Title
Electronic and Optical Properties of Solids with Strong Spin-Orbit Coupling

Author
Barker, Bradford Alan

Publication Date
2018

Electronic and Optical Properties of Solids with Strong Spin-Orbit Coupling

by

Bradford Alan Barker

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 Jeffrey B. Neaton
Professor Mark Asta

Spring 2018

Electronic and Optical Properties of Solids with Strong Spin-Orbit Coupling

Abstract

Electronic and Optical Properties of Solids with Strong Spin-Orbit Coupling

by

Bradford Alan Barker

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Steven G. Louie, Chair

The development of new technology for computing and renewable energy sources requires new insight into the physics governing state-of-the-art materials for these applications. To optimize the search for transistors and solar cells to supplant silicon, it is desirable to have them investigated in advance of their large-scale manufacture. One potentially fruitful avenue of investigation is the study of the electronic and optical properties of materials containing heavy atoms. Such atoms have a large spin-orbit coupling, which can be responsible for relatively exotic physics. Topological insulator materials such as Bi_2Se_3 may have utility in the development of, for example, spin-tronics, in which information may be conveyed without the need for transporting electrical charge.

The details of charges moving through a material, as well as a material absorbing light, require a suitable theoretical treatment. Within the purview of the quantum theory of solids, the technique of many-body perturbation theory gives researchers access to the means of calculating one-particle and two-particle excited states, the exact scenario needed to understand charged excitations and optical absorption, respectively.

In this work, we further extend the ability of the many-body perturbation theory software package of BerkeleyGW to allow for more accurate description of solids containing heavy atoms. Namely, we investigate the properties of materials with large spin-orbit coupling by allowing for the treatment of two-component spinor wavefunctions. In the introduction, we review the physics of one- and two-particle excitations, entirely within a formalism allowing for the presence of spin-orbit coupling. In Chapter 2, we further discuss the implementation of spinor wavefunction functionality in BerkeleyGW and provide many test calculations using materials with varying strengths of spin-orbit coupling, with varying geometries, and including the

metallic system of bulk gold. In Chapter 3, we present a calculation of the quasiparticle bandstructure of β-HgS as a further benchmark material, for which there requires elucidation of the bandstructure topology. We find very close agreement with experiment for both the effective mass and band gap. In Chapter 4, we present the bandstructure of the prototypical topological insulator $\mathrm{Bi_2Se_3}$, finding significant qualitative differences in the bandstructure from a quasiparticle calculation and the more readily available description from Density Functional Theory (DFT). Namely, we find that, in agreement with experiment, the conduction and valence bands are both nearly parabolic, in contrast to the well-known camel-back feature in the valence band of previous DFT calculations. Finally, in Chapter 5, we use DFT calculations to determine the ground-state geometry of Ir dimers adsorbed to graphene and confirm this geometry, that of a horizontal dimer across the so-called bridge sites, by comparing the resulting density of states to that measured by experiment. We find both have a strong central peak near the graphene Dirac point energy.

To my parents,
Kirk David Barker and Margot Schaaf Barker

Contents

List of Figures

List of Tables

Acknowledgments

I would like to thank my parents, Kirk Barker and Margot Barker, who have always given me unconditional support and love.

I would like to thank my brother, Charlie Barker, as well as his wife, Laura, daughter Madeline Barker, and son-to-be. Charlie, as the older brother, has been a great role-model for me, beginning from an early age as a student, and now as a family man. I still recall spending most of my fourth grade bragging to anyone about how good at math and basketball and football my big brother was. Laura has been a perfect sister-in-law, and Maddy and Baby Boy Barker are already so, so special.

I would like to thank my adviser, Steven Louie, for his guidance during my tenure in his research group. Early on, group secretary Katherine mentioned that this was the best in the world, and I gradually came to realize that there was no actual hyperbole in that statement. It took a bit more unexpected work than anyone, really, was happy with, but I believe I delivered on the opportunity given to me.

Speaking of, I always appreciated Katherine de Raadt, former group secretary. She would do her best to keep her off-topic conversations with me to a reasonable level of distraction, but they were usually the best parts of the day when they happened.

In my time in the Louie group, I was privileged to have the opportunity to learn from a large number of top scientists. I began work in the group at the tail end of one of the great generations of graduate students, with my first month being the last week of Cheol-Hwan Park. While still fighting "uphill" in research, I got help from Brad Malone, Jack Deslippe, Oleg Yazyev, Georgy Samsonidze, Manish Jain, Johannes Lischner, and David Strubbe. And then another all-time generation was folded into the group, too, with Derek Vigil-Fowler, Jamal Mustafa, Sinisa Coh, Felipe Jornada, Diana Qiu, Ting Cao, Marco Bernardi, Zhenglu Li, Mauro del Ben, and Gabriel Antonius. Everyone mentioned really helped me out, from honing my chops in Python and Fortran, general many-body theory and solid-state physics, and high-throughput computing.

For direct collaborators, I am honored to thank Jack Deslippe, Oleg Yazyev, Manish Jain, Johannes Lischner, Sinisa Coh, Marco Bernardi, Aaron Bradley, Miguel Moreno Ugeda, and Derek Vigil-Fowler. Extra special thanks to Derek, long-time office-mate and friend. I would not have been able to graduate without his support.

I also must credit Johannes Lischner, Manish Jain, and David Strubbe for their mentorship through a critical point in my graduate school career. They knew that I felt defeated after a long string of failures and were determined that I learn how to tie that string into a rope ladder to get me back on the path to success. I suppose I inherited the title of Most Recognizable Fifth Floor Birge Hall Laugh from Manish, which will now necessarily pass to office-mate Chin-Shen Ong (with whom I will miss

"Night Shift" laughs and physics discussions). I also inherited the title of "Group Stenographer" from David, which was mostly keeping track of Steve's and Sinisa's weekly turns of phrase. Time will tell if this tradition continues, but I will say that, at present (and with a modicum of bias), I am in a three-way tie for fourth place in notable group-meeting quotes, along with Felipe and Feliciano Giustiano. (Steve, as of May 2018, is many horse-lengths ahead at 93 quotes; Marvin L. Cohen, 33; Sinisa, 31; then my cohort with 30. Let us temporarily value this in a similar manner as, say, the h-index.)

Many thanks, again, to all of the members of the Louie and Marvin Cohen groups during my time in graduate school. I enjoyed the most working alongside such amazing, brilliant people, all of you.

I would like to thank my friends, inside and outside of graduate school. My former long-time roommates, Kevin Grosvenor and Kacey Meaker; my current long-time roommate, Austin Hedeman; my social hubs, Sharif Corinaldi, Nicole Carlson, Mike Ramm, Michele Kotiuga, and Jaime Hinton; my Compass comrades, Hilary Jacks, John Haberstroh, Jesse Livezey, and Dimitri Dounas-Frazier. Wonderful people and inspirations, all.

I would like to thank my "West Coast Family," Sal Fahey, Libby Schaaf, Dominic Fahey, Lena Fahey, Kurt Meyer, Chris Schaaf, Bob Shock, Barbara Schaaf, Dick Schaaf, and Mary-Lynn Schaaf. I am happy that my half-joke of Dominic driving me to my commencement ceremony did not actually come to pass, though I did get to see Lena grow up, with her age keeping track of my year in grad school. (No pressure, but I believe we have another set of scientists and engineers, and role models for Maddy and Baby Barker.)

I would like to thank my friends Eric Lewis and Rana McAnear, especially for bearing much of the brunt of my venting about graduate school life, and lifelong friend and current neighbor just down 880, Andrew Fedock. I remember in first grade (we ended up having every grade together as well as both going to the University of Florida), Andy telling me about some exotic school system that exists after college, which can last for years and years, the end of which requires writing a "book" with hundreds of pages, and being a bit in awe. (After a decade of work, it all seems so easy now.) We'll all have a laugh about all of this, as we always do about everything.

I would like to thank my friends made during high school and college, Lauren Egan O'Shea, Chris Mallard (who got in the habit of calling me "Dr. Barker" much, much too early), Sheila Suarez de Flores, Austin Flores, Matthew Yampolsky, Erica Cook Reott, Julie McDonald, Tracy Ploch, Megan Powers, Sarah Joy, Jesse Arost, Chase McGee, Don Burnette, Lindsey Gray, Erica (Bolin) Littauer, and Paden Woodruff. I wish I could say that I gave as much support as I got from you all. Extra special love and thanks to the Steele family and the Boyce family.

School and my teachers have defined so much of my life, and I could not be more proud of my public school education. I would like to thank my elementary school teachers Tom “Mr. Fil” Filipkowski and Cathy Oliver. Cathy (“Mrs. Sellers” at the time) was my teacher or co-teacher for first, third, fourth and fifth grades; it is especially tragic to learn that she passed so shortly after retirement from her beloved Morningside Elementary. Mr. Fil was my co-teacher during fifth grade, where I received my first C on a science assignment. I always wanted to impress Mr. Fil since he was so influential to my big brother, who had him as his sole teacher for third through fifth grades. Mr. Fil was the first to teach me about physics (actually in an “elective” in first grade, and again in fifth), so in a way, it all starts with him teaching me about “the four forces on an airplane.”

By sheer luck, I had so many great middle school teachers. Friends who also were Scorpion Team in Northport Middle School can still recall individual lessons from Maryalice Deptula, “Mr. Cap” Louis Capizola, and Mrs. Webb. I remember Mr. Cap capturing my imagination with his description of particle physics as early as when my elementary school went on field trips to our future middle schools. Ms. Deptula, despite being a masterful English and Reading teacher, also had a few digressions on physics from Einstein and beyond, which was equally influential on me.

In high school, of course John Pillsbury places above all others. He was able to get me to see physics to be a subject unlike any other, at exactly the right time in my life. More than just the subject matter, he taught what I later came to understand as scientific epistemology, and his lessons still contribute to my foundations in understanding that subject. He noticed my talent in the subject and did what he could to nurture my talent like a parent. I would place my success in undergrad squarely on, particularly, him allowing me to teach the last two months of Advanced Placement Physics in his retirement year, perhaps as a part of his legacy in education.

College, of course, derives most of its value from the ability to learn directly from the best. I am more than happy to have learned from Selman Hershfield, John Klauder, Stephen Hill, Rick Field, Douglas Cenzer, and Rick Smith. Dr. Hershfield was my Physics Department faculty adviser, and I was a constant presence in his office immediately after starting at UF. I later took his graduate quantum mechanics sequence and loved it, apart from the lesson on Field Theory and Green’s Functions (’When is that important?’ I ask in 2007 and answer in detail in Chapter 1 of this dissertation). He advised my pair of research projects and helped me transition into graduate school, though it was often the case that we manged my various personal crises. John Klauder is simply the sweetest man on the planet, in addition to being a decorated lecturer and scholar. Professors Stephen Hill, Rick Field, Douglas Cenzer, and Rick Smith were all very encouraging in my effort in pursuing a research career, and I took as many courses from them as possible.

My summer in 2006 spent at the University of Indiana in lovely Bloomington was also a huge influence. My research advisers for the 10 week Research Experience for Undergraduates were Timothy Londergan and David Murdock. David and I kept in touch well into grad school, exchanging similarly droll remarks. Track down his lecture notes for his courses at Tennessee Technology University to see the hammiest humor in physics writing. He had strong opinions on whether the letter "d" in differentials should be italicized or not; in his honor, I use his convention throughout this work. It was devastating to learn of his passing, though Professor Londergan and his former students are assisting with his family to keep his memory alive.

I would also like to thank you for reading this thesis, especially if you are a developer of first-principles materials science software looking for lessons on incorporating the treatment of relativistic effects in your code. I hope that the first two chapters prove helpful and contribute in addition to the references cited therein. I did read the occasional thesis from scientists outside of my own research group when doing background reading on various subjects, so I have always hoped that I will be participating in this tradition but now as the ink on the page.

I am happy that the broader physics community will at long last have access to the developments in the BerkeleyGW code that I was fortunate to contribute. Further features extending the utility of spinor-dependent physics will be developed further by the present generation of students in the Louie Group, though I will be happy to field questions if there are any user issues with the specific features that I implemented, especially if the year is within a decade or two of 2018, the year in which I finally graduated from UC Berkeley.

Chapter 1

Introduction

1.1 Approximate methods for the calculation of the electronic properties of real materials

The Electronic Hamiltonian

The calculation of the electronic properties of materials requires a quantum mechanical description of the systems of interest, be they single molecules, quantum dots, nanotubes, quasi-two dimensional sheets, or crystals. After having a sense of the species of the constituent atoms and roughly their location in space, we might first think to calculate the energy spectra and wavefunctions with the full Hamiltonian of the system. It is a simple matter to define such a many-body Hamiltonian, with n electrons and N ions interacting via the Coulomb potential:

$$H = -\sum_{\mathrm{I}}^{N} \frac{\nabla_{\mathrm{I}}^2}{2M_{\mathrm{I}}} - \sum_{\mathrm{i}}^{n} \frac{\nabla_{\mathrm{i}}^2}{2m_e} + \frac{e^2}{2} \sum_{\mathrm{I,J}\neq\mathrm{I}}^{N,N} \frac{Z_{\mathrm{I}} Z_{\mathrm{J}}}{|\mathbf{R}_{\mathrm{I}} - \mathbf{R}_{\mathrm{J}}|} + \frac{e^2}{2} \sum_{\mathrm{i,j}\neq\mathrm{i}}^{n,n} \frac{1}{|\mathbf{r}_{\mathrm{i}} - \mathbf{r}_{\mathrm{j}}|} - e^2 \sum_{\mathrm{i,I}}^{n,N} \frac{Z_I}{|\mathbf{r}_{\mathrm{i}} - \mathbf{R}_{\mathrm{I}}|}, \tag{1.1}$$

where the N ions have masses $\{M_{\mathrm{I}}\}$ and charges $\{Z_{\mathrm{I}}\}$. Without the need to include the physics of electron-phonon scattering or temperature effects, we may fix the ionic coordinates $\{\mathbf{R}_{\mathrm{I}}\}$ in space and simplify the Hamiltonian for the electrons in the Born-Oppenheimer approximation:

$$H^e = -\sum_{\mathrm{i}}^{n} \frac{\nabla_{\mathrm{i}}^2}{2m_e} - e^2 \sum_{\mathrm{i,I}}^{n,N} \frac{Z_I}{|\mathbf{r}_{\mathrm{i}} - \mathbf{R}_{\mathrm{I}}|} + \frac{e^2}{2} \sum_{\mathrm{i,j}\neq\mathrm{i}}^{n,n} \frac{1}{|\mathbf{r}_{\mathrm{i}} - \mathbf{r}_{\mathrm{j}}|}. \tag{1.2}$$

We may even choose to include terms incorporating relativistic effects to order

c^{-2} with the inclusion of the relativistic Hamiltonian(Ref. [106]):

$$H^{\mathrm{rel}} = -\sum_{\mathrm{i,I}}^{n,N} \frac{1}{8m_e^2c^2}\nabla_{\mathrm{i}}^2 V_{\mathrm{I}}(\mathbf{r}_{\mathrm{i}}) - \sum_{\mathrm{i}}^{n} \frac{\nabla_{\mathrm{i}}^4}{8m_e^3c^2} - \sum_{\mathrm{i,I}}^{n,N} \frac{e^2}{2m_e^2c^2}\mathbf{S}\cdot\mathbf{p}_{\mathrm{i}} \times \nabla_{\mathrm{i}}\left(V_{\mathrm{I}}(\mathbf{r}_{\mathrm{i}})\right) \tag{1.3}$$

The terms on the right-hand side of Eq. 1.3 are the Darwin, relativistic mass correction, and spin-orbit coupling terms, respectively. We use $V_{\mathrm{I}}(\mathbf{r}_{\mathrm{i}})$ to denote the potential from ion I for electron i.

Even for a single helium atom in a vacuum, however, solving such a many-electron Hamiltonian requires further approximations. A complete solution for larger systems with interacting electrons is intractable, and the many-body wavefunction would likely be inscrutable. If, however, we restrict the space of solutions in a manner that allows us to capture much of the desired physics, we may recast this problem into one more readily solved.

The Hartree theory proposes a ground-state wavefunction that is simply a product of single-particle orbitals. The Hartree theory decomposes the many-body Hamiltonian into a single-body Hamiltonian for each orbital by minimizing the Hamiltonian matrix elements, subject to the constraint that the orbitals are normalized[22]. This gives us the single-particle Hamiltonian

$$\begin{aligned}
h^{\mathrm{H}}\phi_{\mathrm{i}}(\mathbf{r}) &= t_{\mathrm{i}}(\mathbf{r})\phi_{\mathrm{i}}(\mathbf{r}) + v^{\mathrm{H}}(\mathbf{r})\phi_{\mathrm{i}}(\mathbf{r}),\\
t_{\mathrm{i}}(\mathbf{r}) &= -\frac{1}{2}\nabla^2 + v^{\mathrm{ion}}(\mathbf{r}),\\
v^{\mathrm{H}}(\mathbf{r}) &= \int \mathrm{d}\mathbf{r}'\frac{n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}.
\end{aligned} \tag{1.4}$$

The term v^{H} is called the Hartree potential, which takes the electron-electron interaction to be merely the direct electrostatic interaction between a single particle and a static charge density $n(\mathbf{r})$ formed by the other electrons in the system. Since the density is constructed from all occupied orbitals, the Hamiltonian must be solved in a self-consistent manner, beginning with a trial set of orbitals with which to construct the density. This is not the correct description of the full electron-electron interaction, but it is a useful first attempt. (In fact, we will see in Sec. 1.3 that we are only a “mass term” away from being able to truly calculate excited-state properties.)

The Hartree-Fock theory, in a similar procedure, requires the ground-state wavefunction to be an antisymmetrized product of orbitals, i.e., a single Slater determinant. This additional requirement allows for the treatment of exchange, a non-local effective electron-electron interaction due to the electrons being indistinguishable. The single-electron Hartree-Fock Hamiltonian is

$$h^{\mathrm{HF}}\phi_{\mathrm{i}}(\mathbf{r}) = h^{\mathrm{H}}\phi_{\mathrm{i}}(\mathbf{r}) - \int \mathrm{d}\mathbf{r}' \frac{n(\mathbf{r},\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\phi_{\mathrm{i}}(\mathbf{r}')\chi_{\mathrm{i}}. \tag{1.5}$$

The additional term on the right-hand side is the "bare exchange," which is non-local. The spinor χ_{i} describes the spin of the state i. The co-density in the bare exchange is defined as $n(\mathbf{r},\mathbf{r}') = e^2 \sum_{\mathrm{j}} \phi^*_{\mathrm{j}}(\mathbf{r}')\phi_{\mathrm{j}}(\mathbf{r})\chi^\dagger_{\mathrm{j}}$. We see, then, that the bare exchange only gives a contribution when the spins of the states i and j are aligned.

The sign of the bare exchange is significant. Electrons with the same spin will keep away from each other due to the need to enforce antisymmetry of the many-body wavefunction, as in the anti-bonding state of molecular hydrogen. There is, then, effectively a cloud of hole states surrounding the electron under consideration, which itself acts positively charged. The interaction with this "exchange-hole" lowers the single-particle energy.

As stated before, Hartree-Fock theory assumes that the ground-state wavefunction is a single Slater determinant. Such a wavefunction however is only an eigenstate of a Hamiltonian neglecting electron-electron interactions. In the presence of interactions, the eigenstate is an appropriate linear combination of Slater determinants. These Slater determinants in principle include an arbitrary number of unoccupied states. The difference in energy of the real system and its description within Hartree-Fock is called the correlation energy. Some approaches in quantum chemistry attempt to solve for the dominant contributions to the sum of Slater determinants comprising the ground-state wavefunction (Ref. [107]); fortunately, there are other techniques that do not require a direct calculation of the many-body wavefunction.

Density Functional Theory

An especially useful – and, in principle, exact – approach for ground-state propertis is Density Functional Theory ("DFT"). Instead of proposing an ansatz for the many-body wavefunction, Density Functional Theory treats the electronic density at the ground state as the most meaningful quantity that describes the system of interest. The Hohenberg-Kohn theorem[46] demonstrates that there exists a universal functional of the density that can be minimized to determine the ground state energy. The Kohn-Sham theorem[60] constructs a fictitious non-interacting system that happens to have the same ground-state density as in the interacting system. Since this system is non-interacting, the many-body wavefunction is a single Slater determinant, and the complicated exchange and correlation effects of the real system are taken into account by some unknown additional potential term, which is, in turn, a functional of the actual density.

The Kohn-Sham equation (neglecting relativistic effects momentarily) is

$$\left(-\frac{1}{2}\nabla^2 + v^{\mathrm{KS}}[n](\mathbf{r})\right)\psi_{\mathrm{i}}^{\mathrm{KS}}(\mathbf{r}) = \epsilon_{\mathrm{i}}\psi_{\mathrm{i}}^{\mathrm{KS}}(\mathbf{r}). \tag{1.6}$$

In a self-consistent procedure, the eigenfunctions are used to construct the ground state charge density:

$$n(\mathbf{r}) = \sum_{\mathrm{i}} f_{\mathrm{i}} \left|\psi_{\mathrm{i}}^{\mathrm{KS}}(\mathbf{r})\right|, \tag{1.7}$$

where f_{i} is the occupation number of the state i.

The Kohn-Sham potential is defined by

$$v^{\mathrm{KS}}[n](\mathbf{r}) = v^{\mathrm{H}}[n](\mathbf{r}) + v^{\mathrm{ext}}(\mathbf{r}) + v^{\mathrm{XC}}[n](\mathbf{r}), \tag{1.8}$$

with the Hartree potential as defined from Eq. 1.4 and the exchange-correlation potential

$$v^{\mathrm{XC}}[n](\mathbf{r}) = \frac{\delta E^{\mathrm{XC}}}{\delta n(\mathbf{r})}. \tag{1.9}$$

If we had access to the exact exchange-correlation functional E^{XC}, this procedure would give exact ground state densities and total energies. In practice, we make due with approximations.

The Local Density Approximation ("LDA") approximates the system as a locally homogeneous electron gas. The exchange energy density is taken to be the Slater exchange[108] and the correlation energy density is parameterized from Monte Carlo simulations of the homogeneous electron gas[18]. Multiple choices for this parameterization exist, such as Perdew-Zunger[90] and Perdew-Wang[92]. It is well-known that there is a cancellation of errors in the exchange and correlation terms[61] that fortuitously leads to a good estimation of binding in many systems. The LDA typically over-binds, underestimating lattice parameters and bond lengths. (An acceptable lattice constant computed within the LDA is within about two percent of that of experiment.) We must note that Wigner constructed by hand a correlation function in 1934[129] that compares admirably to the Quantum Monte Carlo result for realistic densities[22].

The Generalized Gradient Approximation ("GGA") adds in a dependence on the gradient of the density to the exchange-correlation energy density. The parameterization developed by Perdew, Burke, and Ernzerhof ("PBE")[91] is successful and widely used, though it tends to underbind in solids. Some different parameterizations based on PBE attempt to improve the quality of results, such as revPBE[137] ("revised PBE") and PBEsol[93] ("PBE for solids"). Meta-GGA functionals, such as SCAN[112], also include a dependence on kinetic energy.

We must note that the single-particle energies ϵ^{KS} are merely Lagrange multipliers with no intrinsic physical meaning. These energies and the corresponding orbitals ϕ^{KS} are constructed to allow a practical means of calculating the ground state energy and density. An accurate description of the excitations above the ground state requires a different theory whose purpose is the description of such states.

Pseudopotentials

That said, DFT is still a powerful computational tool and merits further description. Certain approaches to DFT explicitly treat all of the electrons in the system being studied. These approaches are called "all-electron" and often require a separation of the states into a core region near the ions, where states are localized, and an interstitial region, where the states behave more like plane-waves.

Such a description would treat, say, $1s$ electrons near a bismuth ion on a similar footing to the chemically active $6p$ electrons. Indeed, the recognition that the set of completed shells isoelectronic with the neighboring noble gas form a hard "core" of states near a bare ion allows they be treated as a single rigid object. This pseudo-ion then has node-free valence states that interact with it via a pseudopotential. (Similar reasoning inspired Hellman to develop the notion of pseudopotentials for chemistry [43] and Enrico Fermi[30] in the context of nuclear matter. A solid understanding of chemistry led J. C. Philips[94] and Marvin L. Cohen to reintroduce the concept to materials science applications.)

This separation into core and valence electrons leads to an effictive potential with a significantly weaker interaction between the valence electrons and the ion. This insight is from the Philips-Kleinman Cancellation Theorem[95]. If we label the single-particle orbitals with many nodes from an all-electron treatment as $|\psi_{\mathrm{i}}\rangle$ and the node-free pseudowavefunction as $|\phi_{\mathrm{i}}\rangle$, we have

$$|\psi_{\mathrm{i}}\rangle = |\phi_{\mathrm{i}}\rangle + \sum_{\mathrm{c}} |f_{\mathrm{c}}\rangle\langle f_{\mathrm{c}}|\phi_{\mathrm{i}}, \tag{1.10}$$

where $|f_{\mathrm{c}}\rangle$ is a core state. For a Schrödinger-like equation (say, the Kohn-Sham equation) with the all-electron state, we have

$$\left(-\frac{\nabla^2}{2m_e} + V(\mathbf{r})\right)\psi_{\mathrm{i}}(\mathbf{r}) = \epsilon_{\mathrm{i}}\psi_{\mathrm{i}}(\mathbf{r}), \tag{1.11}$$

while the pseudo-orbital satisfies the modified equation

$$\left(-\frac{\nabla^2}{2m} + V(\mathbf{r}) + V_{\mathrm{R}}(\mathbf{r})\right)\phi_{\mathrm{i}}(\mathbf{r}) = \epsilon_{\mathrm{i}}\phi_{\mathrm{i}}(\mathbf{r}). \tag{1.12}$$

In principle, we solve for the same eigenvalue ϵ_{i} but with a node-free state, provided we first compute the pseudopotential $V_{\mathrm{ps}} = V + V_{\mathrm{R}}$, with

$$V_{\mathrm{R}} = \sum_{\mathrm{c}} (\epsilon - \epsilon_{\mathrm{c}})|f_{\mathrm{c}}\rangle\langle f_{\mathrm{c}}|. \tag{1.13}$$

The pseudopotential, then, is energy-dependent and non-local, as implied by the outer-product of core states. The energy-dependence, however, is weak, as the core energies are lower by orders of magnitude than the energy range in which we seek to find the valence state eigenvalue ϵ_{i}, usually a range of tens of eV. As expected, this term arising from the core-electron's potential energy term contributes a repulsive interaction on the same order as the interaction of the valence state with the bare ion[70].

There are multiple approaches to pseudopotential construction. Semi-local pseudopotentials are local in the radial coordinate but non-local in the angular coordinates. The Kleinman-Bylander form of pseudopotentials[59] constructs pseudopotentials as having some arbitrary local term v_{Loc} and a sum of fully non-local terms:

$$v_{PS}^{\mathrm{KB}}(\mathbf{r}) = v_{\mathrm{Loc}}(\mathbf{r}) + \sum_{\mathrm{i}} \frac{|\delta V\phi_{\mathrm{i}}\rangle\langle\delta V\phi_{\mathrm{i}}|}{\langle\phi_{\mathrm{i}}|\delta V|\phi_{\mathrm{i}}\rangle}, \tag{1.14}$$

where $\delta V = \left(\epsilon_{\mathrm{i}}^{\mathrm{AE}} - T - v_{\mathrm{Loc}}\right)$ and ϕ_{i} is the wavefunction for state i within the core region, and the energy $\epsilon_{\mathrm{i}}^{\mathrm{AE}}$ refers to that computed from an all-electron calculation. The advantage to this form of pseudopotential is that it simplifies the computation of matrix elements of the psuedopotential, with functions in an N-dimensional basis, from having N^2 non-local integrals involving the radial part of the functions u_{i} within the previous "semi-local" treatment, to having pairs of N local integrals where only one of the pair requires explicit calculation[59].

The index i in Eq. 1.14 is usually taken to be the set of quantum numbers $\{l, m_l\}$ for the atom-like pseudo-orbitals in the core region. The exact form of these wavefunctions can be constructed in numerous ways, such as RRKJ[98], Troullier-Martins[117], HGH[41], or Optimized Norm-Conserving Vanderbilt Pseudopotentials (ONCVPSP)[40]. The latter scheme was devised to improve calculated lattice constants and bulk modulii as well as reduce the plane-wave basis set size required to construct the states[40], so we generally use it in the following chapters.

Pseudopotential generation is in practice a bottleneck for DFT calculations, and the use of pseudopotential databases, such as Pseudo-Dojo[67] is preferable when given the option. However, the need for norm-conserving, fully-relativistic pseudopotentials, often containing “semi-core” states, severely limits options for databases, and pseudopotential generation then proceeds with varyingly sophisticated forms of trial and error. A rigorous approach to determining the various cut-off parameters to generate ONCV pseudopotentials has been developed to construct the “SG15” database[105]. The difficulty in using this method for generating fully-relativistic pseudopotentials, however, is displayed in the results from a similar databse in calculating the DFT band gap of GaAs[104].

Planewaves

Ideal crystals are composed of a repeated unit cell. This discrete symmetry imposes a condition on the wavefunctions in the crystal:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}(\mathbf{r}), \tag{1.15}$$

with $u_{n\mathbf{k}}$ being periodic within a unit cell[58], n its band number, and $\mathbf{k}$ its Bloch wave-vector. In practical calculations, then, it follows that planewaves make a natural choice of basis set for the (“Bloch”) wavefunctions.

The periodic wavefunction u is expanded in this basis set as

$$u_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{1.16}$$

where $\mathbf{G}$ is a reciprocal lattice vector. We require some value of cutoff, which we take to be the number of $\mathbf{G}$-vectors required to suitably describe the pseudopotentials used in the calculation. Since the pseudopotentials are nodeless, this number is usually moderate.

Our task in DFT calculations using plane-waves and pseudopotentials, then, is to determine the plane-wave coefficients $c_{n\mathbf{k}}$ and DFT eigenvalues $\epsilon_{n\mathbf{k}}$.

Spin-Orbit Coupling in DFT

In the presence of spin-orbit coupling, the use of the quantum numbers $\{l, m_l\}$ for the pseudopotential is no longer the natural choice, and we instead use $\{l, j, m_j\}$. This simple substitution is only the beginning of the analysis of incorporating relativistic effects into DFT, however.

Let us first consider a Kohn-Sham equation for a crystalline solid, now incorporating relativistic effects up to order c^{-2}:

$$\begin{aligned} H^{\text{KS+rel.}} = & -\frac{1}{2m_e}\nabla^2 + v^{\text{H}}[n](\mathbf{r}) + v^{\text{ext}}(\mathbf{r}) + v^{\text{xc}}[n](\mathbf{r}) \\ & + v^{\text{Darwin}}(\mathbf{r}) + v^{\text{rel. mass}}(\mathbf{r}) + v^{\text{SOC}}(\mathbf{r}). \end{aligned} \tag{1.17}$$

The Darwin and SOC terms depend on the form of the ionic potentials used, as we are no longer considering interactions of valence states and bare ions, as in Eq. 1.3.

In one approach, we include only the Darwin and mass-correction terms and solve a "scalar relativistic" Kohn-Sham equation. (This is a misnomer, as $\mathbf{L}\cdot\mathbf{S}$ from spin-orbit coupling is also a scalar; nevertheless, this nomenclature persists.) This is adequate for materials containing atoms only from the first three rows (though sometimes we are interested in, for example, spin-orbit in silicon).

We may use the above approach as a first step for a calculation incorporating spin-orbit coupling as a perturbation (see Sec. 1.2). This may be done in a self-consistent procedure, updating the single-particle eigenvalues and eigenfunctions through perturbation theory after completion of each step of a self-consistent cycle ("second-variational method"), or the perturbation may be added at the completion of self-consistency ("non-self-consistent second-variational method")[48]. The second-variational procedure is widely used in DFT software.

The pseudopotential may also incorporate spin-orbit coupling explicitly. In this case, the pseudopotential is parameterized by $\{l, j, m_j\}$, as noted above. This necessitates the Kohn-Sham equation to be solved for two-component spinor wavefunctions:

$$\left(-\frac{1}{2m_e}\nabla^2 + v^{\text{H}}[n](\mathbf{r}) + v^{\text{xc}}[n](\mathbf{r})\right)\delta_{\alpha\alpha}\phi_{n\mathbf{k}\alpha}(\mathbf{r}) + v^{\text{ps}}_{\alpha\beta}\phi_{n\mathbf{k}\beta}(\mathbf{r}) = \epsilon_{n\mathbf{k}}\phi_{n\mathbf{k}\alpha}(\mathbf{r}). \tag{1.18}$$

The indices α and β indicate a component of the two-component spinor. That is, for the spinors $\{\chi_\uparrow, \chi_\downarrow\}$ diagonalizing the Pauli matrix σ_z,

$$\phi_{\text{i},\alpha}(\mathbf{r}) = \phi_{\text{i},\uparrow}(\mathbf{r})\chi_\uparrow + \phi_{\text{i},\downarrow}(\mathbf{r})\chi_\downarrow. \tag{1.19}$$

We are using Einstein summation notation when two quantities with a repeated spin index are multiplied, unless otherwise specified by the use of explicit summation.

We note that we are neglecting spin-dependent interactions that break time-reversal symmetry, such as a Zeeman field, otherwise the exchange correlation functional would itself depend on the spinor indices α and β[78].

Additionally, there are techniques useful in quantum chemistry that incorporate spin-orbit coupling in a different manner. Some quantum chemistry codes solve the

Dirac-Kohn-Sham equation with four-component spinors[103]; this method has been used to calculate (quasiparticle) bandstructures for solids containing Pu and Am[64]. Also, the Dirac Hamiltonian can be expanded to order c^{-2} through an expansion in the term $E/\left(2mc^2 - V\right)$ instead of the usual $(E-V)/2mc^2$; this approach is called Zero-Order Regular Approximation[119], or "ZORA."

