

Optical Spectroscopy and the
Contruction of an Optimal
Wannier Basis with Application to
the Development of *Ab Initio*
Models

A Dissertation Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville

Robert Gerard Van Wesep

December 2016

Copyright © 2016 by Robert Gerard Van Wesep,
All Rights Reserved.

This dissertation is dedicated to my parents, Edward P. and Katherine Van Wesep.

Without their love and support, this work would have been impossible.

Acknowledgements

I would like to thank Prof. Zhenyu Zhang and Dr. G. Malcolm Stocks for their advising early in my graduate career. I owe especial thanks to Prof. Adolfo G. Eguiluz for support, guidance of my research and countless discussions throughout the research that went into this dissertation. I have also stood on the shoulders of many of Prof. Eguiluz's former students and postdoctoral researchers. In particular, Dr. Anton V. Kozhevnikov's work on the Exciting Plus platform before my arrival provided the foundation for virtually all of the calculations and computational innovations in this work. I would also like to thank Joshua Braverman for camaraderie and support day-to-day through many difficult times and Nicholas Sirica for his friendship and many enlightening discussions. I am indebted to the technical support staff at the National Institute for Computational Sciences (NICS), especially Mr. Ryan Hulguin, for assistance in several computational issues. NICS also provided access the *Kraken* and *Darter* supercomputers, without which this research would have been impossible. I am grateful to Prof. Anthony Mezzacappa and Ms. Patti Redd for assistance in obtaining computational resources. I deeply appreciate the financial support I was granted throughout my graduate career, from the UTK/ORNL Distinguished Fellowship Program, the National Science Foundation, the Department of Energy, the Joint Institute for Advanced Materials, the Department of Physics and Astronomy and the Stelson Family. I owe a debt of gratitude to Prof. Hanno Weitering and Ms. Showni Medlin-Crump for assistance in maintaining my financial support to the completion of my degree.

Abstract

Understanding the role of local orbital degrees of freedom in the behavior of solid state systems has long been understood as a key to unraveling the mysteries presented by complex transition metal compounds. A general approach to the many-body problem is density functional theory (DFT) and its time-dependent extension (TDDFT), which provide a realistic representation of the material-dependent symmetry and chemistry of a compound. Calculation of quantities in (TD)DFT are most often performed using the basis of Bloch states, which is not natural for investigating local degrees of freedom. The Wannier basis provides localized orbitals that retain all of the information on the chemistry and symmetry of the groundstate. By transforming the expressions from TDDFT for the density-response functions into the Wannier basis, I have been able to develop novel methods for the investigation into the role of orbital degrees of freedom on the particle-hole excitation spectrum and used these methods to develop a scheme to construct a Wannier basis by tuning the gauge of the transformation to produce the simplest, yet fully accurate, description of the low-energy optical spectrum. Cu_2O [cuprite] provided the foundational test case for this development and the key to optimization proved to be the transference of chemically induced hybridization from energy space into real space. Using FeTe as a test case for correlated materials required understanding its graphene-like optical response, which I explain as a result of orbitally selective, local correlations giving rise to two-dimensional, linearly dispersing band topology. I also interpret the correlations and optical excitations using the Wannier basis. Subsequently, I have extended the

optimization method to FeTe, which led naturally to 1) a “target” subspace of the full Hilbert space that is disentangled from the “rest”, allowing for the construction of minimal, *ab initio* models by utilizing the constrained random-phase approximation (CRPA) and 2) ordered molecular orbitals that correspond to the structural and magnetic phase transition. The results contained in this dissertation will have broad implications for the treatment of long-ranged physics in future theoretical approaches to complex transition metal compounds.

Table of Contents

Introduction	1
1 Theory	12
1.1 Kohn-Sham density-functional theory	15
1.2 The linearized augmented planewave basis	26
1.3 Linear response in time-dependent density-functional theory	34
1.4 The Wannier basis	44
1.5 <i>Ab initio</i> downfolding and the constrained random-phase approximation	56
1.6 Linear response in the Wannier basis	62
1.7 The LDA+ U method	73
2 The optical response of Cu₂O and optimization of the Wannier basis	79
2.1 Optical measurements and calculations in the Bloch basis	85
2.2 Analysis of the optical spectrum involving single-site projected occupied orbitals	90
2.3 Analysis of the optical spectrum involving molecular z^2 orbitals	106
2.4 Implications for other complex transition metal materials	116
3 Local correlations and nanometer particle-holes in the graphene-like optical response of FeTe	117
3.1 Experimental and theoretical overview	120
3.2 Vital importance of local correlations	129

3.3	Mechanism for the graphene-like conductivity	138
3.4	Insight into the mid-infrared using Wannier response	141
3.5	Improving the far-infrared	148
4	Construction of an optimized Hamiltonian for FeTe	152
4.1	Single-site projection	156
4.2	Bonding/anti-bonding projection and orbital ordering	160
4.3	The optimal Wannier basis	170
4.4	Comparison of <i>ab initio</i> models	182
	Conclusion	191
	Bibliography	195
	Appendix	219
A	An orbitally driven collective excitation in Sr_2CuO_3	220
B	Explicit formula for the exchange-correlation energy	229
C	A discussion of spin density-functional theory	232
D	Proof of orthogonality of core orbitals in the LAPW basis	236
E	Explicit expressions for LAPW matrix elements	239
F	Proof of the Runge-Gross Theorem	242
G	Derivation of the Kohn-Sham density response	248
H	Derivation of an expression for the density response in CRPA	253
I	A Multi-Orbital Two-Particle Self-Consistent Method	256
	Vita	267

List of Tables

4.1	Fe d onsite direct (U) and exchange (J) Coulomb parameters in eV for the d - p model.	183
4.2	Te p onsite direct (U) and exchange (J) Coulomb parameters in eV for the d - p model.	183
4.3	Onsite direct (U) and exchange (J) Coulomb parameters in eV for the d model.	184
4.4	Onsite direct (U) and exchange (J) Coulomb parameters in eV for the t_{2g} model.	185

List of Figures

2.1	Cu ₂ O crystal structure.	82
2.2	Calculated dielectric response of Cu ₂ O versus experiment.	87
2.3	Cu ₂ O Wannier decomposed DOS comparison.	93
2.4	Cu ₂ O <i>d</i> -like Wannier orbitals.	95
2.5	Cu ₂ O <i>d</i> -like Wannier orbitals with <i>p</i> hybridization.	96
2.6	DOS due to various Wannier orbitals in Cu ₂ O.	98
2.7	Cu ₂ O <i>d</i> -like Wannier orbitals with rotated local axis.	100
2.8	Cu ₂ O Cu <i>z</i> ² /O <i>p</i> hybridized orbital.	102
2.9	Cu ₂ O convergence of Wannier channels with R summation.	103
2.10	Cu ₂ O optical response decomposed using single-site projected Wannier orbitals.	105
2.11	The Cu ₂ O <i>z</i> ² bonding and anti-bonding orbitals.	108
2.12	Cu ₂ O optical response decomposed using molecular Wannier orbitals.	110
2.13	A momentum space picture of the optical spectrum of Cu ₂ O.	115
3.1	The role of onsite interactions in reproducing the measured optical spectrum of FeTe.	130
3.2	GGA and GGA+ <i>U</i> band structures for FeTe and LaFePO.	134
3.3	Dependence of the electronic structure on <i>U</i>	137
3.4	Two-dimensionality in the band structure for FeTe.	139
3.5	Spatially resolved contributions to the optical spectrum of FeTe.	142
3.6	Real-space visualization of a typical optical transition in FeTe.	143

3.7	Visualization of the long-ranged nature of the particle-hole transitions in FeTe.	145
3.8	Decomposition of the mid-infrared spectrum of FeTe according to classes of Wannier channels.	147
3.9	Low-energy GGA+ U band structure for FeTe and LaFePO.	150
4.1	FeTe single-site projected DOS.	158
4.2	FeTe DOS using a projection with e_g bonding and anti-bonding orbitals.	161
4.3	Band structure for zigzag versus parallel ordering of the e_g orbitals in FeTe and LaFePO.	164
4.4	Optimization of the hole pocket in FeTe.	166
4.5	Optimization of the hole pocket in LaFeAsO.	167
4.6	Ordering pattern for the e_g bonding orbitals in FeTe and LaFeAsO.	169
4.7	Band structure and DOS of the optimal Wannier basis for FeTe.	171
4.8	Bonding and anti-bonding e_g orbitals for FeTe.	173
4.9	t_{2g} orbitals for FeTe.	175
4.10	Optimization of the Wannier basis in FeTe monitored by optics.	178
4.11	The importance of long-ranged particle-hole pairs on the optical spectrum of FeTe.	181
4.12	Evolution of the band structure with hopping for the t_{2g} model.	186
A.1	One-dimensional chain in Sr_2CuO_3	221
A.2	Band structure and contributions of Wannier channels to $\sigma_1(\omega)$ in Sr_2CuO_3	223
A.3	Real space plots of the x^2 - y^2 Wannier functions in Sr_2CuO_3	225
A.4	Dispersion of the loss peak in Sr_2CuO_3	227

Introduction

In any solid-state system, there are two primary length scales to consider. One is the length scale of the atom and its neighbors. The magnitude of this scale is conveniently set by the dimensions of the unit cell. This will typically be on the order of a few angstroms. The other length scale is that of the real-world sample that is studied by experimentalists. This will typically be on the order of a few millimeters or centimeters, that is tens or hundreds of millions times larger than the size of the unit cell. The theoretical approach to solid-state physics approximates the latter using a large volume with periodic boundary conditions, which I will call the “macro-crystal.” This approximate volume will typically be a few nanometers in dimension. This is certainly much smaller than the true crystal, but it still represents a length scale apart from that of the atom. In order to gain a complete theoretical understanding of a solid-state system, these two disparate pictures of the solid, the atom and the crystal, must be reconciled.

The need to explicitly handle physics on the atomic length scale in addition to that on the length scale of the crystal is especially critical for complex transition-metal compounds. These materials display a rich variety of phenomena and present a challenge to scientists who wish to understand them at a fundamental level. These are also the broad class of materials that are under investigation in this dissertation. The degrees of freedom on the atomic scale that are most important in these materials are the d orbitals. These sit on the boundary between delocalized s and p orbitals and localized f orbitals. For systems in which the s and/or p orbitals are most relevant,

the role of the local, atomic length scale can often be left as implicit and one can focus on the physics of long length scales. For systems in which f orbitals are of primary importance, the long length scales can often be placed to the side and the atomic scale can be focused on.

However, d orbitals, and particularly $3d$ orbitals, rest uncomfortably between these two extremes. It is often impossible to consider them entirely in terms of atomic scales. If one treats the physics on these small scales as implicit and focuses on the length scale of the crystal, then vital aspects of the physics are often obscured. Access to these local d orbital degrees of freedom represents the first pillar of the work in this dissertation. The Wannier basis [1–3] provides the ability to elucidate the role of local orbitals on the physics of a solid-state system while not sacrificing the physics on the longer length-scales of the crystal.

The second pillar of the work presented here is time-dependent density-functional theory (TDDFT) [4, 5], which I will use throughout as a tool for calculating particle-hole excitation spectra *from first-principles*. This is of vital importance, because any theoretical calculation must be monitored by and validated with regard to experiment. This is especially true in many-body physics where the problem is so enormously complex and strong assumptions and approximations must always be made. Of particular importance will be optical spectroscopy, which represents a primary window into the electronic structure of materials.

The central quantities in TDDFT linear response are most often written in terms of the Bloch basis. The Bloch basis are the eigenfunctions of the Kohn-Sham effectively noninteracting system used to calculate the groundstate energy and density in density-functional theory (DFT) [6, 7]. In order provide a link between the insight into the local orbital degrees of freedom provided by the Wannier basis and the correspondence to experimentally measured spectroscopic quantities provided by TDDFT, I will rely on a reformulation of these quantities in terms of the Wannier basis. This reformulation was a novel development internal to the group of Adolfo Eguiluz and I relied on these earlier results in this dissertation. However, I found it necessary to

extend those results and much of the investigation of the utility of response formulated in the Wannier basis is work original to this dissertation.

Once the link between Wannier orbitals and spectroscopy has been established, it is possible to build a feedback between these two areas. Using this feedback, I have developed a method for optimizing the Wannier basis with respect to optical spectroscopy which takes advantage of the freedom to choose a “gauge” for the transformation between the Bloch states and Wannier orbitals. This gauge is judged as “optimal” when the simplest and most insightful description of the optical spectrum that retains full accuracy is achieved. Implementing this optimization method for complex transition-metal compounds (specifically Cu_2O and FeTe) I have found that *molecular* orbital degrees of freedom that embody the local chemistry and symmetry are of critical importance. The form of these molecular orbitals highlight the role of *cation-cation bonding chemistry* which are in turn crucial for their role in the structural properties of the materials.

Once a basis of Wannier orbitals has been constructed, it is possible to construct an *ab initio* model for the orbital degrees of freedom covered by the Wannier basis. These degrees of freedom are called the “target” and all other degrees of freedom are the “rest.” The model consists of two parts, a single-particle part that represents the non-interacting reference system and a two-body interaction. The parameters for the single particle part, the site energies and hoppings, are derived directly from the matrix elements of the Kohn-Sham effective Hamiltonian in the Wannier basis. However, the interaction parameters are not simply the bare interactions in the Wannier basis. Instead the interaction between target degrees of freedom must be screened by the degrees of freedom in the rest. A widely used method for calculating these screened interaction parameters from the Kohn-Sham groundstate electronic structure is the constrained random-phase approximation (CRPA) [8–10] and this is the method I have utilized. When I perform the optimization procedure that I have developed using Cu_2O on FeTe , I find that it naturally yields a minimal, three-orbital-per-site model. Minimization of the degrees of freedom in the model is crucial as the

difficulty in solving any many-body model increases dramatically as the number of degrees of freedom increases.

However, there are two striking features of the *ab initio* model that are not frequently present in many models used to study the iron-pnictides and iron-chalcogenides. Both have to do with the long length scale of the problem. First, the hopping parameters that I find cannot be assumed to be zero beyond a few shells of neighbors. Instead, hopping across the entire macro-crystal, at a length scale of at least a few nanometers, must be retained to faithfully reproduce the Kohn-Sham electronic structure. Second, the offsite interaction parameters are not negligible in size and the mathematics of CRPA are such that the interactions will have a long-ranged tail that decreases like $1/R$, R being the distance between orbitals. Thus, though the details of the onsite parameters are sure to be important, the parameters on long, nanometer-scale distance should not be ignored in any solution within the *ab initio* model. This again highlights the coexistence of atomic length scale and the length scale of the crystal, both of which must be considered as important *a priori*.

In the preceding discussion I have given a rough outline of the goal and contents of this dissertation. In the remainder of this introduction I will give a brief road map and summarization of each section. In order to ease navigation, I will tie the points I have brought up concretely to the specific content in each section. I hope that this will make clear the role of each piece in the overall logic of the “big picture” that I have described above. Also, where I find it necessary I will highlight contributions that are my original work distinguished from work of others that I have relied on.

In Chapter 1, I will review and present the main results for the theoretical tools that I will use throughout this dissertation. First among these tools, and the foundation upon which the rest are built, is Kohn-Sham DFT [7], which I will present in Section 1.1. From the perspective of what follows, the most important quantities that result from the solution of the Kohn-Sham DFT problem are the Kohn-Sham eigenvalues and eigenfunctions (which are Bloch states.) The numerical calculation of these quantities requires a basis and I have chosen the APW+LO+lo basis [11]

which is a set of linearized augmented planewaves (LAPW) [12]. The computational platform that provides the foundation for virtually all calculations in this dissertation is the Exciting Plus code [13]. I will describe this basis and the matrices that must be diagonalized to yield the Kohn-Sham eigenvalues and eigenfunctions in Section 1.2.

Kohn-Sham DFT allows for the calculation of the groundstate energy, density and other groundstate quantities. However, essential to this dissertation is the comparison between theoretical calculations and spectroscopic experiments, in particular particle-hole spectroscopies, of which optical spectroscopy will take prime importance. In order to calculate quantities that are directly related to these experiments requires the time-dependent extension of DFT: TDDFT [4, 5] and especially expressions for the linear density response [14]. I will review the basic results from TDDFT and derive the quantities that describe the linear response of a solid-state system in Section 1.3.

Kohn-Sham DFT yields Bloch states and TDDFT linear response uses these to calculate the density response. However, the thrust of this dissertation is specifically the investigation of the role of local orbital degrees of freedom on spectroscopy and the use of optical spectroscopy to guide the optimal construction of a Wannier local orbital basis that can then be used to construct an *ab initio* model. In Section 1.4, I will present the methods that I will be using to construct the Wannier basis. Included among these methods is one that I have developed to construct molecular Wannier orbitals using an extension of the projection method using local orbitals [2] to one using a linear combination of local orbitals. The ability to construct these molecular orbitals will prove to be critical when I describe the optimization of the Wannier basis.

One goal of constructing the Wannier orbitals is the calculation of parameters that would enter into an *ab initio* microscopic model. I have used CRPA to calculate the screened interaction parameters for such a model. In Section 1.5, I will give an overview of the method and present the main mathematical results. In addition, I will define an expression for the density-response for the target degrees of freedom. The latter becomes important as I describe the reformulation of linear response in

TDDFT using the Wannier basis as opposed to the Bloch basis, in Section 1.6. In the basic expressions arrived at there, embodied in Eqs. (1.128) and (1.150), I have relied on unpublished notes of Adolfo Eguiluz. However, in the latter I contributed the insight that the interaction that should be included is the screened interaction rather than the bare interaction.

Section 1.6 also contains several extensions to the mathematics of linear response in the Wannier basis that are original to me and are critical to conducting many of the calculations used in the later chapters. The first is the expression for the Kohn-Sham response in which the sum over lattice vectors has been performed analytically, embodied in Eq. (1.134). This expression was critical since in both Cu_2O and FeTe , the significant contributions from nanometer-scale particle-hole transitions made the numerical calculation of Eq. (1.128) impractical. In addition, I found that both the expressions in Eqs. (1.134) and (1.128) were inadequate for metallic systems, such as FeTe . In order to perform these calculations, I derived alternative expressions to use in calculating the Kohn-Sham response given in Eqs. (1.142) and (1.143).

Rounding out Chapter 1, I present a brief overview of the LDA+ U method [15] in Section 1.7. This method will be crucial for the calculation of the optical response of FeTe in Chapter 3.

Using Cu_2O as a test case in Chapter 2, I will show that the optical response can be used to choose an optimal gauge for the Wannier transformation. In practice, the optimal set of orbitals produces the fewest particle-hole channels as possible to explain the optical spectrum. In Cu_2O , I have found that the orbitals produced using this optimization procedure embody certain vital aspects of the groundstate chemistry, in particular the participation of a z^2 orbital in both Cu-O and Cu-Cu bonding [16]. In this way, I will show that the optimization method I have developed produce a set of orbitals that reveal important aspects of local chemistry and symmetry and the impact on the particle-hole excitation spectrum.

In Section 2.1, I will present the results for the dielectric response calculated using the Bloch basis. There I will show that the TDDFT calculation in the random-phase

approximation (RPA) using the generalized-gradient approximation (GGA) for the groundstate yields remarkable, semi-quantitative agreement with the most prominent features measured by spectroscopic ellipsometry [17–19]. I will also address recent theoretical work on the electronic structure and optical spectrum of Cu_2O and place the GGA/RPA calculation in that context. The agreement with experiment allows me to develop the optimization procedure I have described without concerns that the target optical spectrum does not comport with reality. Thus, Cu_2O provides a reliable benchmark system for testing the optimization scheme.

In Section 2.2, I will present in detail the construction of Wannier orbitals using a “single-site projection.” That is the projection using local orbitals as has been laid out in the literature [2], where trial orbitals on one site are used to construct Wannier orbitals centered on that site. This is in contrast to the projection using a linear combination of local orbitals, the mathematics of which I will present in Section 1.4. Particularly important to the single-site projection is the choice of a local coordinate system that places the z^2 orbital along the O-Cu-O bonding direction. When this choice is made, the z^2 orbital becomes the critical hole for describing the optical spectrum when the spectrum is decomposed using the response calculated in the Wannier basis. However, its contribution only accounts for about half of the total spectral weight.

I will also show in Section 2.2 that particle-holes separated by more than a nanometer are required to reproduce the spectrum. In order to handle this I must use the expression I developed for response in the Wannier basis in which all distances are summed analytically (Eq. (1.134).) This again emphasizes the critical role played by long-ranged physics, even when the physics is described using “localized” orbitals. It is through the ability to use the Wannier basis to represent local orbital degrees of freedom while retaining all of the information about the long-ranged, solid-state physics that the interplay between atomic and nanometer length scales can be revealed.

I am able to achieve an optimal representation of the optical spectrum only when I construct the Wannier orbitals using a linear combination of local orbitals for projection. In Section 2.3, I will describe how these molecular orbitals are constructed and their chemical nature. In particular, I show that the z^2 orbital can be split into independent bonding and anti-bonding orbitals which embody the Cu-Cu bonding chemistry. It is this bonding which has been shown to *stabilize the crystal structure* [16]. I will also show that the transfer of d - d hybridization, which was present in energy space in the DOS for the single-site projection, into real space is a critical part of the nature of the bonding and anti-bonding orbitals.

The ultimate goal of constructing such an optimal set of Wannier orbitals is to use the basis to construct *ab initio* models to which many-body methods beyond DFT can be applied. Cu_2O is not the most suitable system for such a construction. There is little evidence that correlations among the local orbital degrees of freedom are important for the groundstate and models based on local orbitals are not among the widely used theoretical approaches to Cu_2O . The iron-based superconducting compounds, on the other hand, are widely studied using models such as the Hubbard model, which rely implicitly or explicitly on a basis of localized orbitals. As I will show in Chapter 4, I find that FeTe offers the most favorable system for the optimization procedure I have developed for Cu_2O , from a chemical perspective.

However, before optimization with respect to the optical spectrum proceeds, I must first address a problem. Unlike Cu_2O , FeTe is strongly correlated, which has a dramatic impact on the optical spectrum. In Chapter 3, I will give an account of the optical spectrum and electronic structure of FeTe. I will start in Section 3.1 with a brief review of the experimental and theoretical work on FeTe which paint the picture of a strongly correlated, bad metal with an important role played by orbital selectivity. I will focus on the experimental optical [20, 21] and photoemission [22–26] measurements to give a description of what is known about the electronic structure. These measurements will form the basis with which I will evaluate the success of my calculations.

In Section 3.2, I will show by calculating the groundstate with GGA+ U that the electronic structure is changed by local, orbitally selective correlations. These correlations then produce a renormalization of the optical spectrum which reduces its weight in the mid-infrared by about a factor of two. This renormalization is what is needed to bring the calculation into quantitative agreement with experiment, at least in the mid-infrared. I will compare this to the effect of GGA+ U versus GGA for LaFePO as a representative of the iron-pnictides and show that the effect of including local correlations is qualitatively different, with no dramatic renormalization of the spectrum.

As part of this elucidation of the nature of the optical spectrum in FeTe, I have been able to make a number of insights that shed light on the peculiar qualitative features of the spectrum. The experimentally measured optical conductivity is nearly constant over a broad frequency range, from the far-infrared to the mid-infrared [20, 21]. This is both qualitatively and quantitatively similar to the measured optical conductivity of graphite, which in turn can be explained by the electronic structure of graphene [27–29]. Seemingly, FeTe has very little in common with graphene, but based on an analysis of the band structure, I argue in Section 3.3 for a common origin for the flat conductivity in the band structure having to do with two-dimensionality and linearly dispersing bands.

Then in Section 3.4 I use the response in the Wannier basis to go beyond this momentum-space description of the mid-infrared spectrum and I show that the range of particle hole pairs is on the scale of nanometers, as in Cu_2O . This again shows the coexistence of the two length scales I have described. Local (onsite) correlations in the groundstate give rise to an electronic structure that correctly predicts the flat optical conductivity of FeTe, while long-ranged particle-holes must be considered when calculating the optical spectrum. I will also show that the interband transitions present in the far infrared show evidence for the presence of particle-holes on the scale of tens of nanometers. However, the calculation of the optical spectrum using the GGA+ U groundstate deviates sharply from experiment in the far infrared and

in Section 3.5, I propose how subtle changes in the electronic structure could have a further effect on the far infrared. I argue that these changes could result from a better many-body treatment of the correlations and would improve my calculation.

In Chapter 4, I apply the optimization procedure that I laid out in Chapter 2 to the GGA electronic structure of FeTe. In Section 4.1, I present the Wannier basis constructed using a single-site projection. I show that the orbitally resolved electronic structure displays intriguing similarities with Cu_2O . Guided by the results for Cu_2O , I will show that the crucial step is to split some of the d orbitals into independent bonding and anti-bonding orbitals in Section 4.2. This shows the importance of cation-cation bonding in both systems and indicates that this may be an important feature for many other complex transition-metal compounds. Remarkably, the optimization procedure also predicts an ordering pattern for the e_g bonding and anti-bonding orbitals which corresponds to a experimentally observed structural phase transition which is tied to a magnetic ordering unique to FeTe [30]. Though I do not directly study the phase transition in this dissertation, I am able to provide insight into the role of orbital degrees of freedom on the physics of the transition.

In Section 4.3, I will give a detailed account of the properties of the optimal (as defined above) Wannier basis for FeTe. I will show that the elimination of hybridization in energy space between e_g and the remaining t_{2g} orbitals is the key to optimization. The t_{2g} orbitals are the only present near the Fermi level and in particular in the electron and hole pockets. Though I have been able to arrive at the optimal orbital set without reference to TDDFT calculation of the optics, using the calculation of the optical response in the Wannier basis I will show that it is only through optimization via the elimination of hybridization that the simplest description of the low-energy optics as entirely due to the t_{2g} orbitals can be achieved. I will also give numerical evidence for the presence of long-ranged optical charge fluctuations and argue that these should not be ignored in any many-body treatment of a model constructed with the optimal Wannier basis.

In Section 4.4, I will describe and present results for the construction of an *ab initio* model for FeTe. By calculating onsite interaction parameters using CRPA, I will show a trend of increasing orbital selectivity in the interaction parameters as the target space is reduced to the minimal space spanned by the t_{2g} orbitals as well as an increasing relative importance of the Hund’s coupling. These observations support the orbital selectivity of correlations observed in photoemission and the theoretical picture of FeTe, and the iron-based superconducting compounds in general, as Hund metals [31, 32]. However, I also argue that both long-ranged hopping and interactions should not be ignored in any theoretical treatment of the *ab initio* model that I have constructed. Thus, again I show the need to handle both the physics on the length-scale of the atom and the long length-scale of the macro-crystal.

The Appendix generally contains discussions and results that did not belong in the main body of the text. However, I would like to highlight Appendix A as having a different character. Chapters 2, 3 and 4 represent completed work that each correspond to publications being prepared for submission to peer-reviewed journals. Appendix A also represents a high-profile work that will be part of a subsequent publication, but the results are preliminary and do not represent the same level of completeness as the chapters in the main body of the text. In Appendix A, I have applied the same methods of TDDFT linear response in a Wannier basis to the one-dimensional cuprate Sr_2CuO_3 . There I will show that the transition of a hybridized Cu x^2-y^2 /O p orbital along the linear Cu-O chain produces an optical excitation that in turn induces a collective mode that has been measured using EELS [33, 34]. The construction of the critical orbital degrees of freedom represent a valuable extension of the methods I have developed in this dissertation to another class of compounds.

Chapter 1

Theory

In this chapter, I will lay out the theoretical foundation and derive the results that I will use applied to real materials in the subsequent chapters. It will contain a mixture of well established background, novel developments that I inherited from my research group and results that are original to me. In order to make clear what my specific contributions are, I will identify which results I relied on and which I have derived. If there is no such statement or a citation, then that result is general background. I have tried to present this background in some detail so that this dissertation could be as self-contained as possible. There are various results that either diverge too much from the main flow of the text or represent related but tangential topics. These I have placed in the Appendix.

The driving principle in the theoretical development of the methods that I have used in the preparation of this dissertation is to build results on a firm foundation of exact many-body theory wherever possible. This defines a first principles or *ab initio* approach. I also am oriented as much as possible to describing real materials versus models that serve as rough analogs of physical systems. Since the subsequent chapters use calculations of electronic structure and density response, the theory in this chapter is focused on the electronic degrees of freedom. For this reason, from the start I will take the Born-Oppenheimer approximation and consider the nuclei as

fixed. As such they only provide an external potential and do not represent degrees of freedom. Throughout, I will consider positions as determined from either x-ray or neutron diffraction experiments.

With this philosophy in mind, I will begin with the Hamiltonian for a system of N electrons in an external potential, given by

$$\hat{H} = \hat{T} + \hat{U} + \hat{V}. \quad (1.1)$$

Following the derivations in [5], I will write the terms in $\hat{H}$ in first-quantized notation in this section. In this notation, $\hat{T}$ is the kinetic energy given by

$$\hat{T} = -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla_i^2, \quad (1.2)$$

where ∇_i is the gradient operator for coordinate i and m is the mass of the electron. $\hat{U}$ is the interaction potential energy, the Coulomb interaction:

$$\hat{U} = \frac{e^2}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^N \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|}, \quad (1.3)$$

where e is the charge of the electron. Finally, $\hat{V}$ is the external potential:

$$\hat{V} = \sum_{i=1}^N v(\mathbf{x}_i). \quad (1.4)$$

In a perfect crystal, this potential is periodic in lattice vectors $\mathbf{R}$, i.e. $v(\mathbf{x}_i + \mathbf{R}) = v(\mathbf{x}_i)$, and in addition is invariant under transformations of coordinates which constitute the space group symmetries.

The Hamiltonian governs the time evolution of the many-electron wavefunction via the Schrödinger equation:

$$\hat{H}\Psi(\mathbf{x}_1, \dots, \mathbf{x}_N, t) = i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{x}_1, \dots, \mathbf{x}_N, t). \quad (1.5)$$

Through the Schrödinger equation, the Hamiltonian, along with appropriate initial and boundary conditions, determines the full, time-dependent many-body wave function. Thus, the Hamiltonian contains virtually all of the information necessary to determine all of the properties of a system including its dynamics. The problem with the equation as written in Eq. (1.5) is that the number of variables involved in solving this equation is astronomical. For a macroscopic object the number of electrons is $> 10^{23}$. This produces a huge number of degrees of freedom, which make this equation impossible to solve, even numerically on the largest supercomputers that will likely ever be available, even in the distant future. The wave function could not even be stored. In practice, a smaller system with periodic boundary conditions, the macro-crystal, is used in calculations. As an example, Li has three electrons per atom, one atom per unit cell. A typical macro-crystal may contain $20 \times 20 \times 20 = 16,000$ unit cells giving 48,000 electrons in three dimensions leading to 144,000 degrees of freedom. Of course, this is still impossible to solve.

