abstract

Electronic Structure and Transport in Solids from First Principles

by

Jamal Ibrahim Mustafa

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Steven G. Louie, Chair

The focus of this dissertation is the determination of the electronic structure and transport properties of solids. We first review some of the theory and computational methodology used in the calculation of electronic structure and materials properties. Throughout the dissertation, we make extensive use of state-of-the-art software packages that implement density functional theory, density functional perturbation theory, and the GW approximation, in addition to specialized methods for interpolating matrix elements for extremely accurate results. The first application of the computational framework introduced is the determination of band offsets in semiconductor heterojunctions using a theory of quantum dipoles at the interface. This method is applied to the case of heterojunction formed between a new metastable phase of silicon, with a rhombohedral structure, and cubic silicon. Next, we introduce a novel method for the construction of localized Wannier functions, which we have named the optimized projection functions method (OPFM). We illustrate the method on a variety of systems and find that it can reliably construct localized Wannier functions with minimal user intervention. We further develop the OPFM to investigate a class of materials called topological insulators, which are insulating in the bulk but have conductive surface states. These properties are a result of a nontrivial topology in their band structure, which has interesting effects on the character of the Wannier functions. In the last sections of the main text, the noble metals are studied in great detail, including their electronic properties and carrier dynamics. In particular, we investigate the Fermi surface properties of the noble metals, specifically electron-phonon scattering lifetimes, and subsequently the transport properties determined by carriers on the Fermi surface. To achieve this, a novel sampling technique is developed, with wide applicability to transport calculations. Additionally, the generation and transport of hot carriers is studied extensively. The distribution of hot carriers generated from the decay of plasmons is explored over a range of energy, and the transport properties, particularly the lifetimes and mean-free-paths, of the hot carriers are determined. Lastly, we present appendices detailing the implementation of the algorithms developed in this work, along with a useful derivation of the electron-plasmon matrix elements.
For family and friends.
# Table of Contents

List of Figures v

List of Tables viii

I Introduction and Methods 1

1 Introduction 2

2 Methods 4

2.1 The many-body Hamiltonian 4

2.2 Density functional theory 5

2.2.1 Hohenberg-Kohn Theorems 6

2.2.2 Kohn-Sham Formulation 6

2.2.3 Symmetry considerations 7

2.3 Planewave–Pseudopotential Approach 9

2.3.1 Planewave basis 9

2.3.2 Pseudopotentials 9

2.4 Density functional perturbation theory 10

2.5 Quasiparticle properties 11

2.5.1 Quasiparticle energies 13

2.5.2 Quasiparticle lifetimes 15

2.6 Transport Properties 16

2.6.1 Relaxation time approximation 17

2.6.2 Transport coefficients 18

2.7 Wannier functions 19

2.7.1 Localization 19

II Research 20

3 Band offsets in c-Si/Si-XII heterojunctions 21

3.1 Introduction and motivation 21
3.2 Band alignment model ........................................... 21
3.3 Computational method ........................................ 23
3.4 Results .............................................................. 24
   3.4.1 Charge neutrality levels .................................... 24
   3.4.2 Convergence .................................................. 24
   3.4.3 Band alignment ............................................... 25
3.5 Summary ............................................................ 26

4 Automated construction of localized Wannier functions 29
4.1 Introduction and motivation .................................... 29
4.2 Standard approach .............................................. 31
   4.2.1 First step ..................................................... 31
   4.2.2 Second step .................................................. 33
4.3 Alternative approach .......................................... 33
   4.3.1 Comparison to the standard approach .................. 35
   4.3.2 Simplifications .............................................. 35
4.4 Illustration of our approach .................................. 38
   4.4.1 Insight gained from the matrix $W$ ....................... 39
4.5 Summary ............................................................ 42

5 Automated construction of localized Wannier functions for bands with nontrivial topology 44
5.1 Introduction and motivation .................................... 44
   5.1.1 Constructing Wannier functions from a guess .......... 45
5.2 Our approach, the OPFM ....................................... 47
   5.2.1 Selecting the set $\{h\}$ ................................... 48
   5.2.2 Finding the matrix $W$ ..................................... 48
5.3 Examples .......................................................... 49
   5.3.1 Two-dimensional model .................................... 49
   5.3.2 Three-dimensional model .................................. 55
   5.3.3 Density functional theory, Bi$_2$Se$_3$ .................... 58
5.4 Outlook ............................................................. 59

6 Fermi surface properties of noble metals 60
6.1 Introduction ....................................................... 60
6.2 Computational Methods ....................................... 61
6.3 Results ............................................................ 63
6.4 Discussion and conclusion .................................... 69

7 Hot carrier generation and transport in noble metals 70
7.1 Introduction ....................................................... 70
7.2 Theory of SPP-electron coupling .............................. 72
7.3 Computational Methods ........................................... 74
7.4 Results ................................................................. 75
  7.4.1 Hot carrier generation ....................................... 75
  7.4.2 Energy distribution of hot carriers ......................... 77
  7.4.3 Hot carrier mean-free-paths ............................... 81
7.5 Discussion and conclusion ......................................... 82

III  Appendices ......................................................... 85

A  Implementation of the optimized projection functions method 86
  A.1 Spread functional ............................................... 86
  A.2 Codiagonalization algorithm .................................. 87
    A.2.1 Extensions .................................................. 89

B  Tetrahedron method for Fermi surface sampling ................. 92

C  Derivation of the SPP-electron coupling matrix element ........ 95

Bibliography .......................................................... 100
List of Figures

2.1 Sketch of the spectral function obtained within the \( GW \) approximation. In addition to the prominent quasiparticle peak, there are also satellite structures that can be related to additional excitations, such as plasmons. 13

3.1 Crystal structures of rhombohedral silicon (a) and cubic silicon (b). 22

3.2 \( E_{\text{CNL}} \) for c-Si along the lattice vector \( \mathbf{R} = m \mathbf{a}_1 = m_{\frac{a}{2}} (\hat{y} + \hat{z}) \). Also plotted for comparison, the data from Ref. [121]. The charge neutrality level is referenced to the top of the valence band. 25

3.3 \( E_{\text{CNL}} \) for Si-XII for a variety of high symmetry directions. Each color represents a different lattice vector, given in units of the primitive lattice vectors. Notice, that for the various directions \( \mathbf{R} \), the value of \( E_{\text{CNL}} \) is converging to a narrow range with the length of \( \mathbf{R} \) (\( m \) value). The charge neutrality level is referenced to the top of the valence band. 26

3.4 Convergence of \( E_{\text{CNL}} \), for c-Si, along \( \mathbf{R} = m_{\frac{a}{2}} (\hat{y} + \hat{z}) \) with respect to the number of conduction bands included in Eq. (3.2). Each curve is a different \( m \) value. For this test, a 24x24x24 k-point grid was used. 27

3.5 Convergence of \( E_{\text{CNL}} \), for c-Si, along \( \mathbf{R} = m_{\frac{a}{2}} (\hat{y} + \hat{z}) \) with respect to size of the k-point grid used in evaluating the Green’s functions in Eq. (3.2). Each curve is a different \( m \) value. For this test, we used 40 conduction bands. 28

3.6 Band offsets in a c-Si/Si-XII heterojunction as estimated by aligning the charge neutrality levels. 28

4.1 Ratio of \( \Omega_{\text{OPF}} \) and \( \Omega_{\text{GM}} \) as a function of Lagrange multiplier \( \lambda \) on a logarithmic scale. 40

4.2 Maximally localized Wannier functions of the two topmost valence bands in \( \text{LaMnO}_3 \). Isosurfaces of the WFs are shown on the left, looking down along the \( c \) axis. The large green dots are La, medium purple dots are Mn, and small red dots are O. On the right, we show contour plots of the Wannier functions in the plane perpendicular to the \( c \) axis, cutting through the Mn atom. 43
5.1 Band structures of the Kane-Mele model in the normal phase (top) and topological phase (bottom). The part of the bands colored red correspond to $B$-orbital character, while blue corresponds to $A$-orbital character. The thickness of the line corresponds to the spin component, with thicker indicating mostly spin-up, and thinner indicating spin-down. The zero of energy is set to the middle of the gap. ................................................................. 51

5.2 Histograms of the square moduli of the elements of $s^{(k)} - I$ for guess orbitals from Ref. [115] (top panel) and two sets $\{h\}$ within the OPFM (middle and bottom panel). ................................................................. 54

5.3 Plot of the Wannier functions obtained via the OPFM on set $\{0, 1^{st}, 2^{nd}\}$, for the two occupied bands of the Kane-Mele model in the normal insulator phase (top) and the topological insulator phase (bottom). The crosses ($\times$) indicate the $A$ sites, while the plus signs (+) indicate the $B$ sites. Red circles correspond to a component that is positive and blue circles correspond to negative. The size (area) of the circle is proportional to the magnitude of the component $\alpha_1, \alpha_2, \beta_1, \beta_2$. Here, the WFs plotted are for a Kane-Mele model with $\lambda_v = 1$ and $\lambda_R = 0$. . . 56

5.4 Band structure of Bi$_2$Se$_3$. The part of the bands colored red correspond to Se $p$ character, while blue corresponds to Bi $p$ character. The zero of energy is set to the middle of the gap. ................................................................. 59

6.1 Electron-phonon relaxation times (top panel), orbital character (middle panel), and band velocities (bottom panel) of states on the Fermi surface of Cu, Ag, and Au. .................................................................................. 62

6.2 Electron-phonon coupling matrix elements for an initial electronic state of predominantly $sp$ character (top panel) and an initial state of predominantly $d$ character (bottom panel). Shown are the $e$-$ph$ coupling matrix elements (summed over phonon polarizations) versus the character of the final electronic state involved in the $e$-$ph$ scattering process. For each initial state, the average coupling matrix element is also given. ................................................................. 65

6.3 Convergence of the resistivity computed with a uniform $k$-grid (dashed lines) and using Fermi surface importance sampling (solid line with markers). The dotted lines indicate the converged values. ................................................................. 68

7.1 Plasmon dispersion curves. The upper branch describes a bulk plasmon for small wavevector, while the lower branch describes a SPP for larger wavevectors. . . . 70
7.2 Schematic of a metal-dielectric interface under illumination (a), showing the decay of a SPP to a pair of hot electron and hole. The hot carriers are scattered and lose energy through phonon emission processes. The mean free path corresponds to the length of the black arrows, and the excited volume within the metal is shaded in gray. The FM-JDOS (b) of Au (left panel) and Ag (right panel) as a function of SPP energy, shown for several values of $q_p$ along the $\Gamma - L$ direction, up to a value of $q_p = 0.5L$. The SPP wavevector is expressed in units of the $L$ point in the BZ, where $L = (\pi/a)(1, 1, 1)$, with $a$ the lattice constant. The FM-JDOS for photons and SPPs with vanishingly small wavevector corresponds to the $q_p = 0$ case (curve labeled as $\Gamma$). The dashed lines indicate the value of the FM-JDOS low-energy tail for large wavevectors.

7.3 Energy distribution of hot electrons (red) and hot holes (blue) in Au (a) and Ag (b) generated by SPP decay. Each plot refers to a specific SPP energy and wavevector. The right panel shows the GW bandstructures of Au and Ag, respectively in (a) and (b), along with a schematic of the momentum and energy conserving excitation of a hot electron-hole pair resulting from the decay of a SPP with wavevector $q_p$ and energy $E_p$. The shaded area indicates the energy range of the $d$ states. The energy difference between the Fermi energy and the top of the $d$ bands corresponds to the interband transition threshold. The Fermi energy is the zero of the energy axis.

7.4 Relaxation time and scattering rates for HCs in Au (a) and Ag (b) within 5 eV of the Fermi energy. The top panels show the total relaxation time, accounting for both the $e$-$e$ and the $e$-$ph$ interactions. The bottom panels show the scattering rate, expressed as the imaginary part of the self-energy, $\text{Im}\Sigma$, for the $e$-$e$ and $e$-$ph$ interactions. $E_F$ is the zero of the energy axis, as shown by the dashed line. Electrons (holes) possess positive (negative) energies, and the shaded area indicates the energy window spanned by the $d$ states.

7.5 The hot carrier MFPs along the [100], [110], and [111] crystallographic directions in Au (a) and Ag (b). The MFPs for the $d$ states are expanded in the right panels.

7.6 The graph shows two regimes, for intraband $(s-s)$ and interband $(d-s)$ transitions induced by SPPs with energy $E_p < E_{\text{int}}$ or $E_p > E_{\text{int}}$, respectively. The regime with $E_p < E_{\text{int}}$ offers a trade-off between optimal HC generation, moderate HC energies up to $\sim 2$ eV, and superior MFPs and relaxation times compared to more energetic HCs. The dashed line indicates the Fermi energy.

7.7 Plasmon modes in Ag (a) and Au (b) for a variety of different dielectrics.

A.1 The red lines indicate the rows modified by left-multiplication by $R_d^\dagger$. The blue lines indicate the columns modified by right-multiplication by $R$.

B.1 Tetrahedra with a constant energy surface cutting through in each of the cases given in Table B.1.
List of Tables

4.1 Total spread $\Omega^{\text{OPF}}$ computed within our approach and at the global minimum $\Omega^{\text{GM}}$ for all seven materials studied. We also give diagonal and off-diagonal components of spread in each case ($\Omega_\text{D}$ and $\Omega_\text{OD}$). The spreads $\Omega^{\text{OPF}}$ are obtained using the optimal value of $\lambda$ (see Fig. 4.1). The units for the spreads are Å$^2$. In the case of Cr$_2$O$_3$ we wannierize only the topmost 12 bands below the Fermi level, and in the case of LaMnO$_3$ we wannierize the top 2 spin-up bands. In all other cases we wannierize all valence bands. .................................................. 41

5.1 Results of the OPFM applied to the examples in Sec. 5.3. The first column gives the square moduli of $s^{(k)} - I_N$ averaged over k-points and matrix elements. The second column lists the minimal singular value of $s^{(k)}$ over all k-points. The third column show the spread $\Omega^0$ after Löwdin procedure while the fourth column shows the spread $\Omega^{\text{GM}}$ when Löwdin procedure is followed up by the Marzari-Vanderbilt [77] procedure. For three-dimensional cases, the fifth column gives the value for the Chern-Simons $\theta$ term. See text for a description of the sets $\{h\}$ used in the OPFM. ................................................................. 53

6.1 Comparison of relaxation times and velocities computed by averaging on the Fermi surface (values in brackets), versus the same quantities obtained within Drude theory (subscript D). The Fermi surface averaged relaxation times are computed using GW bandstructures, while the band velocities are given for both the DFT and GW bandstructures for comparison. The Drude values are obtained using Eq. (6.3) with resistivity experimental data [19, 57] from single-crystal samples. ................................................................. 66

6.2 Resistivity values (units of $\mu\Omega$ cm) at 300 K computed using GW band velocities and different approximations for the relaxation times. $\tau_{nk}$ are band and k-dependent from ab initio, $\langle \tau_{nk} \rangle$ are Fermi-surface averaged from $\tau_{nk}$, and $\tau_\text{D}$ are empirical Drude relaxation times. Also shown are the computed values from DFT velocities and empirical Drude relaxation times. Experimental data for single crystal (SC) and polycrystalline (PC) samples are shown for comparison. 67
7.1 Square modulus of the coupling matrix element, $|g|^2$, and Feynman diagram for the self-energy of the boson in the lowest-order interaction for the three bosons considered here — plasmons, photons, and SPPs — with electrons in a material. These interactions share a common diagram (right column), consisting of a polarization bubble $\chi$ multiplied by two coupling vertices $g$. The wiggly lines on the two sides of the self-energy loop $g^2\chi$ indicate the incoming and outgoing bosons.

B.1 Various cases of a plane of constant energy cutting through a tetrahedron. Depending on the energy of the isosurface relative to energies at the corners of the tetrahedron, the region of intersection can either be a triangle or a quadrilateral.

B.2 Vertices of the $k$ points that make up the facets in Figure B.1.
Acknowledgments

The journey to obtain my PhD has been a long and difficult one, a worthwhile one. Any successes I have had along the way must be attributed, in part, to my advisors, mentors, and fellow group members. My advisor Steven G. Louie has provided me the opportunity to pursue my own projects and helped me to conduct exciting research.

In addition, two people in particular have contributed greatly to my PhD research and overall graduate student experience. Sinisa Coh, my office mate, has enriched my PhD experience to the point that I consider him not only a colleague, but also a friend. We share an affinity for Wannier functions, and this has led to some very good collaboration and lively conversation. His pragmatic approach to problem solving, where sometimes instead of trying to figure it all out in your head you just go and run the calculations (he calls it the “caveman” approach) is one of the more valuable things I have learned. This approach was the driving force behind the work developing an automated procedure for constructing localized Wannier functions. Much of the development was an enthusiastic back-and-forth: I write some code, do some tests, have some problems, and Sinisa telling me “try this” and “try that” or “use this” code.

Another colleague and friend is Marco Bernardi. Soon after Marco joined the group we began collaborating on the noble metals. Marco is a master of big picture thinking and understanding what problems are interesting and worth exploring. His work ethic and the way he pounds away at a problem is something that has rubbed off on me quite a bit and has helped to really boost productivity.

One of the most valuable aspects of the Louie group is the role that the group members and group meetings play. The members of the group are always the first place to go with your problems and there is always a sense that work is a team effort. During my time at UC Berkeley I had the pleasure of discussing research with some of the brightest minds in our field. I have benefited greatly for working closely with Derek Wayne Vigil-Fowler and Felipe Homrich da Jornada developing our in-house code, BerkeleyGW.

The support I have received has been twofold, those who have helped me at work, and those that have helped me at home. I could not have completed this journey without the help of my wife Tatiana and support of my family as a whole. My wife and I have been together for the entirety of the program and she deserves quite a bit of credit for the support role she has played through this grueling experience. I would also like to mention the fact that during the tail end of my time at UC Berkeley, Tatiana and I welcomed our first child, Dezmen, into the world. In fact, he has taken part in the writing of this document, helping me with some of the typing, although his contributions (banging on the keyboard) needed quite a bit of editing.
Part I

Introduction and Methods
Chapter 1

Introduction

Computational physics has, in many ways, become a third branch of physics, an intermediate branch between theoretical and experimental physics. In addition to its utility in understanding experimental results through simulation, the predictive power of ab initio (first-principles) calculations has played a large role in guiding and interpreting experiments and the discovery of new materials. As the field sets its sights on larger and more complex systems, it becomes necessary to develop better algorithms and more efficient implementations to take advantage of the massive increase in available computing power. Currently, there are many ab initio methods available to the computational physicist: density functional theory (DFT), which has become the standard tool predicting ground state properties; density functional perturbation theory (DFPT) has provided for routine investigation of lattice dynamical properties; the $GW$ approximation has enabled computationally tractable many-body perturbation theory calculations to determine quasiparticle properties. Furthermore, semiclassical approaches to materials properties, such as transport coefficients, coupled with ab initio derived parameters, provide an accurate and predictive framework for materials design.

Chapter 2 is devoted to the review of the theory and computational methods for the determination of electronic structure and transport properties in solids. The chapter covers various approximations that are employed in order to obtain tractable equations from the full quantum mechanical theory of interacting electrons and nuclei. Here we introduce the relevant notation, quantities, and equations used throughout the text.

In Chapter 3 we investigate a heterojunction formed by different phases of silicon. Silicon has a rich phase diagram with a multitude of phases existing over a wide range of pressures and temperatures, in addition to the common cubic silicon (c-Si) phase. One such phase, Si-XII, was first observed less than 2 decades ago in diamond anvil experiments, and more recently as a product of nanoindentation. In some of these latter experiments, I-V measurements were performed to characterize the c-Si/Si-XII interface that results when Si-XII is formed in cubic silicon substrates. In this chapter we describe calculations of the band offsets in c-Si/Si-XII heterojunctions. We find that the heterojunction is of Type I and that the band offsets are estimated to be $\Delta E_v = 0.3$ eV and $\Delta E_c = 0.5$ eV for the valence bands.
and conduction bands, respectively.

In Chapters 4 and 5, we develop a method for the automated construction of maximally localized Wannier functions. Maximally localized Wannier functions are widely used in electronic structure theory for analyses of bonding, electric polarization, orbital magnetization, and for interpolation. The state of the art method for their construction is based on the method of Marzari and Vanderbilt. One of the practical difficulties of this method is guessing functions (initial projections) that approximate the final Wannier functions. Here we present an approach based on optimized projection functions that can construct maximally localized Wannier functions without a guess. We describe and demonstrate this approach on several realistic examples. In Chapter 5 we show that an optimized projection functions method can automatically construct maximally localized Wannier functions even for bands with nontrivial topology. We demonstrate this method on a tight-binding model of a two-dimensional $\mathbb{Z}_2$ topological insulator, on a three-dimensional strong $\mathbb{Z}_2$ topological insulator, as well as on first-principles density functional theory calculated valence states of $\text{Bi}_2\text{Se}_3$. In all cases, the resulting Wannier functions contain large imaginary components and are more extended than those in the topologically trivial phase.

In Chapters 6 and 7, the transport properties and hot carrier generation and dynamics is studied in great detail for the noble metals. Relaxation times employed to study electron transport in metals are typically taken to be constants and obtained empirically. Here, we use fully ab initio calculations to compute the electron-phonon relaxation times of Cu, Ag, and Au, and find that they vary significantly on the Fermi surface, spanning $\sim 15\text{--}45$ fs values that are correlated with the Fermi surface topology. We compute room temperature resistivities in excellent agreement with experiment by combining GW quasiparticle bandstructures, Wannier-interpolated band velocities, and ab initio relaxation times. We introduce an importance sampling scheme to speed up the convergence of resistivity and transport calculations. Hot carriers generated by surface plasmon polaritons (SPPs) in noble metals are promising for application in optoelectronics, plasmonics, and renewable energy. However, existing models fail to explain key quantitative details of SPP-to-hot carrier conversion experiments. Here we develop a quantum mechanical framework and apply first-principles calculations to study the energy distribution and scattering processes of hot carriers generated by SPPs in Au and Ag. We find that the relative positions of the $s$ and $d$ bands of noble metals regulate the energy distribution and mean free path of the hot carriers, and that the electron-phonon interaction controls hot carrier energy loss and transport. Our results prescribe optimal conditions for hot carrier generation and extraction, and invalidate previously employed free-electron-like models. Our work combines density functional theory, GW, and electron-phonon calculations to provide microscopic insight into hot carrier generation and ultrafast dynamics in noble metals.
Chapter 2
Methods

The following sections give brief overviews of some of the theory and methods in the field of condensed matter physics, in particular those used in computational materials science. Since this subject has been developed over many years, the established techniques need only a brief introduction, while further details can be obtained from the extensive literature. For this reason, apart from introducing some equations and expressions and notation for reference throughout the text, much of the details will be deferred to standard references [3, 55, 76, 145, 71, 70, 113].

In the first section of this chapter, we start by describing the central problem in condensed matter physics: solving the many-body Hamiltonian for a system of interacting electrons and nuclei. The following two sections introduce some practical techniques for obtaining approximate solutions to the many-body Hamiltonian. In Sec. 2.4, a theoretical and computational framework to describe lattice dynamical properties is described. Section 2.5 gives a bit more in depth account of quasiparticle approach to electronic structure. Lastly, we give a review of transport in materials within the Boltzmann formalism, specifically the Boltzmann transport equation, the relaxation time approximation, and expressions for various transport coefficients.

2.1 The many-body Hamiltonian

The purpose of this section is to set the stage for the rest of the dissertation. For all intents and purposes, condensed matter physics, and any microscopic description of materials properties, is ultimately concerned with the many-body Hamiltonian in the nonrelativistic limit [76]

\[
\hat{H} = -\sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 + \frac{1}{2} \sum_{i\neq j} \frac{e^2}{|r_i - r_j|} + \sum_{i,I} \frac{Z_I e^2}{|r_i - R_I|} + \frac{1}{2} \sum_{I\neq J} \frac{Z_I Z_J e^2}{|R_I - R_J|}, \quad (2.1)
\]

where \(m_e\) is the electron mass, the \(M_I\) and \(Z_I\) are the mass and atomic number of the \(I\)th ion, \(e\) is the unit of electric charge, and \(\hbar\) is the reduced Planck constant. The many-body
wavefunction describing the electrons and nuclei

$$\Psi = \Psi(r_1, \ldots, r_N, R_I, \ldots, R_M)$$  \hspace{1cm} (2.2)$$

is a function of all the electron and nuclear coordinates \(\{r_i\}\) and \(\{R_I\}\). For materials with heavier atoms, the appropriate relativistic terms such as the spin-orbit interaction need to be included. Since there are \(\sim 10^{23}\) electrons and nuclei in a small sample of material, it is obviously impossible to solve for the many-body wave function directly; approximations must be made in order to make any progress.

The first, and almost universal, approximation to simplify this problem is the Born-Oppenheimer approximation, which separates the many-body wave function into a product of the electronic and nuclear parts

$$\Psi = \Psi_{el}(\{r_i\}) \times \Psi_{nuc}(\{R_I\}).$$  \hspace{1cm} (2.3)$$

This is possible due to fact the the nuclear masses are typically much larger than the electron mass \((M_I \gg m_e)\) so that the electrons basically see the nuclear configuration as static and their respective motion can be decoupled. In this way, the Hamiltonian describing the electronic system can be expressed as (ignoring the kinetic energy of the ions)

$$\hat{H} = -\sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 + \sum_i V_{ext}(r_i) + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|r_i - r_j|} + E_{ion}(\{R_I\}),$$  \hspace{1cm} (2.4)$$

where the external potential seen by the electrons in the system is that of the fixed ions

$$V_{ext}(r) = V_{ion}(r) = \sum_I \frac{Z_I e^2}{|r - R_I|}$$  \hspace{1cm} (2.5)$$

and

$$E_{ion}(\{R\}) = \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J e^2}{|R_I - R_J|}$$  \hspace{1cm} (2.6)$$

is just the energy of the ionic configuration, which depends parametrically on the nuclear coordinates. In the following section, a further approximation is discussed that allows us to deal with the electron-electron Coulomb term by considering electrons moving in the average, or mean-field, of the rest of the electrons rather than explicitly dealing with the two-particle interactions.

### 2.2 Density functional theory

Density functional theory provides a method for finding the ground state of the many-body Hamiltonian. While the formalism, in principle, gives the exact ground state density of the many-body system, necessary approximations to the complex electron-electron interactions must be made. In the following subsections, the foundations for DFT will be reviewed and a practical approach to DFT calculations will be outlined.
2.2.1 Hohenberg-Kohn Theorems

Density functional theory is founded on the seminal work of Hohenberg and Kohn [45], which establishes the theory on the basis of two theorems, the Hohenberg-Kohn (HK) theorems. The first (HK1) proves that there is a one-to-one correspondence between the external potential acting on a system and the ground state charge density. The second (HK2) proves that the energy of the electronic system can be written as a functional of the density $n(r)$

$$E[n] = F[n] + \int d\mathbf{r} V_{\text{ext}}(\mathbf{r}) n(\mathbf{r}), \quad (2.7)$$

where $F[n]$ is a universal functional (independent of the external potential), and that the many-body ground state density minimizes the energy $E[n]$. These theorems were originally shown to be true for a nondegenerate ground state, but have since been generalized to the degenerate case as well [61]. While the HK2 theorem proves that the universal functional $F[n]$ exists, the proof is not constructive and so does not provide an explicit form of the functional. In order to make any progress, an approximation for $F[n]$ must be derived.

2.2.2 Kohn-Sham Formulation

The follow-up work by Kohn and Sham [58] provides a practical route to finding the ground state density as follows: instead of minimizing $E[n]$ with respect to (wrt) to the density in order to find the ground state, consider an auxiliary system of noninteracting electrons described by the Hamiltonian

$$\hat{H}_{\text{KS}} = -\frac{\hbar^2}{2m_e} \nabla^2 + V_{\text{KS}}(\mathbf{r}). \quad (2.8)$$

Eigenstates $\phi_i(\mathbf{r})$ of $\hat{H}_{\text{KS}}$ satisfy

$$\left[-\frac{\hbar^2}{2m_e} \nabla^2 + V_{\text{KS}}(\mathbf{r})\right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}), \quad (2.9)$$

with the $N$ lowest-lying solutions having a corresponding ground state density given by

$$n(\mathbf{r}) = \sum_{i=1}^{N} |\phi_i(\mathbf{r})|^2. \quad (2.10)$$

Taking for granted the Kohn-Sham ansatz, that the ground state density of the interacting system can be expressed in terms of the Kohn-Sham orbitals, we can minimize the energy functional $E[n]$ wrt to the $\phi_i$, rather than the density. In terms of the noninteracting system, we can write the universal functional,

$$F[n] = T_{\text{KS}}[\{\phi_i\}] + E_{\text{H}}[n] + E_{\text{xc}}[n], \quad (2.11)$$
where kinetic energy $T_{KS}$ and Hartree energy $E_H$ are given by

$$T_{KS}\{\phi_i\} = -\frac{1}{2} \sum_i \langle \phi_i | \nabla^2 | \phi_i \rangle,$$

$$E_H[n] = \frac{1}{2} \int d\mathbf{r} d\mathbf{r}' n(\mathbf{r}) n(\mathbf{r}') \frac{|\mathbf{r} - \mathbf{r}'|}{|\mathbf{r} - \mathbf{r}'|}. \tag{2.12}$$

The exchange-correlation term $E_{xc}[n]$ contains all the parts of the universal functional that are not captured by the noninteracting kinetic energy and Hartree energy. Defining the Hartree potential $V_H$ and exchange-correlation potential $V_{xc}$ as

$$V_H(\mathbf{r}) = \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{2.13}$$

$$V_{xc}(\mathbf{r}) = \frac{\delta E_{xc}}{\delta n(\mathbf{r})}, \tag{2.14}$$

we arrive at the Kohn-Sham equations by minimizing Eq. (2.11):

$$\left[ -\frac{1}{2} \nabla^2 + V_{ion}(\mathbf{r}) + \int d\mathbf{r}' \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + V_{xc}(\mathbf{r}) \right] \phi_i(\mathbf{r}) = \epsilon_i \phi_i(\mathbf{r}). \tag{2.15}$$

Within the KS formalism, the noninteracting single-particle kinetic energy and long-range Hartree interaction has been separated out, and so one is better able to make approximations to the exchange-correlation potential. The most common approximation is the local density approximation (LDA), which takes the exchange-correlation (XC) potential $V_{xc}$ to be evaluated at the local density, taking on the electron gas value [17]. Other XC functionals have been developed, including the generalized gradient approximation (GGA) [94], which includes the gradient of the density. Additionally, the formalism has been extended to include electron spin [6] and also relativistic effects [102].