1.2 Spin-Orbit Coupling via perturbation theory

Semi-local pseudopotentials

Correcting eigenvalues by including spin-orbit coupling via perturbation theory when using semi-local pseudopotentials was first treated by Hybertsen and Louie[50]. We assume the pseudopotential includes spin-orbit coupling in the $\{l, m_l\}$ basis, such that

$$v_{\mathrm{ps}}(\mathbf{r}) = \sum_{l} |l\rangle \left[v_l^{\mathrm{ion}}(\mathbf{r}) + v_l^{soc}(\mathbf{r})\mathbf{L}\cdot\mathbf{S}\right]\langle l|. \tag{1.20}$$

With a set of Bloch states $\{|n\mathbf{k}\rangle\}$, we affix a spinor $\chi_s = \{\chi_\uparrow, \chi_\downarrow\}$ to create the now-doubled set of states $\{n\mathbf{k}s\rangle\}$. We then must calculate the matrix elements

$$\langle n_1\mathbf{k}s_1|H^{\mathrm{soc}}|n\mathbf{k}s\rangle = \sum_{\mathbf{G},\mathbf{G}'} \phi^*_{n_1\mathbf{k}}(\mathbf{r}) H^{\mathrm{soc}}_{\mathbf{k}+\mathbf{G},\mathbf{k}+\mathbf{G}'}\phi_{n\mathbf{k}}(\mathbf{r}), \tag{1.21}$$

with

$$H^{\mathrm{soc}} = \sum_{l\neq 0} |l\rangle v_l^{\mathrm{soc}}(\mathbf{r})\mathbf{L}\cdot\mathbf{S}\langle l|. \tag{1.22}$$

The matrix elements are

$$H^{\mathrm{soc}}_{\mathbf{K}_1 s_1;\mathbf{K}s} = \langle s_1|\mathbf{S}|s\rangle \cdot \sum_{l\neq 0} \mathbf{M}_l\, v_l^{\mathrm{soc}}(\mathbf{K}_1, \mathbf{K})\, S(\mathbf{K}-\mathbf{K}_1), \tag{1.23}$$

with $\mathbf{K} = \mathbf{k} + \mathbf{G}$, $\mathbf{K}_1 = \mathbf{k} + \mathbf{G}_1$,

$$\mathbf{M}_l = -i4\pi(2l+1)\frac{\mathrm{d}P_l(\hat{\mathbf{K}}_1\cdot\hat{\mathbf{K}})}{\mathrm{d}\hat{\mathbf{K}}_1\cdot\hat{\mathbf{K}}}\hat{\mathbf{K}}_1\times\hat{\mathbf{K}}, \tag{1.24}$$

where P_l are Legendre polynomials, and

$$v_l^{\mathrm{soc}}(\mathbf{K}_1, \mathbf{K}) = \frac{1}{\Omega_{\mathrm{cell}}}\int_0^\infty \mathrm{d}r\; r^2\, j_l\left(K_1 r\right) V_l^{\mathrm{SOC}}(r) j_l\left(K r\right). \tag{1.25}$$

Kleinman-Bylander pseudopotentials

In order to reduce computation time of the necessary integrals as in the preceding equation, as $N_{\mathbf{G}} = 10^3$ to 10^5, we make use of Kleinman-Bylander potentials. Following Hemstreet[45], we have the non-local pseudopotential in the usual Kleinman-Bylander form, but with the pseudo-atom wavefunctions in the $\{ljm_j\rangle\}$ basis:

$$\begin{aligned}
v_{\text{ps}}(\mathbf{r}) &= v_{\text{Loc}} + \sum_{l,s,j,m_j} |v_{lj}\Phi_{lsjm_j}\rangle\langle v_{lj}\Phi_{lsjm_j}|, \\
v_{lj} &= C_{lj}R_{lj}\delta v_{lj}, \\
C_{lj} &= \langle R_{lj}|\delta v_{lj}|R_{lj}\rangle^{-1/2}, \\
\delta v_{lj} &= v_{lj}(\mathbf{r}) - v_{\text{Loc}}(\mathbf{r}),
\end{aligned} \tag{1.26}$$

where R_{lj} is the radial part of the lj pseudo-atom wavefunction, and Φ_{lsjm_j} is the spin-spherical harmonic. We also construct the functions $\phi_{lj} = R_{lj}\Phi_{lsjm_j}$.

Critically, we can remove the dependence on the quantum numbers $\{j, m_j\}$ by constructing

$$\begin{aligned}
v_l^{\text{ion}} &= (2l+1)^{-1}\left[(l+1)v_{lj=l+1/2} + lv_{lj=l-1/2}\right] \\
v_l^{\text{soc}} &= 2(2l+1)^{-1}\left[v_{lj=l+1/2} - v_{lj=l-1/2}\right]
\end{aligned} \tag{1.27}$$

and expanding Φ_{lsjm_j} in terms of spherical harmonics and spinors. We now can write the (fully-separable) pseudopotential much like in the semi-local case:

$$\begin{aligned}
v_{\text{ps}}(\mathbf{r}) &= v^{\text{Loc}}(\mathbf{r}) + \sum_l |l\rangle\left[U^{\text{ion}} + U^{\text{soc}}\mathbf{L}\cdot\mathbf{S} + W^{\text{soc}}\right]\langle l|, \\
U^{\text{ion}} &= |v_l^{\text{ion}}\rangle\langle v_l^{\text{ion}}| \\
U^{\text{soc}} &= |v_l^{\text{soc}}\rangle\langle v_l^{\text{ion}}| + |v_l^{\text{ion}}\rangle\langle v_l^{\text{soc}}| \\
W^{\text{soc}} &= |v_l^{\text{soc}}\rangle\left[\frac{1}{4}\left(l\left(l+1\right)\right) - \frac{1}{2}\mathbf{L}\cdot\mathbf{S}\right]\langle v_l^{\text{soc}}|.
\end{aligned} \tag{1.28}$$

The term W^{soc} is second-order in $\mathbf{L}\cdot\mathbf{S}$, since v_l^{soc} already has a first-order dependence on this term. We can therefore drop this term to remain at the order c^{-2}.

Rewriting the above, we have

$$\begin{aligned}
v_{\text{ps}}(\mathbf{r}) &= v^{\text{ion}} + v^{\text{soc}}, \\
v^{\text{ion}} &= v^{\text{Loc}}(\mathbf{r}) + \sum_l |l\rangle U^{\text{ion}}\langle l|, \\
v^{\text{soc}} &= \sum_{l\neq 0} |l\rangle U^{\text{soc}}\mathbf{L}\cdot\mathbf{S}\langle l|.
\end{aligned} \tag{1.29}$$

We have now successfully recast the problem in an analogous form as Eq. 1.22. The matrix elements for the four blocks in spin-space are given in Ref. [45]. We note here the result of principal interest, that the analogous plane-wave matrix elements lead to the computation of the integral

$$v_l^{\mathrm{soc}}(\mathbf{K}) = \frac{1}{\Omega_{\mathrm{cell}}} \int_0^\infty \mathrm{d}\, r\; r^2 v_l^{\mathrm{soc}}(r) j_l\,(Kr)\,, \tag{1.30}$$

for each $\mathbf{K}$, instead of the pairs $\mathbf{K}, \mathbf{K}_1$ as in Eq. 1.25.

1.3 Single-particle excitations in the presence of spin-orbit coupling

As mentioned, (static) DFT does not lead to a meaningful treatment of excited states. If we wish to calculate the properties of states formed by the addition or removal of electrons to an already-present ground state, we must use the formalism of many-body perturbation theory.

We begin with an interacting N-particle ground state, $|N, 0\rangle$. Single-particle excited states can be obtained by adding an electron thereby promoting the system to some m'th excited state with N+1 particles, $|N+1, m\rangle$. We may also obtain an excited state by removint an electron from (or "adding a hole into") the system, creating some n'th many-body state excited state with N-1 particles, $|N-1, n\rangle$.

The probability amplitude describing placing an electron into the system at position $\mathbf{r}'$ and time t', and removing it at position $\mathbf{r}$ and time t is given by the Green's function

$$G^{>}(\mathbf{r}, t; \mathbf{r}', t') = -i\langle N, 0|\hat{\psi}(\mathbf{r}, t)\hat{\psi}^{\dagger}(\mathbf{r}', t')|N, 0\rangle\theta(t - t'), \tag{1.31}$$

with the fermion field operator operating via second quantization via the relation $\hat{\psi}_m^{\dagger}(\mathbf{r}) = \sum_{\mathrm{i}} \phi_{\mathrm{i}}^{*}(\mathbf{r}) a_{\mathrm{i}}^{\dagger}$, where ϕ_{i} is the single-particle orbital with the label i, and $a_{\mathrm{i}}^{\dagger}$ is the fermionic creation operator that operates directly on the many-body wavefunction when represented in second quantization.

Similarly, we may consider the process of removing an electron at coordinates $\{\mathbf{r}, t\}$ and then replacing it at coordinates $\{\mathbf{r}', t'\}$, which gives

$$G^{<}(\mathbf{r}, t; \mathbf{r}', t') = i\langle N, 0|\hat{\psi}^{\dagger}(\mathbf{r}', t')\hat{\psi}(\mathbf{r}, t)|N, 0\rangle\theta(t' - t). \tag{1.32}$$

The "time-ordered" Green's function is defined by $G = G^{>} + G^{<}$.

The Hedin-Lundqvist equations[42] relate the polarizability P, screened Coulomb interaction W, and electronic self-energy Σ to each other via the single-particle Green's function G and the vertex function Γ.

We seek to define the Hedin-Lundqvist equations in the presence of spin-orbit coupling. This requires some care, as the end result is subtly distinct from that of Ref. [111], and in Ref. [4], the authors allow for non-Coulombic electron-electron interactions, which are unnecessary to capture the relevant physics. (The present results reduce to the equations derived in [4] in the case when only the charge-charge channel is allowed, but starting from a different unperturbed Hamiltonian, as suggested in Ref. [101].)

We label the fermionic field operators with their (generally non-collinear) spin states with Greek letters and condense space-time coordinates with numerals. The time-ordered Green's function is then

$$G_{\alpha\beta}(11') = -i\langle N,0|T\left[\hat{\psi}_\alpha(1)\hat{\psi}^\dagger_\beta(1')\right]|N,0\rangle, \tag{1.33}$$

where the time-ordering function T gives the appropriate sum of both cases $t > t'$ and $t < t'$ through the use of θ functions:

$$\begin{aligned} G_{\alpha\beta}(11') = &- i\langle N,0|\hat{\psi}_\alpha(1)\hat{\psi}^\dagger_\beta(1')|N,0\rangle\theta(t_1 - t_{1'}) \\ &+ i\langle N,0|\hat{\psi}^\dagger_\beta(1')\hat{\psi}_\alpha(1)|N,0\rangle\theta(t_{1'} - t_1). \end{aligned} \tag{1.34}$$

Similarly,the two-particle Green's function is defined as

$$G^{(2)}_{\alpha\beta\gamma\eta}(11',22') = (-i)^2\langle N,0|T\left[\hat{\psi}_\alpha(1)\hat{\psi}_\beta(1')\hat{\psi}^\dagger_\eta(2')\hat{\psi}^\dagger_\gamma(2)\right]|N,0\rangle. \tag{1.35}$$

We may now define the quantities in terms of the G, v (the usual Coulomb interaction), and some external potential V^{ext}:

$$\begin{aligned} \rho(1) &= -iG_{\alpha\alpha}(11^+) \\ v^{\text{H}}(1) &= \int \text{d}2\; v(12)\rho(2) \\ V(1) &= V^{\text{ext}}(1) + v^{\text{H}}(1), \end{aligned} \tag{1.36}$$

the charge density, Hartree potential, and total potential, respectively. Note that we are using the Einstein summation convention over spin indices, unless when explicitly stated.

Furthermore, we define the quantities that relate a response in the charge distribution and interactions with the external or total potential from

$$\begin{aligned}
P^{\text{red.}}(12) &= \frac{\delta\rho(1)}{\delta V^{\text{ext}}(2)}\\
P(12) &= \frac{\delta\rho(1)}{\delta V(2)}\\
\epsilon(12) &= \frac{\delta V^{\text{ext}}(1)}{\delta V(2)}\\
W(12) &= \int \mathrm{d}3\ \epsilon^{-1}(13)v(32)\\
\Gamma_{\alpha\beta}(12;3) &= -\frac{\delta G^{-1}_{\alpha\beta}(12)}{\delta V(3)},
\end{aligned} \tag{1.37}$$

the reducible polarizability, irreducible polarizability, dielectric function, screened Coulomb interaction, and vertex function, respectively.

We can use the above relations immediately to arrive at the relation between the irreducible polarizability P, G, and Γ:

$$\begin{aligned}
P(12) = \frac{\rho(1)}{V(2)} &= -i\frac{\delta G_{\alpha\alpha}(11)}{\delta V(2)}\\
&= i\int \mathrm{d}3\ G_{\alpha\beta}(13)\frac{\delta G^{-1}_{\beta\gamma}(34)}{\delta V(2)}G_{\gamma\alpha}(41)\\
&= -i\int \mathrm{d}3\ G_{\alpha\beta}(13)G_{\gamma\alpha}(41)\Gamma_{\beta\gamma}(34;2).
\end{aligned} \tag{1.38}$$

Also, the inverse dielectric matrix ϵ^{-1} gives us a Dyson-like equation:

$$\begin{aligned}
\epsilon^{-1}(12) &= \frac{\delta V(1)}{\delta V^{\text{ext}}(2)}\\
&= \frac{\delta\left(v^{\text{H}}(1) + V^{\text{ext}}(1)\right)}{\delta V^{\text{ext}}(2)}\\
&= \delta(12) + \int \mathrm{d}3\ v(13)\frac{\delta\rho(3)}{\delta V^{\text{ext}}(2)}\\
&= \delta(12) + \int \mathrm{d}34\ v(13)\frac{\delta\rho(3)}{\delta V(4)}\frac{\delta V(4)}{\delta V^{\text{ext}}(2)}\\
&= \delta(12) + \int \mathrm{d}34\ v(13)P(34)\epsilon^{-1}(42)
\end{aligned} \tag{1.39}$$

This can be used for the Hedin-Lundqvist equation for the screened Coulomb interaction, $W = \epsilon^{-1}v$:

$$W(12) = v(12) + \int \mathrm{d}34\; v(13)P(34)W(42). \tag{1.40}$$

We now consider the derivation of the equation of motion for the Green's function. Since spin-orbit is a local, one-particle interaction arising from the spin moment of electron experiencing an effective magnetic field from the electric field of the (fixed) ions in the lattice, it and the other relativistic terms of the same order are incorporated into the *one-particle* Hamiltonian h_0:

$$\begin{aligned}
&h_0(\mathbf{x}) = h^{\mathrm{scal}}(\mathbf{x}) + h^{\mathrm{soc}}(\mathbf{x}),\\
&h^{\mathrm{scal}}(\mathbf{x}) = -\frac{1}{2}\nabla^2 + \sum_{\vec{\tau}_{\mathrm{i}}} v^{\text{el-ion}}(\mathbf{x}, \vec{\tau}_{\mathrm{i}}) + v^{\mathrm{Darwin}}(\mathbf{x}) + v^{\text{rel. mass}}(\mathbf{x}),\\
&h^{\mathrm{soc}}_{\alpha\beta}(\mathbf{x}) = \mu_I(\mathbf{x})\sigma^I_{\alpha\beta}
\end{aligned} \tag{1.41}$$

using notation consistent with Ref. [4], in which the capital letters I, J, and K refer to Cartesian coordinates, the Greek letters refer to spin components, and $\mu_I(\mathbf{x})$ is the I'th component of the position-dependent term from the spin-orbit coupling arising from the (rigid) ionic lattice.

The many-body Hamiltonian in second quantized form is

$$\begin{aligned}
\hat{\mathrm{H}}_0 = &\int \mathrm{d}\mathbf{x}_1\; \hat{\psi}^{\dagger}_{\alpha}(\mathbf{x}_1)h^{\mathrm{scal}}(\mathbf{x}_1)\hat{\psi}_{\alpha}(\mathbf{x}_1) + \int \mathrm{d}\mathbf{x}_1\hat{\psi}^{\dagger}_{\alpha}(\mathbf{x}_1)\mu_I(\mathbf{x}_1)\sigma^I_{\alpha\beta}\hat{\psi}_{\beta}(\mathbf{x}_1)\\
&+ \frac{1}{2}\int \mathrm{d}\mathbf{x}_1\; d\mathbf{x}_2\hat{\psi}^{\dagger}_{\alpha}(\mathbf{x}_1)\hat{\psi}^{\dagger}_{\beta}(\mathbf{x}_2)v(\mathbf{x}_1 - \mathbf{x}_2)\hat{\psi}_{\beta}(\mathbf{x}_2)\hat{\psi}_{\alpha}(\mathbf{x}_1).
\end{aligned} \tag{1.42}$$

The field operator $\hat{\psi}_{\alpha}(\mathbf{x}_1)$ acts at the spatial coordinate $\mathbf{r}_1$ and the time coordinate t_1, and $h_0(1)$ and $v(12)$ are understood to mean $h_0(\mathbf{x}_1)$ and $v(\mathbf{r}_1 - \mathbf{r}_2)\delta(t_1 - t_2)$, respectively. (Similarly in all instances where numbers are used as arguments.)

The equation of motion for the field operator is

$$i\frac{\partial}{\partial t_1}\hat{\psi}_{\gamma}(1) = h^{\mathrm{scal}}(1)\hat{\psi}_{\gamma}(1) + \mu_I(1)\sigma^I_{\gamma\alpha}\hat{\psi}_{\alpha}(1) + \int \mathrm{d}2\; \hat{\psi}^{\dagger}_{\alpha}(2)v(12)\hat{\psi}_{\alpha}(2)\hat{\psi}_{\gamma}(1), \tag{1.43}$$

from which it is straightforward but tedious to construct the equation of motion for the Green's function:

$$\left(i\frac{\partial}{\partial t_1} - h^{\mathrm{scal}}(1)\right) G_{\alpha\beta}(11') - \mu_I(1)\sigma^I_{\alpha\gamma}G_{\gamma\beta}(11') \\ + i\int \mathrm{d}2\; v(12)G^{(2)}_{\alpha\gamma\beta\gamma}(12, 1'2) = \delta(11')\delta_{\alpha\beta}. \tag{1.44}$$

We replace the term involving $G^{(2)}$ in the above from the result derived in the Appendix (Eq. A.8):

$$\left(i\frac{\partial}{\partial t_1} - h^{\mathrm{scal}}(1)\right)G_{\alpha\beta}(11') - \mu_I(1)\sigma^I_{\alpha\gamma}G_{\gamma\beta}(11') - v^{\mathrm{H}}(1)G_{\alpha\beta}(11') \\ - i\int \mathrm{d}345\; G_{\alpha\eta}(13)W(15)\Gamma_{\eta\kappa}(34; 5)G_{\kappa\beta}(41') = \delta(11')\delta_{\alpha\beta}. \tag{1.45}$$

We see that we have, as expected, gone beyond the Hartree approximation with the inclusion of the so-called mass term of the electron-electron self-energy,

$$\bar{\Sigma}_{\alpha\beta}(11') = i\int \mathrm{d}234\; G_{\alpha\eta}(12)W(14)\Gamma_{\eta\kappa}(23; 4)G_{\kappa\beta}(31'), \tag{1.46}$$

with the full self-energy

$$\Sigma_{\alpha\beta}(12) = v^H(1)\delta(12)\delta_{\alpha\beta} + \bar{\Sigma}_{\alpha\beta}(12). \tag{1.47}$$

The equation of motion for G now finally can be written as

$$\left(i\frac{\partial}{\partial t_1} - h^{\mathrm{scal}}(1)\right)G_{\alpha\beta}(11') - \mu_I(1)\sigma^I_{\alpha\gamma}G_{\gamma\beta}(11') \\ - \int \mathrm{d}2\; \Sigma_{\alpha\kappa}(12)G_{\kappa\beta}(21') = \delta(11')\delta_{\alpha\beta}, \tag{1.48}$$

and its inverse G^{-1},

$$G^{-1}_{\alpha\beta}(11') = \left[\left(i\frac{\partial}{\partial t_1} - h^{\mathrm{scal}}(1)\right)\delta_{\alpha\beta} - \mu_I(1)\sigma^I_{\alpha\beta}\right]\delta(11') - \Sigma_{\alpha\beta}(11'). \tag{1.49}$$

Neglecting electron-electron interactions, we have the equation of motion for the non-interacting Green's function G^0 and its inverse $G^{(0)-1}$

$$\begin{aligned}
&\left(i\frac{\partial}{\partial t_1} - h^{\mathrm{scal}}(1)\right) G^0_{\alpha\beta}(11') - \mu_I(1)\sigma^I_{\alpha\gamma} G^0_{\gamma\beta}(11') = \delta(11')\delta_{\alpha\beta}, \\
&G^{(0)-1}_{\alpha\beta}(11') = \left[\left(i\frac{\partial}{\partial t_1} - h^{\mathrm{scal}}(1)\right)\delta_{\alpha\beta} - \mu_I(1)\sigma^I_{\alpha\beta}\right]\delta(11').
\end{aligned} \tag{1.50}$$

By comparing Eq. 1.49 with Eq. 1.50, we see

$$G^{-1}_{\alpha\beta}(11') = G^{(0)-1}_{\alpha\beta}(11') + \Sigma_{\alpha\beta}(11'). \tag{1.51}$$

We use Eq. 1.49 to arrive at the Dyson-like equation for the vertex function Γ:

$$\begin{aligned}
\Gamma_{\alpha\beta}(11';2) &= -\frac{\delta G^{-1}_{\alpha\beta}(11')}{\delta V(2)} \\
&= \delta(11')\delta(12)\delta_{\alpha\beta} + \frac{\delta\bar{\Sigma}_{\alpha\beta}(11')}{\delta V(2)} \\
&= \delta(11')\delta(12)\delta_{\alpha\beta} + \int \mathrm{d}34\ \frac{\delta\bar{\Sigma}_{\alpha\beta}(11')}{\delta G_{\gamma\eta}(34)}\frac{\delta G_{\gamma\eta}(34)}{\delta V(2)} \\
&= \delta(11')\delta(12)\delta_{\alpha\beta} - \int \mathrm{d}34\ \frac{\delta\bar{\Sigma}_{\alpha\beta}(11')}{\delta G_{\gamma\eta}(34)} G_{\gamma\kappa}(35)\frac{\delta G^{-1}_{\kappa\zeta}(56)}{\delta V(2)} G_{\zeta\eta}(64) \\
&= \delta(11')\delta(12)\delta_{\alpha\beta} + \int \mathrm{d}3456\ \frac{\delta\bar{\Sigma}_{\alpha\beta}(11')}{\delta G_{\gamma\eta}(34)} G_{\gamma\kappa}(35) G_{\zeta\eta}(64)\Gamma_{\gamma\zeta}(56;2).
\end{aligned} \tag{1.52}$$

To summarize, we have the set of Hedin equations in the presence of spin-orbit coupling when electrons interact with each other solely through the (spin-independent) Coulomb interaction:

$$\begin{aligned}
&P(12) = -i\int \mathrm{d}34\ G_{\alpha\beta}(13)G_{\gamma\alpha}(41)\Gamma_{\beta\gamma}(34;2) \\
&W(12) = v(12) + \int \mathrm{d}34\ v(13)P(34)W(42) \\
&\bar{\Sigma}_{\alpha\beta}(12) = i\int \mathrm{d}34\ G_{\alpha\eta}(13)W(14)\Gamma_{\eta\beta}(32;4) \\
&\Gamma_{\alpha\beta}(11';2) = \delta(11')\delta(12)\delta_{\alpha\beta} + \int \mathrm{d}3456\ \frac{\delta\bar{\Sigma}_{\alpha\beta}(11')}{\delta G_{\gamma\eta}(34)} G_{\gamma\kappa}(35)G_{\zeta\eta}(64)\Gamma_{\gamma\zeta}(56;2).
\end{aligned} \tag{1.53}$$

The Dyson's equation, from inverting Eq. 1.51, is

$$G_{\alpha\beta}(12) = G^{(0)}_{\alpha\beta}(12) + \int \mathrm{d}34\ G^{(0)}_{\alpha\gamma}(13)\Sigma_{\gamma\eta}(34)G_{\eta\beta}(42). \tag{1.54}$$

To lowest order, we ignore vertex corrections, taking $\Gamma_{\alpha\beta}(11';2) = \delta(11')\delta(12)\delta_{\alpha\beta}$, and arriving at the *GW* Approximation, where

$$\begin{aligned} P^{\mathrm{RPA}}(12) &= -iG_{\alpha\beta}(12)G_{\beta\alpha}(21) \\ W(12) &= v(12) + \int \mathrm{d}34\; v(13)P^{\mathrm{RPA}}(34)W(42) \\ \bar{\Sigma}^{\mathrm{GWA}}_{\alpha\beta}(12) &= iG_{\alpha\beta}(12)W(12). \end{aligned} \tag{1.55}$$

The polarizability is said to be computed within the "RPA" for historical reasons. This amounts to disregarding interactions in the polarization bubble. Corrections to this treatment require the treatment of two-particle excitations (Sec. 1.4). The poles of G give

$$E^{\mathrm{QP}} = \epsilon^0 + \Sigma^{\mathrm{GWA}}(\omega = E^{\mathrm{QP}}), \tag{1.56}$$

where ϵ^0 are the energies of the system without any interactions.

Dyson's equation can be represented through Feynman diagrams, with irreducible contributions from the Hartree term v^{H} and the mass term $\bar{\Sigma}$. It is common to renormalize the $G^{(0)}$ by v^{H} and then drop the bar over the mass term $\bar{\Sigma}$. The difference in the poles of G and $G^{(0)}$, then, gives the contribution to the quasiparticle energy from electron-electron interactions beyond the level of the simple Hartree term. Thought of in this manner, the self-energy naturally contains the physics of both the exchange and correlation effects, within the RPA and GW approximations.

1.4 Electron-hole excitations in the presence of spin-orbit coupling

We now wish to apply a similar analysis to two-particle excited states where particle number is conserved and the pair of one-particle excitations are created simultaneously. We are, then, concerned with solving for transitions from $|N,0\rangle$ to $|N,S\rangle$. We restrict our analysis to the case when such an excitation is the result of a process in which the system absorbs a photon. In this case, the spin of the (quasiparticle) electron is the inverse of that of the hole, simplifying the analysis of this process when including spin-orbit coupling. Indeed, we may arrive at the correct solution by noting that apart from this constraint, the process will be identical to the case when neglecting spin, treated as in, e.g., Ref. [111], just with traces over spin in the matrix elements defined in Ref. [99]. We proceed, however, to carry out the full analysis for confirmation.

We find that the full two-particle Green's function $G^{(2)}$ is too general. The analogous two-particle propagator where both particles operate simultaneously is given by L. The relation between L and $G^{(2)}$ is

$$L_{\alpha\beta\gamma\zeta}(12',21') \equiv G_{\alpha\gamma}(12)G_{\beta\zeta}(2'1') - G^{(2)}_{\alpha\beta\gamma\zeta}(12',21'). \tag{1.57}$$

We can introduce the non-interacting propagator as

$$L^{0}_{\alpha\beta\gamma\zeta}(12',21') = G_{\alpha\gamma}(12)G_{\beta\zeta}(2'1'), \tag{1.58}$$

which simply describes a pair of non-interacting particles.

To arrive at a Dyson-like equation for L, we eliminate $G^{(2)}$ using the variational derivative technique of Schwinger, introducing a non-local electrostatic perturbing potential ϕ. In the Appendix, we arrive at the result (Eq. A.5)

$$\frac{\delta G_{\alpha\beta}(11')}{\delta\phi(23)} = -G^{(2)}_{\alpha\gamma\beta\gamma}(13,1'2^{+}) + G_{\alpha\beta}(11')G_{\gamma\gamma}(32^{+}). \tag{1.59}$$

We note that the repeated spinor index ("γ" in the above) is due to the spin-independence of the non-local electrostatic perturbing potential. That is, the electron-hole pair is created in the singlet state, though they may scatter into some arbitrary spin configuration due to the loss of spin conservation in the presence of spin-orbit coupling.

We now can arrive at a self-consistent equation for L:

$$\begin{aligned}
L_{\alpha\beta\gamma\beta}(12',21') &= \frac{\delta G_{\alpha\gamma}(12)}{\delta\phi(1'2')} \\
&= -\int \mathrm{d}34\; G_{\alpha\zeta}(13)\frac{\delta G^{-1}_{\zeta\eta}(34)}{\delta\phi(1'2')}G_{\eta\gamma}(42) \\
&= G_{\alpha\beta}(12')G_{\beta\gamma}(1'2) \\
&\quad + \int \mathrm{d}34\; G_{\alpha\zeta}(13)\frac{\delta\Sigma_{\zeta\eta}(34)}{\delta\phi(1'2')}G_{\eta\gamma}(42) \\
&= L^{(0)}_{\alpha\beta\gamma\beta}(12',21') \\
&\quad + \int \mathrm{d}3456\; G_{\alpha\zeta}(13)\frac{\delta\Sigma_{\zeta\eta}(34)}{\delta G_{\mu\nu}(65)}\frac{\delta G_{\mu\nu}(65)}{\delta\phi(1'2')}G_{\eta\gamma}(42) \\
&= L^{(0)}_{\alpha\beta\gamma\beta}(12',21') \\
&\quad + \int \mathrm{d}3456\; G_{\alpha\zeta}(13)\frac{\delta\Sigma_{\zeta\eta}(34)}{\delta G_{\mu\nu}(65)}L_{\mu\beta\nu\beta}(62',51')G_{\eta\gamma}(42).
\end{aligned} \tag{1.60}$$

We define the Bethe-Salpeter Kernel as

$$\begin{aligned}
K_{\zeta\nu\eta\mu}(35,46) = \frac{\delta\Sigma_{\zeta\eta}(34)}{\delta G_{\mu\nu}(65)} &= \frac{\delta\left(v(37)\rho(7)\delta(34)\delta_{\zeta\eta} + iG_{\zeta\eta}(34)W(3^{+}4)\right)}{\delta G_{\mu\nu}(65)} \\
&= -i\delta(34)\delta_{\zeta\eta}v(37)\frac{\delta G_{\rho\rho}(77^{+})}{\delta G_{\mu\nu}(65)} \\
&\quad + iW(3^{+}4)\frac{\delta G_{\zeta\eta}(34)}{\delta G_{\mu\nu}(65)} \\
&= -i\delta(34)\delta_{\zeta\eta}v(37)\delta(67)\delta(57^{+})\delta_{\mu\rho}\delta_{\nu\rho} \\
&\quad + iW(3^{+}4)\delta(36)\delta(45)\delta_{\zeta\mu}\delta_{\eta\nu} \\
&= -iv(36)\delta(34)\delta(56^{+})\delta_{\zeta\eta}\delta_{\mu\nu} \\
&\quad + iW(3^{+}4)\delta(36)\delta(45)\delta_{\zeta\mu}\delta_{\eta\nu}. \qquad (1.61)
\end{aligned}$$

In the above, we make the approximation that

$$\frac{\delta G_{\zeta\eta}(34)W(3^{+}4)}{\delta G_{\mu\nu}(65)} \approx W(3^{+}4)\frac{\delta G_{\zeta\eta}(34)}{\delta G_{\mu\nu}(65)}, \qquad (1.62)$$

which simplifies the Kernel and is seen to be adequate in practice[99]. We use this to rewrite Eq. 1.60 as

$$\begin{aligned}
&L_{\alpha\beta\gamma\beta}(12',21') - L^{(0)}_{\alpha\beta\gamma\beta}(12',21') \\
&= -i\int \mathrm{d}3456\ L^{(0)}_{\alpha\eta\zeta\gamma}(14,32)v(36)\delta(34)\delta(56^{+})\delta_{\zeta\eta}\delta_{\mu\nu}L_{\mu\beta\nu\beta}(62',51') \\
&\quad + i\int \mathrm{d}3456\ L^{(0)}_{\alpha\eta\zeta\gamma}(14,32)W(3^{+}4)\delta(36)\delta(45)\delta_{\kappa\mu}\delta_{\eta\nu}L_{\mu\beta\nu\beta}(62',51') \\
&= -i\int \mathrm{d}35\ L^{(0)}_{\alpha\zeta\zeta\gamma}(13,32)v(35)L_{\mu\beta\mu\beta}(52',51') \\
&\quad + i\int \mathrm{d}35\ L^{(0)}_{\alpha\eta\zeta\gamma}(15,32)W(3^{+}5)L_{\zeta\beta\eta\beta}(32',51'). \qquad (1.63)
\end{aligned}$$

This Dyson-like equation for the propagator is more readily interpreted through the corresponding diagrams (as in Ref. [111]), where we have the expected exchange and direct terms, $K^{\mathrm{X}} = -iv$ and $K^{\mathrm{D}} = iW$, respectively. Regardless of the details of the ordering of the coordinates, we note that the interactions are the bare and screened Coulomb interaction, which is spin-independent.

The inverted Bethe-Salpeter Equation is, from inspection,

$$L^{-1}_{\alpha\beta\gamma\beta}(12',21') = L^{(0)-1}_{\alpha\beta\gamma\beta}(12',21') - K_{\alpha\beta\gamma\beta}(12',21'). \qquad (1.64)$$

We wish to solve for the poles in the inverted BSE, which requires us to decide on a basis for our excited states. In the orbital basis, the two-particle excited-state wavefunction for state S is

$$\Psi^S_{\alpha\beta}(11') = \sum_{vc\mathbf{k}} A^S_{vc\mathbf{k}} \psi_{c\mathbf{k}\alpha}(1)\psi^*_{v\mathbf{k}\beta}(1') + B^S_{vc\mathbf{k}} \psi_{v\mathbf{k}\alpha}(1)\psi^*_{c\mathbf{k}\beta}(1'). \tag{1.65}$$

In the Tamm-Dancoff approximation (TDA), we take $B^S = 0$ and simplify the wavefunction:

$$\Psi^{S,(\mathrm{TDA})}_{\alpha\beta}(11') = \sum_{vc\mathbf{k}} A^S_{vc\mathbf{k}} \psi_{c\mathbf{k}\alpha}(1)\psi^*_{v\mathbf{k}\beta}(1'). \tag{1.66}$$

The following uses exclusively the TDA to simplify the construction of the results; however the process may be generalized to include the off-diagonal blocks of the BSE allowed by non-zero B^S.