Life would be easier if there was no interaction term $\hat{U}$. In that case the Schrödinger equation would be separable. The much simpler single-particle problem could be solved and then the many-body eigenstates of the Hamiltonian would simply be anti-symmetrized products of eigenfunctions of the single particle problem. This simplification does not necessarily require that the interactions be ignored completely. One can find effective potentials produced by the interaction that are single particle in form. The simplest example is that of the Hartree potential, in which the single electron interacts with the single-particle electron density produced by all of the electrons. If one is willing to accept a non-local potential, one can add the Fock or exchange potential. These two potentials together enter into the Hartree-Fock approximation.

Hartree-Fock still ignores important effects of the interaction, so called correlation effects. The question arises whether we must surrender and tackle the interacting many-body Hamiltonian or if there is another way to solve an effective single-particle problem and still have the ability to treat all of the many-body effects. The answer

turns out to be a qualified yes and is provided by density functional theory (DFT) in its Kohn-Sham form. Kohn-Sham DFT is primarily built on two results. One is the work of Hohenberg and Kohn [6], which establishes the link between external potentials, electron densities and all other properties of the system as functionals of the electron density. The other is the work of Kohn and Sham [7], which provided a way to solve for the groundstate energy and electron density by first solving the eigenvalue equation for an effective Hamiltonian that is non-interacting in form. In this reformulation of the problem the focus shifts from the intractable many-body wavefunction to the single-particle electron density, a simple function of three variables.

Kohn-Sham DFT was not the first attempt to derive an expression for the energy of a many-electron system as a functional of the density alone. The most famous early attempt is the Thomas-Fermi method [35, 36], which seeks to write the energy as an explicit functional of the density. One of the primary problems is the kinetic energy term. In the Thomas-Fermi method, a functional form for the kinetic energy is derived from the expression for the kinetic energy of non-interacting electrons in a uniform gas. One of the most significant contributions of Kohn and Sham [7] was to split the kinetic energy into a part that is non-interacting in form and the rest and to write the non-interacting part as an implicit functional of the density via the so-called Kohn-Sham orbitals. All of this will be made explicit in the next section.

1.1 Kohn-Sham density-functional theory

The first observation to make in formulating DFT is that $\hat{T}$ and $\hat{U}$ are fixed in form for all systems. The only thing that distinguishes sodium from silicon from the iron-pnictides is the external potential. The external potential determines the groundstate wavefunction through the Hamiltonian and Schrödinger equation in Eq. (1.5) and this

in turn determines the groundstate electron density via the relation

$$n_0(\mathbf{x}) = N \int d^3x_1 \dots \int d^3x_{N-1} |\Psi_0(\mathbf{x}, \mathbf{x}_1, \dots, \mathbf{x}_{N-1})|^2, \quad (1.6)$$

with subscript 0 indicating the groundstate throughout. In fact, a more powerful result is true. Given a groundstate electron density, there is a unique external potential that determines that density. This proceeds from the observation that if a potential differs by more than a trivial constant, then the groundstate wavefunction must differ by more than a trivial phase. The proof is one of contradiction, following Hohenberg and Kohn [6]. Assuming that two potentials give the same groundstate wavefunction, the eigenvalue problem for the Hamiltonian gives

$$\langle \Psi_0 | \hat{T} + \hat{U} | \Psi_0 \rangle + \langle \Psi_0 | \hat{V} | \Psi_0 \rangle = E_0 \quad (1.7)$$

and

$$\langle \Psi_0 | \hat{T} + \hat{U} | \Psi_0 \rangle + \langle \Psi_0 | \hat{V}' | \Psi_0 \rangle = E'_0. \quad (1.8)$$

Subtracting gives $\langle \Psi_0 | \hat{V} | \Psi_0 \rangle - \langle \Psi_0 | \hat{V}' | \Psi_0 \rangle = E_0 - E'_0 = \text{const.}$ contradicting our assumption. This means that there is a one-to-one mapping between external potentials and groundstate wavefunctions, but what about the density?

Showing that a similar one-to-one mapping between potentials and groundstate densities exists boils down to showing that the two groundstate wavefunctions produced by the two external potentials cannot produce the same density. If they did, then given the density we would not know which potential it came from. Again, the argument is proof by contradiction [6]. First, there is an inequality proceeding from the Rayleigh-Ritz variational principle [37, 38]:

$$E'_0 < \langle \Psi_0 | \hat{T} + \hat{U} + \hat{V}' | \Psi_0 \rangle = E_0 + \langle \Psi_0 | \hat{V}' - \hat{V} | \Psi_0 \rangle \quad (1.9)$$

and the similar inequality

$$E_0 < \langle \Psi'_0 | \hat{T} + \hat{U} + \hat{V} | \Psi'_0 \rangle = E'_0 - \langle \Psi'_0 | \hat{V}' - \hat{V} | \Psi'_0 \rangle. \quad (1.10)$$

It is necessary here to compute the groundstate expectation value of $\hat{V}$. Inserting the N -particle resolution of the identity yields

$$\langle \Psi_0 | \hat{V} | \Psi_0 \rangle = \sum_{i=1}^N \int d^3x_1 \dots \int d^3x_N v(\mathbf{x}_i) |\Psi_0(\mathbf{x}_1, \dots, \mathbf{x}_N)|^2. \quad (1.11)$$

Since $|\Psi_0(\mathbf{x}_1, \dots, \mathbf{x}_N)|^2$ is invariant under exchange of coordinates, this can be written as

$$\langle \Psi_0 | \hat{V} | \Psi_0 \rangle = N \int d^3x_1 v(\mathbf{x}_1) \int d^3x_2 \dots \int d^3x_N |\Psi_0(\mathbf{x}_1, \dots, \mathbf{x}_N)|^2 = \int d^3x_1 v(\mathbf{x}_1) n_0(\mathbf{x}_1). \quad (1.12)$$

The assumption that $|\Psi_0\rangle$ and $|\Psi'_0\rangle$ give the same density leads to the conclusion that

$$\langle \Psi_0 | \hat{V}' - \hat{V} | \Psi_0 \rangle = \int d^3x [v'(\mathbf{x}) - v(\mathbf{x})] n_0(\mathbf{x}). \quad (1.13)$$

Thus, adding Eq. (1.9) to Eq. (1.10) gives $E'_0 + E_0 < E_0 + E'_0$, which is clearly impossible. $|\Psi_0\rangle$ and $|\Psi'_0\rangle$ cannot give the same density, so the one-to-one mapping between potentials and groundstates and between groundstates and densities combine to provide a one-to-one mapping between potentials and densities.

The existence of this remarkable mapping leads to a number of powerful conclusions. First, since for a given density there is a unique external potential and the external potential determines the wavefunction; the density also determines the expectation value of any observable. In particular, it determines the groundstate expectation value of the Hamiltonian: the groundstate energy. It is useful to define a total energy functional. I will follow the constrained search approach of Levy [39] and Leib [40]. Consider all possible states that have a particular single-particle density

$n(\mathbf{x})$ and write such a state as $|\Psi[n]\rangle$. Then the energy functional is defined using the Rayleigh-Ritz variational principle, minimizing with respect to the constrained set of states:

$$E[n] = \min \langle \Psi[n] | \hat{H} | \Psi[n] \rangle. \quad (1.14)$$

The groundstate energy is given by $\langle \Psi_0 | \hat{H} | \Psi_0 \rangle$, but I have shown that $|\Psi_0\rangle = |\Psi[n_0]\rangle$. Thus, it is the case that

$$\begin{aligned} E[n] &= E_0, & \text{if } n(\mathbf{x}) &= n_0(\mathbf{x}) \\ E[n] &> E_0, & \text{if } n(\mathbf{x}) &\neq n_0(\mathbf{x}). \end{aligned} \quad (1.15)$$

This leads to the second conclusion of Hohenberg and Kohn [6]: that $E[n]$ achieves a global minimum at the groundstate density and nowhere else. Thus, the variation of the energy functional subject to the constraint that there be N electrons vanishes about the groundstate density:

$$\delta \left[E[n] - \mu \int d^3x n(\mathbf{x}) \right]_{n(\mathbf{x})=n_0(\mathbf{x})} = 0, \quad (1.16)$$

where μ is a Lagrange multiplier. This variational principle for the density is what makes DFT worthwhile. Given an energy functional, it provides a way to find the groundstate density and energy. Subsequently, the groundstate density gives the groundstate expectation value of anything else that can be expressed as a functional of the density. From here there are two open problems: how to express the energy functional and how to carry out the variation.

The solution, provided by Kohn and Sham [7], is to go with what we know, namely the non-interacting system. For a non-interacting system the Schrödinger equation is separable and it is sufficient to solve the single-particle eigenvalue equation

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + v(\mathbf{x}) \right] \phi(\mathbf{x}) = \epsilon \phi(\mathbf{x}). \quad (1.17)$$

The many-body groundstate is then built from the solutions to the equation by counting eigenfunctions from lowest to higher eigenvalues until reaching N . These are the occupied eigenfunctions. The density is constructed from the occupied eigenfunctions by summing the complex modulus of them:

$$n_0(\mathbf{x}) = \sum_i^{\text{occ}} |\phi_i(\mathbf{x})|^2. \quad (1.18)$$

Also, the Hohenberg-Kohn theorem still applies to the non-interacting system, so there is a one-to-one mapping between potentials and densities. Thus, if a density for the interacting system is representable by the non-interacting system, then that density determines a unique potential for the non-interacting system that produces the same density for the interacting system. This non-interacting system I will call the Kohn-Sham system.

The kinetic energy functional for the Kohn-Sham system is given by

$$T^{\text{KS}}[n] = -\frac{\hbar^2}{2m} \sum_i^{\text{occ}} \int d^3x \phi_i^*(\mathbf{x}) \nabla^2 \phi_i(\mathbf{x}). \quad (1.19)$$

This is an implicit functional of the density via Eq. (1.18). It is also useful to write the Hartree potential that arises from the interaction explicitly as it is a simple functional of the density:

$$E^{\text{H}}[n] = \frac{e^2}{2} \int d^3x \int d^3x' \frac{n(\mathbf{x})n(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|}. \quad (1.20)$$

Then it is possible to rewrite the energy functional using these terms:

$$E[n] = T^{\text{KS}}[n] + E^{\text{H}}[n] + V[n] + T[n] - T^{\text{KS}}[n] + U[n] - E^{\text{H}}[n]. \quad (1.21)$$

$T[n]$ and $U[n]$ are some complicated functionals of the density, but we can sweep them under the proverbial rug by defining the exchange-correlation energy functional

as the last four terms and then the energy functional is written

$$E[n] = T^{\text{KS}}[n] + E^{\text{H}}[n] + V[n] + E^{\text{xc}}[n]. \quad (1.22)$$

The exchange-correlation energy functional $E^{\text{xc}}[n]$ contains all of the complicated, many-body information about the system. If this functional were known explicitly, then any groundstate density could be found from the minimum principle and thus any groundstate energy. Also, any other functional of the density that could be derived would yield groundstate expectation values of that quantity as well. Alas, no explicit functional of the density is known for a general inhomogeneous system (though a functional is known for the important case of a homogeneous system.) However, the situation is not completely hopeless. An explicit formula can be found for E^{xc} that is an implicit functional of n through the density-density correlation function. A derivation of this formula can be found in [Appendix B](#).

Leaving aside for now the explicit form of $E^{\text{xc}}[n]$ and treating it simply as an unknown functional of the density, it is possible to take the variation of the functional in [Eq. \(1.22\)](#) subject to the constraint and the Kohn-Sham eigenfunctions are normalized to 1:

$$\delta \left[E[n] - \sum_i \epsilon_i \int d^3x |\phi_i(\mathbf{x})|^2 \right] = 0, \quad (1.23)$$

where ϵ_i are suggestively denoted Lagrange multipliers. since the Kohn-Sham kinetic energy functional is an explicit functional of the Kohn-Sham eigenfunctions, it is simpler to carry out the variation of $E[n]$ via the eigenfunctions. This is equivalent to varying with respect to the density, because the eigenfunctions determine the density. The variation of $E[n]$ can be written as

$$\delta E[n] = \sum_i \int d^3x \left[\frac{\delta E}{\delta \phi_i(\mathbf{x})} \delta \phi_i(\mathbf{x}) + \frac{\delta E}{\delta \phi_i^*(\mathbf{x})} \delta \phi_i^*(\mathbf{x}) \right], \quad (1.24)$$

where the functional derivative is defined as

$$\int d^3x \frac{\delta E}{\delta \phi(\mathbf{x})} \delta \phi(\mathbf{x}) = \lim_{\lambda \rightarrow 0} \frac{E[\phi + \lambda \delta \phi] - E[\phi]}{\lambda} \quad (1.25)$$

The functional derivative of the Kohn-Sham kinetic energy functional can be found from this definition:

$$\begin{aligned} T^{\text{KS}}[\phi_i^* + \lambda \delta \phi_i^*, \dots] &= T^{\text{KS}}[\phi_i^*, \dots] - \lambda \frac{\hbar^2}{2m} \int d^3x \delta \phi_i^*(\mathbf{x}) \nabla^2 \phi_i(\mathbf{x}) \\ \implies \frac{\delta T^{\text{KS}}}{\delta \phi_i^*(\mathbf{x})} &= -\frac{\hbar^2}{2m} \nabla^2 \phi_i(\mathbf{x}). \end{aligned} \quad (1.26)$$

$\frac{\delta T^{\text{KS}}}{\delta \phi_i(\mathbf{x})}$ is given by the complex conjugate of this. For the remaining terms that are explicit functionals of the density, it is useful to use the chain rule:

$$\frac{\delta E}{\delta \phi_i^*(\mathbf{x})} = \int d^3x' \frac{\delta E}{\delta n(\mathbf{x}')} \frac{\delta n(\mathbf{x}')}{\delta \phi_i^*(\mathbf{x})} \quad (1.27)$$

The functional derivative of the density with respect to the eigenfunction is given by

$$\frac{\delta n(\mathbf{x}')}{\delta \phi_i^*(\mathbf{x})} = \sum_j \frac{\delta \phi_j^*(\mathbf{x}')}{\delta \phi_i^*(\mathbf{x})} \phi_j(\mathbf{x}') = \delta^3(\mathbf{x} - \mathbf{x}') \phi_i(\mathbf{x}), \quad (1.28)$$

following from linearity and the product rule. It is obvious that $\frac{\delta V}{\delta n(\mathbf{x})} = v(\mathbf{x})$. The functional derivative of the Hartree energy is given by

$$\int d^3x \frac{\delta E^{\text{H}}}{\delta n(\mathbf{x})} \delta n(\mathbf{x}) = \frac{e^2}{2} \int d^3x \int d^3x' \left[\frac{\delta n(\mathbf{x}) n(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} + \frac{n(\mathbf{x}) \delta n(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \right]. \quad (1.29)$$

Dummy variables can be switched yielding

$$\frac{\delta E^{\text{H}}}{\delta n(\mathbf{x})} = e^2 \int d^3x' \frac{n(\mathbf{x}')}{|\mathbf{x} - \mathbf{x}'|} \equiv v^{\text{H}}[n](\mathbf{x}). \quad (1.30)$$

Finally, it is only possible at this point to define an exchange-correlation potential as the functional derivative:

$$v^{\text{xc}}[n](\mathbf{x}) \equiv \frac{\delta E^{\text{xc}}}{\delta n(\mathbf{x})}. \quad (1.31)$$

Thus, the minimum condition on the variation of the total energy functional becomes

$$\sum_i \int d^3x \delta\phi_i^*(\mathbf{x}) \left[-\frac{\hbar^2}{2m} \nabla^2 + v(\mathbf{x}) + v^{\text{H}}[n](\mathbf{x}) + v^{\text{xc}}[n](\mathbf{x}) - \epsilon_i \right] \phi_i(\mathbf{x}) + \text{c.c.} = 0. \quad (1.32)$$

Since this must hold for all possible variations $\delta\phi_i(\mathbf{x})$, one must conclude that the following equation must hold:

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + v(\mathbf{x}) + v^{\text{H}}[n](\mathbf{x}) + v^{\text{xc}}[n](\mathbf{x}) - \epsilon_i \right] \phi_i(\mathbf{x}) = 0, \quad (1.33)$$

along with the conjugate equation.

Equation (1.33) has exactly the form of the eigenvalue equation for non-interacting electrons in an external potential. The conclusion is that if we solve the non-interacting problem with the potential

$$v^{\text{KS}}[n](\mathbf{x}) = v(\mathbf{x}) + v^{\text{H}}[n](\mathbf{x}) + v^{\text{xc}}[n](\mathbf{x}), \quad (1.34)$$

then the groundstate density we find is identical to that of the interacting system. As it should, the groundstate density also provides the groundstate energy. This becomes clearer by rewriting the Kohn-Sham kinetic energy functional as

$$\begin{aligned} T^{\text{KS}}[n] &= \sum_i^{\text{occ}} \int d^3x [\epsilon_i - v^{\text{KS}}[n](\mathbf{x})] \phi_i^*(\mathbf{x}) \phi_i(\mathbf{x}) \\ &= \sum_i^{\text{occ}} \epsilon_i - \int d^3x v^{\text{KS}}[n](\mathbf{x}) n(\mathbf{x}). \end{aligned} \quad (1.35)$$

Then the groundstate energy is given by

$$E_0 = \sum_i^{\text{occ}} \epsilon_i - E^{\text{H}}[n_0] - \int d^3x v^{\text{xc}}[n_0](\mathbf{x}) n_0(\mathbf{x}) + E^{\text{xc}}[n_0]. \quad (1.36)$$

Notice that, contrary to what many people unfamiliar with Kohn-Sham DFT assume, the groundstate energy is not simply the addition of the Kohn-Sham eigenvalues, as the groundstate energy would be for a true non-interacting system. This highlights that the Kohn-Sham system is merely a tool to find true many-body quantities and it has a complicated relationship to the true many-body system.

Throughout the above discussion I have ignored the spin of the electron. This is sufficient if the groundstate shares the spin-rotation symmetry of the original Hamiltonian. However, this need not be the case and this symmetry can be spontaneously broken. To allow for this possibility, it is best to distinguish the spin-up and spin-down densities, as is done in spin DFT (SDFT). As argued by Kohn, Savin and Ullrich [41] before the proofs in [42], for the vast majority of applications the conventional way of extending DFT to include spin are still on firm theoretical grounding, as I summarize in Appendix C. So, I will proceed to briefly present SDFT within the Kohn-Sham method while assuming differentiability of the energy functionals. Since the external potential will almost always be spin independent and the Hartree potential depends only on the total density, the main difference comes in recognizing that the exchange-correlation functional can depend on the spin-up and spin-down densities separately. Then the exchange-correlation potential becomes

$$v_{\sigma}^{\text{xc}}[n_{\uparrow}, n_{\downarrow}](\mathbf{x}) \equiv \frac{\delta E^{\text{xc}}[n_{\uparrow}, n_{\downarrow}]}{\delta n_{\sigma}(\mathbf{x})}, \quad (1.37)$$

where $\sigma = \uparrow, \downarrow$ is a spin index. This produces Kohn-Sham equations for each spin:

$$\left[-\frac{\hbar^2}{2m} \nabla^2 + v(\mathbf{x}) + v^{\text{H}}[n](\mathbf{x}) + v_{\sigma}^{\text{xc}}[n_{\uparrow}, n_{\downarrow}](\mathbf{x}) \right] \phi_{i\sigma}(\mathbf{x}) = \epsilon_{i\sigma} \phi_{i\sigma}(\mathbf{x}). \quad (1.38)$$

The possibility of symmetry breaking also means that $v_\sigma^{\text{xc}}[n_\uparrow, n_\downarrow](\mathbf{x})$ need not be invariant under the same symmetry operations as $v(\mathbf{x})$. Examples include anti-ferromagnetism, spin-density waves and charge-density waves. However, realistically doable calculations of macroscopic solids require some periodicity of the potential, so we shall restrict ourself to these cases. Thus, $v_\sigma^{\text{xc}}[n_\uparrow, n_\downarrow](\mathbf{x})$ will have periodic symmetry:

$$v_\sigma^{\text{xc}}[n_\uparrow, n_\downarrow](\mathbf{x} + \mathbf{R}) = v_\sigma^{\text{xc}}[n_\uparrow, n_\downarrow](\mathbf{x}), \quad (1.39)$$

where $\mathbf{R} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3$, n_1 , n_2 and n_3 being integers and $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ being primitive lattice vectors. The symmetry of the potential ensures that if $\phi(\mathbf{x})$ is a solution of Eq. (1.38), then it must also be a solution of the eigenvalue equation for the translation operator:

$$\hat{D}(\mathbf{R})\phi(\mathbf{x}) = \phi(\mathbf{x} + \mathbf{R}) = \lambda\phi(\mathbf{x}). \quad (1.40)$$

Of course, the eigenvalues of the translation operator are elementary [43]. Using the wave-vector $\mathbf{k}$ as a label, $\lambda_{\mathbf{k}} = e^{i\mathbf{k} \cdot \mathbf{R}}$. Thus, we can label the Kohn-Sham eigenfunctions by $\mathbf{k}$ and see that they satisfy the Bloch theorem [44]:

$$\hat{D}(\mathbf{R})\phi_{\mathbf{k}}(\mathbf{x}) = e^{i\mathbf{k} \cdot \mathbf{R}}\phi_{\mathbf{k}}(\mathbf{x}). \quad (1.41)$$

This allows $\phi_{\mathbf{k}}(\mathbf{x})$ to be written as the product

$$\phi_{\mathbf{k}}(\mathbf{x}) = e^{i\mathbf{k} \cdot \mathbf{R}}u_{\mathbf{k}}(\mathbf{x}), \quad (1.42)$$

where $u_{\mathbf{k}}(\mathbf{x} + \mathbf{R}) = u_{\mathbf{k}}(\mathbf{x})$, i.e. $u_{\mathbf{k}}(\mathbf{x})$ shares the periodicity of the lattice.

Without further information about $\mathbf{k}$, the expression for the Bloch function in Eq. (1.42) is not unique. From $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ it is possible to define vectors $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ via the condition

$$\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}. \quad (1.43)$$

Using $\mathbf{b}_1$, $\mathbf{b}_2$ and $\mathbf{b}_3$ it is also possible to define a set of reciprocal lattice vectors

$$\mathbf{G} = n_1\mathbf{b}_1 + n_2\mathbf{b}_2 + n_3\mathbf{b}_3, \quad n_1, n_2, n_3 \in \mathbb{Z}. \quad (1.44)$$

Then an equivalent Bloch function could be written

$$\phi_{\mathbf{k}}(\mathbf{x}) = e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{R}} e^{-i\mathbf{G}\cdot\mathbf{R}} u_{\mathbf{k}}(\mathbf{x}) = e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{R}} u_{\mathbf{k}+\mathbf{G}}(\mathbf{x}), \quad (1.45)$$

because $e^{-i\mathbf{G}\cdot\mathbf{R}}$ is also periodic upon translation by $\mathbf{R}$. So, it is convenient to adopt the conventional restriction that $\mathbf{k}$ should be closer to the origin than to any other reciprocal lattice vector. All of these $\mathbf{k}$ -vectors lie in what is defined as the first Brillouin zone and this choice guarantees a unique set of Bloch eigenfunctions.

From the Kohn-Sham eigenvalue equation it is possible to find an equation for $u_{\mathbf{k}}(\mathbf{x})$:

$$\left[-\frac{\hbar^2}{2m}(\mathbf{i}\mathbf{k} + \nabla)^2 + v_{\sigma}^{\text{KS}}[n_{\uparrow}, n_{\downarrow}](\mathbf{x}) \right] u_{\mathbf{k}\sigma}(\mathbf{x}) = \epsilon_{\mathbf{k}\sigma} u_{\mathbf{k}\sigma}(\mathbf{x}). \quad (1.46)$$

This equation too will have an eigenvalue spectrum, labeled by j : the band index. So,

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + v_{\sigma}^{\text{KS}}[n_{\uparrow}, n_{\downarrow}](\mathbf{x}) \right] \phi_{\mathbf{k}j\sigma}(\mathbf{x}) = \epsilon_{\mathbf{k}j\sigma} \phi_{\mathbf{k}j\sigma}(\mathbf{x}) \quad (1.47)$$

is the fully labeled Kohn-Sham equation in a periodic system. From Eq. (1.46), one can deduce a couple of insights that will help in approaching the solution to the Kohn-Sham equation. First, it is only necessary to find solutions for the equation at each $\mathbf{k}$ -point and this can be done independently from the solution at the other $\mathbf{k}$ -points. The final solution must be self-consistent, so integration over the Brillouin zone is required, but in each iteration on the way to a self-consistent solution each $\mathbf{k}$ -point is independent. Second, $u_{\mathbf{k}j\sigma}(\mathbf{x})$ is periodic and all the necessary information on the eigenfunction is contained within it, so the eigenvalue equation only needs to be solved in a single unit-cell with periodic boundary conditions.

1.2 The linearized augmented planewave basis

Solving the eigenvalue equation on a computer requires a basis. Formally complete bases include the position basis or planewave basis. On the computer, only a finite subset of these bases can be used, a real-space grid for the former and planewaves with wave-numbers below a maximum value for the latter. However, the requirement that the valence single-electron wavefunctions be orthogonal to the core wavefunctions means that the former have a rapidly varying value close to the atomic nuclei, as I will discuss in more detail below. In particular this means that a prohibitively large number of planewaves would be required to represent the valence wavefunctions. Thus, the planewave basis is always accompanied by a pseudopotential method that utilizes pseudo-wavefunctions that are smooth near in the atomic core region [45–50].

One can preserve an all-electron method and avoid the use of pseudopotentials. In practice, this is achieved by using bases that are not necessarily complete, but can provide a variational ansatz. Example methods include linearized muffin-tin orbitals (LMTO) [51–54] and linear combinations of atomic orbitals (LCAO) [55–57], though in practice pseudopotentials are also used in modern LCAO implementations [58]. Though LMTO and LCAO bases are able to represent wavefunctions that are atomic-like better than planewaves, they relinquish the ability to represent wavefunctions that planewaves are better suited for, such as those at high-energy. An attempt to maintain the utility of both the LMTO and planewave bases is linearized augmented planewave (LAPW) basis [11, 12, 59, 60]. This is actually a family of bases, but all seek to combine the benefits of muffin-tin orbitals near the atomic centers and planewaves in the space between atoms while remaining an all-electron method.

Here I will note that there is another way to retain the usefulness of all-electron wavefunctions and the computational utility of pseudopotentials, the projector augmented wave method [61, 62]. This method uses a linear transformation between all-electron and pseudo-wavefunctions to allow for the use of the latter for solving the Kohn-Sham equations while retaining access to the former. However, in this

dissertation I have used the LAPW basis, which remains the gold standard for accuracy [12].

Though historically the LAPW basis was derived [59, 60] from the nonlinear APW method of Slater [63, 64], I believe that it is instructive to examine the solution to the Kohn-Sham eigenvalue equations using a plane-wave basis before constructing the LAPW basis. Completeness of the plane-wave basis implies that the Bloch eigenfunctions are exactly representable as

$$\phi_{\mathbf{k}j\sigma}(\mathbf{x}) = \frac{1}{\sqrt{\Omega}} \sum_{\mathbf{G}} c_{\mathbf{k}j\sigma,\mathbf{G}} e^{-i(\mathbf{k}+\mathbf{G})\cdot\mathbf{x}}, \quad (1.48)$$

where Ω is the volume of the macro-crystal. Using this expression, the equation for the eigenfunctions becomes a matrix eigenvalue equation for the $c_{\mathbf{k}j\sigma,\mathbf{G}}$ coefficients:

$$\sum_{\mathbf{G}'} \left[\frac{\hbar^2}{2m} (\mathbf{k} + \mathbf{G})^2 \delta_{\mathbf{G}\mathbf{G}'} + v_{\sigma,\mathbf{G}-\mathbf{G}'}^{\text{KS}} \right] c_{\mathbf{k}j\sigma,\mathbf{G}'} = \epsilon_{\mathbf{k}j\sigma} c_{\mathbf{k}j\sigma,\mathbf{G}}, \quad (1.49)$$

where

$$v_{\sigma,\mathbf{G}-\mathbf{G}'}^{\text{KS}} = \frac{1}{\Omega} \int d^3x e^{i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{x}} v_{\sigma}^{\text{KS}}[n_{\uparrow}, n_{\downarrow}](\mathbf{x}) \quad (1.50)$$

is the lattice Fourier transform of the Kohn-Sham potential and its functional dependence on the spin densities are to be understood.

The problem with using planewaves as a basis is that the potential due to the nuclei goes like $1/r$ close to the nucleus locations. This means that the Fourier transform $v_{\sigma,\mathbf{G}-\mathbf{G}'}^{\text{KS}}$ decays slowly with increasing $|\mathbf{G} - \mathbf{G}'|$. This in turn means that the $c_{\mathbf{k}j\sigma,\mathbf{G}}$ with distantly separated values of $\mathbf{G}$ are coupled. The upshot is that the matrix that is to be diagonalized must be prohibitively large. Since the problem lies with the potential near the nucleus, the regions near the nuclei should be treated specially. Partitioning the space into muffin-tin spheres centered on the nuclei is the first step towards this treatment. The spheres do not cover the entire unit cell, the rest making up the interstitial region. So, a position is either in one of the muffin-tins, in which case I will write it as $\mathbf{x} = \mathbf{x}_{\alpha} + \mathbf{r}$, where $\mathbf{x}_{\alpha}$ is the position of nucleus α .

Otherwise, the position is in the interstitial, so it remains written as $\mathbf{x}$. The reason for introducing $\mathbf{r}$ is that this vector can be written in spherical coordinates with the nucleus at the origin.