It is important to note that the energy eigenvalues that show up in Eq. (2.15) cannot rigorously be interpreted as quasiparticle energies; rather, they are merely Lagrange multipliers that enforce orthonormality among the KS orbitals while minimizing the energy. Interpreting the KS eigenvalues as quasiparticle energies leads to the well known “band gap problem”, where the calculated gaps of semiconductors and insulators can be underestimated by up to a factor of $\sim 2$. This failure of DFT to predict band gaps should not come as a surprise, since it is a ground state theory.

### 2.2.3 Symmetry considerations

While the Born-Oppenheimer approximation has greatly simplified Eq. (2.1), allowing us to consider the electronic system for a fixed configuration of the ions [Eq. (2.4)], there is still the issue of sheer number of electrons ($\sim 10^{23}$). In the case of crystals, the material is made
up of a lattice of repeating unit cells so that \( V_{\text{ion}}(\mathbf{r}) \) is periodic. The unit cell is defined by a set of lattice vectors \( \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\} \) and the periodicity of the potential can be expressed as
\[
V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r}),
\]
where \( \mathbf{R} \) is a translation vector of the form
\[
\mathbf{R} = n_1\mathbf{a}_1 + n_2\mathbf{a}_2 + n_3\mathbf{a}_3,
\]
and the \( n \)’s are integers.

In addition to the real space (direct) lattice, a crystal is also characterized by a reciprocal lattice defined by the vectors \( \mathbf{G} \) satisfying
\[
e^{i\mathbf{G} \cdot \mathbf{R}_n} = 1
\]
for all \( \mathbf{R}_n \). The set of all vectors \( \mathbf{G} \) can be expressed as
\[
\mathbf{G}_m = m_1\mathbf{b}_1 + m_2\mathbf{b}_2 + m_3\mathbf{b}_3
\]
in terms of the reciprocal lattice vectors \( \mathbf{b}_i \), which are defined through the relation
\[
\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}
\]
and can be explicitly constructed as
\[
\mathbf{b}_1 = \frac{2\pi}{\Omega_{\text{cell}}} \mathbf{a}_2 \times \mathbf{a}_3, \quad \mathbf{b}_2 = \frac{2\pi}{\Omega_{\text{cell}}} \mathbf{a}_3 \times \mathbf{a}_1, \quad \mathbf{b}_3 = \frac{2\pi}{\Omega_{\text{cell}}} \mathbf{a}_1 \times \mathbf{a}_2,
\]
where \( \Omega_{\text{cell}} \) is the volume of the primitive unit cell (\( \Omega_{\text{cell}} = |\mathbf{a}_1 \cdot \mathbf{a}_2 \times \mathbf{a}_3| \)).

For a single-particle Hamiltonian with a periodic potential term, Bloch’s theorem shows that the eigenstates must be of the form
\[
\psi_{nk}(\mathbf{r}) = e^{i\mathbf{k} \cdot \mathbf{r}} u_{nk}(\mathbf{r})
\]
where \( u_{nk}(\mathbf{r}) \) is periodic wrt to lattice translations \( \mathbf{R} \), and \( \mathbf{k} \) is a point in the Wigner-Seitz cell of the reciprocal lattice (the first Brilloin Zone). We adopt the convention that the Bloch functions are normalized to the primitive unit cell. In this way, we need only solve for the electronic wavefunctions in a single cell, with periodic boundary conditions.

In addition to translational symmetry, crystals also possess space group symmetries. The point group (rotational) symmetries associated with the space group reduces the part of the Brilloin zone (BZ) that must be considered. Only wavefunctions at \( \mathbf{k} \) points within this irreducible wedge must be calculated, leading to a further reduction in computational cost.
2.3 Planewave–Pseudopotential Approach

In this section we introduce the planewave-pseudopotential approach commonly used to solve the KS equations. Planewaves and pseudopotentials are discussed together because in many ways they are closely related. In the previous sections, the external potential is taken as the potential due to a periodic arrangement of nuclei. This periodicity naturally suggests the use of a planewave basis of the form $\exp(iG_m \cdot r)$, since these basis functions are also cell periodic. One drawback of the planewave basis is the large number of planewaves necessary to capture the large oscillations in the valence wavefunctions near the nuclei. Pseudopotentials replace the bare nuclei with a potential that gives nodeless (smoother) valence wavefunctions that can be represented by significantly fewer planewaves.

2.3.1 Planewave basis

In many ways, the planewave basis is the simplest basis one could consider for representing Bloch functions. Considering planewaves with wavevectors $G_m$ on the reciprocal lattice provides a complete basis to expand the Bloch functions, consistent with the translational symmetry in a crystal. By Bloch’s theorem, the crystal wavefunctions can be expressed in terms of planewaves as

$$\psi_{nk}(r) = \sum_G c_{nk}(G)e^{i(k+G)\cdot r},$$

(2.23)

where the $c_{nk}(G)$ are the expansion coefficients of the period part of the wavefunction $u_{nk}(r)$. In practice, for computation purposes, we only expand the wavefunctions with planewaves up to some $G_{\text{max}}$ corresponding to a kinetic energy cutoff

$$E_{\text{cut}} = \frac{\hbar^2 |G_{\text{max}}|^2}{2m_e}.$$  

(2.24)

2.3.2 Pseudopotentials

Pseudopotentials replace the bare ion with the nucleus together with the core electrons, while preserving the scattering properties of the ion. This is physically motivated by the fact that the core electrons are typically inert, and only the valence electrons participate in the chemical bonding. There is also a computational incentive, not only are there fewer electrons to consider, but the valence-electron wavefunctions are much smoother and can be represented with fewer planewaves.

Construction of a pseudopotential starts with the Schrödinger equation for the isolated atom. The radial part, $\chi_{nl}(r)$, of the wavefunction $\psi_{nlm}(r)$ for an electron in a spherically symmetric potential $V(r)$ satisfies the radial Schrödinger for a particular angular momentum state $l$, expressed in Hartree atomic units as

$$\left[ -\frac{1}{2}\frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + V(r) \right] \chi_{nl}(r) = \epsilon_{nl} \chi_{nl}(r).$$

(2.25)
The wavefunction \( \psi_{nlm}(r) = \chi_{nl}(r)Y_l^m(\theta, \phi) \), with \( Y_l^m \) the spherical harmonics. In practice, Eq. (2.25) is solved using density functional theory.

Norm-conserving pseudopotentials [42] are constructed so that the following conditions are satisfied (for each angular momentum channel \( l \)):

1. The valence eigenvalues of the pseudopotential are the same as the all-electron potential.
   \[ \epsilon_i^{PP} = \epsilon_i^{AE} \]

2. The valence wavefunctions agree with the all-electron wavefunctions beyond some chosen core radius.
   \[ \chi_i^{PP}(r) = \chi_i^{AE}(r) \text{ for } r > r_c \]

3. The total integrated charge within the core radius is the same.
   \[ \int_0^{r_c} dr \chi_i^{PP}(r) = \int_0^{r_c} dr \chi_i^{AE}(r) \]

These conditions guarantee that the scattering properties of the pseudopotential will be the same (for some range of energy around \( \epsilon_i^{AE} \)) as the all-electron potential. Pseudopotentials constructed in this way are "transferrable" in the sense that they will work in a variety of chemical environments.

While pseudopotentials are typically constructed by considering only the valence (lowest energy) states, for atoms that have \( d \) electrons, ignoring the \( s \) and \( p \) states with the same principle quantum number \( n \) generally leads to pseudopotentials with poor transferability. In addition, as will be discussed in Sec. 2.5, inclusion of the semicore states is necessary to accurately compute the self-energy, which includes electron exchange. The exchange term depends on the spatial overlap of the wavefunctions, which is large for states with the same \( n \), even though they may be well separated in terms of energy. For the noble metals (studied in Chapters 6 and 7), one must also consider the semicore \( s \) state along with the valence \( s \) state. This can be dealt with using a multi-reference pseudopotentials [90], which construct a potential for the angular momentum channel from multiple states in order that all the conditions for norm conserving pseudopotentials are satisfied.

### 2.4 Density functional perturbation theory

Up until this point, we have only considered the electronic system for a fixed configuration \( \{ R_I \} \) of ions. The total energy of a solid within the Born-Oppenheimer approximation can be determined as a function of the nuclear coordinates [i.e. via the terms \( V_{\text{ion}} \) and \( E_{\text{ion}}(\{ R_I \}) \) in Eq. (2.4)]. Lattice dynamical properties are computed within density function perturbation theory (DFPT), which computes derivatives of the total energy with respect to perturbations of the ions in the crystal lattice. First derivatives of the total energy with respect to the nuclear coordinates are obtained using the Hellman-Feynman theorem, which yields

\[
\frac{\partial E(\{ R \})}{\partial R_I} = \int d\mathbf{r} \frac{\partial V_{\text{ion}}(\mathbf{r})}{\partial R_I} n(\mathbf{r}) + \frac{\partial E_{\text{ion}}(\{ R \})}{\partial R_I},
\]
while the second derivative is obtained via Eq. (2.26), yielding
\[
\frac{\partial^2 E(\{R\})}{\partial R_I \partial R_J} = \int \mathrm{d}r \frac{\partial^2 V_{\text{ion}}(r)}{\partial R_I \partial R_J} n(r) + \int \mathrm{d}r \frac{\partial V_{\text{ion}}(r)}{\partial R_I} \frac{\partial n(r)}{\partial R_J} + \frac{\partial^2 E_{\text{ion}}(\{R\})}{\partial R_I \partial R_J}. \tag{2.27}
\]
Since the derivatives of the external potential \(V_{\text{ion}}(r)\) and the energy of the ionic configuration \(E_{\text{ion}}(\{R_I\})\) can be readily computed, all that is needed is the derivative of the electron density wrt the \(\{R_I\}\).

Within KS-DFT, the density is constructed from the KS orbitals, so the derivative is simply
\[
\partial R_I n(r) = 2 \Re \sum_i \phi_i^*(r) \partial R_I \phi_i(r), \tag{2.28}
\]
with the notation \(\partial R_I\) representing \(\partial / \partial R_I\). The derivative of the KS orbitals wrt variations of the \(\{R_I\}\) is computed using the following [along with Eq. (2.28)] interdependent equations:
\[
(H_{\text{KS}} - \epsilon_i \partial R_I) |\phi_i\rangle = - (\partial R_I V_{\text{KS}} - \partial R_I \epsilon_i) |\phi_i\rangle, \tag{2.29}
\]
\[
\partial R_I \epsilon_i = \langle \phi_i | \partial R_I V_{\text{KS}} | \phi_i \rangle, \tag{2.30}
\]
\[
\partial R_I V_{\text{KS}}(r) = \partial R_I V_{\text{ion}}(r) + \int \mathrm{d}r' \frac{\partial R_I n(r')}{|r - r'|} + \frac{\mathrm{d}V_{xc}(n)}{\mathrm{d}n} \partial R_I n(r). \tag{2.31}
\]

Once the second derivatives of the total energy have been determined by self-consistently solving Eqs. (2.28), (2.29), (2.4), and (2.31) for the change in the electron density, we can construct the dynamical matrix. The phonon normal modes and frequencies can be obtained by solving
\[
\det \left| \frac{1}{\sqrt{M_I M_J}} \frac{\partial^2 E(\{R\})}{\partial R_I \partial R_J} - \omega^2 \right| = 0. \tag{2.32}
\]

Another result of the DFPT calculation is the self-consistent perturbation potential \(\partial V_{\nu q}\) (here written in the eigenmode basis). Matrix elements of this perturbation potential give the electron-phonon coupling matrix elements
\[
g_{\nu mn}(k, q) = \langle mk + q | \partial V_{\nu q} | nk \rangle. \tag{2.33}
\]
The matix elements \(g_{\nu mn}(k, q)\) allow for the computation of a variety of properties involving the electron-phonon interaction, including, but not limited to: transport properties, thermal expansion, temperature dependence of the band gap, and superconductivity.

### 2.5 Quasiparticle properties

The quasiparticle properties of materials are typically computed with many-body perturbation theory (MBPT) [80, 71, 30]. Within this formalism, the single-particle Green’s
function \( G(\mathbf{r}, \mathbf{r}', \omega) \) describes the propagation of particle-like excitations in the many-body system. The poles of the Green’s function represent the energies of the single-particle-like excitations from the many-body ground state. The spectrum of quasiparticles is given by the spectral function \( A(\omega) = 1/\pi |\text{Im} \, G(\omega)| \). So, once \( G \) has been determined, basically all the properties of quasielectrons and quasiholes can be determined, in addition to expectation values of single-particle operators such as the density and total energy of the many-body system.

The Green’s function formalism employs the concept of self-energy in order to describe interactions. The interacting Green’s function \( G \) can be obtained from the noninteracting \( G_0 \) (ie., from DFT, or some other mean-field theory) via the Dyson equation for \( G \)

\[
G = G_0 + G_0 \Sigma G,
\]

which yields

\[
G^{-1} = G_0^{-1} - \Sigma.
\]

The noninteracting \( G_0 \) can be constructed from the single-particle Bloch functions according to

\[
G_0(\mathbf{r}, \mathbf{r}', \omega) = \sum_{nk} \frac{\psi_{nk}(\mathbf{r})\psi_{nk}(\mathbf{r}')}{\omega - \epsilon_{nk} + i\delta_{nk}}.
\]

Representing the Dyson equation [Eq. (2.35)] in state-space yields

\[
G_{nk}(\omega) = \frac{1}{\omega - \epsilon_{nk} - \Sigma_{nk}(\omega)}.
\]

So, with the mean-field wavefunctions and energies, the quasiparticle properties can be obtained after computing the self-energy.

The most common and successful application of the self-energy formalism has been realized via the \( GW \) approximation [44]

\[
\Sigma = iGW,
\]

which is obtained by expanding the self-energy \( \Sigma \) in powers of the screened Coulomb interaction \( W \), and taking the lowest-order term. The screened Coulomb interaction is obtained by screening the bare Coulomb potential \( v \) with the dielectric function of the material

\[
W = \frac{v}{\varepsilon}.
\]

We construct the dielectric function within the Adler-Wiser formulation [1, 136] through the polarizability in the random phase approximation (RPA)

\[
\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}; 0) = \sum_n^{\text{occ}} \sum_{n'}^{\text{unocc}} \sum_k M_{nn'}(\mathbf{k}, \mathbf{q}, \mathbf{G}) M^*_{n'n}(\mathbf{k}, \mathbf{q}, \mathbf{G}') \frac{E_{nk+q} - E_{n'k}}{E_{nk+q} - E_{n'k}}
\]
Much like the development of DFT, once Hybertsen and Louie [47] developed a workable implementation of using the GW approximation, it has become the state-of-the-art method for computing quasiparticle properties. In practice we perform so-called “one-shot”, or $G_0W_0$ calculations, where the Green’s function $G_0$ and screened interaction $W_0$ is computed using the single-particle wavefunctions from DFT in order to evaluate the self-energy. A more detailed treatment of the computational methodology employing the GW approximation can be found in Ref. [24].

The spectral function for a state $|n\mathbf{k}\rangle$ is given by

$$A_{nk}(\omega) = \frac{1}{\pi} \frac{|\text{Im } \Sigma_{nk}(\omega)|}{\left(\omega - \epsilon_{nk} - \text{Re } \Sigma_{nk}(\omega)\right)^2 + \left(\text{Im } \Sigma_{nk}\right)^2},$$

(2.42)

which describes a quasiparticle with a Lorentzian lineshape. The location of the peak gives the quasiparticle energy (see Sec. 2.5.1) and its width is related to the lifetime (see Sec. 2.5.2).

### 2.5.1 Quasiparticle energies

Quasiparticle energies are obtained by solving

$$E_{nk} = \epsilon_{nk} + \text{Re } \Sigma_{nk}(E_{nk}),$$

(2.43)
where the self-energy is evaluated at the quasiparticle energy \( E_{nk} \). For the purpose of computing the quasiparticle energies, the self-energy can be split into two terms: the screened exchange (SX) and Coulomb hole (CH) terms,

\[
\Sigma = \Sigma_{\text{SX}} + \Sigma_{\text{CH}}. \tag{2.44}
\]

Within the generalized plasmon-pole (GPP) approximation, the matrix elements of the screened exchange part of the self-energy can be expressed as

\[
\langle nk|\Sigma_{\text{SX}}(E)|n'k'\rangle = -\sum_{n''} \sum_{qGG'} M_{n'n}^*(k, -q, -G) M_{n'n'}(k, -q, -G') \times \left[ \delta_{GG'} + \Omega_{GG'}^2(q) \left( 1 - i \tan \phi_{GG'}(q) \right) \right] v(q + G'),
\]

and those for the Coulomb hole part are

\[
\langle nk|\Sigma_{\text{CH}}(E)|n'k'\rangle = \frac{1}{2} \sum_{n''} \sum_{qGG'} M_{n'n}^*(k, -q, -G) M_{n'n'}(k, -q, -G') \times \frac{\Omega_{GG'}^2(q) \left( 1 - i \tan \phi_{GG'}(q) \right)}{\tilde{\omega}_{GG'}(q)(E - E_{n''k-q} - \tilde{\omega}_{GG'}^2(q))} v(q + G'),
\]

with

\[
M_{n'n'}(k, q, G) = \langle nk + q|e^{i(q+G)\cdot r}|k\rangle
\]
\[
\Omega_{GG'}(q) = \frac{\sqrt{4\pi \rho(0)}}{\rho(G - G')} \rho(G - G')
\]
\[
\tilde{\omega}_{GG'}(q) = \frac{|\lambda_{GG'}(q)|}{\cos \phi_{GG'}(q)}
\]
\[
|\lambda_{GG'}(q)|e^{i\phi_{GG'}(q)} = \frac{\Omega_{GG'}(q)}{\delta_{GG'} - \epsilon_{GG'}^{-1}(q; 0)}
\]

In the above expressions for the constituents in Eqs. (2.45) and (2.46), the \( \rho(G) \) are the Fourier components for the charge density and \( \omega_p \) is the usual plasma frequency, \( \sqrt{4\pi \rho(0) e^2/m} \).

An important note to make is that the density enters into the expression for the dielectric matrix through the plasma frequency. The use of semicore pseudopotentials means there is quite a bit of core charge; this is not included in determining the plasma frequency. It has been suggested [108] that tightly bound semicore states do not contribute to the polarizability and thus should not be included when determining the plasma frequency. For transition metal oxides, this has been shown to give GPP dielectric functions that agree better with full-frequency calculations [108].
In addition to the GPP approximation, we construct $\Sigma$ in the diagonal approximation, whereby the quasiparticle energies are obtained, as corrections to the mean-field energies

$$E_{nk} = \epsilon_{nk} + \langle nk | \Sigma(E_{nk}) | nk \rangle - V_{nk}^{xc}$$  \hspace{1cm} (2.47)

In practice, the self-energy is very nearly linear, and so the quasiparticle energy can be determined self-consistently using Newton’s method

$$E_{nk} = \epsilon_{nk} + Z_{nk}(\epsilon_{nk} - V_{nk}^{xc})$$  \hspace{1cm} (2.48)

where $Z_{nk}$ is the quasiparticle renormalization factor, given by

$$Z_{nk} = \frac{d\Sigma_{nk}/dE}{1 - d\Sigma_{nk}/dE}.$$  \hspace{1cm} (2.49)

### 2.5.2 Quasiparticle lifetimes

The noninteracting Green’s function expressed in terms of Bloch states can be written

$$G_{nk}(0)(\omega) = \frac{1}{\omega - \epsilon_{nk} + i\delta_{nk}},$$  \hspace{1cm} (2.50)

whose Fourier transform to the time domain is given by

$$G_{nk}^{(0)}(t) = e^{i\epsilon_{nk}t}.$$  \hspace{1cm} (2.51)

Considering that the mean-field energies $\epsilon_{nk}$ are real, the Green’s function describes particles that do not decay but have a phase factor that oscillates for all time. On the other hand, including the complex self-energy, the quasiparticle energies ($E_{nk} = \epsilon_{nk} + \Sigma_{nk}$) will have a complex component and the Green’s function will decay with time, describing quasiparticles with a finite lifetime that is inversely proportional to the imaginary part of the self-energy.

These two properties can be seen in Figure 2.1: the spectral function within mean-field theory shows a delta function centered at the mean-field energy, describing an infinitely long-lived particle, while the quasiparticle peak has a finite width. The lifetime of a quasiparticle is directly related to the width of the quasiparticle peak in the spectral function. Within quasiparticle theory, the peak has a Lorentzian shape with full width at half maximum $2 \text{Im} \Sigma$. The lifetime can be computed from the self-energy according to

$$\tau_{nk}^{-1} = \frac{2}{\hbar} \text{Im} \Sigma_{nk}(\epsilon_{nk}).$$  \hspace{1cm} (2.52)

The self-energy computed with the GPP approximation does not yield physical quasiparticle lifetimes; the frequency dependence of $\Sigma$ must be considered. This can be done using the contour deformation method, or by evaluating the self-energy at frequencies on the real line. We evaluate the self-energy at the mean-field energies for computational convenience, and also since this correctly yields quasiparticles on the Fermi surface with $\text{Im} \Sigma = 0$. 
2.6 Transport Properties

To discuss the transport properties of solids, we must first have an understanding of the dynamics of Bloch electrons \([3, 144]\). For an electron in the state \(|k\rangle\), its mean velocity is given by (assuming vanishing Berry curvature)

\[
v_k = \frac{1}{\hbar} \frac{\partial E_k}{\partial k}.
\]

Strictly speaking, the expression for the velocity [Eq. (2.53)] holds for a wavepacket constructed from Bloch states centered around \(k\). In the presence of externally applied fields (electric and magnetic), the wavevector of the state evolves according to

\[
\hbar \dot{k} = -e \left( \mathbf{E} + \frac{v_k}{c} \mathbf{c} \times \mathbf{B} \right).
\]

The transport properties of materials are well understood within a semiclassical approach to kinetic theory for a gas of electrons. A gas can be characterized by its probability distribution function \(f(k, r, t)\), which describes the number of electrons with wavevector \(k\) at position \(r\) (at a particular time \(t\)) through the relation

\[
dN = \frac{2}{(2\pi)^3} f(k, r) \, dk \, dr.
\]

The dynamics of the gas are governed by the Boltzmann equation, which gives an expression for the time evolution of \(f\)

\[
\frac{\partial f}{\partial t} = \left( \frac{\partial f}{\partial t} \right)_{\text{diff.}} + \left( \frac{\partial f}{\partial t} \right)_{\text{coll.}} + \left( \frac{\partial f}{\partial t} \right)_{\text{fields}}.
\]

The diffusion term is

\[
\left( \frac{\partial f}{\partial t} \right)_{\text{diff.}} = -v_k \cdot \nabla f(k, r),
\]

while the evolution of \(f\) due to the externally applied fields is

\[
\left( \frac{\partial f}{\partial t} \right)_{\text{fields}} = -\mathbf{k} \cdot \nabla f(k, r) = -\frac{e}{\hbar} \left( \mathbf{E} + \frac{v_k}{c} \mathbf{c} \times \mathbf{B} \right).
\]

The collision term is the most difficult to work out. Generically, elastic scattering can be expressed in terms of some scattering transition probability \(W_{kk'}\) (assuming \(W_{kk'} = W_{k'k}\)) as

\[
\left( \frac{\partial f}{\partial t} \right)_{\text{coll.}} = \int d\mathbf{k}' W_{kk'} [f(\mathbf{k}, \mathbf{r}) - f(\mathbf{k}', \mathbf{r})].
\]

At steady-state conditions, the transport properties do not change with time, and are obtained by solving \(\frac{\partial f}{\partial t} = 0\). If we assume the steady-state distribution does not deviate much from the equilibrium \(f^{(0)}\), we can write

\[
f(k, r) = f^{(0)}(k, r) + f^{(1)}(k, r),
\]
where the $f^{(0)}(k, r)$ is the local equilibrium distribution for a spatially varying temperature $T(r)$ and chemical potential $\mu(r)$ (with $\beta(r) = \frac{1}{k_B T(r)}$)

$$f^{(0)}(k) = \frac{1}{e^{\beta(r)[E(k) - \mu(r)]} + 1}. \quad (2.61)$$

Inserting the expression for $f$ in Eq. (2.60) into Eq. (2.56) for the steady-state ($\frac{\partial f}{\partial t} = 0$), and rearranging terms, keeping only those linear in $f^{(1)}$ and the field, and assuming uniform scattering ($\nabla_r f^{(1)} = 0$), yields the linearized Boltzmann equation

$$v_k \cdot \nabla_k f^{(0)} + \frac{e E}{\hbar} \cdot \nabla_k f^{(0)} = \left( \frac{\partial f^{(1)}}{\partial t} \right)_{\text{coll.}} - \frac{e}{\hbar c} (v_k \times B) \cdot \nabla_k f^{(1)}. \quad (2.62)$$

Once $f = f^{(0)} + f^{(1)}$ is obtained from Eq. (2.62), the charge and energy currents can be computed as

$$j(r) = \frac{2}{(2\pi)^3} e \int dk v_k f(k, r), \quad (2.63)$$

and

$$Q(r) = \frac{2}{(2\pi)^3} \int dk E_k v_k f(k, r), \quad (2.64)$$

respectively. Explicit expressions for these quantities will be given in Sec. 2.6.2. Generalizations to the multiband case can be found in Refs. [3, 144].

### 2.6.1 Relaxation time approximation

In order to deal with the collision term, we assume a simple time dependence, where the deviation of $f$ from the equilibrium distribution (after any perturbation is removed) decays with some characteristic constant relaxation time $\tau_k$,

$$\left( \frac{\partial f}{\partial t} \right)_{\text{coll.}} = - \frac{f(k, r) - f^{(0)}(k, r)}{\tau_k} = - \frac{f^{(1)}(k, r)}{\tau_k}. \quad (2.65)$$

By evaluating the derivatives $\nabla_r f^{(0)}(k, r)$ and $\nabla_k f^{(0)}(k, r)$, defining

$$f^{(1)} \equiv - \frac{\partial f^{(0)}}{\partial E} \phi(k, r), \quad (2.66)$$

and factoring out the common $\frac{\partial f^{(0)}}{\partial E}$ term, Eq. (2.62) reduces to the following expression for $\phi(k, r)$

$$T(r)v_k \cdot \nabla_r \left( \frac{E_k - \mu(r)}{T(r)} \right) + e \mathcal{E} \cdot v_k = \frac{e}{\hbar c} (v_k \times B) \cdot \nabla_k \phi(k, r) + \frac{\phi(k, r)}{\tau_k}. \quad (2.67)$$
2.6.2 Transport coefficients

Now that the linearized Boltzmann equation has been developed within the relaxation time approximation [Eq. (2.67)], expressions for the various transport coefficients will be given.

Isothermal conductivity

The isothermal conductivity can be derived from Eq. (2.67) with a constant temperature and $B = 0$, yielding

$$\phi_n(k) = e\tau_{nk}\mathbf{E} \cdot \mathbf{v}_{nk}, \quad (2.68)$$

which gives

$$f_n^{(1)}(k, r) = \left( -\frac{\partial f^{(0)}}{\partial E} \right) e\tau_{nk}\mathbf{E} \cdot \mathbf{v}_{nk}. \quad (2.69)$$

Plugging Eq. (2.69) into Eq. (2.63), with the understanding that the $f^{(0)}$ term does not contribute to the current, yields

$$j(r) = 2\left( \frac{2\pi}{3} \right) e^2 \sum_n \int dk \left( -\frac{\partial f^{(0)}}{\partial E} \right) \tau_{nk}\mathbf{v}_{nk}\mathbf{v}_{nk}\mathbf{E}. \quad (2.70)$$

The conductivity relates the applied electric field to the induced current via $j = \sigma \mathbf{E}$, so that the conductivity derived from Eq. (2.70) is

$$\sigma_{\alpha\beta} = 2\left( \frac{2\pi}{3} \right) e^2 \int dE \left( -\frac{\partial f^{(0)}}{\partial E} \right) \Sigma_{\alpha\beta}(E), \quad (2.71)$$

For the following subsections, we define the transport distribution function

$$\Sigma_{\alpha\beta}(E) \equiv \frac{1}{\Omega_{cell}} \sum_{nk} v_{nk,\alpha} v_{nk,\beta} \tau_{nk} \delta(E_{nk} - E) \quad (2.72)$$

which allows the transport coefficients to be expressed in a compact form, such as conductivity

$$\sigma_{\alpha\beta} = e^2 \int dE \left( -\frac{\partial f^{(0)}}{\partial E} \right) \Sigma_{\alpha\beta}(E). \quad (2.73)$$

Thermal conductivity and Thermopower

Making use of the transport distribution function [Eq. (2.72)], the thermal conductivity (due to electrons) and thermopower can be expressed as

$$[\sigma S]_{\alpha\beta} = \frac{e}{T} \int dE \left( -\frac{\partial f^{(0)}}{\partial E} \right) (E - \mu) \Sigma_{\alpha\beta}(E), \quad (2.74)$$

$$\kappa_{\alpha\beta} = \frac{1}{T} \int dE \left( -\frac{\partial f^{(0)}}{\partial E} \right) (E - \mu)^2 \Sigma_{\alpha\beta}(E) \quad (2.75)$$

respectively.
2.7 Wannier functions

Wannier functions (WFs), in the simplest formulation, are the Fourier transform of the Bloch functions. A Wannier function for the $n$th band, and centered in the cell located by $\mathbf{R}$, can be constructed via

$$|\mathbf{R}n\rangle = \frac{\Omega_{\text{cell}}}{(2\pi)^3} \int_{\text{BZ}} d\mathbf{k} e^{-i\mathbf{k} \cdot \mathbf{R}} |\psi_{nk}\rangle,$$

where $\Omega_{\text{cell}}$ is the volume of the real-space primitive cell. Wannier functions have the following properties:

- translation by a lattice vector $\mathbf{R}'$
  $$w_{n\mathbf{R}}(\mathbf{r}) = w_{n\mathbf{R}+\mathbf{R}'}(\mathbf{r} + \mathbf{R}')$$
- orthonormal basis for the $n$th band
  $$\langle \mathbf{R}n | \mathbf{R}'n \rangle = \delta_{\mathbf{R}\mathbf{R}'}$$
  $$\psi_{nk} = \sum_{\mathbf{R}} e^{i\mathbf{k} \cdot \mathbf{R}} |\mathbf{R}n\rangle$$
- gauge freedom
  $$\psi_{nk}(\mathbf{r}) \rightarrow e^{i\phi_{nk}(\mathbf{k})} \psi_{nk}(\mathbf{r})$$
  does not change the above properties of the WFs

Generalized Wannier functions are constructed from Bloch functions while allowing for arbitrary unitary transformations among the Bloch functions $\psi_{nk} \rightarrow \sum_{m} u_{mn}^{(k)} |\psi_{mk}\rangle$:

$$|\mathbf{R}n\rangle = \frac{\Omega_{\text{cell}}}{(2\pi)^3} \int_{\text{BZ}} d\mathbf{k} e^{-i\mathbf{k} \cdot \mathbf{R}} \sum_{m} u_{mn}^{(k)} |\psi_{mk}\rangle.$$  

(2.77)

2.7.1 Localization

Because of the gauge freedom in the definition of the WFs, one can construct WFs that are localized in real-space by choosing phases so that the Bloch functions are smooth in $\mathbf{k}$-space. In fact, it has been shown that exponentially localized WFs can always be found for normal insulators (see Ch. 5 for the case of topological insulators). This is typically achieved through the maximal localization procedure [77, 117] to obtain maximally localized Wannier functions (MLWFs). Details of WF construction, and an automated procedure for doing so, are covered in Chapters 4 and 5.
Part II
Research
Chapter 3

Band offsets in c-Si/Si-XII heterojunctions

3.1 Introduction and motivation

There are at least 10 structurally characterized phases of silicon within the pressure range 0–248 GPa, with hints pointing to the existence of several more [50, 133, 89, 46, 143, 81, 23, 43]. Some of these phases are metastable at room temperature and ambient pressure [34]. One such phase, Si-XII, has a rhombohedral unit cell with an 8 atom basis [23] (see Fig. 3.1b) and can be formed using a diamond anvil cell [23, 43] and by nanoindentation of cubic silicon (c-Si) wafers [105]. An interesting aspect of this phase is the presence of five-membered rings, which are expected to have an impact on its electronic properties. Recent I-V measurements performed on Si-XII samples, created through nanoindentation of c-Si substrates, have shown this material to be a semiconductor [106], supporting ab initio calculations that predict a narrow gap semiconductor [73]. Additionally, GW-BSE calculations yield a very interesting absorption spectrum for Si-XII, with indications that this material is a good absorber of light in the optical range [72], possibly pointing to potential usefulness in solar energy applications. Understanding the band offsets at the interface between c-Si and Si-XII is central to the description of the transport properties of this system and to the evaluation of possible device applications. In this work we investigate the band offsets between c-Si and Si-XII to learn more about the physics and nature of these interfaces. This work has been published as Ref. [86] in the literature.