In the orbital basis, then, we have [99],

$$\begin{aligned} L^{(0)-1}_{\alpha\beta\gamma\beta}(12';21';\omega) &= \sum_{vc\mathbf{k}} M^{\alpha\beta}_{cv\mathbf{k}}(11')\left(\omega - E_{cv\mathbf{k}}\right) M^{\beta\gamma\,*}_{cv\mathbf{k}}(2'2) - M^{\gamma\beta}_{vc\mathbf{k}}(22')\left(\omega + E_{cv\mathbf{k}}\right) M^{\beta\alpha\,*}_{vc\mathbf{k}}(1'1) \\ L^{-1}_{\alpha\beta\gamma\beta}(12',21';\omega) &= \sum_{S} \left[\Psi^S_{\alpha\beta}(11')\left(\omega - \Omega^S\right) \Psi^{S}_{\beta\gamma}{}^*(2'2)) - \Psi^S_{\gamma\beta}(22')\left(\omega + \Omega^S\right) \Psi^{S}_{\beta\alpha}{}^*(1'1)\right], \end{aligned} \tag{1.67}$$

with $M^{\alpha\beta}_{vc\mathbf{k}}(11') = \psi_{c\mathbf{k}\alpha}(1)\psi^*_{v\mathbf{k}\beta}(1')$ and $E_{cv\mathbf{k}} = E_{c\mathbf{k}} - E_{v\mathbf{k}}$.

We solve the inverted BSE, enforcing the solution $\omega = \Omega_S$ and using orthonormality of the excited-state wavefunctions, to arrive at the eigenvalue equation for the excited state S, with energy Ω_S and envelope function $A^S_{vc\mathbf{k}}$:

$$\left(E_{c\mathbf{k}} - E_{v\mathbf{k}}\right) A^S_{vc\mathbf{k}} + \sum_{v'c'\mathbf{k}'} \left(K^{\mathrm{X}}_{vc\mathbf{k},v'c'\mathbf{k}'} + K^{\mathrm{D}}_{vc\mathbf{k},v'c'\mathbf{k}'}\right) A^S_{v'c'\mathbf{k}'} = \Omega^S A^S_{vc\mathbf{k}}.$$

This is exactly the form of the eigenvalue equation from the BSE when neglecting spin-orbit coupling, with the exception that we have twice as many valence and conduction states. Further, the direct product of states $|v\mathbf{k}\rangle \bigotimes |c\mathbf{k}\rangle$ no longer has an irreducible representation as the sum of eigenstates of S^2, as it no longer commutes with the many-body Hamiltonian.

The matrix elements of K^X and K^D are

$$\begin{aligned} K^{\mathrm{X}}_{vc\mathbf{k},v'c'\mathbf{k}'} &= -i \int \mathrm{d}12\; \psi^*_{c'\mathbf{k}'\eta}(1)\psi_{v'\mathbf{k}'\eta}(1) v(12) \psi_{c\mathbf{k}\mu}(2)\psi^*_{v\mathbf{k}\mu}(2), \\ K^{\mathrm{D}}_{vc\mathbf{k},v'c'\mathbf{k}'} &= i \int \mathrm{d}12\; \psi^*_{c'\mathbf{k}'\mu}(1)\psi_{c\mathbf{k}\mu}(1) W(34) \psi_{v'\mathbf{k}'\nu}(2)\psi^*_{v\mathbf{k}\nu}(2). \end{aligned} \tag{1.68}$$

The spin-dependence of the Kernel, and thus the BSE, reduces as expected to a trace over the spin coordinates in the computation of the relevant generalized charge density matrix elements.

Chapter 2

The Implementation of Spinor Wavefunction Functionality in BerkeleyGW

2.1 Introduction

Solid state physics and materials research is increasingly focusing its attention on materials containing heavy elements. Such materials have large spin-orbit coupling, exceeding 1 eV for atoms from the fifth and sixth rows of the periodic table. The standard approach to investigate the ground state electronic structure of these materials is Density Functional Theory (DFT)[60, 46]. Despite its widespread use to compute bandstructures, it is important to recall that the Kohn-Sham eigenvalues of DFT do not have a rigorous physical meaning (apart from the highest occupied molecular orbital), resulting in the well-known band gap problem of DFT. To compute excited-state properties such as bandstructures and absorption spectra, we must go beyond DFT and use many-body perturbation theory approaches, such as the GW[42, 49] and GW-BSE methods[99].

For materials with weak spin-orbit coupling, quasiparticle bandstructures incorporating spin-orbit coupling can be computed from the change in eigenvalues through conventional perturbation theory in which the spin-orbit Hamiltonian is constructed using orbitals neglecting spin[50]. This approach, “GW+SOC,” has been successfully used in ab initio calculations of diamond- and zinc-blende-structure semiconductors[73], and many other systems.

Despite the success of perturbation theory in computing the changes of eigenvalues for materials with weak spin-orbit coupling, there is a clear need for a nonperturbative first-principles treatment of materials with strong spin-orbit coupling. For materials with weak spin-orbit coupling, the Kohn-Sham bandstructure obtained neglecting spin-orbit coupling is qualitatively similar to the quasiparticle bandstructure including spin-orbit coupling, apart from a smaller band gap, absence of spin-orbit splittings, and some differences in bandwidths. However, some materials containing heavy elements, such as Bi_2Se_3[134] and β-HgS[8], have DFT bandstructures that differ significantly when including or neglecting spin-orbit coupling. In cases such as these, we use a fully-relativistic treatment from the outset, starting with the calculation of two-component spinor Kohn-Sham states and then using them to calculate excited-state properties. This first-principles method also allows for capturing the effect of the renormalization of the spin-orbit coupling strength[101], along with improved band gaps.

2.2 Spinors and the BerkeleyGW software package

In the frequency domain, the one-partice Green's function is

$$G_{\alpha\beta}\left(\mathbf{r}_1, \mathbf{r}_2; \omega\right) = \sum_{n\mathbf{k}} \frac{\phi_{n\mathbf{k}\alpha}(\mathbf{r}_1)\phi^*_{n\mathbf{k}\beta}(\mathbf{r}_2)}{\omega - E_{n\mathbf{k}} \pm i\eta}, \tag{2.1}$$

where $\phi_{n\mathbf{k}\alpha}$ is a quasiparticle wavefunction for band n at k-point $\mathbf{k}$ with z-axis spin projection α and energy $E_{n\mathbf{k}}$.

In practice, we find that the Kohn-Sham wavefunctions $\phi^{\mathrm{KS}}_{n\mathbf{k}\alpha}$ are usually good approximations to the actual quasiparticle wavefunctions[49], so we approximate the Green's function as

$$G^{\mathrm{KS}}_{\alpha\beta}\left(\mathbf{r}_1, \mathbf{r}_2; \omega\right) = \sum_{n\mathbf{k}} \frac{\phi^{\mathrm{KS}}_{n\mathbf{k}\alpha}(\mathbf{r}_1)\phi^{\mathrm{KS}*}{}_{n\mathbf{k}\beta}(\mathbf{r}_2)}{\omega - \epsilon_{n\mathbf{k}} \pm i\eta} \tag{2.2}$$

We evaluate the quasiparticle energies with a one-shot "G_0W_0" procedure:

$$E_{n\mathbf{k}} = \epsilon_{n\mathbf{k}} + \sum_{\alpha\beta} \langle n\mathbf{k}\alpha| \left(\Sigma_{\alpha\beta}\left(E_{n\mathbf{k}}\right) - V^{\mathrm{xc}}\delta_{\alpha\beta}\right) |n\mathbf{k}\beta\rangle. \tag{2.3}$$

We note that V^{xc} has only spin-diagonal components in the absence of time-reversal symmetry-breaking fields. The Kohn-Sham eigenfunctions are written in a plane-wave basis as

$$\phi^{\mathrm{KS}}_{n\mathbf{k}\alpha}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n\mathbf{k}\alpha} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}. \tag{2.4}$$

Since the non-magnetic polarizability has its physical origins from density fluctuations arising from the spin-independent Coulomb interaction, the form of the polarizability is identical to the case in which spin-orbit is neglected, apart from the traces in the computation of the matrix elements, the doubled number of both valence and conduction bands, and any differences in eigenfunctions and eigenvalues. In many cases, the latter can be taken to be small, in which case one may calculate the polarizability using the Kohn-Sham eigenfunctions and eigenvalues from a non-relativistic DFT calculation[101, 2]. However, in this work, we use the Kohn-Sham eigenfunctions and eigenvalues from fully-relativistic DFT calculations.

The polarizability matrix for real frequencies may be constructed from the Kohn-Sham eigenfunctions and eigenvalues as[1, 130, 110]

$$\begin{aligned} P_{\mathbf{GG}'}(\mathbf{q},\omega) =& \frac{1}{N_{\mathbf{k}}} \sum_{n}^{\mathrm{occ}} \sum_{n'}^{\mathrm{emp}} \sum_{\mathbf{k}} M^{*}_{nn'}(\mathbf{k},\mathbf{q},\mathbf{G}) M_{nn'}(\mathbf{k},\mathbf{q},\mathbf{G}') \\ & \times \frac{1}{2}\left[\frac{1}{\epsilon_{n\mathbf{k}+\mathbf{q}} - \epsilon_{n'\mathbf{k}} - \omega \mp i\eta} + \frac{1}{\epsilon_{n\mathbf{k}+\mathbf{q}} - \epsilon_{n'\mathbf{k}} + \omega \pm i\eta}\right]. \end{aligned} \tag{2.5}$$

with η an infinitesimal positive number.

The matrix elements are

$$M_{nn'}(\mathbf{k},\mathbf{q},\mathbf{G}) = \sum_{\alpha} \langle n\mathbf{k}+\mathbf{q}\alpha | e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}} | n'\mathbf{k}\alpha\rangle, \tag{2.6}$$

where the spin index α is explicitly traced over. We may compute these matrix elements for all $\mathbf{G}$ by multiplying the Fourier transforms of the wavefunctions, for a spin component common to both wavefunctions, computing the inverse Fourier transform of this product[26, 31], and then tracing over spin index:

$$M_{nn'}(\mathbf{k},\mathbf{q},\{\mathbf{G}\}) = \sum_{\alpha\alpha'} \mathrm{FFT}^{-1}\left(\phi^{*}_{n\mathbf{k}+\mathbf{q}\alpha}(\mathbf{r})\phi_{n'\mathbf{k}\alpha'}(\mathbf{r})\right)\delta_{\alpha\alpha'}. \tag{2.7}$$

We note that self-energy operator $\Sigma = iGW$ inherits the spin-dependence from the Green's function[4], but the process of taking matrix elements reduces this spin-dependence to taking traces over the spinor components of the wavefunction. This is readily seen by considering the matrix elements Σ, which separate into two terms, with the screened-exchange ("SX") coming from the poles of G, and the coulomb-hole ("COH") coming from the poles of W[49]. The matrix elements of Σ^{SX}:

$$\begin{aligned} \langle n\mathbf{k}\alpha | \Sigma^{\mathrm{SX}}_{\alpha\beta}(\omega) | n\mathbf{k}\beta\rangle = & -\sum_{n''}^{\mathrm{occ}} \sum_{\mathbf{qGG}'} M^{*}_{n''n}(\mathbf{k},-\mathbf{q},-\mathbf{G}) M_{n''n}(\mathbf{k},-\mathbf{q},-\mathbf{G}') \\ & \times \epsilon^{-1}_{\mathbf{GG}'}(\mathbf{q};\omega - E_{n''\mathbf{k}-\mathbf{q}}) v(\mathbf{q}+\mathbf{G}'). \end{aligned} \tag{2.8}$$

The Coulomb-hole term ("COH") requires similar traces in the computation of the matrix elements and is otherwise unaffected by spin-orbit coupling:

$$\begin{aligned}\langle n\mathbf{k}\alpha|\Sigma^{\mathrm{COH}}_{\alpha\beta}(\omega)|n\mathbf{k}\beta\rangle = -\sum_{n''}\sum_{\mathbf{qGG'}} & M^{*}_{n''n}(\mathbf{k},-\mathbf{q},-\mathbf{G})M_{n''n}(\mathbf{k},-\mathbf{q},-\mathbf{G'}) \\ & \times \frac{1}{\pi}\int \mathrm{d}\omega' \frac{\mathrm{Im}\,\epsilon^{-1}_{\mathbf{G},\mathbf{G'}}(\mathbf{q},\omega')}{\omega-\epsilon_{n''\mathbf{k}}-\omega'+i\eta}. \end{aligned} \tag{2.9}$$

The spin dependence of $\Sigma_{\alpha\beta}(\omega)$ merits further consideration. We may recast the self-energy operator as sum of products of operators, $\Sigma_{\alpha\beta}(\omega) = A(\omega)\times \mathrm{I}_{\alpha\beta} + B(\omega)\times\sigma^{x}_{\alpha\beta} + C(\omega)\times\sigma^{y}_{\alpha\beta} + D(\omega)\times\sigma^{z}_{\alpha\beta}$, for some appropriate choice of operators A, B, C, and D. We use this decomposition simply to illustrate that due to the inclusion of spin-orbit coupling entirely as a one-electron interaction, the electron-electron self-energy generally allows for spin scattering even when mediated solely by the Coulomb interaction. This is in contrast to the argument proposed in Ref. [4], which uses spin-dependent interactions that conserve spin. (The existence of further symmetries of course may introduce relevant selection rules that forbid certain spin scatterings.) This is a consequence of spin no longer being a good quantum number when spin-orbit coupling is included into the Hamiltonian that propagates excited states. This physics is absent in V^{xc}, which is in a sense the DFT approximation to the self-energy.

To simplify the calculation of matrix elements of Σ^{COH}, a generalized plasmon pole model (GPP)[136, 49] treats the integrand in the integral over ω' as sufficiently sharply peaked that they may be replaced by delta functions. In this case, only the static dielectric function needs to be explicitly computed. The Hybertsen-Louie GPP[49] is justified through the use of a sum rule that is derived from a double-commutator of a one-particle non-interacting Hamiltonian (as in the RPA) with charge density operators, and is therefore independent of spin-orbit coupling.

To calculate the optical absorption spectrum of a material, we may first try and evaluate the imaginary part of the macroscopic dielectric function within the independent-particle approximation. Since the dipole operator is spin-independent, we may readily determine, using the usual expression[132, 22], the dielectric function to be

$$\epsilon_2(\omega) = \frac{8\pi^2e^2}{\omega^2}\sum_{v c\mathbf{k}}\left|\boldsymbol{\lambda}\cdot\sum_{\alpha}\langle v\mathbf{k}\alpha|\mathbf{v}|c\mathbf{k}\alpha\rangle\right|^2 \delta\left(\omega-\left(E_{c\mathbf{k}}-E_{v\mathbf{k}}\right)\right), \tag{2.10}$$

where $\boldsymbol{\lambda}$ is the direction of light polarization and $\sum_{\alpha}\langle v\mathbf{k}\alpha|\mathbf{v}|c\mathbf{k}\alpha\rangle$ is the matrix element of the velocity operator. Usually the numerator of the prefactor is "$16\pi^2e^2$," but that is with the spin degree of freedom summed over. There is only a single spin degree of freedom due to the spin-independence of the velocity operator giving non-zero matrix elements only when the electron and hole have opposite spins.

More accurate calculations of the absorption spectrum require the inclusion of excitonic effects, which we calculate by solving for the excitonic eigenfunctions and eigenvalues of the Bethe-Salpeter Equation (BSE).

With the excitonic envelope functions $A^S_{vc\mathbf{k}}$ and energies Ω^S, we compute the imaginary part of the macroscopic dielectric function:

$$\epsilon_2(\omega) = \frac{8\pi^2 e^2}{\omega^2} \sum_S \left| \sum_{vc\mathbf{k}\alpha} A^S_{vc\mathbf{k}} \boldsymbol{\lambda} \cdot \langle v\mathbf{k}\alpha|\mathbf{v}|c\mathbf{k}\alpha\rangle \right|^2 \delta\left(\omega - \Omega^S\right). \tag{2.11}$$

The most concerning challenge is the increase in the number of matrix elements we must calculate. Compared to a comparable calculation without spin, the number of both valence and conduction states double. Taking the ratio of the scaling of the matrix element calculation with system size[26], we find an increase in computation time by

$$\frac{2^2 N^2 2\log(2N)}{N^2 \log N} = 8(1 + \log_N 2), \tag{2.12}$$

where the additional factor of 2 in the numerator comes from having to compute the inverse-FFT of two-component spinors. Since, at best, we are increasing the cost of matrix element calculations by roughly a factor of 8, we must make use of symmetries of the Brillouin Zone to allow for converged calculations within reasonable computational cost. Further, as we have assumed time-reversal symmetry, we may consider how that may be exploited to ease computation time.

2.3 Spinor wavefunctions and symmetries

Symmetries, No spin

We first briefly review the use of symmetries for wavefunctions without spin, as this is also necessary when using spinor wavefunctions. We use the symmetries of the Brillouin Zone to store only the necessary wavefunction coefficients within the irreducible wedge, a fraction of the full Brillouin Zone, provided that more symmetries exist than the identity. When computing the polarizability and self-energy, we then unfold to “little group of the q-vector” except when computing off-diagonal matrix elements of the self-energy[26].

A (three-dimensional) crystal structure belongs to one of 230 space groups, each allowing some transformation $\{\mathrm{R}|\boldsymbol{\tau}\}$, a combination of rotation followed by translation by a fraction of a lattice vector, that leaves the structure invariant. Transformations for which $\boldsymbol{\tau}$ are not zero are called non-symmorphic symmetries. Diamond and zinc-blende structures contain non-symmorphic symmetries. The ionic potential $V_{ion}(\mathbf{r})$, thus the Hamiltonian, inherits the symmetry of the lattice, so we may consider the action of these transformations on the Bloch wavefunctions $\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}$.

We consider the action of a rotation R followed by a fractional translation $\boldsymbol{\tau}$. A coordinate vector $\mathbf{r}$ is transformed as $\mathbf{r}' = \mathrm{R}\mathbf{r} + \boldsymbol{\tau}$. A function is transformed according to the convention

$$\begin{aligned} &P_{\{\mathrm{R}|\boldsymbol{\tau}\}}f\left(\{\mathrm{R}|\boldsymbol{\tau}\}\mathbf{r}\right) = P_{\{\mathrm{R}|\boldsymbol{\tau}\}}f\left(\mathrm{R}\mathbf{r} + \boldsymbol{\tau}\right) = f(\mathbf{r}), \\ &P_{\{\mathrm{R}|\boldsymbol{\tau}\}}f\left(\mathbf{r}\right) = f\left(\mathrm{R}^{-1}\mathbf{r} - \mathrm{R}^{-1}\boldsymbol{\tau}\right). \end{aligned} \tag{2.13}$$

Thus

$$\begin{aligned} P_{\{\mathrm{R}|\boldsymbol{\tau}\}}\psi_{n\mathbf{k}}(\mathbf{r}) &= u_{n\mathbf{k}}(\mathrm{R}^{-1}\mathbf{r} - \mathrm{R}^{-1}\boldsymbol{\tau})e^{i\mathbf{k}\cdot\left(\mathrm{R}^{-1}\mathbf{r} - \mathrm{R}^{-1}\boldsymbol{\tau}\right)} \\ &= u_{n\mathbf{k}}(\mathrm{R}^{-1}\mathbf{r} - \mathrm{R}^{-1}\boldsymbol{\tau})e^{i\mathrm{R}\mathbf{k}\cdot(\mathbf{r}-\boldsymbol{\tau})}. \end{aligned} \tag{2.14}$$

where we use the property that scalars formed in dot products are invariant under rotation of both vectors. We drop the phase factor $e^{-i\mathrm{R}\mathbf{k}\cdot\boldsymbol{\tau}}$, as this is common among all bands at a given k-point.

We expand the periodic part of the Bloch function in plane-waves:

$$\begin{aligned} u_{n\mathbf{k}}\left(\mathrm{R}^{-1}\mathbf{r} - \mathrm{R}^{-1}\boldsymbol{\tau}\right) &= \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G})e^{i\mathbf{G}\cdot\left(\mathrm{R}^{-1}\mathbf{r} - \mathrm{R}^{-1}\boldsymbol{\tau}\right)} \\ &= \sum_{\mathbf{G}} c_{n\mathbf{k}}(G)e^{i\mathrm{R}\mathbf{G}\cdot(r-\boldsymbol{\tau})} \\ &= \sum_{\mathrm{R}^{-1}\mathbf{G}} c_{n\mathbf{k}}(\mathrm{R}^{-1}\mathbf{G})e^{-i\mathbf{G}\cdot\boldsymbol{\tau}}e^{i\mathbf{G}\cdot\mathbf{r}}. \end{aligned} \tag{2.15}$$

Substituting in to the previous equation and reordering the summation over G-vectors, we have

$$\begin{aligned} &P_{\{\mathrm{R}|\boldsymbol{\tau}\}}\psi_{n\mathbf{k}}(\mathbf{r}) = \tilde{u}_{n\mathrm{R}\mathbf{k}}(\mathbf{r})e^{i\mathrm{R}\mathbf{k}\cdot\mathbf{r}}, \\ &\tilde{u}_{n\mathrm{R}\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} \tilde{c}_{n\mathrm{R}\mathbf{k}}(\mathrm{R}^{-1}\mathbf{G})e^{-i\mathbf{G}\cdot\boldsymbol{\tau}}e^{i\mathbf{G}\cdot\mathbf{r}}. \end{aligned} \tag{2.16}$$

(The use of the $\tilde{c}$ refers to the change in the function u when evaluated at $\mathbf{r}$ instead of $\mathrm{R}^{-1}\mathbf{r} - \mathrm{R}^{-1}\boldsymbol{\tau}$.)

This allows us to use the usual relation when using symmetries to unfold the Brillouin Zone from an irreducible wedge:

$$c_{n\mathbf{k}}(\mathbf{G}) \to \tilde{c}_{n\mathrm{R}\mathbf{k}}(\mathrm{R}^{-1}\mathbf{G})e^{-i\mathbf{G}\boldsymbol{\tau}}. \tag{2.17}$$

Symmetries, with spin

We now extend the above discussion for the case of spinor wavefunctions, where

$$\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathbf{k}\uparrow}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}\chi_\uparrow + u_{n\mathbf{k}\downarrow}(\mathbf{r})e^{i\mathbf{k}\cdot\mathbf{r}}\chi_\downarrow. \tag{2.18}$$

The periodic functions $u_{n\mathbf{k}\uparrow}$ and $u_{n\mathbf{k}\downarrow}$ are spatial and thus transform according to the above. However, the spinor itself rotates according to the rules of transformation for elements of the group SU(2):

$$P_{\{\mathrm{R}|\boldsymbol{\tau}\}}\psi_{n\mathbf{k}}(\mathbf{r}) = u_{n\mathrm{R}\mathbf{k}\uparrow}(\mathbf{r})e^{i\mathrm{R}\mathbf{k}\cdot\mathbf{r}}e^{-i\hat{n}_\mathrm{R}\cdot\vec{\sigma}\theta_\mathrm{R}/2}\chi_\uparrow + u_{n\mathrm{R}\mathbf{k}\downarrow}(\mathbf{r})e^{i\mathrm{R}\mathbf{k}\cdot\mathbf{r}}e^{-i\hat{n}_\mathrm{R}\cdot\vec{\sigma}\theta_\mathrm{R}/2}\chi_\downarrow, \tag{2.19}$$

where $\hat{n}$ and θ are the unit-axis and angle (about the axis $\hat{n}$) that recreates the action of the symmetry operation R. We readily arrive at the rule for transforming two-component spinor Bloch functions:

$$P_{\{\mathrm{R}|\boldsymbol{\tau}\}}\psi_{n\mathbf{k}}(\mathbf{r}) = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) - in_z\sin\left(\frac{\theta}{2}\right) & (-n_y - in_x)\sin\left(\frac{\theta}{2}\right) \\ (-n_y + in_x)\sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) + in_z\sin\left(\frac{\theta}{2}\right) \end{pmatrix} \begin{pmatrix} u_{n\mathrm{R}\mathbf{k}\uparrow}(\mathbf{r})e^{i\mathrm{R}\mathbf{k}\cdot\mathbf{r}} \\ u_{n\mathrm{R}\mathbf{k}\downarrow}(\mathbf{r})e^{i\mathrm{R}\mathbf{k}\cdot\mathbf{r}} \end{pmatrix}, \tag{2.20}$$

where n_i is the i'th Cartesian component of $\hat{n}$.

The task, then, is to determine $\hat{n}$ and θ for each symmetry R used in unfolding the Brillouin Zone.

2.4 Axis and Angle Extraction

The Euler axis/angle form of the matrix in the Cartesian basis that rotates a vector (in R^3) by some angle θ about a unit vector $\hat{n}$ can be recreated by consideration of the generators of the Lie group for SO(3). We simply restate the end result[77]:

$$\mathrm{R}\left(\hat{n},\theta\right) = \begin{pmatrix} c + n_\mathrm{x}^2\left(1-c\right) & -n_\mathrm{z}s + n_\mathrm{x}n_\mathrm{y}\left(1-c\right) & n_\mathrm{y}s + n_\mathrm{x}n_\mathrm{z}\left(1-c\right) \\ n_\mathrm{z}s + n_\mathrm{y}n_\mathrm{x}\left(1-c\right) & c + n_\mathrm{y}^2\left(1-c\right) & -n_\mathrm{x}s + n_\mathrm{y}n_\mathrm{z}\left(1-c\right) \\ -n_\mathrm{y}s + n_\mathrm{z}n_\mathrm{x}\left(1-c\right) & n_\mathrm{x}s + n_\mathrm{z}n_\mathrm{y}\left(1-c\right) & c + n_\mathrm{z}^2\left(1-c\right) \end{pmatrix} \tag{2.21}$$

with $c = \cos\theta$ and $s = \sin\theta$.

A first attempt to extract the values of n_{x}, n_{y}, n_{z}, and θ for any given rotation matrix might make use of the following:

$$\begin{aligned} \cos\theta &= \frac{1}{2}\left(Tr\left(\mathrm{R}\right) - 1\right), \\ n_{\mathrm{k}} &= \epsilon_{\mathrm{ijk}}\left(-\left[\mathrm{R}\right]_{\mathrm{ij}} + \left[\mathrm{R}\right]_{\mathrm{ji}}\right)/2\sin\theta. \end{aligned} \tag{2.22}$$

However, the division by $\sin\theta$ is singular for rotations of $\theta = 0 \pm \eta$ or $\theta = \pi \pm \eta$, where η is the machine precision for floating-point numbers. Further, if $Tr\left(\mathrm{R}\right) = 1 + \eta$, the use of arccos to find θ is also subject to failure, given sufficient round-off error.

More general extraction algorithms attempt to remove singularities by making use of a reparameterization of the above rotation matrix in terms of a four-component quaternion. A unit quaternion q, with transpose $q^T = (q_1, q_2, q_3, q_4)$, can be parameterized by $\hat{n}$ and θ from

$$q^T = (n_1 \sin\frac{\theta}{2}, n_2 \sin\frac{\theta}{2}, n_3 \sin\frac{\theta}{2}, cos\frac{\theta}{2}). \tag{2.23}$$

We now rewrite the rotation matrix above as[77]

$$\mathrm{R} = \begin{pmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2\left(q_1 q_2 + q_3 q_4\right) & 2\left(q_1 q_3 - q_2 q_4\right) \\ 2\left(q_2 q_1 - q_3 q_4\right) & -q_1^2 + q^2 - q_3^2 + q_4^2 & 2\left(q_2 q_3 + q_1 q_4\right) \\ 2\left(q_3 q_1 + q_2 q_4\right) & 2\left(q_3 q_2 - q_1 q_4\right) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{pmatrix}. \tag{2.24}$$

We extract the parameters $\hat{n}$ and θ by using Markley's[76] modification to Shepperd's algorithm[124]. We first construct an auxiliary matrix $\mathbf{X}$:

$$\mathbf{X} = \begin{pmatrix} 1 + 2\mathrm{R}_{11} - Tr\left(\mathrm{R}\right) & \mathrm{R}_{21} + \mathrm{R}_{12} & \mathrm{R}_{31} + \mathrm{R}_{12} & \mathrm{R}_{23} - \mathrm{R}_{32} \\ \mathrm{R}_{12} + \mathrm{R}_{21} & 1 + 2\mathrm{R}_{22} - Tr\left(\mathrm{R}\right) & \mathrm{R}_{32} + \mathrm{R}_{23} & \mathrm{R}_{31} - \mathrm{R}_{13} \\ \mathrm{R}_{13} + \mathrm{R}_{31} & \mathrm{R}_{23} + \mathrm{R}_{32} & 1 + 2\mathrm{R}_{33} - Tr\left(\mathrm{R}\right) & \mathrm{R}_{12} - \mathrm{R}_{21} \\ \mathrm{R}_{23} - \mathrm{R}_{32} & \mathrm{R}_{31} - \mathrm{R}_{13} & \mathrm{R}_{12} - \mathrm{R}_{21} & 1 + Tr\left(\mathrm{R}\right) \end{pmatrix}. \tag{2.25}$$

Then we compute the norms of each column $\mathbf{x}^{\mathrm{i}}$. We use the column with the largest norm to compute the quaternion q, from

$$q = \pm\mathbf{x}^{\mathrm{i}} / \left|\mathbf{x}^{\mathrm{i}}\right|, \tag{2.26}$$

by construction[76]. The positive and negative branches of solution come from the double-cover of SO(3) by SU(2), the latter of which is parameterized by the four real quaternion components q^{i} instead of the usual two complex components for spin. We arbitrarily choose to use the positive branch, which is adequate for materials science applications.

Markley demonstrated that this algorithm produces errors that are bounded to the order of round-off error[76]. We note that methods of further optimizing quaternion extraction algorithms appears to be an active area of research; for instance, Ref. [15]

The set of rotation matrices R for a crystalline system are usually stored in the basis of lattice vectors in *ab initio* codes, as it allows these matrices (up to 48) to be written with nine integers. In this case, we must transform the rotation matrices in the lattice basis, $\mathrm{R}^{\mathrm{lat}}$, to the rotation matrix in the Cartesian basis. If we form a matrix A out of the three lattice vectors $\mathbf{a}_1$, $\mathbf{a}_2$, and $\mathbf{a}_3$ as

$$A = \begin{pmatrix} a_{1x} & a_{2x} & a_{3x} \\ a_{1y} & a_{2y} & a_{3y} \\ a_{1z} & a_{2z} & a_{3z} \end{pmatrix}, \tag{2.27}$$

this transformation is

$$\mathrm{R}^{\mathrm{cart}} = A\mathrm{R}^{\mathrm{lat}}A^{-1}. \tag{2.28}$$

If instead we decide to use the reciprocal lattice vectors $\mathbf{b}_1$, $\mathbf{b}_2$, $\mathbf{b}_3$ to construct the matrix B in a fashion as in the above, we make use of the identity $B^{\mathrm{T}}A = 2\pi I$ to write

$$\mathrm{R}^{\mathrm{cart}} = \left(B^{\mathrm{T}}\right)^{-1}\mathrm{R}^{\mathrm{lat}}B^{\mathrm{T}}. \tag{2.29}$$

This latter choice is beneficial if the matrices A and B are in fact stored as their transposes, as some codes do.

Finally, we note that in the presence of inversion symmetries, each rotation R can be followed by the inversion N to form an “improper rotation” $S = N\mathrm{R}$. However, if both spatial inversion and time-reversal commute with the Hamiltonian under consideration, the (spinor) wavefunction is a simultaneous eigenstate of both symmetries. Thus in the presence of only time-reversal symmetric terms in the Hamiltonian, the wavefunctions are unaffected by inversion, apart from perhaps an overall phase factor. We identify improper rotations by the identity $\det(S) = -1$, and if detected, use only the rotation part R of S to transform the spinor components of the wavefunction. In the absence of time-reversal symmetry, however, we must use the appropriate magnetic point group (indeed, for all symmetries), but such consideration is beyond the scope of this work.

Time Reversal and Inversion symmetries

We may try to use time reversal symmetry to compress spinor wavefunctions, as can be done in the case of single-component wavefunctions in the presence of inversion symmetry[26]. With spin, the Bloch function $\phi_{n\mathbf{k}}(\mathbf{r})$ for a given band is generally a non-collinear spinor

$$\phi_{n\mathbf{k}}(\mathbf{r}) = \phi_{n\mathbf{k}\uparrow}(\mathbf{r})\chi_{\uparrow} + \phi_{n\mathbf{k}\downarrow}(\mathbf{r})\chi_{\downarrow}. \tag{2.30}$$

In the presence of inversion and time-reversal symmetries, we again have the eigenvalue equation

$$\mathrm{PT}\phi_{n\mathbf{k}} = \lambda\phi_{n\mathbf{k}}, \tag{2.31}$$

but now the time-reversal symmetry operation is $\mathrm{T} = -i\sigma_y K$, with K being the complex conjugate operation. Operating explicitly on the left-hand side, we have

$$\mathrm{PT}\phi_{n\mathbf{k}\uparrow}(\mathbf{r})\chi_{\uparrow} + \mathrm{PT}\phi_{n\mathbf{k}\downarrow}(\mathbf{r})\chi_{\downarrow} = -\phi^*_{n\mathbf{k}\downarrow}(-\mathbf{r})\chi_{\uparrow} + \phi^*_{n\mathbf{k}\uparrow}(-\mathbf{r})\chi_{\downarrow}. \tag{2.32}$$

The term $\phi^*_{n\mathbf{k}\uparrow}$ in the planewave basis is

$$\begin{aligned}\phi^*_{n\mathbf{k}\uparrow}(-\mathbf{r}) &= \sum_{\mathbf{G}} c^*_{n\mathbf{k}\uparrow}(\mathbf{G})e^{i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}} \\ &= \sum_{\mathbf{G}} \tilde{c}_{n\mathbf{k}\uparrow}(\mathbf{G})e^{i(\mathbf{G}+\mathbf{k})\cdot\mathbf{r}} \\ &= \phi_{n\mathbf{k}\uparrow}(\mathbf{r}),\end{aligned} \tag{2.33}$$

within the phase factor relating the coefficients $\tilde{c}$ to c. A similar result holds for the "down" component.

We are not able to reduce the amount of information needed to describe the wavefunctions, since we just have the result that the spinor components switch places, apart from a sign difference. However, this result may be used to reduce the calculation of matrix elements in the following manner (Ref. [58]) :

$$\langle n\mathbf{k}\downarrow|e^{i\mathbf{G}\cdot\mathbf{r}}|m\mathbf{k}\downarrow\rangle = \langle n\mathbf{k}\uparrow|e^{-i\mathbf{G}\cdot\mathbf{r}}|m\mathbf{k}\uparrow\rangle. \tag{2.34}$$

In this case,

$$\mathrm{Tr}_{\alpha}\left(\langle n\mathbf{k}\alpha|e^{i\mathbf{G}\cdot\mathbf{r}}|m\mathbf{k}\alpha\rangle\right) = 2\left(\langle n\mathbf{k}\uparrow|\cos\left(\mathbf{G}\cdot\mathbf{r}\right)|n\mathbf{k}\uparrow\rangle\right). \tag{2.35}$$

While this reduction in the number of terms reduces the prefactor in the additional number of matrix element calculations the inclusion of spin necessitates, this property is not used in BerkeleyGW at present.