Now, the introduction of $\mathbf{r}$ allows the expansion of planewaves in spherical harmonics:

$$\frac{1}{\sqrt{\Omega}} e^{i(\mathbf{k}+\mathbf{G})\cdot(\mathbf{x}_\alpha+\mathbf{r})} = \frac{4\pi}{\sqrt{\Omega}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{x}_\alpha} \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} i^\ell j_\ell(|\mathbf{k}+\mathbf{G}|r) Y_{\ell m}^*(\widehat{\mathbf{k}+\mathbf{G}}) Y_{\ell m}(\hat{\mathbf{r}}), \quad (1.51)$$

where j_ℓ is a spherical Bessel function and $Y_{\ell m}$ is a spherical harmonic. The object of performing this expansion is to find an expression for the Kohn-Sham eigenfunction inside the muffin-tin that does not require truncating the sum over $\mathbf{G}$. Instead, analytic summation over $\mathbf{G}$ produces a sum over ℓ and m that can be truncated. Upon summation over $\mathbf{G}$, the Kohn-Sham eigenfunction inside muffin-tin α becomes

$$\phi_{\mathbf{k}j\sigma}(\mathbf{x}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} \tilde{f}_{\mathbf{k}j\sigma}^{\alpha\ell m}(r) Y_{\ell m}(\hat{\mathbf{r}}), \quad \mathbf{x} \in \text{MT}_\alpha, \quad (1.52)$$

where $\tilde{f}_{\mathbf{k}j\sigma}^{\alpha\ell m}(r)$ is given formally by

$$\tilde{f}_{\mathbf{k}j\sigma}^{\alpha\ell m}(r) = \frac{4\pi}{\sqrt{\Omega}} \sum_{\mathbf{G}} c_{\mathbf{k}j\sigma,\mathbf{G}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{x}_\alpha} \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} i^\ell j_\ell(|\mathbf{k}+\mathbf{G}|r) Y_{\ell m}^*(\widehat{\mathbf{k}+\mathbf{G}}). \quad (1.53)$$

Actually using this formula would run into the same problem of truncation in $\mathbf{G}$ as before, so the goal of this exercise is to find some other expression for $\tilde{f}_{\mathbf{k}j\sigma}^{\alpha\ell m}(r)$.

If the Kohn-Sham equations had already been solved, then there would be a way to find $\tilde{f}_{\mathbf{k}j\sigma}^{\alpha\ell m}(r)$. Plugging Eq. (1.52) in the KS equation in Eq. (1.47), one finds a set

of equations for $\tilde{f}_{\mathbf{k}j\sigma}^{\alpha\ell m}(r)$:

$$-\frac{\hbar^2}{2mr^2} \frac{d}{dr} \left(r^2 \frac{d\tilde{f}_{\mathbf{k}j\sigma}^{\alpha\ell m}}{dr} \right) + \frac{\hbar^2 \ell(\ell+1)}{2mr^2} \tilde{f}_{\mathbf{k}j\sigma}^{\alpha\ell m}(r) + \sum_{\substack{\ell'm' \\ \ell''m''}} v_{\sigma, \alpha \ell' m'}^{\text{KS}}(r) \left[\int d\Omega Y_{\ell m}^*(\hat{\mathbf{r}}) Y_{\ell' m'}(\hat{\mathbf{r}}) Y_{\ell'' m''}(\hat{\mathbf{r}}) \right] \tilde{f}_{\mathbf{k}j\sigma}^{\alpha\ell m}(r) = \epsilon_{\mathbf{k}j\sigma} \tilde{f}_{\mathbf{k}j\sigma}^{\alpha\ell m}(r), \quad (1.54)$$

where $\int d\Omega$ is an integral over solid angles and the KS potential has been expanded in spherical harmonics:

$$v_{\sigma}^{\text{KS}}[n_{\uparrow}, n_{\downarrow}](\mathbf{x}) = \sum_{\ell=0}^{\infty} \sum_{m=-\ell}^{\ell} v_{\sigma, \alpha \ell' m'}^{\text{KS}}(r) Y_{\ell m}(\hat{\mathbf{r}}), \quad \mathbf{x} \in \text{MT}_{\alpha}. \quad (1.55)$$

The boundary condition on this equation is that $\phi_{\mathbf{k}j\sigma}(\mathbf{x})$ (and perhaps, but not necessarily its first derivatives) must be continuous at the surface of the muffin-tin sphere. Again, if we had a solution to the Kohn-Sham equations, we could solve these equations. But this would imply that we already know $\tilde{f}_{\mathbf{k}j\sigma}^{\alpha\ell m}(r)$! So, this is not the way forward.

Instead, I will side-step all of these issues and simply make an ansatz for $\tilde{f}_{\mathbf{k}j\sigma}^{\alpha\ell m}(r)$. The hope is that a reasonable ansatz should be good enough, because variational methods are often quite forgiving. We can be assured that, given the limitations of the ansatz, we will find the closest solution to the exact one. So, in place of the complicated eigenvalue problem of Eq. (1.54), I will make use of the following, much simpler equation:

$$-\frac{\hbar^2}{2mr^2} \frac{d}{dr} \left(r^2 \frac{df_{\alpha\ell}}{dr} \right) + \frac{\hbar^2 \ell(\ell+1)}{2mr^2} f_{\alpha\ell}(r) + v_{\sigma, \alpha 00}^{\text{KS}}(r) f_{\alpha\ell}(r) = E_L f_{\alpha\ell}(r). \quad (1.56)$$

In this equation, the Kohn-Sham eigenvalue has been replaced by a parameter E_L called the linearization energy. In addition, $f_{\alpha\ell}(r)$ represents a solution only in the presence of the spherical part of the Kohn-Sham potential inside the muffin-tin. The non-spherical parts have not disappeared, because they will be retained in

the solution of the eigenvalue equations using this ansatz. The combination of the $f_{\alpha\ell}(r)$ radial functions in the muffin-tins, the planewaves in the interstitial and the boundary condition of continuity define the APW functions. The hope is that these radial functions will offer sufficient variational freedom while allowing considerable simplification of the equations that must be solved.

However, the radial functions that are solutions of Eq. (1.56) do not offer sufficient variational freedom alone. In the original formulation of LAPW, the radial functions $\frac{df_{\alpha\ell}(r;E_L)}{dr}$ were included to increase variational freedom and linear combinations between these and the $f_{\alpha\ell}(r;E_L)$ were sought that made the APW functions continuously differentiable across the boundary [59]. However, these functions differ significantly from those that are found in a non-linear APW method [65] and so have some trouble representing the valence states. In a method introduced by Sjöstedt, Nordström and Singh [65], another set of local orbitals are added to the basis set. These are formed by linear combinations of $f_{\alpha\ell}(r;E_L)$ and $\frac{df_{\alpha\ell}(r;E_L)}{dr}$ chosen such that the orbitals vanish at the muffin tin boundary. Using an APW basis with this addition reduced the size of the basis needed for convergence [11, 65]. This new basis has been dubbed “APW+LO+lo,” “LO” for the semi-core local orbitals and “lo” for the valence local orbitals.

So, in order to have sufficient variational freedom in the basis used to expand the Kohn-Sham Bloch states, I must include local orbitals for both the semi-core and valence. Inclusion of local orbitals amounts to including more radial functions additional to the APW radial functions given by Eq. (1.56). The radial functions for the local orbitals are constructed from the APW radial function, its first derivative and an APW radial function for a semi-core linearization energy for LO local orbitals. Orthogonality of the local orbitals to the other LAPW basis functions applicable to both the core and semi-core is proved in Appendix D. Rather than enumerate the APW functions and their derivatives, I will enumerate the radial functions after linear combinations have been taken to satisfy the boundary conditions for the local orbitals. The index λ will indicate this enumeration, running from 1 to $N_\lambda(\alpha)$ depending on

the type of atom in muffin tin α , and for concreteness the APW radial function will be given by $\lambda = 1$.

Also, since the Exciting Plus code [13] only uses the APW+LO+lo basis to calculate the spin independent Kohn-Sham equations as a first variational step with spin being included as a second variational step [66], I will drop the spin index in what follows. In that case, the Kohn-Sham eigenfunctions in the APW+LO+lo variational basis are given by:

$$\phi_{\mathbf{k}j}(\mathbf{x}) = \sum_{\mathbf{G}} c_{\mathbf{k}j,\mathbf{G}} \zeta_{\mathbf{k}+\mathbf{G}}(\mathbf{x}) + \sum_{\alpha\ell m} \sum_{\lambda=2}^{N_{\lambda}(\alpha)} c_{\mathbf{k}j,\lambda\alpha\ell m} \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \zeta_{\lambda\alpha\ell m}(\mathbf{x} - \mathbf{R}) \quad (1.57)$$

There are two types of functions in this expansion. The first are the APW functions

$$\zeta_{\mathbf{k}+\mathbf{G}}(\mathbf{x}) = \begin{cases} \frac{1}{\sqrt{\Omega}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{x}}, & \mathbf{x} \in \text{I} \\ \sum_{\ell=0}^{\ell_{\max}} \sum_{m=-\ell}^{\ell} A_{\mathbf{k}+\mathbf{G}}^{\alpha\ell m} f_{\lambda=1,\alpha\ell}(r) Y_{\ell m}(\hat{\mathbf{r}}), & \mathbf{x} = \mathbf{x}_{\alpha} + \mathbf{r} \in \text{MT}_{\alpha} \end{cases} \quad (1.58)$$

where “I” denotes the interstitial. The matching coefficients $A_{\mathbf{k}+\mathbf{G}}^{\alpha\ell m}$ are not independent and are instead given by

$$f_{\alpha\ell}(R_{\alpha}) A_{\mathbf{k}+\mathbf{G}}^{\alpha\ell m} = \frac{4\pi}{\sqrt{\Omega}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{x}_{\alpha}} i^{\ell} j_{\ell}(|\mathbf{k} + \mathbf{G}| R_{\alpha}) Y_{\ell m}^*(\widehat{\mathbf{k} + \mathbf{G}}), \quad (1.59)$$

where R_{α} is the radius of muffin-tin α . This equation only has a non-trivial solution if $f_{\alpha\ell}(R_{\alpha}) \neq 0$. This expansion will be cut off at some finite number of $\mathbf{G}$ ’s and it turns out that far fewer $\mathbf{G}$ ’s are required than if the expansion were done in planewaves. The second type of functions in the expansion Eq. (1.57) are the local orbitals:

$$\zeta_{\mathbf{k},\lambda\alpha\ell m}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \text{I} \\ f_{\lambda\alpha\ell}(r) Y_{\ell m}(\hat{\mathbf{r}}), & \mathbf{x} = \mathbf{x}_{\alpha} + \mathbf{r} \in \text{MT}_{\alpha} \end{cases} \quad (1.60)$$

However, there is a complication that I have implicitly introduced by allowing the APW basis functions and valence local orbitals to have discontinuous first derivatives. The problem lies with the variational expression for T^{KS} that I wrote in Eq. (1.19). If I use that expression when the orbitals have the possibility of containing a discontinuous first derivative, then the adjoint of the Kohn-Sham equations in Eq. (1.33) no longer satisfy the condition for the energy to be a minimum embodied in Eq. (1.32). This is because in writing that the complex conjugate of the displayed expression in Eq. (1.32) is all that remains of the minimization condition, I implicitly used Green's second identity. Namely, I used

$$\int_V d^3x \phi_i^*(\mathbf{x}) \nabla^2 \phi_i(\mathbf{x}) = \int_V d^3x \phi_i(\mathbf{x}) \nabla^2 \phi_i^*(\mathbf{x}) + \int_{\partial V} dS [\phi_i^*(\mathbf{x}) \nabla \phi_i(\mathbf{x}) \cdot \mathbf{n} - \phi_i(\mathbf{x}) \nabla \phi_i^*(\mathbf{x}) \cdot \mathbf{n}], \quad (1.61)$$

where $\int_{\partial V} dS$ is the surface integral over the boundary of V , and assumed that the directional derivatives $\nabla \phi_i(\mathbf{x}) \cdot \mathbf{n}$ for any bounded region V that I wished to choose were continuous. However, with the APW+LO+lo basis this is no longer true when that region is one of the muffin-tin spheres.

Schlosser and Marcus [67] offered a solution which repairs the variational expression for the energy. Adapted for the current case [12], the Kohn-Sham kinetic energy functional must now read

$$T^{\text{KS}}[n] = \frac{\hbar^2}{2m} \sum_i^{\text{occ}} \left\{ - \int d^3x \phi_i^*(\mathbf{x}) \nabla^2 \phi_i(\mathbf{x}) + \sum_{\alpha} \int_{\text{MT}_{\alpha}} dS \phi_i^*(\mathbf{x}) \left[\frac{\partial \phi_i}{\partial r} \Big|_{\text{MT}} - \frac{\partial \phi_i}{\partial r} \Big|_{\text{I}} \right] \right\}. \quad (1.62)$$

In practical computation, such as in the Exciting Plus code [13] I use in this dissertation, a more symmetric form is preferred in the interstitial region. Using

Green's first identity in the interstitial, I can rewrite Eq. (1.62) as

$$T^{\text{KS}}[n] = \frac{\hbar^2}{2m} \sum_i^{\text{occ}} \left\{ - \sum_{\alpha} \int_{\text{MT}_{\alpha}} d^3x \phi_i^*(\mathbf{x}) \nabla^2 \phi_i(\mathbf{x}) + \int_{\text{I}} d^3x \nabla \phi_i^*(\mathbf{x}) \cdot \nabla \phi_i(\mathbf{x}) \right. \\ \left. + \sum_{\alpha} \int_{\text{MT}_{\alpha}} dS \phi_i^*(\mathbf{x}) \frac{\partial \phi_i}{\partial r} \Big|_{\text{MT}} \right\}. \quad (1.63)$$

Written this way, the surface kinetic energy contribution can be easily computed from the radial functions $f_{\lambda\alpha\ell}(r)$. In particular, since I am using the symmetric form in the interstitial, I do not have to worry about computing the radial derivatives approaching the boundary from the interstitial side.

In deriving the elements of the matrix that must be diagonalized in order to find the Kohn-Sham eigenfunctions and eigenvalues, it is useful to view the $c_{\mathbf{k}j,\mathbf{G}}$ and $c_{\mathbf{k}j,\lambda\alpha\ell m}$ coefficients as elements of a vector $[\mathbf{c}_{\mathbf{k}j}^{\text{APW}}, \mathbf{c}_{\mathbf{k}j}^{\text{LO}}]^{\top}$. Then the Kohn-Sham Hamiltonian can be written as a block matrix

$$\mathbf{H}^{\text{KS}}(\mathbf{k}) = \begin{bmatrix} \mathbf{H}^{\text{APW,APW}}(\mathbf{k}) & \mathbf{H}^{\text{APW,LO}}(\mathbf{k}) \\ \mathbf{H}^{\text{LO,APW}}(\mathbf{k}) & \mathbf{H}^{\text{LO,LO}}(\mathbf{k}) \end{bmatrix}. \quad (1.64)$$

In Appendix E, I present some explicit expressions for these matrix elements. Though these matrix elements are significantly more complex than those for the planewave basis, the final matrix is small enough to allow practical computation without the use of pseudopotentials. This makes the APW+LO+lo basis both practical and accurate.

With the eigenvalue equation in the APW+LO+lo basis in hand the Kohn-Sham eigenvalues and eigenfunctions can be computed. These are not just useful for calculating groundstate properties, but they can also be used to calculate the two-particle excitation spectrum using TDDFT linear response, which I will cover in the next section.

1.3 Linear response in time-dependent density-functional theory

The equivalent in TDDFT to the Hohenberg-Kohn theorem in ground state DFT is the Runge-Gross theorem [4]. This theorem states that, given some conditions that are often benign, the time-dependent density evolving under the influence of a time-dependent potential coupled to the density can only be produced by that potential or one differing from it by only a purely time-dependent function. In addition, any other density can only be produced by a potential differing by more than such a time-dependent function.

That is, consider the expectation value of the density in an initial state:

$$n(\mathbf{x}, t_0) = \langle \Psi(t_0) | \hat{n}(\mathbf{x}) | \Psi(t_0) \rangle \quad (1.65)$$

and evolve this expectation value forward in time by the Hamiltonian

$$\hat{H}(t) = \hat{T} + \hat{W} + \int d^3x v(\mathbf{x}, t) \hat{n}(\mathbf{x}). \quad (1.66)$$

If we use the time-evolution operator generated by $\hat{H}(t)$, $\hat{U}(t, t_0)$, the time-dependent density can be written in the Schrödinger picture using the time-evolved state $|\Psi(t)\rangle = \hat{U}(t, t_0)|\Psi(t_0)\rangle$ as

$$n(\mathbf{x}, t) = \langle \Psi(t) | \hat{n}(\mathbf{x}) | \Psi(t) \rangle. \quad (1.67)$$

Now consider the same initial density, but evolve it with the Hamiltonian

$$\hat{H}'(t) = \hat{T} + \hat{W} + \int d^3x v'(\mathbf{x}, t) \hat{n}(\mathbf{x}) \quad (1.68)$$

yielding

$$n'(\mathbf{x}, t) = \langle \Psi'(t) | \hat{n}(\mathbf{x}) | \Psi'(t) \rangle, \quad (1.69)$$

where $|\Psi'(t)\rangle = \hat{U}'(t, t_0)|\Psi(t_0)\rangle$ and $\hat{U}'(t, t_0)$ is the time-evolution operator generated by $\hat{H}'(t)$. The conclusion of the Runge-Gross theorem is that if $n'(\mathbf{x}, t) = n(\mathbf{x}, t)$, then $v'(\mathbf{x}, t) = v(\mathbf{x}, t) + c(t)$, and if $n'(\mathbf{x}, t) \neq n(\mathbf{x}, t)$, then $v'(\mathbf{x}, t) \neq v(\mathbf{x}, t) + c(t)$.

Thus, this theorem proves the existence of a one-to-one correspondence between time-dependent densities and equivalent time-dependent potentials. Of course, this theorem is most useful if the initial state $|\Psi(t_0)\rangle$ is the ground state before the potential's time-dependence is turned on. Then the Hohenberg-Kohn theorem gives us the one-to-one correspondence between static potentials and densities. Then we consider the evolution of the ground state density as the potential varies in time. I present a proof of the Runge-Gross Theorem in Appendix F and show how this proof can be extended to prove the van Leeuwen theorem [68]. The latter leads to the time-dependent analog of the Kohn-Sham equations.

The equation governing the time dependence of the Kohn-Sham Bloch states is

$$\left[i\hbar \frac{\partial}{\partial t} + \frac{\hbar^2}{2m} \nabla^2 - v_{\text{KS}}[n](\mathbf{x}, t) \right] \phi_{\mathbf{k}j\sigma}(\mathbf{x}, t) = 0, \quad (1.70)$$

where the time-dependent Kohn-Sham potential contains three terms, as in ground-state DFT:

$$v_{\text{KS}}(\mathbf{x}, t) = v(\mathbf{x}, t) + v_{\text{H}}[n](\mathbf{x}, t) + v_{\text{xc}}[n](\mathbf{x}, t). \quad (1.71)$$

The first is the external potential which includes both the static (in our case crystal) potential and an arbitrary time-dependent driving potential. The second is the Hartree potential due to a time-dependent density:

$$v_{\text{H}}[n](\mathbf{x}, t) = e^2 \int d^3x' \frac{n(\mathbf{x}', t)}{|\mathbf{x} - \mathbf{x}'|}. \quad (1.72)$$

The last term is the time-dependent analog of $v_{\text{xc}}[n](\mathbf{x})$ in groundstate Kohn-Sham DFT that contains all effects of exchange and correlation. However, one must take care not to take this analogy too far. In the groundstate formalism, the definition of the exchange-correlation potential was given in terms of the functional derivative

of the exchange-correlation energy. The analogous definition for the time-dependent case is complicated by the need to respect the causality of the density response to a perturbation. This is resolved by defining an exchange-correlation action on the Keldysh contour [69].

Vital to this dissertation is the connection between the theoretical results presented within TDDFT and spectroscopic experimental results. In order to draw this connection I must be able to calculate how the system responds to external probes. In addition, I must be able to handle time-dependent probes as all spectroscopic measurements will be made at certain frequencies or energies. Often it is possible to make the approximation that the external probe is weak enough that the response of the system only depends linearly on the external probe. In this case, the behavior of the system under the perturbation is described by the response function, which can be calculated using only the knowledge of the system being probed and does not depend on the details of the external probe.

The general theory of linear response was laid out on the landmark paper of Kubo [70], with which one can find expressions for the response of virtually any quantity to virtually any probe that is weak enough. That work focused on systems at finite temperature, but the formalism is easily applied to zero temperature and standard derivations of linear response are widely available. I will follow to some extent the textbook presentation in Fetter and Walecka [71], which I have adapted to TDDFT linear response. However, I must draw a clear distinction between the standard derivations of linear response and linear response in TDDFT. Standard derivations are generally oriented to many-body perturbation theory which relies on diagrammatic expansions of the quantities derived in linear response theory. On the other hand, TDDFT relies on the density, external potential and exchange-correlation functional as basic quantities.

Of course, this opens the question of just how to apply the formulation of the theory of linear response, which has no knowledge of DFT or TDDFT, to calculations performed using the latter theory. Using the eigenvalues and eigenfunctions from

Kohn-Sham DFT, it is straightforward to make a reasonable, if not entirely rigorous, proposal as to what the proper expressions for response functions are. This was first done for the optical response of semiconductor surfaces by Ando [72, 73] and later for the resonant photoemission of Ba and Ce by Zangwill and Soven [74]. However, the more formal presentation for the density response of Gross and Kohn [14] sets the origin of time for TDDFT linear response. And later the work of Petersilka, Gossmann and Gross [75] put the theory on a rigorous foundation, almost 20 years after the work of Ando. Regardless of the historical path, I will present a derivation of linear response that I hope will be familiar to those readers with little experience with TDDFT and highlight the elements peculiar to TDDFT along the way.

TDDFT linear response primarily focuses on the time-dependent response on the density to an external probe that couples to the density. Thus I will start by considering the time evolution of the electron density, which can be found from the time evolution of the Kohn-Sham noninteracting system. Since the Kohn-Sham groundstate is a Slater determinant of the occupied Kohn-Sham orbitals, the time evolution of this ground state is also given by Eq. (1.70). Thus, an explicit expression for the time-dependent density in the Schrödinger picture would be

$$n(\mathbf{x}, t) = \langle \Phi_{0\text{KS}}(t) | \hat{n}(\mathbf{x}) | \Phi_{0\text{KS}}(t) \rangle. \quad (1.73)$$

An important special case is to find an expression for the density when the time-dependent part of the external potential is in some sense small. The idea is to split all of the small time-dependent terms away from the static terms which define the groundstate Kohn-Sham DFT problem: $n(\mathbf{x}, t) = n_0(\mathbf{x}) + n_1(\mathbf{x}, t)$, $v(\mathbf{x}, t) = v_0(\mathbf{x}) + v_1(\mathbf{x}, t)$, $v_{\text{H}}[n](\mathbf{x}, t) = v_{\text{H}0}[n](\mathbf{x}) + v_{\text{H}1}[n](\mathbf{x}, t)$ and $v_{\text{xc}}[n](\mathbf{x}, t) = v_{\text{xc}0}[n](\mathbf{x}) + v_{\text{xc}1}[n](\mathbf{x}, t)$.

The time evolution of the Kohn-Sham groundstate in the absence of the time-dependent perturbation can be written in terms of a Kohn-Sham Hamiltonian:

$$|\Phi_{0\text{KS}}(t)\rangle = e^{-\frac{i}{\hbar}\hat{H}_{\text{KS}0}(t-t_0)}|\Phi_{0\text{KS}}(t_0)\rangle. \quad (1.74)$$

I will write the Hamiltonian in second-quantized notation [71] as

$$\hat{H}_{\text{KS}0} = \int d^3x \left[\frac{\hbar}{2m} \hat{\psi}^\dagger(\mathbf{x}) \nabla^2 \hat{\psi}(\mathbf{x}) + v_{\text{KS}0}(\mathbf{x}) \hat{n}(\mathbf{x}) \right], \quad (1.75)$$

where $\hat{n}(\mathbf{x}) = \hat{\psi}^\dagger(\mathbf{x})\hat{\psi}(\mathbf{x})$ is the electron density operator in second quantization. Since the perturbation depends on time, no such exponentiation of a Hamiltonian is possible. However, I will make the ansatz that the time-evolution operator in the presence of the perturbation factors as one would expect, into the form

$$\hat{U}(t, t_0) = e^{-\frac{i}{\hbar}\hat{H}_{\text{KS}0}(t-t_0)}\hat{U}_1(t, t_0). \quad (1.76)$$

Though straightforward exponentiation is not possible, I can make use of the differential equation for $|\Phi_{0\text{KS}}(t)\rangle$:

$$i\hbar \frac{\partial}{\partial t} |\Phi_{0\text{KS}}(t)\rangle = [\hat{H}_{\text{KS}0} + \hat{H}_{\text{KS}1}(t)] |\Phi_{0\text{KS}}(t)\rangle, \quad (1.77)$$

where

$$\hat{H}_{\text{KS}1}(t) = \int d^3x [v_1(\mathbf{x}, t) + v_{\text{H}1}[n](\mathbf{x}, t) + v_{\text{xc}1}[n](\mathbf{x}, t)] \hat{n}(\mathbf{x}) = \int d^3x v_{\text{KS}1}[n](\mathbf{x}, t) \hat{n}(\mathbf{x}). \quad (1.78)$$

By writing $|\Phi_{0\text{KS}}(t)\rangle$ in terms of the time-evolution operator, I find that

$$\begin{aligned} & \left[\hat{H}_{\text{KS}0} e^{-\frac{i}{\hbar}\hat{H}_{\text{KS}0}t} \hat{U}_1(t, 0) + e^{-\frac{i}{\hbar}\hat{H}_{\text{KS}0}t} i\hbar \frac{\partial}{\partial t} \hat{U}_1(t, 0) \right] |\Phi_{0\text{KS}}(0)\rangle \\ &= \left[\hat{H}_{\text{KS}0} e^{-\frac{i}{\hbar}\hat{H}_{\text{KS}0}t} \hat{U}_1(t, 0) + \hat{H}_{\text{KS}1}(t) e^{-\frac{i}{\hbar}\hat{H}_{\text{KS}0}t} \hat{U}_1(t, 0) \right] |\Phi_{0\text{KS}}(0)\rangle. \end{aligned} \quad (1.79)$$

This equation is satisfied if $\hat{U}_1(t, t_0)$ in turn satisfies the following

$$i\hbar \frac{\partial}{\partial t} \hat{U}_1(t, t_0) = \hat{H}_{\text{KS}1(\text{I})}(t) \hat{U}_1(t, t_0), \quad (1.80)$$

where I have made the definition

$$\hat{H}_{\text{KS}1(\text{I})}(t) \equiv e^{\frac{i}{\hbar} \hat{H}_{\text{KS}0} t} \hat{H}_{\text{KS}1}(t) e^{-\frac{i}{\hbar} \hat{H}_{\text{KS}0} t} \quad (1.81)$$

and I have used the fact that $\hat{U}_1(t, t_0) = \hat{U}_1(t, 0) \hat{U}_1(0, t_0)$. Solving Eq. (1.80) iteratively and keeping only terms linear in $\hat{H}_{\text{KS}1}(t)$, I find an approximate solution for $\hat{U}_1(t, t_0)$:

$$\hat{U}_1(t, t_0) \approx \hat{1} - \frac{i}{\hbar} \int_{t_0}^t dt' \hat{H}_{\text{KS}1(\text{I})}(t'). \quad (1.82)$$

With the time-evolution operator in hand, I can find an expression for the time-dependent density:

$$\begin{aligned} n(\mathbf{x}, t) &= \langle \Phi_{0\text{KS}}(t_0) | \hat{U}(t_0, t) \hat{n}(\mathbf{x}) \hat{U}(t, t_0) | \Phi_{0\text{KS}}(t_0) \rangle \\ &= \langle \Phi_{0\text{KS}}(t_0) | e^{\frac{i}{\hbar} \hat{H}_{\text{KS}0}(t-t_0)} \hat{n}(\mathbf{x}) e^{-\frac{i}{\hbar} \hat{H}_{\text{KS}0}(t-t_0)} | \Phi_{0\text{KS}}(t_0) \rangle \\ &\quad + \frac{i}{\hbar} \langle \Phi_{0\text{KS}}(t_0) | \int_{t_0}^t dt' \hat{H}_{\text{KS}1(\text{I})}(t') e^{\frac{i}{\hbar} \hat{H}_{\text{KS}0}(t-t_0)} \hat{n}(\mathbf{x}) e^{-\frac{i}{\hbar} \hat{H}_{\text{KS}0}(t-t_0)} | \Phi_{0\text{KS}}(t_0) \rangle \\ &\quad - \frac{i}{\hbar} \langle \Phi_{0\text{KS}}(t_0) | e^{\frac{i}{\hbar} \hat{H}_{\text{KS}0}(t-t_0)} \hat{n}(\mathbf{x}) e^{-\frac{i}{\hbar} \hat{H}_{\text{KS}0}(t-t_0)} \int_{t_0}^t dt' \hat{H}_{\text{KS}1(\text{I})}(t') | \Phi_{0\text{KS}}(t_0) \rangle \\ &= n_0(\mathbf{x}) + \frac{i}{\hbar} \int_{t_0}^t dt' \langle \Phi_{0\text{KS}}(0) | [\hat{H}_{\text{KS}1(\text{I})}(t'), \hat{n}_{(I)}(\mathbf{x}, t)]_- | \Phi_{0\text{KS}}(0) \rangle, \end{aligned} \quad (1.83)$$

where I have used the fact that the density is static in the absence of the perturbation and I have made the definition

$$\hat{n}_{(I)}(\mathbf{x}, t) \equiv e^{\frac{i}{\hbar} \hat{H}_{\text{KS}0} t} \hat{n}(\mathbf{x}) e^{-\frac{i}{\hbar} \hat{H}_{\text{KS}0} t}. \quad (1.84)$$

From Eq. (1.83), I can identify the time-dependent response of the density to the perturbation in the *Kohn-Sham potential*:

$$n_1(\mathbf{x}, t) = -\frac{i}{\hbar} \int_{t_0}^t dt' \int d^3x' \langle \Phi_{0\text{KS}}(0) | [\hat{n}_{(I)}(\mathbf{x}, t), \hat{n}_{(I)}(\mathbf{x}', t')]_- | \Phi_{0\text{KS}}(0) \rangle v_{\text{KS}1}[n](\mathbf{x}', t'). \quad (1.85)$$

The Kohn-Sham density response function is defined as the integral kernel that relates the linear response of the density to a perturbation in the Kohn-Sham potential:

$$n_1(\mathbf{x}, t) = \int_{-\infty}^{\infty} dt' \int d^3x' \chi^{\text{KS}}(\mathbf{x}t, \mathbf{x}'t') v_{\text{KS}1}[n](\mathbf{x}', t') \quad (1.86)$$

Then, Eq. (1.85) provides an equation that can be used to calculate this response function:

$$\chi^{\text{KS}}(\mathbf{x}t, \mathbf{x}'t') = -\frac{i}{\hbar} \Theta(t - t') \langle \Phi_{0\text{KS}} | [\hat{n}_{(I)}(\mathbf{x}, t), \hat{n}_{(I)}(\mathbf{x}', t')]_- | \Phi_{0\text{KS}} \rangle, \quad (1.87)$$

where I have made the identification $|\Phi_{0\text{KS}}(0)\rangle = |\Phi_{0\text{KS}}\rangle$.