3.2 Band alignment model

There are a multitude of theories to describe interfaces in semiconductor heterostructures and methods to quantify band offsets [82]. These include earlier works such as the electron affinity rule [2], a more recent Green’s function method [122, 123, 121], and explicit interface calculations [130]. We make use of the Green’s function method because of its ability to
describe band offsets in terms of bulk properties. Also, since the detailed interfacial orientations and structures of the observed c-Si/Si-XII junctions have not been determined, the goal of the current study is to provide a general and broad analysis of the band offsets.

The Green’s function method employed for the band offsets is due to Tersoff. The details of the theory are explained in Refs. [122], [123], and [121], but a brief overview will be given here. Tersoff’s theory is born out of the study of Schottky barriers and the concept of states in the gap of the semiconductor that are induced by the metallic states, so called metal induced gap states (MIGS) [66]. In contrast, for semiconductor heterojunctions, the concept of the MIGS is replaced by that of virtual gap states (ViGS) [82]. The ViGS are states with complex wavevector $\mathbf{k}$ that exist in the gap of a semiconductor. These states are characterized as either “valence-like” or “conduction-like” via the cell-averaged Green’s function, allowing for propagation by a lattice vector $\mathbf{R}$,

$$
G(E, \mathbf{R}) = \sum_{n\mathbf{k}} \frac{e^{i\mathbf{k} \cdot \mathbf{R}}}{E_{nk} - E}.
$$

The empty valence-like states carry a positive charge while filled conduction-like states carry a negative charge. The charge neutrality level $E_{\text{CNL}}$ is such that all the valence-like states are filled and all the conduction-like states are empty. This occurs when the magnitude of the contributions from the valence bands and conduction bands are equal

$$
|G_v(E_{\text{CNL}}, \mathbf{R})| = |G_c(E_{\text{CNL}}, \mathbf{R})|.
$$

Applications of this theory have been quite successful in predicting the band offsets of a variety of semiconductor heterojunctions [123]. Other applications have made use of variants of the Tersoff theory. For example, in a study of group III-nitrides and III-/II-oxides, the charge neutrality level is calculated as a Brillouin zone average of the gap provided.

Figure 3.1: Crystal structures of rhombohedral silicon (a) and cubic silicon (b).
only a set of relevant bands are included [111]. Although a more rigorous approach to the
calculation of band offsets at interfaces is possible [130], where the offsets are determined by
the discontinuity in the planar averaged electrostatic potential at the interface, an explicit
interface calculation was not performed in this work because the details of the interface
are not known experimentally. In Ref. [106], Ruffel et al. perform I-V measurements on
nanoindented Si-XII in c-Si substrates. In the regions where the Si-XII is present, there is
also a small amount of Si- III, but, since the (nanoindented) region is so small, the authors
make the reasonable claim that it is highly likely that there exists continuous paths formed by
Si-XII grains. This suggests that a bulk method be used because of the various orientations
of interfaces that make up the c-Si/Si-XII heterojunction.

3.3 Computational method

For the band structure calculations, density functional theory (DFT) was used within
the local density approximation (LDA) as implemented in the Quantum ESPRESSO [35]
electronic structure package. We used a norm-conserving pseudopotential for silicon and
the wavefunctions were expanded in a planewave basis [48]. It is well known that the gap
is underestimated within DFT, which will affect the location of the calculated $E_{CNL}$. Our
band structure calculations for c-Si yield a semiconductor with a gap of 0.50 eV and for
Si-XII we find a semimetal with an overlap of 0.28 eV. In order to correct for this deficiency,
the conduction bands are rigidly shifted to match the experimental gap of 1.11 eV (at 300K)
for c-Si and to match a gap of 0.24 eV Si-XII, calculated by Malone et al. [73] using the ab
initio GW approach [47] for the electron self-energy (as implemented in the BerkeleyGW
package [24]).

In order to evaluate Eq. (3.1), in addition to the electronic spectra, a choice must be
made with respect to the propagation lattice vector $\mathbf{R}$. For zincblende materials, Tersoff
argues that the appropriate choice for the direction of $\mathbf{R}$, for a disordered interface, is along
[100] (along the zig-zag bonded chains, see Ref. [122]), since this is the direction of the
strongest propagation of charge disturbances [54]. In Cartesian coordinates, vectors along
this direction are written as $m\alpha^2(1,1,0)$, with $m$ an integer and $\alpha$ the lattice constant. Tersoff
suggested $m = 4 - 8$ are the appropriate values for Si. Si-XII, having a rhombohedral unit
cell, does not admit this particular choice of $\mathbf{R}$, so the appropriate choice of $\mathbf{R}$ must be
derived from other means. For our purposes, we searched a sufficiently large subset of the
space of possible directions to determine the possible choices of the charge neutrality level.
Although knowing the “correct” direction would be ideal, a tight range of values of the $E_{CNL}$
would be sufficient to give a good estimate.

The procedure for calculating $E_{CNL}$ is as follows: first a self-consistent calculation is
performed to determine the ground state charge density, then a band structure calculation is
performed on a fine grid of $k$-points, and lastly $E_{CNL}$ is determined by finding the energy at
which the valence and conduction contributions to the Green’s function are equal [Eq. (3.2)].
For c-Si, we used a kinetic energy cutoff of 50 Ry for the self-consistent field (SCF) calculation
and Brillouin zone integrations were performed on a $10 \times 10 \times 10$ grid, parameters which converge the total energy to within 1 meV per atom. For Si-XII, the kinetic energy cutoff used was, again, 50 Ry, which makes sense since the same pseudopotential was used. Si-XII is predicted to be a semimetal within DFT, and so smearing was used to perform the Brillouin zone integrations, specifically the “cold” smearing scheme of Marzari and Vanderbilt [79], to enable convergence of the total energy with a reasonable sized k-point sampling. We find that a smearing parameter of 0.03 eV and a $6 \times 6 \times 6$ k-point grid were sufficient. In both cases, the band structures were calculated using the theoretical lattice constant. For c-Si we calculate an equilibrium lattice constant of 5.40 Å while for Si- XII the equilibrium lattice parameters are $a = 5.70$ Å and $\alpha = \beta = \gamma = 109.9^\circ$ (see Fig. 3.1). A detailed account of the internal parameters of the Si-XII unit cell has been given by Pfrommer et al. [96].

3.4 Results

3.4.1 Charge neutrality levels

The charge neutrality level $E_{\text{CNL}}$ was calculated for both c-Si and Si-XII. In the case of c-Si, the value of $E_{\text{CNL}}$ was found to be 0.41 eV (see Fig. 3.2), which is in agreement with the value of 0.36 eV obtained by Tersoff. We determined the charge neutrality level by taking note of the fact that $E_{\text{CNL}}$ has converged with the length of $R$. As mentioned in Section 3.3, this choice of $R$ for c-Si is the appropriate choice for the disordered interface. To determine the charge neutrality level for Si-XII, we examine a variety of directions and notice that the values of $E_{\text{CNL}}$ have converged from values in the range $0.04 - 0.24$ eV to $0.06 - 0.10$ eV, allowing us to estimate $E_{\text{CNL}}$ at 0.08 eV (wrt the valence band maximum) with an uncertainty of about 0.02 eV.

It is worth commenting on the relative convergence of the charge neutrality level of the two systems. The distance the tails of the ViGS penetrate a semiconductor depends on the band gap: shorter for larger gaps and longer for smaller gaps. In c-Si the gap is larger and so $E_{\text{CNL}}$ converges with smaller $R$ while for Si-XII the gap is smaller, the ViGS penetrate further, and larger $R$ vectors are needed to converge $E_{\text{CNL}}$.

3.4.2 Convergence

Calculating the Green’s function in Eq. (3.1) requires a sampling of the band structure over the Brillouin zone. We tested a variety of grids to obtain convergence of $E_{\text{CNL}}$ not only wrt $k$-points but also bands. An example of the convergence of $E_{\text{CNL}}$, for c-Si, wrt bands is given in Fig. 3.4; it is clear that convergence is achieved rapidly with bands, becoming sufficiently stable after the inclusion of approximately 16 conduction bands. The convergence with respect to k-point sampling is given in Fig. 3.5, where it is apparent that convergence is achieved with a $k$-point grid size of $24 \times 24 \times 24$. These tests suggest that for c-Si the charge neutrality level is converged to within 0.01 eV with a $24 \times 24 \times 24$ k-point grid and 20
bands. The same tests were performed for Si-XII, yielding a converged value (within 0.01 eV) for $E_{\text{CNL}}$, for each $\mathbf{R}$, with a 24×24×24 k-point grid and 40 bands. To solve Eq. (3.2), $G_v(E, \mathbf{R})$ and $G_c(E, \mathbf{R})$ were evaluated for energies between the valence band maximum and conduction band minimum and the zero of $|G_v(E, \mathbf{R})| - |G_c(E, \mathbf{R})|$ was determined by linear interpolation.

### 3.4.3 Band alignment

Aligning the charge neutrality levels of the two semiconductors allows us to estimate the band offsets in the case of zero interfacial dipole. The assumption of zero interfacial dipole is motivated by the fact that the interface is between two phases of the same material that only differ in structure but have similar bonding characteristics, as opposed to a heterojunction between very different semiconductors. The values of $E_{\text{CNL}}$ for c-Si and Si-XII are 0.41 eV and 0.08 eV respectively. For the charge neutrality levels we calculated, the resulting valence band offset is $\Delta E_v = 0.3$ eV and the conduction band offset is $\Delta E_c = 0.5$ eV; Fig. 3.6 shows...
Figure 3.3: $E_{\text{CNL}}$ for Si-XII for a variety of high symmetry directions. Each color represents a different lattice vector, given in units of the primitive lattice vectors. Notice, that for the various directions $\mathbf{R}$, the value of $E_{\text{CNL}}$ is converging to a narrow range with the length of $\mathbf{R}$ ($m$ value). The charge neutrality level is referenced to the top of the valence band.

3.5 Summary

To summarize, we performed ab initio calculations of the band structures of c-Si and Si-XII. After adjusting the conduction bands to correct the bandgap problem that exists in density functional theory, the band structures were used to determine the intrinsic charge neutrality level, $E_{\text{CNL}}$, using the method developed by Tersoff [Eq. (3.1)]. To determine the band offsets at the c-Si/Si-XII interface, the charge neutrality levels were aligned and offsets are taken from the resulting band diagram. We conclude that the c-Si/Si-XII interface is of Type I with $\Delta E_v = 0.3$ eV and $\Delta E_c = 0.5$ eV. 

the estimated band diagram for the c-Si/Si-XII interface. These offsets indicate that the heterojunction is of Type I, which agrees with the rectifying behavior seen by Ruffell et al. [106].
Figure 3.4: Convergence of $E_{\text{CNL}}$, for c-Si, along $\mathbf{R} = m\frac{e}{2}(\hat{y} + \hat{z})$ with respect to the number of conduction bands included in Eq. (3.2). Each curve is a different $m$ value. For this test, a 24x24x24 k-point grid was used.
Figure 3.5: Convergence of $E_{\text{CNL}}$, for c-Si, along $\mathbf{R} = m_{\frac{2}{3}}(\hat{y} + \hat{z})$ with respect to size of the k-point grid used in evaluating the Green’s functions in Eq. (3.2). Each curve is a different $m$ value. For this test, we used 40 conduction bands.

Figure 3.6: Band offsets in a c-Si/Si-XII heterojunction as estimated by aligning the charge neutrality levels.
Chapter 4

Automated construction of localized Wannier functions

4.1 Introduction and motivation

Within the quasiparticle approximation, the electronic states of a crystal can be described in terms of single-particle Bloch functions $\psi_{nk}(r)$. These functions are eigenstates of the crystal Hamiltonian, and can be labeled by their band index $n$ and crystal momentum $k$. Wannier functions (WFs) provide an alternative representation in which an entire band of electrons is described by a single function $|R_n\rangle$ localized in or near the unit cell labeled by the lattice vector $R$. In their simplest form, WFs are obtained from the Bloch functions via the Fourier transformation

$$
|R_n\rangle = \frac{V}{(2\pi)^3} \int_{BZ} dke^{-ik\cdot R} |\psi_{nk}\rangle,
$$

where $V$ is the volume of the real-space primitive cell. The definition of WFs is not unique because there is a gauge freedom in the right-hand side of Eq. (4.1). Namely, at each $k$ point and for each $n$, one can change the overall phase of the Bloch state $|\psi_{nk}\rangle$. In fact, one often considers an even more general gauge choice which allows an arbitrary unitary transformation of a set of $N$ bands at each $k$ point,

$$
|\psi_{nk}\rangle \to \sum_m u_{mn}^{(k)} |\psi_{mk}\rangle.
$$

We focus here on the case when these $N$ bands are isolated from the rest. The choice of gauge is now expressed through a $k$-dependent $N \times N$ unitary matrix $u^{(k)}$.

When Wannier functions are localized in real space they have a wide use in the electronic structure community. An extensive review of maximally localized Wannier functions (ML-WFs) and their properties and applications can be found in Ref. [78]. For example, they have been used in the description of electronic polarization [137] and orbital magnetization,
in addition to being used for interpolation of band structures and matrix elements [140, 36, 88] and electron transport calculations [99].

For this reason, one often uses the gauge freedom \( u^{(k)} \) so that the corresponding WFs are localized. As a general consequence of the Fourier transform, the localization of the WFs \(|Rn\rangle\) in \(r\) space will depend on the smoothness of the gauge \( u^{(k)} \) in \(k\) space. If the \( \psi_{nk}(r) \) are chosen with random overall complex phases [which often happens if \( \psi_{nk}(r) \) are acquired numerically by diagonalizing a \( k \)-dependent Hamiltonian matrix separately for each \( k \) point] the WFs obtained from Eq. (4.1) need not be localized. However, if matrices \( u^{(k)} \) are chosen so that Bloch states are smooth in \( k \) space (smooth gauge), the corresponding WFs will be localized in \( r \) space.

The idea of maximally localized Wannier functions and a procedure for obtaining them from a set of composite bands was introduced by Marzari and Vanderbilt [77] for isolated bands and later extended to the case of entangled bands [117]. Maximally localized Wannier functions are constructed by choosing a gauge \( u^{(k)}_{mn} \) for Eq. (4.2) that minimizes the spread functional

\[
\Omega = \sum_n \left[ \langle r^2 \rangle_n - \bar{r}_n^2 \right],
\]

where

\[
\langle r^2 \rangle_n = \langle 0n| r^2 |0n \rangle, \tag{4.4}
\]
\[
\bar{r}_n = \langle 0n| r |0n \rangle. \tag{4.5}
\]

Here the spread functional \( \Omega \) is written in terms of the Wannier functions \(|0n\rangle\). Usually there exists a global minimum of \( \Omega \) corresponding to a unique choice of \( u^{(k)} \) (up to translation of the WFs and their overall complex phase), but in some cases there are multiple solutions [77].

Using the general form of Eq. (4.1) including the \( u^{(k)} \) matrix in Eq. (4.2), the spread can be recast in terms of the Bloch states. More specifically, \( \Omega \) can be expressed only as a function of the overlaps of the periodic parts of the Bloch functions at neighboring \( k \) points \( k \) and \( k + b \),

\[
m^{(k,b)}_{ij} = \langle u_{ik}| u_{jk+b} \rangle. \tag{4.6}
\]

See Appendix A.1 for an explicit definition of \( \Omega \) in terms of \( m^{(k,b)} \),

\[
\Omega = \Omega[m^{(k,b)}]. \tag{4.7}
\]

Here we only note that spread \( \Omega \) can be decomposed into three parts: the invariant part, which does not depend on the gauge, and the diagonal part and the off-diagonal parts which do,

\[
\Omega[m^{(k,b)}] = \Omega_I + \Omega_D + \Omega_{OD}. \tag{4.8}
\]
The procedure for minimizing $\Omega$, outlined in Refs. [77] and [117], is implemented in the Wannier90 code [83] and has become the standard method for obtaining localized WFs. A notable drawback in the standard approach that we address in this chapter is that one often needs to provide a good initial guess of the MLWFs to find the global minimum of $\Omega$. In this work we demonstrate a modified procedure, in which localized Wannier functions are constructed as a linear combination of physically based atom-centered orbitals without requiring an initial guess, as in the standard approach [77]. This work has been published as Ref. [85] in the literature. This is achieved by finding optimal projection functions (OPFs) so that the resulting Wannier functions obtained via projection (as in Section 4.2.1) are as localized as possible. This OPF method could, for example, be used in constructing material properties databases, such as the database of the Materials Project [49], by providing a simple localized Hamiltonian that could serve as a descriptor for the electronic structure of a material. We present the theoretical approach and numerical methods in Sections 4.2 and 4.3. Several realistic materials are investigated in Sections 4.4 to illustrate our approach for constructing localized WFs.

Schemes beyond the standard implementation [77, 117, 83] have been developed by others to improve the construction of MLWFs and their properties. The inclusion of unoccupied anti-bonding states has been shown [125, 126] to give more localized Wannier functions, but at the expense of a chemical picture of the occupied states. Additionally, constraints on the $u^{(k)}$ matrices can be imposed in order to construct localized Wannier functions that possess all the space group symmetries of the crystal [107].

4.2 Standard approach

Here we summarize the main result of Ref. [77] for a two-step construction of maximally localized Wannier functions. In the first step of minimizing the spread functional $\Omega$ one needs to guess orbitals $g_j(r)$ with roughly the same orbital characters and real-space location $r_j$ as the target MLWFs. This choice is often done based on an intuitive understanding of the band structure of the crystal under investigation. Given a choice of $g_j(r)$ close to target WFs, one constructs the gauge for which spread functional $\Omega$ is near its global minimum (better choices give $\Omega$ closer to the global minimum). In the second step, this initial gauge choice is iteratively optimized until $\Omega$ reaches a global minimum. In practice, the second step usually reduces the spread $\Omega$ only by 20% or less.

4.2.1 First step

Now we describe the first step of this procedure in the simple case of a single band of states $\psi_k(r)$. Given a localized function $g(r)$ approximating the target MLWF at the origin, we first project it onto the Bloch state $\psi_k(r)$ at each $k$

$$a^{(k)} = \langle \psi_k | g \rangle .$$ (4.9)
Now we rotate the phase of Bloch state $\psi_k(\mathbf{r})$ so that the relative phase of rotated Bloch state and $g(\mathbf{r})$ is zero for all $k$ points,

$$|\psi_k\rangle \rightarrow a^{(k)}(a^{(k)*}a^{(k)})^{-1/2} |\psi_k\rangle.$$  \hspace{1cm} (4.10)

It is easy to check that if the initial guess $g(\mathbf{r})$ were a true target MLWF, inserting these rotated Bloch states into Eq. (4.1) would give back the target MLWF. However, since $g(\mathbf{r})$ is only an approximation, the spread $\Omega$ of the rotated Bloch states is not exactly at the global minimum. For a good guess $g(\mathbf{r})$, however, the spread should be close to the global minimum.

Following the procedure in the case of a single band, we now generalize it to the case of $N$ composite bands. First, we choose a set of $N$ localized orbitals $g_j(\mathbf{r})$ that are approximately equal to the $N$ target MLWFs, $|g_j\rangle \approx |0_j\rangle$. \hspace{1cm} (4.11)

Here we choose for convenience $|g_j\rangle$ to be close to the MLWFs near the origin ($\mathbf{R} = \mathbf{0}$), but in principle any other $\mathbf{R}$ can be chosen.

Next we compute the overlap between all $N$ Bloch bands and $N$ initial guesses for the WFs, $a_{ij}^{(k)} = \langle \psi_{ik} | g_j \rangle$. \hspace{1cm} (4.12)

Unlike the case of a single band, $a^{(k)}$ for an isolated group of $N$ Bloch bands is a $N \times N$ matrix, so that Eq. (4.10) generalizes to

$$|\psi_{ik}\rangle \rightarrow \sum_j u_{ji}^{(k)} |\psi_{jk}\rangle = \sum_j a_{jl} \left[ (a^\dagger a)^{-1/2} \right]_{li} |\psi_{jk}\rangle.$$  \hspace{1cm} (4.13)

Here, to simplify notation, we suppress the $k$ dependence of $a$. The inverse square root on the right-hand side is the matrix square root of $(a^\dagger a)^{-1}$. For further simplification, define $u_x$, for an arbitrary matrix $x$, as $u_x \equiv x(x^\dagger x)^{-1/2}$. \hspace{1cm} (4.14)

The matrix $u_x$ is unitary by construction. In fact, it is the closest unitary approximation of $x$.

In practice, the unitary matrix $u_x$ is obtained via the singular value decomposition (SVD) of $x$. If $x = z dv$ is a SVD with $z$ and $v$ unitary, and $d$ diagonal, then $u_x$ is simply $zv$. Given this notation, the earlier gauge transformation from Eq. (4.13) now reads

$$|\psi_{ik}\rangle \rightarrow \sum_j [u_a^{(k)}]_{ji} |\psi_{jk}\rangle.$$  \hspace{1cm} (4.15)

As in the case of a single band, if trial orbitals $g_j(\mathbf{r})$ were chosen close enough to the target MLWF, the gauge $u_a^{(k)}$ from Eq. (4.15) will by definition give a spread close to the global minimum,

$$\Omega = \Omega \left[ u_a^{(k)*} t_{l(k,b)} u_a^{(k+b)} \right].$$  \hspace{1cm} (4.16)
Here we implicitly wrote $\Omega$ in terms of overlap matrices $m^{(k,b)}$ and used gauge transformation of Bloch states from Eq. (4.2) to get transformation of the overlap matrix $m^{(k,b)}$ defined in Eq. (4.6).

### 4.2.2 Second step

The initial gauge $u_a^{(k)}$ can be further improved in the second step by rotating, at each $k$ point, the gauge from $u_a^{(k)}$ to $u_a^{(k)}v^{(k)}$ with an appropriate choice of $k$-dependent matrices $v^{(k)}$. The spread functional $\Omega$ is minimized using the method of steepest descent. The gradient is determined by calculating the derivative of the spread with respect to the unitary matrices $v^{(k)}$ and then following the path along the direction which minimizes $\Omega$. Written more formally, the second step of the standard procedure finds a set of unitary $N \times N$ matrices

$$v^{(k)} \in U(N, N), \text{ one for each } k,$$

that

$$\text{minimize } \Omega\left[ v^{(k)}u_a^{(k)\dagger}m^{(k,b)}u_a^{(k+b)\dagger}v^{(k+b)}\right].$$

Quite generally, the global minimization of a function using the steepest descent algorithm is bound to work well when one starts near the global minimum. Otherwise it is quite possible for the algorithm to get stuck in a local minimum. In other words, the second step of the procedure will arrive at the true MLWFs as long as the initial guesses $g_j(r)$ in the first step are close enough. It is this issue that we aim to address in the following sections of this chapter: how to automatically construct a gauge that is guaranteed to be close to the global minimum.

### 4.3 Alternative approach

In our approach, instead of choosing $N$ functions $g_j(r)$ that are close to the $N$ target MLWFs, we start with a larger set of $M$ functions ($M \geq N$) labeled $h_j(r)$. These functions $h_j$ will be chosen so that any MLWF near the origin ($R = 0$) can approximately be written as a linear combination of $h_j$. In other words, the space spanned by $h_j$ must approximately contain, as a subset, the space spanned by the MLWFs near the origin,

$$\text{Span}(|h_j\rangle) \supset \text{Span}(|0n\rangle).$$

The requirement on $|h_j\rangle$ is significantly less restrictive than that on $|g_j\rangle$ in the standard approach. In fact, the requirement Eq. (4.19) should be rather easily satisfied. Since we expect MLWFs to be linear combinations of atomiclike valence electrons, we can simply choose $h_j$ to be a set of atom-centered atomic orbitals for each atom in the crystal basis and for each relevant atomiclike orbital in the valence (some combination of $s$, $p$, $d$, $f$ atomic orbitals, depending on the valence). If nominal valence atomiclike orbitals are not enough
to satisfy Eq. (4.19) (which might happen for example in material under extreme pressure), one can always include atomic orbitals with higher radial and orbital quantum numbers.

In the case of covalently bonded materials, a specific target MLWF might have its center on a covalent bond at the edges of the primitive unit cell. If this is the case, then we can expand the set \( h_j(r) \) by including the periodic images of a few atoms in the crystal basis, so that in the end, for each unique covalent bond, both atoms forming the bond are included in \( h_j(r) \).

Since the functions \( h_j \) satisfy Eq. (4.19), it is possible to approximate the MLWFs as linear combinations of \( h_j \). Formally, it is possible to find a semiunitary rectangular \( M \times N \) matrix \( W \) such that the functions

\[
|\bar{g}_j\rangle = \sum_{i=1}^{M} W_{ij} |h_i\rangle \quad (4.20)
\]

are close to the target MLWFs. (Since \( W \) is rectangular, the \( N \) functions \( \bar{g}_j \) are linear combinations of \( M \) functions \( h_j \).) Thus obtaining approximate MLWFs is equivalent to finding the matrix \( W \). We shall call these \( \bar{g}_j \) optimized projection functions (OPFs).

To measure the closeness of \( \bar{g}_j \) to the target MLWFs, we need to express spread \( \Omega \) in terms of \( W \). Therefore, we first need a projection of \( \bar{g}_j \) into Bloch states. Since \( \bar{g}_j \) depends on \( W \), it is more convenient to first project \( h_j \) onto the Bloch states, yielding the \( N \times M \) projection matrix

\[
A^{(k)}_{ij} = \langle \psi_i | h_j \rangle \quad (4.21)
\]

Given \( A^{(k)} \) we can compute the overlap matrix between the \( \bar{g}_j \) and the Bloch states,

\[
\bar{a}_{ij}^{(k)} = \langle \psi_i | \bar{g}_j \rangle = \sum_{l=1}^{M} \langle \psi_i | h_l \rangle W_{lj}, \quad (4.22)
\]

or, in short,

\[
\bar{a}^{(k)} = A^{(k)} W. \quad (4.23)
\]

Here we adopted the convention that small (\( N \times N \)) square matrices are written with lowercase Latin letters, while rectangular (\( N \times M \)) or large square matrices (\( M \times M \)) are denoted by uppercase Latin letters.

Now we are ready to express \( \Omega \) in terms of \( W \). Combining Eq. (4.16) and Eq. (4.23) yields

\[
\Omega = \Omega \left[ u_{AW}^{(k)} u_{AW}^{(k+b)} \right]. \quad (4.24)
\]

To draw comparison with Eqs. (4.17) and (4.18), in our approach the process of constructing MLWFs is equivalent to finding

\[
W \in \mathcal{U}(M, N), \text{ a single matrix} \quad (4.25)
\]
that
\[ \text{minimizes } \Omega \left[ u_{AW}^{(k)†} m_{(k,b)}^{(k+b)} u_{AW}^{(k+b)} \right]. \] (4.26)

Once the \( W \) which minimizes Eq. (4.26) is found, we use the matrices \( u_{AW}^{(k)} \) to rotate Bloch states at each \( k \) point into a smooth gauge. In most of the concrete cases studied, the spread of the Wannier functions corresponding to this gauge is within 1% of the global minimum (this is discussed further in Section 4.4) and therefore there is no need to improve the gauge further. However, in principle one could run the second step of the standard procedure to bring spread to its true global minimum and thus obtain maximally localized Wannier functions.

Now we will compare our approach to the standard method in more detail, outlining both the advantages and disadvantages of our approach. We also discuss the approximations that are made to implement an algorithm to construct the \( W \) matrix.

### 4.3.1 Comparison to the standard approach

The procedure for constructing MLWFs by generating OPFs [Eqs. (4.25) and (4.26)] has several advantages compared to the standard procedure [Eqs. (4.17) and (4.18)]. First, OPF construction is given by a single matrix \( W \), instead of a set of \( u^{(k)} \) matrices, one at each \( k \) point. For this reason, as will be shown in Section 4.3.2, one can more directly solve Eq. (4.26) without using the method of steepest descent; rather, an iterative procedure is used to construct \( W \) as a product of large unitary transformations (Givens rotations). Therefore, this procedure is less likely to get stuck in a local minimum. The second advantage of OPF construction is that the \( W \) matrix itself has a lot of chemical information encoded in it. For example, one can see directly from \( W \) the contribution of the various atomic orbitals to each OPF and thus the corresponding Wannier functions. We discuss this point on concrete examples in Section 4.4. Third, the use of a single matrix might make it easier to impose constraints such as crystal symmetry.