2.5 Benchmarking with simple test systems

Introduction

We present results for seven different materials with a wide range of spin-orbit coupling (SOC) strengths. The diamond and zincblende semiconductors Si, Ge, and

	E_{cut} (Ry)	a_0^{relaxed}(Å)	$a_0^{\text{exp.}}$(Å)
Si	120	5.48	5.47
Ge	120	5.63	5.66
GaAs	350	5.61	5.65
GaSb	350	6.09	6.10
CdSe	200	4.30	4.30
Au	72	4.08	4.08

Table 2.1: The kinetic energy cutoffs, calculated lattice parameters, and experimental lattice parameters. The pseudopotentials for Ge, Sb, Cd, and Au contain the full shell of the semicore states (e.g., $4s^24p^64d^{10}$ for Sb)[116].

The experimental lattice parameters are from Ref. [72].

GaAs are technologically important materials with weak SOC. GaSb has a strong SOC, and the spin-orbit splitting is of similar magnitude to the band gap. CdSe has a wurtzite structure and a significant SOC (approximately 30 times larger than wurtzite GaN, for example). We also study Au, being a prototypical metal with strong SOC.

Computational Details

We compute mean-field wavefunctions and eigenvalues from Density Functional Theory[60, 46]. For the exchange-correlation energy, we employ the Perdew-Zunger parameterization of the LDA[90]. We generate fully-relativistic pseudopotentials using the Optimized Norm-Conserving Vanderbilt Pseudopotential (ONCVPSP) scheme[40] with parameters from the Pseudo-Dojo pseudopotential database[67]. The pseudopotentials for Au, Bi, Cd, Ga, Ge, and Sb contain the full shell of the semicore states (e.g., $5s^25p^65d^{10}$ for Bi) for accurate calculation of the bare exchange[116]. All DFT calculations are carried out with the Quantum ESPRESSO software package[35].

We first determine the equilibrium lattice constants and atom positions. Table 2.1 shows that all relaxed lattice constants are in very good agreement with experimental measurements.

Next, the quasiparticle energies are computed with the one-shot "G_0W_0" approach, using the Hybertsen-Louie Generalized Plasmon Pole model[49, 136] for the inverse dielectric matrix. For the case of bulk Au, we also calculated the quasiparticle band structure in the Godby-Needs Plasmon Pole Model[36] and found differences of 50 meV or smaller in the energy, in the range 6 eV above and below the Fermi energy. Table 2.2 summarizes our parameters for the empty state summations, the

	$\mathbf{k}$-grid	ϵ_{cut} (Ry)	Empty States
Si	8x8x8	20	800
Ge	8x8x8	25	600 Chi, 1000 COH
GaAs	8x8x8	20	1002
GaSb	8x8x8	20	1002
CdSe	6x6x4	20	996
Au	8x8x8	50	2018

Table 2.2: The values of the Brillouin Zone sampling, the screened Coulomb cutoff, and the number of empty states used in the sums for both the polarizability ("Chi") and the Coulomb-hole ("COH") term in the self-energy. For Si, Ge, GaAs, and GaSb, we use the same parameters as from Ref. [73], and for Au, Ref. [86].

	$\mathbf{k}_{\text{fine}}$ grid	N_v	N_c	Broadening (meV)
GaAs	12x12x12	6	8	150
GaSb	12x12x12	6	8	100

Table 2.3: The values of the Brillouin Zone sampling of the fine grid, the number of valence and conduction bands used as the basis for the BSE, and the Gaussian broadening of the delta function.

k-point sampling, and the plane-wave cutoff for the dielectric matrix. We use the static remainder method to improve convergence with the number of empty states in the Coulomb-hole summation[27]. We verified that performing one-shot or "G_0W_0" evaluation of the self energy in the band-diagonal approximation yields quantitatively accurate bandstructures for these test systems.

The k-point sampling and number of bands used in constructing the GW-BSE Hamiltonian are summarized in Table 2.3. All excited-state calculations are carried out with the BerkeleyGW program adapted for use with spinor wavefunctions.

Si

Fig. 2.1 shows the bandstructure of Si calculated within fully-relativistic LDA ("FR-LDA")and fully-relativistic GW ("FR-GW"). We find the FR-LDA band gap to be 0.445 eV, with the valence band maximum at the Γ-point and the conduction band minimum along the $\Gamma - X$ line. The FR-GW gap has the value 1.22 eV, with the

	FR-LDA	FR-GW	Experiment
E_g (eV)	0.445	1.22	1.22^a - 1.23^b
$E(\Gamma_{7c})$-$E(\Gamma_{8v})$ (eV)	2.46	3.22	3.34^c
$\Delta^{\mathrm{SOC}}(\Gamma, v)$ (eV)	0.047	0.049	0.044^d
$\Delta^{\mathrm{SOC}}(\Gamma, c)$ (eV)	0.034	0.035	0.030 - 0.040^c
$\Delta^{\mathrm{SOC}}(L, v)$ (eV)	0.031	0.032	0.030^c
$\Delta^{\mathrm{SOC}}(L, c)$ (eV)	0.014	0.014	–

Table 2.4: The band gap and spin-orbit splitting for Si, computed at the FR-LDA and FR-GW levels, compared to experiment. The fundamental band gap from experiment is reported with Zero-Point Renormalization corrections. ^{a} Ref. [14], ^{b} Ref. [13], ^{c} Ref. [80], ^{d} Ref. [72].

valence band maximum and conduction band minimum occuring at the same points as in FR-LDA. The measured band gap is 1.17 eV at low temperature; correcting for the zero-point electron-phonon renormzlization ("ZPR") yields a gap of 1.22[14] to 1.23 eV[13], in excellent agreement with our FR-GW result. Table 2.4 shows that the calculated spin-orbit splittings from FR-LDA and FR-GW have excellent agreement with experiment.

In the absence of spin-orbit coupling, the band gaps increase slightly to 0.46 eV in the LDA and 1.23 eV within (scalar-relativistic, or "SR") GW. The SR-GW direct gap at Γ is also slightly larger at 3.26 eV, compared to 3.22 eV within FR-GW. The FR-GW and SR-GW quasiparticle bandstructures in Fig. 2.1 are nearly identical, due to the SOC strength in Si.

Ge

Fig. 2.2 shows the bandstructure of Ge from FR-LDA and FR-GW. Previous DFT calculations often found an incorrect negative band gap for Ge[6, 116, 62, 63]. With the use of the fully-relativistic ONCVPSP pseudopotential with $3s3p3d$ semicore states, however, we find a small but positive direct gap at the Γ-point of 0.15 eV (Table 2.5).

The FR-LDA (indirect) band gap is computed to be 0.13 eV, with the top of the valence band at Γ and the bottom of the conduction band at L. Self-energy corrections at the FR-GW level increase this gap to 0.743 eV. The experimental gap is 0.744 eV[72], increasing to 0.79 eV[13] when ZPR is taken into account, in good agreement with the FR-GW result. Table 2.5 shows the calculated spin-orbit

	FR-LDA	FR-GW	Experiment[a]
E_g (eV)	0.129	0.743	0.79
$E(\Gamma_{6c})$ - $E(\Gamma_{8v})$ (eV)	0.154	0.960	0.90
$\Delta^{\text{SOC}}(\Gamma, v)$ (eV)	0.307	0.302	0.297
$\Delta^{\text{SOC}}(\Gamma, c)$ (eV)	0.221	0.208	0.200
$\Delta^{\text{SOC}}(L, v)$ (eV)	0.194	0.192	0.228
$\Delta^{\text{SOC}}(L, c)$ (eV)	0.095	0.082	0.09[b]

Table 2.5: The band gap and spin-orbit splittings for Ge, computed at the FR-LDA and FR-GW levels,compared to experiment. [a] Ref. [72], [b] Ref. [20]

splittings. The splittings in FR-GW are smaller than in FR-LDA by few meV and are in good agreement with experimental data.

The band gap in the SR-GW approach increases to 0.842 eV. The direct gap at Γ in FR-GW is calculated to be 0.960 eV, smaller than the SR-GW result of 1.05 eV. The increases in these gaps compared to FR-GW are due to the moderately strong SOC in Ge. The difference in the direct band gaps between SR-GW and FR-GW is well-approximated by the result from usual perturbation theory for the atomic spin-orbit splitting, in which a $p_{3/2}$ state shifts upward in energy by $\Delta^{\text{SOC}}/3$ and $p_{1/2}$ downward by $2\Delta^{\text{SOC}}/3$. We need only the SR-GW gap E_g^{SR} and the FR-LDA spin-orbit splitting $\Delta^{\text{SOC}}(\Gamma, v)$ for the top of the valence band for the estimation

$$E_g^{\text{FR}} = E_g^{\text{SR}} - \frac{1}{3}\Delta^{\text{SOC}}(\Gamma, v). \tag{2.36}$$

This estimation yields 0.95 eV, which agrees with the calculated value within 10 meV.

GaAs

Fig. 2.3 shows the bandstructure for GaAs calculated with FR-LDA and FR-GW. We find an FR-LDA band gap of 0.55 eV and a FR-GW gap of 1.49 eV, compared to the 0 K gap of 1.52 eV, from experiment[72]. (Correcting for ZPR, the experimental band gap increases to 1.57 eV[65]). Table 2.6 shows the calculated spin-orbit splittings. We find that the FR-GW splittings are in excellent agreement with experiment.

We also calculate the bandstructure from SR-LDA and SR-GW. We find that SR-LDA band gap is slightly larger than that of FR-LDA, at 0.652 eV. The SR-GW band gap was found to be 1.599 eV. The atomic SOC perturbation theory estimation

	FR-LDA	FR-GW	Experiment
E_g (eV)	0.55	1.49	1.57^a
$\Delta^{\text{SOC}}(\Gamma, v)$ (eV)	0.320	0.340	0.340
$\Delta^{\text{SOC}}(\Gamma, c)$ (eV)	0.189	0.171	0.171
$\Delta^{\text{SOC}}(L, v)$ (eV)	0.198	0.209	0.22
$\Delta^{\text{SOC}}(L, c)$ (eV)	0.08	0.07	0.05

Table 2.6: The band gap and spin-orbit splitting for GaAs, computed at the FR-LDA and FR-GW levels, compared to experiment. Experimental data is from Ref. [72] unless otherwise specified. ^{a} Ref. [65].

of the FR-GW (direct) band gap, 1.49 eV, agrees with the calculated value within less than 10 meV.

Fig. 2.4 shows the absorption spectrum of GaAs from FR-GW-BSE and SR-GW-BSE. Both methods yield similar spectra, apart from a small shift of 40 meV in the absorption onset and a difference in amplitude of the absorption peak at 3 eV. The shift in the absorption onset is due to the smaller direct band gap within FR-GW as compared to the SR-GW result. The 3 eV peak has been measured to be split into a pair of peaks due to SOC, but the splitting (200 meV) is on the order of the resolution of the calculation (150 meV) and is thus obscured.

GaSb

Figure 2.5 shows the FR-LDA, FR-GW, SR-LDA, and SR-GW bandstructures. We compute an FR-LDA band gap of 0.135 eV and a FR-GW direct gap of 0.819 eV, compared to the low temperature gap of 0.822 eV from experiment[72]. Table 2.7 shows the calculated spin-orbit splittings, which are in good agreement with experiment. However, at the FR-GW, as well as SR-GW, level, we find the conduction band minimum to be located at the L point instead of the Γ point, despite experimental evidence of a direct band gap at the Γ-point in GaSb[72].

In the SR-LDA, the direct band gap at Γ is larger than FR-LDA, at 0.358 eV, and in the SR-GW, 1.065 eV. As in FR-GW, the fundamental band gap is predicted to be indirect in SR-GW, from Γ to L, with a value of 1.024 eV. The over 200 meV disagreement of the band gap between SR-GW and experiment indicates the necessity of treating spin-orbit coupling for GaSb. The direct gap is again well-approximated by applying the perturbation theory estimate, 0.818 eV.

Figure 2.6 shows the absorption spectrum for GaSb calculated with the SR-GW-BSE and FR-GW-BSE methods. The absorption spectrum of GaSb has significant

	FR-LDA	FR-GW	Experiment[1]
$E(\Gamma_{6c})$-$E(\Gamma_{8v})$ (eV)	0.135	0.819	0.822
$E(L_{6c})$-$E(\Gamma_{8v})$ (eV)	0.245	0.779	0.907
$\Delta^{\mathrm{SOC}}(\Gamma, v)$ (eV)	0.740	0.733	0.756
$\Delta^{\mathrm{SOC}}(\Gamma, c)$ (eV)	0.225	0.203	0.213
$\Delta^{\mathrm{SOC}}(L, v)$ (eV)	0.424	0.422	0.430
$\Delta^{\mathrm{SOC}}(L, c)$ (eV)	0.12	0.09	0.13

Table 2.7: The band gap and spin-orbit splittings for GaSb, computed at the FR-LDA and FR-GW levels, compared to experiment. Experimental data is from Ref. [72]

differences when including SOC. First, the absorption onset is shifted by 190 meV due to the large difference in the quasiparticle band gap when including (0.822 eV) or neglecting SOC (1.065 eV). Also, we can clearly resolve the 2.3 eV peak splitting into the so-called E_1 and $E_1 + \Delta$ peaks with the inclusion of SOC. When split, the pair of peaks have a reduced peak height. (See section 2.5.) The E_1 and $E_1 + \Delta$ peak placements at 2.18 eV and 2.54 eV agree well with the experimental[139] spectrum placements of 2.18 eV and 2.62 eV, respectively.

CdSe

Figure 2.7 shows the bandstructure of CdSe computed at the FR-LDA, FR-GW, SR-LDA, and SR-GW levels. We compute the FR-LDA band gap of 0.58 eV and the FR-GW gap of 1.85 eV, compared to the low temperature gap of 1.84 eV from experiment[44]. Table 2.8 shows the spin-orbit and crystal field ("CF") splittings, with FR-GW showing excellent agreement with experiment and much improved compared to FR-DFT.

The SR-GW gap is larger due to the neglect of the large spin-orbit splitting, with a value of 1.986 eV. The change in the band gap due to the inclusion of spin-orbit coupling is well-approximated by perturbation theory, though the spin-orbit splitting differs by over 30 meV whether using the value computed in LDA or GW (Table 2.8). The band gap is estimated within a few tens of meV when using the FR-LDA value of the spin-orbit splitting, giving 1.862 eV. The bandstructure (Fig. 2.7) shows little change when including or neglecting SOC, apart from the moderate spin-orbit splitting at the top of the valence band.

	FR-LDA	FR-GW	Experiment
E_g (eV)	0.584	1.845	1.84
$\Delta^{\text{SOC}}(\Gamma, v)$ (eV)	0.372	0.405	0.429
$\Delta^{\text{CF}}(\Gamma, v)$ (eV)	0.036	0.026	0.026

Table 2.8: The band gap and spin-orbit splitting for CdSe, computed at the FR-LDA and FR-GW levels, compared to experiment. The spin-orbit (SOC) and crystal field (CF) splitting refers to the states at the top of the valence band at Γ. Experimental data is from Ref. [44]

Au

Figure 2.8 shows the bandstructure of Au computed from FR-LDA, FR-GW, SR-LDA, and SR-GW. The inclusion of spin-orbit coupling changes the degeneracy of the occupied *s*-*d* hybridized states. The bandstructures show slight changes upon inclusion of the electronic self-energy. Notably, the valence states shift downward in energy and the conduction states shift upward. Table 2.9 shows that the interband gaps are generally improved with FR-GW, particularly near the Fermi level. The quasiparticle energies are largely similar whether using the Hybertsen-Louie or the Godby-Needs GPP model, within an energy range of 6 eV above or below the Fermi level.

2.6 Performance

We give a comparison in the performance of BerkeleyGW for the representative case of GaAs with and without spinor wavefunctions. We see in Table 2.10 that the calculation of the wavefunctions for the self-energy matrix elements (“DFT Coarse”) takes four times longer, in accordance with expectation from having to double the number of bands and double the size of each band, for the spin-up and spin-down components. The calculation of the wavefunctions for the basis of the BSE Hamiltonian (“DFT Fine”) is more rapid, since the bottleneck in generating these wavefunctions is the number of k-points and not the number of bands. Calculation of the dielectric matrix (“Epsilon”) sees an increase in cost of only 2.5, far less than the increase in cost of generating the matrix elements alone. The size of the dielectric matrix is the same when including or disregarding spin, so the matrix inversion step allows for an overall retention of relative performance. The calculation of quasiparticle energies (“Sigma”), however, is closer to the expected increase in cost, at a factor of 4.1. The costs of constructing the BSE Kernel (“Kernel”) and solving the eigenvectors and

	FR-LDA	FR-GW	Experiment
Γ_{6+}	-10.17	-10.22	–
Γ_{8+}	-5.69	-6.05	-5.09[a], -6[b], -6.01[c]
Γ_{7+}	-4.58	-4.89	-4.45[a], -4.6[b], -4.68[c]
Γ_{8+}	-3.29	-3.67	-3.55[a], -3.65[b], -3.71[c]
Γ_{7-}	13.91	14.46	16[c], 15.9[d]
Γ_{6-}	17.26	17.81	18.8[c]
L_{6+}	-7.84	-8.11	-7.8[b]
$L_{4,5+}$	-5.80	-6.21	-6.23[b], -6.2[c]
L_{6+}	-4.69	-5.08	-4.88[b], -5[c]
L_{6+}	-2.56	-2.87	-3.2[c]
$L_{4,5+}$	-1.90	-2.19	-2.3[c], -2.5[e]
L_{6-}	-1.32	-1.26	-1[e], -1[f], -1.01[g], -1.1[h]
L_{6+}	3.09	3.44	3.6[e], 3.65[f], 3.56[g], 3.4[h]

Table 2.9: The FR-LDA and FR-GW eigenvalues for Au, as compared to experiment. [a] Ref. [82], [b] Ref. [5], [c] Ref. [23], [d] Ref. [51], [e] Ref. [115], [f] Ref. [52], [g] Ref. [21], [h] Ref. [74].

Step	No. CPUs	CPU Hours (no spin)	CPU Hours (spinor)	Ratio
DFT Coarse	1024	162	650	4.0
DFT Fine	1728	173	490	2.8
Epsilon	864	864	2160	2.5
Sigma	864	2760	11232	4.1
Kernel	600	560	3600	6.4
Absorption	600	48	720	15.0

Table 2.10: Comparison of performance of BerkeleyGW when disregarding spin and when using spinor wavefunctions, as seen in calculations of GaAs.

Step	Wall time, no spin (s)	Wall time spinor (s)	Ratio
I/O	138	560	4.0
Interp. WFN	57	240	4.2
Interp. Kernel	27	274	10.1
Diag.	53	3013	56.8

Table 2.11: Comparison of performance of Absorption executable in BerkeleyGW when disregarding spin and when using spinor wavefunctions, as seen in calculations of GaAs.

eigenvalues ("Absorption") have the largest increases, at 6.4 and 15.0, respectively. The Kernel code requires the calculation of three sets of matrix elements, an increase in cost partially offset by time spent on the better-scaling routines such as I/O. We discuss the Absorption code performance in more detail below.

The Absorption code has four main routines: I/O, interpolation of the wavefunctions, interpolation of the kernel matrix elements, and diagonalization. We see the performance of each when disregarding spin and when using spinor wavefunctions in Table 2.11. The I/O necessarily has an increase in cost of a factor of 4, from the increase in the size of the wavefunction files. Similarly, the interpolation of the wavefunctions is also nearly 4. The interpolation of the kernel matrix elements increases by a factor of 10.1, less than an estimated increase of 16. The diagonalization sees an increased cost by a factor of 56.8, close to the expected factor of 64.

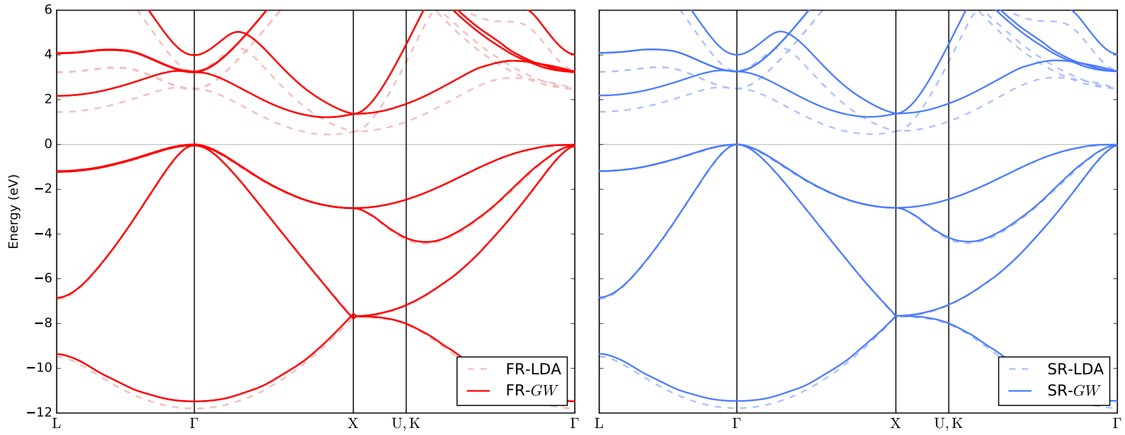

Figure 2.1: The electronic bandstructure of Si. Left-most figure: Fully-relativistic ("FR") LDA and *GW* in dashed and solid lines, respectively. Right-most figure: Scalar-relativistic ("SR") LDA and *GW* in dashed and solid lines, respectively; spin-orbit coupling is neglected.

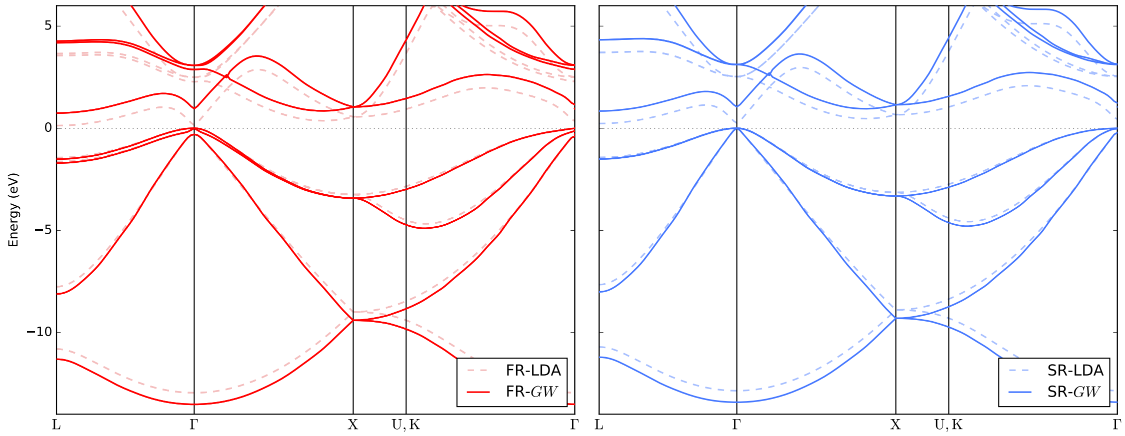

Figure 2.2: The electronic bandstructure of Ge. Left-most figure: Fully-relativistic ("FR") LDA and *GW* in dashed and solid lines, respectively. Right-most figure: Scalar-relativistic ("SR") LDA and *GW* in dashed and solid lines, respectively; spin-orbit coupling is neglected.

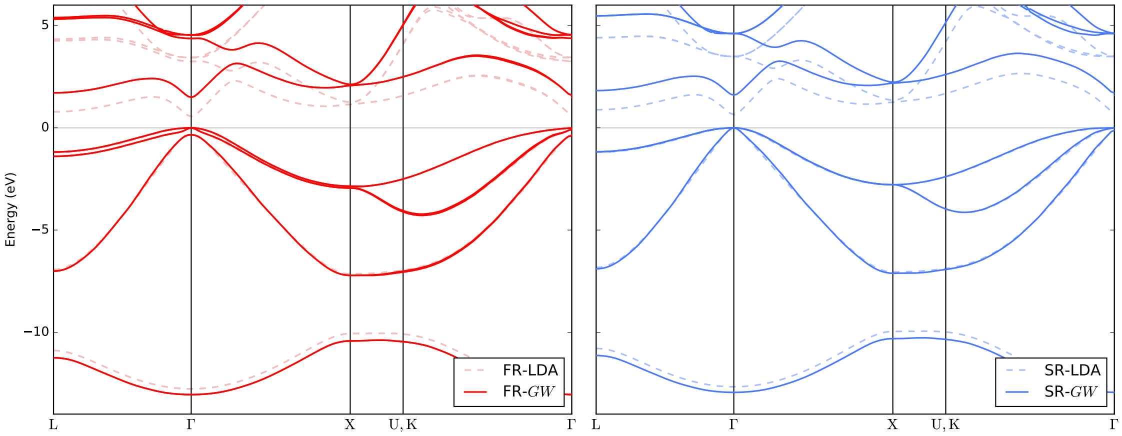

Figure 2.3: The electronic bandstructure of GaAs. Left-most figure: Fully-relativistic ("FR") LDA and GW in dashed and solid lines, respectively. Right-most figure: Scalar-relativistic ("SR") LDA and GW in dashed and solid lines, respectively; spin-orbit coupling is neglected.

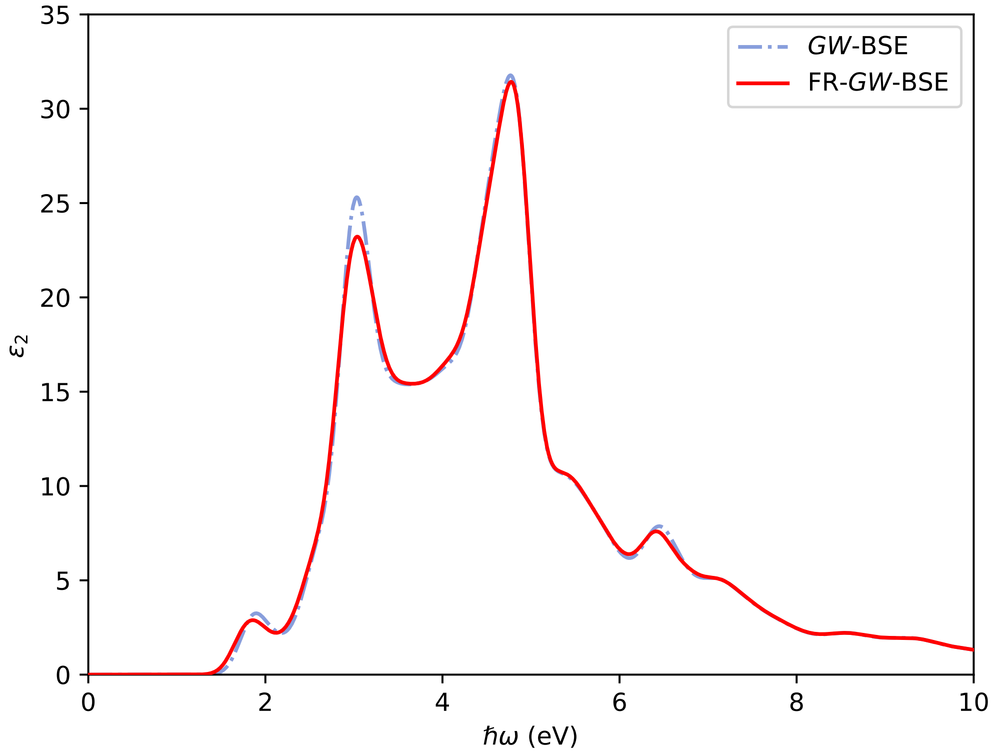

Figure 2.4: The absorption spectra of GaAs, calculated at the (FR-)GW-BSE level. Spin-orbit is neglected (included) in the blue (red) curve.

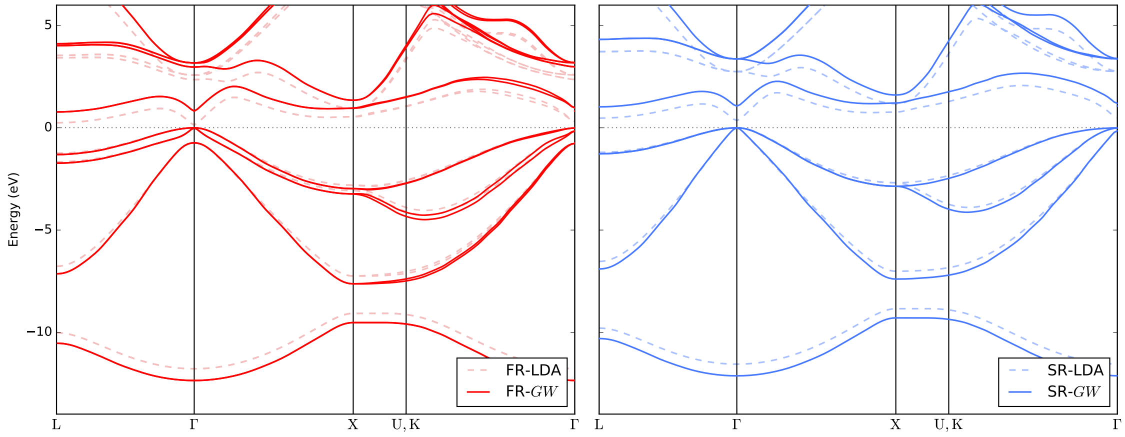

Figure 2.5: The electronic bandstructure of GaSb. Left-most figure: Fully-relativistic ("FR") LDA and GW in dashed and solid lines, respectively. Right-most figure: Scalar-relativistic ("SR") LDA and GW in dashed and solid lines, respectively; spin-orbit coupling is neglected.

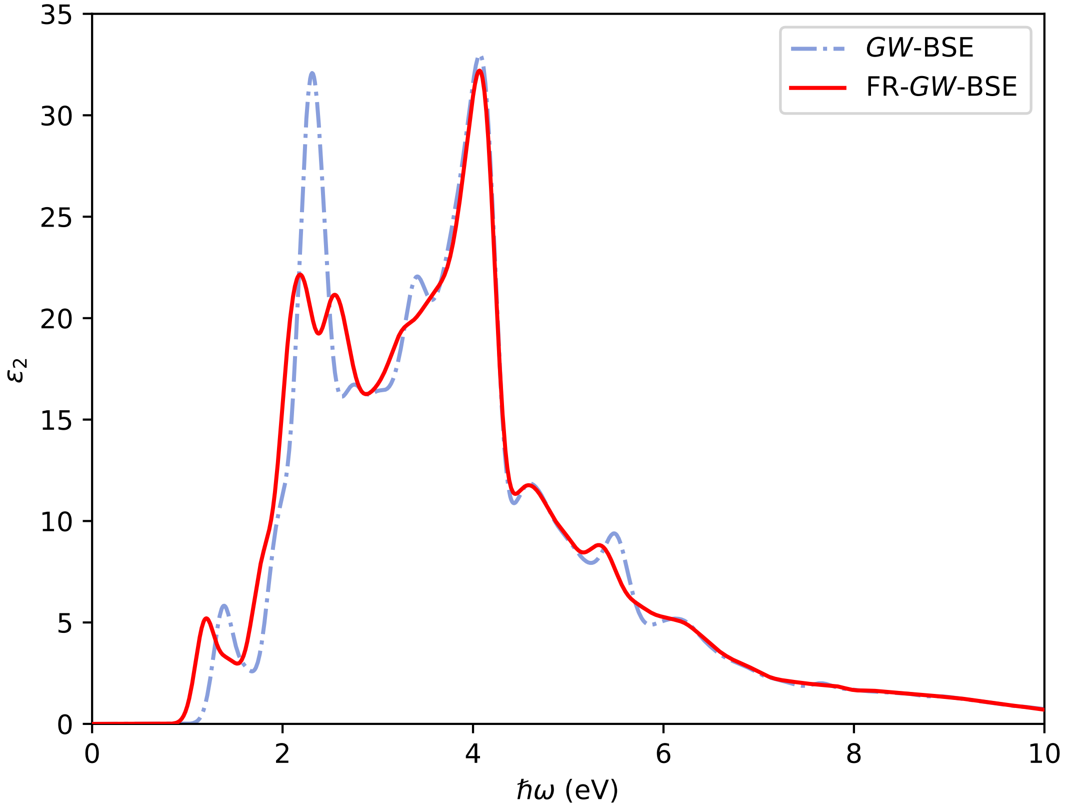

Figure 2.6: The absorption spectra of GaSb, calculated at the (FR-)GW-BSE level. Spin-orbit is neglected (included) in the blue (red) curve.

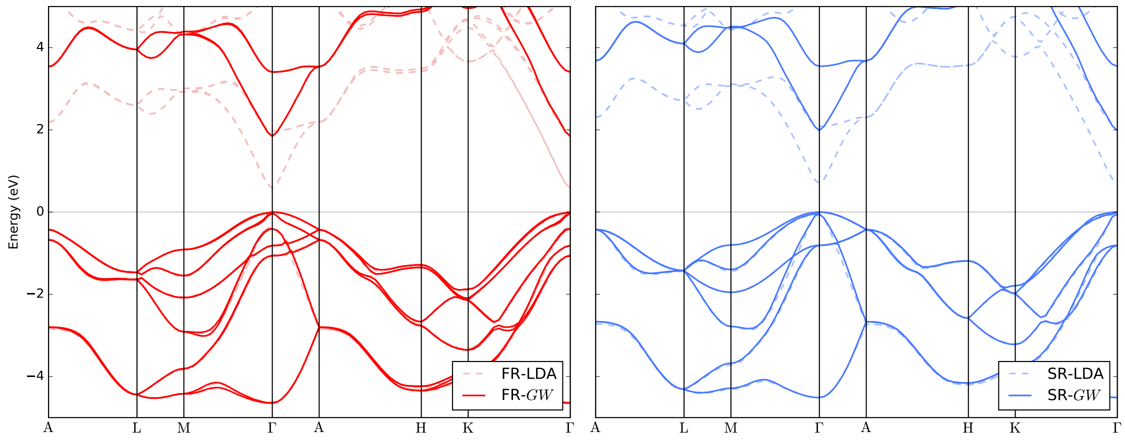

Figure 2.7: The electronic bandstructure of CdSe. Right-most figure: Fully-relativistic (“FR”) LDA and GW in dashed and solid lines, respectively. Left-most figure: Scalar-relativistic (“SR”) LDA and GW in dashed and solid lines, respectively; spin-orbit coupling is neglected.