A definition similar to Eq. (1.86) applies to the true density response function, that is the integral kernel relating the linear response of the density to a perturbation in the *external potential*:

$$n_1(\mathbf{x}, t) = \int_{-\infty}^{\infty} dt' \int d^3x' \chi(\mathbf{x}t, \mathbf{x}'t') v_1(\mathbf{x}', t'). \quad (1.88)$$

This density response is what is actually measured experimentally, but the Kohn-Sham density response function is much easier to calculate using Eq. (1.87). What we would like is to find an equation relating the two response functions. The trick is to relate the perturbation to the Kohn-Sham potential with the perturbation to the external potential. In order to achieve this, I can use the fact provided by the Runge-Gross theorem that there is a one-to-one correspondence between the density and potential along with the definition for the functional derivative to write $v_{\text{KS}1}[n](\mathbf{x}, t)$

in terms of the perturbation of the external potential and the density response:

$$\begin{aligned}
v_{\text{KS1}}[n](\mathbf{x}, t) &= v_1(\mathbf{x}, t) + \int_{-\infty}^{\infty} dt' \int d^3x' \left[\frac{\delta v_{\text{H}}(\mathbf{x}, t)}{\delta n(\mathbf{x}', t')} + \frac{\delta v_{\text{xc}}(\mathbf{x}, t)}{\delta n(\mathbf{x}', t')} \right] n_1(\mathbf{x}', t') \\
&= v_1(\mathbf{x}, t) + \int_{-\infty}^{\infty} dt' \int d^3x' \left[\frac{e^2}{|\mathbf{x} - \mathbf{x}'|} \delta(t - t') + f^{\text{xc}}(\mathbf{x}t, \mathbf{x}'t') \right] n_1(\mathbf{x}', t'),
\end{aligned} \tag{1.89}$$

where I have defined the exchange-correlation kernel $f^{\text{xc}}(\mathbf{x}t, \mathbf{x}'t') \equiv \frac{\delta v_{\text{xc}}(\mathbf{x}, t)}{\delta n(\mathbf{x}', t')}$.

Folded up into f^{xc} are all of the complicated many-body effects due to interaction between particles and holes contained in the density response function in Eq. (1.88). Now it is a simple matter to plug Eq. (1.88) into Eq. (1.89) and then Eq. (1.89) into Eq. (1.86). Comparing the result to Eq. (1.88) allows us to identify the TDDFT linear response equation:

$$\begin{aligned}
\chi(\mathbf{x}t, \mathbf{x}'t') &= \chi^{\text{KS}}(\mathbf{x}t, \mathbf{x}'t') \\
&+ \int dt_1 d^3x_1 dt_2 d^3x_2 \chi^{\text{KS}}(\mathbf{x}t, \mathbf{x}_1t_1) \left[\frac{e^2}{|\mathbf{x}_1 - \mathbf{x}_2|} \delta(t_1 - t_2) + f^{\text{xc}}(\mathbf{x}_1t_1, \mathbf{x}_2t_2) \right] \chi(\mathbf{x}_2t_2, \mathbf{x}'t').
\end{aligned} \tag{1.90}$$

This equation has a similar form to the Dyson equation for the density response in diagrammatic perturbation theory, but there are important differences that distinguish linear response in TDDFT. First, the Kohn-Sham response occupies the place that would be held by the polarization in perturbation theory. The polarization has a complicated diagrammatic expansion which contains all effects beyond RPA, i.e. those involving particle-hole coupling. In contrast, the Kohn-Sham response needs no such expansion to be calculated and is simply given in the same form as the RPA polarization by Eq. (1.87). Second, f^{xc} does not appear in the Dyson equation in perturbation theory, where only the bare Coulomb interaction is present, and all of the many-body effects that are beyond RPA are contained in it. So, what TDDFT does versus diagrammatic perturbation theory (with the Kohn-Sham system used

as the non-interacting reference system) is take all of the effects contained in the polarization beyond the non-interacting term and place them into f^{xc} .

More useful for comparing to what is actually measured experimentally is the Fourier transform of the response function. Defining this Fourier transform involves recognizing two symmetries. First, is the fact that the response function is unchanged if both time variables are translated so it is a function only of the difference in time. Second, the response function is unchanged if both spatial variables are translated by *the same* lattice vector. These symmetries are summarized by the relations

$$\chi(\mathbf{x}t, \mathbf{x}'t') = \chi(\mathbf{x}, \mathbf{x}'; t - t') = \chi(\mathbf{x} + \mathbf{R}, \mathbf{x}' + \mathbf{R}; t - t'). \quad (1.91)$$

These symmetries apply equally to the Kohn-Sham density response and imply that either can be written in terms of Fourier components:

$$\chi(\mathbf{x}t, \mathbf{x}'t') = \frac{1}{\Omega} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega(t-t')} \sum_{\mathbf{q}} \sum_{\mathbf{G}\mathbf{G}'}^{\text{BZ}} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} \chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{x}'}, \quad (1.92)$$

where $\mathbf{q}$ is restricted to the first Brillouin zone. Eq. (1.92) is the inverse double Fourier transform. The forward transform is given by

$$\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) = \frac{1}{\Omega} \int_{-\infty}^{\infty} d(t-t') \int_{\Omega} d^3x d^3x' e^{i\omega(t-t')} e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} \chi(\mathbf{x}t, \mathbf{x}'t') e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{x}'}. \quad (1.93)$$

Thus, the more relevant response equation is given by transforming Eq. (1.90) according to Eq. (1.93):

$$\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) = \chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}}(\mathbf{q}, \omega) + \sum_{\mathbf{G}_1\mathbf{G}_2} \chi_{\mathbf{G}\mathbf{G}_1}^{\text{KS}}(\mathbf{q}, \omega) \left[\frac{4\pi e^2}{|\mathbf{q} + \mathbf{G}_1|^2} \delta_{\mathbf{G}_1\mathbf{G}_2} + f_{\mathbf{G}_1\mathbf{G}_2}^{\text{xc}}(\mathbf{q}, \omega) \right] \chi_{\mathbf{G}_2\mathbf{G}'}(\mathbf{q}, \omega), \quad (1.94)$$

where $f_{\mathbf{G}_1\mathbf{G}_2}^{\text{xc}}(\mathbf{q}, \omega)$ is given by the same definition as in Eq. (1.93).

The response function describes the charge density fluctuations of the system. Via the fluctuation-dissipation theorem, the imaginary part of this function will produce

the excitation spectrum of any scattering experiment coupled to the density. This includes both hard X-ray scattering, which couples to the density via the diamagnetic term in the Hamiltonian when electromagnetic fields are included (called Thompson scattering), and electron energy-loss spectroscopy, which couples via the electron-electron interaction term. Of particular importance to this dissertation is the optical response of the system. However, a distinction must be made between optical response and e.g. X-ray scattering. In X-ray scattering the observed response is linear in the *external* electromagnetic field: the applied X-rays. In optics, the observed response is linear in the *self-consistent* electromagnetic field. That is the sum of both the external, applied field and the fields produced by the charge and currents in the material once the system has achieved a steady state. The quantity of interest in this case is the so-called “macroscopic” dielectric function. This quantity can also be found from the density response in Eq. (1.94). The expression is [5]

$$\epsilon(\omega) = 1 - \lim_{\mathbf{q} \rightarrow \mathbf{0}} \frac{4\pi e^2}{|\mathbf{q}|^2} \bar{\chi}_{00}(\mathbf{q}, \omega), \quad (1.95)$$

where $\bar{\chi}$ satisfies a modified version of Eq. (1.94) with $v_{\mathbf{G}}(\mathbf{q}) = \frac{4\pi e^2}{|\mathbf{q} + \mathbf{G}|^2}$ replaced by

$$\bar{v}_{\mathbf{G}}(\mathbf{q}) = \begin{cases} 0 & \mathbf{G} = \mathbf{0} \\ \frac{4\pi e^2}{|\mathbf{q} + \mathbf{G}|^2} & \mathbf{G} \neq \mathbf{0} \end{cases}. \quad (1.96)$$

Equation (1.95) still requires the inversion of a matrix equation in reciprocal space in order find $\bar{\chi}_{00}(\mathbf{q}, \omega)$. This is a reflection of the physical reality that the self-consistent field and responding density can have variations on the scale of the unit cell even though the perturbing field has an effectively infinite wavelength. If these short length-scale variations are strong enough, they can lead to crystal local-field effects in the response of the system. It is often true that in the optical limit these effects are small, but this must always be checked numerically.

If crystal local-field effects are small enough to be negligible, then $\chi_{\mathbf{G}\mathbf{G}'}$ is approximately diagonal. In this case, and in the RPA, Eq. (1.95) is approximated by

$$\epsilon(\omega) = 1 - \lim_{\mathbf{q} \rightarrow \mathbf{0}} \frac{4\pi e^2}{|\mathbf{q}|^2} \chi_{\mathbf{00}}^{\text{KS}}(\mathbf{q}, \omega). \quad (1.97)$$

This equation only requires the calculation of $\chi_{\mathbf{00}}^{\text{KS}}$, which has a considerably simpler form. In Appendix G, I derive an expression for the Kohn-Sham response, which is given by

$$\begin{aligned} \chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}}(\mathbf{q}, \omega) = \lim_{\eta \rightarrow 0} \frac{1}{\Omega} \sum_{\sigma} \sum_{\mathbf{k}} \sum_{jj'}^{\text{BZ}} & \langle \mathbf{k}, j, \sigma | e^{-i(\mathbf{q}+\mathbf{G}) \cdot \mathbf{x}} | \mathbf{k} + \mathbf{q}, j', \sigma \rangle \frac{f_{\mathbf{k}, j, \sigma} - f_{\mathbf{k}+\mathbf{q}, j', \sigma}}{\epsilon_{\mathbf{k}, j, \sigma} - \epsilon_{\mathbf{k}+\mathbf{q}, j', \sigma} + \hbar(\omega + i\eta)} \\ & \times \langle \mathbf{k} + \mathbf{q}, j', \sigma | e^{i(\mathbf{q}+\mathbf{G}') \cdot \mathbf{x}} | \mathbf{k}, j, \sigma \rangle, \end{aligned} \quad (1.98)$$

where $|\mathbf{k}, j, \sigma\rangle$ is a Bloch state. The form of χ^{KS} is a sum over particle-hole channels. In Eq. (1.98) the particles and holes are Bloch states, which are not very useful if we want to analyze the spectrum produced by χ^{KS} in terms of localized degrees of freedom. However, the basic form of χ^{KS} and a sum over channels will become especially important when I present linear response in the Wannier basis in Section 1.6. But first, I must introduce the Wannier basis, which I will do in the next section.

1.4 The Wannier basis

The ability to construct local orbitals that fully respect the Kohn-Sham DFT electronic structure is vital to the work in this dissertation. As I have shown in Section 1.1, the electronic structure in Kohn-Sham DFT is most naturally represented in terms of Bloch functions that are spread throughout the macro-crystal. That is, if an electron is placed in a Bloch state and the question is asked, “What is the probability of finding the electron in a particular location?” the answer is that the

probability is roughly equal no matter where that location is in the macro-crystal. If I wish to shed light on the role of local chemistry and symmetry in the physics of materials and construct models with interactions defined in real space, then I must have access to orbitals that are contained to a particular region. That is, an electron placed in such an orbital would almost certainly be found within a small ($\sim 1 a_0$) radius of a particular atom. These local orbitals are the Wannier functions and under certain conditions they are a unitary transformation from a subset the Bloch basis. Though the Wannier orbitals are concentrated around a single atom, I will show in Chapters 2, 3 and 4 that significant long-ranged tails that do spread throughout the macro-crystal are important for both optical excitations and single-particle hopping.

In this section, I will begin by giving a brief review of the history of Wannier functions and the various approaches to constructing them from the Bloch states. Key to defining the transformation between the two bases is the choice of gauge and I will present the method of choice used in this dissertation: the projection method, particularly that detailed in Anisimov et al. [2]. I have extended this method to include projection using a linear combination of local orbitals. This will prove a critical development for the ability it gives to construct molecular orbitals that are crucial in the optimization of the Wannier basis for Cu_2O in Chapter 2 and FeTe in Chapter 4.

Wannier functions have an almost 80 year history dating back to the foundational paper of Wannier [1]. They were used by Slater in the same year in the presentation of his theory of ferromagnetism [76] and later in his treatment of the perturbed periodic potential [77]. Löwdin was the first to note connections between his orthonormalized atomic orbitals and the orbitals of Wannier [78], which in turn has a close connection to the method used to construct Wannier functions in this dissertation. More than 20 years after the original presentation of Wannier, Kohn laid out in detail the formal mathematical properties of Wannier functions for a one-dimensional crystal, including the proof of the exponential decay of the orbitals [79]. Des Cloizeaux later generalized the formal theory of Wannier functions to two- and three-dimensions, focusing on

the connection between the symmetries in the space group and the transformation properties of the orbitals [80]. However, he left the exponential localization problem unaddressed. This problem is considerably more complex beyond one-dimension [3], but it does not pose any practical limitation on the work in this dissertation.

However, localization of Wannier orbitals, and in particular the ability of the imposition of localization to yield a method that provides a unique (though not the *only*) choice of gauge for the transformation from a subset of Bloch states to Wannier orbitals, led to Marzari and Vanderbilt developing “maximally localized” Wannier functions 60 years after Wannier’s initial paper [81]. This led to an explosion in the use of Wannier functions and they have gone from a little used tool to a standard part of solid state physics. This includes the particular application of Wannier construction to calculations using the LAPW basis [82–84]. Though maximally localized Wannier functions yield one route to a unique transformation, it is not the only method of constructing a Wannier basis that has wide spread use. In particular, the projection method is favored by many [85–88]. Ku et al. highlighted the importance of the Wannier basis embodying the crystal symmetries and the projection method is a practical way of respecting these symmetries [86]. Symmetry also plays a key role in other methods for Wannier construction [89–91] and will be of central importance in this dissertation, which utilizes the projection method.

No matter what method of construction is used, the expression for the transformation from a set of Bloch states $\psi_{j\mathbf{k}}(\mathbf{x})$ to a set of Wannier orbitals $w_n(\mathbf{x} - \mathbf{R})$ is given by

$$w_{\nu\sigma}(\mathbf{x} - \mathbf{R}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}}^{\text{BZ}} e^{-i\mathbf{k}\cdot\mathbf{R}} \sum_j a_{\nu j}(\mathbf{k}) \phi_{\mathbf{k}j\sigma}(\mathbf{x}), \quad (1.99)$$

where ν is a composite index that includes α , ℓ and m . α labels the sites of the atoms, ℓ is the total angular momentum quantum number and in practical calculations m labels real harmonic functions. Orthonormality and completeness of the Wannier orbitals requires that the $a_{nj}(\mathbf{k})$ coefficients are elements of a unitary matrix at each point $\mathbf{k}$. In the projection method, the initial set of Wannier functions formed using

trial local orbitals is not orthonormal. Orthonormalization is done as a second step via the Löwdin method [78]. This procedure still allows for the Wannier basis to span a space that does not coincide with a subspace spanned by some subset of the Bloch basis. That is, the Wannier orbitals can and often are orthonormal without being complete. Completeness requires that there are the same number of Wannier orbitals as Bloch states within a certain energy range for every $\mathbf{k}$ in the first Brillouin zone.

In all practical situations, the band index j is a subset of all the Bloch bands. In most cases this subset is defined via an energy window. Conventionally, this is either a user-defined energy range and all Bloch states with eigenvalue sitting in that range for a particular $\mathbf{k}$ are included, or a fixed range of bands such that a number for the lower band and one for the upper band are defined as input parameters and all bands in between are included. The bands indices are arranged in ascending order according to energy eigenvalue. Of course, picking an energy window or a band range are not the only ways to define the subset of bands for construction. A band or several bands can be excluded in that range according to some criterion or the band range for each $\mathbf{k}$ can be determined algorithmically. I have had to utilize the latter type of definition in the construction of optimized orbitals, as I will describe in Chapter 4. The selection of the band range is the most critical factor in determining and constraining what type of Wannier orbitals will be produced. They will largely determine the degree of localization and the chemistry that the orbitals embody in real space. But for now I will take the definition of the band range as given and discuss the projection method.

I'll begin by defining our pre-orthogonalized Wannier functions using a projection of the Bloch states associated with bands within some energy interval:

$$\tilde{w}_{\nu\sigma}(\mathbf{x} - \mathbf{R}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}}^{\text{BZ}} e^{-i\mathbf{k}\cdot\mathbf{R}} \sum_{j=N_1(\mathbf{k})}^{N_2(\mathbf{k})} \langle \phi_{\mathbf{k}j\sigma} | \varphi_{\nu} \rangle \phi_{\mathbf{k}j\sigma}(\mathbf{x}) \quad (1.100)$$

The bands included in the sum depend on $\mathbf{k}$ because a band may be inside the desired energy interval for some wave vectors and not for others. φ_{ν} is a local orbital,

introduced in Section 1.2, with a radial part determined by the Schrödinger equation with a fixed energy parameter given in Eq. (1.56) and the boundary condition that the function must go to zero at the muffin tin radius and an angular part equal to a real harmonic in a coordinate system with the nucleus at its center. These local orbitals can be expressed explicitly as

$$\varphi_{\nu=\{\alpha\lambda\ell m\}}(\mathbf{r}) = \sum_{m_1=-\ell}^{\ell} f_{\alpha\lambda\ell}(r) \mathcal{R}_{\ell;mm_1} Y_{\ell m_1}(\hat{\mathbf{r}}), \quad (1.101)$$

where $\mathcal{R}$ transforms spherical harmonics with a global z -axis into real harmonics in a local coordinate system which is in general rotated with respect to the global coordinate system separately for each site α and λ is the index in the APW+LO+lo basis that chooses the “lo” local orbitals. Whereas the global coordinate system is defined with respect to the crystallographic axes of the unit cell, the local coordinate system is chosen to respect the local chemistry and symmetry around each atom. In general, the local axes will differ from site to site. This will become particularly important when I must define a chemistry-respecting z^2 orbital in Cu_2O by rotating the z -axis, as I describe in Chapter 2.

Using Eq. (1.101) expansion one can find the inner product in Eq. (1.100):

$$\begin{aligned} \langle \phi_{\mathbf{k}j\sigma} | \varphi_{\nu=\{\alpha\lambda\ell m\}} \rangle &= \int d^3r \sum_{\ell'm'\lambda'} C_{\mathbf{k}j\sigma}^{*\alpha\lambda'\ell'm'} f_{\alpha\lambda'\ell'}(r) Y_{\ell'm'}^*(\hat{\mathbf{r}}) \sum_{m_1=-\ell}^{\ell} f_{\alpha\lambda\ell}(r) \mathcal{R}_{\ell;mm_1} Y_{\ell m_1}(\hat{\mathbf{r}}) \\ &= \sum_{\ell'm'\lambda} \sum_{m_1=-\ell}^{\ell} C_{\mathbf{k}j\sigma}^{*\alpha\lambda'\ell'm'} \mathcal{R}_{\ell;mm_1} \int_0^{R_{\text{MT}}} dr r^2 f_{\alpha\lambda'\ell'}(r) f_{\alpha\lambda\ell}(r) \int d\Omega Y_{\ell'm'}^*(\hat{\mathbf{r}}) Y_{\ell m_1}(\hat{\mathbf{r}}) \\ &= \sum_{\lambda'} \sum_{m_1=-\ell}^{\ell} C_{\mathbf{k}j\sigma}^{*\alpha\lambda'\ell m_1} \mathcal{R}_{\ell;mm_1} \int_0^{R_{\text{MT}}} dr r^2 f_{\alpha\lambda'\ell}(r) f_{\alpha\lambda\ell}(r) \end{aligned} \quad (1.102)$$

Since the various radial functions in the APW+LO+lo basis are not necessarily orthogonal, the integrals $\int_0^{R_{\text{MT}}} dr r^2 f_{\alpha\lambda'\ell}(r) f_{\alpha\lambda\ell}(r)$ must be calculated numerically, which is part of the Exciting Plus [13] LAPW code that I have used.

Now we must orthogonalize these Wannier functions which means we must examine the quantity

$$\begin{aligned}
\langle \tilde{w}_\nu(\mathbf{R}) | \tilde{w}_{\nu'}(\mathbf{R}') \rangle &= \int d^3r \tilde{w}_\nu^*(\mathbf{r} - \mathbf{R}) \tilde{w}_{\nu'}(\mathbf{r} - \mathbf{R}') \\
&= \int d^3x \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} e^{i\mathbf{k}\cdot\mathbf{R}} e^{-i\mathbf{k}'\cdot\mathbf{R}'} \sum_{j=N_1(\mathbf{k})}^{N_2(\mathbf{k})} \langle \varphi_\nu | \phi_{\mathbf{k}j\sigma} \rangle \phi_{\mathbf{k}j\sigma}^*(\mathbf{x}) \sum_{j'=N_1(\mathbf{k}')}^{N_2(\mathbf{k}')} \langle \phi_{\mathbf{k}'j'\sigma'} | \varphi_{\nu'} \rangle \phi_{\mathbf{k}'j'\sigma'}(\mathbf{x}) \\
&= \frac{1}{N} \sum_{\mathbf{k}\mathbf{k}'} e^{i\mathbf{k}\cdot\mathbf{R}} e^{-i\mathbf{k}'\cdot\mathbf{R}'} \sum_{j=N_1(\mathbf{k})}^{N_2(\mathbf{k})} \langle \varphi_\nu | \phi_{\mathbf{k}j\sigma} \rangle \sum_{j'=N_1(\mathbf{k}')}^{N_2(\mathbf{k}')} \langle \phi_{\mathbf{k}'j'\sigma'} | \varphi_{\nu'} \rangle \int d^3x \phi_{\mathbf{k}j\sigma}^*(\mathbf{x}) \phi_{\mathbf{k}'j'\sigma'}(\mathbf{x}) \\
&= \frac{1}{N} \sum_{\mathbf{k}} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} \sum_{j=N_1(\mathbf{k})}^{N_2(\mathbf{k})} \langle \varphi_\nu | \phi_{\mathbf{k}j\sigma} \rangle \langle \phi_{\mathbf{k}j\sigma} | \varphi_{\nu'} \rangle.
\end{aligned} \tag{1.103}$$

I will proceed with orthogonalization using the Löwdin procedure [2, 78]. First, the overlap matrix is defined by

$$O_{\nu\nu',\sigma}(\mathbf{k}) \equiv \sum_{j=N_1(\mathbf{k})}^{N_2(\mathbf{k})} \langle \varphi_\nu | \phi_{\mathbf{k}j\sigma} \rangle \langle \phi_{\mathbf{k}j\sigma} | \varphi_{\nu'} \rangle \tag{1.104}$$

Using the overlap matrix, one can define an inverse square root matrix:

$$S_{\nu\nu',\sigma}(\mathbf{k}) \equiv [O_{\nu\nu'}(\mathbf{k})]^{-1/2}. \tag{1.105}$$

Critically, the overlap matrix is a Hermitian matrix, which is obvious from the definition in Eq. (1.104), and thus the inverse square root matrix is also Hermitian. Using these definitions, new Wannier functions are defined as

$$w_{\nu\sigma}(\mathbf{x} - \mathbf{R}) = \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}} \sum_{\nu'} S_{\nu'\nu,\sigma}(\mathbf{k}) \sum_{j=N_1(\mathbf{k})}^{N_2(\mathbf{k})} \langle \phi_{\mathbf{k}j\sigma} | \varphi_{\nu'} \rangle \phi_{\mathbf{k}j\sigma}(\mathbf{x}) \tag{1.106}$$

The reason for making all of these definitions is to produce orthonormalized Wannier functions, so let us check to make sure that our new Wannier functions give the desired

result. The evaluation of the inner product of two of the Wannier functions defined in Eq. (1.106) proceeds in much the same way as the calculation in Eq. (1.103):

$$\begin{aligned}
\langle w_{\nu\sigma}(\mathbf{R}) | w_{\nu'\sigma}(\mathbf{R}') \rangle &= \frac{1}{N} \sum_{\mathbf{k}} e^{i\mathbf{k} \cdot (\mathbf{R} - \mathbf{R}')} \sum_{j=N_1(\mathbf{k})}^{N_2(\mathbf{k})} \sum_{\nu''} S_{\nu''\nu,\sigma}^*(\mathbf{k}) \langle \varphi_{\nu''} | \phi_{\mathbf{k}j\sigma} \rangle \sum_{\nu'''} S_{\nu'''\nu',\sigma}(\mathbf{k}) \langle \phi_{\mathbf{k}j\sigma} | \varphi_{\nu'''} \rangle \\
&= \frac{1}{N} \sum_{\mathbf{k}} e^{i\mathbf{k} \cdot (\mathbf{R} - \mathbf{R}')} \sum_{\nu''\nu'''} S_{\nu''\nu',\sigma}(\mathbf{k}) S_{\nu'''\nu',\sigma}(\mathbf{k}) \sum_{j=N_1(\mathbf{k})}^{N_2(\mathbf{k})} \langle \varphi_{\nu''} | \phi_{\mathbf{k}j\sigma} \rangle \\
&= \frac{1}{N} \sum_{\mathbf{k}} e^{i\mathbf{k} \cdot (\mathbf{R} - \mathbf{R}')} \sum_{\nu''\nu''',\sigma} S_{\nu''\nu',\sigma}(\mathbf{k}) O_{\nu''\nu''',\sigma}(\mathbf{k}) S_{\nu'''\nu',\sigma}(\mathbf{k}) \\
&= \frac{1}{N} \sum_{\mathbf{k}} e^{i\mathbf{k} \cdot (\mathbf{R} - \mathbf{R}')} [\mathbf{SOS}]_{\nu\nu',\sigma}(\mathbf{k}) = \frac{1}{N} \sum_{\mathbf{k}} e^{i\mathbf{k} \cdot (\mathbf{R} - \mathbf{R}')} [\mathbf{SS}^{-1}\mathbf{S}^{-1}\mathbf{S}]_{\nu\nu',\sigma}(\mathbf{k}) \\
&= \frac{1}{N} \sum_{\mathbf{k}} e^{i\mathbf{k} \cdot (\mathbf{R} - \mathbf{R}')} \delta_{\nu\nu'} = \delta_{\mathbf{R}\mathbf{R}'} \delta_{\nu\nu'},
\end{aligned} \tag{1.107}$$

where I have used boldface to denote matrices. Thus, making the definition in Eq. (1.106) produces Wannier functions that are orthonormal.

Of central importance to orthonormalizing the Wannier functions was the definition of the inverse square root matrix in Eq. (1.105). This was a formal definition with the inverse square root defined using the conventional definition of a function of a matrix by its Taylor series. Of course, this formal definition is not at all useful for computation. Thus, the next task is to compute the inverse square root matrix in such a way that it can be practically applied. This can be done by diagonalizing the overlap matrix and taking the inverse square root of the eigenvalues. This is because a function of a diagonal matrix is given by applying the function to each element of

the matrix. The proof is as follows:

$$\begin{aligned}
&\text{Let } \mathbf{S} = \mathbf{O}^{-1/2} \text{ denote } (\mathbf{S}^{-1})^2 = (\mathbf{S}^2)^{-1} = \mathbf{O} \\
&\text{Let } \mathbf{O} = \mathbf{U}\mathbf{D}\mathbf{U}^\dagger, \text{ then } [\mathbf{S}, \mathbf{O}] = 0 \implies \mathbf{S} = \mathbf{U}\mathbf{D}'\mathbf{U}^\dagger \\
&\quad \mathbf{U}\mathbf{D}'\mathbf{U}^\dagger = \mathbf{S} = \mathbf{O}^{-1/2} = [\mathbf{U}\mathbf{D}\mathbf{U}^\dagger]^{-1/2} \\
&\quad [\mathbf{U}\mathbf{D}'\mathbf{U}^\dagger\mathbf{U}\mathbf{D}'\mathbf{U}^\dagger]^{-1} = \mathbf{U}\mathbf{D}\mathbf{U}^\dagger \\
&\quad [\mathbf{U}\mathbf{D}'^2\mathbf{U}^\dagger]^{-1} = \mathbf{U}\mathbf{D}\mathbf{U}^\dagger \\
&\quad \mathbf{U}(\mathbf{D}'^2)^{-1}\mathbf{U}^\dagger = \mathbf{U}\mathbf{D}\mathbf{U}^\dagger \implies \mathbf{D}' = \mathbf{D}^{-1/2} \\
&\therefore \mathbf{S} = \mathbf{U}\mathbf{D}^{-1/2}\mathbf{U}^\dagger
\end{aligned} \tag{1.108}$$

The discussion of Wannier functions thus far has been in the context of the single-site projection method. In many places in this dissertation we will encounter compounds and situations for which an atomic-sphere-centric, ionic picture is not natural and instead the local degrees of freedom appropriate for understanding the physics are molecular in nature. So, rather than be restricted to projecting onto local orbitals such that each Wannier function is associated with one particular orbital, I would instead like to have the freedom to project onto a linear combination of local orbitals. I would like any such linear combination to preserve the orthonormality of the set of Wannier functions that we have constructed thus far. Hence, I wish to find a unitary transformation:

$$w_{\nu\sigma}^{\text{LC}}(\mathbf{r} - \mathbf{R}) = \sum_{\nu'\mathbf{R}'} \mathcal{U}_{\nu\mathbf{R}\nu'\mathbf{R}'} w_{\nu'\sigma}(\mathbf{r} - \mathbf{R}'). \tag{1.109}$$

The specific form of this transformation would be dictated by the chemistry of the molecular orbital that one is trying to construct. An admissible orbital would have to respect the chemistry and symmetry of the compound. In both Cu_2O as I describe in Chapter 2 and the iron-based superconductors as I describe in Chapter 4, the chemistry of cation-cation bonding is vitally important and the molecular

orbitals that I construct are linear combinations that reflect bonding and anti-bonding configurations of single-site projected orbitals.