There are however some disadvantages to the OPF construction approach. First, the spread \( \Omega \) in Eq. (4.26) depends nonlinearly on \( W \) since it appears under the matrix inverse square root in \( u_{AW}^{(k)} \). In fact, Taylor expansion of the inverse square root leads to a power series in all positive integer powers of \( W \). Second, since we do not want to rely on a steepest decent method, minimization of the diagonal part of the spread [\( \Omega_D \) in Eq. (4.8)] becomes nontrivial.

In the following section we introduce two simplifications to Eq. (4.26) which deal with these two disadvantages of OPF and allow for an efficient numerical construction of OPFs in all the cases studied.

### 4.3.2 Simplifications

The following two subsections describe two simplifications that turn minimization of Eq. (4.26) into a numerically efficient form.
Linearizing $u_{AW}$

The first simplification in minimizing the spread $\Omega$ from Eq. (4.26) is to expand it to the leading order in $W$. Explicitly writing $u_{AW}$ in terms of its definition [Eq. (4.14)] and ignoring $k$ index for the moment,

$$u_{AW} = AW (W^\dagger A^\dagger AW)^{-1/2}.$$  \hfill (4.27)

For $W$ which minimizes Eq. (4.26) we expect

$$W^\dagger A^{(k)} A^{(k)} W \approx I_N,$$  \hfill (4.28)

for all $k$ since the OPFs approximately overspan the space of MLWFs ($I_N$ is the $N \times N$ identity matrix). Therefore, at least near the optimal value of $W$, we are justified in Taylor expanding $u_{AW}$ around $W^\dagger A^\dagger AW$ close to the identity ($I_N$),

$$u_{AW} = AW \left[ I_N - \frac{1}{2} (W^\dagger A^\dagger AW - I_N) + \ldots \right].$$  \hfill (4.29)

Therefore, to lowest order, $u_{AW} \approx AW$. Restoring unitarity we can replace $A$ with $U_A$, thus obtaining a unitary approximation to $U_{AW}$,

$$u_{AW} \approx U_A W.$$  \hfill (4.30)

Here $U_A$ has been constructed according to the Löwdin orthonormalization procedure given by Eq. (4.14). We follow here the notation we introduced earlier so that $U_A$ with upper case $U$ is a rectangular $N \times M$ matrix (while $u_{AW}$ with lowercase $u$ is a square $N \times N$ matrix). We also note here that Eq. (4.30) is exact if $W$ were a square matrix.

Inserting Eq. (4.30) into Eq. (4.26) we find that construction of OPFs is equivalent to finding a rectangular matrix $W \in U(M,N)$ that

minimizes $\Omega \left[ W^\dagger U_A^{(k)} m^{(k,b)} U_A^{(k+b)} W \right].$  \hfill (4.31)

Here, $U_A^\dagger m U_A$ is identified as the enlarged ($M \times M$) overlap matrices projected into the space of $M$ orbitals $h_j$.

In most cases, the $W$ that minimizes Eq. (4.31) also satisfies Eq. (4.28), which then justifies the Taylor expansion of $u_{AW}$. However, occasionally this is not the case (for example, in strongly covalent materials with a lot of symmetry). Therefore, we will introduce a Lagrange multiplier $\lambda$ to Eq. (4.31), which imposes condition Eq. (4.28). With this modification, we now seek matrix $W$ and $\lambda$ at a saddle point of the Lagrangian,

$$\mathcal{L}(W; \lambda) = \Omega \left[ W^\dagger U_A^{(k)} m^{(k,b)} U_A^{(k+b)} W \right] + \lambda w \sum_{k} \sum_{i=1}^{N} \left[ \left[ W^\dagger A^{(k)} A^{(k)} W \right]_{ii} - 1 \right]^2.$$  \hfill (4.32)

For convenience we rescaled the Lagrange multiplier $\lambda$ so that $\lambda = 1$ corresponds to a situation where the relative importance of the first and second term in the Lagrangian $\mathcal{L}$ are equal ($w$ is defined as $w = \sum_b w_b$ and $w_b$ are $k$-point weights appearing in the definition of $\Omega$; see Appendix A.1).
Replacing $\Omega$ with $\Omega_{I,OD}$

Now we show that within our approach one can replace, in Eq. (4.26), the total spread $\Omega$ with $\Omega_{I,OD}(= \Omega_I + \Omega_{OD})$ thus ignoring diagonal part of the spread $\Omega_D$.

We now examine how the diagonal and off-diagonal spread depend on the gauge transformation written in the Wannier space. The most general gauge transformation of Bloch states is given by Eq. (4.2) and it involves an arbitrary unitary transformation of the states at each $k$ point in the Brillouin zone. In the Wannier space, this same gauge transformation corresponds to the unitary mixtures of WF's among all unit cells,

$$|0n\rangle \rightarrow \sum_\mathbf{P} \sum_m u_{mn}^{(\mathbf{P})}|\mathbf{P}m\rangle.$$  \(4.33\)

Here the matrix $u_{mn}^{(\mathbf{P})}$ is the Fourier transform of the matrix $u_{mn}^{(k)}$ in Eq. (4.2). A gauge transformation for which $u_{mn}^{(\mathbf{P})}$ is nonzero only for $\mathbf{P} = 0$ we will call an intracell gauge transformation, since it involves only mixtures of the WFs in the same unit cell.

Let us now start from a set of MLWFs in the home cell $|0n\rangle$ and see what is the effect of the intracell gauge transformation on $\Omega_D$ and $\Omega_{OD}$. First we will express the diagonal and off-diagonal spread in terms of the WFs \[77\],

$$\Omega_D = \sum_n \sum_{\mathbf{R} \neq 0} |\langle \mathbf{R}n|\mathbf{r}|0n\rangle|^2,$$

$$\Omega_{OD} = \sum_{m \neq n} \sum_{\mathbf{R}} |\langle \mathbf{R}m|\mathbf{r}|0n\rangle|^2.$$  \(4.34\),  \(4.35\)

Since the MLWFs are exponentially localized, we expect that the dominant term of a gauge dependent spread $\Omega_D + \Omega_{OD}$ will be the $\mathbf{R} = 0$ term. Since the $\mathbf{R} = 0$ term appears only in $\Omega_{OD}$, it will dominate over $\Omega_D$ for an intracell gauge transformation.

Let us return now back to the optimization problem Eq. (4.26) in question. By construction, the OPFs $\bar{g}_j$ approximately overspan the space of MLWFs near the origin; in other words, they are related by an intracell gauge transformation. Therefore, we are justified in ignoring the diagonal part of the spread $\Omega_D$ in Eq. (4.26).

With this simplification, the problem of finding MLWFs is reduced to finding a rectangular semiunitary matrix $W$ and a real number $\lambda$ which are at a saddle point of the Lagrangian,

$$\mathcal{L}(W, \lambda) = \Omega_{I,OD} \left[ W^\dagger U_A^{(k)\dagger} m^{(k,b)} U_A^{(k+b)\dagger} W \right] + \lambda w \sum_k \sum_{i=1}^N \left| \left[ W^\dagger A^{(k)\dagger} A^{(k)} W \right]_{ii} - 1 \right|^2.$$  \(4.36\)

Inserting here an explicit definition of $\Omega_{I,OD}$ (see Appendix A.1) and ignoring the constant term and the $1/N$ prefactor, we obtain

$$\mathcal{L}(W, \lambda) = - \sum_{k,b} w_b \sum_{i=1}^N \left| \left[ W^\dagger U_A^{(k)\dagger} m^{(k,b)} U_A^{(k+b)\dagger} W \right]_{ii} \right|^2 + \lambda w \sum_k \sum_{i=1}^N \left| \left[ W^\dagger A^{(k)\dagger} A^{(k)} W \right]_{ii} - 1 \right|^2.$$  \(4.37\)
Let us now define the following two quantities that are independent of $W$ and $\lambda$:

\begin{align*}
M^{(k,b)} &= U_A^{(k)} m_{(k,b)} U_A^{(k+b)}, \\
S^{(k)} &= A^{(k)} A^{(k)} - I_M.
\end{align*}

(4.38)

(4.39)

With this simplification, the Lagrangian Eq. (4.37) now simply reads

\[ \mathcal{L}(W, \lambda) = \sum_{\alpha} t^{(\alpha)} \sum_{i=1}^{N} \left| \left[ W^\dagger X^{(\alpha)} W \right]_{ii} \right|^2. \]

(4.40)

Here $X^{(\alpha)}$ stands for a collection of $M^{(k,b)}$ and $S^{(k)}$ matrices. The $t^{(\alpha)}$ are the weights associated with the matrices $X^{(\alpha)}$, with a weight $-w_b$ for the $M^{(k,b)}$ matrices and a weight $\lambda w$ for the $S^{(k)}$ matrices. Therefore, we have reduced a problem of finding MLWFs to the problem of codiagonalizing a set of large ($M \times M$) square matrices $X^{(\alpha)}$ with a single (i.e. $k$-point independent) rectangular ($M \times N$) matrix $W$. A mathematically similar approach for a square matrix $W$ has been used in Ref. [40] to find MLWFs of a localized system.

In Appendix A.2 we present a numerically efficient algorithm for minimizing Eq. (4.40), largely following Refs. [16] and [116]. In the following section, we illustrate the OPF procedure and empirically validate the approximations discussed above.

\section{4.4 Illustration of our approach}

We now illustrate the OPF procedure on a variety of systems with chemical bonding ranging from ionic to covalent. For predominantly ionic materials we choose NaCl, Cr$_2$O$_3$, and LaMnO$_3$. The last two cases have additional complexity because they have magnetic and orbital order on the transition metals. For predominantly covalently bonded materials we choose cubic silicon (c-Si), strongly distorted silicon with 20 atoms in the primitive unit cell (Si-20 from Ref. [139]), cubic GaAs, and SiO$_2$ in the ideal $\beta$-cristobalite structure.

We computed Bloch wave functions for all seven compounds within the density-functional theory and plane wave pseudopotential approach as implemented in the Quantum ESPRESSO package [35]. The atomic potentials were replaced with ultrasoft [131] pseudopotentials from the GBRV [33] library. For the plane wave cutoff, we used 40 and 200 Ry for the wave functions and charge density, respectively. All calculations are done with experimental lattice parameters. In the case of Cr$_2$O$_3$ we sampled the Brillouin zone on a uniform $6 \times 6 \times 6$ $k$-point grid and for all other cases we used a $4 \times 4 \times 4$ grid.

Using the Bloch wave functions, we computed the overlap matrices $m^{(k,b)}$ between the neighboring Bloch states and the overlaps $A^{(k)}$ between the Bloch states and atomiclike functions that approximately overspan the space of MLWFs. For predominantly ionic materials in our test (NaCl, Cr$_2$O$_3$, and LaMnO$_3$), $A^{(k)}$ includes projections of Bloch states into all valence atomiclike functions for all atoms in the primitive unit cell. For covalently bonded materials (c-Si, Si-20, GaAs, and SiO$_2$) some Wannier function centers lie on the edge of
the primitive unit cell (see Section 4.3), so we included in $A^{(k)}$ projections onto atoms near the edge of the cell. Failing to include these additional projections in the case of c-Si yields Wannier functions at the computational unit cell boundary with spreads two times larger than if we include the additional projections.

We also checked the opposite case by overspanning the space of MLWFs even further by including orbitals into $A^{(k)}$ that are nominally not in valence (for example, d-orbitals in the case of cubic silicon). In this case, the final spread for the WFs for the occupied valence band complex is unaffected and the matrix elements of $W$ corresponding to these additional orbitals is small, as expected.

Given matrices $m^{(k,b)}$ and $A^{(k)}$ and a choice of the parameter $\lambda$ we now find matrix $W$ (i.e., OPFs) that minimizes Lagrangian from Eq. (4.40) using the algorithm described in Appendix A.2. Given $W$, we construct the $u^{(k)}_{AW}$ to rotate Bloch states into a smooth gauge as described in Section 4.3. The smoothness of this gauge is quantified by first computing the spread $\Omega^{OPF}$ from the rotated overlap matrices in Eq. (4.26) and then comparing it to the spread $\Omega^{GM}$ at the global minimum. (We define $\Omega^{GM}$ to be a spread of the Wannier functions after running both steps of the standard procedure for obtaining MLWFs. For convenience, in the first step of finding the global minimum, we do not guess the initial projections but instead project into the OPFs obtained from our approach.)

Figure 4.1 shows, for all seven cases studied, the ratio of the spread $\Omega^{OPF}$ and $\Omega^{GM}$ as a function of $\lambda$ on a logarithmic scale. In all cases, the spread $\Omega^{OPF}$ is nearly insensitive to the value of $\lambda$ over several orders of magnitude. For example, in the case of GaAs or LaMnO$_3$ spread $\Omega^{OPF}$ is nearly the same for $0.01 < \lambda < 100$. In the worst case scenario (c-Si), the spread is still nearly the same for $0.1 < \lambda < 2$. Therefore, even though in principle one may need to vary $\lambda$ to find an optimal value of spread, in practice, $\lambda \sim 1$ is usually a good enough choice.

In each of the seven test cases, the spread $\Omega^{OPF}$ is only just 1% larger than at a global minimum ($\Omega^{GM}$). In the worst case situation (Si-20), the spread is only 6% larger than at a global minimum. As mentioned earlier in Section 4.3, this spread could be reduced further by starting from OPFs as initial projections and running the second step of the standard procedure.

We give numerical values of $\Omega^{OPF}$ and $\Omega^{GM}$ in Table 4.1 along with a decomposition of spread into diagonal and off-diagonal components. From here we find an additional validation of two simplifications discussed in Section 4.3.2. First, Table 4.1 shows that linearization of $u_{AW}$ is justified since the off-diagonal component of the spread $\Omega^{OPF}$ and $\Omega^{GM}$ is nearly the same. Second, replacing $\Omega$ with $\Omega_{L,OD}$ (thus, ignoring diagonal spread) is justified within our approach since diagonal spread of $\Omega^{OPF}$ and $\Omega^{GM}$ are both very small compared to the total spread.

### 4.4.1 Insight gained from the matrix $W$

To demonstrate the kind of insight that can be gained from analyzing the $W$ matrix, we analyze here in more detail case of LaMnO$_3$ and Cr$_2$O$_3$. In both cases, $s$ and $p$ orbitals on
Figure 4.1: Ratio of $\Omega^{OPF}$ and $\Omega^{GM}$ as a function of Lagrange multiplier $\lambda$ on a logarithmic scale.
Table 4.1: Total spread $\Omega^{\text{OPF}}$ computed within our approach and at the global minimum $\Omega^{\text{GM}}$ for all seven materials studied. We also give diagonal and off-diagonal components of spread in each case ($\Omega_D$ and $\Omega_{OD}$). The spreads $\Omega^{\text{OPF}}$ are obtained using the optimal value of $\lambda$ (see Fig. 4.1). The units for the spreads are Å$^2$. In the case of Cr$_2$O$_3$ we wannierize only the topmost 12 bands below the Fermi level, and in the case of LaMnO$_3$ we wannierize the top 2 spin-up bands. In all other cases we wannierize all valence bands.

<table>
<thead>
<tr>
<th></th>
<th>$\Omega^{\text{OPF}}$</th>
<th>$\Omega^{\text{GM}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total</td>
<td>Components</td>
</tr>
<tr>
<td></td>
<td>$\Omega_D$</td>
<td>$\Omega_{OD}$</td>
</tr>
<tr>
<td>c-Si</td>
<td>6.51</td>
<td>0.00</td>
</tr>
<tr>
<td>Si-20</td>
<td>103.91</td>
<td>0.05</td>
</tr>
<tr>
<td>GaAs</td>
<td>7.25</td>
<td>0.02</td>
</tr>
<tr>
<td>SiO$_2$</td>
<td>9.39</td>
<td>0.00</td>
</tr>
<tr>
<td>Cr$_2$O$_3$</td>
<td>36.04</td>
<td>0.10</td>
</tr>
<tr>
<td>LaMnO$_3$</td>
<td>14.89</td>
<td>0.15</td>
</tr>
<tr>
<td>NaCl</td>
<td>4.05</td>
<td>0.00</td>
</tr>
</tbody>
</table>

the neighboring oxygen atoms outside the computational unit cell are included in $A^{(k)}$ in order to complete the octahedral coordination of the Cr and Mn atoms.

We studied LaMnO$_3$ in its low temperature ($\lesssim$ 135K) A-AFM phase characterized by ferromagnetic ordering of the Mn spins in-plane and antiferromagnetic order between planes [26]. In addition to the magnetic order, Mn $d$ states are orbitally ordered, oxygen octahedra are tilted and Jahn-Teller distorted. In the following, we focus only on the two topmost spin-polarized bands below the Fermi level. Analyzing the $W$ matrix we see that the Wannier functions for the two topmost bands in LaMnO$_3$ are dominantly composed of rotated $d_{z^2}$ components on Mn that are oriented perpendicular to each other. This can be seen also by analyzing the $W$ matrix for these two WFs,

$$|1\rangle \approx 0.5 |\text{Mn1}; d_{z^2}\rangle + 0.6 |\text{Mn1}; d_{xy}\rangle$$
$$|2\rangle \approx 0.6 |\text{Mn2}; d_{z^2}\rangle - 0.5 |\text{Mn2}; d_{xy}\rangle .$$

Figure 4.2 shows a plot of these WFs for the top bands with isosurfaces in the left panels and contour plots in the right panels. The contours are plotted in the plane perpendicular to the $c$ axis, cutting through the Mn atom.

Furthermore, The $W$ matrix shows hybridization of the Mn $d$ states with the oxygen $p$ states, with the corresponding elements of $W$ having a magnitude of approximately 0.2 (three times smaller than for Mn $d$ states). The contribution of the $p$-like lobes (colored red) can be seen in the right panels of Fig. 4.2 as the large lobes near the center.
Now we analyze the case of Cr$_2$O$_3$ in more detail. Cr$_2$O$_3$ is an antiferromagnetic insulator with four Cr atoms in the primitive unit cell. Therefore, we expect each Cr$^{3+}$ ion to nominally have three occupied $d$ orbitals of same spin. These three occupied $d$ orbitals on four Cr ions form a complex of $3 \times 4 = 12$ isolated bands that make up the topmost valence bands. Again, analyzing the $W$ matrix we obtained within our approach we find that each of the twelve WFs is a particular linear combination of all five $d$ orbitals, all having the same spin component along the $z$ axis. In fact, there is a large degeneracy regarding the particular combination of $d$ orbitals that make up the WFs. For example, even slight change in $\lambda$ from 1 to 2 gives different linear combinations of $d$ orbitals, while the spread remains nearly the same (see Fig. 4.1). This observation is consistent with the fact that the choice of MLWFs is not always unique. This was first suggested in Ref. [77] for the case of LiCl. There it was found that an arbitrary rotation of the $sp^3$ orbitals on chlorine atoms has no effect on the total spread $\Omega$.

4.5 Summary

We present an automated procedure for constructing maximally localized Wannier functions for an isolated group of bands. The extension of our method to the case of entangled bands will be the subject of future work. Instead of having to guess functions (initial projections) that approximate the MLWFs as in Ref. [77], our approach only requires as input a set of functions that overspan the space of MLWFs. In practice, this can rather easily be achieved by selecting an appropriate set of valence atomiclike functions.
Figure 4.2: Maximally localized Wannier functions of the two topmost valence bands in LaMnO$_3$. Isosurfaces of the WFs are shown on the left, looking down along the $c$ axis. The large green dots are La, medium purple dots are Mn, and small red dots are O. On the right, we show contour plots of the Wannier functions in the plane perpendicular to the $c$ axis, cutting through the Mn atom.
Chapter 5

Automated construction of localized Wannier functions for bands with nontrivial topology

5.1 Introduction and motivation

A useful representation of the occupied states in a periodic insulator is the Wannier function. Wannier functions (WFs) provide a localized real-space description of the extended Bloch states [78]. In particular, WFs give a chemical picture of the bonding nature of a material, an alternative real-space formalism for many quantities, and can also be used for interpolating various physical properties on a fine mesh in the Brillouin zone [132, 36]. For examples, WFs can be used to compute electronic polarization, orbital magnetization, the component of isotropic magnetoelectric coupling, and various transport properties.

However, an exponentially localized Wannier function representation does not exist for insulators with a non-zero Chern number \( C \) \([124, 15]\). Insulators with a non-zero Chern number are called integer quantum Hall insulators (or Chern insulators) and are characterized with a non-zero Hall conductance \( \sigma = C e^2/h \). (Three-dimensional insulators are characterized by a triplet of Chern numbers.)

In the past several years there has been significant interest in a group of materials related to the Chern insulator. These are called \( \mathbb{Z}_2 \) topological insulators (TIs). In two dimensions these \( \mathbb{Z}_2 \) topological insulators can be seen as topologically equivalent to two copies of a Chern insulator, one with \( C = 1 \) and another with \( C = -1 \),

\[
H_{TI} = H_{+1} \oplus H_{-1}.
\]  

(5.1)

Therefore the Chern number for a \( \mathbb{Z}_2 \) topological insulator is zero \([15]\) which guarantees that it allows exponentially localized WFs (\( C \) is additive over bands so in this case we have \( C = 1 - 1 = 0 \)).

However, WFs of \( \mathbb{Z}_2 \) topological insulators do not preserve time reversal (TR) symmetry \([115]\) even though the underlying Hamiltonian \( H_{TI} \) itself is time-reversal symmetric. This
can be seen by realizing that constructing TR-preserving WFs would be equivalent to constructing WFs individually for the bands given by $H_{+1}$ and $H_{-1}$ separately, which is not possible as bands with non-zero $C$ don’t have exponentially localized WFs. Therefore, the only possibility for constructing a smooth gauge of a compound system $(H_T I)$ is to break TR symmetry in the gauge by mixing eigenstates of $H_{+1}$ with those of $H_{-1}$.

5.1.1 Constructing Wannier functions from a guess

Generalized WFs [77] are obtained as the Fourier transform of the Bloch states $\psi_{mk}$ (here we consider the case of three dimensions)

$$|R_n\rangle = \frac{V}{(2\pi)^3} \int dke^{-ik\cdot R} \sum_m u_{mn}^{(k)} |\psi_{mk}\rangle,$$

where $u_{mn}^{(k)}$ is an arbitrary unitary matrix that mixes different bands for a given $k$-point, $R$ is a translation vector, and $n$ is an integer running over the number of bands considered. This gauge freedom can be used to construct WFs with minimal possible spatial extent. These so-called maximally localized Wannier functions (MLWFs) minimize the spread functional $\Omega$,

$$\Omega = \sum_n \left( \langle r^2 \rangle_n - \langle r \rangle^2_n \right),$$

with

$$\langle r^2 \rangle_n = \langle 0n | r^2 | 0n \rangle,$$

$$\langle r \rangle_n = \langle 0n | r | 0n \rangle.$$

Due to the properties of the Fourier transform, localization of Wannier function $|0n\rangle$ in real space is equivalent to the smoothness of Bloch states $\sum_m u_{mn}^{(k)} |\psi_{mk}\rangle$ in $k$-space.

Within the standard approach [77] MLWFs are constructed for a set of $N$ composite bands by first guessing a set of $N$ localized orbitals $g_n(r)$ that are close to the $N$ target Wannier functions,

$$|g_n\rangle \approx |0n\rangle.$$

Now given a set of Bloch states $|\psi_{mk}\rangle$ that are potentially not smooth in $k$-space (or equivalently its WFs are not localized) one can try smoothing them by first projecting them into these guess orbitals $g_n$

$$a_{mn}^{(k)} = \langle \psi_{mk} | g_n \rangle$$

and then constructing the unitary matrices $u^{(k)}$ via the L"{o}wdin orthonormalization procedure [67],

$$u^{(k)} = a^{(k)} \left[ a^{(k)\dagger} a^{(k)} \right]^{-1/2}.$$

If the overlap matrix appearing above under the inverse square root

$$s^{(k)} \equiv a^{(k)\dagger} a^{(k)}$$

(5.9)
has large singular values then this procedure is well-defined and matrices $u^{(k)}$ constructed in this manner can be used to rotate Bloch states into a smooth gauge $u^{(k)} | \psi \rangle$.

It is trivial to show that if $|g_n\rangle$ are MLWFs, that singular values of $s^{(k)}$ all equal 1 and this procedure gives back rotated Bloch states $u^{(k)} | \psi \rangle$ that correspond to the original MLWFs. Therefore, one can hope that if the $|g_n\rangle$ are at least close to MLWFs the resulting rotated Bloch states again correspond to localized—but not necessarily maximally localized—WFs. Given this starting point one can use procedure from Ref. [77] to arrive at MLWFs if needed.

On the other hand if the $|g_n\rangle$ are not close to MLWFs then the resulting rotated Bloch states need not be smooth. This is easily seen in the case of a single band. In this case if the orbital $|g\rangle$ fails to capture the character of the Bloch state $|\psi_k\rangle$ for some $k$-point then the complex number $a^{(k)}$ will have a small norm at that $k$. If the norm of $a^{(k)}$ is exactly zero for some $k$ then the procedure involving a negative square root in Eq. (5.8) is ill-defined as it involves division by zero. However, if the norm of $a^{(k)}$ is small but non-zero then the procedure is potentially numerically unstable as small noise in $a^{(k)}$ might get amplified when taking the negative square root of $s = a^{(k)} a^{(k \dagger)}$. This analysis generalizes to the case of multiple occupied bands ($N > 1$) in the following way. If one of the rotated $N$ Bloch states are not captured well by guess $|g_n\rangle$ then the overlap matrix (which is now a full $N \times N$ matrix) will have one small singular value and again the process of taking the negative square root of $s^{(k)}$ in Eq. (5.8) will be ill-defined or unstable.

In fact this is precisely the way in which the Löwdin procedure fails if one tries to apply it to the case of a Chern insulator [124]. For any localized trial orbital $|g_n\rangle$ the overlap matrix $s^{(k)}$ for a Chern insulator will not have at least one zero singular value at least at one point in the Brillouin zone. This will also happen in the case of a $Z_2$ topological insulators if one chooses trial orbitals that form a time-reversal symmetric pair. In Ref. [115] it was recognized that projecting trial orbitals that break TR symmetry is necessary to ensure that all singular values of $s^{(k)}$ are nonzero everywhere in the Brillouin zone. In practice this approach still requires an initial guess of orbitals that approximate the target WF. This guessing is somewhat harder than in the case of a non-topological insulator since it must break TR and potentially some other crystalline symmetries. In Ref. [115] this was achieved for a tight-binding model by an educated guess of trial orbitals based on the orbital character of the bands at the band inversion points and symmetries present in the model.

Another approach for constructing WF for $Z_2$ topological insulators was introduced in Ref. [135]. This approach relies on constructing a smooth gauge in a closely related non-topological insulator phase and then transporting that gauge to the TI of interest by following a path in the parameter space that explicitly breaks time-reversal symmetry (and potentially other symmetries such as inversion). This parameter space has to be chosen for each system at hand by adding terms to the Hamiltonian that break TR (and potentially crystalline) symmetry and keep the electron band gap open.

In the following sections we will present a method that can automatically construct WF for topologically nontrivial insulators.
5.2 Our approach, the OPFM

In the previous chapter we introduced the optimized projection functions method that allows automatic construction of MLWFs. We will now give a brief review of the optimized projection functions method (OPFM) [85] and then discuss why this approach is suitable for constructing WFs for $Z_2$ topological insulators.

As opposed to the standard approach, which requires $N$ trial orbitals for $N$ composite electron bands, in the OPFM one selects a larger set of $M > N$ orbitals $h_i(r)$ that approximately span the space of $N$ Wannier functions in a home cell,

$$\text{Span}(|h_i\rangle) \supseteq \text{Span}(|0n\rangle).$$

(5.10)

This can easily be achieved by including in $\{ h \}$ valence atomic orbitals.

Given a set of projection orbitals $\{ h \}$, we use the OPFM to find a semiunitary $M \times N$ matrix $W$ such that the $N$ orbitals

$$|\tilde{g}_j\rangle = \sum_{i=1}^{M} W_{ij} |h_i\rangle$$

(5.11)

are as close as possible to localized WFs. Given the functions $\tilde{g}_j$ one can construct the smooth gauge by first expanding the original functions into Bloch states,

$$A^{(k)}_{mn} = \langle \psi_{n(k)} | h_i \rangle$$

(5.12)

and then rotating them into the optimal subset,

$$\epsilon^{(k)}_{mn} = A^{(k)} W$$

(5.13)

which can then be used in the L"owdin procedure.

Now we discuss why the OPFM is suitable for constructing WFs in topological insulators. In $Z_2$ topological insulators, the spin-orbit interaction induces a band-inversion between states of different orbital character. For example, in the case of Bi$_2$Se$_3$ the topologically nontrivial state is induced by a band inversion at the $\Gamma$ point between Se and Bi states. Therefore one can expect that the MLWFs for the occupied bands in this system will contain a mixture of both Se and Bi states. Guessing such a mixture is nontrivial for several reasons. First, as we will show later, this mixture includes complex imaginary components. Second, the mixture typically contains a contribution from more than two atoms. Third, the mixture must break all relevant symmetries which enforce the topologically nontrivial state. However, our OPFM can find this mixture since $\{ h \}$ in the case of Bi$_2$Se$_3$ can include both Se and Bi atomic orbitals as $M$, the number of orbitals in set $\{ h \}$, can be larger than the number of electron bands $N$.

In the following sections we adopt the notation we introduced in Ref. [85] where square $N \times N$ matrices are represented by lowercase letters (e.g. $a^{(k)}$ and $u^{(k)}$), and larger rectangular $N \times M$, $M \times N$, or square $M \times M$ matrices are represented by uppercase letters (e.g. $A^{(k)}$ and $W$).
5.2.1 Selecting the set \{h\}

Now we will discuss a choice of orbitals \(h_i\) that satisfy the condition given in Eq. (5.10). Mathematically speaking, without knowing anything about chemical bonding in the insulator of interest, one would have to include in set \{h\} all atomic orbitals on all atoms in the crystal to guarantee a complete basis for expansion of the WFs. Luckily, in an ionic or a covalent insulator it is enough to choose set \{h\} to include only valence atomic orbitals, as they are typically the ones forming atomic bonds. In addition, since WFs are typically exponentially localized it is enough to choose orbitals in the home cell, and possibly few atoms in the neighboring unit cells. (For example, as discussed in Ref. [85] in the case of cubic silicon one has to include in set \{h\} atomic orbitals centered on two atoms in the basis as well as three neighboring atoms, so that, for each of the four Si-Si bonds, both Si atoms forming a particular bond are included in the set \{h\}.)