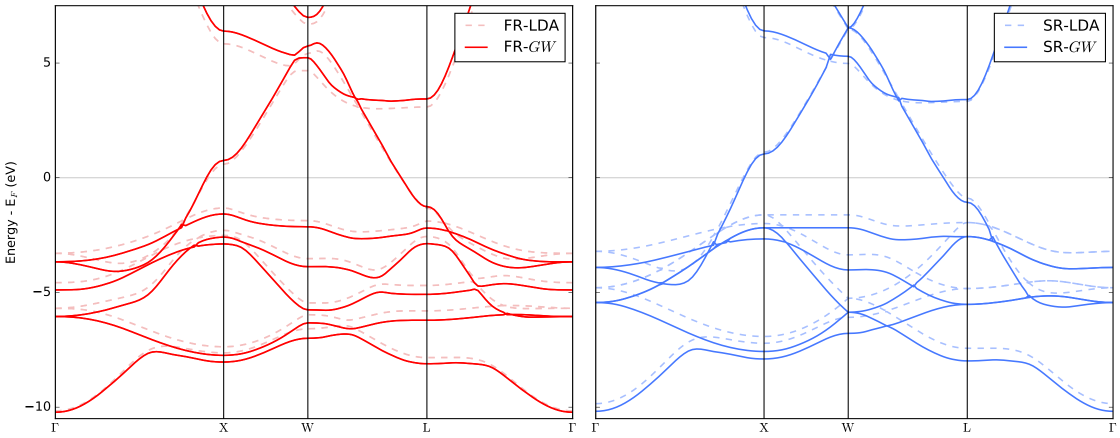

Figure 2.8: The electronic bandstructure of Au. Left-most figure: Fully-relativistic (“FR”) LDA and GW in dashed and solid lines, respectively. Right-most figure: Scalar-relativistic (“SR”) LDA and GW in dashed and solid lines, respectively; spin-orbit coupling is neglected.

Chapter 3

The electronic structure of β-HgS via *GW* calculations

3.1 Introduction

Metacinnabar, or β-HgS, has a zincblende structure with electrons near the band edge experiencing large spin-orbit coupling. It is in the mercury chalcogenide family, along with HgSe and HgTe. These latter two zincblende solids have been demonstrated to have a semimetallic, α-Sn-like bandstructure, with parabolic electron and hole bands joining at the Γ-point. For α-Sn, HgSe, and HgTe, the bands near the Fermi level at the Γ-point have an inverted bandstructure: The usual order of states for a zincblende system (from lowest to highest in energy) is Γ_{7v}, Γ_{8v}, Γ_{6c}, while these materials place Γ_6 lowest in energy, then Γ_7, then Γ_8. Furthermore, the four-fold degenerate Γ_8 splits into a pair of parabolic electron and hole bands, degenerate at the Γ-point. This behavior in the bandstructure is in agreement in the literature, before[99, 114] and after[101] the implementation of the fully-relativistic treatment of *GW* calculations.

The bandstructure for β-HgS, however, has seen considerable disagreement between different theoretical calculations and experiment. The α-Sn-like inverted bandstructure was proposed to be consistent with measurements[133]. Reflectivity measurements indicated a plasma edge at 0.10 eV[133], suggesting a metallic or semimetallic nature. Absorption measurements indicated an onset of interband transitions at 0.25 eV (after free-carrier-like absorption), understood as the onset of transitions to a partially occupied, zero-gap parabolic conduction band[133]. In the absence of ARPES measurements, however, we do not definitively have a definitive experimental description of the band structure and the true ordering of the states.

Density functional theory calculations, however, indicated a small indirect band gap and a "doubly inverted" ordering of the states, Γ_{6v}, Γ_{8v}, Γ_{7c}[25]. Compared to HgSe and HgTe, there is a further interchange of ordering between Γ_8 and Γ_7, with Γ_8 now being fully occupied even away from the Γ-point. The semimetallic state is only recreated in DFT with the neglect of spin-orbit coupling (SOC). As the spin-orbit strength within the Hg $5d$ subshell is 1.86 eV[84], this is not a physical result. Based on the fully-relativistic DFT bandstructure, β-HgS has been predicted to be a nontrivial Z_2 insulator much like strained HgTe, with highly anisotropic topologically protected Dirac surface states along the [001] direction[123, 122].

Topological properties based on the ordering of bands near the Fermi energy need to be confirmed from calculations that are more accurate than DFT[121]. Many-body perturbation theory, in the *GW* approximation, provides physically accurate excited state properties such as the electronic bandstructure. *GW* calculations in which the spin-orbit coupling Hamiltonian is applied as a perturbation to the quasiparticle states ("GW+SOC") indicated a band ordering of Γ_{8v}, Γ_{7v}, Γ_{6c}[31, 114], which is much like that of CdTe, but with the p-d-hybridized orbitals Γ_{8v} and Γ_{7v} switched, due to the strength of SOC in the Hg d states. This ordering of states yields a topologically trivial band gap, as the bandstructure can be adiabatically deformed to that of the topologically trivial CdTe[7] without closing the bulk band gap by, e.g., tuning the atomic spin-orbit parameters of Hg and S.

GW calculations in which fully-relativistic pseudopotentials and spinor wavefunctions were used in the construction of the self-energy ("FR-*GW*"), thereby treating spin-orbit coupling non-perturbatively, yielded the same band ordering as in DFT[101, 104], but with the energy difference between the Γ_6 and Γ_8 states much reduced when using an LDA exchange-correlation functional as the starting point, and a slightly larger energy difference when calculated with the GGA exchange-correlation functional[104].

The disagreement in the various computed quasiparticle bandstructure topologies is curious, since the fully-relativistic LDA ("FR-LDA") bandstructures are all in agreement despite the different choices of basis sets. In this work, we present the FR-*GW* quasiparticle bandstructure using a plane-wave basis and approximating the electron-ion interaction within the fully-relativistic Local Density Approximation (LDA)[90] of the pseudopotential method. In Section 3.2 we review the method of (one-shot) FR-*GW* in a plane-wave basis. In Section 3.3 we show that, contrary to the results of the other FR-*GW* calculations in the literature, the states at the Γ-point are ordered Γ_8, Γ_7, then Γ_6, in increasing energy.

3.2 Methods

In computing the quasiparticle eigenvalues, we solve for the poles of Dyson's equation,

$$E_{n\mathbf{k}} = \epsilon_{n\mathbf{k}} + \langle n\mathbf{k}\alpha| \left(\Sigma_{\alpha\beta}\left(E_{n\mathbf{k}}\right) - V^{xc}\delta_{\alpha\beta}\right) |n\mathbf{k}\beta\rangle. \tag{3.1}$$

We note that in this work, we evaluate the polarizability by explicitly calculating non-zero imaginary frequencies, and the integration over frequency in the Coulomb-hole self-energy is performed using the contour deformation method[37, 12, 39].

3.3 Calculations

We construct the Kohn-Sham energies and orbitals using Quantum ESPRESSO[35], with fully-relativistic pseudopotentials for Hg and S generated from the Optimized Norm-Conserving Vanderbilt Pseudopotential method[40] with parameters adapted from the Pseudo-Dojo pseudopotential database[67]. The Hg pseudopotential includes the $5s^2 5p^6 5d^{10}$ semicore states as valence for accurate calculation of the bare exchange matrix elements [116]. We use a plane-wave cutoff of 200 Ry and an 8x8x8 k-point grid in calculating the charge density and relaxed structural geometry. We find that the relaxed lattice parameter to be identical to the experimental value of 5.85 Å.

The conventional band gap is defined to be $E_0 = E(\Gamma_6) - E(\Gamma_8)$, which is positive for usual zincblende materials but negative for inverted systems. For the usual zincblende materials (e.g. GaAs or CdTe), this is equal to the true band gap. Likewise, the spin-orbit splitting $\Delta^{\mathrm{SOC}} = E(\Gamma_8) - E(\Gamma_7)$ is also defined to be positive for the usual zincblende materials. However, the Hg $5d$ orbitals contribute to these states significantly, as p-d orbital hybridization is allowed for tetrahedral symmetries[127]. The spin-orbit coupling of these states, then, is dominated by the contribution of the Hg $5d$ atomic states. Since the $5d_{5/2}$ state contributes to the Γ_7 and the $5d_{3/2}$, to Γ_8, Δ^{SOC} will be negative, as the spin-orbit splitting in these states (1.86 eV[84]) dominates that of the S $6p_{3/2}$ and $6p_{1/2}$ states (< 0.10 eV[83]).

The fully-relativistic LDA bandstructure is shown in Fig. 3.1. At Γ, spin-orbit coupling breaks the degeneracy at the Fermi level and opens a small spin-orbit gap, with the Γ_7 states unoccupied and the four-fold Γ_8 states occupied. The small indirect gap is 0.10 eV.

We then use the LDA bands to calculate the dielectric function and quasiparticle eigenvalues with the BerkeleyGW[26] excited-state code adapted for compatibility with spinor wavefunctions. We use a dielectric matrix cutoff of 35 Ry and use 2000 empty states for both the polarizability and Coulomb-hole sums. The error in the band gap due to the use of 2000 empty states in the Coulomb-hole sum is estimated to be 7 meV. We use an 8x8x8 **q**-point grid for the dielectric function, as this grid has been shown to be sufficient for accurate calculations of the band gap for the similarly sized Ge (Sec. 2.5).

The gap at Γ has also been calculated with the frequency dependence of the self-energy treated by the COHSEX and GPP[136] approximations, in addition to explicit calculation of the frequency dependence. We treat the $\mathbf{q} \rightarrow 0$ limit with the dual grid technique, appropriate for a semiconductor even with a small gap[26]. The results from the GPP and the explicit calculation of the frequency dependence include the contributions from the static-remainder method[27] to more accurately simulate an infinite sum in the COH term.

We calculate the quasiparticle energies within FR-*GW* approach at discrete points along the Γ-L and Γ-X lines, with explicit treatment of the frequency dependence. In units of the length of these paths in the Brillouin Zone, the points are 1/8, 1/4, 1/2, and 3/4, along with the Γ, L, and X points. The bandstructure is then plotted with cubic splines, as the usual approach in which the LDA bandstructure is used for a linear interpolation based on the Bloch states relies on the LDA bandstructure being qualitatively similar to the quasiparticle bandstructure[26]. These calculations are expedited by the use of the "static subspace" method[12], in which we find the eigenvectors of $\epsilon_{\mathbf{GG}'}(\mathbf{q}, \omega = 0)$ for each **q**-point and use a fraction of that basis in the calculation of the self-energy matrix elements. We find that using the lowest 10 percent of the static dielectric matrix eigenvectors ($N_{basis} = 117$) gives a converged fundamental band gap within 6 meV. We use 15 imaginary frequencies for contour deformation method, which is converged within 3 meV as compared to the use of 25 frequencies.

We find that the states near the Fermi energy are ordered Γ_8, Γ_7, then Γ_6, with the latter forming the conduction band minimum. This is in contrast to the other reported FR-*GW* calculations in the literature[101, 104], where the states are ordered, as in DFT, Γ_{6v}, Γ_{8v}, then Γ_{7c}. We find our ordering in agreement with other *GW* calculations in which spin-orbit coupling is added perturbatively[31, 114]. We note that the authors in Ref. [101] constructed their polarizability by using the states and eigenvalues from a DFT calculation ignoring spin-orbit coupling. The bandstructure in this case has a vanishing energy gap at Γ, in contrast to the small gap when including SOC. The vanishing gap will yield a divergence in the dielectric function for $\mathbf{q} \rightarrow 0$, giving an overestimation of screening and a smaller correction to the band gap than would be otherwise given.

Moreover, our calculation of the band gap and electron effective mass were found to agree well with the literature. The band gap was estimated to be -0.11 eV from a Shubnikov-de Haas measurement of transition metal-doped samples of β-HgS, with

	E_0 (eV)	Δ^{SOC} (eV)	E_g (eV)
DFT	-0.54	-0.10	0.10
COHSEX	0.92	-0.12	0.80
GPP	0.37	-0.15	0.22
Explicit Frequency	0.23	-0.13	0.10

Table 3.1: Interband gaps near the Fermi energy for β-HgS. At the DFT level (in the LDA), The four-fold Γ_8 states are higher in energy than the two-fold s-like Γ_6 states. With the inclusion of self-energy effects at any level (Hartree-Fock, static-COHSEX, GPP, or Explicit Frequency), the Γ_8 states are lower in energy than the now-unoccupied Γ_6 states.

the carrier concentration extrapolated to zero.[29]. The model used to interpret the experiment, however, depends only on the absolute value of the band gap. With the view that the band gap is in fact positive, our computed value of 0.10 eV is in almost perfect agreement. Similarly, we compute the electron effective mass to be $0.07m_e$. This is in the range of values cited for the transition metal-doped β-HgS samples, $0.04m_e$ to $0.07m_e$[29]. The electron effective mass estimated from reflectivity measurements gives the value $0.07m_e$[133].

We confirmed that the topologically trivial ordering of bands was reproduced within static-COHSEX and GPP FR-GW. The ordering of states was indeed the same in all cases. The band gaps of interest are summarized in Table 3.1.

We also sought to confirm the ordering of the states by performing an off-diagonal static-COHSEX calculation to determine if the assumption that the LDA orbitals are nearly identical to the quasiparticle wavefunctions breaks down, apart from a mere reordering of states. Care must be taken for small gap semiconductors with large spin-orbit coupling that the self-energy is calculated with the appropriate basis set, especially when the ordering of the quasiparticle energies changes relative to LDA.

The static-COHSEX approximation is a static approximation to the self-energy operator. The frequency-independence of the static-COHSEX operator allows us to use a completion relation to remove the sum over empty states in the Coulom-hole term in the self-energy:

$$\langle n\mathbf{k}\alpha|\Sigma^{\mathrm{COH}}_{\alpha\beta}(\omega=0)|m\mathbf{k}\beta\rangle = \frac{1}{2}\sum_{\mathbf{q}\mathbf{G}\mathbf{G}'}\langle n\mathbf{k}\alpha|e^{i(\mathbf{G}'-\mathbf{G})\cdot\mathbf{r}}\delta_{\alpha\beta}|m\mathbf{k}\beta\rangle\left[\epsilon^{-1}_{\mathbf{G}\mathbf{G}'}(\mathbf{q};\omega=0)-\delta_{\mathbf{G}\mathbf{G}'}\right]v(\mathbf{q}+\mathbf{G}'). \tag{3.2}$$

	E_0 (eV)	Δ^{SOC} (eV)	E_g (eV)	Basis set
Present Work (FR-G_0W_0)	0.23	-0.13	0.10	Plane-wave (PW)
G_0W_0+SOC[a]	0.18	-0.12	0.06	Gaussian+PW
hybrid-QSGW[b]	0.37	-0.07	0.31	LMTO
FR-G_0W_0 [c]	-0.02	-0.19	0.02	FLAPW
FR-G_0W_0 [d]	-0.02	-0.10	0.02	Plane-wave

Table 3.2: Comparison of results for interband gaps near the Fermi energy from the literature. [a] Ref. [31], [b] Ref. [114], [c] Ref. [101], [d] Ref. [104].

Thus static-COHSEX allows for rapid calculation of matrix elements of an approximate form of the self-energy operator. We estimate the m'th *LDA* wavefunction's contribution to the first-order correction to the static-COHSEX wavefunctions from

$$\psi^{(1)}_{n\mathbf{k}}(\mathbf{r}) = \sum_{m\neq n} U^{(1)}_{nm\mathbf{k}} \phi^{\text{LDA}}_{m\mathbf{k}}(\mathbf{r}), \tag{3.3}$$

with

$$U^{(1)}_{nm\mathbf{k}} = \frac{\langle n\mathbf{k}\alpha|\Sigma^{\text{COHSEX}}_{\alpha\beta} - V^{xc}\delta_{\alpha\beta}|m\mathbf{k}\beta\rangle}{\epsilon_n - \epsilon_m}. \tag{3.4}$$

Nonzero values of $|U^{(1)}_{nm\mathbf{k}}|$ indicate a contribution from the m'th LDA band to the n'th static-COHSEX band and we must fully diagonalize Dyson's equation to arrive at the true static-COHSEX wavefunctions. In a rather extreme case, the values of U may be distributed in such a manner that the quasiparticle states are re-ordered LDA states. We look, then, for any possible further reordering of the Γ-point states after the inclusion of off-diagonals.

We plot in Fig. 3.3 the values of $|U^{(1)}_{nm\mathbf{k}}|$, with the rows n and columns m. We see that only the Γ_6 and Γ_7 states show non-zero contribution to the others' quasiparticle wavefunctions to first order. However, this contribution is at most still less than 0.1 percent of the zero-order contribution, so the LDA bands are indeed good approximations to the quasiparticle states.

To gain further physical insight, we calculate the bandstructure in DFT but now with non-relativistic and scalar-relativistic pseudopotetials. Scalar-relativistic pseudopotentials neglect spin-orbit coupling but include the physics of the Darwin term and the relativistic mass correction, which are both significant for s-states. The non-relativistic pseudopotentials neglect even these. In Fig. 3.4 we show that the non-relativistic pseudopotentials give a bandstructure typical for a II-VI or III-V zincblende material with light atoms, where an s-like Γ_1 state is placed E_g above occupied Γ_{15} states. Incorporating the Darwin and relativistic mass-correction terms, however, lower the Γ_1 below the previous band gap, and it then becomes occupied (Fig. 3.5). The Γ_{15} states then split in a manner to form parabolic conduction and valence bands that are degenerate at Γ, predicting a zero-gap semimetallic state. Incorporating spin-orbit coupling breaks this degeneracy, and the states can be identified as belonging to either the Γ_7 or Γ_8 representations.

We may now imgagine tuning first the band gap E_g and then the relativistic corrections Δ^{rel}, in analogy to Ref. [131], starting from a DFT-like calculation neglecting first all relativistic effects. A sufficient increase in E_g places the Γ_1 state high enough so that tuning Δ^{rel} lowers it but not by an amount sufficient to place it among the occupied states. Tuning Δ^{rel} further turns on spin-orbit coupling and splits the occupied Γ_{15} states into the Γ_7 and Γ_8 states, leading to the appropriate value of the band gap. We note that in practice, by beginning with states and eigenvalues from FR-DFT, we have in this framework tuned Δ^{rel} first and then opened the gap with the use of FR-GW, with the physics of the final bandstructure being insensitive to the ordering of which parameters were employed.

3.4 Conclusion

We calculate the bandstructure of β-HgS near the Fermi energy with the FR-GW method and find that the order of the states differs from that predicted by DFT calculations. While DFT gives an ordering of the bands Γ_{6v}, Γ_{8v}, Γ_{7c}, we find the quasiparticle bandstructure is ordered Γ_{8v}, Γ_{7v}, Γ_{6c}. The difference in the ordering changes the prediction of the topological phase of the material, with DFT giving a false-positive identification of the material as a Z_2 insulator.

Our calculated value of the band gap is 0.10 eV, which compares favorably to the value of 0.11 eV from experiment[29], and our calculated value of the electron effective mass is $0.07m_e$, which agrees well with the range of values cited in experiment of 0.04-$0.07m_e$[133, 29]. This close agreement with experiment supports the conclusion that our presented calculation is sufficiently accurate. We then confirmed that the ordering is reproducible via other approaches of computing GW, the COHSEX and GPP methods. Our calculation of the possible deviation of the LDA basis from the true quasiparticle states indicated that the use of the LDA states is accurate to within 0.1 percent, so self-consistency in the wavefunctions would not appreciably change our results.

The reversal of the band inversion predicted in DFT can be understood in terms of the band gap problem. The underestimation of the band gap places the s-like Γ_6 state too low in energy, such that it becomes occupied due to the strength of its change in energy upon incorporation of the relativistic mass correction and Darwin

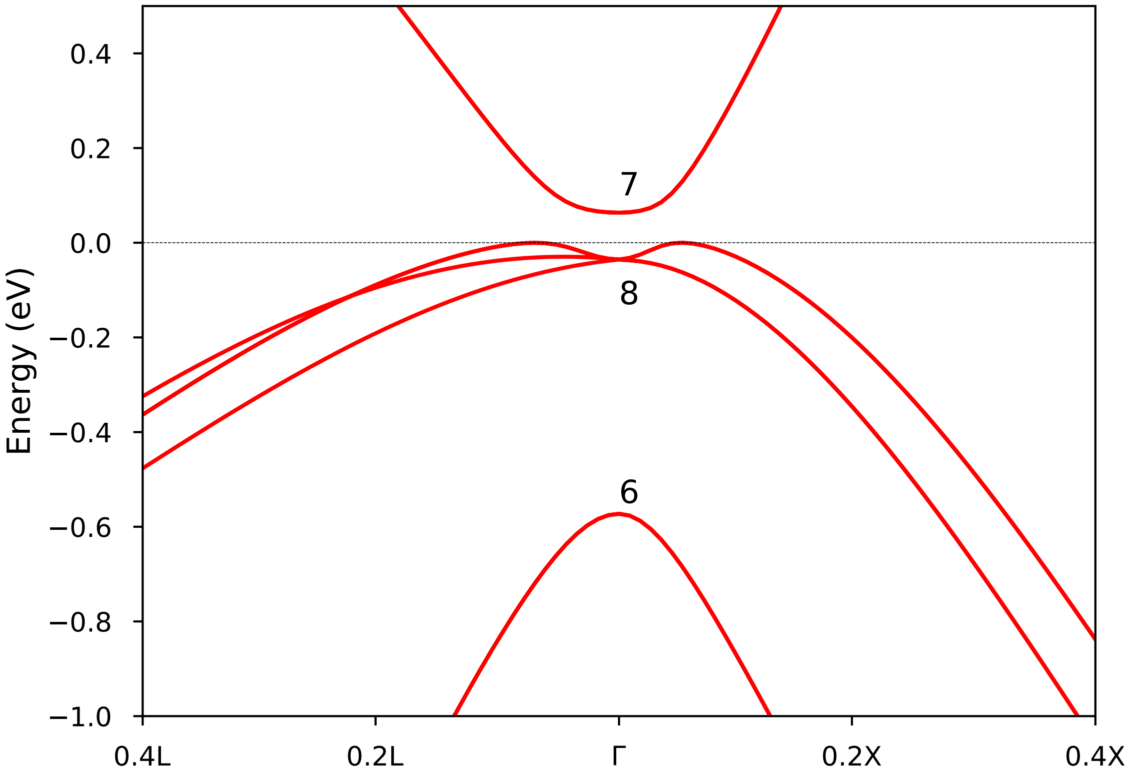

Figure 3.1: The fully-relativistic LDA bandstructure of β-HgS. The states, from low energy to high, are Γ_6, Γ_8, Γ_7.

terms captured in scalar- and fully-relativistic pseudopotentials. *GW* calculations raise its energy sufficiently to place it above the Fermi level.

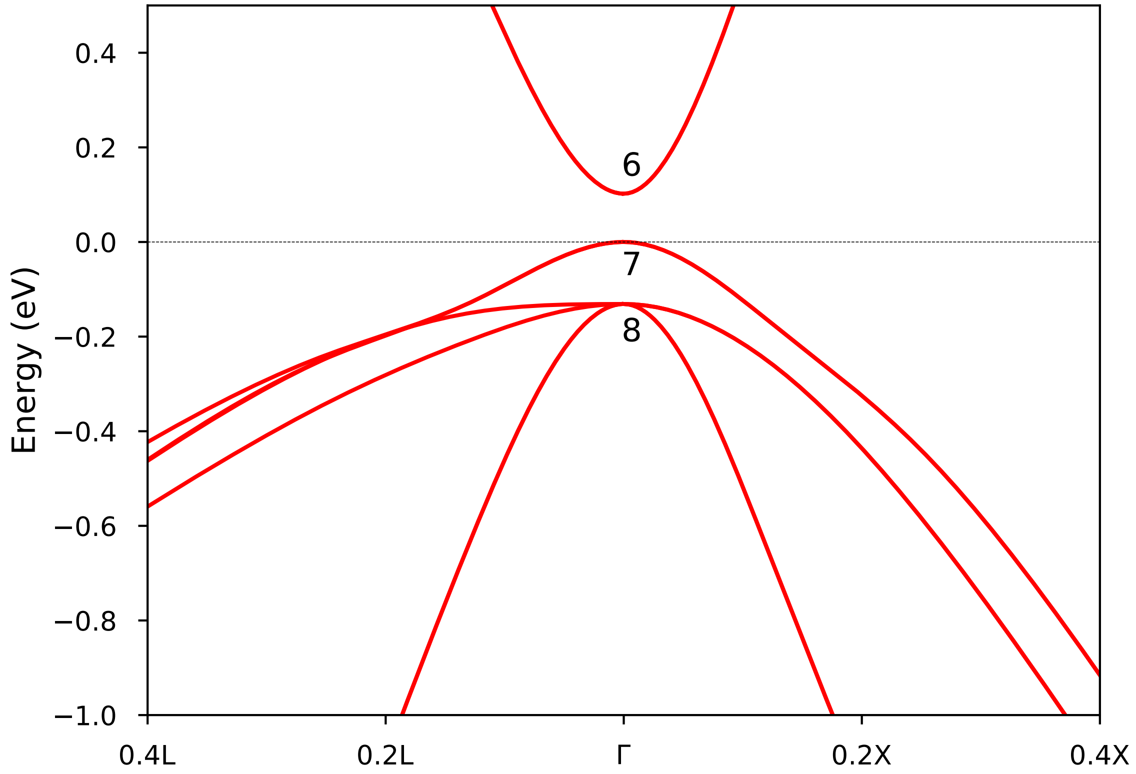

Figure 3.2: The quasiparticle bandstructure of β-HgS, computed at the FR-GW level using the contour deformation method. The bands at the Γ point are labelled by their irreducible representation.

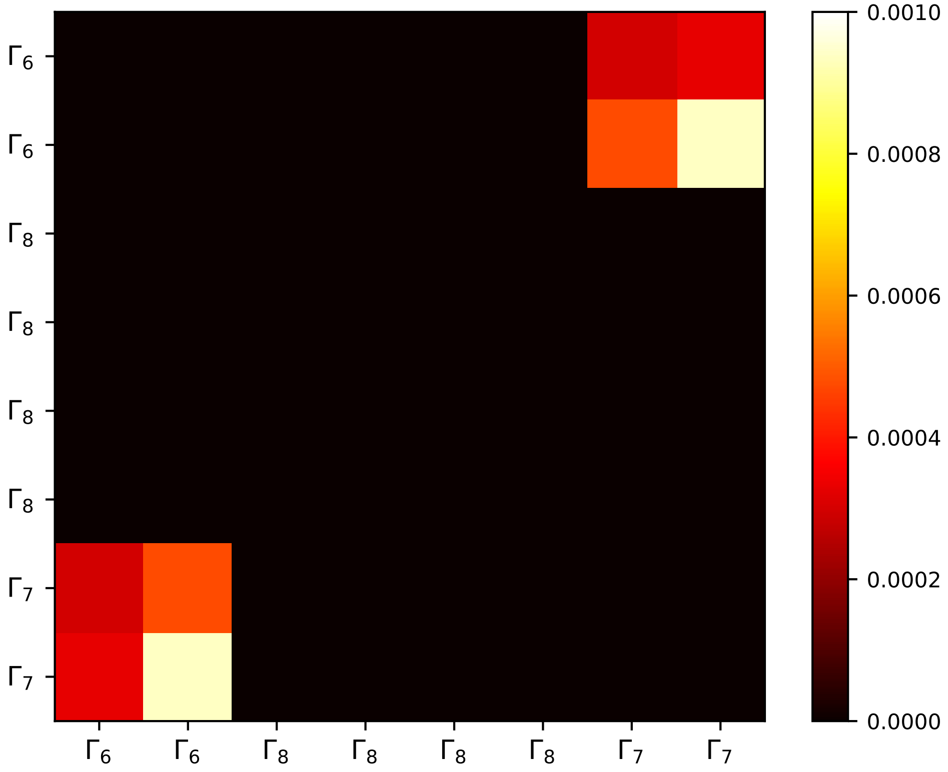

Figure 3.3: $U^{(1)}_{nm\mathbf{k}}$, the first-order correction to each COHSEX wavefunction, within the LDA basis. The n'th COHSEX state (columns) is the n'th LDA state, plus the sum of each contribution of the remaining LDA states (rows), the magnitude of which is indicated by the heat-map. The maximum off-diagonal contribution is less than 0.1 percent.

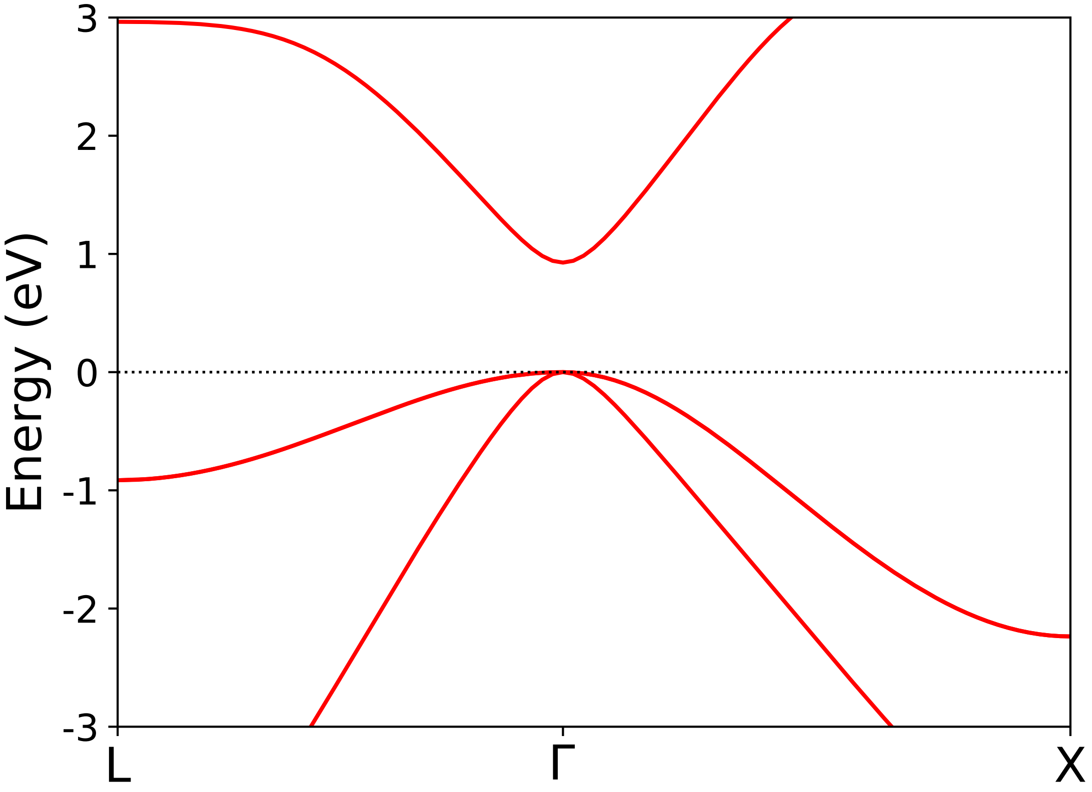

Figure 3.4: The non-relativistic LDA bandstructure of β-HgS. The states at the Γ point are ordered Γ_{15}, Γ_1, from occupied to unoccupied, as in a conventional zincblende structure.

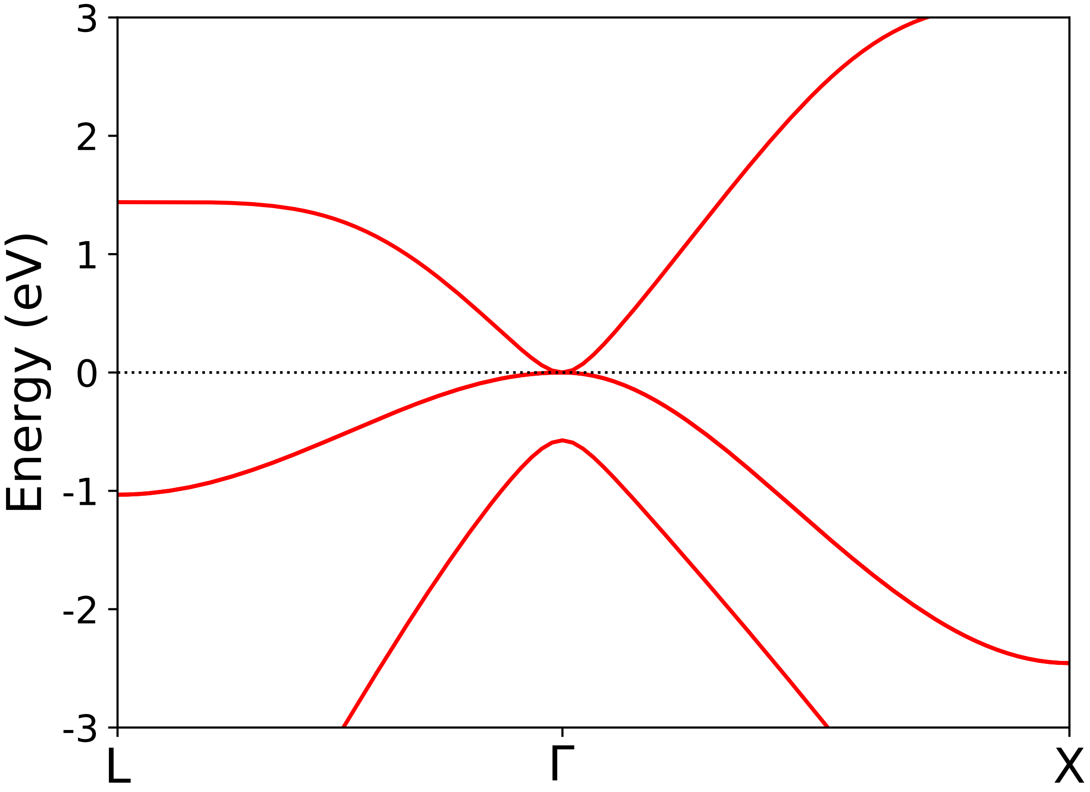

Figure 3.5: The scalar-relativistic LDA bandstructure of β-HgS. The degenerate states at $E_{\mathrm{F}} = 0$ belong to the Γ_{15} representation, and the lower occupied state belongs to Γ_1.