Writing the transformation as in Eq. (1.109) is simple enough, but what I would like to have for practical computation is a projection onto a *linear combination of local orbitals* in order to produce our set of Wannier functions *rather than taking a linear combination after the projection has already been performed*. It turns out that some care must be taken in finding the appropriate linear combination of local orbitals. Let us start by plugging Eq. (1.106) into Eq. (1.109):

$$w_{\nu\sigma}^{\text{LC}}(\mathbf{x} - \mathbf{R}) = \sum_{\nu'\mathbf{R}'} \mathcal{U}_{\nu\mathbf{R}\nu'\mathbf{R}'} \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}'} \sum_{\nu''} S_{\nu''\nu',\sigma}(\mathbf{k}) \sum_{j=N_1(\mathbf{k})}^{N_2(\mathbf{k})} \langle \phi_{\mathbf{k}j\sigma} | \varphi_{\nu''} \rangle \phi_{\mathbf{k}j\sigma}(\mathbf{x}) \quad (1.110)$$

First, it is convenient to write the elements of the unitary transformation as a product of two terms:

$$\mathcal{U}_{\nu\mathbf{R}\nu'\mathbf{R}'} = \mathcal{U}_{\nu\nu'} \mathcal{U}_{\mathbf{R}-\mathbf{R}'}, \quad (1.111)$$

where I have explicitly enforced periodicity by writing $\mathbf{R} - \mathbf{R}'$. If both $\mathcal{U}_{\nu\nu'}$ and $\mathcal{U}_{\mathbf{R}-\mathbf{R}'}$ are elements of a unitary matrix then $\mathcal{U}_{\nu\mathbf{R}\nu'\mathbf{R}'}$ will still form a unitary matrix. With some simple manipulation using the unitarity of $\mathcal{U}_{\nu\nu'}$, I can arrive at the following:

$$\begin{aligned} w_{\nu\sigma}^{\text{LC}}(\mathbf{x} - \mathbf{R}) &= \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}} \sum_{\nu'\nu''\nu'''} \mathcal{U}_{\nu''\nu'''}^* S_{\nu''\nu',\sigma}(\mathbf{k}) \mathcal{U}_{\nu\nu'} \\ &\quad \times \sum_{j=N_1(\mathbf{k})}^{N_2(\mathbf{k})} \langle \phi_{\mathbf{k}j\sigma} | \sum_{\nu''''\mathbf{R}'} e^{i\mathbf{k}\cdot(\mathbf{R}-\mathbf{R}')} \mathcal{U}_{\mathbf{R}-\mathbf{R}'} \mathcal{U}_{\nu''''\nu'''} | \varphi_{\nu''''} \rangle \phi_{\mathbf{k}j\sigma}(\mathbf{x}) \quad (1.112) \\ &= \frac{1}{\sqrt{N}} \sum_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}} \sum_{\nu'} S_{\nu'\nu,\sigma}^{\text{LC}}(\mathbf{k}) \sum_{j=N_1(\mathbf{k})}^{N_2(\mathbf{k})} \langle \phi_{\mathbf{k}j\sigma} | \varphi_{\nu'}^{\text{LC}}(\mathbf{k}) \rangle \phi_{\mathbf{k}j\sigma}(\mathbf{x}), \end{aligned}$$

where I have reordered the dummy indices and made the change of variables $\mathbf{R}' \rightarrow \mathbf{R}' + \mathbf{R}$ to produce the last line.

In Eq. (1.112) I have made a couple of identifications in order to make it clear that this equation is of the same form as the definition in Eq. (1.106). First, the inverse

square root matrix is transformed:

$$S_{\nu'\nu,\sigma}^{\text{LC}}(\mathbf{k}) \equiv \sum_{\nu''\nu'''} \mathcal{U}_{\nu'''\nu'}^* S_{\nu'''\nu'',\sigma}(\mathbf{k}) \mathcal{U}_{\nu\nu''}. \quad (1.113)$$

Since the $\mathcal{U}_{\nu\nu'}$ form a unitary matrix, this is clearly just a change of basis transformation and I can identify the new matrix as the inverse square root matrix for the linear combination Wannier functions. Second, the linear combination of local orbitals that yields the linear combination of Wannier functions in Eq. (1.109) is given by

$$|\varphi_{\nu}^{\text{LC}}(\mathbf{k})\rangle \equiv \sum_{\mathbf{R}'\nu'} e^{-i\mathbf{k}\cdot\mathbf{R}'} \mathcal{U}_{\mathbf{0}-\mathbf{R}'} \mathcal{U}_{\nu\nu'} |\varphi_{\nu'}\rangle = \sum_{\mathbf{R}'\nu'} e^{-i\mathbf{k}\cdot\mathbf{R}'} \mathcal{U}_{\nu\mathbf{0}\nu'\mathbf{R}'} |\varphi_{\nu'}\rangle. \quad (1.114)$$

The expression in the last line of Eq. (1.112) is superior to that in Eq. (1.110) because the form of Eq. (1.112) is nearly identical to Eq. (1.106). That means that the same procedure and algorithms for constructing single-site projected orbitals can be used for projection using a linear combination of local orbitals. Also, in Eq. (1.110) the matrix $\mathcal{U}_{\nu\mathbf{R}\nu'\mathbf{R}'}$ must be exactly unitary to maintain orthonormality. Though formally the same matrix appears in Eq. (1.114), the practical definition of the linear combination of local orbitals can be considerably more liberal. In particular, the matrix need not necessarily be unitary and can even be rectangular, as long as the orbitals can be orthonormalized. I have found this to be the case for both Cu_2O in Chapter 2 and the iron-based compounds in Chapter 4.

I would also like to draw attention to a couple of technical points that are important for the computation of the Wannier functions defined by Eq. (1.112). In addition to the unitary transformation used in Eq. (1.109), the linear combination of local orbitals picks up a $\mathbf{k}$ dependent phase that is essential for obtaining a true Wannier function. However, since in the algorithm that I am following from the Exciting Plus code [13] the construction of the Wannier functions is performed independently at each $\mathbf{k}$ point, this $\mathbf{k}$ dependent phase does not introduce any additional complication. Also important is that for practical computation it

unnecessary to have the $S_{\nu'\nu,\sigma}(\mathbf{k})$ before obtaining the $S_{\nu'\nu,\sigma}^{\text{LC}}(\mathbf{k})$. One can simply form the linear combination in Eq. (1.114) and obtain the $S_{\nu'\nu,\sigma}^{\text{LC}}(\mathbf{k})$ via the same Löwdin procedure used in the single-site projection.

Now that I have presented the projection method, I believe that it is necessary in the remainder of this section to discuss the completeness of the Wannier basis. I have already stated that if the $a_{\nu j}(\mathbf{k})$ in Eq. (1.99) are elements of a unitary matrix at each $\mathbf{k}$, then the resulting orbitals are orthonormal and complete. This is elementary to show. The question is, under what conditions does the projection method, which is a way of constructing the $a_{\nu j}(\mathbf{k})$, produce orbitals that are orthonormal and complete. First, what I mean by completeness is that the Wannier basis is complete over the bands onto which they are projected. Bloch states for bands outside of the energy window for projection will obviously not be any linear combination of Wannier orbitals. Second, I was able to show that if the $S_{\nu\nu',\sigma}(\mathbf{k})$ matrix exists for each $\mathbf{k}$, then the resulting Wannier basis is orthonormal. As is clear from Eq. (1.108), the necessary and sufficient condition for the existence of $S_{\nu\nu',\sigma}(\mathbf{k})$ is that $O_{\nu\nu',\sigma}(\mathbf{k})$ does not have any zero eigenvalues. A zero eigenvalue in $O_{\nu\nu',\sigma}(\mathbf{k})$ would indicate that there is some linear dependency among the Wannier orbitals.

However, assuming that the Löwdin procedure works and I am able to produce an orthonormal set of Wannier orbitals, there is no guarantee that the resulting $a_{\nu j}(\mathbf{k})$ will be a unitary matrix. There are three cases to examine having to do with the relative number of ν 's, which I will denote as $\mathcal{N}$, and the number of j 's onto which I am projecting, which I will denote as $\mathcal{M}$. The final point, which I will show carefully in detail below, is that if $\mathcal{M} = \mathcal{N}$ then the Wannier basis will be complete over the subspace of bands onto which they are projected.

Now I will proceed to show that this is true by taking advantage of mathematical results from linear algebra. First, I will note that, if I have an orthonormal set of Wannier orbitals, then the rows of $a_{\nu j}(\mathbf{k})$ will form an orthonormal basis. Since there are $\mathcal{N}$ rows, this basis will span an $\mathcal{N}$ -dimensional space. Since the Bloch states are

solutions of the Kohn-Sham equations, they will also form an orthonormal basis, so they will span an $\mathcal{M}$ -dimensional space.

If $\mathcal{M} < \mathcal{N}$, then the row space of $a_{\nu j}(\mathbf{k})$ will have a smaller dimension than the column space. Since I am assuming that the Wannier orbitals are projected onto the bands under consideration, the row space will be a subspace of the column space. Then orthonormality of the Bloch states will be preserved and the rank of $a_{\nu j}(\mathbf{k})$ will be equal to $\mathcal{M}$. However, by the rank-nullity, this means that the left nullspace of $a_{\nu j}(\mathbf{k})$ has dimension $\mathcal{N} - \mathcal{M} \neq 0$. This in turn implies that there must be some linear dependency among the rows $a_{\nu j}(\mathbf{k})$, i.e. the Wannier orbitals will be linear dependent. This contradicts my assumption that the Löwdin procedure worked. Therefore, if the Löwdin procedure works, then it must be true that $\mathcal{M} \geq \mathcal{N}$.

If $\mathcal{M} > \mathcal{N}$, then the row space of $a_{\nu j}(\mathbf{k})$ will have a larger dimension than the column space. In this case, the column space will be a subspace of the row space and the orthonormality of the Wannier orbitals will be preserved. That means that the rows of $a_{\nu j}(\mathbf{k})$ are linearly independent and the rank will be equal to $\mathcal{N}$. But now the nullspace will be equal to $\mathcal{M} - \mathcal{N} \neq 0$ and the columns of $a_{\nu j}(\mathbf{k})$ will be linearly dependent. This means that, if I attempt to write the Bloch states in my energy window as a linear combination of the Wannier orbitals I will end up with linearly dependent Bloch states. In this case the Wannier orbitals are clearly not complete.

Finally, if $\mathcal{M} = \mathcal{N}$, then $a_{\nu j}(\mathbf{k})$ is a square matrix. Since I am assuming that the Löwdin procedure worked, the rows of $a_{\nu j}(\mathbf{k})$ are linearly independent and the rank will be equal to $\mathcal{N}$. This is equal to the number of columns $\mathcal{M}$, so the nullity is zero. This, combined with the orthonormality of the rows means that $a_{\nu j}(\mathbf{k})$ is a unitary matrix, so the resulting Wannier basis will be both orthonormal and complete. The final conclusion is that if the Löwdin procedure works the necessary and sufficient condition for an orthonormal and complete Wannier basis is that $\mathcal{M} = \mathcal{N}$.

1.5 *Ab initio* downfolding and the constrained random-phase approximation

The combination of the mathematics of linear response presented in Section 1.3 with the Wannier basis as constructed in Section 1.4 allows for the implementation of the constrained random-phase approximation (CRPA) as a method of *ab initio* downfolding. The CRPA takes advantage of two properties of most strongly correlated materials [92]. One, the properties of these systems is controlled by only a small subset of the degrees of freedom near the Fermi level. Two, these degrees of freedom are typically isolated from the rest in energy space. I will call the degrees of freedom near the Fermi level, like the *d*-orbitals in FeTe, the “target”. All of the other degrees of freedom I will call the “rest”.

The advantage of separating the degrees of freedom in this way is that a microscopic theory for the target can be formulated with only the implicit presence of the rest in terms of modified interaction parameters, though I have shown that one cannot completely forget about the rest in Appendix H. In the early history of attempts to estimate these parameters from first principles, and largely to this day, the model was chosen from the outset and the parameters calculated for that model. This model was either the Hubbard or Anderson model and the onsite interactions were the focus [8–10, 93]. One of the earliest attempts was the constrained LDA [93–95], which sought to estimate these interaction by calculating the energy difference between a state where an electron is moved from one orbital to another and the ground state.

However, this dissertation is primarily concerned with the CRPA [8–10]. The CRPA has a singular advantage over the constrained LDA. The expression for the interaction parameters has a clear origin in terms of the functional field integral expression for the trace and the derivation of an effective model for the target from performing a partial trace over the degrees of freedom in the rest [96, 97]. The degrees

of freedom in the rest are integrated out and the resulting effective action is written entirely in terms of quantities in the target.

The effect of the rest on the target is approximated to be entirely in the form of screening. The interactions between degrees of freedom in the target are reduced by the polarization of the degrees of freedom in the rest. In effect, the rest is transformed into a polarizable medium and only the polarization function is required to include its impact on the target. When the RPA polarization is used, then this is the CRPA. Though, I show in Appendix H that recovering response functions still requires considering the response of the rest, which is present even after downfolding.

The theoretical power of formulating CRPA in terms of the functional field integral, at least conceptually, is that any correlation function defined in the target Hilbert space can be calculated [98]. Moreover, any diagrammatic expression or equation of motion for the correlation function expressed in terms of the full Hilbert space can be used by simply replacing the bare Coulomb interaction with the screened interaction from CRPA and replacing internal sums over all of the degrees of freedom to sums over the target degrees of freedom. Since all quantities that are actually measured are written in terms of correlation functions, the complete solution of the microscopic theory is at least potentially attainable. Though the formulation of CRPA in terms of the functional field integral is conceptually useful, a full derivation would require an extensive diversion into the mathematics of the functional renormalization group without any change to the final results. As this has no bearing on any of the calculations I will be performing, I will instead present the more conventional formulation of CRPA.

CRPA has found considerable utility in the calculation of interactions for the transition metals [99–102], transition metal oxide perovskites [100, 101], iridates [103], graphene and graphite [104] and the iron pnictides and chalcogenides [105–107]. Though most of the work has focused on the onsite interactions, Nakamura, Arita and Imada [105] showed that the offsite interactions drop off as $\sim 1/R$, as can be expected from the form of the expression for the interactions. There is evidence

from a study of the role of magnetism in opening the band gap in the iridates that a value for interaction within $\approx 10\%$ of the CRPA calculated values give remarkably good results. Systematic application to a large number of diverse systems will help to firmly establish the reliability of CRPA calculated interactions.

The most common derivations for the screened interaction in CRPA starts with the polarization [10], which for my purposes is expressed as the Kohn-Sham response function in Eq. (1.98). This identification is valid since I am considering the RPA for which $f^{\text{xc}} = 0$. The division between target and rest is then simple and takes advantage of the sums over j and j' in Eq. (1.98). The Kohn-Sham response is the sum of two parts:

$$\chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}}(\mathbf{q}, \omega) = \chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}, \text{t}}(\mathbf{q}, \omega) + \chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}, \text{r}}(\mathbf{q}, \omega). \quad (1.115)$$

Both terms are given by expressions similar to Eq. (1.98) only for the first term the sums over j and j' are restricted so that both indices are for bands in the target. For CRPA applied to the ultimate construction of a microscopic theory in a downfolded, Wannier orbital basis, these bands will be precisely the bands used in the construction of Wannier functions described in Section 1.4. Though this correspondence is only strictly possible when the transformation between the Bloch and Wannier bases is unitary, an approximate microscopic theory for the target can be obtained even when there is remaining entanglement with the rest [108]. The second term in Eq. (1.115) consists of all the remaining terms in the sums over j and j' , including those that have either j or j' in the target bands, but not both.

The screened interaction in the RPA is given by

$$W_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) = v_{\mathbf{G}}(\mathbf{q})\delta_{\mathbf{G}\mathbf{G}'} + \sum_{\mathbf{G}_1\mathbf{G}_2} v_{\mathbf{G}}(\mathbf{q})\delta_{\mathbf{G}\mathbf{G}_1}\chi_{\mathbf{G}_1\mathbf{G}_2}^{\text{KS}}(\mathbf{q}, \omega)W_{\mathbf{G}_2\mathbf{G}'}(\mathbf{q}, \omega), \quad (1.116)$$

where $v_{\mathbf{G}}(\mathbf{q})$ are the Fourier components of the Coulomb interaction. This is a matrix equation in reciprocal space at each $\mathbf{q}$ and ω so that its solution is given by

$$\mathbf{W}(\mathbf{q}, \omega) = [\mathbf{1} - \mathbf{v}(\mathbf{q})\chi^{\text{KS}}(\mathbf{q}, \omega)]^{-1}\mathbf{v}(\mathbf{q}), \quad (1.117)$$

where boldface indicates matrices. By using Eq. (1.115) in Eq. (1.117), I can derive an alternative expression for the RPA screened interaction:

$$\begin{aligned} \mathbf{W}(\mathbf{q}, \omega) &= [\mathbf{1} - \mathbf{v}(\mathbf{q})\chi^{\text{KS}, \text{t}}(\mathbf{q}, \omega) - \mathbf{v}(\mathbf{q})\chi^{\text{KS}, \text{r}}(\mathbf{q}, \omega)]^{-1}\mathbf{v}(\mathbf{q}) \\ &= [(\mathbf{1} - \mathbf{v}(\mathbf{q})\chi^{\text{KS}, \text{r}}(\mathbf{q}, \omega))(\mathbf{1} - [\mathbf{1} - \mathbf{v}(\mathbf{q})\chi^{\text{KS}, \text{r}}(\mathbf{q}, \omega)]^{-1}\mathbf{v}(\mathbf{q})\chi^{\text{KS}, \text{t}}(\mathbf{q}, \omega))]^{-1}\mathbf{v}(\mathbf{q}) \\ &= [\mathbf{1} - [\mathbf{1} - \mathbf{v}(\mathbf{q})\chi^{\text{KS}, \text{r}}(\mathbf{q}, \omega)]^{-1}\mathbf{v}(\mathbf{q})\chi^{\text{KS}, \text{t}}(\mathbf{q}, \omega)]^{-1}[(\mathbf{1} - \mathbf{v}(\mathbf{q})\chi^{\text{KS}, \text{r}}(\mathbf{q}, \omega)]^{-1}\mathbf{v}(\mathbf{q}) \\ &= [\mathbf{1} - \mathbf{W}^{\text{c}}(\mathbf{q}, \omega)\chi^{\text{KS}, \text{t}}(\mathbf{q}, \omega)]^{-1}\mathbf{W}^{\text{c}}(\mathbf{q}, \omega), \end{aligned} \quad (1.118)$$

where I have defined the constrained screened interaction given by a similar equation to Eq. (1.116):

$$W_{\mathbf{G}\mathbf{G}'}^{\text{c}}(\mathbf{q}, \omega) = v_{\mathbf{G}}(\mathbf{q})\delta_{\mathbf{G}\mathbf{G}'} + \sum_{\mathbf{G}_1\mathbf{G}_2} v_{\mathbf{G}}(\mathbf{q})\delta_{\mathbf{G}\mathbf{G}_1}\chi_{\mathbf{G}_1\mathbf{G}_2}^{\text{KS}, \text{r}}(\mathbf{q}, \omega)W_{\mathbf{G}_2\mathbf{G}'}^{\text{c}}(\mathbf{q}, \omega). \quad (1.119)$$

The key result to all of this is that if I consider W^{c} to be the bare interaction for the downfolded target system, then I am able to reproduce the screened interaction for the full system (target and rest) using Eq. (1.118) where the polarization is the RPA polarization produced by the target degrees of freedom only. At least for W , the constrained interaction W^{c} together with the degrees of freedom in the target is able to fully reproduce the RPA result *in exactly the same way* as the bare interaction v together with the degrees of freedom in the full system.

Another way to see this is to examine an alternative expression for the screened interaction:

$$W_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega) = v_{\mathbf{G}}(\mathbf{q})\delta_{\mathbf{G}\mathbf{G}'} + v_{\mathbf{G}}(\mathbf{q})\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, \omega)v_{\mathbf{G}'}(\mathbf{q}). \quad (1.120)$$

Is it possible find an alternative expression for W in this form that involves the density response calculated using only target degrees of freedom? If W^c can indeed be considered at the bare interaction for the target degrees of freedom, then the density response in the target (in the RPA) should be given by

$$\chi_{\mathbf{G}\mathbf{G}'}^t(\mathbf{q}, \omega) = \chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}, t}(\mathbf{q}, \omega) + \sum_{\mathbf{G}_1 \mathbf{G}_2} \chi_{\mathbf{G}\mathbf{G}_1}^{\text{KS}, t}(\mathbf{q}, \omega) W_{\mathbf{G}_1 \mathbf{G}_2}^c(\mathbf{q}, \omega) \chi_{\mathbf{G}_2 \mathbf{G}'}^t(\mathbf{q}, \omega). \quad (1.121)$$

In order to see if this does the trick I will first rearrange Eq. (1.118) and expand it in a series:

$$\begin{aligned} \mathbf{W}(\mathbf{q}, \omega) &= \mathbf{W}^c(\mathbf{q}, \omega) + \mathbf{W}^c(\mathbf{q}, \omega) \chi^{\text{KS}, t}(\mathbf{q}, \omega) \mathbf{W}(\mathbf{q}, \omega) \\ &= \mathbf{W}^c(\mathbf{q}, \omega) + \mathbf{W}^c(\mathbf{q}, \omega) \chi^{\text{KS}, t}(\mathbf{q}, \omega) \mathbf{W}^c(\mathbf{q}, \omega) \\ &\quad + \mathbf{W}^c(\mathbf{q}, \omega) \chi^{\text{KS}, t}(\mathbf{q}, \omega) \mathbf{W}^c(\mathbf{q}, \omega) \chi^{\text{KS}, t}(\mathbf{q}, \omega) \mathbf{W}^c(\mathbf{q}, \omega) + \dots \\ &= \mathbf{W}^c(\mathbf{q}, \omega) + \mathbf{W}^c(\mathbf{q}, \omega) [\chi^{\text{KS}, t}(\mathbf{q}, \omega) + \chi^{\text{KS}, t}(\mathbf{q}, \omega) \mathbf{W}^c(\mathbf{q}, \omega) \chi^{\text{KS}, t}(\mathbf{q}, \omega) + \dots] \mathbf{W}^c(\mathbf{q}, \omega) \\ &= \mathbf{W}^c(\mathbf{q}, \omega) + \mathbf{W}^c(\mathbf{q}, \omega) \chi^t(\mathbf{q}, \omega) \mathbf{W}^c(\mathbf{q}, \omega). \end{aligned} \quad (1.122)$$

The final line is the exact analog of Eq. (1.120) only the bare interaction for the full system v is replaced by the constrained interaction appropriate for the target W^c and the density response for the full system χ is replaced by the density response for the target χ^t .

Though it is impressive that the full RPA screened interaction can be reproduced by an RPA calculation conducted entirely in the target, it suffers from the fact that W is not an observable of the system and so is not actually measured in experiment. An important question to address is to what extent the actual physical observables calculated in the target can reproduce those calculated using the full system. In Appendix H, I treat the case of a particularly relevant physical observable: the density response χ . Though the expression for χ given by the last line in Eq. (H.1) represents

a more complicated relationship between $\chi^\dagger$ and χ than maybe expected, the central quantity for low-energy excitations is still $\chi^\dagger$.

The terms in the Hamiltonian that contain the interaction parameters are only one part of the total model Hamiltonian. The other main part is the single-particle term which describes the noninteracting reference system. The parameters for this term are the so-called “hoppings.” These have a particularly simple expression when one uses Wannier functions for the orbital degrees of freedom. The definition of the hopping parameters is given by

$$t_{\nu\nu',\sigma\sigma'}(\mathbf{R} - \mathbf{R}') \equiv \int d^3x w_{\nu\sigma}^*(\mathbf{x} - \mathbf{R}) \left[-\frac{\hbar^2}{2m} \nabla^2 + v^{\text{KS}}[n](\mathbf{x}) \right] w_{\nu'\sigma'}(\mathbf{x} - \mathbf{R}'). \quad (1.123)$$

The integral in Eq. (1.123) does not need to be evaluated explicitly. Instead it is simpler to use the definition for the Wannier function given by Eq. (1.99) together with the fact the the Kohn-Sham eigenfunctions satisfy Eq. (1.47). Then the expression for the hopping parameters becomes

$$t_{\nu\nu',\sigma\sigma'}(\mathbf{R} - \mathbf{R}') = \delta_{\sigma\sigma'} \frac{1}{N} \sum_{\mathbf{k}}^{\text{BZ}} e^{i\mathbf{k} \cdot (\mathbf{R} - \mathbf{R}')} \sum_j \epsilon_{\mathbf{k}j\sigma} a_{\nu j}^*(\mathbf{k}) a_{\nu'j}(\mathbf{k}). \quad (1.124)$$

Equation (1.121) implies a solution for $\chi^\dagger$ as an inversion in reciprocal space. If the target is written in terms of a Wannier basis and the degrees of freedom are localized, as is the case in this dissertation, then an inversion in reciprocal space is likely not natural. The need to include significant crystal local-field effects due to the orbital degrees of freedom may necessitate a large number of reciprocal lattice vectors (large matrices in Eq. (1.121)) in order to achieve an accurate result. Instead, it would be more natural to invert in the space of Wannier orbitals. This is one of the goals of the next section, in which I present the formalism of TDDFT linear response in the Wannier basis.

1.6 Linear response in the Wannier basis

One of the primary applications of the Wannier basis in this dissertation is to calculate the charge response of the orbital degrees of freedom spanning the Hilbert subspace defined by the transformation to Wannier orbitals. That is the subspace of desired bands near the Fermi level. To this end, it is useful to start from the Kohn-Sham density response function calculated using the Bloch basis. In what follows, I am relying heavily on unpublished notes by Adolfo Eguiluz.

Let us consider one of the Bloch matrix elements in Eq. (1.98). I will use the formula for transforming to Wannier orbitals from Bloch states in order to find an expression for the Bloch matrix elements in terms of Wannier orbitals:

$$\begin{aligned} \langle \mathbf{k}, j, \sigma | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} | \mathbf{k} + \mathbf{q}, j', \sigma \rangle &= \frac{1}{N} \sum_{\mathbf{R}\mathbf{R}'} e^{-i\mathbf{k}\cdot\mathbf{R}} e^{i(\mathbf{k}+\mathbf{q})\cdot\mathbf{R}'} \sum_{\nu\nu'} a_{\nu j}(\mathbf{k}) a_{\nu' j'}^*(\mathbf{k} + \mathbf{q}) \\ &\times \langle w_{\nu\sigma}(\mathbf{R}) | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} | w_{\nu'\sigma}(\mathbf{R}') \rangle. \end{aligned} \quad (1.125)$$

Note that in order to perform this transformation the Wannier basis must be complete basis. The space that the Wannier orbitals span will always be a subspace of the full single particle Hilbert space. Thus, the transformation in Eq. (1.125) only makes sense if all of the Bloch states labeled by j span the same subspace as the Wannier basis. That is they must be related by a unitary transformation.

I can use periodic symmetry to eliminate one of the sums over $\mathbf{R}$ in Eq. (1.125), which represents a significant simplification of the expression. The first thing to notice is that I can perform a change of variables in the integral implicit in the bracket on the right-hand side of Eq. (1.125) to derive the identity

$$\langle w_{\nu\sigma}(\mathbf{R}) | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} | w_{\nu'\sigma}(\mathbf{R}') \rangle = e^{-i\mathbf{q}\cdot\mathbf{R}} \langle w_{\nu\sigma}(\mathbf{0}) | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} | w_{\nu'\sigma}(\mathbf{R}' - \mathbf{R}) \rangle. \quad (1.126)$$

Now I can make the change of dummy variables $\mathbf{R} \rightarrow \mathbf{R}' - \mathbf{R}$ in Eq. (1.125). The resulting summand will only depend on $\mathbf{R}$, so I can perform the sum over $\mathbf{R}'$: $\sum_{\mathbf{R}'} 1 =$

N . The final result includes only one sum over $\mathbf{R}$:

$$\begin{aligned} \langle \mathbf{k}, j, \sigma | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} | \mathbf{k} + \mathbf{q}, j', \sigma \rangle &= \sum_{\mathbf{R}} e^{i(\mathbf{k}+\mathbf{q})\cdot\mathbf{R}} \sum_{\nu\nu'} a_{\nu j}(\mathbf{k}) a_{\nu' j'}^*(\mathbf{k} + \mathbf{q}) \\ &\times \langle w_{\nu\sigma}(\mathbf{0}) | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} | w_{\nu'\sigma}(\mathbf{R}) \rangle. \end{aligned} \quad (1.127)$$

As a practical matter, this simplification makes many calculations in this dissertation possible that otherwise would be computationally prohibitive. Plugging Eq. (1.127) into Eq. (1.98) and making several definitions that I will provide shortly, I can arrive at an expression for the Kohn-Sham response function in terms of sums over Wannier indices:

$$\chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}}(\mathbf{q}, \omega) = \sum_{\sigma} \sum_{\lambda\lambda'} A_{\lambda,\sigma}(\mathbf{q} + \mathbf{G}) \tilde{\chi}_{\lambda\lambda',\sigma}^{\text{KS}}(\mathbf{q}, \omega) A_{\lambda',\sigma}^*(\mathbf{q} + \mathbf{G}'), \quad (1.128)$$

where $\lambda = \{\nu, \nu'; \mathbf{R}\}$, $\lambda' = \{\nu_1, \nu_1'; \mathbf{R}'\}$ are compound indices enumerating Wannier particle-hole transitions with one orbital designated by ν in the unit cell with its origin at $\mathbf{0}$ and the other designated by ν' in the unit cell with its origin at $\mathbf{R}$.