However, bonding in the case of \(Z_2\) TIs is more involved than in a typical insulator. As we will show in Sec. 5.3, presence of spin-orbit induced band inversion induces an intricate bonding network so that some WFs extends over more than two atoms (as in the case of the bonds in silicon) and thus one needs to use a somewhat larger set \{h\} than in a conventional covalent material. However, in all cases we tested, it was enough to include in \{h\} orbitals in the home-cell along with the orbitals in a single neighboring cell.

5.2.2 Finding the matrix \(W\)

In Ref. [85] the problem of finding \(W\) that minimizes the WF spread \(\Omega\) was reduced to minimizing the Lagrangian

\[
\mathcal{L}(W, \lambda) = \bar{\Omega}_{I,OD}(W) + \lambda w \sum_{k} \sum_{i=1}^{N} \left| \left[ W^\dagger S^{(k)} W \right]_{ii} \right|^2,
\]

where we define the large overlap matrix

\[
S^{(k)} \equiv A^{(k)\dagger} A^{(k)}.
\]

The first term in Eq. (5.14) approximates the sum of the invariant and offdiagonal parts of the spread \(\Omega\). However, this approximation is valid only when rotated overlap matrix \(W^\dagger S^{(k)} W\) is close to the identity matrix (see discussion in Ref. [85]). For simply bonded insulators this condition is enforced by the second term in Eq. (5.14).

While strictly speaking the entire matrix \(W^\dagger S^{(k)} W\) should be close to the identity matrix, the second term in Eq. (5.14) only penalizes the deviation of diagonal elements of \(W^\dagger S^{(k)} W\) away from 1. This simplification is adequate for the case of simply bonded insulators where only a small number of atoms are needed to span the space of WFs centered in the home cell. However, for the case of \(Z_2\) TIs, one needs to use a somewhat larger set \{h\} and
this simplification is insufficient since now some contributions to a WF could potentially be duplicated by more than one element in the set \( \{h\} \).

Therefore, in the following, we will construct \( W \) by minimizing the following Lagrangian that penalizes the offdiagonal elements of \( W^\dagger S^{(k)} W \) as well,

\[
\mathcal{L}(W, \lambda) = \bar{\Omega}_{\text{I,OD}}(W) + \lambda \sum_k \sum_{i=1}^N \sum_{j=1}^N |[W^\dagger (S^{(k)} - I_M) W]_{ij}|^2.
\] (5.16)

We describe the algorithm to minimize such Lagrangian in App. A.2.

We confirmed that with this extended Lagrangian WF spreads for conventional insulators investigated in Ref. [85] are unaffected. For example, the initial spread in the case of cubic silicon changes by less than 0.2% when offdiagonal elements are included in \( \mathcal{L} \). The only difference with respect to Ref. [85] is that now—with the Lagrangian from Eq. (5.16)—it does not matter whether some WFs can be represented by orbitals from \( \{h\} \) in more than one way, at a small additional cost in the computational time. For example, in the case of the cubic silicon if we include 6 (instead of 3) neighboring atoms in the set \( \{h\} \) so that three out of four Si-Si bonds can be represented in duplicated ways, the total initial spread is changed only by 0.6%.

We note here that it is numerically straightforward to construct an arbitrary set \( \{h\} \) given a set of orbitals \( h_i \) in the home cell. For a particular orbital \( h_i(r) \) on the basis atom, a projection onto another orbital given by the same orbital but translated by lattice vector \( R \) is simply,

\[
\langle \psi_{nk}|h_i(r-R)\rangle = e^{-ik\cdot R} \langle \psi_{nk}|h_i(r)\rangle
\] (5.17)
by the virtue of Bloch’s theorem.

### 5.3 Examples

In the following subsections, we apply the optimized projection functions method to three examples of \( \mathbb{Z}_2 \) topological insulators. The first is the Kane-Mele model [53], which is a two-dimensional tight-binding model on the honeycomb structure. The second is a three-dimensional tight-binding model of a strong topological insulator that was introduced in Ref. [100]. The third example is a realistic case of a three-dimensional strong topological insulator (Bi\(_2\)Se\(_3\)) as calculated within the density functional theory approach. [138, 142]

#### 5.3.1 Two-dimensional model

The Kane-Mele model is a two-dimensional model of a \( \mathbb{Z}_2 \) topological insulator. It contains four electron bands, two of which are considered to be occupied. It is defined on a honeycomb structure that can be described in terms of the hexagonal lattice with primitive
lattice vectors $\mathbf{a}_{1,2} = \frac{a}{2}(\sqrt{3}\hat{y} \pm \hat{x})$, and a basis of two sites, $A$ and $B$, located at $\mathbf{r}_A = a\hat{y}/\sqrt{3}$ and $\mathbf{r}_B = 2a\hat{y}/\sqrt{3}$, respectively. In what follows, we choose $a = 1\ \text{Å}$ for convenience.

The Kane-Mele Hamiltonian is

$$ H = t \sum_{\langle ij \rangle} c_i^\dagger c_j + i\lambda_{\text{SO}} \sum_{\langle\langle ij \rangle\rangle} \nu_{ij} c_i^\dagger s^z c_j + i\lambda_{\text{R}} \sum_{\langle ij \rangle} c_i^\dagger (s \times \hat{d}_{ij}) c_j + \lambda_v \sum_i \xi_i c_i^\dagger c_i. \quad (5.18) $$

We suppressed spin indices on the electron creation and annihilation operators. Symbol $\langle ij \rangle$ indicates a sum over nearest neighbors and $\langle\langle ij \rangle\rangle$ indicates a sum over next-nearest neighbors. The first term in the Hamiltonian is the nearest neighbor hopping, with hopping strength $t$ that we set equal to 1 for convenience (i.e., all energies are in units of $t$). The second term describes spin-dependent second nearest neighbor hopping, which emulates a spin-orbit interaction. Here, $\nu_{ij}$ takes on the value $\pm 1$ depending on the sign of $(\hat{d}_1 \times \hat{d}_2)_z$, where $\hat{d}_1$ and $\hat{d}_2$ are the unit vectors along the bonds traversed as the electron hops from site $j$ to $i$, and $s^z$ is the Pauli spin matrix. The third term describes nearest neighbor Rashba coupling, where $\hat{d}_{ij}$ is the unit vector along the bond from $j$ to $i$. Lastly, the fourth term introduces a staggered on-site potential ($\xi_i = \pm 1$) between the $A$ and $B$ sublattices; we choose $\xi_i$ so that the on-site potential is negative on the $B$ sublattice and the occupied bands in the normal phase have dominant $B$ character. In the following we set the staggered on-site term $\lambda_v = 1$ and the Rashba term $\lambda_{\text{R}} = 0.5$. Increasing the strength of the spin-orbit term $\lambda_{\text{SO}}$ tunes the model from describing the normal to the topological insulator phase, with the transition at $\lambda_{\text{SO}} \approx 0.27$. For calculations in the topological phase we use $\lambda_{\text{SO}} = 0.6$.

The Kane-Mele model is solved using the PythTB\(^a\) package with a basis of two orbitals per site, one each for spin-up and spin-down,

$$ |A; \uparrow_z\rangle, \quad |A; \downarrow_z\rangle, \quad |B; \uparrow_z\rangle, \quad |B; \downarrow_z\rangle. \quad (5.19) $$

In Figure 5.1 we plot the band structures in both the normal phase ($\mathbb{Z}_2$ even) and the topological phase ($\mathbb{Z}_2$ odd). The bands are colored according to the character of the state, with red corresponding to a state of $B$-orbital character and blue corresponding to $A$-orbital character, and gray indicating a mixture. The thickness of the line indicate the spin of the state along the $z$-axis, with thicker corresponding to mostly spin-up and thinner line corresponding to mostly spin-down; an intermediate thickness indicates mixed spin state due to the Rashba-like term. In the topological phase, there is a clear inversion of the character of the states near the $K$ and $K'$ points.

**Selecting the set $\{h\}$**

Our results of the OPFM applied to the case of a Kane-Mele model are shown at the top of Table 5.1 along with the result from previous work [115]. In the previous work the following guess orbitals were used to construct WFs for the Kane-Mele model in the topological phase:

\(^a\)http://physics.rutgers.edu/pythtb/
Figure 5.1: Band structures of the Kane-Mele model in the normal phase (top) and topological phase (bottom). The part of the bands colored red correspond to $B$-orbital character, while blue corresponds to $A$-orbital character. The thickness of the line corresponds to the spin component, with thicker indicating mostly spin-up, and thinner indicating spin-down. The zero of energy is set to the middle of the gap.
\(|A; \uparrow_x\rangle\) and \(|B; \downarrow_x\rangle\). Note that spins here point in plane \((x)\) while the basis functions have spins pointing perpendicular to the plane \((z)\).

To construct localized WFs for the Kane-Mele model using the OPFM we consider several different sets \(\{h\}\) of basis functions. The smallest set consisted of the four orbitals on the two atoms in the home cell defined in Eq. (5.19). As expected, this small set is unable to capture the extended nature of the WF in the \(\mathbb{Z}_2\) topological insulator. The fact that this set is too small is numerically indicated with a small minimal singular value of \(s^{(k)} = W^{\dagger} S^{(k)} W\) (its value is only 0.03, not given in Table 5.1) which then results in an ill-defined Löwdin orthonormalization procedure. This procedure produces a gauge with a WF spread \(\Omega^0 = 0.333\), significantly higher than that of a MLWF \((\Omega^{GM} = 0.189)\). If we try smoothening this gauge further with the Marzari-Vanderbilt procedure [77] it remains stuck in a local minimum as the spread is only slightly reduced to \(\Omega^{GM} = 0.319\).

Since the figure of merit (i.e. minimal singular value of \(s^{(k)}\)) was small for this set \(\{h\}\) we decided to use the OPFM with larger sets \(\{h\}\). The next set we considered, labeled \(\{0, 1^{st}\}\), includes—in addition to orbitals in the home cell—orbitals on their four first-neighboring atoms. An even larger set we tried \(\{0, 1^{st}, 2^{nd}\}\) includes both first and second nearest neighboring atoms.

As soon as we include first or second neighbor atoms into the set \(\{h\}\) the resulting minimal singular value of \(s^{(k)}\) increases from 0.03 to 0.40 and 0.71, respectively for the two sets, and the resulting spread \(\Omega^0\) decreases. Final spread \(\Omega^{GM}\) agrees with previous result [115] up to numerical precision.

Table 5.1 contains also the average distance between the overlap matrix \(s^{(k)}\) and the identity matrix. However, that averaged quantity masks the fact that the overlap matrix is typically different from an identity matrix only in a small part of the Brillouin zone where inversion occurs (see band characters near \(K\) and \(K'\) points in Fig. 5.1). Therefore for the purpose of presentation we give in Fig. 5.2 the distribution of \(|(s^{(k)} - I)_{ij}|^2\) over all \(k\)-points and all its matrix elements \(ij\). Note that this quantity is the same as the second term in Eq. (5.16).

As can be seen from Fig. 5.2 in all cases distance of \(s^{(k)}\) from identity is small for nearly all \(k\)-points (note that the vertical scale is logarithmic). However, in the case of the guess orbitals from Ref. [115] there is a fraction of \(k\) points for which matrix elements of \(s^{(k)}\) are quite far away from identity matrix (up to 0.7). These \(k\)-points correspond to the small part of the Brillouin zone with inverted bands. However, singular values of \(s^{(k)}\) are large enough (smallest one is 0.11) so that the Löwdin procedure is well behaved even for the guess orbital from Ref. [115]. In the case of the OPFM the deviation of \(s^{(k)}\) from identity is significantly smaller (the maximum value is only 0.1 for the cluster \(\{0, 1^{st}, 2^{nd}\}\)).

We also confirmed that OPFM gives automatically good spread as one varies \(\lambda_{SO}\) through the transition from the topological all the way to the normal phase.

For this two-dimensional model we constructed the projection matrices on a \(15^2\) \(k\)-point grid. We find that the optimal value for the Lagrange multiplier \((\lambda)\) is 0.03. We also note that while in the case of normal insulators studied previously one can often initialize \(W\) with
Table 5.1: Results of the OPFM applied to the examples in Sec. 5.3. The first column gives the square moduli of $s^{(k)} - I_N$ averaged over $k$-points and matrix elements. The second column lists the minimal singular value of $s^{(k)}$ over all $k$-points. The third column show the spread $\Omega^0$ after Löwdin procedure while the fourth column shows the spread $\Omega^{GM}$ when Löwdin procedure is followed up by the Marzari-Vanderbilt [77] procedure. For three-dimensional cases, the fifth column gives the value for the Chern-Simons $\theta$ term. See text for a description of the sets $\{h\}$ used in the OPFM.

|                           | Average $| (s^{(k)} - I_N)_{ij} |^2$ | Min. sing $(s^{(k)})$ | Spread $(\AA^2)$ | $\frac{\Omega^0}{\Omega^{GM}}$ | Chern-Simons $\theta$ |
|---------------------------|-----------------------------------|------------------------|-------------------|-----------------|------------------|
| **Two-dimensional model** |                                    |                        |                   |                 |                  |
| Previous work, Ref. [115] | 0.148                             | 0.11                   | 0.212             | 0.189           |                  |
| OPFM using set $\{0, 1^{st}\}$ | 0.017                             | 0.40                   | 0.244             | 0.189           |                  |
| OPFM using set $\{0, 1^{st}, 2^{nd}\}$ | 0.006                             | 0.71                   | 0.207             | 0.189           |                  |
| **Three-dimensional model** |                                    |                        |                   |                 |                  |
| OPFM using set $\{0, 1\}$   | 0.0184                            | 0.0472                 | 0.142             | 0.135           | 0.96$\pi$       |
| OPFM using set $\{0, 1, 2, 3\}$ | 0.0133                            | 0.1022                 | 0.142             | 0.135           | 0.96$\pi$       |
| **Density functional theory, Bi$_2$Se$_3$** | |                        |                   |                 |                  |
| Previous work, Ref. [21]    | 0.0068                            | 0.0003                 | 109.80            | 95.84           | 0.32$\pi$       |
| Previous work, Ref. [135]   | 0.0057                            | 0.0002                 | 126.12            | 95.85           | 0.35$\pi$       |
| OPFM using set $\{0, 1\}$   | 0.0026                            | 0.0001                 | 133.43            | 95.84           | 0.34$\pi$       |
| OPFM using set $\{0, 1, 2, 3\}$ | 0.0017                            | 0.0240                 | 310.29            | 95.83           | 0.34$\pi$       |
Figure 5.2: Histograms of the square moduli of the elements of $s^{(k)} - I$ for guess orbitals from Ref. [115] (top panel) and two sets $\{h\}$ within the OPFM (middle and bottom panel).

The identity matrix, in the case of $\mathbb{Z}_2$ TIs we sometimes need to start off the minimization from a random matrix so that the initial $W$ breaks all symmetries. (We confirmed that, in the case of normal insulators studied in Ref. [85], starting minimization procedure from a random matrix does not affect the final spread. For example, in the case of cubic silicon the total spread is unaffected within numerical precision.)

Analysis of the WF

As expected, we find that the WF in the $\mathbb{Z}_2$ topological insulator case extends well beyond the home cell. This finding is expected since band inversion usually occurs over a small region in $k$-space. To analyze the extent of the Wannier function in more detail we show in Fig. 5.3 the MLWFs for the Kane-Mele model in the real space. We write the WF amplitude on a particular sublattice $j = \{A, B; R\}$ in the crystal as,

$$\begin{align*}
(\alpha_1 + i\alpha_2) |j \uparrow\rangle + (\beta_1 + i\beta_2) |j \downarrow\rangle.
\end{align*}$$

(5.20)
These amplitudes can be computed from the projections of Bloch states into basis functions $(A^{(k)})$ and the smooth gauge for the WFs $u^{(k)}$ as,

$$\langle h_j | R_n \rangle = \langle h_j | \sum_{m,k} u^{(k)}_{mn} | \psi_{mk} \rangle = \sum_k A^{(k)\dagger} u^{(k)}$$

Figure 5.3 shows the amplitudes $\alpha_1, \alpha_2, \beta_1,$ and $\beta_2$ on each site $j$ in the crystal for both occupied bands (labelled #1 and #2). The size of the circles are proportional to the absolute value of the magnitude of $\alpha_1, \alpha_2, \beta_1, \beta_2$ while color denotes their sign (red for positive and blue for negative). The cross symbols ($\times$) denote $A$ sites while plus symbols ($+$) denote $B$ sites.

In the normal phase, both WFs are centered near the $B$ site in the home cell, with components of opposite sign on the first nearest neighbors, and small components on second nearest neighbors.

The topological phase has WFs that are centered near different sites ($A$ and $B$ in the home cell), with both being a mixture of spin-up and spin-down. Most importantly, in the topological phase the WF amplitude extends well beyond the home cell into the first and second nearest neighbors. The beyond-home-cell component of the WF in addition has a significant imaginary part.

Therefore, from here we confirm once again that the set $\{h\}$ in the case of a $\mathbb{Z}_2$ topological insulator must include orbitals beyond those in the home cell.

### 5.3.2 Three-dimensional model

We now turn to the model of a three-dimensional strong $\mathbb{Z}_2$ topological insulator. A simple model of such an insulator is given in Ref. [100] by constructing a higher dimensional insulator with a non-zero second Chern number and then restricting the model to three dimensions. Similarly as in the case of a Kane-Mele model this model consists of four orbitals in the basis and two occupied electronic bands. One difference with respect to the Kane-Mele model is that the only hopping terms in the model are either between orbitals in the same unit cell, or between the first neighboring cells. The Kane-Mele model includes hopping to second nearest neighboring cells as well.

As is done for the Kane-Mele model, here we performed OPFM with sets $\{h\}$ of increasing size. Once again we find that as soon as a neighboring cell is included in the set $\{h\}$, the OPFM procedure produces a smooth gauge. Similarly, as in the case of Kane-Mele model, we find that with larger sets $\{h\}$ the minimal singular value of the overlap matrix is increased. Here we adopt a notation by which the set $\{0, 1\}$ corresponds to orbitals in the home cell as well as neighboring cells translated along the first lattice vector. Similarly, the set $\{0, 1, 2, 3\}$ corresponds to orbitals in the home cell as well as those translated along all three lattice vectors.

In addition to the quantities reported in Table 5.1 for the Kane-Mele model, here we also report the value of the Chern-Simons orbital magneto-electric coupling $\theta$. [100, 28] The $\theta$ term takes on the value $0$ or $\pi$ (modulo $2\pi$) in the normal and topological phase, respectively.
Figure 5.3: Plot of the Wannier functions obtained via the OPFM on set \{0, 1^{st}, 2^{nd}\}, for the two occupied bands of the Kane-Mele model in the normal insulator phase (top) and the topological insulator phase (bottom). The crosses (×) indicate the A sites, while the plus signs (+) indicate the B sites. Red circles correspond to a component that is positive and blue circles correspond to negative. The size (area) of the circle is proportional to the magnitude of the component $\alpha_1, \alpha_2, \beta_1, \beta_2$. Here, the WFs plotted are for a Kane-Mele model with $\lambda_v = 1$ and $\lambda_R = 0$. 
However, these values would be obtained only in the limit of infinitely dense \( k \)-meshes, as the discretized expression for \( \theta \) we used is not gauge invariant (gauge invariant discrete form of \( \theta \) is unknown, as far as we are aware). On a finite mesh the calculated value of \( \theta \) in the topological phase is typically smaller than \( \pi \) and it converges very slowly to \( \pi \) as the \( k \)-mesh gets denser. We used a 20\(^3\) mesh of \( k \) points for this calculation. Here the value of the Lagrange multiplier (\( \lambda \)) is 1.

The expressions for \( \theta \) in terms of WFs given in Ref. [21] clearly shows that \( \theta \) must be 0 if the WFs are purely real. Therefore, just as in the case of two-dimensional model, the WFs in the three-dimensional topological insulator must contain large imaginary components so that \( \theta \) can be non-zero (\( \theta = \pi \) modulo \( 2\pi \) to be precise).

Table 5.1 contains some of the results of the OPFM applied to the three-dimensional model. In the normal phase of that model (not shown in the Table 5.1) our OPFM finds projection functions that well approximate the WFs even when we use the set \( \{0\} \) with orbitals only in the home cell. The minimal singular value of the \( s^{(k)} \) matrix is close to identity (0.92) and the spread after the Löwdin procedure agrees with the spread at the global minimum within the first four non-zero significant digits.

In the topological phase of the model, the set \( \{0\} \) results in an overlap matrix with a very small singular value \( (10^{-29}) \) but with an inclusion of a larger set \( \{h\} \) all figures of merit improve, as in the case of the Kane-Mele model. Therefore we conclude that even in the case of three-dimensional models the WF in a \( \mathbb{Z}_2 \) topological insulator extends well beyond the home cell.

We note here that for relatively small sets such as \( \{0,1\} \) the minimal singular value is rather small (0.0472); however, the resulting final spread is very close to the spread at the global minimum and the value of \( \theta \) is close to \( \pi \). The minimal singular value increases to 0.1022 in the set \( \{0,1,2,3\} \). We also tried using an even larger set where orbitals are translated both in positive and negative direction of the lattice vector \( \{0,1,\bar{1},2,\bar{2},3,\bar{3}\} \) and we find that the minimal singular value increases to 0.514. Despite having a larger minimal singular value, \( \theta \) and \( \Omega_{GM} \) computed from this set are up to numerical precision equal to those obtained using a much smaller set, \( \{0,1\} \). Therefore we conclude that the set \( \{0,1\} \) is adequate for this system even though it yields a somewhat small minimal singular value of the overlap matrix (0.0472).

To further test our method we generalized the tight-binding model from Ref. [100] to higher number of bands. The model from Ref. [100] was constructed from \( 2n \)-dimensional Clifford algebra where \( n = 2 \). If we use \( n = 3 \) or \( n = 4 \) algebras and again perform dimensional reduction to three dimensions, the resulting tight-binding model will have \( 2^{n-1} \) occupied bands out of \( 2^n \) bands. This means that in the \( n = 3 \) case we have 8-band model with 4 occupied bands, while with \( n = 4 \) we have 16-band model with 8 occupied bands. Applying the OPFM to these models using the set \( \{0,1,\bar{1},2,\bar{2},3,\bar{3}\} \) again produces a smooth gauge. While the minimal singular value in the 4-band model \( (n = 2) \) discussed earlier is 0.514, with 8-band model \( (n = 3) \) it is 0.356, and with 16-band model \( (n = 4) \) it is 0.360.
5.3.3 Density functional theory, Bi$_2$Se$_3$

We now turn from the tight-binding models to some realistic calculations based on density functional theory. As an example of a prototypical strong 3D TI we use Bi$_2$Se$_3$ [138, 142]. Its crystal structure is described by a rhombohedral lattice within the $D_{3d}^5$ space group. The material is made up of units of quintuple layers of Bi and Se. Each of the five layers in the quintuple forms a hexagonal sheet in plane. The topological phase is realized due to the strong spin-orbit coupling causing a band inversion of Se $p$ and Bi $p$ character around the Γ point [142]. This inversion is evident in Fig. 5.4.

To construct localized WFs for Bi$_2$Se$_3$, we first perform fully relativistic density functional theory calculations with the Quantum ESPRESSO package [35]. The ground state properties are obtained using a $6^3$ $k$-point grid and a kinetic energy cutoff of 60 Ry. The projection matrices $A_{mn}^{(k)}$ are obtained on a $12^3$ $k$-point grid by projecting the top 28 valence bands into atomic Bi and Se $s$ and $p$ orbitals. We use Eq. (5.17) to construct projections into orbitals translated by a lattice vector. In all calculations for Bi$_2$Se$_3$ we used the value $\lambda = 1$ for the Lagrange multiplier.

Again we consider several sets $\{h\}$ generated by translating the basis atoms by different lattice vectors, as using only orbitals in the home cell once again gave very small ($10^{-6}$) minimal singular value of $s^{(k)}$. As soon as a neighboring cells are included in the set $\{h\}$ the minimal singular values increase as well as the spread $\Omega^G$. We used the same translation vectors as in the three-dimensional model case: $\{0, 1\}$ and $\{0, 1, 2, 3\}$. Here again 0 represents orbitals in the home cell while non-zero integers 1, 2, 3 represent translations along the three equivalent rhombohedral lattice vectors. For completeness, we note that we chose as basis atoms those for which the reduced coordinates in the rhombohedral frame are as small as possible (between $-\frac{1}{2}$ and $\frac{1}{2}$).

The results of OPFM in the case of Bi$_2$Se$_3$ are shown in Table 5.1 along with the results from previous work. One of the previous works [21] guessed WFs by trying out various initial projections that break symmetries while the other [135] found it by constructing a path in parameter space that breaks time-reversal and inversion symmetry. In both earlier works the Bloch states were projected into hydrogen-like orbitals. We find very good agreement in both $\theta$ and $\Omega^{GM}$ between our approach and two earlier works. The computed value for $\theta$ in all three cases is close to $\theta \approx 0.3\pi$ since we used a relatively small $k$-point grid (it was $12^3$). With larger $k$-grids $\theta$ converges towards $\pi$.

We note here that the minimal singular value of $s^{(k)}$ using relatively large set $\{0, 1, 2, 3\}$ in OPFM is still somewhat small (0.0240) even though it is two orders of magnitude larger than those in previous works [21, 135]. Nevertheless $\Omega^{GM}$ agrees well with each other in all cases and the value of $\theta$ is what is expected for a three-dimensional strong $\mathbb{Z}_2$ topological insulator. (Some of the difference between the minimal singular values in these approaches might originate from use of hydrogen-like projection functions in Refs. [21, 135])
Figure 5.4: Band structure of Bi$_2$Se$_3$. The part of the bands colored red correspond to Se $p$ character, while blue corresponds to Bi $p$ character. The zero of energy is set to the middle of the gap.

5.4 Outlook

In this chapter, we described a procedure for automated construction of maximally localized Wannier functions for topologically nontrivial set of bands. We expect that this method can be applied to any topological insulator, either protected by time-reversal symmetry, or by crystalline symmetry, as long as there exists a localized representation, i.e. as long as the first Chern numbers are all zero. Similarly, we expect that this method could be applied to topologically nontrivial bands not only of electrons, but also of phonons, photons, cold atoms, or other particles.
Chapter 6

Fermi surface properties of noble metals

6.1 Introduction

Copper, silver, and gold are noble metals with broad application in electronics, power generation, catalysis, and plasmonics. They have attracted interest since the early days of solid state theory as their electronic structure deviates from the free-electron model that applies to the alkali metals. The Fermi surface of noble metals is not spherical as in free-electron theory, but is deformed due to the proximity of the \(d\) bands to the free-electron-like \(sp\) band [114, 104, 52, 41]. Electron scattering processes at the Fermi surface are of particular relevance for noble metal applications as they regulate charge and heat transport [145]. At room temperature in relatively pure metals, scattering with phonons [113, 70] is dominant, while scattering with defects and impurities is important at low temperature and in alloys or samples with low purity.

Transport in metals can be understood heuristically with the Drude theory [3], which assumes free electrons with a constant (that is, band- and \(k\)-independent) relaxation time. Even in noble metals, where important deviations are expected from the Drude theory, resistive losses and optical experiments are routinely interpreted using constant relaxation times [56, 51, 3]. More advanced models, such as state-of-the-art ab initio calculations of resistivity and other transport properties [99, 69], typically employ density functional theory (DFT) bandstructures combined with constant relaxation times inferred from experiment or estimated heuristically.

However, the relaxation time of an electron in a Bloch state depends, in general, on the band and crystal momentum \(k\), a point that has so far been often neglected. Accurate calculations of electron relaxation times are computationally costly as they require fine Brillouin zone (BZ) sampling [8, 9, 10], prompting adoption of simplified schemes that employ either a constant relaxation time [93] or an average scattering strength [110, 109]. An exception are recent transport calculations in two-dimensional materials [92, 62] accounting for the band
and \( k \) dependence of the relaxation times by leveraging the lower computational cost of BZ sampling in two dimensions.

As striking new experimental findings emerge in noble metals [57, 29, 91], predictive theories are needed to study electron scattering in these materials. For example, recent experiments from Kim et al. [57] show a remarkable resistivity drop upon doping a single crystal of Ag with a small fraction of Cu, at odds with the intuition that the resistivity should increase upon alloying due to enhanced impurity scattering. While the constant relaxation time approach may be too simplistic, at present it is not known whether more refined theories are necessary to improve quantitative understanding of the noble metals.

6.2 Computational Methods

Here we compute electron-phonon (\( e \)-ph) relaxation times and GW band velocities of Cu, Ag, and Au on fine \( k \)-point grids. We develop a linear interpolation approach to sample \( k \)-points directly on the Fermi surface, and use it to correlate the relaxation times and velocities to Fermi surface topology. We find that the \( e \)-ph relaxation times are far from constant on the Fermi surface. They vary by a factor of 2–3 between their minima and maxima, located at Fermi surface regions with different curvatures. The band velocity maps on the Fermi surface exhibit anti-correlation with the \( e \)-ph relaxation time maps. We investigate approximations to compute the resistivity of Cu, Ag and Au at room temperature, and find excellent agreement (within \( \sim \)10\%) with experiment only when ab initio relaxation times and GW band velocities are used. Our Fermi surface sampling approach dramatically speeds up the convergence of resistivity calculations, and may be extended to study heat transport and thermoelectric effects in materials.

We carry out DFT calculations at experimental lattice constants (Cu: 3.61 Å, Ag: 4.09 Å, Au: 4.08 Å) using the local density approximation (LDA) [95] and a plane-wave basis with the Quantum ESPRESSO code [35]. Scalar-relativistic norm-conserving pseudopotentials [128] (including semicore \( s \) and \( p \) states) are employed, using kinetic energy cutoffs of 240 Ry for Cu and 180 Ry for Ag and Au and shifted \( 10^3 \) \( k \)-point grids, to determine the ground state charge density. Spin-orbit effects are neglected. The quasiparticle energies are computed within the \( G_0W_0 \) and generalized plasmon pole approximations [44, 47, 65] using the BerkeleyGW package [24]. Our GW calculations are well converged using a 50 Ry cutoff for the dielectric matrix, which was evaluated on a \( 8^3 \) \( k \)-point grid for interband transitions (finite \( q \)) and a \( 16^3 \) \( k \)-point grid for intraband transitions (\( q \to 0 \)), and using \( \sim \)1000 unoccupied bands, consistent with previous studies [75, 119, 103].