Chapter 4

The Quasiparticle Bandstructure of Bi_2Se_3

4.1 Introduction

The narrow bandgap semiconductor Bi_2Se_3 is a prototypical topological insulator, where the surface states are protected against scattering by time-reversal symmetry[3]. While Bi_2Se_3 has a small bandgap compared to other semiconductors, its value of 0.2 to 0.3 eV is large relative to the set of experimentally verified topological insulators, which range from under 10 meV to 0.3 eV[3], though some systems are predicted to have non-trivial topology and a bandgap on the order of 1 eV[53, 109].

The large spin-orbit splitting of the Bi 6*p* electrons inverts the positive and negative parity *p*-like states (from the Bi 6*p* and Se 4*p* orbitals) near the band gap, creating a nontrivial value of the Z_2 topological index[134]. The band gap is caused by the level-repulsion of the now-inverted states at the Γ-point, mixing the character of the conduction and valence states within a neighborhood of Γ[131].

4.2 Summary of previous experimental results

The electronic structure of materials can be determined experimentally by various means. Angle-resolved photoemission spectroscopy (ARPES) gives direct access to the excited quasiparticle states emitted from a particular surface, allowing for the experimental determination of the quasiparticle bandstructure for a two-dimensional projection of some path in the Brillouin zone. Scanning tunneling spectroscopy (STS) allows for the measurement of the density of states, which can give insight into the van Hove singularities and band edges of a material. Optical transmission passes infrared or optical beams through a bulk sample of the material of interest, sweeping through frequencies until the beams reach the absorption edge. Magneto-optics can further resolve details of the structure of the band edges. We note that ARPES and STS are measurements of surfaces, while optical transmission measures the optical gap. The optical gap is approximately the quasiparticle gap when the excitonic binding energy is sufficiently small, as is often the case for three-dimensional semiconductors.

ARPES measurements of n-doped Bi_2Se_3 indicate a "camel-backed" or "M"-shaped bulk valence band and a parabolic bulk conduction band, along with a Dirac cone from the topologically non-trivial surface states. The presence of band-bending, highly doped samples via Se vacancies, and access only to surface bandstructures makes interpreting ARPES results for bulk properties difficult. However, with a wide range of photon energies and surfaces, Necheav, et al. found a direct gap of 0.33 eV through multiple ARPES measurements[87]. Further, STS measurements by Kim, et al, estimate a band gap of 0.3 eV.

Optical transmission from Black, et al. [9] indicate a room-temperature optical gap of 0.36 eV for Bi_2Se_3. This group also measured optical gaps (also at room temperature) for the similar materials As_2Te_3, Sb_2S_3, Sb_2Se_3, Sb_2Te_3, Bi_2S_3, and Bi_2Te_3, with all values agreeing well with reported results through present: 1, 1.7, 1.2, 0.3, 1.3, and 0.15 eV, respectively. We note that contemporaries often measured an optical gap in the range of 0.15 - 0.2 for Bi_2Se_3[72].

Magneto-optic measurements from Piot, et al.[88] and Martinez[79] confirm the nearly parabolic features of the valence and conduction bands near the bulk band gap, as suggested by ARPES[87]. However, these measurements find a small band gap of 0.2 eV.

4.3 Summary of previous theoretical calculations

DFT bandstructure

While a difference of 0.1 eV in reported energy gaps for most systems would not be significant, this error is on the order of the gap of Bi_2Se_3, and is either 50 percent or 33 percent of the value. This scale of discrepancy is also within the usual systematic error of the *GW* method, as well. We will now discuss that there is significant difficulty in the calculation of the DFT bandstructure when using pseudopotentials, as well.

Bi_2Se_3 has been found to have a fully-relativistic DFT band gap value that is extremely sensitive to the choice of functional and the use of relaxed or experimental geometries[100], with a plane-wave basis pseudopotential code, Abinit[38]. The position of the conduction band minimum was found to vary in the LDA depending on the use of experimental or relaxed coordinates. The gap at Γ, which in DFT is not the fundamental gap, for the relaxed structure is about 100 meV larger in the LDA than the experimental structure. The gap at Γ has the opposite behavior in the PBE, where relaxation reduces the gap by approximately the same amount [100]. The LDA and PBE bandstructures with the experimental coordinates also disagree somewhat, with LDA having a more deeply camel-backed valence band and thus a larger gap at Γ by roughly 100 meV.

With an all-electron approach to DFT, there is still some sensitivity. Nechaev, et al. found band gaps at Γ for unrelaxed LDA (0.5 eV), relaxed LDA (0.4 eV), unrelaxed GGA (0.4 eV), and relaxed GGA (0.3 eV). In either case of LDA or GGA, relaxing reduces the band gap, unlike in the pseudopotential method. The other trends are similar, though with a systematically smaller value of the gap by about 100 meV.

GW quasiparticle bandstructure

The first approach to calculating the quasiparticle bandstructure used the "GW+SOC" method, where calculations were first performed by ignoring spin-orbit coupling. After the quasiparticle energies are calculated in this manner, the spin-orbit coupling is added as a perturbation, which can be solved to arbitrary order in perturbation theory through exact diagonalization. Yazyev et al. performed such a calculation with BerkeleyGW, beginning from an LDA-based plane-wave pseudopotential mean-field calculation, and found a direct gap of 0.33 eV, with the valence and conduction bands being parabolic[131].

An additional calculation using the GW+SOC approach, using the all-electron code SPEX[33], though, found a vanishing band gap. Further, the bands appeared to become linear, which is unexpected for the bulk. The FR-GW calculation in this work, however, gave the reasonable values of 0.2 eV. Interestingly, the GW+SOC calculation of Bi_2Te_3 was within 10 meV of the FR-GW result[2].

The FR-GW calculation from Fœrster et al. begins with a Gaussian basis and an LDA exchange-correlation functional and arrives at a direct bulk band gap at Γ, of 0.2 eV[32].

A plane-wave pseudopotential calculation with a PBE functional arrives at a FR-GW gap of 0.36 eV[66] using the YAMBO code[75]. The quasiparticle bandstructure in this work suggests a direct gap at Γ, though the resolution is not fine enough to determine if the bands have a parabolic dispersion.

4.4 Quasiparticle bandstructure as computed within BerkeleyGW with spinor functionality

Due to the sensitivity of this system on the functional and the geometry, in our study of the bulk band gap of Bi$_2$Se$_3$, we use the experimental geometry. For consistency with the majority of previous calculations in the literature, we use the LDA functional. We use a Brillouin zone sampling of 8x8x8 in constructing the charge density as well as the dielectric function. We use a 160 Ry cutoff for the planewave basis for our wavefunctions and a 25 Ry cutoff for the dielectric function. The polarizability ("Chi") summation uses 1000 unoccupied bands, and the Coulomb-hole ("COH") summation, 1254 bands.

We compute the bandstructure in the neighborhood of the Γ-point by computing quasiparticle energies at the Γ-point and at particular points along the Γ-to-L and the Γ-to-Z high-symmetry lines. The Γ-to-L direction has energies computed at $\frac{1}{16}L$, $\frac{1}{8}L$, $\frac{3}{16}L$, and $\frac{1}{4}L$. The Γ-to-Z direction has energies computed at $\frac{1}{16}Z$, $\frac{1}{8}Z$, $\frac{3}{16}Z$, $\frac{1}{4}Z$, $\frac{1}{2}Z$, and Z. The whole Γ-to-Z line is represented as it is a much shorter path in the Brillouin zone than the Γ-to-L line. We then plot interpolated curves as estimates to the quasiparticle bandstructure. The LDA bandstructure interpolated in a similar fashion shows a high agreement with the bandstructure calculated explicitly at each k-point. The FR-GW bandstructure along the Γ-to-Z line suggests that the band-diagonal approximation is not generally sufficient, as is apparent from the appearance of small dimples in both the conduction and valence bands in a very narrow region about Γ (Fig. 4.1).

We find a direct bulk band gap of 0.38 eV in the band-diagonal approximation, which is in good agreement with values reported by ARPES[87] (0.332 eV) as well as STS[57] (0.3 eV). Optical measurements of the gap, however, report a smaller value of 0.2 eV[88] but confirm a direct band gap at Γ.

We further investigate the effect of including band-offdiagonals in the calculation of the self-energy matrix elements:

$$E_{n\mathbf{k}} = \epsilon_{l\mathbf{k}}\delta_{lm} + \mathrm{eig}\left(\langle l\mathbf{k}\alpha| \left(\Sigma_{\alpha\beta}\left(E_{p\mathbf{k}}\right) - V^{\mathrm{xc}}\delta_{\alpha\beta}\right)|m\mathbf{k}\beta\rangle\right), \tag{4.1}$$

where "eig" denotes the eigenvalues of the matrix constructed from the self-energy in the Kohn-Sham orbital basis, the band n is a member of the set of bands spanned by the indices l and m, and the energy $E_{p\mathbf{k}}$ at which the self-energy operator is either the row or column, as the difference in eigenvalues from this choice and a truly Hermitian construction is found to be well under 1 meV.

We find that the choice of the four valence bands and two conduction bands near the Fermi energy is sufficient to correct the deficiencies in the bandstructure when using the LDA eigenfunctions as the quasiparticle wavefunctions (Fig. 4.2). The necessity of calculating a matrix for the self-energy can be seen by noting the strength of the level repulsion – and character of the wavefunctions – depends on the band gap value[131]. When changing the gap, as in a *GW* calculation, the wavefunctions in the region where the character is inverted necessarily change along with the extent of the region in the bandstructure with inverted orbital character.

We note that in evaluating the effects of expanding the basis for the quasiparticle wavefunctions, we evaluate the self-energy at the frequencies consistent with the previous basis. That is, the first correction to the basis diagonalizes the self-energy evaluated at the Kohn-Sham eigenvalues. The matrix elements in the diagonal basis, then, are not strictly quasiparticle energy eigenvalues. However, since there are no off-diagonal terms for the self-energy at the Γ-point, the self-consistently calculated quasiparticle energies must match exactly at the Γ-point. Assuming that the off-diagonal matrix elements of the self-energy for the states away from the Γ point are weakly sensitive to corrections to the Kohn-Sham eigenvalues, we may simply take the eigenvalues E^0 calculated from diagonalizing $\Sigma(\omega = \epsilon^{\mathrm{KS}})$ and rigidly shift them such that $E^{\mathrm{QP}}_{n\mathbf{k}} \approx E^0_{n\mathbf{k}} + (E^1_{n\Gamma} - E^0_{n\Gamma})$. The eigenvalues $E^1_{n\mathbf{k}}$ are the corrections to E^0 by way of Newton's method, which itself is the lowest-order correction to evaluating the self-energy operator at the Kohn-Sham eigenvalues. Equivalently, we may keep the valence bands fixed and just correct the conduction bands with $E^1_{\mathbf{ck}} \approx E^0_{\mathbf{ck}} + \Delta$, where Δ is the increase of the band gap from using values E^0_Γ to E^1_Γ.

Since the conduction and valence bands are now unambiguously parabolic, we compute the effective mass. We calculate an effective mass of $0.19m_e$ for the holes and $0.14m_e$ for the electrons. This compares favorably with the experimentally determined effective masses, from magneto-optics, of $0.14m_e$ for electrons and holes, in Ref. [88]. We note that our determination of effective masses agree despite the discrepancy in the value of the band gap.

Finally, we also calculate the band gap at the Γ-point through the use of the full-frequency treatment of the dielectric function, via the contour deformation method[37] and a low rank approximation[12]. We used 15 imaginary frequencies, 200 eigenvectors in a reduced basis scheme, corresponding to roughly 10 percent of the full spectrum, a frequency spacing of 0.25 Ry, with frequencies calculated out to 10 Ry. The dielectric cutoff was again 25 Ry, with 1000 empty bands in the Chi summation, and the Coulomb-hole summation had 1254 empty bands. We found a larger gap than in the Hybertsen-Louie GPP, with a value of 0.408 eV. Usually, redistribution of the weight of the screening from a single frequency gives a lower gap. This increase of the gap relative to GPP is understood as a consequence of the inverted nature of the bandstructure.

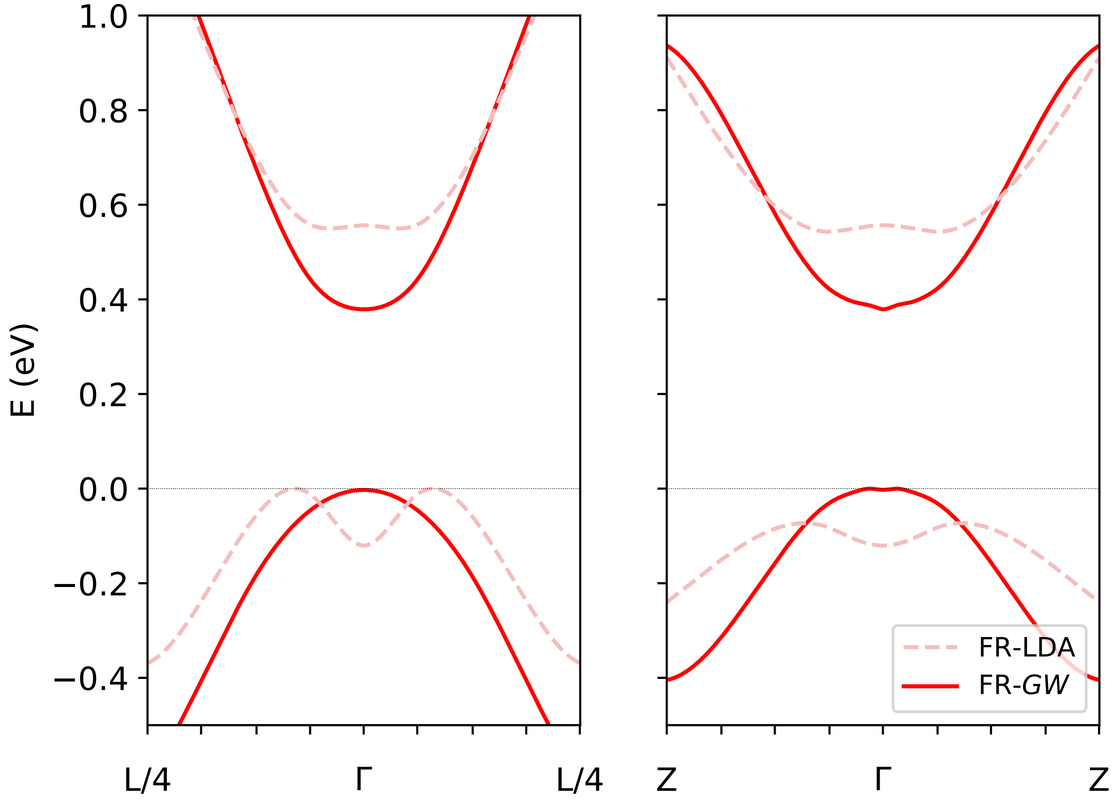

Figure 4.1: The electronic bandstructure of $\mathrm{Bi_2Se_3}$ along the Γ to L and Γ to Z directions, including spin-orbit coupling. The quasiparticle bandstructure, computed from FR-GW, is in solid lines, and the FR-LDA bandstructure is in dashed lines. The bandstructure is strongly camel-backed in the FR-LDA but becomes parabolic in the FR-GW.

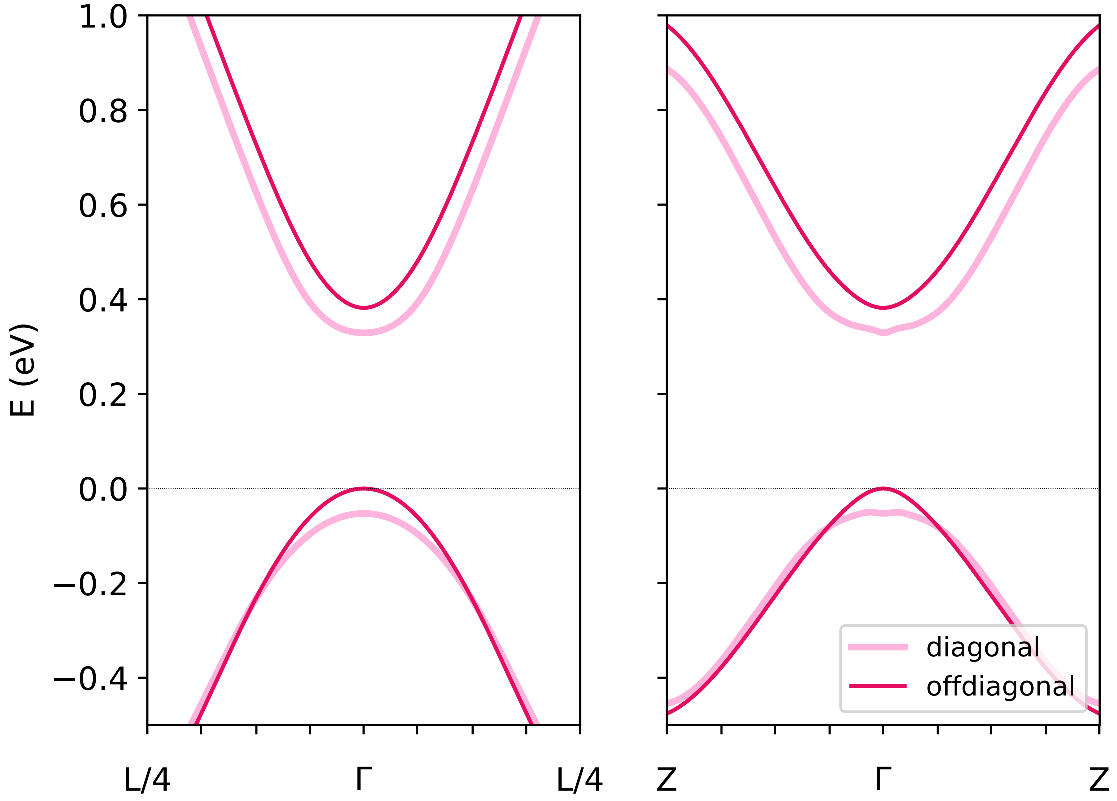

Figure 4.2: The electronic bandstructure of Bi_2Se_3 along the Γ to L and Γ to Z directions, including spin-orbit coupling. The quasiparticle bandstructure computed from FR-GW with off-diagonal entries in the self-energy is in dark magenta, and the quasiparticle bandstructure without off-diagonals is in lighter magenta and arbitrarily shifted downward by 0.05 eV for clarity. Besides the change in band curvature, note that the small, spurious dimples present in the diagonal calculation along the Γ to Z direction are eliminated when including off-diagonal matrix elements in the self-energy.

Chapter 5

The Geometry and Electronic Structure of Iridium adsorbed on Graphene

5.1 Introduction

The sublattice symmetry of graphene results in a Dirac-like low-energy Hamiltonian about the K and K' points in its Brillouin Zone, which gives it the remarkable electronic properties of ultrarelativistic chiral Dirac fermions, such as ballistic transport and the anomalous Integer Quantum Hall Effect[17]. Much experimental and theoretical work has been devoted to the manipulation of graphene to modify its electronic structure in novel ways.

A subset of this work is to investigate how adatoms and other adsorbates modify various electronic properties of graphene. The electronic structure of graphene with a dozen different adatoms were calculated in Ref. [19], finding slight modification to the electronic density of states (DOS) with alkali adatoms and considerable modification to the DOS with transition metal adatoms. In particular, the linear DOS in the vicinity of the Dirac point was largely obscured by the *d*-like atomic states from transition metal adatoms Ti, Fe, Pd, and Au[19].

Theoretical studies of Pd dimers adsorbed onto graphene have shown potential for hydrogen storage, as well. Pairs of hydrogen atoms prefer to dissociate and bond with the Pd dimers instead of forming a hydrogen bond[69]. Similar results have been reported for larger Pd clusters[11].

Further theoretical work[16, 71, 135, 68, 28, 96] has focused on enhancing the weak spin-orbit coupling of graphene by using adatoms to break symmetries, as well as the intrinsic contribution of the atomic spin-orbit coupling from sufficiently heavy metallic adatoms. In particular, calculations with Au[71], W[135], Os[47], Ir[47], and In[126] adatoms suggested relatively large spin-orbit gaps with non-trivial band topologies.

Attempts at realizing these exotic electronic phases with In adatoms[54] and W adatoms[102] have shown null results, however. Initial theoretical work suggested that the adsorption of In onto the hollow site of graphene allows for a fortuitous cancellation of the Rashba spin-orbit coupling[126], thus allowing an induced quantum spin Hall effect. However, enhancement of the Coulomb scattering from the defect rendered any possible non-trivial topology unobservable. Refined first-principles scattering calculations using the quasiparticle band structure with spin-orbit coupling treated non-perturbatively[102] have supported this interpretation.

Additionally, the work with Os and Ir adatoms noted that monomers form spin moments of 0.45 μ_B and 0.30 μ_B, respectively[47], breaking the time-reversal invariance required for the quantum spin Hall effect. Dimerizing the transition metal atoms with Cu was proposed to eliminate magnetic moments in the system and thus preserve time-reversal symmetry[47].

Our experimental colleagues, Aaron J. Bradley and Miguel Ugeda-Moreno in the Michael F. Crommie group at UC Berkeley, used vapor deposition to adsorb Ir atoms onto graphene and analyzed the electronic features through STS measurements. They found a tendency of the Ir atoms to form small clusters with an average size of two atoms per cluster, and a large peak in the LDOS at the Dirac point. Based on these experimental results, we perform density functional calculations to further understand the geometry of the adsorbed Ir atoms as well as their electronic structure. We present evidence that these Ir dimer adsorbates bind in a horizontal configuration along the carbon-carbon bond sites, in contrast with the prediction of vertical dimers binding to the hollow site in a ring of carbon atoms[47]. We find that the $5d$ states from the Ir atoms contribute to the peak in the LDOS.

5.2 Experimental results

We briefly summarize the experimental results by our collaborators. All STM and STS measurements were carried out at T = 7K in an ultra-high vacuum (UHV) environment using a homebuilt scanning tunneling microscope. Graphene was grown on Pt(111), and Ir was deposited on cold samples using an Omicron e-beam evaporator. The maximum temperature of the sample during Ir evaporations ranged between 8K and 12K, with different temperatures leading to qualitatively similar coverages.

Fig. 5.2 (upper left) shows a typical STM image of Ir clusters on bare Pt(111) and graphene / Pt(111). The graphene patch is outlined with a dotted green line. The Ir clusters appear as orange protrusions, with the clusters on the graphene patch appearing significantly larger than those on the bare Pt(111) surface.

Unlike on Pt(111), the size of the clusters on the graphene surface is not homogeneous, likely indicating a varying number of atoms per cluster. STS curves from three different Ir clusters on graphene are shown in Fig. 5.2, upper right. On each of these clusters, there is a feature which occurs at around +0.33 V, which is similar in energy to the Dirac point of graphene on Pt(111)[113, 118], as well as the onset of an unoccupied Pt(111) surface state[128]. The STS curves also exhibit a feature at higher energy, centered at +1.2 V.

A conductance map taken at $V_{bias} = +0.33$ V, the energy of the large STS feature, is shown in Fig. 5.2, lower left. As can be seen, a number of clusters exhibit the spectral feature at +0.33 V.

To better understand the distribution of clusters on the surface, we estimate the atom density (Fig. 5.2, lower right), which plots the number of clusters as a function of the area of a graphene island for a given set of evaporation parameters. The black dashed line is the expected number of clusters if all were indeed single Ir atoms, equal to the density of Ir atoms on the bare Pt(111) surface. The red dashed line is the expected number of clusters if all were dimers. The majority of our data points fall closer to the dimer line, leading us to believe that Ir adatoms are mobile on graphene at temperatures as low as 8K and cluster with a small number of atoms, likely forming dimers and trimers. The mobility of Ir adatoms on graphene may also be affected by the superlattice formed between graphene and the Pt(111).

5.3 Calculation methods

We perform *ab initio* calculations using Quantum ESPRESSO [35]. We construct a 4x4 supercell for the single adatom and vertically-oriented dimer systems, and a 5x5 supercell for the horizontally-oriented dimer, with a vacuum of 15 Å in all cases, as measured from the topmost Ir atom in the system. We use Vanderbilt ultrasoft fully relativistic pseudopotentials[120][24], a 40 Ry planewave kinetic energy cutoff, and a 381 Ry charge density planewave cutoff. We use LDA pseudopotentials for C and Ir, which in the context of graphene adsorbed on a Ir (111) surface gives similar results as compared with van der Waals functionals [10], and is the choice consistent with Ref. [47]. The metallic bonding of Ir to graphene is adequately described in the LDA, despite the systematic overestimation of binding. We use the Methfessel-Paxton smearing, appropriate for metals, of width 0.005 Ry. The Monkhorst-Pack grid for sampling the Brillouin zone was 9x9x1 when constructing the charge density. For details on the DOS calculations, see section 5.5.

We find a theoretical carbon-carbon distance of 1.417 Å, which agrees well with the experimental value of 1.42 Å, by relaxing the primitive unit cell for graphene. We then calculate the binding energy of a single adatom according to $\Delta E = E_{\text{tot}} - E_{\text{gr}} - E_{\text{Ir}}$, where E_{tot} is the energy of the Ir adatom-on-graphene system, E_{gr} is the energy of an isolated graphene supercell, and E_{Ir} is the energy of an isolated Ir atom. For the adsorbed dimer systems , we calculate the binding energy per Ir atom: $\Delta E = \frac{1}{2}(E_{\text{tot}} - E_{\text{gr}} - 2E_{\text{Ir}})$, with the E_{tot} now referring to the Ir dimer-on-graphene system. The value of E_{gr} is scaled appropriately for the size of the supercell under consideration.

We note that we are ignoring the Pt(111) substrate. We expect that the graphene-Pt substrate interaction will not by itself introduce a gap in the graphene Dirac cone[138], with the substrate merely introducing hybridization effects that obscure the usual linear dispersion and also raising the Fermi level. We implicitly take into consideration the dominant physical effect of the substrate, which is to fix the positions of the carbon atoms in the graphene sheet. We therefore do not relax the carbon atoms when determining the relaxed coordinates of the Ir atoms. In comparing our theoretical calculations to experiment, we have different values of Fermi energy due to the neglect of the Pt substrate, so we instead choose to compare theoretical and experimental spectra near the estimated Dirac point energy.

Hopping matrix elements between Ir atoms in neighboring supercells, computed in Wannier90[85], were found to be at most 2 meV, indicating that the wavefunctions are well-contained in a single 4x4 supercell. The small intercellular interaction was confirmed by a 6x6 supercell calculation for the $H^1 - H^2$ system, as the adsorption energy per Ir atom changes by only 7 meV, relative to the 5x5 supercell.

5.4 Theoretical geometry

Single adatom

We relaxed the Ir adatom in the bridge ("B"), hollow ("H"), and top ("T") sites (shown in Fig. 5.1). We find that the Ir atom favors the bridge site, in agreement with previous calculations using PBE and van der Waals functionals[89], in addition to calculations of single-adatom Pd.[19]. Ref [47] finds a preference for the H site using LDA and not including corrections to the dipole interaction with periodic images in the out-of-plane direction. We reproduce this result when we perform the single atom relaxation neglecting the correction to the dipole interaction. Table 5.1

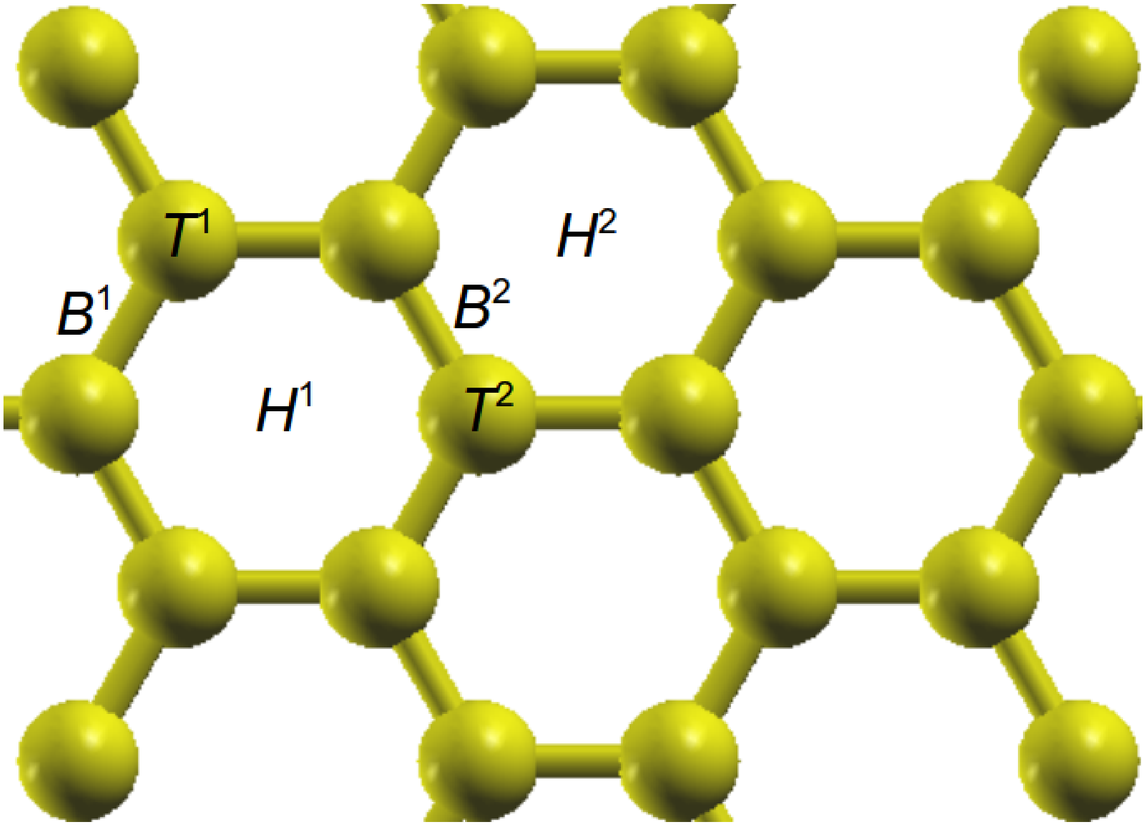

Figure 5.1: The adsorption sites considered. Bridge (B^1, B^2), Hollow (H^1, H^2), and Top (T^1, T^2). In the monomer and vertical dimer cases, only B^1, H^1, and T^1 are considered and are referred to as B, H, and T, respectively.

summarizes the binding heights h, adsorption energies ΔE, diffusion barries E_{Diff}, magnetic moment per supercell $\mathbf{m}$, and effective charge Q for the single Ir adsorption. The estimate of the effective charge indicates a donation of electrons from Ir to the graphene sheet, while the magnetic moment per supercell of nearly 1 μ_{B} is consistent with an effective charge of five, indicating an uncompensated additional spin at the Ir atom.

We approximate the diffusion barrier across a high-symmetry point as $E_{\mathrm{Diff}} = E_{\mathrm{a}}^{\max} - \Delta E$, the difference in energy from the lowest-energy geometry to the energy of the site under consideration[19]. The most appropriate approximation of the diffusion path would be one that is calculated by the nudged elastic band method to explore the full energy surface for all possible directions. However, we assume that

Config.	h (Å)	ΔE (eV)	E_{Diff} (eV)	$\mathbf{m}(\mu_{\mathrm{B}})$	$Q(e)$
Bridge	1.95	1.86	–	0.95	4.88
Hollow	1.70	1.13	0.73	0.61	5.32
Top	1.93	1.37	0.49	0.99	4.55

Table 5.1: For the Hollow, Bridge, and Top configurations, we calculate the heights h, the binding energy ΔE of the single Ir atom adsorbed to graphene in the LDA, the diffusion barrier, the magnetic moment in the supercell, and the estimation of the local charge of the Ir atom.

diffusion is likely to occur nearly along high-symmetry directions, and approximate the barrier to diffusion along a direction simply to be the difference in energy of the initial and final configurations. The estimated diffusion barriers of 0.73 eV and 0.49 eV for the top and hollow sites, respectively, are well-above room temperature and are on the order of that of Fe adsorbed on graphene[19].

We can qualitatively determine whether adatoms on graphene tend to form single adatoms or clusters by comparing the binding energy per atom ΔE to the bulk cohesive energy per atom E_{c}. A ratio of $\Delta E/E_{\mathrm{c}} \sim 1$ indicates a strong adsorption energy of the adatom and hence a tendency to bind to graphene as an adatom; the opposite limit indicates a tendency to cluster. However, the LDA overestimates cohesive energies for solids, particularly $5d$ transition metals[125], so we instead use the experimental value of $E_{\mathrm{c}} = 6.94$ eV[125]. With the adsorption energy at the bridge site $\Delta E = 1.86$ eV, we estimate the ratio to be 0.268. Though this ratio may still be overestimated due to the use of the LDA for the calculation of the binding energy, we compare this result to the value of 0.307 reported for Al[19], which is also known experimentally to form clusters with three atoms when adsorbed to graphite[34][81]. Based on this, we would propose that Ir tends to cluster on graphene, which, as shown in Section 5.2 is seen in experiment.

Dimer

The horizontal dimer at the bridge sites B^1 and B^2 (see Fig. 5.1) has the largest binding energy for any of the considered dimer configurations, so it is the most likely orientation for the observed two-atom clusters at low temperature. The bonding distances for the horizontal dimers are longer than in isolation, 2.23 Å[55], and are on the order of the C-C bond length (2.46 Å). Indeed, we see that the bridge-site dimer has the shortest Ir-Ir distance, explaining its favorable adsorption energy, as it is stretched from its free value the least.

Config.	$h_{1=2}$ (Å)	$d_{\text{Ir-Ir}}$ (Å)	ΔE (eV)	E_{Diff} (eV)	$\mu_x(\mu_{\text{B}})$	$\mu_y(\mu_{\text{B}})$	$\mu_z(\mu_{\text{B}})$	$Q_{1=2}(e)$
B^1 – B^2	2.05	2.32	3.36	–	0.00	1.06	1.29	4.99
H^1 – H^2	1.88	2.46	2.85	1.02	0.00	0.00	1.32	5.68
T^1 – T^2	2.10	2.45	3.19	0.34	0.23	0.23	1.28	5.04

Table 5.2: For the Hollow, Bridge, and Top configurations in which the Ir atoms are placed at the B/H/T 1 and B/H/T 2 sites, we calculate the relaxed z-coordinates of each Ir atom, the adsorption energy ΔE of the horizontally-oriented Ir dimer adsorbed to graphene in the LDA, the estimate of the diffusion energy, the cartesian components of the magnetic moment in the supercell, and the local charge for each Ir atom.