The amplitude for the particle-hole is given by

$$A_{\lambda,\sigma}(\mathbf{q} + \mathbf{G}) = \int_{\Omega} d^3x w_{\nu\sigma}^*(\mathbf{x}) e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} w_{\nu'\sigma}(\mathbf{x} - \mathbf{R}). \quad (1.129)$$

Intuitively, the particle-hole amplitude describes how the particle-hole couples to the perturbing field. As an example, if I were to examine the optical limit, $\mathbf{G} = \mathbf{0}$ and $\mathbf{q} \rightarrow \mathbf{0}$, and consider a λ that describes two d orbitals that respect inversion symmetry residing on the same site, then the amplitude in Eq. (1.129) would be zero. This means that the amplitudes will correctly describe selection rules. The integral in Eq. (1.129) is over the entire macro-crystal and performing it directly would be no simple matter. Instead, the numerical calculation in Exciting Plus [13] computes the Wannier matrix elements using matrix elements in the Bloch basis, which can be

evaluated over a unit cell after appropriate transformation. The formula is

$$A_{\lambda,\sigma}(\mathbf{q} + \mathbf{G}) = \frac{1}{N} \sum_{\mathbf{k}}^{\text{BZ}} e^{-i(\mathbf{k}+\mathbf{q})\cdot\mathbf{R}} \sum_{jj'} a_{\nu j}^*(\mathbf{k}) a_{\nu' j'}(\mathbf{k} + \mathbf{q}) \langle \mathbf{k}, j, \sigma | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} | \mathbf{k} + \mathbf{q}, j', \sigma \rangle. \quad (1.130)$$

The Kohn-Sham density response written in Wannier particle-hole indices is

$$\begin{aligned} \tilde{\chi}_{\lambda\lambda',\sigma}^{\text{KS}}(\mathbf{q}, \omega) = & \frac{1}{\Omega} \sum_{\mathbf{k}}^{\text{BZ}} e^{i(\mathbf{k}+\mathbf{q})\cdot(\mathbf{R}-\mathbf{R}')} \sum_{jj'} a_{\nu j}(\mathbf{k}) a_{\nu' j'}^*(\mathbf{k} + \mathbf{q}) \\ & \times \frac{f_{\mathbf{k},j,\sigma} - f_{\mathbf{k}+\mathbf{q},j',\sigma}}{\epsilon_{\mathbf{k},j,\sigma} - \epsilon_{\mathbf{k}+\mathbf{q},j',\sigma} + \hbar(\omega + i\eta)} a_{\nu_1 j}^*(\mathbf{k}) a_{\nu_1' j'}(\mathbf{k} + \mathbf{q}). \end{aligned} \quad (1.131)$$

The fraction is familiar from the Bloch expression in Eq. (1.98) and serves the same purpose here: to place simple poles at the non-interacting particle-hole excitation energies. However, in addition to this are the products $a_{\nu j}(\mathbf{k}) a_{\nu' j'}^*(\mathbf{k} + \mathbf{q})$. In this context, one can view $a_{\nu j}(\mathbf{k})$ as an amplitude that measures the amount of the Wannier orbital class labeled by ν that is in the Bloch state labeled by $j, \mathbf{k}$. Then the product $a_{\nu j}(\mathbf{k}) a_{\nu' j'}^*(\mathbf{k} + \mathbf{q})$ describes a kind of joint availability for the Wannier particle-hole. In order for a term to contribute to Eq. (1.131), all of the $a_{\nu j}(\mathbf{k})$ appearing in the expression must be finite. For example, if the orbital class ν does not have any content in the band j across the Brillouin zone, then the terms with the band index j will not contribute to the sum in Eq. (1.131).

Calculating the response written in the form of Eq. (1.128) requires performing two sums over lattice vectors $\mathbf{R}$. For particle-hole transitions that are long-ranged, the amplitudes A_λ are significant even for large values of $\mathbf{R}$ and the number of transitions increases with distance. This makes performing the sum in Eq. (1.128) computationally difficult due to the enormous number of terms and matrix elements that must be calculated. To mitigate these difficulties, I will present the derivation of an alternative expression in which the sum over lattice vectors has been performed analytically, eliminating the computational difficulty. In addition, the alternative

expression provides a limit in which we move from a picture of localized Wannier particle-holes to a picture that is more naturally viewed in terms of momentum space.

In the context of optics, the result in which we have summed over $\mathbf{R}$ allows us the possibility of describing the optical excitations in terms of pockets in the Brillouin zone. This is of considerable utility when analyzing the spectrum of Cu_2O (see Section 2.3 and in particular Fig. 2.13.) Combined with the expression in Eq. (1.128), where the Kohn-Sham density response is decomposed into Wannier particle-hole transitions, we are also able to evaluate whether such a momentum-space picture is appropriate or not.

The derivation that follows represents work that is original to me. I will begin with the expression for the Bloch matrix elements in Eq. (1.125). From there I will part ways with the derivation of Eq. (1.128). In that derivation, I used periodicity to eliminate one of the sums over $\mathbf{R}$. In this derivation I will sum analytically over $\mathbf{R}$ instead. Using the relation

$$\frac{1}{\sqrt{N}} \sum_{\mathbf{R}} e^{-i\mathbf{k}\cdot\mathbf{R}} |w_{\nu}(\mathbf{R})\rangle = \sum_j a_{\nu j}(\mathbf{k}) |\mathbf{k}, j\rangle, \quad (1.132)$$

I find the following expression for $\mathbf{R}$ -summed Wannier matrix elements:

$$\begin{aligned} \frac{1}{N} \sum_{\mathbf{R}\mathbf{R}'} e^{-i\mathbf{k}\cdot\mathbf{R}} e^{i(\mathbf{k}+\mathbf{q})\cdot\mathbf{R}'} \langle w_{\nu}(\mathbf{R}) | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} | w_{\nu'}(\mathbf{R}') \rangle \\ = \sum_{jj'} a_{\nu j}^*(\mathbf{k}) a_{\nu' j'}(\mathbf{k} + \mathbf{q}) \langle \mathbf{k}, j | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} | \mathbf{k} + \mathbf{q}, j' \rangle \quad (1.133) \\ \equiv M_{\nu\nu'}(\mathbf{k}, \mathbf{q}, \mathbf{G}). \end{aligned}$$

Now I am able to write the response decomposed into classes of Wannier channels where for each class of particle-hole transitions, enumerated by the pair of indices

$\{\nu, \nu'\}$, all distances separating the particle and hole are included:

$$\chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}}(\mathbf{q}, \omega) = \frac{1}{\Omega} \sum_{\mathbf{k}}^{\text{BZ}} \sum_{\substack{\nu\nu' \\ \nu_1\nu_1'}} M_{\nu\nu'}(\mathbf{k}, \mathbf{q}, \mathbf{G}) \bar{\chi}_{\nu\nu'\nu_1\nu_1'}^{\text{KS}}(\mathbf{k}, \mathbf{q}, \omega) M_{\nu_1\nu_1'}^*(\mathbf{k}, \mathbf{q}, \mathbf{G}'), \quad (1.134)$$

where

$$\bar{\chi}_{\nu\nu'\nu_1\nu_1'}^{\text{KS}}(\mathbf{k}, \mathbf{q}, \omega) = \sum_{jj'} a_{\nu j}(\mathbf{k}) a_{\nu' j'}^*(\mathbf{k} + \mathbf{q}) \frac{f_{\mathbf{k}, j} - f_{\mathbf{k} + \mathbf{q}, j'}}{E_{\mathbf{k}, j} - E_{\mathbf{k} + \mathbf{q}, j'} + \hbar\omega + i\eta} a_{\nu_1 j}^*(\mathbf{k}) a_{\nu_1' j'}(\mathbf{k} + \mathbf{q}) \quad (1.135)$$

is the $\mathbf{R}$ -summed analogue of Eq. (1.131).

From here, it is instructive to examine the optical limit. In the $\mathbf{q} \rightarrow \mathbf{0}$ limit, the matrix elements defined in Eq. (1.133) become

$$\begin{aligned} \lim_{\mathbf{q} \rightarrow \mathbf{0}} M_{\nu\nu'}(\mathbf{k}, \mathbf{q}, \mathbf{0}) &= \sum_j a_{\nu j}^*(\mathbf{k}) a_{\nu' j}(\mathbf{k}) - i\mathbf{q} \cdot \sum_{jj'} a_{\nu j}^*(\mathbf{k}) a_{\nu' j'}(\mathbf{k}) \langle \mathbf{k}, j | \mathbf{x} | \mathbf{k}, j' \rangle \\ &= \sum_j a_{\nu j}^*(\mathbf{k}) a_{\nu' j}(\mathbf{k}) + \mathbf{q} \cdot \sum_{jj'} a_{\nu j}^*(\mathbf{k}) a_{\nu' j'}(\mathbf{k}) \frac{\langle \mathbf{k}, j | \mathbf{p} | \mathbf{k}, j' \rangle}{E_{\mathbf{k}, j'} - E_{\mathbf{k}, j}}. \end{aligned} \quad (1.136)$$

If the Wannier orbitals are either completely occupied or completely empty, then the first term vanishes and only the second remains. The second term contains the matrix elements of the momentum operator familiar as a part of the traditional theory of optical properties, but it also contains the combination $a_{\nu j}^*(\mathbf{k}) a_{\nu' j'}(\mathbf{k})$, which is unique to response using Wannier orbitals. The sum over j and j' is a sum over interband transitions and the transitions that contribute the most to the matrix elements will (i) have maximal $|a_{\nu j}(\mathbf{k})|$ and $|a_{\nu' j'}(\mathbf{k})|$ and (ii) have a large value for $|\langle \mathbf{k}, j | \mathbf{p} | \mathbf{k}, j' \rangle|$.

$\bar{\chi}^{\text{KS}}$ has a simple optical limit in the case of fully occupied or fully empty Wannier orbitals, namely

$$\lim_{\mathbf{q} \rightarrow \mathbf{0}} \bar{\chi}_{\nu\nu'\nu_1\nu_1'}^{\text{KS}}(\mathbf{k}, \mathbf{q}, \omega) = \sum_{jj'} a_{\nu j}(\mathbf{k}) a_{\nu' j'}^*(\mathbf{k}) \frac{f_{\mathbf{k}, j} - f_{\mathbf{k}, j'}}{E_{\mathbf{k}, j} - E_{\mathbf{k}, j'} + \hbar\omega + i\eta} a_{\nu_1 j}^*(\mathbf{k}) a_{\nu_1' j'}(\mathbf{k}). \quad (1.137)$$

Again we see the combination $a_{\nu j}^*(\mathbf{k})a_{\nu' j'}(\mathbf{k})$, so that this part of the expression for the optical response in terms of Wannier orbitals will also be strongly suppressed if either of these coefficients are small. Thus, for a particular class of Wannier particle-hole transitions defined by the combination $\{\nu, \nu'\}$, the combination $a_{\nu j}^*(\mathbf{k})a_{\nu' j'}(\mathbf{k})$ appears to fourth order, serving to exaggerate the enhancement or suppression due to this product. This means that by examining the values of $|a_{\nu j}(\mathbf{k})|$, we are able to eliminate interband transitions. In addition, when examining them in concert with the Wannier decomposition of the response, we can deduce which pockets in the Brillouin zone are likely to produce particular optical features and make assignments of peaks according to orbitals. This will become important in the interpretation of the optical spectrum of Cu_2O (see Fig. 2.13.)

When calculating the contribution from particular Wannier channels to the total response calculated either in the Bloch basis or by including all channels for a metallic system, a problem arises as a result of the Wannier basis including orbitals which are neither fully occupied nor fully empty. This can be seen by examining the expression for a single Wannier channel.

$$A_\lambda(\mathbf{q} + \mathbf{G})\tilde{\chi}_{\lambda\lambda'}^{\text{KS}}(\mathbf{q}, \omega)A_{\lambda'}^*(\mathbf{q} + \mathbf{G}'), \quad (1.138)$$

The problem lies in the calculation of the matrix elements, most apparent when examining them in the form given in Eq. (1.130). For an insulating system with, say, orbital ν being occupied and ν' being empty, it is necessarily the case that, at a particular point $\mathbf{k}$, the Bloch state marked by j is occupied and that marked by j' is empty. Then this Bloch matrix element would be included in the calculation of the response in the Bloch basis and is also included in calculating the Wannier matrix element. Conversely, if both ν and ν' are occupied (or empty) then the Bloch matrix should not be included and this can be easily enforced by eliminating these type of Wannier transitions in the calculation of the Wannier matrix elements.

However, if we have a metallic system and some of the orbitals are partially occupied, then there is no good criterion for eliminating Wannier transitions when calculating the matrix elements and Bloch matrix elements that make no contribution to the total result will end up included in the calculation of the contribution from a particular channel. The solution is to eliminate these spurious Bloch matrix elements from the calculation of the Wannier matrix elements in such a way that if all transitions are included the total result will be reproduced.

The derivation of just such an expression is also original to me. I will start from the equation for χ^{KS} in Eq. (1.98). I would like the information contained in the quantity $f_{\mathbf{k},j} - f_{\mathbf{k}+\mathbf{q},j'}$, which excludes particle-particle and hole-hole transitions, to be included in both $\tilde{\chi}_{\lambda\lambda'}^{\text{KS}}$ and A_λ . This is facilitated by writing it as

$$f_{\mathbf{k},j} - f_{\mathbf{k}+\mathbf{q},j'} = \text{sign}[f_{\mathbf{k},j} - f_{\mathbf{k}+\mathbf{q},j'}] (|f_{\mathbf{k},j} - f_{\mathbf{k}+\mathbf{q},j'}|^{1/3})^3. \quad (1.139)$$

I can then write Eq. (1.98) as

$$\begin{aligned} \chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}}(\mathbf{q}, \omega) &= \sum_{\mathbf{k}_1}^{\text{BZ}} \sum_{j_1 j_1'} \delta_{\mathbf{k}_1 \mathbf{k}_2} \delta_{j_1 j_2} \delta_{j_1' j_2'} |f_{\mathbf{k}_1, j_1} - f_{\mathbf{k}_1 + \mathbf{q}, j_1'}|^{1/3} \langle \mathbf{k}_1, j_1 | e^{-i(\mathbf{q} + \mathbf{G}) \cdot \mathbf{x}} | \mathbf{k}_1 + \mathbf{q}, j_1' \rangle \\ &\times \frac{1}{\Omega} \sum_{\mathbf{k}_2}^{\text{BZ}} \sum_{j_2 j_2'} \frac{\text{sign}[f_{\mathbf{k}_2, j_2} - f_{\mathbf{k}_2 + \mathbf{q}, j_2'}] |f_{\mathbf{k}_2, j_2} - f_{\mathbf{k}_2 + \mathbf{q}, j_2'}|^{1/3}}{E_{\mathbf{k}_2, j_2} - E_{\mathbf{k}_2 + \mathbf{q}, j_2'} + \hbar\omega + i\eta} \\ &\times \sum_{\mathbf{k}_3}^{\text{BZ}} \sum_{j_3 j_3'} \delta_{\mathbf{k}_3 \mathbf{k}_2} \delta_{j_3 j_2} \delta_{j_3' j_2'} |f_{\mathbf{k}_3, j_3} - f_{\mathbf{k}_3 + \mathbf{q}, j_3'}|^{1/3} \langle \mathbf{k}_3 + \mathbf{q}, j_3' | e^{i(\mathbf{q} + \mathbf{G}') \cdot \mathbf{x}} | \mathbf{k}_3, j_3 \rangle. \end{aligned} \quad (1.140)$$

Using the orthonormality and completeness of the Wannier basis, one can write the delta functions in the forms $\delta_{\mathbf{k}_1 \mathbf{k}_2} = \sum_{\mathbf{R} - \mathbf{R}'} e^{-i(\mathbf{k}_1 - \mathbf{k}_2) \cdot (\mathbf{R} - \mathbf{R}')}$ and $\delta_{j_1 j_2} = \sum_{\nu} a_{\nu j_1}^*(\mathbf{k}) a_{\nu j_2}(\mathbf{k})$. This allows me to split χ^{KS} into contributions from Wannier channels:

$$\chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}}(\mathbf{q}, \omega) = \sum_{\lambda\lambda'} A_\lambda^{\text{met}}(\mathbf{q} + \mathbf{G}) \tilde{\chi}_{\lambda\lambda'}^{\text{KS, met}}(\mathbf{q}, \omega) A_{\lambda'}^{\text{met}*}(\mathbf{q} + \mathbf{G}'), \quad (1.141)$$

where the quantities involved are now calculated using the following expressions:

$$A_{\lambda}^{\text{met}}(\mathbf{q} + \mathbf{G}) = \frac{1}{N} \sum_{\mathbf{k}} e^{-i(\mathbf{k}+\mathbf{q}) \cdot \mathbf{R}} \sum_{jj'} |f_{\mathbf{k},j} - f_{\mathbf{k}+\mathbf{q},j'}|^{1/3} \times a_{\nu j}^*(\mathbf{k}) a_{\nu' j'}(\mathbf{k} + \mathbf{q}) \langle \mathbf{k}, j | e^{-i(\mathbf{q}+\mathbf{G}) \cdot \mathbf{x}} | \mathbf{k} + \mathbf{q}, j' \rangle \quad (1.142)$$

and

$$\tilde{\chi}_{\lambda\lambda'}^{\text{KS, met}}(\mathbf{q}, \omega) = \frac{1}{\Omega} \sum_{\mathbf{k}} e^{i(\mathbf{k}+\mathbf{q}) \cdot (\mathbf{R}-\mathbf{R}')} \sum_{jj'} a_{\nu j}(\mathbf{k}) a_{\nu' j'}^*(\mathbf{k} + \mathbf{q}) \times \frac{\text{sign}[f_{\mathbf{k},j} - f_{\mathbf{k}+\mathbf{q},j'}] |f_{\mathbf{k},j} - f_{\mathbf{k}+\mathbf{q},j'}|^{1/3}}{E_{\mathbf{k},j} - E_{\mathbf{k}+\mathbf{q},j'} + \hbar\omega + i\eta} \times a_{\nu_1 j}^*(\mathbf{k}) a_{\nu_1' j'}(\mathbf{k} + \mathbf{q}). \quad (1.143)$$

I have found it necessary to use the modified expressions for response in the Wannier basis Eqs. (1.142) and (1.143) to decompose the optical response of FeTe into orbital particle-hole channels in Section 4.3. Failing to do so results in contributions from each channel separately being far too large in relation to the total spectrum, though including all channels results in massive cancellation and a reproduction of the total spectrum. I believe that this is completely general and will be true for any metallic system that will have orbitals that are partially occupied. When orbitals are completely occupied or empty it is always possible to restrict the Wannier response to channels pairing occupied to empty and excluding unphysical ones. For metallic systems, this requires the modified expressions Eqs. (1.142) and (1.143). In this dissertation, I have only investigated the optical response as represented by χ^{KS} for a metallic system. The impact of partially occupied orbitals in channel-by-channel contributions to χ remains an open question.

Response in the Wannier basis also allows one to conduct CRPA calculations in which the downfolded subspace spanned by the Wannier orbitals is not clearly separable as a subset of bands. Though I derived the expressions in what follows

independently, similar refinements to CRPA have been done before [109, 110]. In typical CRPA calculations, the Kohn-Sham density response is split into the sum of a portion calculated using transitions entirely in the target subspace, designated by $\chi_{\mathbf{G}\mathbf{G}'}^{\text{KS},\text{t}}(\mathbf{q}, \omega)$, and another calculated using all of the remaining transitions, designated by $\chi_{\mathbf{G}\mathbf{G}'}^{\text{KS},\text{r}}(\mathbf{q}, \omega)$. $\chi_{\mathbf{G}\mathbf{G}'}^{\text{KS},\text{r}}(\mathbf{q}, \omega)$ is then used to calculate the screened interaction between orbitals in the target subspace. In my extension of the method, the ultimate target space is spanned by only a subset of Wannier orbitals. In that case, I add another quantity, designated $\chi_{\mathbf{G}\mathbf{G}'}^{\text{KS},\text{wr}}(\mathbf{q}, \omega)$, to $\chi_{\mathbf{G}\mathbf{G}'}^{\text{KS},\text{r}}(\mathbf{q}, \omega)$ in the calculation of the screened interaction. $\chi_{\mathbf{G}\mathbf{G}'}^{\text{KS},\text{wr}}(\mathbf{q}, \omega)$ is computed using Eq. (1.128) where either λ or λ' have a Wannier index outside of the ultimate target subspace. Thus, $\chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}}(\mathbf{q}, \omega)$ is split into three parts:

$$\chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}}(\mathbf{q}, \omega) = \chi_{\mathbf{G}\mathbf{G}'}^{\text{KS},\text{t}}(\mathbf{q}, \omega) + \chi_{\mathbf{G}\mathbf{G}'}^{\text{KS},\text{wr}}(\mathbf{q}, \omega) + \chi_{\mathbf{G}\mathbf{G}'}^{\text{KS},\text{r}}(\mathbf{q}, \omega), \quad (1.144)$$

where the first two terms are expressed using Eq. (1.128), the last term is computed in the Bloch basis and the last two terms and used to compute the screened interaction.

The reason I am able to do this is that the mathematics of CRPA only rely on being able to split χ^{KS} into two parts and it does not depend on using the Bloch basis to do this splitting. As long as W^c is only used as the interaction for the degrees of freedom in the target, everything is still consistent. If there are target orbitals at every site in the macrocrystal, then the $\mathbf{R}$ -summed Wannier response in Eq. (1.134) is appropriate to use in Eq. (1.144) as then $\chi^{\text{KS},\text{t}}$ will include transitions for orbitals for every unit cell. However, even that is not a restriction. In principle one could sum in $\mathbf{R}$ in only two dimensions but not in the third. In a layered compound this may mean that one keeps only orbital degrees of freedom in a single layer of a layered compound. Nakamura, Yoshimoto, Nohara and Imada [109] have dubbed this “dimensional downfolding”. In the most extreme example, $\chi^{\text{KS},\text{t}}$ may only include onsite transitions and then the target consists of orbitals on a single atom. This would

be an appropriate scheme for calculating the interaction parameters in dynamical mean-field theory [110].

All of the preceding discussion was regarding and made use of the Kohn-Sham response function in the Wannier basis. What about the true density response which, after all, represents a true physical observable. If I consider the Wannier basis as spanning the target space that would be appropriate for CRPA downfolding, then I argue that the appropriate quantity to consider is the part of the density response due to the degrees of freedom in the target, given by

$$\chi_{\mathbf{G}\mathbf{G}'}^t(\mathbf{q}, \omega) = \chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}, t}(\mathbf{q}, \omega) + \sum_{\mathbf{G}_1\mathbf{G}_2} \chi_{\mathbf{G}\mathbf{G}_1}^{\text{KS}, t}(\mathbf{q}, \omega) [W_{\mathbf{G}_1\mathbf{G}_2}^c(\mathbf{q}, \omega) + f_{\mathbf{G}_1\mathbf{G}_2}^{\text{xc}, c}(\mathbf{q}, \omega)] \chi_{\mathbf{G}_2\mathbf{G}'}^t(\mathbf{q}, \omega). \quad (1.145)$$

where I have included a term $f^{\text{xc}, c}$ which represents the exchange-correlation kernel appropriate for the target degrees of freedom alone as opposed to the f^{xc} appropriate for the full Hilbert space. Just as v had to be replaced by W^c to include screening from the rest, so f^{xc} must be replaced by $f^{\text{xc}, c}$. Of course, if the rest is treated in the RPA and the only contribution to f^{xc} is considered to be from the target degrees of freedom, then the approximation $f^{\text{xc}} = f^{\text{xc}, c}$ has been made. If $f^{\text{xc}, c} = 0$, then the target is being treated in the RPA.

I would like to find an expression for χ^t decomposed into Wannier channels in much the same way as for $\chi^{\text{KS}, t}$ in Eq. (1.128). To this end, using unpublished notes of Adolfo Eguiluz, I will expand Eq. (1.145) in a Dyson series and plug in the

expression Eq. (1.128) for $\chi^{\text{KS}, \text{t}}$:

$$\begin{aligned}
\chi_{\mathbf{G}\mathbf{G}'}^{\text{t}}(\mathbf{q}, \omega) &= \sum_{\sigma\sigma'} \sum_{\lambda\lambda'} A_{\lambda,\sigma}(\mathbf{q} + \mathbf{G}) \delta_{\sigma\sigma'} \tilde{\chi}_{\lambda\lambda',\sigma}^{\text{KS}}(\mathbf{q}, \omega) A_{\lambda',\sigma}^*(\mathbf{q} + \mathbf{G}') \\
&\quad + \sum_{\mathbf{G}_1\mathbf{G}_2} \sum_{\sigma\sigma'} \sum_{\lambda\lambda'} \sum_{\sigma_1\sigma_2} \sum_{\lambda_1\lambda_2} A_{\lambda,\sigma}(\mathbf{q} + \mathbf{G}) \delta_{\sigma\sigma_1} \tilde{\chi}_{\lambda\lambda_1,\sigma}^{\text{KS}}(\mathbf{q}, \omega) A_{\lambda_1,\sigma_1}^*(\mathbf{q} + \mathbf{G}_1) \\
&\quad \times [W_{\mathbf{G}_1\mathbf{G}_2}^{\text{c}}(\mathbf{q}, \omega) + f_{\mathbf{G}_1\mathbf{G}_2}^{\text{xc}, \text{c}}(\mathbf{q}, \omega)] \\
&\quad \times A_{\lambda_2,\sigma_2}(\mathbf{q} + \mathbf{G}_2) \delta_{\sigma_2\sigma'} \tilde{\chi}_{\lambda_2\lambda',\sigma'}^{\text{KS}}(\mathbf{q}, \omega) A_{\lambda',\sigma'}^*(\mathbf{q} + \mathbf{G}') + \dots \\
&= \sum_{\sigma\sigma'} \sum_{\lambda\lambda'} A_{\lambda,\sigma}(\mathbf{q} + \mathbf{G}) \left\{ \delta_{\sigma\sigma'} \tilde{\chi}_{\lambda\lambda',\sigma}^{\text{KS}}(\mathbf{q}, \omega) \right. \\
&\quad + \sum_{\sigma_1\sigma_2} \sum_{\lambda_1\lambda_2} \delta_{\sigma\sigma_1} \tilde{\chi}_{\lambda\lambda_1,\sigma}^{\text{KS}}(\mathbf{q}, \omega) \sum_{\mathbf{G}_1\mathbf{G}_2} A_{\lambda_1,\sigma_1}^*(\mathbf{q} + \mathbf{G}_1) [W_{\mathbf{G}_1\mathbf{G}_2}^{\text{c}}(\mathbf{q}, \omega) + f_{\mathbf{G}_1\mathbf{G}_2}^{\text{xc}, \text{c}}(\mathbf{q}, \omega)] \\
&\quad \times A_{\lambda_2,\sigma_2}(\mathbf{q} + \mathbf{G}_2) \delta_{\sigma_2\sigma'} \tilde{\chi}_{\lambda_2\lambda',\sigma'}^{\text{KS}}(\mathbf{q}, \omega) + \dots \left. \right\} A_{\lambda_1,\sigma_1}^*(\mathbf{q} + \mathbf{G}_1).
\end{aligned} \tag{1.146}$$

Now I can make the following definitions:

$$W_{\lambda\lambda',\sigma\sigma'}^{\text{c}}(\mathbf{q}, \omega) \equiv \sum_{\mathbf{G}_1\mathbf{G}_2} A_{\lambda,\sigma}^*(\mathbf{q} + \mathbf{G}_1) W_{\mathbf{G}_1\mathbf{G}_2}^{\text{c}}(\mathbf{q}, \omega) A_{\lambda',\sigma'}(\mathbf{q} + \mathbf{G}_2) \tag{1.147}$$

and

$$f_{\lambda\lambda',\sigma\sigma'}^{\text{xc}, \text{c}}(\mathbf{q}, \omega) \equiv \sum_{\mathbf{G}_1\mathbf{G}_2} A_{\lambda,\sigma}^*(\mathbf{q} + \mathbf{G}_1) f_{\mathbf{G}_1\mathbf{G}_2}^{\text{xc}, \text{c}}(\mathbf{q}, \omega) A_{\lambda',\sigma'}(\mathbf{q} + \mathbf{G}_2). \tag{1.148}$$

These are the interaction terms written explicitly as coupling Wannier orbitals. As W^{c} and $f^{\text{xc}, \text{c}}$ are quantities that only have any meaning inside of the target space, I consider the quantities as written on the left hand sides of Eq. (1.147) and Eq. (1.148) to be more meaningful than the reciprocal space expressions. Now it is possible to define the density response for the target in Wannier indices as the summation of the

Dyson series in the curly brackets in Eq. (1.146):

$$\begin{aligned} \tilde{\chi}_{\lambda\lambda',\sigma\sigma'}^t(\mathbf{q}, \omega) &\equiv \delta_{\sigma\sigma'} \tilde{\chi}_{\lambda\lambda',\sigma}^{\text{KS}}(\mathbf{q}, \omega) \\ &+ \sum_{\sigma_1\sigma_2} \sum_{\lambda_1\lambda_2} \delta_{\sigma\sigma_1} \tilde{\chi}_{\lambda\lambda_1,\sigma}^{\text{KS}}(\mathbf{q}, \omega) [W_{\lambda_1\lambda_2,\sigma_1\sigma_2}^c(\mathbf{q}, \omega) + f_{\lambda_1\lambda_2,\sigma_1\sigma_2}^{\text{xc},c}(\mathbf{q}, \omega)] \tilde{\chi}_{\lambda_2\lambda',\sigma_2\sigma'}^t(\mathbf{q}, \omega). \end{aligned} \quad (1.149)$$

The density response due to the target degrees of freedom in reciprocal space can then be written

$$\chi_{\mathbf{G}\mathbf{G}'}^t(\mathbf{q}, \omega) = \sum_{\sigma} \sum_{\lambda\lambda'} A_{\lambda,\sigma}(\mathbf{q} + \mathbf{G}) \tilde{\chi}_{\lambda\lambda',\sigma}^t(\mathbf{q}, \omega) A_{\lambda',\sigma}^*(\mathbf{q} + \mathbf{G}'). \quad (1.150)$$

This is the analogous expression to Eq. (1.128), which was for the Kohn-Sham response, but for the true density response.