The \( e \)-ph matrix elements are computed on fine \( k \)-point grids with a procedure detailed in Ref. [10]. Briefly, we use density functional perturbation theory [5] to determine the phonon dispersions and displacement patterns and \( e \)-ph coupling matrix elements on a coarse \( 4\times4\times4 \) \( q \)-point grid. The bandstructures, phonon dispersions, and \( e \)-ph matrix elements are interpolated using maximally localized Wannier functions [77, 117, 83] constructed from Bloch
Figure 6.1: Electron-phonon relaxation times (top panel), orbital character (middle panel), and band velocities (bottom panel) of states on the Fermi surface of Cu, Ag, and Au.
states on a $12 \times 12 \times 12$ \textit{k}-point grid. All quantities are interpolated using the EPW code [88, 36] with our recently developed approach [8, 9, 10]. The band velocities are calculated in the Wannier representation [140] using GW eigenvalues. For $e$-ph and resistivity calculations, we use the tetrahedron method [60, 68, 12] to obtain the Fermi surface as a collection of triangular facets, whose centroids constitute an ultra-fine \textit{k}-point grid employed in the calculations to sample the Fermi surface. Following Mahan [70], the $e$-ph transport relaxation scattering rates $\Gamma_{nk}$ (and their inverse, the relaxation times $\tau_{nk} = \Gamma_{nk}^{-1}$) are computed with perturbation theory, due to the interaction with phonons:

$$\Gamma_{nk} = \frac{2\pi}{\hbar} \sum_{mq\nu} |g_{nm,\nu}(k,q)|^2 (1 - \cos \theta_{k,k+q})$$

$$\times [(N_{\nu q} + f_{mk+q})\delta(\epsilon_{nk} - \omega_{\nu q} - \epsilon_{mk+q}) + (N_{\nu q} + 1 - f_{mk+q})\delta(\epsilon_{nk} + \omega_{\nu q} - \epsilon_{mk+q})]$$

(6.1)

where $g_{nm,\nu}(k,q)$ are \textit{e}–\textit{ph} coupling matrix elements for an electron in Bloch state $|n_k\rangle$ (with quasiparticle energy $\epsilon_{nk}$) that scatters into a Bloch state $|m_{k+q}\rangle$ (with quasiparticle energy $\epsilon_{mk+q}$) due to a phonon with polarization $\nu$, wavevector $q$, and frequency $\omega_{\nu q}$. Here, the $(1 - \cos \theta_{k,k+q})$ factor accounts for a forward-backward scattering asymmetry, given that $\theta_{k,k+q}$ is the scattering angle between the vectors $k$ and $k + q$. The two terms in square brackets correspond to phonon absorption and emission, respectively, and the temperature dependence stems from the occupation factors $f_{nk}$ and $N_{\nu q}$, for electrons and phonons respectively (in this work, both the electrons and the phonons are at 300 K). The conductivity tensor from Boltzmann transport theory within the relaxation-time approximation is expressed as the Fermi surface integral [145, 70]

$$\sigma_{\alpha\beta} = \frac{2}{(2\pi)^3} \frac{e^2}{\hbar} \sum_n \int dS_F \frac{v_{nk,\alpha} v_{nk,\beta}}{|v_{nk}|} \tau_{nk}$$

(6.2)

where $v_{nk,\alpha}$ are Cartesian components ($\alpha = x, y, z$) of the band velocities and $\tau_{nk}$ \textit{e}–\textit{ph} relaxation times. The area elements $dS_F$ correspond to the area of the triangular facets that make up the Fermi surface within our interpolation scheme. The resistivity $\rho$ is obtained by averaging the trace of the inverse of the conductivity tensor, i.e., $\rho = \text{Tr}[\sigma^{-1}]/3$. We note that Eq. (6.2) is valid at temperatures much lower than the Fermi temperature $E_F/k_B$, as is the case for noble metals at room temperature.

### 6.3 Results

The computed Fermi surfaces (see Fig. 6.1) exhibit well-known [114, 104, 52, 41] topological features, including open-orbit regions with upward curvature (known as “necks”) near the $L$ points of the BZ, spherical regions known as the “belly”, as well as flatter regions near the $K$ points between the necks, and “bulges” where the Fermi surface approaches the BZ edge at the $X$ points. The Fermi surface of Ag is noticeably different from those of Cu...
and Au, with much smaller necks and less pronounced bulges, and an overall more spherical shape due to the $d$ bands being lower in energy compared to Cu and Au.

Figure 6.1 shows maps of the $e$-ph relaxation times and GW band velocities calculated at $k$-points on the Fermi surface. For all three materials, we find large variations in the relaxation times, with the necks and bulges exhibiting shorter relaxation times than the belly and flat regions between the necks. The relaxation times vary the most in Cu and Au, with a factor of $\sim 3$ difference between the minima (located near the $L$ and $X$ points) and maxima (located near the $K$ point), while the variation in Ag is smaller, only a factor of $\sim 2$ difference between minima and maxima. The band velocities show an opposite trend, so that Fermi surface regions with longer (shorter) relaxation times are associated with smaller (larger) band velocities. Cu and Au exhibit moderate variations in the velocities, while Ag shows a large difference. The difference between the minimal velocities at the necks and the maximal velocities at the belly is a factor of $\sim 2$ in Cu and Au and a factor of $\sim 4$ in Ag.

We investigate the physical origin of the variation of the relaxation times and band velocities on the Fermi surface. By mapping the orbital character of the electronic states on the Fermi surface (see Fig. 6.1), we find a correlation between the character of the state and its relaxation time. Electronic states with a predominantly $sp$ character, as found in the necks and bulges of the Fermi surface, are associated with shorter relaxation times, while states with predominantly $d$ character exhibit longer relaxation times.

In Fig. 6.2, we explore quantitatively the origin of the anisotropic relaxation times for the case of Ag, by choosing initial states with $sp$ and $d$ character respectively, and computing the coupling matrix elements $|g|^2$ connecting these initial $k$-points to all other final $k$-points on the Fermi surface. These matrix elements are the ones entering the calculation of the relaxation times for each of the initial states within our approach [Eq. (6.2)], and we now seek to categorize them based on the character of the initial and final states in the $e$-ph scattering processes.

For an initial state near $X$ with dominant $sp$ character, the coupling, as measured by the sum over phonon polarizations of the matrix element between the initial and final states, $\sum_\nu |g_\nu|^2$, is strongest to $d$-like final states along the [110] directions. We find that coupling to other final states with $d$ character is in general weaker and strongly anisotropic, so that different groups of final $d$-like states exhibit $|g|^2$ values distributed in multiple sets (see Fig. 6.2). As the fraction of $sp$ character of the final state increases, the coupling decreases monotonically.

For an initial state near the $K$ point with dominant $d$ character, coupling is strong to one set of $d$-like states (along [110]), and weak to most other $d$-like states. Two distinct sets of $|g|^2$ values are found for such $d$-$d$ coupling (see Fig. 6.2). As the $sp$ character of the final state increases, different final states exhibit different trends: for some states the coupling becomes stronger, while for others it becomes weaker. Overall, the average $|g|^2$ is greater by a factor of 1.2 for the initial $sp$ state near $X$, thus explaining the higher scattering rate (which is proportional to $|g|^2$) and shorter relaxation time compared to $d$ states, as discussed above. While further investigation is needed to more completely understand the role of the character of the initial and final state and the perturbation potential, our results demon-
Figure 6.2: Electron-phonon coupling matrix elements for an initial electronic state of predominantly $sp$ character (top panel) and an initial state of predominantly $d$ character (bottom panel). Shown are the $e$-$ph$ coupling matrix elements (summed over phonon polarizations) versus the character of the final electronic state involved in the $e$-$ph$ scattering process. For each initial state, the average coupling matrix element is also given.

strate that the e-ph matrix elements are highly anisotropic and depend strongly both on the character of the initial and final state and on the wavevector connecting the initial and final state through the perturbation potential induced by the phonons. The combination of these effects results in the anisotropic relaxation times found here. The variation in the band velocities is easier to explain. It can be understood by considering the curvature of the Fermi surface: flatter regions have smaller velocities compared to more curved regions, given that the latter are associated to a higher gradient of the band at the Fermi energy.

To highlight the difference between the ab initio and empirical data, we compare in Table 6.1 Fermi-surface-averaged relaxation times $\langle \tau_{nk} \rangle$ and velocities $\langle v_{nk} \rangle$, obtained by averaging the ab initio data in Fig. 6.1, with values derived from the free-electron Drude model:

$$\tau_D = \frac{m^*}{ne^2\rho}, \quad v_D = \frac{\hbar(3\pi^2n)^{1/3}}{m^*}.$$  \hspace{1cm} (6.3)

where $\tau_D$ and $v_D$ are the empirical Drude relaxation time and free electron velocity, respectively. The ab initio average quantities $\langle X \rangle$ (with $X = \tau, v$) are computed over the triangular facets that compose the Fermi surface, using $\sim$10,000 $k$-points and weighing by the facet

\^Ref. [19]  
\^Ref. [57]  
\^Using estimated SC resistivity
Table 6.1: Comparison of relaxation times and velocities computed by averaging on the Fermi surface (values in brackets), versus the same quantities obtained within Drude theory (subscript D). The Fermi surface averaged relaxation times are computed using GW bandstructures, while the band velocities are given for both the DFT and GW bandstructures for comparison. The Drude values are obtained using Eq. (6.3) with resistivity experimental data [19, 57] from single-crystal samples.

<table>
<thead>
<tr>
<th></th>
<th>$\tau$ (fs)</th>
<th></th>
<th>$v$ ($10^8$ cm/s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\langle \tau_{nk} \rangle$</td>
<td>$\tau_D$</td>
<td>$\langle v_{nk}^{\text{DFT}} \rangle$</td>
</tr>
<tr>
<td>Cu</td>
<td>37</td>
<td>27$^a$</td>
<td>1.13</td>
</tr>
<tr>
<td>Ag</td>
<td>30</td>
<td>41$^b$</td>
<td>1.45</td>
</tr>
<tr>
<td>Au</td>
<td>24</td>
<td>30$^c$</td>
<td>1.38</td>
</tr>
</tbody>
</table>

area: $\langle X \rangle = \frac{\sum_t A_t X_{k_c}}{\sum_t A_t}$, (6.4)

with $A_t$ the facet area for a triangle $t$ and $k_c$ its centroid. The Drude relaxation times $\tau_D$ are calculated using the resistivity $\rho$ measured on single-crystal samples [19, 57] (see Table 6.2); we further obtain the charge density $n$ using 1 free electron per unit cell and take the effective mass $m^* = m_e$, consistent with a simple Drude treatment of noble metals [3]. We find that $\langle \tau_{nk} \rangle$ of Cu is 50% larger than the Drude value, while $\langle \tau_{nk} \rangle$ of Ag and Au are 15–25% smaller than their respective Drude relaxation times. The magnitude of the band velocities averaged over the Fermi surface show the opposite behavior, with $\langle v_{nk} \rangle$ smaller in Cu and larger in Ag and Au than the Drude value. These results challenge the accuracy of the widely used empirical Drude model. In particular, since the tabulated Drude relaxation times are constants, they miss the complex trends and significant variation on the Fermi surface found here. The Drude relaxation times should thus be regarded as mere model parameters used to interpret transport and optical experiments in the absence of ab initio data.

Next, we use our results to compute the resistivity of Cu, Ag, and Au, comparing a range of approximations for the relaxation times and velocities. Table 6.2 shows the resistivities calculated with GW band velocities in combination with i) full band- and $k$-resolved relaxation times $\tau_{nk}$, ii) Fermi-surface-averaged ab initio relaxation times $\langle \tau_{nk} \rangle$, and iii) empirical Drude relaxation times $\tau_D$. For comparison, we also compute the resistivities using the common hybrid approach which combines DFT (as opposed to GW) band velocities with Drude relaxation times. We compare our calculated resistivities with experiments at 293 K for single-crystal (SC) samples. Polycrystalline (PC) experimental data are also given. {For Cu and Ag, resistivity measurements on SC samples [19, 57] show a ~10% reduction compared to PC samples (see Table 6.2). We assume a similar reduction in the case of SC Au in order to obtain an estimate of the SC resistivity, which we could not find in the literature.} Since in pure SC samples at room temperature the source of resistivity is almost exclusively $e$-ph
Table 6.2: Resistivity values (units of $\mu\Omega\text{cm}$) at 300 K computed using GW band velocities and different approximations for the relaxation times. $\tau_{nk}$ are band and $k$-dependent from ab initio, $\langle \tau_{nk} \rangle$ are Fermi-surface averaged from $\tau_{nk}$, and $\tau_D$ are empirical Drude relaxation times. Also shown are the computed values from DFT velocities and empirical Drude relaxation times. Experimental data for single crystal (SC) and polycrystalline (PC) samples are shown for comparison.

<table>
<thead>
<tr>
<th></th>
<th>Computed $\rho$</th>
<th>Experimental $\rho$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\tau_{nk}$</td>
<td>$\langle \tau_{nk} \rangle$</td>
</tr>
<tr>
<td>Cu</td>
<td>1.51</td>
<td>1.51</td>
</tr>
<tr>
<td>Ag</td>
<td>1.81</td>
<td>1.83</td>
</tr>
<tr>
<td>Au</td>
<td>2.24</td>
<td>2.24</td>
</tr>
</tbody>
</table>

scattering, our approach using $e$-$ph$ relaxation times is justified.

We first discuss the data in Table 6.2 obtained with our best approximation, namely ab initio $\tau_{nk}$ and GW velocities. In the case of Cu, the agreement with SC experiment is excellent (within 1%). For Ag and Au we find resistivities that are larger by $\sim$10% than experiment. We attribute this small discrepancy to a number of sources, including the $\sim$0.1 eV accuracy of the quasiparticle energies obtained via the GW method, and a $\sim$5% uncertainty in the relaxation times depending on whether the cosine factor in Eq. (6.1) is employed or not [70]. Using the average relaxation times $\langle \tau_{nk} \rangle$ with GW band velocities gives nearly identical resistivities as the band- and $k$-resolved relaxation times, thus suggesting that approximate schemes using a constant relaxation time should employ Fermi-surface-averaged ab initio data rather than empirical data.

For all three materials, we find that the band velocities obtained from the GW quasiparticle bandstructures are $\sim$10% larger than those obtained via DFT (see Table 6.1), consistent with previous studies in Au [103] showing large effect from GW on the $d$ band energies and $sp$ band width. For Ag and Au, the smaller DFT velocities, combined with Drude relaxation times that are larger than the ab initio values, yield resistivities in fortuitous agreement with experiment when using the DFT velocity plus $\tau_D$ approach, a result of compensation of errors. For Cu, on the other hand, the DFT velocity plus $\tau_D$ approach gives a resistivity almost 50% higher than experiment. Our results highlight the predictive character of calculations combining GW bandstructures and velocities with ab initio $e$-$ph$ relaxation times.

We close by discussing the advantages of the Fermi surface importance sampling scheme introduced here. The standard scheme [99, 69] to compute transport properties employs a modified form of Eq. (6.2) summing velocities and relaxation times over a uniform $k$-point grid and integrating over energy with a $(-\partial f/\partial E)$ weighting factor [3]. Eq. (6.2) con-

$^a$Ref. [19]
$^b$Ref. [57]
$^c$Estimated from PC resistivity
verts this expression to a Fermi surface integral, assumes importance sampling with properly weighted \( k \)-points on the Fermi surface, and employs the lowest order Sommerfeld expansion, \((-\partial f/\partial E) \approx \delta(E - E_F)\), which is valid for \( T \ll T_F/k_B \), where \( T_F \) is the Fermi temperature (\( \sim 10,000 \) K in metals) and \( E_F \) is the Fermi energy.

Fig. 6.3 compares the convergence of the resistivity with respect to number of \( k \)-points for uniform and importance sampling. Since we focus only on the convergence behavior with respect to number of \( k \)-points, the resistivity for illustration in Fig. 6.3 is determined using DFT velocities and \( \tau_D \), though the trends are general. We find that the importance sampling scheme reaches convergence for \( \sim 10^4 \) \( k \)-points and approaches monotonically the converged value. On the other hand, the commonly employed uniform grid requires \( \sim 10^7 \) \( k \)-points to converge and fluctuates considerably with increasing number of \( k \)-points as it approaches convergence. Since calculations of \( e\)-ph relaxation times are computationally expensive, the importance sampling scheme provides great advantages when ab initio relaxation times are employed. For the noble metal cases discussed here, importance sampling decreases the number of \( k \)-points in the \( e\)-ph relaxation time calculation by a factor of \( 10^3 \). When higher order terms in the Sommerfeld expansion are needed, e.g., to compute high temperature electrical resistivity or thermal conductivity and thermopower, additional isoenergy surfaces (at \( E_F \pm dE \)) can be computed with a small computational overhead.
6.4 Discussion and conclusion

In conclusion, we show that the $\epsilon$-ph relaxation times and band velocities vary considerably on the Fermi surface of Cu, Ag, and Au. The correlation between relaxation times, band velocities and Fermi surface topology found here may extend to other classes of materials with complex Fermi surfaces. Our work points out the shortcomings of employing empirical relaxation times, DFT bands, and uniform $k$-grids to compute transport properties in materials. It further suggests that even energy-dependent (but not $k$-dependent) relaxation times, also commonly employed in transport studies, would miss the complex interplay of bandstructure and scattering processes. These observations may extend to defect and impurity scattering, which will be the subject of future investigations. While complex non-equilibrium approaches are being explored for carrier dynamics, we argue that accurate ab initio state-dependent relaxation times may yield close agreement with experiment in many cases of practical interest, even within the semiclassical Boltzmann relaxation time approach.
Chapter 7

Hot carrier generation and transport in noble metals

7.1 Introduction

Surface plasmon polaritons (SPPs) are electron collective excitations generated by light at the interface between a metal and a dielectric [4]. A schematic diagram for the dispersion relation of the plasmon modes at the interface between a uniform electron gas and a idealized dielectric (constant dielectric function) can be found in Fig. 7.1. The upper branch represents a bulk plasmon for small wavevectors and approaching the light-line in the dielectric for larger wavevectors. The lower branch represents a SPP, which asymptotes to the surface

\[
\omega^2 = \omega_p^2 + c^2 k_x^2
\]

\[
\omega = \frac{ck_x}{\sqrt{\epsilon_d}}
\]

Figure 7.1: Plasmon dispersion curves. The upper branch describes a bulk plasmon for small wavevector, while the lower branch describes a SPP for larger wavevectors.
plasmon energy at larger wavevectors. Modeling of SPPs is dominated by approaches that employ classical electromagnetism and account empirically for the properties of the materials supporting the SPP. Yet, the inherently quantum mechanical nature of SPPs [120, 29] becomes manifest in their scattering and decay processes in bulk and nanoscale materials. A key example is the decay of SPPs to electron-hole pairs, a process whose crucial importance is two-fold: first, it is a main energy loss mechanism of SPPs in metals, currently limiting the applicability of plasmonics [134, 56]; second, it leads to generation of hot electrons and holes with application in several branches of applied physics [14]. Recent experiments reported the extraction of such SPP-induced hot carriers (HCs) in metals before they thermalize — typically over nanometer lengths and sub-picosecond times — by injecting the HCs over a Schottky barrier into a semiconductor or oxide [20, 18, 127], or by transferring them into surface adsorbates to perform chemical reactions with large activation barriers [84, 63, 14].

The vast literature on HCs in noble metals generated by intense light pulses [31, 129, 25] (instead of SPPs), together with recent calculations on SPP-to-HC conversion [118, 74], help interpret these recent experiments on SPP-induced HCs.

Lack of predictive and accurate quantum mechanical approaches to study HCs generated by SPPs or photons has led to ambiguity in the microscopic interpretation of experiments involving HCs in Au and Ag, two mainstay materials in plasmonics [134]. In particular, recent studies indicate that SPP decay excites electronic transitions from occupied states close to the Fermi energy ($E_F$) [84], implying that most of the SPP energy goes into hot electrons rather than hot holes, and that nanocrystals with less than 10 nm diameter are necessary to obtain HCs with significant energy [38] (e.g., 1–2 eV away from $E_F$). The mean free paths (MFPs) of HCs in noble metals appear to be short [84], though a quantitative estimate of the MFPs and their dependence on crystal direction is not available. In addition, it is often assumed that HC scattering and thermalization are dominated by Auger processes [84, 25], since thermalization induced by phonons occurs on a slower time scale [25]. As shown in this work, all these conclusions must be revised. This work has been published as Ref. [10] in the literature.

Here, we develop a quantum mechanical framework to study the energy distribution of HCs generated by SPPs or photons in Au and Ag, and employ ab initio calculations of the electron-phonon (e-ph) and electron-electron (e-e) interactions to study the MFPs and relaxation times of HCs within 5 eV of the Fermi energy. Our approach is free of experimental input, and combines density functional theory (DFT) [76], GW [47], and ab initio e-ph calculations [88]. Our ability to use extremely fine grids for Brillouin zone (BZ) integrations allows us to resolve HC scattering with unprecedented accuracy [8]. We find that the interband transition threshold (between $d$ and $s$ states) defines two regimes for HC generation and transport. The decay of SPPs with energy lower than the interband threshold leads to generation of long-lived HCs with long isotropic MFPs of up to 40 nm and energy within 1–2 eV of $E_F$. On the other hand, decay of SPPs with energy higher than the interband threshold leads to generation of short-lived hot holes in $d$ states with anisotropic and short (sub-5 nm) MFPs, and hot electrons with only less than 1 eV energy above $E_F$. The regime characterized by SPP energy below the interband threshold is better suited to employ HCs
### Table 7.1: Square modulus of the coupling matrix element, $|g|^2$, and Feynman diagram for the self-energy of the boson in the lowest-order interaction for the three bosons considered here — plasmons, photons, and SPPs — with electrons in a material. These interactions share a common diagram (right column), consisting of a polarizability bubble $\chi$ multiplied by two coupling vertices $g$. The wiggly lines on the two sides of the self-energy loop $g^2\chi$ indicate the incoming and outgoing bosons.

<table>
<thead>
<tr>
<th>Interaction</th>
<th>Matrix element</th>
<th>Diagram</th>
</tr>
</thead>
<tbody>
<tr>
<td>Plasmon-electron</td>
<td>$</td>
<td>g^{PL}</td>
</tr>
<tr>
<td>Photon-electron</td>
<td>$</td>
<td>g^{PP}_{n,n',k}</td>
</tr>
<tr>
<td>SPP-electron</td>
<td>$</td>
<td>g^{SPP}_{n,n',k}</td>
</tr>
</tbody>
</table>

in applications requiring long MFPs, and allows one to optimize HC generation by tuning the SPP energy. These results represent an important first step to understand SPP decay and energy loss, and to control HC generation and transport in bulk and nanostructured noble metals.

### 7.2 Theory of SPP-electron coupling

Coupling of SPPs or photons to electrons in materials can be described in the framework of many-body perturbation theory [71]. The lowest-order Feynman diagram for this coupling process, shown in Table 7.1, describes a boson (here, SPP or photon) coupling to the electron gas through the electronic polarizability $\chi$ and a coupling matrix element $g$. This diagram is analogous to the lowest-order coupling of phonons to electrons [112, 36]. The decay rate $\Gamma(\mathbf{q}_p, E_p)$ for a boson of momentum $\hbar \mathbf{q}_p$ and energy $E_p$ to electron-hole pairs is proportional to the imaginary part of the self-energy [71]:

$$
\Gamma(\mathbf{q}_p, E_p) = \frac{2\pi}{\hbar} \sum_{n,n',k} |g_{n,n',k}(\mathbf{q}_p, E_p)|^2 \frac{1}{\pi} \text{Im} \left[ \frac{f_{n,k} - f_{n',k+\mathbf{q}_p}}{E_p(\mathbf{q}_p) - (E_{n',k+\mathbf{q}_p} - E_{nk}) - i\eta} \right]
$$

where $-E_{nk}$ and $E_{n',k+\mathbf{q}_p}$ are, respectively, the quasiparticle energies of the hot hole in a state with band $n$ and crystal momentum $-k$, and hot electron in a state with band $n'$ and crystal momentum $k + \mathbf{q}_p$, as generated in the boson decay process. In addition, $f_{n,k}$ and $f_{n',k+\mathbf{q}_p}$ are Fermi occupations, $\eta$ is a small broadening parameter, and the sum is contributed by electronic transitions from occupied to empty states that differ by $\hbar \mathbf{q}_p$ in crystal momentum and $E_p$ in energy.

For coupling matrix elements with a weak band and $k$ dependence, the decay rate of a SPP or photon is proportional to the imaginary part of the electronic polarizability, $\text{Im} \chi(\mathbf{q}_p, E_p)$. 
Here, we employ a quantity closely related to $\text{Im} \chi(q_p, E_p)$, the finite momentum joint density of states (FM-JDOS) $J(q_p, E_p)$, defined as the number of states per unit energy separated by momentum $\hbar q_p$ and energy $E_p$ in the quasiparticle bandstructure:

$$J(q_p, E_p) = \frac{1}{\pi} \text{Im} \chi(q_p, E_p) = \frac{1}{\pi} \text{Im} \left[ \sum_{n,n',k} \frac{f_{n,k} - f_{n',k + q_p}}{E_p(q_p) - (E_{n',k + q_p} - E_{n,k}) - i\eta} \right]. \quad (7.2)$$

The FM-JDOS is a measure of the phase space available for the decay of a SPP or photon with momentum $\hbar q_p$ and energy $E_p$. Greater FM-JDOS values correspond to greater SPP decay and HC generation rates.

The coupling matrix elements in Eq. 7.1 can be derived with a range of theoretical approaches. We first discuss the coupling of electrons to plasmons – i.e., collective electronic excitations in the absence of light – since theories developed for plasmons have been occasionally applied to SPPs [39]. The so-called Landau damping [97] treatment of the electron gas derives the decay rate $\Gamma_p$ of a bulk plasmon in the weak coupling limit via the zeros of the complex macroscopic dielectric function $\epsilon_M$ of the metal close to the real energy axis, i.e., $\epsilon_M(q_p, E_p + i \hbar \Gamma_p(q_p, E_p)) = 0$. Due to the bosonic nature of plasmons, the decay rate $\Gamma(q_p, E_p)$ has the form given in Eq. 7.1, with coupling matrix elements [97] $|g_{\text{PL}}|^2 = 2\pi e^2 E_p/q_p^2$. In spite of its common use, Landau damping is rarely discussed in terms of the diagram in Table 7.1 and the coupling $g_{\text{PL}}$, though this viewpoint elucidates its physical origin.

The coupling matrix elements for SPPs and photons can be derived by second quantization of their vector potentials [64] (see Sec. 7.3). For photons, the coupling is given by the well-known dipole matrix elements $|g_{n,n',k}^{\text{PH}}|^2 = (\hbar c/m)^2 \left[ 1/(2\epsilon E_p V) \right] |e_{q_p} \cdot p_{n,n',k}(q_p = 0)|^2$, where $V$ is the volume, $e_{q_p}$ is the polarization vector, and $p_{n,n',k}(q_p)$ the transition dipole matrix element. For SPPs at a metal-dielectric interface, we apply a vector potential second quantization procedure [87, 27] (see Sec. 7.3) to obtain the SPP-electron coupling $|g_{n,n',k}^{\text{SPP}}|^2$ given in Table 7.1. The SPP-electron coupling has a form analogous to that for photons, with two important differences. First, the volume $V$ is replaced by the SPP-field volume $V_{\text{SPP}} = A \cdot L_z$, where $A$ is the metal-dielectric interface area and $L_z$ the decay length of the SPP in the metal. Since $L_z$ is of order 10 nm at visible wavelength for noble metals [4] and thus much shorter than the light absorption depth, the electric field associated with the SPP is concentrated to a small volume $V_{\text{SPP}}$ at the metal-dielectric interface [4], resulting in enhanced local generation of HCs from SPPs compared to light. Second, the SPP wavevector $q_p$ is in general comparable to the size of the BZ, while for photons $q_p \approx 0$. The finite SPP wavevector can significantly alter the phase space available for SPP decay, as dictated by the $q_p$ dependence of the FM-JDOS. We remark that additional momentum can be transferred by phonons and defects assisting the SPP scattering process.

For both SPP and photon decay, the FM-JDOS regulates the phase space for HC generation. Since the experimental conditions and the material-specific SPP dispersion relation dictate the SPP energy $E_p$ and wavevector $q_p$ [4, 101, 141], we study the FM-JDOS as a function of $q_p$ and $E_p$ and arbitrarily treat them for the purpose of this work as inde-
pendent variables; HC generation from light is obtained as the specific limit $q_p = 0$. We note that the SPP decay rate $\Gamma_p(q_p, E_p)$ could also be computed by evaluating the coupling matrix elements explicitly for fixed experimental conditions and dispersion relation [4] $E_p(q_p) = \hbar q_p \sqrt{(\epsilon_M + \epsilon_p)/(\epsilon_M \epsilon_D)}$. This approach would enable first-principles calculations of plasmonic losses and the MFPs of SPPs in materials [134, 56].

### 7.3 Computational Methods

We carry out ab initio calculations on Au, Ag, and Cu in the face-centered cubic (FCC) structure with lattice parameters of 7.72, 7.71, and 6.82 bohr, respectively. The ground-state electronic structure is computed within the local density approximation (LDA) [95] of DFT using the QUANTUM ESPRESSO code [35]. We use norm-conserving pseudopotentials (which include semicore $s$ and $p$ states) to describe the core-valence interaction [128], together with a plane-wave basis set with kinetic energy cutoff of 60 Ry for Ag and Au, and 240 Ry for Cu. The $e-e$ (i.e., GW) and $e-ph$ self-energies are computed separately, and then combined together.

The real and imaginary parts of the GW self-energy [47] are computed using the Berkeley-GW code [24]. The real part $\text{Re}(\Sigma_{nk}^{GW})$ is obtained with a generalized plasmon-pole calculation [47] on a $12 \times 12 \times 12$ $k$-point grid and then interpolated using Wannier functions (see below). Kinetic energy cutoffs of 50 Ry and 120 Ry are used, respectively, for the screened and bare Coulomb interactions, together with $\sim 1000$ empty bands and a $8 \times 8 \times 8$ $q$-point grid for the dielectric screening. The DFT eigenvalues $\epsilon_{nk}$ are corrected with GW self-energies to obtain the quasiparticle bandstructures, $E_{nk} = \epsilon_{nk} + \text{Re}(\Sigma_{nk}^{GW})$, used in all calculations in this work. The imaginary part of the GW self-energy, $\text{Im}(\Sigma_{nk}^{GW})$, is computed using full-frequency GW calculations [24]. Here, $\text{Im}(\Sigma_{nk}^{GW})$ are evaluated on-shell at the LDA eigenvalues $\epsilon_{nk}$, and then plotted versus the corresponding GW eigenvalues $E_{nk}$. We use $20 \times 20 \times 20$ $k$-point grids to converge $\text{Im}(\Sigma_{nk}^{GW})$ within $\sim 10$ meV. In addition, we use kinetic energy cutoffs of 20 Ry and 120 Ry, respectively, for the screened and bare Coulomb interactions, together with $\sim 100$ empty bands and a $20 \times 20 \times 20$ $q$-point grid for the dielectric screening.