The estimated diffusion barrier is computed for the dimer as a whole, defined from $E_{\text{Diff}} = 2\,(E_{\text{a}}^{\text{max}} - \Delta E)$. The hollow site has an estimated diffusion barrier of 1.02 eV, indicates a strong preference to diffuse through the top site, with an estimated diffusion barrier of 0.34 eV, which may allow a pathway to form larger clusters. The estimated charge per Ir atom, 4.99, is slightly larger than its single-adatom value, 4.88, indicating that less charge is transferred to the graphene sheet due to the presence of the dimer bond. The magnetic moment per unit-cell is consequently increased to over 1 μ_{B} per Ir atom. The magnetic moment now has a large in-plane component in addition to the out-of-plane component, due to the in-plane dimer bond defining an additional axis for the magnetic moment.

The binding energies per Ir atom in the vertical configuration are nearly identical for the B, H, and T positions, with 3.11 eV for B and T, and 3.14 eV for H, giving small diffusion barriers of 0.06 eV, comparable to room temperature. The dimer bond lengths are nearly the same as the isolated dimer bonding length of 2.17 Å. The distances from the bottom-most Ir atom in the dimer from the graphene sheet (h_b in Table 5.3) are larger than the single Ir adatom case.

Config.	h_b (Å)	$d_{\text{Ir-Ir}}$ (Å)	ΔE (eV)	E_{Diff} (eV)	$\mathbf{m}_b(\mu_{\text{B}})$	$\mathbf{m}_t(\mu_{\text{B}})$	$\mathbf{m}(\mu_{\text{B}})$	$Q_b(e)$	$Q_t(e)$
Bridge	2.41	2.19	3.11	0.06	0.58	0.67	2.07	5.08	5.11
Hollow	2.29	2.19	3.14	–	0.54	0.77	2.00	5.03	5.13
Top	2.40	2.18	3.11	0.06	0.58	0.68	2.05	5.05	5.08

Table 5.3: For the Hollow, Bridge, and Top configurations, we calculate the relaxed z-coordinates, the adsorption energy ΔE of the vertically-oriented Ir dimer adsorbed to graphene in the LDA, the estimate of the diffusion energy, the (out-of-plane) magnetic moments for each Ir atom, the total magnetic moment in the supercell, and the local charge for each Ir atom.

5.5 Electronic Structure

Single adatom

We calculate the DOS of pristine graphene in a primitive unit cell, with a 96x96x1 Monkhorst-Pack grid, which is dense enough to resolve the zero DOS at the Dirac point energy. We then calculate the DOS for the single adatom atop a 4x4 supercell of graphene with a 24x24x1 grid, giving an overall sampling density equivalent to that of pristine graphene. We use the peaks in the DOS from the weakly-hybridized, low-lying sp^2 states to match with that of pristine graphene (Fig. 5.3) and thereby provide an estimate to the Dirac point.

For comparison with the experimental STS spectra, we calculate the LDOS by integrating the charge density 4 Å above the Ir atoms. We use Gaussian functions with a smearing of 75 meV as an approximation to the delta function in energy. The actual tip to atom distance is estimated at 5 to 10 AA.

The LDOS (Fig. 5.4) shows a significant peak placed about 1 eV above the Dirac point energy, with a much weaker peak at the Dirac point. This indicates that the experimentally observed peaks (Fig. 5.2) do not come from single adatoms. By comparing the LDOS with the partial density of states resolving the *s*, *p*, and *d* Ir states (Fig. 5.5), we see that the Ir *s*- and *p*-states give the most significant contribution to the large peak, with the *d*-states being localized below the 4 Å simulated tip placement.

Dimer

We similarly compute the DOS, LDOS, and PDOS for the adsorbed dimer system. Again, we see in the DOS (Fig. 5.6) that the low-lying sp^2 states from graphene are only weakly hybridized with the Ir states and therefore allow us to estimate the Dirac point energy.

The LDOS places a large peak near the estimated Dirac point energy (Fig. 5.7), as observed in experiment. We are therefore confident that the structure of the Ir clusters are indeed a pair of bridge-site Ir adatoms in the horizontal configuration. Comparing the LDOS with the PDOS (Fig. 5.8), we see that the features of the LDOS appear to arise from both the *s*- and *p*-states. (Though the *d*-states contribute the most states, they do not extend out appreciably to 4 Å.)

Fig. 5.9 directly compares the measured $\mathrm{d}I/\mathrm{d}V$ spectrum with the calculated LDOS, plotted in arbitrary units. We see good agreement with the placement of the central peak, which we understand to be the Ir *s*- and *p*-states. The higher-energy feature also seems to be reproduced, though within a significantly red-shifted energy range, likely due to DFT underestimating such excited-state features. The peaks below the estimated Dirac point energy that are present in the calculated LDOS are suppressed in the measured spectrum. Otherwise, we generally have a satisfactory agreement between experiment and theory.

5.6 Conclusion

We find that Ir atoms deposited on graphene on Pt(111) tend to form clusters, even at low temperatures. By calcuting the areal density of the observed clusters on the graphene flake, we estimate the typical size of the clusters to be composed of two Ir atoms. Through *ab initio* density functional calculations, we predict that the Ir dimer is oriented horizontally, at the "Bridge" sites of the graphene lattice. We find peaks in the experimental and theoretical LDOS near the estimated energies of the Dirac points and attribute them to the Ir dimer *s*- and *p*-states. We further computed the LDOS for the single adatom and found it to have significantly different peak locations and amplitudes than that of the clusters. In either the dimer or single adatom case, each contributes nearly 1 μ_B per Ir atom per supercell, though the spin density is highly localized around the Ir atoms.

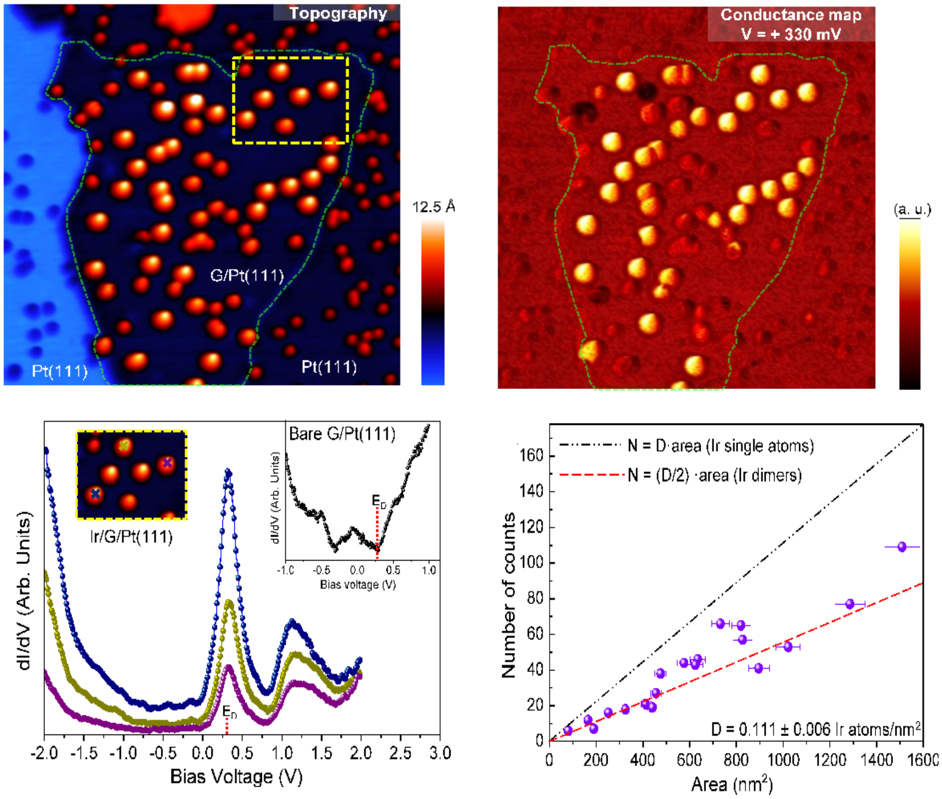

Figure 5.2: Top left: The STM topography of Ir clusters on bare Pt(111) and graphene on Pt(111). Top right: The conductance map of the same region as the top light. Bottom left: The conductance maps for the indicated Ir clusters in the inset. Bottom right: The number of clusters as a function of the area of the graphene island. Measurements and figures by Aaron J. Bradley, Miguel Moreno Ugeda, and Michael F. Crommie.

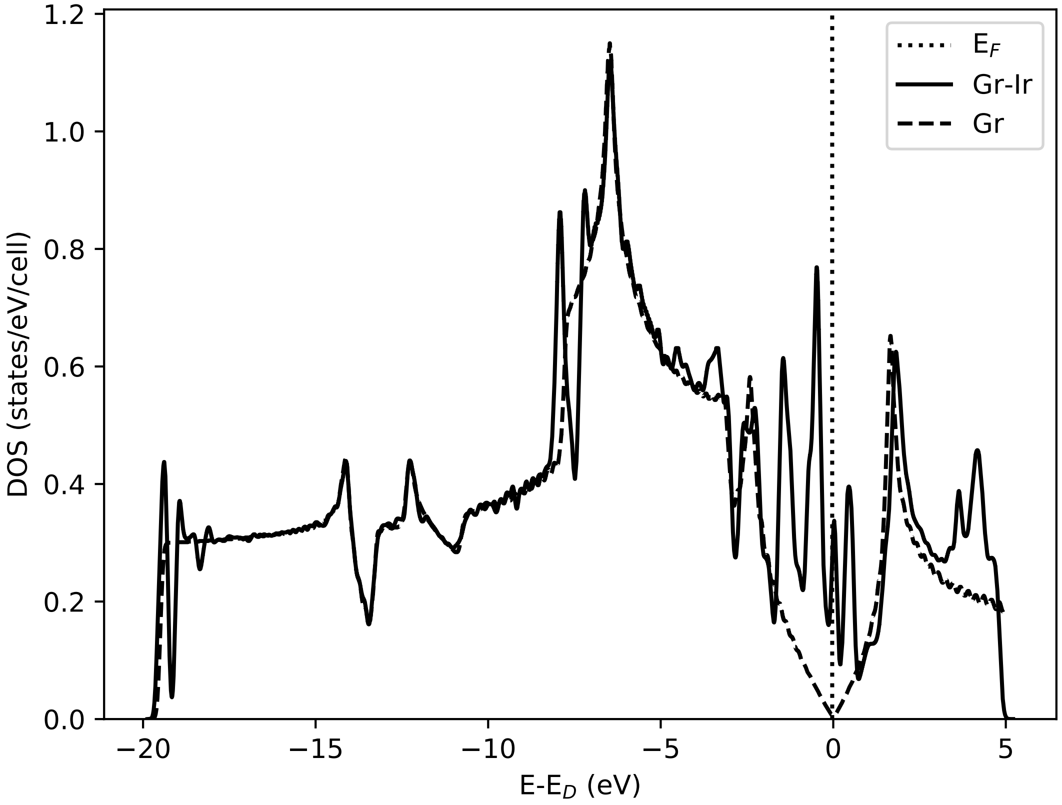

Figure 5.3: The comparison between the DOS for pristine graphene and for the single Ir adatom on graphene. The peaks from about -10 to -15 eV are matched to provide an estimate for the Dirac point energy.

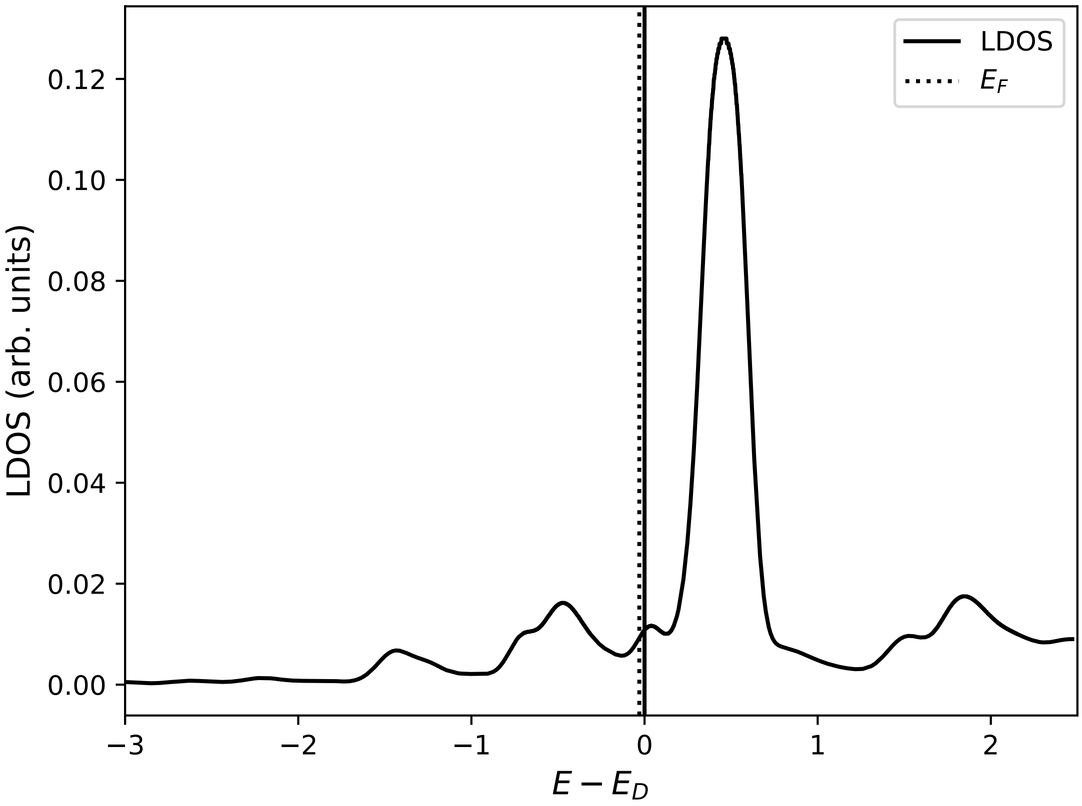

Figure 5.4: The LDOS for the single Ir adatom on graphene, measured 4 Å from the Ir atom.

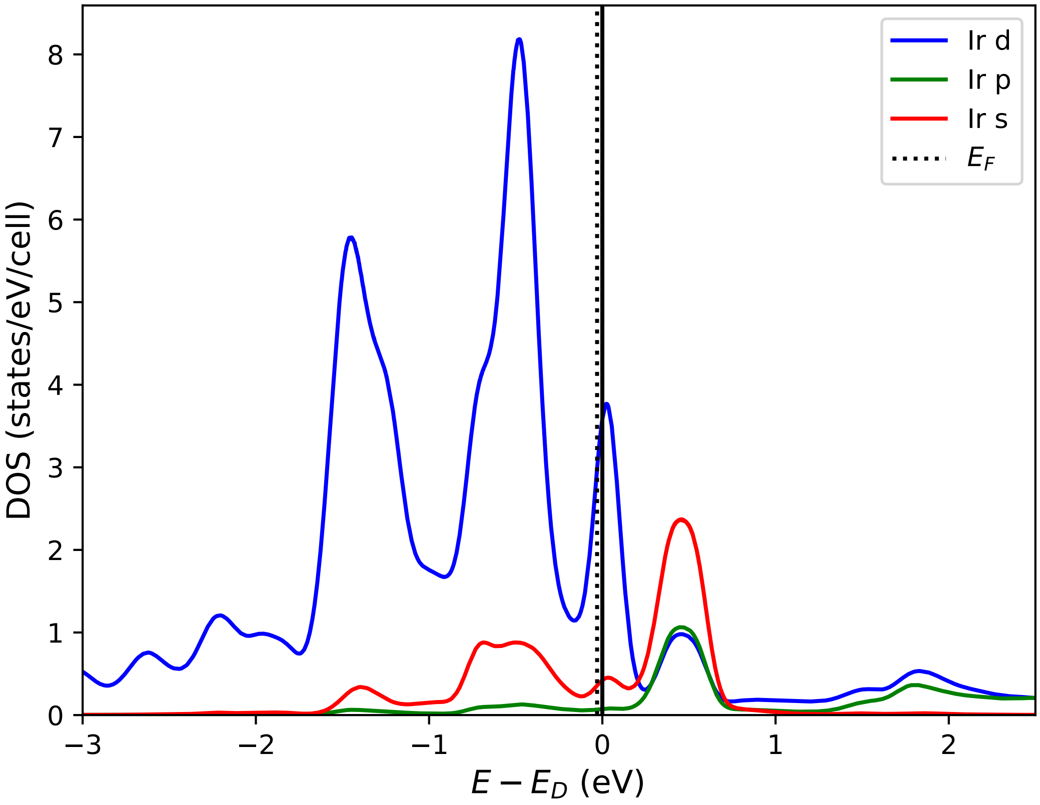

Figure 5.5: The PDOS for the single Ir adatom on graphene. We see large peaks from the localized Ir *d*-states, and the large peaks from the *s*- and *p*-states above the estimated Dirac point energy.

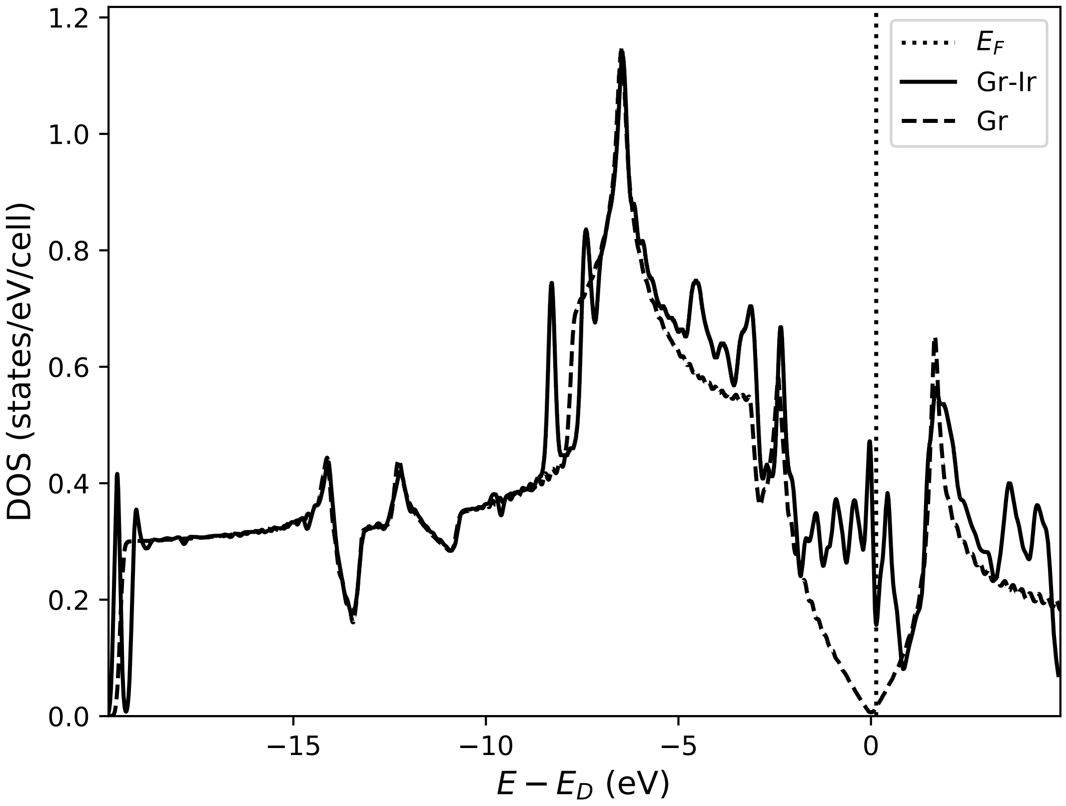

Figure 5.6: The comparison of the DOS for a pristine sheet of graphene and an Ir dimer adsorbed on graphene, in a 5x5 supercell. The peaks from -10 to -15 eV are matched to determine an estimate for the Dirac point energy.

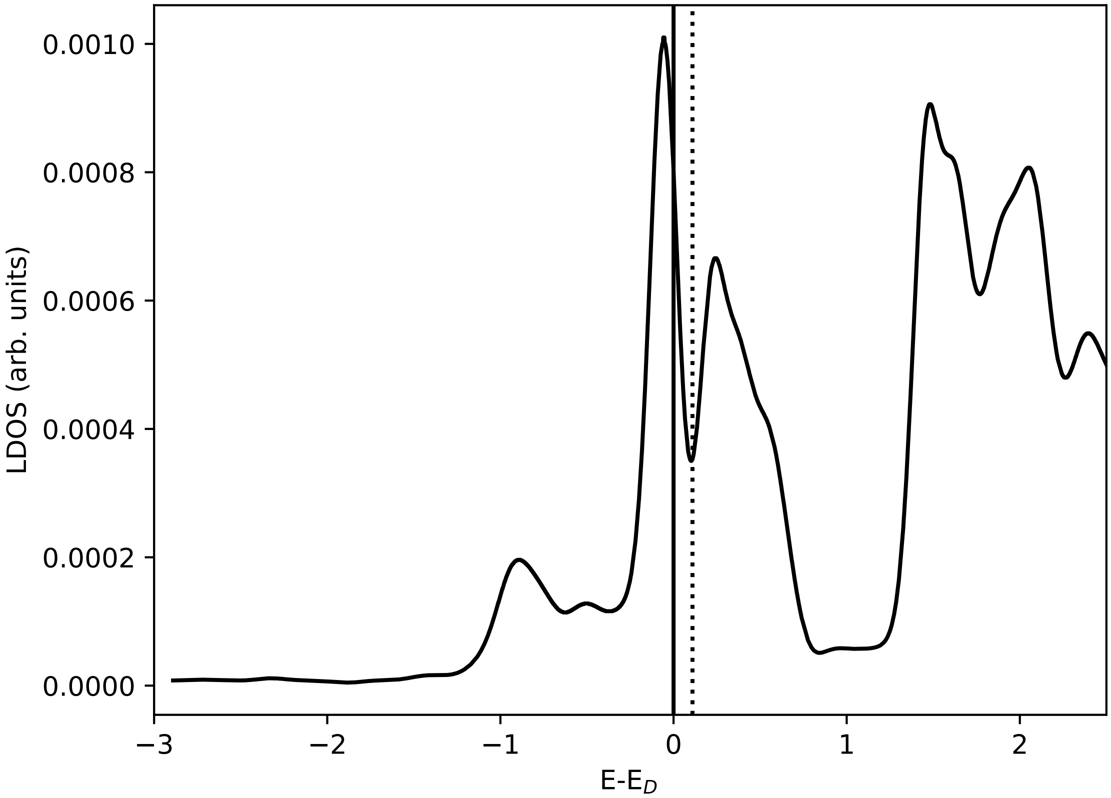

Figure 5.7: The LDOS for the Ir dimer on graphene, with the zero of energy as the estimated Dirac point energy. The states are 4 Å above the Ir atoms.

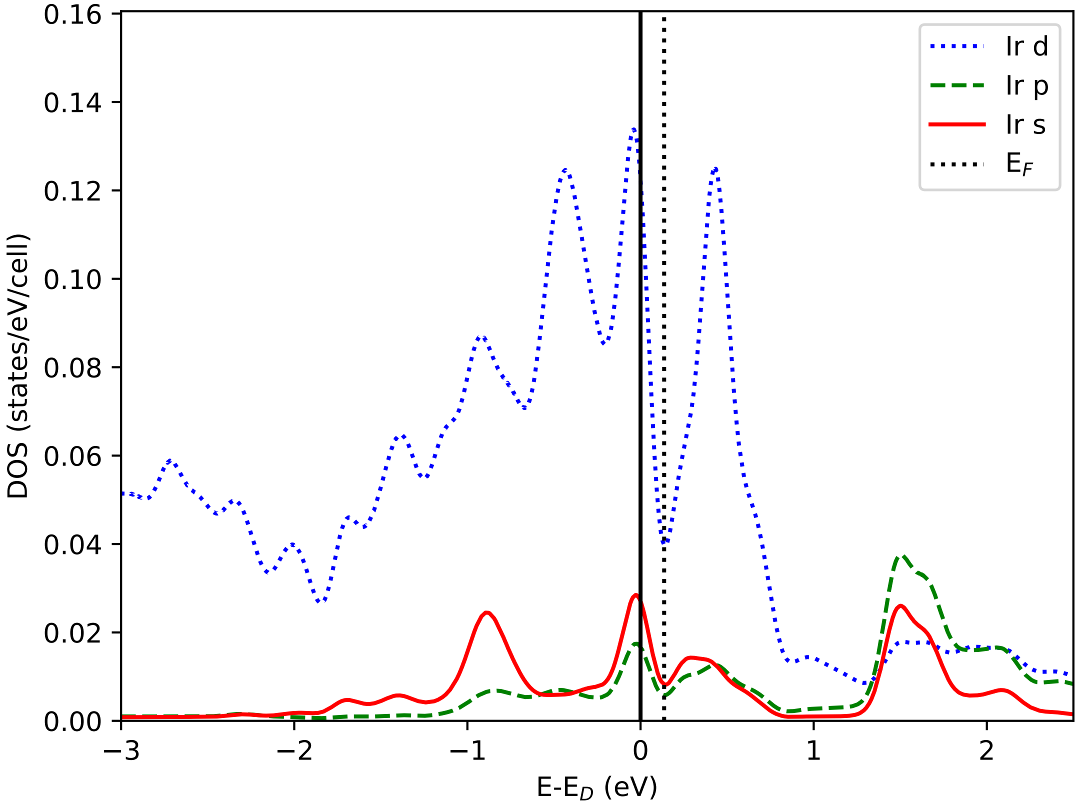

Figure 5.8: The PDOS for the Ir dimer on graphene, with the zero of energy as the estimated Dirac point energy.

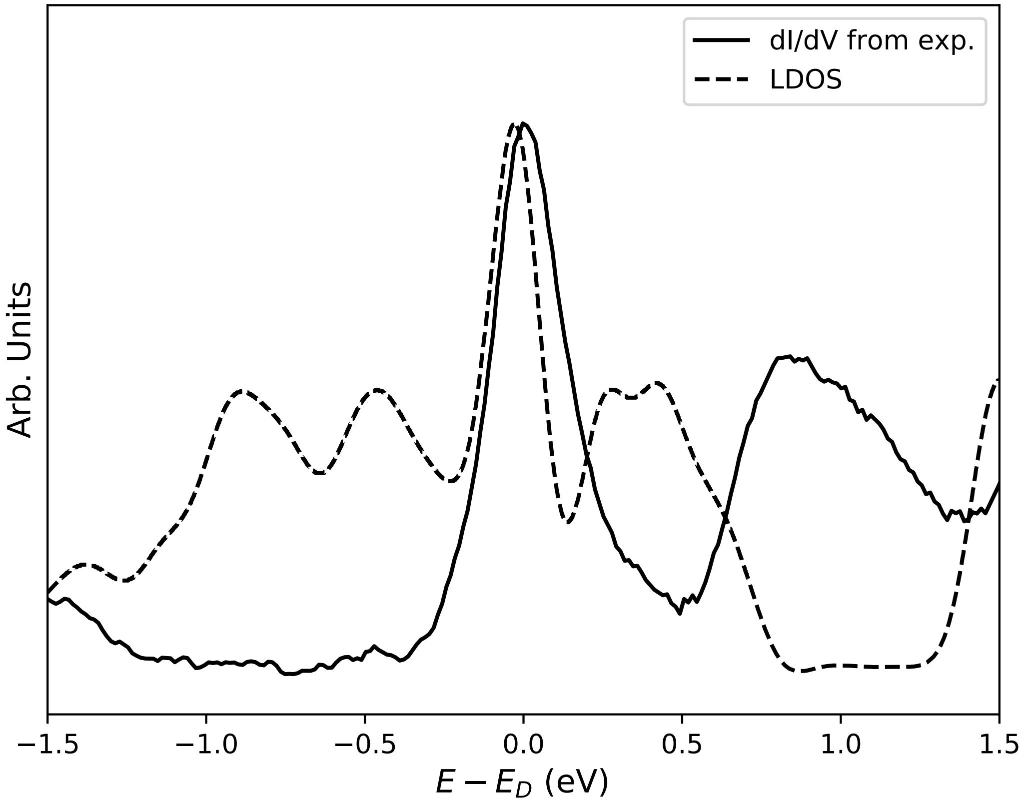

Figure 5.9: Experimentally determined dI/dV for an Iridium cluster on graphene on a Pt(111) substrate (left). Calculated LDOS, 4 Å above a horizontally-oriented Ir dimer across the graphene *B* sites. The energy is plotted relative to the estimated energy of the Dirac point.

Appendix A

Variational Derivative of G

In Many-Body Perturbation Theory, we arrive at the need to relate the two-particle Green's Function to the one-particle Green's Function. This can be done through Feynman diagrams or the Schwinger approach; we choose the latter option, which simply requires taking a few derivatives.

In the Schwinger approach, we introduce a perturbing potential, ϕ, to the system. We take this potential to be spatially non-local so that we can use the final result in both the derivation of the Hedin equations as well as the Bethe-Salpeter equation. The potential also is entirely electrostatic in nature. It thus introduces the perturbing term in our Hamiltonian,

$$H^1(t) = \int \mathrm{d}2\mathrm{d}3\ \hat{\psi}^\dagger_\gamma(2^+)\phi(23)\delta(t-t_2)\delta(t-t_3)\delta_{\gamma\zeta}\hat{\psi}_\zeta(3). \tag{A.1}$$

We now need to take the Green's Function to the interaction picture picture, where the ground state is now $|\Phi_0^N\rangle$, by introducing the time development operator

$$\hat{S} = \exp\left\{-i\left(\int_{-\infty}^{+\infty} \mathrm{d}t\ H^1(t)\right)\right\}. \tag{A.2}$$

The Green's Function now is

$$G_{\alpha\beta}(11') = -i\ \frac{\langle\Phi_0^N|T\left[\hat{S}\hat{\psi}_\alpha(1)\hat{\psi}^\dagger_\beta(1')\right]|\Phi_0^N\rangle}{\langle\Phi_0^N|T\left[\hat{S}\right]|\Phi_0^N\rangle}. \tag{A.3}$$

A variation in the perturbing potential, $\delta\phi$, will create a variation in the Green's Function

$$\begin{aligned}\delta G_{\alpha\beta}(11') = (-i)^2 \int \mathrm{d}23\; \delta\phi(23)\; &\frac{\langle \Phi_0^N | T\left[\hat{S}\hat{\psi}_\gamma^\dagger(2^+)\hat{\psi}_\gamma(3)\hat{\psi}_\alpha(1)\hat{\psi}_\beta^\dagger(1')\right]|\Phi_0^N\rangle}{\langle \Phi_0^N | T\left[\hat{S}\right]|\Phi_0^N\rangle} \\ - (-i)G_{\alpha\beta}(11')\int \mathrm{d}23\; \delta\phi(23)\; &\frac{\langle \Phi_0^N | T\left[\hat{S}\hat{\psi}_\gamma^\dagger(2^+)\hat{\psi}_\gamma(3)\right]|\Phi_0^N\rangle}{\langle \Phi_0^N | T\left[\hat{S}\right]|\Phi_0^N\rangle}. \end{aligned} \tag{A.4}$$

Recognizing that the first term on the right-hand side in Eq. A.4 is a two-particle Green's Function $G^{(2)}$, apart from the multiplication by the $\delta\phi$ and the integration over its coordinates, we have the variational derivative

$$\frac{\delta G_{\alpha\beta}(11')}{\delta\phi(23)} = -G^{(2)}_{\alpha\gamma\beta\gamma}(13, 1'2^+) + G_{\alpha\beta}(11')G_{\gamma\gamma}(32^+). \tag{A.5}$$

In deriving the self-energy, we introduce a local perturbing potential $\phi(2)$, so we just enforce $\phi(2) = \phi(23)\delta(23)$, and arrive at

$$\frac{\delta G_{\alpha\beta}(11')}{\delta\phi(2)} = -G^{(2)}_{\alpha\gamma\beta\gamma}(12, 1'2^+) + G_{\alpha\beta}(11')G_{\gamma\gamma}(22^+). \tag{A.6}$$

By recognizing $G_{\gamma\gamma}(22^+)$ as the density we see clearly that the variational derivative with respect to the local perturbing potential is some additional term beyond the Hartree approximation. We note that in the context of non-collinear spins, this result was first derived in Ref. [4], though now generalized to include non-local perturbations. (We note that time-reversal symmetry-breaking magnetic perturbations can be included in the previous derivation; this is done in [56].

We may eliminate ϕ altogether (thereby implicitly taking the $\phi \to 0$ limit) in Eq. A.6 by using the total potential $V = \phi + v^{\mathrm{H}}$, Eqs. 1.36, 1.4, and 1.37:

$$\begin{aligned}\frac{\delta G_{\alpha\beta}(11')}{\delta\phi(2)} &= -\int \mathrm{d}34\; G_{\alpha\eta}(13)\frac{\delta G^{-1}_{\eta\kappa}(34)}{\delta\phi(2)}G_{\kappa\beta}(41') \\ &= -\int \mathrm{d}345\; G_{\alpha\eta}(13)\frac{\delta V(5)}{\delta\phi(2)}\frac{\delta G^{-1}_{\eta\kappa}(34)}{\delta V(5)}G_{\kappa\beta}(41') \\ &= -\int \mathrm{d}345\; G_{\alpha\eta}(13)\epsilon^{-1}(25)\Gamma_{\eta\kappa}(34;5)G_{\kappa\beta}(41'). \end{aligned} \tag{A.7}$$

This result is critical in the derivation of the Hedin-Lundqvist equations, where we may now make the replacement

$$
\begin{aligned}
&\int \mathrm{d}2\; v(12) G^{(2)}_{\alpha\gamma\beta\gamma}(12, 1'2) \\
&= i \int \mathrm{d}2\; v(12)\rho(2) G_{\alpha\beta}(11') + \int \mathrm{d}2345\; v(12)\epsilon^{-1}(25) G_{\alpha\eta}(13)\Gamma_{\eta\kappa}(34;5) G_{\kappa\beta}(41') \\
&= i v^{\mathrm{H}}(1) G_{\alpha\beta}(11') + \int \mathrm{d}345\; G_{\alpha\eta}(13) W(15)\Gamma_{\eta\kappa}(34;5) G_{\kappa\beta}(41'). \qquad \text{(A.8)}
\end{aligned}
$$

Appendix B

Excitons with spin-orbit coupling as a perturbation

The considerable increase in cost when using spinor wavefunctions in the Kernel and Absorption executables in the BerkeleyGW software motivates us to determine the inclusion of spin-orbit coupling in the excitonic states as a perturbation. This has already been applied to the system $\mathrm{MoS_2}$ to tremendous success in Ref. [97], where the derivation is first outlined. We consider the derivation in more detail and in a more general context in the following.