A key point to finding the expressions Eqs. (1.149) and (1.150) is that the inversion that must take place to find the density response in the target is not longer done in reciprocal space as in Eq. (1.145) but in Wannier space. For situations where the physics is expected to be local, it is possible that the numerical cost could be reduced considerably in going from Eqs. (1.145) to (1.150). In addition, if one wishes to consider models for f^{xc} that take into account interactions between local orbital degrees of freedom, then the response equation in Eq. (1.149) is much more natural to consider. In that case one would devise an $f_{\lambda\lambda',\sigma\sigma'}^{\text{xc},c}(\mathbf{q}, \omega)$ directly rather than using Eq. (1.148). It is likely that coming up with an f^{xc} by considering interactions between orbitals versus the abstract $f_{\mathbf{G}\mathbf{G}'}^{\text{xc},c}(\mathbf{q}, \omega)$ will be more physically intuitive.

1.7 The LDA+ U method

The only many-body method departing from Kohn-Sham TDDFT in this dissertation is that of using LDA+ U to calculate the ground state and to then use this ground state for calculation of the linear density response. LDA+ U methods have had great success

in extending Kohn-Sham DFT into situations where by itself it does poorly [15]. The canonical examples are the Mott-Hubbard and charge-transfer insulators [111]. At best, LSDA will predict Mott-Hubbard insulators to have a gap that is much too small [112, 113]. This is because LDA functionals make a local approximation based on the homogeneous electron gas in which electrons are efficiently screened, but in systems with localized d and f electrons the screening is much less efficient [95]. In such situations it is appropriate to treat interactions among localized degrees of freedom differently than they are treated in LDA, GGA or other local DFT functionals based on the homogeneous electron gas.

This need to handle interactions among local orbitals is the rationale behind introducing the LDA+ U method. In LDA+ U , an orbitally dependent, Hubbard-like interaction is introduced that is only among d orbitals (or other localized orbitals) on the same atomic site. The Hartree-Fock approximation is then used to find an orbitally dependent, effective single particle contribution to the Kohn-Sham equation. This contribution depends on the density matrix rather than the total density, as in Kohn-Sham DFT. This contribution also introduces a non-local potential as it includes exchange between the orbitals whereas Kohn-Sham DFT contains only a local potential. The result is an inexpensive way of modifying DFT that is able to open gaps in Mott-Hubbard and charge transfer insulators as well as correct other systems in which local density functionals treat localized degrees of freedom inadequately.

In LDA+ U , a term is added to the energy functional which is given by [15]

$$E^U(\{n_{m,\sigma}\}) = \frac{1}{2} \left[U \sum_{mm',\sigma} n_{m,\sigma} n_{m',\bar{\sigma}} + (U - J) \sum_{\substack{mm',\sigma \\ m \neq m'}} n_{m,\sigma} n_{m',\sigma} \right], \quad (1.151)$$

where U and J are parameters representing the screened direct and exchange interactions, respectively, $n_{m,\sigma}$ is the occupation number of orbital m and spin σ and $\bar{\sigma}$ is the spin opposite to σ . Unlike the expression for the energy in DFT, Eq. (1.151) is not a universal functional of the total density, but rather a function of the orbital

occupation numbers parameterized by U and J . As a result, the contribution to the potential arising from the energy function in Eq. (1.151) is given by the derivative of the energy with respect to the occupation numbers. This results in an *orbitally dependent* potential versus the *local* potential in Kohn-Sham DFT. Thus the orbitally dependent potential is given by

$$v_{m,\sigma}^U(\{n_{m,\sigma}\}) = U \sum_{m'} n_{m',\bar{\sigma}} + (U - J) \sum_{\substack{m' \\ m \neq m'}} n_{m',\sigma}. \quad (1.152)$$

However, Eq. (1.151) cannot be simply added to the energy functional from LDA (or another DFT functional like GGA), because the LDA functional includes both the Hartree term and contributions from exchange and correlation which include the effects of electron-electron interactions. Thus, a “double-counting” term must be subtracted to account for the contribution from LDA. In implementations of LDA+ U , this term has not been derived directly from the LDA functional, but rather one of a couple empirical ansatzes have been used.

The first proposal by Anisimov, Zaanen and Andersen [114], the so-called “around mean field (AMF)” approximation, assumed that the contribution from the additional LDA+ U terms (that is Eq. (1.151) minus the double-counting) should vanish in the limit of uniform occupation. This accounts for the observation that in LDA, localization, which can lead to spin and orbital polarization, is controlled by the scale exchange interaction (given roughly by J) rather than the direct interaction (given roughly by U). In general U is an order of magnitude larger than J so that spin or orbital polarization will not be favored in LDA and instead the occupancies will be uniform. Thus, Eq. (1.151) is replaced in the AMF approximation by

$$E^{U, \text{AMF}}(\{n_{m,\sigma}\}) = \frac{1}{2} \left[U \sum_{mm',\sigma} (n_{m,\sigma} - \bar{n})(n_{m',\bar{\sigma}} - \bar{n}) + (U - J) \sum_{\substack{mm',\sigma \\ m \neq m'}} (n_{m,\sigma} - \bar{n})(n_{m',\sigma} - \bar{n}) \right], \quad (1.153)$$

where $\bar{n}$ is the average orbital occupancy. If the orbitals are all equally occupied, then $n_{m,\sigma} = \bar{n}$ for all m and σ and the desired limit is achieved.

Later, Anisimov et al. [115] observed that the AMF expression for the energy in Eq. (1.153) may not be the best. Instead they pointed out that the LDA gives values for total energies that agree well with experiment even though the orbital energies given by the Kohn-Sham eigenvalues agree poorly with measured single-particle energies. To a certain extent this should be expected as DFT is a theory oriented towards calculating the exact electron density and functionals of that density, such as the total energy. The Kohn-Sham eigenvalues do not have a meaning in terms of single particle energies. However, there is a true error that is made in LDA (and GGA, etc.) As argued by Perdew et al. [116], the exact exchange-correlation functional as a function of total particle number has a discontinuous first derivative when the particle number crosses an integer value. As a consequence, the exchange-correlation potential, and in turn the Kohn-Sham eigenvalues, have a discontinuity upon crossing an integer number of particles. This will account for some of the discrepancy between good total energies in LDA on one hand and poor single particle energies on the other.

Now, if we view the transition metal ion in a solid state system as approximately an atom in a open system such that it can exchange particles with its environment, but has no other interactions with its surroundings. Then, the discontinuities in the the exchange-correlation potential for this effective atom become vital to determining its energy levels. When one then considers the effective atom again as part of the solid-state system, then change in energy levels versus LDA will alter the electronic structure. If one assumes spherical symmetry, the energy of the effective atom can be found exactly and is given by [117]

$$E^{\text{atom}}(N_{\uparrow}, N_{\downarrow}) = \frac{1}{2} [UN(N-1) - JN_{\uparrow}(N_{\uparrow}-1) - JN_{\downarrow}(N_{\downarrow}-1)], \quad (1.154)$$

where $N_\sigma = \sum_m n_{m,\sigma}$ and $N = \sum_\sigma N_\sigma$. This is the double-counting term for the “fully localized limit (FLL),” under the assumption that the LDA energy and the exact energy for the effective atom is approximately equal:

$$E^{U, \text{FLL}}(\{n_{m,\sigma}\}) = E^U(\{n_{m,\sigma}\}) - E^{\text{atom}}(N_\uparrow, N_\downarrow). \quad (1.155)$$

Equation (1.155) also restores the discontinuity in the potential, as should be the case for the exact exchange-correlation potential. The contribution to the potential from Eq. (1.155) is, after a little algebra, given by

$$v_{m,\sigma}^{U, \text{FLL}}(n_{m,\sigma}) = (U - J)(1/2 - n_{m,\sigma}). \quad (1.156)$$

To see the discontinuity, consider a solid-state system with N^{solid} electrons where N^{solid} is effectively infinite. If we consider the energy of the solid as a function of N^{solid} and then consider the situation in which we add a single electron in the orbital labeled by m and σ . Then the maximum occupied Kohn-Sham eigenvalue jumps by [115, 116]

$$\Delta E = [E(N^{\text{solid}} + 1) - E(N^{\text{solid}})] - [E(N^{\text{solid}}) - E(N^{\text{solid}} - 1)] = U - J. \quad (1.157)$$

Thus, an effectively infinitesimal increase of the number of electrons above an integer value increases the maximum Kohn-Sham eigenvalue by a finite amount, yielding a discontinuity.

The AMF and FLL approximations for the double counting term represent two extremes. One represents the contribution from LDA as that of an approximately uniform system and the other as an effective atom. The true value lies somewhere in between these two limits. Petukhov et al. [118] devised a way to choose a linear interpolation constant such that the total energy functional remains variational in the charge density. This is done by insuring that the combination of Eq. (1.151) and the double-counting give a net zero contribution at the self-consistent density. Thus

the energy is the same as in LDA, but the LDA+ U contributions to the potential result in the density being different from the LDA density, resulting in a change in the electronic structure. In this way, the LDA+ U becomes a DFT functional, albeit an empirical one, that respects the Hohenberg-Kohn theorems. However, the LDA+ U method is still not a Kohn-Sham method because of the orbitally dependent, non-local potential. In general, in this dissertation I will use the interpolated double-counting correction, though the conclusions do not depend on this choice.

Chapter 2

The optical response of Cu_2O and optimization of the Wannier basis

The primary goal of the work presented in this dissertation is the construction of a Wannier basis that provides an “optimal description of physics” with a direct correspondence to experimentally measurable quantities, with a focus on optical spectroscopy. By “optimal” I generally mean that the fewest number of Wannier particle-hole channels will be needed to describe the optical spectrum, producing the simplest possible description of optics. By optimizing with respect to an experimentally measurable quantity, I can be secure in the knowledge that the physics embodied in the orbitals reflects realities. This criterion for optimization stands in contrast to other criteria for optimization that have been used, such as maximal localization [3, 81, 119], approaches that emphasize the symmetry of the resulting orbital [80, 82, 86, 91] or maximal similarity [120]. The criterion of producing an optimal description of the optics has the advantage of being unbiased in the sense that I am not enforcing an *a priori* assumption of what the orbitals should look like. Instead, I am evaluating the orbitals based on the real-world physics they can reproduce.

The reason to construct such an optimal Wannier basis is two-fold. First, by providing a simple description of optical spectroscopy, I am able to provide insight into the role of local chemistry and symmetry and orbital degrees of freedom on the two-particle excitation spectrum. Second, the resulting orbitals can then be used to develop microscopic models that provide access to other physical observables and allow the inclusion of many-body corrections beyond the approximations used in the construction of the orbitals (such as GGA to calculate the Kohn-Sham groundstate) and the optimization of the orbitals (such as RPA in the TDDFT response). Cu_2O is not the most natural system for the construction of such a model, so I will not do so in this chapter. Instead, I apply the optimization scheme presented in this chapter to FeTe and construct a model for that system in Chapter 4.

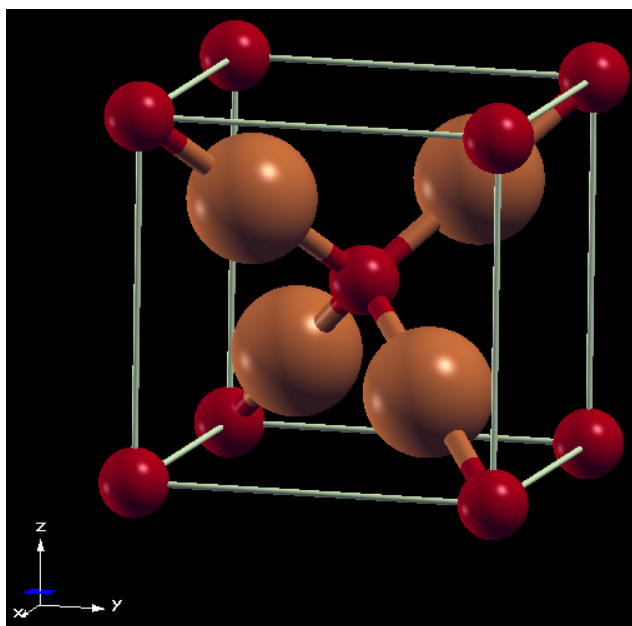
Essential in providing a link between local orbital physics embodied in the Wannier basis and optical is the transformation from the Bloch basis to a Wannier basis, introduced in Section 1.4 in combination with a reformulation of TDDFT linear response in the Wannier basis, introduced in Section 1.6. With regards to the latter, the work that I present here represents a continuation of unpublished work within the research group of Adolfo Eguiluz that preceded my entry. However, in developing the ability to calculate the optical response of Cu_2O I had to make several developments, including the ability to construct molecular orbitals using linear combinations of local orbitals given in Eq. (1.114) and the ability to sum analytically over lattice vectors in the Kohn-Sham response using Eq. (1.134). These developments were required to provide a link between local chemistry and symmetry and optics. With this link established, feedback has been able to develop in which improvement in the Wannier basis improves description of the optical features and monitoring the optics indicates ways to improve the orbitals.

Of course, this feedback between optics and orbitals is only as valuable as the quality of the results produced for a given level of approximation in TDDFT. As I will show in Section 2.1, the combination of GGA for generating the Kohn-Sham ground-state and RPA for the calculation of the Kohn-Sham response displays a remarkable

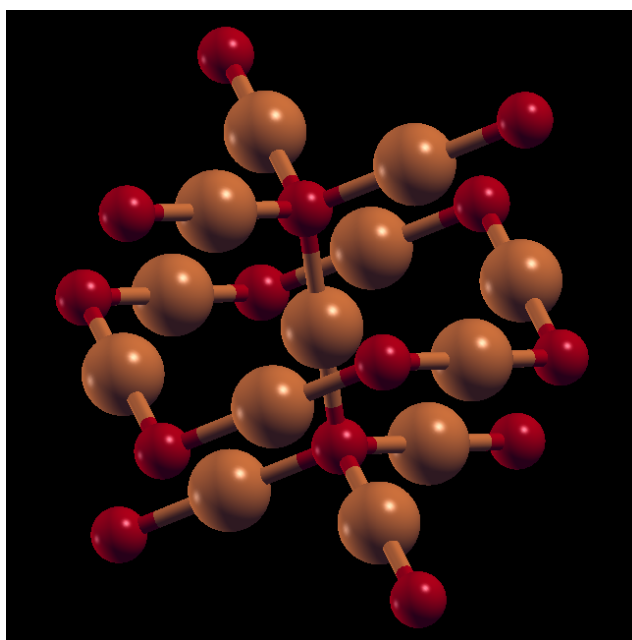
agreement with spectroscopic ellipsometry experiments. This is despite statements in the recent literature that GGA is insufficient for the groundstate [19] and the response can only be calculated with the inclusion of excitonic interactions [121]. I will also address these previous findings and reconcile them with the GGA/RPA calculations in Section 2.1. But first, I would like to present an argument that Cu_2O as a physical system is suitable as a test case aside from the agreement with experiment that I have been able to achieve.

Development of methods and procedures for optimizing the Wannier basis with respect to optical spectroscopy requires a test subject that I can use as an analog for a broad class of complex, transition metal compounds. Such compounds provide a rich variety of often poorly understood phenomena that are subject of intense current research. As such, they are the primary focus of this dissertation. As a test case for this investigation, the copper-oxide semiconductor Cu_2O provides an ideal subject. It has long been used as a standard material to study copper oxides [122] and, in contrast to many other transition metal oxides, Cu_2O is not strongly correlated and can be adequately treated using single-particle theory [122, 123]. In addition the non-trivial chemistry [16, 124] provides a rich environment for studying the local orbital degrees of freedom via the Wannier basis. The suitability of Cu_2O as a test case for my investigations is supported by picture of the material developed through a long history of experimental and theoretical investigations.

Full elucidation as to the suitability of Cu_2O as a test subject begins with the crystal structure, which is illustrated in Fig. 2.1. The structure consists of an FCC lattice of Cu atoms and a BCC lattice of O atoms displaced from the origin by (0.5, 0.5, 0.5) in units of the cubic lattice spacing [125, 126]. Thus, there are two formula units per primitive, cubic unit cell, as shown in Fig. 2.1a. The Cu atoms have two, linearly coordinated O neighbors and the O have four Cu neighbors. In an ionic picture of Cu_2O , the d -orbital shell would indeed be completely filled and neutral excitations involving the d -orbital degrees of freedom would be strongly suppressed.



(a)



(b)

Figure 2.1: Cu_2O crystal structure. Shown in (a) is the unit cell of Cu_2O showing the Cu atoms in an FCC arrangement and the O atoms in a BCC arrangement. However, the structure may be easier to understand as two interpenetrating Cu-O networks, which are highlighted in the selection of the crystal structure in (b). It is the cation-cation bonding that holds these networks in place against electrostatic repulsion.

However, it has long been known that the crystal structure of Cu_2O is unstable in the ionic picture [127], because the structure consists of two, interpenetrating Cu-O networks, shown in Fig. 2.1b, that would electrostatically repel in the absence of a chemical bonding mechanism. Zuo et al. [16] experimentally observed a z^2 -orbital-shaped depletion of charge density indicating that the d -orbital shell is not completely filled. In addition, the z^2 orbital is associated with covalent Cu-Cu bonding, which serves stabilize the structure. Filippetti and Fiorentini [124] disputed the covalency of the bonding, but via DFT calculations with a self-interaction correction, they were able to provide evidence for bonding due to incomplete charge transfer from Cu to O with the excess charge being delocalized into the space between Cu and O. Though there are some subtleties in the nature of the cation-cation bonding, I have found that the GGA calculations that I have performed are sufficient for its existence.

The combination of strong bonding chemistry between Cu and O and non-trivial chemistry between Cu and Cu is a vital aspect that helps to make Cu_2O a good test subject for the optimization scheme I am presenting. The bonding between transition metal cations and anions is a general feature of all complex transition metal compounds and the subtle bonding between transition metal atoms promises to be important in a broad class of such materials. I will show this in particular for the iron pnictides and chalcogenides in Chapter 4.

However, non-trivial local chemistry is not the only feature that one should seek in finding an ideal test subject for optimizing orbitals with respect to two-particle spectroscopy. I am primarily interested in the d -orbital degrees of freedom as these are the most important orbitals involved in producing interesting physics in the complex transition metal compounds. Two-particle spectroscopy involves transitions in which a hole is left behind in the occupied states with a particle created in the unoccupied states. Thus, to use two-particle spectroscopy most effectively as a window into d -orbital physics there should be either d orbitals spanning the Fermi level in a metal, or d orbitals on both sides of the gap in a insulator or semiconductor.

The existence of d -orbital degrees of freedom on both sides of the gap in Cu_2O implied by the observation of the z^2 orbital hole [16] has been confirmed by independent experimental investigations and supported by *ab initio* theory. The one-particle spectrum as measured by photoemission and inverse photoemission shows that the valence and conduction band can be reliably reproduced by even relatively primitive band structure calculations [122, 128–131]. This allowed a reliable assignment to peaks below and above the gap to d -orbital states.

Using resonant inelastic x-ray scattering (RIXS) Hu et al. [123] were also able to deduce that the spectrum was dominated by interband, d - d excitations. This was possible because of the relevant enhancement of the d - d process via the matrix element between the core hole and the conduction band. Of course, this Raman-type process favors onsite excitations thanks to the same matrix element and so the measured excitations were assigned to onsite d - d transitions. This differs from optical excitations, which are the focus of my investigation. Regardless, the RIXS measurements confirm the existence of d states on both sides of the gap.

The final reason why Cu_2O provides an ideal test subject for the development of optimization methods for constructing a Wannier orbital basis is a pragmatic one. Namely, I find a good agreement between theory and experiment when I perform calculations in TDDFT in the RPA on top of the GGA Kohn-Sham groundstate. In addition, I find that crystal local-field effects can be ignored, meaning that I only need to examine χ^{KS} , avoiding the inversion involved in calculating χ and the resulting, complicating effects of long-ranged screening. This allows a clean and simple comparison to be made to experiment and ensures that the orbitals that I construct have a concrete correspondence to reality (excepting the exciton series in the optical gap [132–138], which RPA is unable to reproduce). The achievement of an accurate calculation of the optical response for Cu_2O using the Bloch basis is so important that I will expand on it in detail in the next section.

2.1 Optical measurements and calculations in the Bloch basis

My ability to use the TDDFT calculation in the RPA is vital to my use of Cu_2O as a subject of investigation, but this seems to contradict the fact that Cu_2O has presented some considerable theoretical challenges. One issue that challenges first principles theory is the value of the optical gap in Cu_2O . Here, local approximations to the exchange correlation functional (such as the GGA which I utilize) suffer from the well known gap problem, underestimating both the photoemission and optical gaps [139]. This has served as a problem for the theoretical investigation of the optical spectrum as much of the experimental and theoretical interest has focused on the exciton series at the edge the gap [132–138]. There have been several attempts to fix the optical gap in Cu_2O , all closer to the experimental value of 2.4 eV and some in very good agreement. These include the use of hybrid functionals that incorporate non-local exchange [123, 140], “one-shot” GW (G_0W_0) on top of LDA [121], LDA+ U [19] and hybrid functionals [141], and so-called “self-consistent” GW [121, 142, 143].

The failure of GGA band structure to reproduce the experimentally measured optical gap opens the work I am presenting here up to a potential challenge. My intent is to use a calculation corresponding to experiment to test the role of local orbital degrees of freedom in the optical excitations, but this calculation is bound to be wrong for the value of the gap. This is especially true because I am using the RPA and do not have to consider crystal local-field effects, so the optical gap is precisely the smallest energy difference between Bloch eigenvalues for a particular $\mathbf{k}$.

The first thing to note in addressing this issue is that I am not concerned as much with features below the onset of the main optical features (the exciton series) as I am with the most prominent features in the dielectric response in the electron-volt range above this onset. Thus, it is essential to gain a full picture of the spectrum for energies of several electron-volts above the gap as measured experimentally. Several highly accurate spectroscopic ellipsometry measurements have been performed for Cu_2O

and all have been in very good agreement with each other [17–19]. The ellipsometry measurements have an advantage over earlier reflectivity measurements [138] in that they do not rely on Kramers-Kronig analysis and so they must be taken as the more reliable measurements of the dielectric response. The red dots shown in Fig. 2.2 are the ellipsometry data from Ito et al. [17]. This measurement reveals several sharp features at energies between ≈ 2.5 -5 eV. In particular, these are a small peak at ≈ 2.7 eV, a peak at ≈ 3.4 eV and a somewhat shorter peak at ≈ 4.2 eV.

The calculation I have performed in TDDFT (RPA) on top of the Kohn-Sham groundstate in GGA, as shown in Fig. 2.2 by the black curve, displays features in good, semi-quantitative agreement with the experimental measurement. There are some differences, as should be expected. The small peak at ≈ 2.7 eV in the experiment is a shoulder in my calculation, the peak at ≈ 3.4 eV in the experiment is lower by ≈ 0.2 eV in my calculation and both the ≈ 3.4 eV and 4.2 eV peaks in the experiment are about 10-20% lower than my calculation. Regardless of these minor differences, the agreement with experiment is quite good for the most prominent features above the optical onset. This agreement allows me to target these features in the optimization of the Wannier basis, as I will begin to discuss in the next section. However for completeness, I must first address the results from extensive theoretical investigations in relation to the GGA/RPA results I have obtained, which I will do in the remainder of this section.

The results of the current GGA/RPA calculation should be measured against other recent calculations in order to assess the relative quality of the agreement with experiment. For the most part, recent, first-principles, many-body theoretical studies calculated the gap by reading the energy differences between occupied and unoccupied states in the band structure. This is in effect taking the RPA on top of the calculated effective single-particle electronic structure. As I have stated, there are many methods that are able to improve the value of the gap. However, when the optical spectrum is calculated with a band structure from *GW* [145] (see the blue curve in Fig. 2.2 as an example) then it does not even agree qualitatively with experiment.

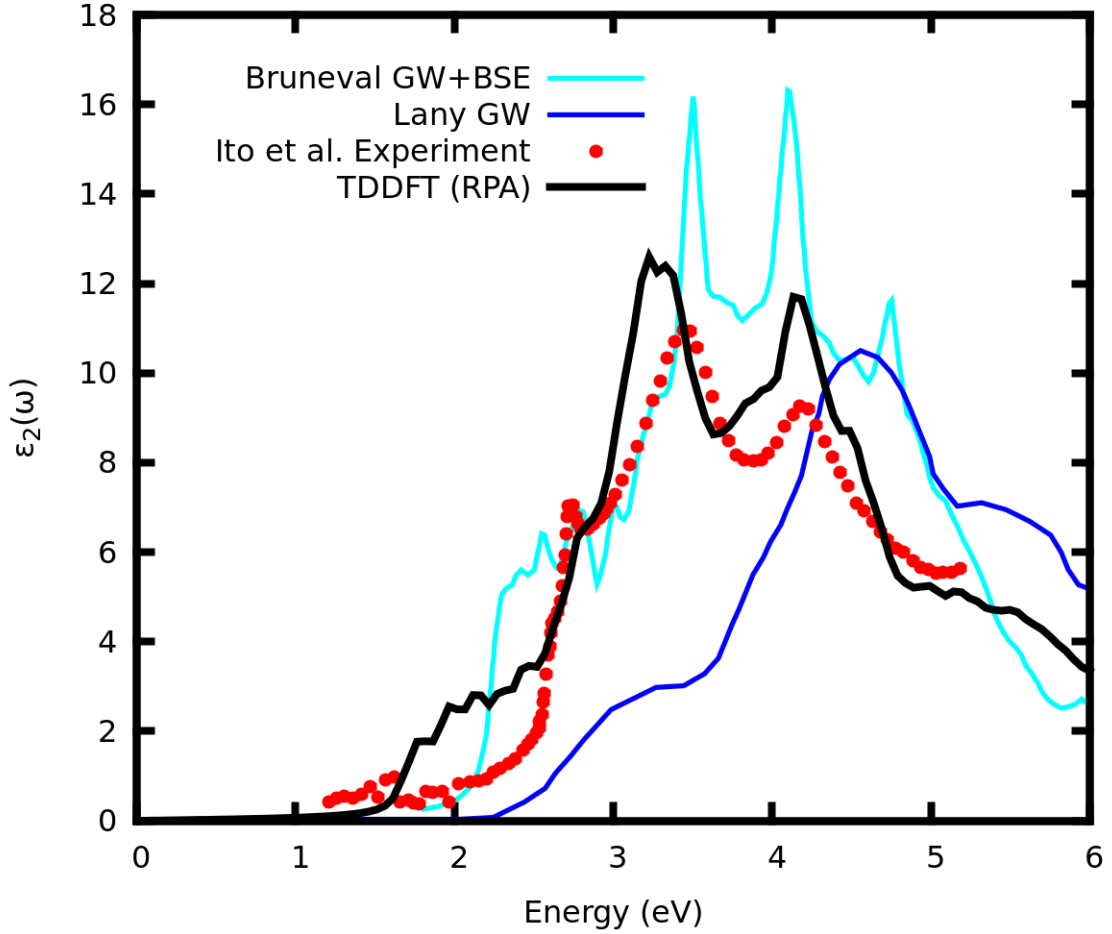


Figure 2.2: Calculated dielectric response of Cu_2O versus experiment. My TDDFT calculation of $\epsilon_2(\omega)$ using the Bloch basis in RPA (black line) together with experimental data from Ito et al. [17]. The calculation agrees well with the two prominent peaks between $\approx 3\text{--}4.5$ eV., though it differs from experiment in the region that contains the exciton series below 2.5 eV. It is expected that this series would not be produced because the calculation is done in the RPA. A self-energy correction via the GW can be introduced to widen the gap, but this requires an excitonic interaction to be introduced in the solution for the two-particle response via the Bethe-Salpeter Equation (BSE). The results of just such a calculation from Bruneval [144] is shown in blue. For comparison, the GW calculation of Lany [145] is shown in magenta. Clearly the Lany calculation is far from experiment. Bruneval’s calculation is much closer and has a superior optical gap to my own calculation. Also, the energy location of the ≈ 3.5 eV peak is closer to the experimental value. However, the strength of the peaks and the sharpness of the lineshapes are far too high and there is a peak at ≈ 4.8 eV that is absent from experiment. Thus, my calculation represents a better agreement with experiment for the most prominent peaks.

The solution lies in the need to compensate for the self-energy insertion used to open the band gap. As Bruneval et al. showed [121], if the gap is opened using an *ad hoc* GW type approximation for the self-energy which is used to shift the values of the Kohn-Sham eigenvalues, then it is essential that excitonic effects must be included in the two-particle problem. Otherwise, the optical features between ≈ 2.5 -5 eV are not accurately reproduced. The results of this calculation [144] are shown with the cyan curve in Fig. 2.2. Though the inclusion of an excitonic interaction can correct the problem introduced by the self-energy correction, the authors also claim that this is the only way to adequately describe the optical spectrum. This assertion cannot be accurate in light of the good agreement with experiment that I find in TDDFT (RPA).

I believe part of the resolution to this apparent contradiction lies in the effect on the band structure of opening the gap to achieve better agreement with experiment. When the gap is pushed wider by, e.g. QSGW, the occupied states that I will show are most relevant are largely unchanged while the unoccupied states are shifted upward in an approximately rigid fashion [142]. The effect that this would have on the spectrum as I have calculated it would be to shift my spectrum higher, ruining the agreement with experiment. This is what Bruneval et al. presents, using the G_0W_0 band structure, but then using an RPA dielectric function with this band structure as an input. That this gives a poor reproduction with experiment is in agreement with the conclusion from my calculation.

Yet, the conclusion of Bruneval et al. is that RPA is insufficient. This is certainly true when it comes to describing the optical gap and the exciton series at the edge of that gap, as is clear in my calculation. However, the numerical evidence in Fig 2.2 shows that RPA is sufficient (and apparently superior) for describing the peaks between ≈ 2.5 -5 eV. The resolution to this apparent discrepancy lies in relation of corrections to the self-energy that widen the gap to the excitonic interactions in the two-particle problem. If the self-energy corrections can be viewed as being introduced as in the theory of Baym and Kadanoff [146], then these corrections induce

an excitonic interaction in the calculation of the response function. The current GGA/RPA calculation represents a strict TDDFT calculation (as opposed to that of Bruneval et al. which departs from TDDFT by shifting the Kohn-Sham eigenvalues) and thus does not contain any *GW*-like self-energy corrections. As such it would be inappropriate to include an excitonic interaction in the calculation of the optical response. Indeed, the calculation agrees with experiment without them. However, when Bruneval et al. introduced a *GW* correction to the self-energy to widen the band gap, it requires that they use an excitonic interaction in the Bethe-Salpeter equation, just as they conclude.