The imaginary part of the $e-ph$ self-energy, $\text{Im}(\Sigma_{nk}^{e-ph})$, is computed using the EPW code [88]. We compute the self-consistent potential and Kohn-Sham states on a $12 \times 12 \times 12$ $k$-point grid using DFT, and lattice-dynamical properties with DFPT [5] on a $4 \times 4 \times 4$ $q$-point grid. The $e-ph$ matrix elements are first computed on these coarse grids, and then obtained on significantly finer grids using an interpolation procedure based on Wannier functions as implemented in the EPW code [88, 36]. Maximally localized Wannier functions [77, 117] are obtained starting from five $d$ orbitals on the metal atoms and two $s$ orbitals, each in a tetrahedral interstitial point of the FCC lattice, for a total of 7 wannierized bands (we skip the semicore states). The Wannier interpolated and DFT bandstructures agree within 5 meV in an energy window up to 5 eV above the Fermi energy. The Bloch-to-Wannier rotation matrices are then used to interpolate the GW bandstructures, which are used to
compute \( \text{Im}(\Sigma_{n\mathbf{k}}^{e-ph}) \). The fine grids for the calculation of \( \text{Im}(\Sigma_{n\mathbf{k}}^{e-ph}) \) consist in a \( 40 \times 40 \times 40 \) \( \mathbf{k} \)-point grid, and up to 1 million random phonon \( \mathbf{q} \)-points in the BZ. Such fine grids allow us to converge \( \text{Im}(\Sigma_{n\mathbf{k}}^{e-ph}) \) within 1 meV. Further details of our approach to compute the \( e \)-\( ph \) self-energy are discussed in Ref. [8].

The FM-JDOS calculations are carried out using Wannier interpolated GW bandstructures, a \( 100 \times 100 \times 100 \) \( \mathbf{k} \)-point grid, and a small Lorentzian broadening \( \eta \approx 25 \text{ meV} \). The energy distribution of the generated HCs in Fig. 7.3 are obtained by counting the momentum and energy-conserving transitions in Eq. (7.2) as a function of occupied band (for hot holes) and empty band (for hot electrons) quasiparticle energies.

The MFPs are obtained using velocities computed from the slope of the GW bandstructure along the given high-symmetry direction, and total relaxation times \( \tau_{n\mathbf{k}} \) combining the \( e \)-\( ph \) and \( e \)-\( e \) relaxation times with Matthiessen’s rule. The total relaxation times are obtained by combining \( \text{Im}(\Sigma_{n\mathbf{k}}^{e-ph}) \) computed on grids of over 100 \( \mathbf{k} \) points along each high-symmetry direction, and \( \text{Im}(\Sigma_{n\mathbf{k}}^{GW}) \) computed on a \( 20 \times 20 \times 20 \) \( \mathbf{k} \)-grid and then averaged over \( \mathbf{k} \) points to generate an energy dependent \( \text{Im}(\Sigma_{n\mathbf{k}}^{GW}) \) at each energy \( E_{n\mathbf{k}} \) for which a calculation of \( \text{Im}(\Sigma_{n\mathbf{k}}^{e-ph}) \) has been performed.

### 7.4 Results

#### 7.4.1 Hot carrier generation

To study HC generation from SPPs at a metal-dielectric interface (see Fig. 7.2a), we compute the FM-JDOS for Au and Ag, as shown in Fig. 7.2b, for SPP energies up to 5 eV and increasing SPP wavevectors along the \( \Gamma - L \) direction in the BZ (i.e., for SPPs propagating along the [111] crystallographic direction). Our calculations are carried out using GW bandstructures on very fine grids in the BZ (see Sec. 7.3).

Detailed features of the bandstructures of Au and Ag, shown in Fig. 7.3, regulate HC generation and transport in these materials. The main features of the Au and Ag bandstructures are the presence of relatively dispersionless, occupied \( d \) bands with large associated electronic density of states (DOS), straddled by a free-electron-like \( s \) band. The top of the \( d \) bands is located at energy \( E_{\text{int}} \) below the Fermi energy, where \( E_{\text{int}} \) is the threshold for inter-band (\( d \) to \( s \)) electronic transitions [22, 51]. From our GW calculations, we find \( E_{\text{int}} \) values of 2.3 eV in Au and 3.7 eV in Ag (see Fig. 7.3), in excellent agreement with experiment [56, 22, 51]. The free-electron-like \( s \) band spans an energy window between \( E_{\text{int}} \) below and over 4 eV above \( E_F \), and hence dominates the electronic properties of Au and Ag near \( E_F \).

The FM-JDOS in Fig. 7.2b follows similar trends in Au and Ag, with deviations due to the different \( E_{\text{int}} \) values in the two materials. We first discuss the FM-JDOS with \( \mathbf{q}_p = 0 \), which quantifies the phase space for HC generation for both SPPs with very small wavevectors and photons. The FM-JDOS curve with \( \mathbf{q}_p = 0 \) \( (\Gamma) \) has a zero value for SPP energies up to \( E_{\text{int}} \), indicating the absence of possible SPP decay into HCs in this energy range. For
SPP energies greater than $E_{\text{int}}$, the FM-JDOS increases linearly with energy indicating that a large phase space opens up for HC generation. In this regime, SPP damping is strong, and HC generation is intense and dominated by formation of hot $d$ holes collecting most of the energy from the SPP, as discussed below.

For small SPP wavevectors along the $\Gamma-L$ direction, we find a peak in the FM-JDOS corresponding to SPP decay through intraband transitions within the $s$ band. The energy position of the peak increases linearly with SPP wavevector, and for large enough wavevectors merges with the FM-JDOS above $E_{\text{int}}$. In Au, the intraband peak is below $E_{\text{int}}$ for $q_p$ values up to $\sim 0.2 \ L$, and in Ag for $q_p$ values up to $\sim 0.3 \ L$ due to the higher $E_{\text{int}}$ value in Ag. The presence of an intraband peak associated with $s-s$ transitions can be understood...
by examining the energy difference between states with momentum $k$ and $k + q_p$ in the free-electron-like $s$ band:

$$E(k + q_p) - E(k) = \frac{\hbar^2}{2m}(2k \cdot q_p + q_p^2).$$

For small wavevector $q_p$, the term quadratic in $q_p$ can be neglected, while the term linear in $q_p$ increases by approximately $\hbar^2 \pi^2 \sqrt{3}/(10ma^2) \approx 1$ eV ($a$ is the lattice constant) for an increase of $q_p$ by a $1/10$ the $\Gamma - L$ distance. The computed FM-JDOS confirms these trends, as seen by the increase in the energy of the intraband peak by $\sim 1$ eV in going from $0.1L$ to $0.2L$ and from $0.2L$ to $0.3L$. In this small wavevector regime, generation of HCs is appreciable only if the SPP energy and wavevector are matched to the intraband FM-JDOS peak, namely if $E_p(q_p) \approx 2k \cdot q_p$ for a set of $k$ points in the BZ. When this condition is met, intense HC generation from intraband transitions can occur, resulting in an almost equal apportioning of the SPP energy between hot electrons and holes, as discussed below.

For SPP wavevectors larger than $\sim 0.3L$, the FM-JDOS resembles that for $q_p = 0$ except from a flat tail that extends to low SPP energy, as shown by the dashed lines in Fig. 7.2. This tail is due to generation of HCs from intraband transitions with large transferred momentum from the decay of SPPs with wavevector comparable to the size of the BZ. The FM-JDOS above $0.5L$ are unchanged and nearly identical to the $0.5L$ case.

To summarize, we find two HC generation regimes. For SPPs with energy $E_p < E_{int}$, optimal generation can be achieved using SPPs with relatively small wavevector (less than $\sim 1/5$ the BZ size) that are matched with the FM-JDOS peak, while HC generation is relatively weak for SPPs with larger momentum. For SPPs with energy $E_p > E_{int}$, HC generation is intense regardless of SPP momentum. We find identical trends for SPPs with wavevector in the $\Gamma - X$ high-symmetry direction, i.e., propagating along the [100] direction. Removal of the electron gas approximation used in simplified treatments of SPPs in noble metals [38] is necessary to capture the delicate interplay found here between the energetics of the $s$ and $d$ states and intraband ($s-s$) and interband ($d-s$) transitions.

### 7.4.2 Energy distribution of hot carriers

Fig. 7.3 shows the energy distribution of HCs generated by specifically assumed SPPs with energy increasing in small steps from $\sim 1$ to $5$ eV and a range of wavevectors in the $\Gamma - L$ direction, along with the GW bandstructures (see Sec. 7.3). From now on throughout the chapter, all HC energies are referenced to the Fermi energy ($E_F$), and the energy of the hot hole is $-E$, i.e., the direction of increasing energy in the bandstructure is downward for holes and upward for electrons. The energy distribution histograms identify the HCs generated in the two regimes discussed above. For SPPs with energy lower than $E_{int}$, hot holes and electrons are generated with a roughly uniform energy distribution as a result of the $s-s$ intraband transitions. The HC energy distributions shown in Fig. 7.3 in this energy range pertain to SPPs with energy and wavevectors matched to the FM-JDOS intraband
Figure 7.3: Energy distribution of hot electrons (red) and hot holes (blue) in Au (a) and Ag (b) generated by SPP decay. Each plot refers to a specific SPP energy and wavevector. The right panel shows the GW bandstructures of Au and Ag, respectively in (a) and (b), along with a schematic of the momentum and energy conserving excitation of a hot electron-hole pair resulting from the decay of a SPP with wavevector $q_p$ and energy $E_p$. The shaded area indicates the energy range of the $d$ states. The energy difference between the Fermi energy and the top of the $d$ bands corresponds to the interband transition threshold. The Fermi energy is the zero of the energy axis.

peaks (see Fig. 7.2). While the details of the energy distribution in this regime are sensitive to the SPP energy $E_p$, the overall trends indicate that HCs are generated with an average energy of $\sim E_p/2$.

For SPPs with energy greater than $E_{\text{int}}$, we find a change to a different HC generation regime. At the onset of the interband transitions for SPP energy of 2.3 eV in Au and 3.7 eV in Ag, upon SPP decay the hot holes absorb most of the SPP energy, creating a spike in the hot hole population at the top of the $d$ band. On the other hand, hot electrons have only modest energies in this regime. At higher SPP energies of $\sim 5$ eV, the HC population becomes approximately uniformly distributed over a wide energy range, although a peak
in the hot $d$ hole population is still present in Ag due to its higher $E_{\text{int}}$. These trends are common to both Au and Ag, and we have verified that Cu, a material less commonly used in plasmonics [56] and not discussed here, follows identical HC generation and energy distribution trends as Au and Ag.

**Hot carrier scattering processes.** While ample experimental [129, 22] and theoretical [25] data exists on HCs in noble metals, characterization of the MFPs and relaxation times of HCs using ab initio theory has been limited by the absence of accurate $e$-$ph$ calculations including all phonon modes over the entire BZ. First-principles calculations combining $e$-$ph$ and $e$-$e$ interactions are necessary to characterize the anisotropic and energy-dependent MFPs of HCs in materials, as shown in our recent work on HCs in semiconductors [8]. We compute the scattering rate (and its inverse, the relaxation time) for the $e$-$ph$ and $e$-$e$ interactions from the imaginary part of the respective self-energies, $\text{Im}\Sigma^{e-ph}$ and $\text{Im}\Sigma^{e-e}$, and resolve these quantities for each electronic state with band $n$ and $k$-point on very fine BZ grids (see Sec. 7.3 and Ref. [8]). The total relaxation times $\tau_{nk}$, combining the $e$-$ph$ and $e$-$e$

![Figure 7.4: Relaxation time and scattering rates for HCs in Au (a) and Ag (b) within 5 eV of the Fermi energy. The top panels show the total relaxation time, accounting for both the $e$-$e$ and the $e$-$ph$ interactions. The bottom panels show the scattering rate, expressed as the imaginary part of the self-energy, $\text{Im}\Sigma$, for the $e$-$e$ and $e$-$ph$ interactions. $E_F$ is the zero of the energy axis, as shown by the dashed line. Electrons (holes) possess positive (negative) energies, and the shaded area indicates the energy window spanned by the $d$ states.](image-url)
interactions, are obtained as \( \tau_{nk} = \left[ \frac{2}{\hbar} \left( \text{Im} \Sigma_{nk}^{e-e} + \text{Im} \Sigma_{nk}^{e-ph} \right) \right]^{-1} \).

Figure 3 shows the total relaxation times \( \tau_{nk} \) for HC's in Au and Ag with energy up to 5 eV. The total relaxation times decay rapidly away from \( E_F \) following a volcano-shaped trend, with the peak of the volcano centered close to \( E_F \). In Ag, we also find a peak in the relaxation times at the top of the \( d \) bands, in agreement with recent photoemission experiments [91]. Analysis of the \( e \)-\( ph \) and \( e \)-\( e \) scattering rates (see the bottom panels in Fig. 7.4) highlights the differences between the energy windows spanned by the \( s \) and \( d \) states, and elucidates the origin of the relaxation time trends. In the energy range spanned by the \( s \) states, the \( e \)-\( e \) scattering rates (i.e., \( \text{Im} \Sigma^{e-e} \)) form a parabola with minimum value of zero at \( E_F \) as predicted by Fermi liquid theory [25], while the \( e \)-\( ph \) scattering rates are relatively constant and exhibit a minimum at the onset of the \( d \) states and a maximum 1–2 eV above \( E_F \). The \( e \)-\( e \) scattering rates become greater than the \( e \)-\( ph \) \( \sim 2 \) eV away from \( E_F \), thus indicating that Auger and impact ionization processes included in the \( e \)-\( e \) term dominate HC scattering only 2 eV or more away from \( E_F \), while \( e \)-\( ph \) dominates HC scattering within 2 eV of \( E_F \). Combining the two scattering mechanisms leads to total relaxation times with a broad maximum centered at \( E_F \) and a rapid decay 1–2 eV away from \( E_F \), as the parabolic \( e \)-\( e \) scattering rate becomes dominant (see Fig. 7.4). In the energy window spanned by the \( d \) bands, the large increase in the electronic DOS causes very strong \( e \)-\( ph \) scattering, and the localized nature of the \( d \) states leads to deviations from the free-electron parabolic trends for \( e \)-\( e \) scattering. The total relaxation times show that hot holes arising from \( d \) states lose energy on a sub-5 fs time scale, thus making such hot \( d \) holes challenging to extract before thermalization. Finally, our calculations yield comparable time scale and scattering rate for \( e \)-\( ph \) and \( e \)-\( e \) scattering, in contrast with the common heuristic assumption that scattering by phonons occurs on a much slower time scale [25]. We thus conclude that previous models including only \( e \)-\( e \) scattering [25] and missing the important \( e \)-\( ph \) component are incomplete for understanding HC relaxation times and photoemission linewidths in noble metals.

To validate our results, we compare our relaxation times at the Fermi energy in Au and Ag with those inferred from carrier transport measurements. Kopitzki [59] combined the Drude model with room temperature resistivity measurements, and obtained semiempirical relaxation times of 29 fs for Au and 41 fs for Ag, in close agreement with those obtained by Ashcroft and Mermin with the same approach [3]. Averaging our relaxation times over a small energy window around \( E_F \) yields relaxation times of 24 fs for Au and 27 fs for Ag. While the agreement is excellent for Au and good for Ag, our data shows a large spread in the relaxation times at \( E_F \), a feature not captured by the Drude model and stemming from the Fermi surface anisotropy. It is therefore puzzling that the Drude relaxation times [51, 59] are widely used in plasmonics [134, 56], given that their physical meaning is questionable when applied to non-equilibrium situations involving SPPs. We conclude that further investigation on this point is necessary, and remark that ours is the first truly ab initio determination of the relaxation times on the Fermi surface of noble metals, to be discussed in detail elsewhere.
7.4.3 Hot carrier mean-free-paths

Next, we turn our attention to HC transport. Fig. 7.5 shows the MFP for the three crystallographic directions [100], [110], and [111] within 5 eV of the Fermi energy. The MFPs are obtained by multiplying the total relaxation times by the carrier velocities computed from the slope of the GW bandstructures (see Sec. 7.3). Our computed MFPs are of order 10−40 nm in the energy window spanned by the $s$ states, and much shorter (1−5 nm) for the $d$ states. The MFPs of the $s$ states exhibit an energy dependence with the volcano shape also seen in the relaxation times, with deviations coming from the different band velocities along each direction. The MFPs of hot holes with 1−2 eV energy are nearly isotropic, while the MFPs of hot holes with less than 1 eV energy are longer for the [100] and [110] directions.

![Figure 7.5](image)

Figure 7.5: The hot carrier MFPs along the [100], [110], and [111] crystallographic directions in Au (a) and Ag (b). The MFPs for the $d$ states are expanded in the right panels.
The MFPs of hot electrons are *highly* anisotropic due to the absence of electronic states up to 4 eV above $E_F$ in the [111] direction and at energy greater than 1−2 eV in the [100] direction. The longest hot electron MFPs are in the [110] direction up to 4 eV above $E_F$. We note that the energy derivative of the MFPs at $E_F$ is negative, consistent with the positive thermopower in Au and Ag [13]. In the energy window spanned by the $d$ states, the MFPs are of order 5 nm in the [111] direction and 1 nm in the other directions, and thus highly anisotropic with longer MFPs in the [111] direction. The isotropic MFPs for $s$ holes and anisotropic for $d$ holes are consistent with findings in ballistic electron energy microscopy (BEEM) experiments in noble metals, where it was found that $d$ holes injected by a STM tip propagate along narrow cones and span small volumes, while $s$ holes span large volumes isotropically [11].

### 7.5 Discussion and conclusion

The data obtained so far allow us to formulate optimal conditions to generate HCs and extract or utilize them before thermalization. Fig. 7.6 summarizes HC generation and transport in Au and Ag by defining two regimes for HC generation from intraband ($E_p < E_{int}$) and interband ($d - s$)

---

![Figure 7.6](image-url)

**Figure 7.6:** The graph shows two regimes, for intraband ($s - s$) and interband ($d - s$) transitions induced by SPPs with energy $E_p < E_{int}$ or $E_p > E_{int}$, respectively. The regime with $E_p < E_{int}$ offers a trade-off between optimal HC generation, moderate HC energies up to $\sim 2$ eV, and superior MFPs and relaxation times compared to more energetic HCs. The dashed line indicates the Fermi energy.
and interband ($E_p > E_{\text{int}}$) transitions induced by SPP decay. The energy distribution, relaxation time, and MFP of the generated HCs are dramatically different in these two regimes. For $E_p < E_{\text{int}}$, the HCs consist of $s$ holes and $s$ electrons with average energies up to $E_{\text{int}}/2$ (i.e., 1–2 eV) generated by weakly damped SPPs due to the small FM-JDOS values. HC generation can be optimized by matching the energy and momentum of the SPP to the FM-JDOS intraband peak described above and shown in Fig. 7.2. This can be achieved by changing the dielectric in order to modify the range of wavevectors and energies of the SPPs, as can be seen, for example, in Fig. 7.7. The HCs in this energy range possess relatively long MFPs of $\sim$10–40 nm, and can thus be extracted before thermalization over $\sim$50 nm lengths in ideally pure samples. This regime is thus optimal for HC generation and extraction. For hot electrons, we predict that the (110) surface would be the best suited to extract HCs due to the longer MFPs in the [110] direction. For hot holes with $\sim$1–2 eV energy, the (100) and (110) surfaces are predicted to enable optimal extraction. Our results further suggest that Ag may be better suited than Au for HC generation due to the wider energy window for intraband transitions, thus motivating studies of HC generation from SPPs in Ag. On the other hand, the strong SPP damping regime with $E_p > E_{\text{int}}$ is non-optimal for extracting HCs, since in spite of strong HC generation the SPP energy is mostly transferred to short-lived hot $d$ holes with $\sim$1 nm MFP, while hot electrons possess only low energy of less than 1 eV. These conditions prevent HC extraction in this regime, unless a nm-thick
thin film or nanostructured metal is employed.

Our data further show that nanostructuring the metal to sub-10 nm size is not required to obtain energetic HCs, and not needed to extract them unless HCs with more than 2 eV energy are desired. In addition, for SPP with energy lower than $E_{\text{int}}$, the SPP energy is equally distributed between hot electrons and holes, and energy loss is dominated by $e$-ph rather than Auger scattering. These findings address the misconceptions emerged in recent work as outlined above, and show that optimal HC generation is possible by carefully tuning the SPP energy and wavevector at noble metal-dielectric interfaces. Finally, we note that our approach is for now limited to SPP at metal-dielectric interfaces, while localized surface plasmons (LSP) need to be treated separately due to their different nature [98]. As such, our results cannot address directly studies on metallic nanostructures. We believe that the method presented in our work can be extended to study hot carrier generation from LSP modes, which will be the subject of future investigation.

In summary, we establish a theoretical framework to study SPP damping and HC generation and transport on the same footing using many-body perturbation theory. Our first-principles calculations can accurately describe SPP-induced generation and ultrafast scattering of HCs in noble metals of use in plasmonics, photocatalysis, photovoltaics, and optoelectronics. Our work highlights the interplay of the $s$ and $d$ bands in noble metals, and prescribe optimal experimental conditions for generation and extraction of HCs. Our approach paves the way to first-principles calculations of SPP losses in materials [134].
Part III

Appendices
Appendix A

Implementation of the optimized projection functions method

A.1 Spread functional

Here we express components of the spread $\Omega$, corresponding to $N$ composite bands, as a function of the overlap matrices $m^{(k,b)}$ following Ref. [77],

$$
\Omega_I = \frac{1}{N_k} \sum_{k,b} w_b \left( N - \sum_{ij} \left| m_{ij}^{(k,b)} \right|^2 \right),
$$

$$
\Omega_D = \frac{1}{N_k} \sum_{k,b} w_b \sum_i \left( -\text{Im} \ln m_{ii}^{(k,b)} - b \cdot \mathbf{r}_i \right)^2,
$$

$$
\Omega_{OD} = \frac{1}{N_k} \sum_{k,b} w_b \sum_{i \neq j} \left| m_{ij}^{(k,b)} \right|^2.
$$

The $w_b$ are the weights of the $b$ vectors connecting neighboring $k$ points (see Sec. 2.1 of Ref. [83]), while

$$
\mathbf{r}_i = \frac{1}{N_k} \sum_{k,b} w_b b \text{Im} \ln m_{ii}^{(k,b)}
$$

is the center of the $i$th Wannier function.

We note that the diagonal and offdiagonal parts of the spread depend only on the diagonal and offdiagonal components of the overlap matrices, respectively. Combining the invariant and offdiagonal parts of the spread gives an expression that depends only on the diagonal components of the overlap matrices,

$$
\Omega_{I,OD} = \Omega_I + \Omega_{OD} = \frac{1}{N_k} \sum_{k,b} w_b \sum_{i=1}^N \left[ 1 - \left| m_{ii}^{(k,b)} \right|^2 \right].
$$
A.2 Codiagonalization algorithm

In the main text, the construction of localized Wannier functions is recast into the following mathematical problem. Given a set of large $M \times M$ matrices $X^{(\alpha)}$, we wish to find a single rectangular semiunitary $M \times N$ matrix $W$ such that the set of small $N \times N$ matrices $W^{\dagger}X^{(\alpha)}W$ minimize the Lagrangian $L$, defined in Eq. (4.40),

$$\sum_\alpha t^{(\alpha)} \sum_{i=1}^{N} \left| \left[ W^{\dagger}X^{(\alpha)}W \right]_{ii} \right|^2.$$  \hspace{1cm} (A.4)

We parametrize the semiunitary matrix $W$ as follows. First, we define $W$ to be first $N$ columns of an $M \times M$ unitary matrix $\tilde{W}$. Second, we iteratively parametrize the enlarged matrix $\tilde{W}$ as a product (post-multiplication) of Givens rotations \[37\],

$$\tilde{W} = \prod_{l=1}^{L} \prod_{i=1}^{N} \prod_{j=i+1}^{M} R_l[i, j, \theta, \phi].$$  \hspace{1cm} (A.5)

Here integer $l$ denotes a particular iteration in the expansion.

A Givens rotation $R[i, j, \theta, \phi]$ is the most general unitary matrix that acts only on $i$th and $j$th rows and columns. Therefore, we parametrize $R[i, j, \theta, \phi]$ with two angles $\theta$ and $\phi$ as a matrix equal to the identity matrix for all elements except for the $ii$, $ij$, $ji$, and $jj$ elements,

$$\begin{pmatrix} R_{ii} & R_{ij} \\ R_{ji} & R_{jj} \end{pmatrix} = \begin{pmatrix} \cos \theta & e^{i\phi} \sin \theta \\ -e^{-i\phi} \sin \theta & \cos \theta \end{pmatrix}.$$  \hspace{1cm} (A.6)

The only diagonal elements of $X^{(\alpha)}$ affected by $R[i, j, \theta, \phi]$ are $X_{ii}^{(\alpha)}$ and $X_{jj}^{(\alpha)}$. Therefore there is no need to include in Eq. (A.5) cases when both $i$ and $j$ are larger than $N$, since that operation will have no effect on the Lagrangian. In addition, we don’t consider cases when $j < i$ since that transformation is captured by $j > i$. With this parametrization an arbitrary unitary matrix $\tilde{W}$ can be approximated to an arbitrary precision with large enough number of iterations, $L$.

Let us now see how does a single Givens rotation affect the Lagrangian. For a Givens rotation $R[i, j, \theta, \phi]$, the sum of the weighted square moduli of the diagonal elements ($ii$ and $jj$) of a set of rotated matrices $R^{\dagger}X^{(\alpha)}R$ are \[116\]

$$\sum_\alpha t^{(\alpha)} \left| \left[ R^{\dagger}X^{(\alpha)}R \right]_{ii} \right|^2 = x^T Q x + p^T x + c$$  \hspace{1cm} (A.7)

$$\sum_\alpha t^{(\alpha)} \left| \left[ R^{\dagger}X^{(\alpha)}R \right]_{jj} \right|^2 = x^T Q x - p^T x + c$$  \hspace{1cm} (A.8)

where

$$x^T = (\cos 2\theta, \sin 2\theta \cos \phi, \sin 2\theta \sin \phi).$$  \hspace{1cm} (A.9)
is a vector with unit norm by construction. The coefficients of the quadratic forms above [Eqs. (A.7) and (A.8)] depend only on the $ii$, $ij$, $ji$, and $jj$ components of the $X^{(a)}$ matrices

$$Q = \sum_{\alpha} t^{(a)} \Re [z^{(a)} z^{(a)*}]$$

$$p = \sum_{\alpha} t^{(a)} \Re \left[ \left( X_{ii}^{(a)} + X_{jj}^{(a)} \right)^* z^{(a)} \right]$$

(A.10)

$$c = \sum_{\alpha} \frac{1}{4} t^{(a)} \left| X_{ii}^{(a)} + X_{jj}^{(a)} \right|^2$$

where

$$z^{(a)} = \frac{1}{2} \begin{bmatrix} X_{ii}^{(a)} - X_{jj}^{(a)} \\ - (X_{ij}^{(a)} + X_{ji}^{(a)}) \\ i \left( X_{ij}^{(a)} - X_{ji}^{(a)} \right) \end{bmatrix}.$$ 

(A.11)

We now consider two cases. First, if $j \leq N$ both the $ii$ and $jj$ diagonal elements enter the Lagrangian $L$ so we need to find $x$ that minimizes the sum of Eqs. (A.7) and (A.8),

$$\sum_{\alpha} t^{(a)} \left| \left[ R^t X^{(a)} R \right]_{ii} \right|^2 + t^{(a)} \left| \left[ R^t X^{(a)} R \right]_{jj} \right|^2 = 2x^t Q x + 2c.$$ 

(A.12)

This is a quadratic programming problem with the constraint that $|x| = 1$. Here, the Lagrangian is simply minimized for $x$ that is the normalized eigenvector corresponding to the minimal eigenvalue of $Q$. For numerical stability, if the first component of $x$ (i.e., $\cos 2\theta$) happens to be negative we choose $-x$ instead of $x$.

In the second case ($j > N$), only the $ii$ diagonal components enters the Lagrangian $L$ so we need to find $x$ that minimizes Eq. (A.7),

$$x^t Q x + p^t x + c.$$ 

(A.13)

The solution of this problem is discussed in Ref. [116] within the context of matrix codiagonalization. However we find the general quadratic programming solution from Ref. [32] more numerically stable. Following Ref. [32], we first find the minimal eigenvalue $\lambda_{\text{min}}$ of the quadratic eigenvalue problem (QEP)

$$\left( \chi^2 A_2 + \chi A_1 + A_0 \right) x = 0,$$ 

(A.14)

with

$$A_2 = I_3$$

$$A_1 = -2Q$$

$$A_0 = Q^2 - \frac{1}{4} pp^t.$$ 

(A.15)
The QEP is linearized by introducing

\[ \tilde{x} = \begin{pmatrix} \chi x \\ x \end{pmatrix} \tag{A.16} \]

yielding a generalized eigenvalue problem

\[ A \tilde{x} = \chi B \tilde{x}, \tag{A.17} \]

with

\[ A = \begin{pmatrix} A_1 & A_0 \\ -I_3 & 0 \end{pmatrix} \]
\[ B = \begin{pmatrix} A_2 & 0 \\ 0 & I_3 \end{pmatrix}. \tag{A.18} \]

This generalized eigenvalue problem we solve using standard linear algebra techniques. The solution \( x \) that minimizes Eq. A.13 depends on whether \( \chi_{\text{min}} \) is in the spectrum of \( Q \) or not.

If \( \chi_{\text{min}} \) is not in the spectrum (i.e. not an eigenvalue) of \( Q \) then the solution is \( x = (Q - \chi_{\text{min}} I)^{-1}(-p/2) \). If \( \chi_{\text{min}} \) is an eigenvalue of \( Q \) we first define

\[ u := (Q - \chi_{\text{min}} I)^{+}(-p/2). \tag{A.19} \]

Here symbol \( ^{+} \) denotes a matrix pseudoinverse. A nontrivial solution to Eq. A.13 exists only when the following conditions are satisfied:

\[ (Q - \chi_{\text{min}} I)u = -p/2 \quad \text{and} \quad |u| \leq 1. \tag{A.20} \]

Finally, if \( |u| = 1 \), then the solution is \( x = u \). Otherwise (\( |u| < 1 \)) the solution is \( x = u + \xi \). Here \( \xi \) is an eigenvector of \( Q \) corresponding to \( \chi_{\text{min}} \) chosen so that \( |\xi|^2 = 1 - |u|^2 \).