We start with a set of N spinless orbitals $\phi_{n\mathbf{k}}(\mathbf{r})$ and affix a two-component spinor χ. As the full wavefunction must now be orthogonal to the others, we must use both χ_1 and χ_2 for any of the original orbitals. The states $\phi_{n\mathbf{k}}(\mathbf{r})\chi_1$ and $\phi_{n\mathbf{k}}(\mathbf{r})\chi_2$ are now considered to be two distinct states belonging to a set of $2N$ spin-orbitals. (The pair χ_1 and χ_2 are any orthogonal pair of spinors, e.g., the up and down spinors.) We use the convention in which each band index for the set of the $2N$ orbitals is unique and the spin indices α, β, etc., refer to the α'th component of a single two-component spin-orbital. With this convention, the aforementioned spin-orbitals would be denoted $\phi_{m\mathbf{k}\alpha}(\mathbf{r})$ and $\phi_{p\mathbf{k}\beta}(\mathbf{r})$, and the energies $E_{m\mathbf{k}}$ and $E_{p\mathbf{k}}$ would be degenerate and equal to the previous $E_{n\mathbf{k}}$.

In second quantized form, the single-particle electronic state $\phi_{m\mathbf{k}\alpha_e}(\mathbf{r}_e)$ is generated from

$$\phi_{m\mathbf{k}}(\mathbf{r})\chi_{\alpha_e} = \phi_{m\mathbf{k}\alpha_e}(\mathbf{r}_e) = \langle \mathbf{r}_e | b^{\dagger}_{m\mathbf{k}\alpha_e} | N, 0\rangle, \tag{B.1}$$

with $|N, 0\rangle$ being the many-body ground state.

For completeness, we define hole states from

$$\phi^{*}_{m\mathbf{k}}(\mathbf{r})\chi_{\alpha_h} = \phi^{*}_{m\mathbf{k}\alpha_h}(\mathbf{r}_h) = \langle \mathbf{r}_h | a^{\dagger}_{m\mathbf{k}\alpha_h} | N, 0\rangle. \tag{B.2}$$

We treat the hole's spin to be a coordinate independent from the electron, although the two can technically be related through the application of the time-reversal symmetry.

We assume that we have already used the set of N spinless orbitals to construct the BSE Hamiltonian and have found the $N_S = N_v \times N_c \times N_{\mathbf{k}}$ eigenvectors $A^S_{vc\mathbf{k}}$ that diagonalize it. (The N_v and N_c states are the valence and conduction subsets of the original N spinless orbitals, respectively. Afterward, however, we consider all states to be in the doubled, spin-degenerate set.) We now consider the excitonic wavefunction constructed from single-particle spin-orbitals.

In the Tamm-Dancoff approximation, the wavefunction for an exciton in state S is written as

$$|S\rangle \chi_{\alpha_e}\chi_{\alpha_h} = \sum_{vc\mathbf{k}} A^S_{vc\mathbf{k}} a^\dagger_{v\mathbf{k}\alpha_h} b^\dagger_{c\mathbf{k}\alpha_e} |N,0\rangle. \tag{B.3}$$

In real-space, the wavefunction is

$$\Psi^S_{\alpha_e\alpha_h}(\mathbf{r}_e,\mathbf{r}_h) = \sum_{vc\mathbf{k}} A^S_{vc\mathbf{k}} \phi^*_{v\mathbf{k}\alpha_h}(\mathbf{r}_h)\phi_{c\mathbf{k}\alpha_e}(\mathbf{r}_e). \tag{B.4}$$

Note that the wavefunction has the same value throughout the (four-dimensional) $\alpha_e \bigotimes \alpha_h$ space. The spinor is decoupled from the orbital coordinates, so we may recreate the usual singlet and triplet states from the irreducible representations of the total spin. Indeed, so far we are only re-enacting the process in Rohlfing-Louie[99], without the added benefit of simplifying the form of the BSE Kernel. However, we wish to preserve the basis of the product of single-particle spinors to consider matrix elements of the spin-orbit coupling Hamiltonian.

The spin-orbit coupling Hamiltonian acts on a single-particle state locally, so when we operate on the excitonic states, we consider

$$H^{\mathrm{SOC}}_{\alpha_e\beta_e;\,\alpha_h\beta_h}(\mathbf{r}_e,\mathbf{r}_h) = \sum_{I\neq J=(e,h)} H^{\mathrm{SOC}}_{\alpha_I\beta_I}(\mathbf{r}_I)\delta_{\alpha_J\beta_J}. \tag{B.5}$$

In second quantized form, we operate with

$$\begin{aligned} \hat{H}^{\mathrm{SOC}}_{\gamma_e\kappa_e;\gamma_h\kappa_h} &= \sum_{n\mathbf{q};n'\mathbf{q}'} H^{\mathrm{SOC}}_{n\mathbf{q};n'\mathbf{q}'} \left(\delta_{\gamma_e\kappa_e} a_{n\mathbf{q}\gamma_h} a^\dagger_{n'\mathbf{q}'\kappa_h} + \delta_{\gamma_h\kappa_h} b^\dagger_{n\mathbf{q}\gamma_e} b_{n'\mathbf{q}'\kappa_e}\right) \\ H^{\mathrm{SOC}}_{n\mathbf{q};n'\mathbf{q}'} &= \int \mathrm{d}^3\mathbf{r}\,\phi^*_{n'\mathbf{q}'\alpha}(\mathbf{r}) H^{\mathrm{SOC}}_{\alpha\beta}(\mathbf{r})\phi_{n\mathbf{q}\beta}(\mathbf{r}). \end{aligned} \tag{B.6}$$

In the basis of excitonic states, the matrix elements of $\hat{H}^{\mathrm{SOC}}$ are

$$\langle S'|\chi^\dagger_{\alpha_e}\chi^\dagger_{\alpha_h}\hat{H}^{\mathrm{SOC}}_{\gamma_e\kappa_e;\,\gamma_h\kappa_h}\chi_{\beta_h}\chi_{\beta_e}|S\rangle. \tag{B.7}$$

Expanding the excitonic states of an arbitrary matrix element gives

$$\begin{aligned}\langle S'|H^{\mathrm{SOC}}|S\rangle = &\sum_{vc\mathbf{k};v'c'\mathbf{k}'} {A^{S'}}^{*}_{v'c'\mathbf{k}'} A^{S}_{vc\mathbf{k}} H^{\mathrm{SOC}}_{n\mathbf{q};n'\mathbf{q}'} \langle 0| b_{c'\mathbf{k}'\alpha_e} a_{v'\mathbf{k}'\alpha_h} b^{\dagger}_{n\mathbf{q}\gamma_e} b_{n'\mathbf{q}'\kappa_e} a^{\dagger}_{v\mathbf{k}\beta_h} b^{\dagger}_{c\mathbf{k}\beta_e} |0\rangle \delta_{\gamma_h\kappa_h} \\ &+ \sum_{vc\mathbf{k};v'c'\mathbf{k}'} {A^{S'}}^{*}_{v'c'\mathbf{k}'} A^{S}_{vc\mathbf{k}} H^{\mathrm{SOC}}_{n\mathbf{q};n'\mathbf{q}'} \langle 0| b_{c'\mathbf{k}'\alpha_e} a_{v'\mathbf{k}'\alpha_h} a_{n\mathbf{q}\gamma_h} a^{\dagger}_{n'\mathbf{q}'\kappa_h} a^{\dagger}_{v\mathbf{k}\beta_h} b^{\dagger}_{c\mathbf{k}\beta_e} |0\rangle \delta_{\gamma_e\kappa_e}.\end{aligned} \tag{B.8}$$

By inspection, we have

$$\langle S'|H^{\mathrm{SOC}}|S\rangle = \left(\sum_{vc c'\mathbf{k}} {A^{S'}}^{*}_{vc'\mathbf{k}} A^{S}_{vc\mathbf{k}} H^{\mathrm{SOC}}_{c'\mathbf{k};c\mathbf{k}}\right) - \left(\sum_{vv'c\mathbf{k}} {A^{S'}}^{*}_{v'c\mathbf{k}} A^{S}_{vc\mathbf{k}} H^{\mathrm{SOC}}_{v\mathbf{k};v'\mathbf{k}}\right). \tag{B.9}$$

This expression reduces to that used in Ref. [97] when considering only diagonal matrix elements ($S = S'$) and additionally imposing $v' = v$ and $c' = c$:

$$\langle S|H^{\mathrm{SOC}}|S\rangle = \sum_{vc\mathbf{k}} \left|A^{S}_{vc\mathbf{k}}\right|^2 \left(H^{\mathrm{SOC}}_{c\mathbf{k};c\mathbf{k}} - H^{\mathrm{SOC}}_{v\mathbf{k};v\mathbf{k}}\right). \tag{B.10}$$

Thus the matrix may be constructed and subsequently diagonalized as in 1.2, once one solves for the excitonic eigenstates from the BSE.

Bibliography

[1] Stephen L. Adler. “Quantum Theory of the Dielectric Constant in Real Solids”. In: *Phys. Rev.* 126 (2 1962), pp. 413–420.

[2] Irene Aguilera, Christoph Friedrich, and Stefan Blügel. “Spin-orbit coupling in quasiparticle studies of topological insulators”. In: *Phys. Rev. B* 88 (16 2013), p. 165136.

[3] Yoichi Ando. “Topological insulator materials”. In: *Journal of the Physical Society of Japan* 82.10 (2013), p. 102001.

[4] F. Aryasetiawan and S. Biermann. “Generalized Hedin’s Equations for Quantum Many-Body Systems with Spin-Dependent Interactions”. In: *Phys. Rev. Lett.* 100 (11 2008), p. 116402.

[5] A. Baalmann et al. In: *Ann. Israeli Phys. Soc.* 6 (1983), p. 351.

[6] Giovanni B. Bachelet and Niels E. Christensen. “Relativistic and core-relaxation effects on the energy bands of gallium arsenide and germanium”. In: *Phys. Rev. B* 31 (2 1985), pp. 879–887.

[7] B Andrei Bernevig, Taylor L Hughes, and Shou-Cheng Zhang. “Quantum spin Hall effect and topological phase transition in HgTe quantum wells”. In: *Science* 314.5806 (2006), pp. 1757–1761.

[8] S Biering and P Schwerdtfeger. “A comparative density functional study of the high-pressure phases of solid ZnX, CdX, and HgX (X= S, Se, and Te): Trends and relativistic effects”. In: *The Journal of chemical physics* 137.3 (2012), p. 034705.

[9] J Black et al. “Electrical and optical properties of some M2v- bN3vi- b semiconductors”. In: *Journal of Physics and Chemistry of Solids* 2.3 (1957), pp. 240–251.

[10] R. Brako et al. “Graphene on the Ir(111) surface: from van der Waals to strong bonding”. In: *New Journal of Physics* 12 (2010), p. 113016.

[11] I Cabria et al. “Adsorption and dissociation of molecular hydrogen on palladium clusters supported on graphene”. In: *The Journal of Physical Chemistry C* 116.40 (2012), pp. 21179–21189.

[12] R. Car et al. “Dielectric band structure of crystals: General properties and calculations for silicon”. In: *Phys. Rev. B* 24 (2 1981), pp. 985–999.

[13] Manuel Cardona. “Electronphonon interaction in tetrahedral semiconductors”. In: *Solid State Communications* 133.1 (2005), pp. 3 –18.

[14] Manuel Cardona and M. L. W. Thewalt. “Isotope effects on the optical spectra of semiconductors”. In: *Rev. Mod. Phys.* 77 (4 2005), pp. 1173–1224.

[15] Aleksandr Cariow and Galina Cariowa. “A Hardware-Efficient Approach to Computing the Rotation Matrix from a Quaternion”. In: *arXiv:1609.01585* (2016).

[16] A. H. Castro Neto and F. Guinea. “Impurity-Induced Spin-Orbit Coupling in Graphene”. In: *Phys. Rev. Lett.* 103 (2 2009), p. 026804.

[17] A. H. Castro Neto et al. “The electronic properties of graphene”. In: *Rev. Mod. Phys.* 81 (1 2009), p. 109.

[18] David M Ceperley and BJ Alder. “Ground state of the electron gas by a stochastic method”. In: *Physical Review Letters* 45.7 (1980), p. 566.

[19] Kevin T. Chan, J. B. Neaton, and Marvin L. Cohen. “First-principles study of metal adatom adsorption on graphene”. In: *Phys. Rev. B* 77 (23 2008), p. 235430.

[20] James R. Chelikowsky and Marvin L. Cohen. “Nonlocal pseudopotential calculations for the electronic structure of eleven diamond and zinc-blende semiconductors”. In: *Phys. Rev. B* 14 (2 1976), pp. 556–582.

[21] An-Ban Chen and B. Segall. “Piezooptical response associated with the interconduction band transitions in the noble metals”. In: *Solid State Communications* 18.1 (1976), pp. 149 –152.

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

[23] R. Courths et al. “Electronic structure of gold: An angle-resolved photoemission study along the Λ line”. In: *Phys. Rev. B* 34 (6 1986), pp. 3577–3585.

[24] Andrea Dal Corso. “Pseudopotentials periodic table: From H to Pu”. In: *Computational Materials Science* 95 (2014), pp. 337–350.

[25] A. Delin. “First-principles calculations of the II-VI semiconductor β-HgS: Metal or semiconductor”. In: *Phys. Rev. B* 65 (15 2002), p. 153205.

[26] Jack Deslippe et al. "BerkeleyGW: A massively parallel computer package for the calculation of the quasiparticle and optical properties of materials and nanostructures". In: *Computer Physics Communications* 183.6 (2012), pp. 1269–1289.

[27] Jack Deslippe et al. "Coulomb-hole summations and energies for G W calculations with limited number of empty orbitals: A modified static remainder approach". In: *Physical Review B* 87.16 (2013), p. 165124.

[28] Jun Ding et al. "Engineering quantum anomalous/valley Hall states in graphene via metal-atom adsorption: An ab-initio study". In: *Phys. Rev. B* 84 (19 2011), p. 195444.

[29] K Dybko et al. "Band structure of β-HgS from Shubnikov–de Haas effect". In: *Physica B: Condensed Matter* 256 (1998), pp. 629–632.

[30] Enrico Fermi et al. "Motion of neutrons in hydrogenous substances". In: *Ricerca sci* 7 (1936), pp. 13–52.

[31] Andrzej Fleszar. "Dielectric response in semiconductors: theory and applications". PhD thesis. University of Trieste, 1985.

[32] Tobias Förster, Peter Krüger, and Michael Rohlfing. "Two-dimensional topological phases and electronic spectrum of Bi 2 Se 3 thin films from G W calculations". In: *Physical Review B* 92.20 (2015), p. 201404.

[33] Christoph Friedrich, Stefan Blügel, and Arno Schindlmayr. "Efficient implementation of the G W approximation within the all-electron FLAPW method". In: *Physical Review B* 81.12 (2010), p. 125102.

[34] Eric Ganz, Klaus Sattler, and John Clarke. "Scanning tunneling microscopy of Cu, Ag, Au and Al adatoms, small clusters, and islands on graphite". In: *Surface Science* 219.1 (1989), pp. 33 –67.

[35] Paolo Giannozzi et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials". In: *Journal of Physics: Condensed Matter* 21.39 (2009), p. 395502.

[36] R. W. Godby and R. J. Needs. "Metal-insulator transition in Kohn-Sham theory and quasiparticle theory". In: *Phys. Rev. Lett.* 62 (10 1989), pp. 1169–1172.

[37] R. W. Godby, M. Schlüter, and L. J. Sham. "Self-energy operators and exchange-correlation potentials in semiconductors". In: *Phys. Rev. B* 37 (17 1988), pp. 10159–10175.

[38] X Gonze et al. "A brief introduction to the ABINIT software package". In: *Zeitschrift für Kristallographie.(Special issue on Computational Crystallography.)* 220 (2005), pp. 558–562.

[39] Marco Govoni and Giulia Galli. "Large Scale GW Calculations". In: *J. Chem. Theory Comput.* 11.6 (2015). PMID: 26575564, pp. 2680–2696.

[40] D. R. Hamann. "Optimized norm-conserving Vanderbilt pseudopotentials". In: *Phys. Rev. B* 88 (8 2013), p. 085117.

[41] C Hartwigsen, Sephen Gœdecker, and Jürg Hutter. "Relativistic separable dual-space Gaussian pseudopotentials from H to Rn". In: *Physical Review B* 58.7 (1998), p. 3641.

[42] Lars Hedin and Stig Lundqvist. "Effects of Electron-Electron and Electron-Phonon Interactions on the One-Electron States of Solids". In: *Solid State Physics* 23 (1970), pp. 1 –181.

[43] Hans Hellmann. "A new approximation method in the problem of many electrons". In: *The Journal of Chemical Physics* 3.1 (1935), pp. 61–61.

[44] K. H. Hellwege and O. Madelung. *Numerical Data and Functional Relationships in Science and Technology.* Vol. 17a and 22a. Landolt Brnstein, New Series, Group III. Springer, 1982.

[45] LA Hemstreet, CY Fong, and JS Nelson. "First-principles calculations of spin-orbit splittings in solids using nonlocal separable pseudopotentials". In: *Physical Review B* 47.8 (1993), p. 4238.

[46] P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas". In: *Phys. Rev.* 136 (3B 1964), B864–B871.

[47] Jun Hu et al. "Giant Topological Insulator Gap in Graphene with $5d$ Adatoms". In: *Phys. Rev. Lett.* 109 (26 2012), p. 266801.

[48] William P Huhn and Volker Blum. "One-hundred-three compound band-structure benchmark of post-self-consistent spin-orbit coupling treatments in density functional theory". In: *Physical Review Materials* 1.3 (2017), p. 033803.

[49] Mark S. Hybertsen and Steven G. Louie. "Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies". In: *Phys. Rev. B* 34 (8 1986), pp. 5390–5413.

[50] Mark S Hybertsen and Steven G Louie. "Spin-orbit splitting in semiconductors and insulators from the ab initio pseudopotential". In: *Physical Review B* 34.4 (1986), p. 2920.

[51] R. C. Jaklevic and L. C. Davis. "Band signatures in the low-energy-electron reflectance spectra of fcc metals". In: *Phys. Rev. B* 26 (10 1982), pp. 5391–5397.

[52] R. C. Jaklevic and John Lambe. "Experimental study of quantum size effects in thin metal films by electron tunneling". In: *Phys. Rev. B* 12 (10 1975), pp. 4146–4160.

[53] Wei-xiao Ji et al. "Giant gap quantum spin Hall effect and valley-polarized quantum anomalous Hall effect in cyanided bismuth bilayers". In: *New Journal of Physics* 18.8 (2016), p. 083002.

[54] Zhenzhao Jia et al. "Transport study of graphene adsorbed with indium adatoms". In: *Phys. Rev. B* 91 (8 2015), p. 085411.

[55] Joseph L Jules and John R Lombardi. "Transition metal dimer internuclear distances from measured force constants". In: *The Journal of Physical Chemistry A* 107.9 (2003), pp. 1268–1273.

[56] K Karlsson and F Aryasetiawan. "Spin-wave excitation spectra of nickel and iron". In: *Physical Review B* 62.5 (2000), p. 3006.

[57] Sunghun Kim et al. "Surface Scattering via Bulk Continuum States in the 3D Topological Insulator Bi_2Se_3". In: *Phys. Rev. Lett.* 107 (5 2011), p. 056803.

[58] Charles Kittel. *Quantum theory of solids.* Wiley, 1987.

[59] Leonard Kleinman and DM Bylander. "Efficacious form for model pseudopotentials". In: *Physical Review Letters* 48.20 (1982), p. 1425.

[60] W. Kohn and L. J. Sham. "Self-Consistent Equations Including Exchange and Correlation Effects". In: *Phys. Rev.* 140 (4A 1965), A1133–A1138.

[61] W Kohn and P Vashishta. "General density functional theory". In: *Theory of the Inhomogeneous Electron gas.* Springer, 1983, pp. 79–147.

[62] Takao Kotani and Mark Van Schilfgaarde. "All-electron GW approximation with the mixed basis expansion based on the full-potential LMTO method". In: *Solid State Communications* 121.9 (2002), pp. 461–465.

[63] Wei Ku and Adolfo G. Eguiluz. "Band-Gap Problem in Semiconductors Revisited: Effects of Core States and Many-Body Self-Consistency". In: *Phys. Rev. Lett.* 89 (12 2002), p. 126401.

[64] Andrey Kutepov et al. "Electronic structure of Pu and Am metals by self-consistent relativistic G W method". In: *Physical Review B* 85.15 (2012), p. 155129.

[65] P. Lautenschlager et al. “Interband critical points of GaAs and their temperature dependence”. In: *Phys. Rev. B* 35 (17 1987), pp. 9174–9189.

[66] Abdullahi Lawal et al. “First-principles many-body comparative study of Bi2Se3 crystal: A promising candidate for broadband photodetector”. In: *Physics Letters A* 381.35 (2017), pp. 2993 –2999.

[67] Kurt Lejaeghere et al. “Reproducibility in density functional theory calculations of solids”. In: *Science* 351.6280 (2016). eprint: `http://science.sciencemag.org/content/351/6280/aad3000.full.pdf`.

[68] Yuanchang Li et al. “Topological insulators in transition-metal intercalated graphene: The role of d electrons in significantly increasing the spin-orbit gap”. In: *Phys. Rev. B* 87 (24 2013), p. 245127.

[69] I Lopez-Corral et al. “Hydrogen adsorption on palladium dimer decorated graphene: a bonding study”. In: *international journal of hydrogen energy* 37.8 (2012), pp. 6653–6665.

[70] Steven G Louie and Marvin L Cohen. *Conceptual foundations of materials: a standard model for ground-and excited-state properties*. Vol. 2. Elsevier, 2006.

[71] Dongwei Ma, Zhongyao Li, and Zhongqin Yang. “Strong spinorbit splitting in graphene with adsorbed Au atoms”. In: *Carbon* 50.1 (2012), pp. 297 –305.

[72] Otfried Madelung. *Semiconductors: Data Handbook*. Springer Berlin Heidelberg, 2004.

[73] Brad D Malone and Marvin L Cohen. “Quasiparticle semiconductor band structures including spinorbit interactions”. In: *Journal of Physics: Condensed Matter* 25.10 (2013), p. 105503.

[74] D.v.d. Marel et al. “Unoccupied band critical point energies of noble metals determined with Bremsstrahlung isochromat spectroscopy”. In: *Solid State Communications* 50.1 (1984), pp. 47 –50.

[75] Andrea Marini et al. “Yambo: an ab initio tool for excited state calculations”. In: *Computer Physics Communications* 180.8 (2009), pp. 1392–1403.

[76] F Landis Markley. “Unit quaternion from rotation matrix”. In: *Journal of guidance, control, and dynamics* 31.2 (2008), pp. 440–442.

[77] FL Markley and GM Lerner. “Three-Axis Altitude Determination Methods”. In: *Spacecraft Altitude Determination and Control* 1 (1978), pp. 410–428.

[78] Richard M Martin. *Electronic structure: basic theory and practical methods*. Cambridge university press, 2004.

[79] G Martinez et al. “Determination of the energy band gap of Bi 2 Se 3”. In: *Scientific reports* 7.1 (2017), p. 6891.

[80] D R Masovic, F R Vukajlovic, and S Zekovic. “Local-pseudopotential calculation for optical properties and photoemission valence-band spectrum of silicon”. In: *Journal of Physics C: Solid State Physics* 16.35 (1983), p. 6731.

[81] V. Maurice and P. Marcus. “STM study of sputter-deposited Al clusters in chemical interaction with graphite (0001) surfaces”. In: *Surface Science* 275.1 (1992), pp. 65 –74.

[82] K. A. Mills et al. “Angle-resolved photoemission determination of Λ-line valence bands in Pt and Au using synchrotron radiation”. In: *Phys. Rev. B* 22 (2 1980), pp. 581–592.

[83] C. E. Moore. *Atomic Energy Levels.* Vol. 1. Washington, DC: Circular of the National Buraeu of Standards, 1958.

[84] C. E. Moore. *Atomic Energy Levels.* Vol. 3. Washington, DC: Circular of the National Buraeu of Standards, 1958.

[85] Arash A. Mostofi et al. “wannier90: A tool for obtaining maximally-localised Wannier functions”. In: *Computer Physics Communications* 178.9 (2008), pp. 685 –699.

[86] Jamal I. Mustafa et al. “Ab initio”. In: *Phys. Rev. B* 94 (15 2016), p. 155105.

[87] I. A. Nechaev et al. “Evidence for a direct band gap in the topological insulator Bi_2Se_3 from theory and experiment”. In: *Phys. Rev. B* 87 (12 2013), p. 121111.

[88] M. Orlita et al. “Magneto-Optics of Massive Dirac Fermions in Bulk Bi_2Se_3”. In: *Phys. Rev. Lett.* 114 (18 2015), p. 186401.

[89] Igor A Pašti et al. “Atomic adsorption on pristine graphene along the Periodic Table of Elements-From PBE to non-local functionals”. In: *arXiv:1710.08985* (2017).

[90] J. P. Perdew and Alex Zunger. “Self-interaction correction to density-functional approximations for many-electron systems”. In: *Phys. Rev. B* 23 (10 1981), pp. 5048–5079.

[91] John P Perdew, Kieron Burke, and Matthias Ernzerhof. “Generalized gradient approximation made simple”. In: *Physical review letters* 77.18 (1996), p. 3865.

[92] John P Perdew and Yue Wang. “Accurate and simple analytic representation of the electron-gas correlation energy”. In: *Physical Review B* 45.23 (1992), p. 13244.

[93] John P Perdew et al. "Restoring the density-gradient expansion for exchange in solids and surfaces". In: *Physical Review Letters* 100.13 (2008), p. 136406.

[94] James C Phillips. "Energy-band interpolation scheme based on a pseudopotential". In: *Physical Review* 112.3 (1958), p. 685.

[95] James C Phillips and Leonard Kleinman. "New method for calculating wave functions in crystals and molecules". In: *Physical Review* 116.2 (1959), p. 287.

[96] Zhenhua Qiao et al. "Quantum anomalous Hall effect in graphene from Rashba and exchange effects". In: *Phys. Rev. B* 82 (16 2010), p. 161414.

[97] Diana Y Qiu, H Felipe, and Steven G Louie. "Optical spectrum of MoS 2: many-body effects and diversity of exciton states". In: *Physical review letters* 111.21 (2013), p. 216805.

[98] Andrew M Rappe et al. "Optimized pseudopotentials". In: *Physical Review B* 41.2 (1990), p. 1227.

[99] Michael Rohlfing and Steven G. Louie. "Electron-hole excitations and optical spectra from first principles". In: *Phys. Rev. B* 62 (8 2000), pp. 4927–4944.

[100] I. P. Rusinov, I. A. Nechaev, and E. V. Chulkov. "Theoretical study of influencing factors on the dispersion of bulk band-gap edges and the surface states in topological insulators Bi2Te3 and Bi2Se3". In: *Journal of Experimental and Theoretical Physics* 116.6 (2013), pp. 1006–1017.

[101] R. Sakuma et al. "*GW* calculations including spin-orbit coupling: Application to Hg chalcogenides". In: *Phys. Rev. B* 84 (8 2011), p. 085144.

[102] Flaviano Jose dos Santos et al. "The impact of complex adatom-induced interactions on quantum spin Hall phases". In: *arXiv:1712.07827* (2017).

[103] Trond Saue and Trygve Helgaker. "Four-component relativistic Kohn–Sham theory". In: *Journal of computational chemistry* 23.8 (2002), pp. 814–823.

[104] Peter Scherpelz et al. "Implementation and Validation of Fully Relativistic GW Calculations: SpinOrbit Coupling in Molecules, Nanocrystals, and Solids". In: *Journal of Chemical Theory and Computation* 12.8 (2016), pp. 3523–3544.

[105] Martin Schlipf and François Gygi. "Optimization algorithm for the generation of ONCV pseudopotentials". In: *Computer Physics Communications* 196 (2015), pp. 36–44.

[106] Ramamurti Shankar. *Principles of quantum mechanics.* Springer Science & Business Media, 2012.

[107] C David Sherrill and Henry F Schaefer III. “The configuration interaction method: Advances in highly correlated approaches”. In: *Advances in quantum chemistry.* Vol. 34. Elsevier, 1999, pp. 143–269.

[108] John C Slater. “Statistical exchange-correlation in the self-consistent field”. In: *Advances in quantum chemistry.* Vol. 6. Elsevier, 1972, pp. 1–92.

[109] Zhigang Song et al. “Quantum spin Hall insulators and quantum valley Hall insulators of BiX/SbX (X= H, F, Cl and Br) monolayers with a record bulk band gap”. In: *NPG Asia Materials* 6.12 (2014), e147.

[110] Catalin-Dan Spataru. “Electronic Excitations in Solids and Novel Materials”. PhD thesis. University of California at Berkeley, 2004.

[111] G. Strinati. “Application of the Green’s functions method to the study of the optical properties of semiconductors”. In: *La Rivista del Nuovo Cimento (1978-1999)* 11.12 (1988), pp. 1–86.

[112] Jianwei Sun, Adrienn Ruzsinszky, and John P Perdew. “Strongly constrained and appropriately normed semilocal density functional”. In: *Physical review letters* 115.3 (2015), p. 036402.

[113] Peter Sutter, Jerzy T. Sadowski, and Eli Sutter. “Graphene on Pt(111): Growth and substrate interaction”. In: *Phys. Rev. B* 80 (24 2009), p. 245411.

[114] A. Svane et al. “Quasiparticle band structures of β-HgS, HgSe, and HgTe”. In: *Phys. Rev. B* 84 (20 2011), p. 205205.

[115] Paul Szczepanek and R. Glosser. “Piezo-optical constants of gold”. In: *Solid State Communications* 15.8 (1974), pp. 1425 –1429.

[116] Murilo L Tiago, Sohrab Ismail-Beigi, and Steven G Louie. “Effect of semicore orbitals on the electronic band gaps of Si, Ge, and GaAs within the GW approximation”. In: *Physical Review B* 69.12 (2004), p. 125212.

[117] Norman Troullier and José Luís Martins. “Efficient pseudopotentials for plane-wave calculations”. In: *Physical review B* 43.3 (1991), p. 1993.

[118] M. M. Ugeda et al. “Point Defects on Graphene on Metals”. In: *Phys. Rev. Lett.* 107 (11 2011), p. 116803.

[119] E v Van Lenthe, JG Snijders, and EJ Baerends. “The zero-order regular approximation for relativistic effects: The effect of spin–orbit coupling in closed shell molecules”. In: *The Journal of chemical physics* 105.15 (1996), pp. 6505–6516.

[120] David Vanderbilt. “Soft self-consistent pseudopotentials in a generalized eigenvalue formalism”. In: *Phys. Rev. B* 41 (11 1990), pp. 7892–7895.

[121] Julien Vidal et al. “False-positive and false-negative assignments of topological insulators in density functional theory and hybrids”. In: *Physical Review B* 84.4 (2011), p. 041109.

[122] Fran çois Virot et al. “Engineering Topological Surface States: HgS, HgSe, and HgTe”. In: *Phys. Rev. Lett.* 111 (14 2013), p. 146803.

[123] Fran çois Virot et al. “Metacinnabar (β-HgS): A Strong 3D Topological Insulator with Highly Anisotropic Surface States”. In: *Phys. Rev. Lett.* 106 (23 2011), p. 236806.

[124] S W. Shepperd. “Quaternion from Rotation Matrix”. In: JGC-1 (July 1978).

[125] R. E. Watson et al. “Local-density approximation: Cohesion in the transition metals and s→d promotion in the transition-metal atoms”. In: *Phys. Rev. B* 43 (2 1991), pp. 1455–1462.

[126] Conan Weeks et al. “Engineering a Robust Quantum Spin Hall State in Graphene via Adatom Deposition”. In: *Phys. Rev. X* 1 (2 2011), p. 021001.

[127] S-H Wei and Alex Zunger. “Role of metal d states in II-VI semiconductors”. In: *Physical Review B* 37.15 (1988), p. 8958.

[128] J. Wiebe et al. “Unoccupied surface state on Pt(111) revealed by scanning tunneling spectroscopy”. In: *Phys. Rev. B* 72 (19 2005), p. 193406.

[129] E. Wigner. “On the Interaction of Electrons in Metals”. In: *Phys. Rev.* 46 (11 1934), pp. 1002–1011.

[130] Nathan Wiser. “Dielectric Constant with Local Field Effects Included”. In: *Phys. Rev.* 129 (1 1963), pp. 62–69.

[131] Oleg V. Yazyev et al. “Quasiparticle effects in the bulk and surface-state bands of Bi_2Se_3 and Bi_2Te_3 topological insulators”. In: *Phys. Rev. B* 85 (16 2012), p. 161101.

[132] Peter Y Yu and Manuel Cardona. *Fundamentals of semiconductors: physics and materials properties.* Graduate texts in physics. Berlin: Springer, 2010.

[133] R Zallen and Michael Slade. “Plasma edge and band structure of cubic HgS”. In: *Solid State Communications* 8.16 (1970), pp. 1291–1294.

[134] Haijun Zhang et al. “Topological insulators in Bi2Se3, Bi2Te3 and Sb2Te3 with a single Dirac cone on the surface”. In: *Nature physics* 5.6 (2009), p. 438.

[135] Hongbin Zhang et al. “Electrically Tunable Quantum Anomalous Hall Effect in Graphene Decorated by 5d Transition-Metal Adatoms”. In: *Phys. Rev. Lett.* 108 (5 2012), p. 056802.

[136] S. B. Zhang et al. “Evaluation of quasiparticle energies for semiconductors without inversion symmetry”. In: *Phys. Rev. B* 40 (5 1989), pp. 3162–3168.

[137] Yingkai Zhang and Weitao Yang. “Comment on Generalized gradient approximation made simple”. In: *Physical Review Letters* 80.4 (1998), p. 890.

[138] Qin Zhou et al. “Imprint of transition metal *d* orbitals on a graphene Dirac cone”. In: *Phys. Rev. B* 88 (23 2013), p. 235431.

[139] Stefan Zollner et al. “Temperature dependence of the dielectric function and the interband critical-point parameters of GaSb”. In: *Phys. Rev. B* 43 (5 1991), pp. 4349–4360.