Once the \( x \) is found for a given \((i, j)\) in either of the two approaches we determine the corresponding angles \((\theta, \phi)\) from Eq. (A.9) and construct the Givens rotation \( R[i, j, \theta, \phi] \). Next we update at each iteration the matrix \( \tilde{W} \) according to the postmultiplication parametrization from Eq. (A.5),

\[ \tilde{W} \to \tilde{W} R. \tag{A.21} \]

This iterative procedure over \( i, j, \) and \( l \) continues until the Lagrangian converges.

### A.2.1 Extensions

In the previous section we described an implementation of the optimized projection functions method. This subsection describes how to incorporate the offdiagonals of \( W^{\dagger} \left( S^{(k)} - I_M \right) W \) into Lagrangian.
As before, we construct the semiunitary $W$ as the $M \times N$ submatrix of a square $M \times M$ unitary matrix $\tilde{W}$. The matrix $\tilde{W}$ is written as a product (post-multiplication) of Givens rotations,

$$\tilde{W} = \prod_{l=1}^{L} \prod_{i=1}^{N} \prod_{j=i+1}^{M} R_l[i, j, \theta, \phi]. \quad (A.22)$$

A Givens rotation $R[i, j, \theta, \phi]$ ($R_l[i, j, \theta, \phi]$) is a unitary planar rotation that only acts on the $i$th and $j$th columns (rows) of a matrix (see Figure A.1). The matrix $R[i, j, \theta, \phi]$ is identity except the $ii$, $ij$, $ji$, and $jj$ elements,

$$\begin{pmatrix} R_{ii} & R_{ij} \\ R_{ji} & R_{jj} \end{pmatrix} = \begin{pmatrix} \cos \theta & e^{i\phi} \sin \theta \\ -e^{-i\phi} \sin \theta & \cos \theta \end{pmatrix}. \quad (A.23)$$

Again we consider two cases (see Figure A.1). If $j \leq N$, the $ij$ and $ji$ elements enter the Lagrangian $\mathcal{L}$. Just like the $ii$ and $jj$ components of the transformed matrices can be written in a quadratic form, so too can the offdiagonal $ij$ and $ji$ elements,

$$\sum_{\alpha} t^{(\alpha)} \left( \left| [R^l X^{(\alpha)} R]_{ij} \right|^2 + \left| [R^l X^{(\alpha)} R]_{ji} \right|^2 \right) = x^T Q x + c. \quad (A.24)$$
The matrix $Q$ is symmetric, and its independent components are:

\[
2Q_{11} = |X^{(a)}_{ij}|^2 + |X^{(a)}_{ji}|^2
\]
\[
2Q_{22} = |X^{(a)}_{ii}|^2 + |X^{(a)}_{jj}|^2 - 2 \text{Re} \left[ X^{(a)}_{ij} X^{(a)*}_{ji} + X^{(a)}_{ii} X^{(a)*}_{jj} \right]
\]
\[
2Q_{33} = |X^{(a)}_{ii}|^2 + |X^{(a)}_{jj}|^2 + 2 \text{Re} \left[ X^{(a)*}_{ij} X^{(a)}_{ji} - X^{(a)*}_{ii} X^{(a)}_{jj} \right]
\]
\[
2Q_{12} = \text{Re} \left[ \left( X^{(a)}_{ii} - X^{(a)}_{jj} \right) \left( X^{(a)*}_{ij} + X^{(a)*}_{ji} \right) \right]
\]
\[
2Q_{13} = \text{Re} \left[ \left( X^{(a)}_{ii} - X^{(a)}_{jj} \right) \text{Im} \left( X^{(a)}_{ij} - X^{(a)}_{ji} \right) + \left(-X^{(a)}_{ij} + X^{(a)}_{ji}\right) \text{Im} \left( X^{(a)}_{ii} - X^{(a)}_{jj} \right) \right]
\]
\[
2Q_{23} = 2 \text{Im} X^{(a)}_{ji} \text{Re} X^{(a)}_{ij} - 2 \text{Im} X^{(a)}_{ij} \text{Re} X^{(a)}_{ji}.
\] (A.25)

The term $c$ can be ignored as it does not depend on $x$. For the other offdiagonal matrix elements ($ik$, $ki$, $jk$, and $kj$, with $k \neq i \neq j \neq k$), the sum

\[
|X^{(a)}_{ik}|^2 + |X^{(a)}_{ki}|^2 + |X^{(a)}_{jk}|^2 + |X^{(a)}_{kj}|^2
\] (A.26)

is conserved and therefore does not affect the variation of the Lagrangian.

In the case of $j > N$, the $ij$ and $ji$ elements are outside of the $N \times N$ submatrix and they therefore do not enter the Lagrangian. However, in this case when $j > N$, the sum of the square moduli of the $ik$ and $ki$ offdiagonal elements,

\[
\sum_{\alpha} t^{(a)} \left( |\left[R^\dagger X^{(a)} R\right]_{ik}|^2 + |\left[R^\dagger X^{(a)} R\right]_{ki}|^2 \right) = p^T x + c
\] (A.27)

is not conserved and it therefore must be included in minimization of $L$. The coefficient $p$ of the term linear in $x$ can be expressed as

\[
p_1 = \sum_{\alpha} \frac{1}{2} \left( |X^{(a)}_{ik}|^2 - |X^{(a)}_{jk}|^2 + |X^{(a)}_{ki}|^2 - |X^{(a)}_{kj}|^2 \right)
\]
\[
p_2 = \sum_{\alpha} - \text{Re} \left[ X^{(a)}_{ik} X^{(a)*}_{jk} + X^{(a)*}_{ki} X^{(a)}_{kj} \right]
\]
\[
p_3 = \sum_{\alpha} \left( \text{Im} X^{(a)}_{jk} \text{Re} X^{(a)}_{ik} - \text{Im} X^{(a)}_{ik} \text{Re} X^{(a)}_{jk} - \text{Im} X^{(a)}_{kj} \text{Re} X^{(a)}_{ki} + \text{Im} X^{(a)}_{ki} \text{Re} X^{(a)}_{kj} \right)
\] (A.28)

For each $k$ (such that $k \neq i \neq j \neq k$) we have a term as in Eq. (A.27), with $p = p^{(k)}$, so that in $L$ we include the term $\sum_k p^{(k)} x$. The terms in Eq. (A.24) and Eq. (A.27) are easily added to the minimization algorithm described in the previous section.
Appendix B

Tetrahedron method for Fermi surface sampling

The tetrahedron method has been developed over the years to efficiently perform Brillouin zone integrations. Many properties of materials can be expressed as an integral taking one of the following three forms [140]:

\[ I^{(I)} = \sum_{n} \int_{BZ} \frac{dk}{(2\pi)^3} F_{nn}(k) \theta(E_{nk} - E_F) \]  
\[ I^{(II)} = \sum_{n} \int_{BZ} \frac{dk}{(2\pi)^3} F_{nn}(k) \delta(E_{nk} - E_F) \]  
\[ I^{(III)}(\omega) = \sum_{vc} \int_{BZ} \frac{dk}{(2\pi)^3} F_{vc}(k) \delta(\hbar\omega - (E_{ck} - E_{vk})) \].

The tetrahedron method is commonly used in the evaluation of integrals involving delta functions since the integrand can be linearly interpolated and evaluated and has been implemented in many codes in order to compute properties like the density of states. The transport calculations in Chapter 6 are an example of type II, which, after transforming to an integral over energy isosurfaces, can be expressed as

\[ I^{(II)} = \sum_{n} \int dE \frac{dS_E}{(2\pi)^3} \frac{F_{nn}(k)}{v_{nk}}. \]  

In this section, we describe in detail the application of the tetrahedron method for determining constant energy surfaces in the Brillouin zone for evaluating of integrals of type II. In what follows, we focus on a specific tetrahedron. A tetrahedron is made up of 4 \( k \) points, with the associated quantity of interest, such as the band energy \( E_{nk} \), at each vertex. The energies are ordered so that \( E_1 < E_2 < E_3 < E_4 \). Depending on the energy of the isosurface, the plane will cut the tetrahedra in one of 4 ways (see Table B.1). In cases (b) and (d), the section of the isosurface is given by the triangle, while in case (c) the quadrilateral can
Table B.1: Various cases of a plane of constant energy cutting through a tetrahedron. Depending on the energy of the isosurface relative to energies at the corners of the tetrahedron, the region of intersection can either be a triangle or a quadrilateral.

(a) $E < E_1$ or $E > E_4$: the isosurface does not intersect the tetrahedron
(b) $E < E_2$ and $E > E_1$: the isosurface intersects the tetrahedron in a triangle
(c) $E < E_3$ and $E > E_2$: the isosurface intersects the tetrahedron in a quadrilateral
(d) $E < E_4$ and $E > E_3$: the isosurface intersects the tetrahedron in a triangle

Figure B.1: Tetrahedra with a constant energy surface cutting through in each of the cases given in Table B.1.

be broken into two triangles. By considering all the tetrahedra that tessellate the Brillouin zone, the collection of triangles that are obtained by considering those tetrahedra with that contain

Linear interpolation gives the $k_\alpha$, $k_\beta$, $k_\gamma$, $k_\delta$ that define the isoenergy surface in terms
Table B.2: Vertices of the $\mathbf{k}$ points that make up the facets in Figure B.1.

(b)  \hspace{1.5cm} (c)  \hspace{1.5cm} (d)

$\mathbf{k}_\alpha = f_{2,1}\mathbf{k}_2 + f_{1,2}\mathbf{k}_1$
$\mathbf{k}_\beta = f_{3,1}\mathbf{k}_3 + f_{1,3}\mathbf{k}_1$
$\mathbf{k}_\gamma = f_{4,1}\mathbf{k}_4 + f_{1,4}\mathbf{k}_1$

$\mathbf{k}_\alpha = f_{4,2}\mathbf{k}_4 + f_{2,4}\mathbf{k}_2$
$\mathbf{k}_\beta = f_{3,2}\mathbf{k}_3 + f_{2,3}\mathbf{k}_2$
$\mathbf{k}_\gamma = f_{3,1}\mathbf{k}_3 + f_{1,3}\mathbf{k}_1$
$\mathbf{k}_\delta = f_{4,1}\mathbf{k}_4 + f_{1,4}\mathbf{k}_1$

$\mathbf{k}_\beta = f_{4,2}\mathbf{k}_4 + f_{2,4}\mathbf{k}_2$
$\mathbf{k}_\gamma = f_{4,3}\mathbf{k}_4 + f_{3,4}\mathbf{k}_3$
$\mathbf{k}_\delta = f_{4,1}\mathbf{k}_4 + f_{1,4}\mathbf{k}_1$

of the $\mathbf{k}$ points and energies $E_{nk}$ at the vertices of the tetrahedron. By defining

$$f_{n,m} \equiv \frac{\omega - \omega_m}{\omega_n - \omega_m},$$

these vertices can be expressed compactly as given in Table B.2. An isosurface is constructed as follows. For a given energy $E$, we determine all the tetrahedra for which either (b), (c), or (d) in Table B.1 is satisfied (i.e., the isosurface intersects the tetrahedra). The isosurface is linearly interpolated inside the tetrahedra and decomposed into triangles (see Figure B.1) with vertices given by the $\mathbf{k}$ points in Table B.2.

The tetrahedron method lends itself to visualization using standard graphics libraries, such as OpenGL. In fact, graphics rendering makes heavy use of tetrahedra and triangles for constructing volumes and surfaces. Representing the Fermi surface, for example, by a collection of triangles as explained above, allows for easy rendering. Not only can images (snapshots) be produced, as shown in Figure 6.1, but fully 3-dimensional objects can be constructed and manipulated within a graphical user interface window. This capability offers a unique analysis tool in which data can be visualized.
Appendix C

Derivation of the SPP-electron coupling matrix element

The derivation of the plasmon-electron matrix element within the Landau damping approximation is given in Ref. [97] and will not be discussed. Here, we derive the coupling matrix elements for the photon-electron and SPP-electron interactions given in Table 7.1. We use SI units throughout the derivations. The photon case is well known [64] and discussed here as a starting point for the SPP case. The second quantized vector potential \( \hat{A}(r,t) \) for a photon with momentum \( \hbar \mathbf{q}_p \), energy \( E_{qp} = \hbar \omega_p \), and polarization unit vector \( \mathbf{e}_{qp} \), can be obtained by associating the amplitude with a destruction operator \( \hat{a}_{qp} \) and its complex conjugate with a creation operator \( \hat{a}_{qp}^\dagger \). We use the vector potential with unit amplitude since we are interested in the coupling matrix element, and thus the transition rate per unit of incident power. It reads [64]:

\[
\hat{A}_{qp}(r,t) = f_{qp} \left[ \mathbf{e}_{qp} \hat{a}_{qp} e^{i(q_p \cdot r - \omega_p t)} + \mathbf{e}_{qp}^* \hat{a}_{qp}^\dagger e^{-i(q_p \cdot r - \omega_p t)} \right] \quad (C.1)
\]

where \( f_{qp} \) is a normalization constant. Using the Coulomb gauge \( \nabla \cdot \mathbf{A}_{qp} = 0 \), the second quantized electric field can be obtained from \( \hat{E} = -\partial \hat{A}/\partial t \).

The normalization constant \( f_{qp} \) is determined by equating the energy of the classical and quantized fields. The classical (cycle-averaged) energy \( \langle U_{qp} \rangle \) for a photon with momentum \( \hbar \mathbf{q}_p \) in a homogeneous and isotropic material with dielectric constant \( \epsilon \) is \( \langle U_{qp} \rangle = V \epsilon E_{qp}^2 / 2 \), where \( V \) is the volume of the material. This energy needs to be equal to its quantum counterpart, \( \hbar \omega_{qp} (n_{qp} + 1/2) \). In addition, \( \langle U_{qp} \rangle \) can be obtained by averaging the quantized electric field \( \hat{E}_{qp} = -\partial \hat{A}_{qp}/\partial t \) over the quantum state \( |n_{qp}\rangle \) with \( n_{qp} \) photons. We get:

\[
\langle U_{qp} \rangle = \hbar \omega_{qp} \left( n_{qp} + \frac{1}{2} \right) \quad (C.2)
\]

\[
\langle U_{qp} \rangle = V \epsilon \frac{n_{qp} \langle \hat{E}_{qp}^2 | n_{qp} \rangle}{2} = 2V \epsilon \omega_{qp}^2 f_{qp}^2 \left( n_{qp} + \frac{1}{2} \right) \quad (C.3)
\]
where in the last equation we used the time derivative of Eq. (C.1) for the electric field and the commutation rules for boson creation and annihilation operators, \([\hat{a}_{q\mathbf{p}}, \hat{a}^\dagger_{q\mathbf{p}}] = 1\).

Equating Eqs. (C.2) and (C.3) yields a normalization constant \(f_{q\mathbf{p}} = [\hbar/(2\epsilon \omega_{q\mathbf{p}})]^{1/2}\), and a vector potential:

\[
\hat{\mathbf{A}}_{q\mathbf{p}}(\mathbf{r}, t) = \left( \frac{\hbar}{2\epsilon \omega_{q\mathbf{p}}} \right)^{1/2} \left[ \mathbf{e}_{q\mathbf{p}} \hat{a}_{q\mathbf{p}} e^{i(q\mathbf{p} \cdot \mathbf{r} - \omega_{q\mathbf{p}} t)} + \mathbf{e}_{q\mathbf{p}}^* \hat{a}^\dagger_{q\mathbf{p}} e^{-i(q\mathbf{p} \cdot \mathbf{r} - \omega_{q\mathbf{p}} t)} \right]
\]

(C.4)

The photon-electron interaction Hamiltonian is \(H_{\text{int}} = \frac{\varepsilon}{m} \hat{\mathbf{A}}_{q\mathbf{p}} \cdot \hat{\mathbf{p}}, \) as obtained by expanding the kinetic energy \(|\hat{\mathbf{p}} + e\hat{\mathbf{A}}_{q\mathbf{p}}|^2/(2m)\) to first order in the vector potential. The second quantized fields associated with the electronic Bloch states are:

\[
\hat{\Psi}(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{n, \mathbf{k}} u_{n, \mathbf{k}}(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}} \hat{c}_{n, \mathbf{k}}
\]

\[
\hat{\Psi}^\dagger(\mathbf{r}) = \frac{1}{\sqrt{V}} \sum_{n', \mathbf{k}'} u_{n', \mathbf{k}'}^*(\mathbf{r}) e^{-i\mathbf{k}' \cdot \mathbf{r}} \hat{c}^\dagger_{n', \mathbf{k}'}
\]

where \(\hat{c}_{n, \mathbf{k}}\) and \(\hat{c}^\dagger_{n', \mathbf{k}'}\) are, respectively, fermion annihilation and creation operators satisfying the anticommutation rule \(\hat{c}_{n, \mathbf{k}} \hat{c}^\dagger_{n', \mathbf{k}'} + \hat{c}^\dagger_{n', \mathbf{k}'} \hat{c}_{n, \mathbf{k}} = \delta_{n, n'} \delta_{\mathbf{k}, \mathbf{k}'}\). The second quantized interaction Hamiltonian can be written as:

\[
H' = \int d\mathbf{r} \hat{\Psi}^\dagger(\mathbf{r}) H_{\text{int}} \hat{\Psi}(\mathbf{r})
\]

(C.5)

\[
= \frac{\varepsilon}{m} \sum_{n, n', \mathbf{k}, \mathbf{k}'} \left( \frac{\hbar}{2\epsilon \omega_{q\mathbf{p}}} \right)^{1/2}
\]

(C.6)

Crystal momentum conservation requires \(\mathbf{k}' = \mathbf{k} \pm \mathbf{q}_\mathbf{p}\), where the plus and minus sign apply for the first and second integral above, respectively. The interacting Hamiltonian becomes:

\[
H' = \sum_{n, n', \mathbf{k}} \left( \frac{\hbar e^2}{2V m^2 \epsilon \omega_{q\mathbf{p}}} \right)^{1/2} \left[ \mathbf{e}_{q\mathbf{p}} \cdot \mathbf{p}_{n, n', \mathbf{k}}(\mathbf{q}_\mathbf{p}) e^{-i\omega_{q\mathbf{p}} t} \hat{c}^\dagger_{n', \mathbf{k} + \mathbf{q}_\mathbf{p}} \hat{c}_{n, \mathbf{k}} \hat{a}_{q\mathbf{p}} + h.c. \right]
\]

(C.7)

where \(h.c.\) is the hermitian conjugate, and we defined the momentum matrix element as the integral \(\mathbf{p}_{n, n', \mathbf{k}}(\mathbf{q}_\mathbf{p}) = \frac{1}{\Omega} \int d\mathbf{r} u_{n, \mathbf{k} + \mathbf{q}_\mathbf{p}}^*(\mathbf{r}) (\mathbf{e}_{q\mathbf{p}} \cdot \hat{\mathbf{p}}) u_{n, \mathbf{k}}(\mathbf{r})\) over the unit cell volume \(\Omega\), as commonly used in the study of optical absorption [7]. We obtain the photon decay rate \(\Gamma_{q\mathbf{p}}\) by applying Fermi’s golden rule. Photon absorption occurs with a transition rate \(\Gamma_{q\mathbf{p}}^{\text{abs}}\) between an initial state \(|0; 1\mathbf{q}_\mathbf{p}\rangle\), consisting of the ground electronic state at room temperature and one photon with momentum \(\mathbf{q}_\mathbf{p}\), and a final state \(|S; 0\rangle\) with an electron in the Bloch state \(|n', \mathbf{k} + \mathbf{q}_\mathbf{p}\rangle\), a hole in the Bloch state \(|n, \mathbf{k}\rangle\), and no photons. The absorption rate is given by:

\[
\Gamma_{q\mathbf{p}}^{\text{abs}} = \frac{2\pi}{\hbar} \left| \langle S; 0 | H' | 0; 1\mathbf{q}_\mathbf{p} \rangle \right|^2 \delta(\hbar \omega_{q\mathbf{p}} + E_{n, \mathbf{k}} - E_{n', \mathbf{k} + \mathbf{q}_\mathbf{p}})
\]

(C.8)
Only the first term in Eq. (C.7) gives a non-zero contribution, and we get:

\[
\Gamma_{q_p}^{\text{abs}} = \frac{2\pi}{\hbar} \sum_{n,n',k} \left( \frac{\hbar e^2}{2V m^2 \epsilon \omega_{q_p}} \right) |e_{q_p} \cdot p_{n,n',k}(q_p)|^2 f_{n,k} (1 - f_{n',k+q_p}) \delta(\hbar \omega_{q_p} + E_{n,k} - E_{n',k+q_p})
\]

The inverse process of photon emission from an initial state |S; 0\rangle to a final state |0; I_{q_p}\rangle has a rate \( \Gamma_{q_p}^{\text{em}} \) given by:

\[
\Gamma_{q_p}^{\text{em}} = \frac{2\pi}{\hbar} \sum_{n,n',k} \left( \frac{\hbar e^2}{2V m^2 \epsilon \omega_{q_p}} \right) |e_{q_p} \cdot p_{n,n',k}(q_p)|^2 f_{n',k+q_p} (1 - f_{n,k}) \delta(\hbar \omega_{q_p} + E_{n,k} - E_{n',k+q_p})
\]

The net rate \( \Gamma_{q_p} = \Gamma_{q_p}^{\text{abs}} - \Gamma_{q_p}^{\text{em}} \) is the decay rate of a photon with momentum \( \hbar q_p \) and energy \( E_p = \hbar \omega_{q_p} \) into a hot electron-hole pair. Writing the delta function as a Lorentzian with broadening \( \eta \rightarrow 0 \), namely \( \delta(x) = \frac{1}{\pi} \text{Im} \frac{1}{x - i\eta} \), we obtain the decay rate as:

\[
\Gamma_{q_p} = \frac{2\pi}{\hbar} \sum_{n,n',k} \left( \frac{\hbar e^2}{2V m^2 \epsilon \omega_{q_p}} \right) |e_{q_p} \cdot p_{n,n',k}(q_p)|^2 \frac{1}{\pi} \text{Im} \left[ \frac{f_{n,k} - f_{n',k+q_p}}{E_p(q_p) - (E_{n',k+q_p} - E_{n,k}) - i\eta} \right]
\]

This expression is identical to Eq. (7.1) with the squared coupling matrix elements \( |g_{n,n',k}^{\text{PH}}|^2 \) in Table 7.1. Since for photons \( q_p \approx 0 \) compared to the size of the BZ, the squared coupling matrix elements are:

\[
|g_{n,n',k}^{\text{PH}}|^2 = \left( \frac{\epsilon \hbar}{m} \right)^2 \frac{1}{2\epsilon E_p V} |e_{q_p} \cdot p_{n,n',k}(q_p = 0)|^2
\]

as shown in Table 7.1. We note that for slowly varying matrix elements in the BZ, the photon decay rate is proportional to the zero-momentum JDOS commonly used in the theory of the optical properties of solids [7], i.e., \( \Gamma_{q_p} \propto J(q_p = 0, E_p) \).

The derivation of the coupling matrix elements \( |g_{n,n',k}^{\text{SPP}}|^2 \) for SPPs follows similar steps to the photon case discussed above. Following Nkoma et al. [87], we study the case of a SPP generated at the interface between a metal with frequency-dependent dielectric function \( \epsilon_M(\omega) \) and a dielectric material with frequency-independent dielectric constant \( \epsilon_p \). A similar second quantization procedure was also developed by Elson et al. [27]. It is well known that the intensity of the SPP wave decays exponentially in the direction perpendicular to the dielectric-metal interface [4, 101]. To study HC generation, we focus here on the SPP field in the metal, and its associated vector potential. Since SPP possess transverse magnetic (TM) polarization, their wavevector can be taken as real in the \( xy \) plane — here, the plane of the metal-dielectric interface, with area \( A \) — and purely imaginary in the \( z \) direction normal to the interface. With the metal in the \( z > 0 \) space, and the \( x \)-axis parallel to the propagation direction of the SPP wave, the SPP wavevector is \( q_p = (q_{p,x}, 0, i q_{p,z}) \). While in the main text we define \( q_p = (q_{p,x}, 0, 0) \) as the propagation wavevector in the \( xy \) plane, for the purpose of this derivation \( q_p \) is defined differently. Here, \( q_{p,x} \) is a real positive number since the
electron-induced dissipation effects we aim to find would be encoded in its imaginary part, and \( q_{p,z} \) is related to the decay length \( L_z \) of the SPP field intensity by \([101]\) \( q_{p,z} = 1/(2L_z) \). The relative magnitude of the components of \( \mathbf{q}_p \) are determined by the dielectric properties of the metal and dielectric \([4, 101]\), which impose \( q_{p,x}/q_{p,z} = |\sqrt{\varepsilon_D/\varepsilon_M}| \) (here, the real parts of \( \varepsilon_D \) and \( \varepsilon_M \) should be employed for consistency). Similarly, the components of the unit polarization vector \( \mathbf{e}_{q_p}^{TM} \) are related by \([101]\) \( (\mathbf{e}_{q_p}^{TM})_x/(\mathbf{e}_{q_p}^{TM})_z = |\sqrt{\varepsilon_D/\varepsilon_M}| \).

The vector potential of the SPP wave in the metal, with unit amplitude as in the photon case, can be written in second quantized form as:

\[
\mathbf{A}_{q_p}(\mathbf{r}, t) = f_{q_p} \left[ \mathbf{e}_{q_p}^{TM} \hat{a}_{q_p} e^{i(\mathbf{q}_p \cdot \mathbf{r} - \omega_{p}t)} + (\mathbf{e}_{q_p}^{TM})^\dagger \hat{a}_{q_p}^\dagger e^{-i(\mathbf{q}_p \cdot \mathbf{r} - \omega_{p}t)} \right] \tag{C.10}
\]

where, as in the photon case, the normalization constant \( f_{q_p} \) can be determined by energy considerations. We write the classical electric field associated with the SPP in the metal as \( \mathbf{E} = \mathbf{E}_q e^{i(\mathbf{q}_p \cdot \mathbf{r} - \omega_{p}t)} \). Using a result in Ref. \([87]\), the cycle-averaged energy in the SPP field — which accounts for contributions from both the field in the dielectric and the metal — can be written as \( \langle U_{q_p} \rangle = A\varepsilon_{eff}|E_0|^2/(4\,q_{p,z}) \), where we defined an effective dielectric constant:

\[
\varepsilon_{eff}(\omega) = \frac{\varepsilon_M(\omega)[\varepsilon_D + \varepsilon_M(\omega)\varepsilon_D]}{2} + \frac{\omega}{2} \frac{\partial \varepsilon_M(\omega)}{\partial \omega}
\]

We use the correspondence principle for a quantum state \( |n_{q_p}\rangle \) with \( n_{q_p} \) SPP quanta. Similar to the photon case, the total energy in the SPP mode \( \hbar \omega_{q_p}(n_{q_p} + 1/2) \) can be equated to its classical counterpart, while the square modulus of the electric field \( |E_{q_p}|^2 \) can be replaced with its quantum mechanical average \( \langle n_{q_p} | \mathbf{E}_{q_p}^2 | n_{q_p} \rangle \), with \( \mathbf{E}_{q_p} = -\partial \mathbf{A}_{q_p}/\partial t \). We obtain the two equations:

\[
\langle U_{q_p} \rangle = \hbar \omega_{q_p} \left( n_{q_p} + \frac{1}{2} \right) \tag{C.11}
\]

\[
\langle U_{q_p} \rangle = \frac{\varepsilon_{eff} A}{4\,q_{p,z}} \left| n_{q_p} \right| \left( \mathbf{E}_{q_p}^2 \right) \left| n_{q_p} \right| e^{(2\,q_{p,z})} = \left( \frac{A}{q_{p,z}} \right) \varepsilon_{eff} \omega_{q_p}^2 f_{q_p}^2 \left( n_{q_p} + \frac{1}{2} \right) \tag{C.12}
\]

By equating Eqs. \( C.11 \) and \( C.12 \), and using \( q_{p,z} = 1/(2\,L_z) \), we obtain the normalization constant \( f_{q_p} = [\hbar/(2AL_z\varepsilon_{eff} \omega_{q_p})]^{1/2} \), and thus the SPP vector potential:

\[
\mathbf{A}_{q_p}(\mathbf{r}, t) = \left( \frac{\hbar}{2AL_z\varepsilon_{eff} \omega_{q_p}} \right)^{1/2} \left[ \mathbf{e}_{q_p}^{TM} \hat{a}_{q_p} e^{i(\mathbf{q}_p \cdot \mathbf{r} - \omega_{p}t)} + (\mathbf{e}_{q_p}^{TM})^\dagger \hat{a}_{q_p}^\dagger e^{-i(\mathbf{q}_p \cdot \mathbf{r} - \omega_{p}t)} \right] \tag{C.13}
\]

This vector potential is similar to the photon case, Eq. \((C.4)\). The key differences are the decay in the \( z \) direction for the SPP field, which introduces an effective SPP volume \( AL_z \) in the vector potential, the appearance of an effective dielectric constant for the medium, and the fact that the polarization is TM with a direction imposed by the properties of the metal and dielectric material.

Since the form of the SPP vector potential is completely analogous to the one for photons,
the decay rate of SPPs to hot electron-hole pairs can be carried out following the same steps as in the photon case discussed above, with the substitutions $V \rightarrow AL_z$, $e_{q_p} \rightarrow e_{q_p}^{TM}$, and allowing for a finite momentum $[4]$ $q_p$ with magnitude $q_p = (E_p/hc)\sqrt{\epsilon_M\epsilon_D}/(\epsilon_M + \epsilon_D)$.

This leads to the SPP-electron coupling matrix element in Table 7.1:

$$|g_{n,n',k}^{SPP}(q_p, E_p)|^2 = \left(\frac{e\hbar}{m}\right)^2 \frac{1}{2\epsilon_{eff} E_p AL_z} \left|e_{q_p}^{TM} \cdot p_{n,n',k}(q_p)\right|^2.$$  (C.14)
Bibliography


[138] Y. Xia et al. “Observation of a large-gap topological-insulator class with a single Dirac cone on the surface”. In: Nat. Phys. 5.6 (June 2009), pp. 398–402. ISSN: 1745-2473. DOI: 10.1038/nphys1274. URL: http://dx.doi.org/10.1038/nphys1274.


[142] Haijun Zhang et al. “Topological insulators in Bi2Se3, Bi2Te3 and Sb2Te3 with a single Dirac cone on the surface”. In: Nat. Phys. 5.6 (June 2009), pp. 438–442. ISSN: 1745-2473. DOI: 10.1038/nphys1270. URL: http://dx.doi.org/10.1038/nphys1270.