Thus, the apparently discrepant views on the optical response of Cu_2O can be resolved. Though the GGA/RPA calculation suffers from an inadequate description of the gap, I am most concerned with the peaks at somewhat higher energy, so TDDFT in RPA is sufficient. Bruneval et al. were concerned with these peaks, but also desired a better description of the gap, so RPA is insufficient and excitonic interactions must be included. In addition, though correct energy location and intensity of the peaks are necessary, I am primarily concerned with the nature of the orbital channels that make up the peaks. For this purpose, I believe TDDFT in RPA provides the best foundation for analysis.

In addition, when the TDDFT (RPA) results and Bruneval et al. results with a *GW*-like self-energy and excitonic interactions included in the Bethe-Salpeter equation (BSE) are compared, as shown in Fig. 2.2 by the black and cyan curves respectively, it is apparent that the RPA result is actually superior. The peaks at ≈ 3.4 eV and 4.2 eV are about 30-40% higher in the BSE result versus 10-20% higher in my RPA result and the peaks in the BSE spectrum are much narrower than experiment. Certainly, my calculation must benefit to a greater degree from some cancellation of errors in order to achieve superior agreement with experiment, but the result make Cu_2O an ideal test case for investigation of local orbital degrees of freedom with regard to the optical spectrum.

2.2 Analysis of the optical spectrum involving single-site projected occupied orbitals

The remarkable agreement between the optical spectrum as measured by ellipsometry and the current TDDFT (RPA) calculation provides a direct connection to experiment that can be used to benchmark the methods I have developed to investigate the role of local orbital degrees of freedom in the optical excitations. The local orbital degrees of freedom are intimately connected to the local chemistry and symmetry of the material. Vital to the discussion of local chemistry and symmetry and its role in electronic structure is the transformation between the Bloch states and Wannier orbitals.

The equation relating the Bloch basis to the Wannier basis is given in Eq. (1.99). Other than the unitarity requirement for the $a_{\nu j}$ coefficients, one is free to choose them and this amounts to choosing a gauge for the Wannier transformation. The method for choosing the gauge for the transformation from Bloch to Wannier functions in the projection method using Eqs.(1.106) or (1.112) may be described in terms of making of three decisions. First is the desired symmetry of the Wannier orbitals, which is determined by the choice of trial orbitals. Second is the band range, which defines how many Bloch states are to be included at each point in the first Brillouin zone. Third is the local coordinate system for the trial orbitals used in the projection. The goal in making choices along these three lines is to use the enormous freedom that they offer to find an optimal set of orbitals.

These three decisions are interrelated and cannot be considered entirely separately. In Section 1.4, I introduced the method of projecting onto local orbitals to produce a set of $a_{\nu j}$ coefficients. The particular choice of local orbitals with respect to a local axis defines the symmetry of the resulting Wannier orbitals and greatly restricts the available energies for choosing an energy window. Certain orbital degrees of freedom will naturally sit in certain bands. Obviously, placing a p orbital in a largely d band (as determined by e.g. the amount of $\ell = 2$ content from the LAPW basis) will not

result in a successful projection in that linear dependencies will arise among the set of Wannier orbitals, making them unsuitable as a basis. Thus, choosing a set of local orbitals to use in a projection provides some restrictions on the gauge for the Wannier transformation.

The most important result of a particular choice of local orbitals is that they will largely determine the symmetry of the final Wannier orbital. For example, projecting using d local orbitals, which respect spherical symmetry, on a particular site will tend to result in Wannier orbitals that transform the same way as those orbitals under the symmetry operations of the crystal. In particular, for a cubic crystal with the conventional coordinate system, the t_{2g} orbitals will transform among themselves under the site symmetries and likewise for the e_g orbitals. It is important that the Wannier orbitals are influenced in this way by the local orbitals in the context of calculating optical matrix elements as it is the symmetry of the orbitals which determines the selection rules that should be respected. Also from a conceptual point of view, it is important for the Wannier orbital to share symmetry properties with atomic-like local orbitals so that it remains meaningful to call them “ d orbitals” at all.

However, I would like to make clear to what extent the symmetry of the Wannier orbital corresponds to that of the local trial orbital from which they are derived and what symmetry properties one should seek. First, it should be obvious that they will not have exactly the same symmetry of the local orbitals which reflect spherical symmetry. The crystal breaks spherical symmetry and the Wannier orbitals should only respect that reduced symmetry of the crystal. Second, the Wannier orbitals should provide some representation of the local symmetry operations and this representation should be as reduced as possible. Though asking the Wannier orbitals to produce irreducible representations is likely too restrictive, separating the orbitals into classes with respect to symmetry will aid in the final decomposition into orbital channels.

Though the symmetry of the trial orbitals significantly restrict which projections will be successful, there is considerable freedom with regard to the band ranges used in the projection. This represents the second primary choice for the projection to produce Wannier orbitals. The choice of band range is concretely the choice of the range over which the index j runs in Eq. (1.99). In the case of the occupied orbitals in Cu_2O , one natural starting point is to project five d orbitals per Cu site into the predominantly Cu- d -character band, which is largely disentangled from the predominantly O- p -character band. One could also project 5 d orbitals on each Cu site and 3 p orbitals on each O site into a large band range that includes both the p - and d -character bands. As I will show, this choice will reflect a decision as to how the hybridization between p and d orbitals manifests.

In order to present the effects of the choice of band range independent of the choice of local axis, which I will discuss later, I will choose the highest symmetry axis as a reasonable *a priori* choice. This is the c -axis of the cubic unit cell. With this choice of local coordinate system, the d orbitals break into degenerate manifolds of t_{2g} and e_g orbitals due to the cubic symmetry of Cu_2O , as shown in Fig. 2.3. In this case, if I were to find that any one of the orbitals in the t_{2g} or e_g manifold are important for reproducing the optical signal then all of them would have to be included.

The difference in the DOS for the two projection schemes, one window for both p and d orbitals and two windows which separates p and d , is illustrated in Fig. 2.3. The appearance of d orbital content between -7.5 eV and -5.0 eV and p orbital content between -3.0 eV and the Fermi level in Fig. 2.3a is a manifestation of hybridization between p and d orbitals. In restricting the energy windows into which the orbitals are projected such that the window between -7.5 eV and -5.0 eV consists entirely of O p orbitals and the window between -5.0 eV and the Fermi level consists entirely of Cu d orbitals (Fig. 2.3b), hybridization between p and d orbitals in the DOS is eliminated by construction. Of course, this hybridization is physical and reflects the chemistry of the compound, so the hybridization has to go somewhere. In exiting

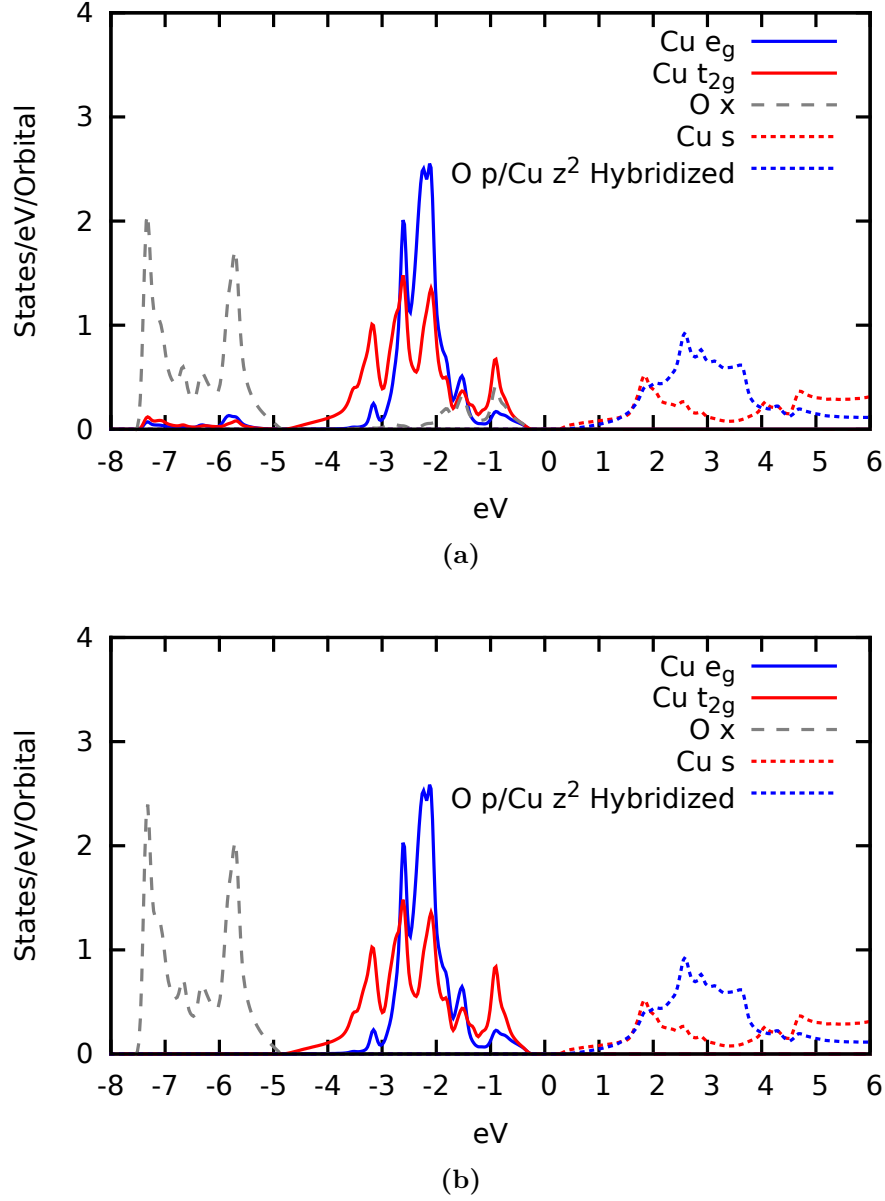


Figure 2.3: Cu₂O Wannier decomposed DOS comparison. In both (a) and (b) I projected five d orbitals per Cu site and three p orbitals per O site into the occupied band with the local coordinate axes coinciding with the cubic axes. In (a) all of the occupied orbitals are defined with the same energy window running from -8 eV to 0 eV, whereas in (b) there are two windows, one for the p orbitals running from -8 eV to ≈ -4.8 eV and one for the d orbitals running from ≈ -4.8 eV to 0 eV. Dividing the energy window into two separate ones for p and d orbitals eliminates hybridization between p and d orbitals with respect to the DOS, with the hybridization becoming manifest in real space.

the DOS, it will manifest as hybridization tails in the real-space representation of the Wannier orbitals.

This is illustrated in Figs. 2.4 and 2.5. The orbitals shown in Fig. 2.4 correspond to the DOS in Fig. 2.3a. There is clear hybridization between Cu d and O p orbitals in the DOS and since it appears there it does not appear in the real-space representation of the Wannier orbitals. However, the d orbitals are not necessarily atomic-like. They are still affected by the presence of the O p orbitals because they must be orthogonal to them by construction. This is especially apparent in the t_{2g} orbitals in Figs. 2.4c, 2.4d and 2.4e. The positive and negative lobes of these orbitals would be the same size for atomic orbitals, but they are distorted because of the Cu-O bonding chemistry.

When the energy windows used for projecting the Wannier orbitals are restricted so that there is one window for the d orbitals and one for the p orbitals, as in Figs. 2.3b and 2.5, the evidence of hybridization between p and d disappears from the DOS and appears instead in the real space representations of the orbitals. However, we see that all of the orbitals show a similar level of hybridization, which means that describing nature of the local chemistry would require consideration of all of the orbitals on the same footing. As I will show later, a choice of the local axis for projection that better reflects the chemistry of Cu₂O results in a clear classification of the orbitals according to their chemical nature. But first, I must determine whether the choice of energy window producing the orbitals in Fig. 2.5 is appropriate.

Without any consideration of how the orbitals are going to be used or what physical quantities are of interest, there is no reason to prefer a large p and d window or two separate windows, one for p and one for d . However, with regards to my criterion of explaining the optical spectrum as simply and with as few channels as possible, the choice of a large p and d window is not optimal. Though projecting onto a d -only occupied window means that the resulting Wannier orbitals are less atomic-like, they retain d -like symmetry which makes calling them d orbitals meaningful. In addition and crucially, it also reduces the number of channels involved when examining the

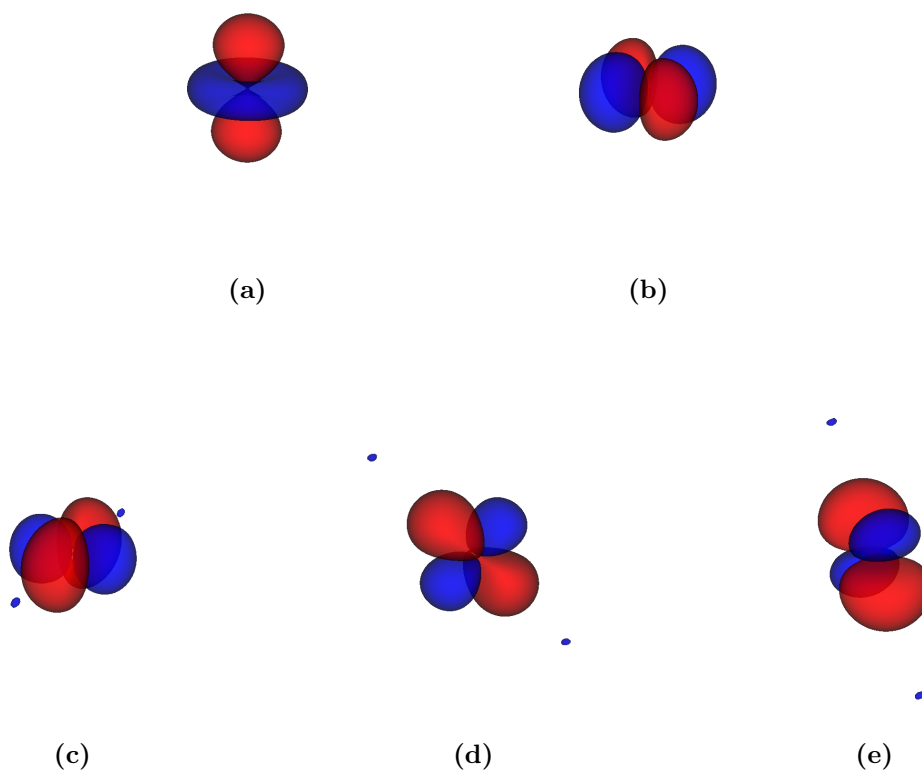


Figure 2.4: Cu_2O d -like Wannier orbitals. I produced these orbitals using an occupied window large enough to encompass all Cu d and O p orbitals. The local axis for the orbitals are aligned with the cubic axis. Blue indicates a positive value and red a negative value. They are, in alphabetical order of the labels, the z^2 , x^2-y^2 , xy , yz and xz orbitals. Due to the cubic symmetry of Cu_2O and my choice of local axis along the cubic axis, these are classified as e_g (z^2 and x^2-y^2) and t_{2g} (xy , yz and xz). Though hybridization with the O p orbitals is small for all orbitals, it is largest for the t_{2g} orbitals. Also, the central lobes are most distorted from the atomic-like symmetry for the t_{2g} orbitals. The negative lobes are larger than the positive whereas for atomic-like orbitals they would be the same size.

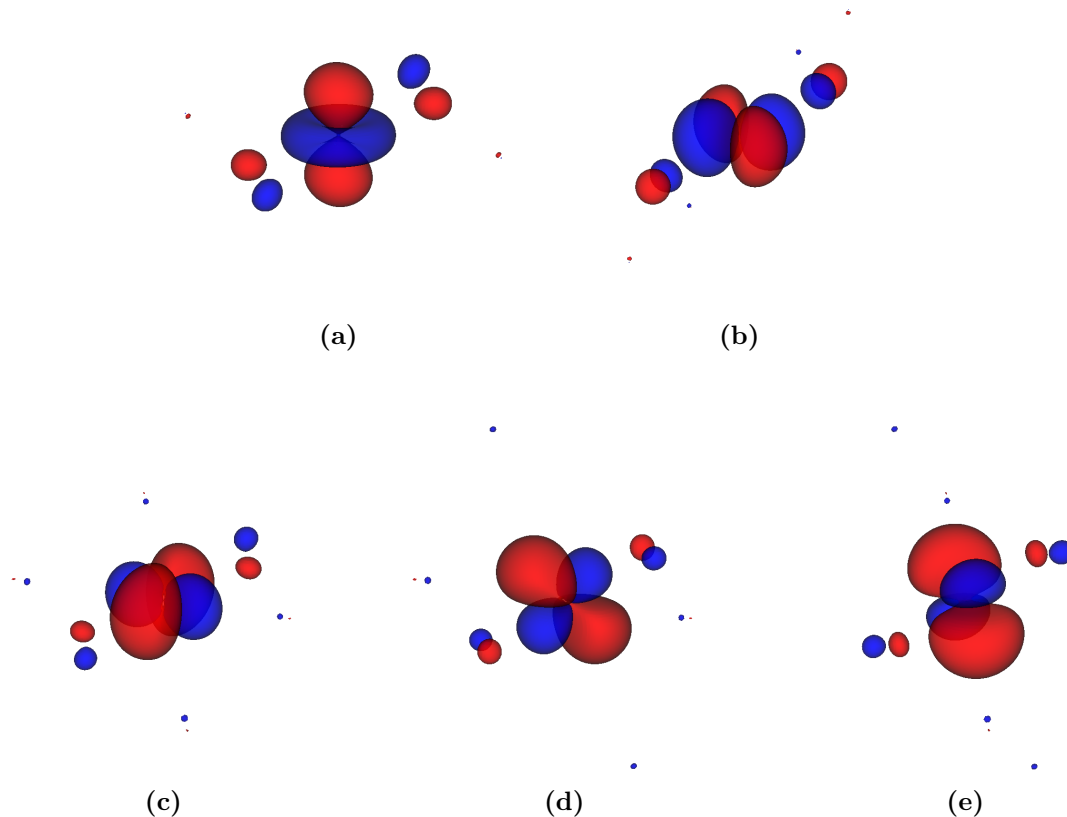


Figure 2.5: Cu_2O d -like Wannier orbitals with p hybridization. I produced these orbitals using an occupied window large enough to encompass only the Cu d orbitals. The local axis for the orbitals are aligned with the cubic axis. Blue indicates a positive value and red a negative value. They are, in alphabetical order of the labels, the z^2 , x^2-y^2 , xy , yz and xz orbitals. By restricting the d orbitals to their own window, separated from the p orbitals, I have pushed the hybridization from energy space to real space. This hybridization in real space manifests as p orbital tails on the neighboring O sites. All orbitals exhibit hybridization, but it is stronger for the e_g orbitals. However, a more judicious choice of local axis, along the O-Cu-O direction can place most of the p hybridization with the new z^2 orbital. The cubic axis is thus not the ideal from a chemical perspective.

decomposed optical signal, because the p Wannier orbitals are confined to energies far enough below the unoccupied states to be ignored. Thus, I will restrict the projection of the Cu d orbitals into an energy window that only includes the occupied bands that participate in producing the calculated spectrum in Fig. 2.2, because they encompass the highest occupied states. This window is labeled as “d-window” in Fig. 2.6.

The third primary choice to make in constructing the Wannier basis is the local coordinate system with respect to which the local trial orbitals are defined, which amounts to setting the rotation matrix $\mathcal{R}_{\ell,mm'}$ in Eq. (1.101). Examination of the nature of p - d hybridization leads to a superior choice of local axes compared to the axes along the edges of the unit cell. The d orbital lobes distort due to the presence of O sites due to chemistry. This distortion is visible in Figs. 2.4 and 2.5. The lobes represent where the orbital density is largest and if there is large density near a neighboring site. The total overlap between that orbital and orbitals on its neighboring sites must vanish due to orthonormality and that means that the orbitals must change shape with respect to their atomic-like counterparts to accommodate each other. Orbital overlap is also connected to chemical bonding. If there were no overlap, then there would be no bonding. So, one guess as to how I would construct orbitals in such a way as one or a few orbitals are primarily involved in bonding is to rotate the local coordinate system using $\mathcal{R}_{\ell,mm'}$ in Eq. (1.101) so that the lobes of one of the orbitals are pointing towards the bonded neighbors.

By examining the local chemistry of Cu_2O , I am led to a route which allows one orbital to be primarily involved in the Cu-O bonding. Each Cu atom is bonded strongly to its linearly coordinated oxygen neighbors. Placing the local z -axis along the O-Cu-O direction (the three-fold inversion axis) ensures that the z^2 orbital will have maxima in the direction of the O atoms. Remarkably, this configuration to the z^2 orbital has been observed experimentally [16], which lends further weight to this choice. In turn, it will have the greatest hybridization with the O p orbitals, a feature

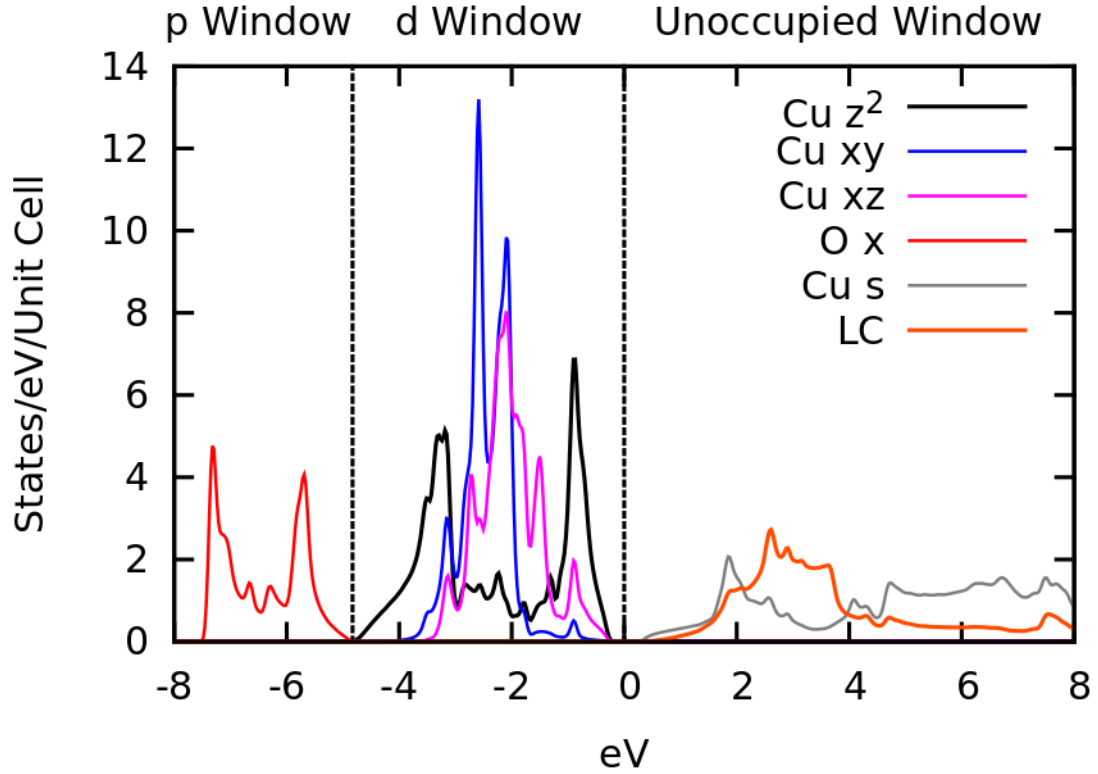


Figure 2.6: DOS due to various Wannier orbitals in Cu_2O . The energy windows into which I projected the orbitals are marked. Restricting d orbitals to their own window places hybridization with $\text{O } p$ orbitals in real space. The local z -axis is along the O-Cu-O direction so that the $\text{Cu } z^2$ orbital is aligned with the neighboring O atoms. The z^2 orbital shows a distinctive two-peak structure, which is qualitatively different from the other d orbitals. This structure is related to Cu-Cu bonding and anti-bonding states, in other words offsite hybridization with other $\text{Cu } d$ orbitals. The orbitals constructed with a linear combination of local orbitals are marked “LC.” A linear combination of local orbitals was necessary for these unoccupied states because of their delocalized nature. They span the entire unit cell and are not associated with any particular site.

observed experimentally [16]. Thus, the choice of the local z -axis along the O-Cu-O direction is appealing from the point of view of local chemistry. The dominant bonding in Cu_2O is between Cu-O nearest neighbors.

By isolating the z^2 orbital as having the strongest overlap with O sites I am able to reduce the number of orbitals participating in the bonding. Consequently, I am also able to reduce the number of origin states for particle-hole excitations involving the chemically active orbitals. Thus, this choice promises to be superior for a simple description of the optical spectrum. The choice of local axis in this particular case also suggests a general guiding principle for selecting the local axis in transition-metal oxides and related compounds: *choose the local axis such that the maxima of the fewest number of d orbitals are directed toward the neighboring anion sites.*

From a local symmetry perspective, choosing the local z -axis along the O-Cu-O direction produces a z^2 orbital that will be a basis function for the a_{1g} representation of the three-fold inversion site-symmetry group. In addition, it offers the possibility of a single, a_{1g} orbital—the z^2 orbital—being the only orbital that makes a significant impact on the optical signal. This is an example of reducing the representation of symmetries as much as possible to aid in classification of the orbitals.

The orbitals resulting from this choice of local axis are shown in real space in Fig. 2.7. The z^2 orbital in Fig. 2.7a is clearly the most strongly distorted by the presence of the O atoms and participates in a σ anti-bonding relationship with the O nearest neighbors. The yz and xz orbitals in Figs. 2.7b and 2.7c Also show strong hybridization with O nearest neighbors and exist in a π anti-bonding relationship. In contrast, the x^2-y^2 and xy orbitals in Figs. 2.7d show almost no hybridization with O nearest neighbors. Thus, choosing the local axis along the O-Cu-O bonding direction produces three subclasses of orbitals, with the z^2 in a class by itself.

With the symmetry of the trial orbitals, the band range and the local axis chosen for the occupied band, I have determined the Wannier basis spanning the single-particle states in that band. However, since I am interested in using the Wannier orbital basis to calculate particle-hole excitation spectra, it is necessary that I project

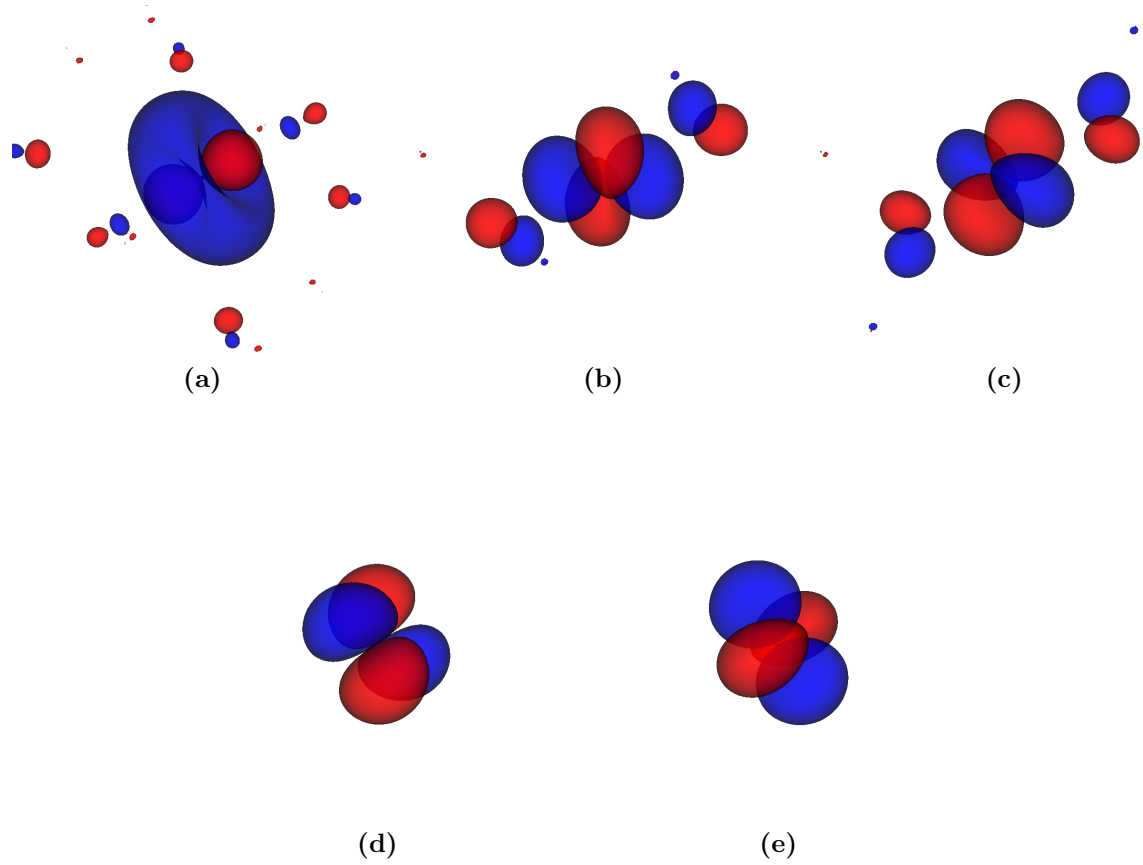


Figure 2.7: Cu_2O d -like Wannier orbitals with rotated local axis. I produced these orbitals using an occupied window large enough to encompass only the Cu d orbitals. The local axis for the orbitals are aligned with the O-Cu-O direction. Blue indicates a positive value and red a negative value. They are, in alphabetical order of the labels, the z^2 , x^2-y^2 , xy , yz and xz orbitals. By choosing a local z -axis long the O-Cu-O direction according to guidance from local chemistry and symmetry, I am able to produce orbitals that are qualitatively different chemically. The z^2 orbital in (a) is clearly the most hybridized and is heavily distorted by the presence of the oxygens.

orbitals not only onto the occupied band, but also the unoccupied band. Four of these unoccupied orbitals per unit cell are Cu 4s orbitals. I found that these can be projected into the unoccupied band using the same, single-site projection method as described for the occupied orbitals. However, in order to construct a set of Wannier functions that do not exhibit any linear dependencies, more than four bands are required. As a result, the unoccupied band structure produced with only Cu 4s orbitals does not faithfully reproduce the Bloch band structure and the optical spectrum produced with only these orbitals is ruined.

I found that the solution to this problem required the projection of three additional orbitals. For these orbitals, unlike the 4s orbitals, the single-site projection method is insufficient. This is because the required orbitals are strongly hybridized between O p and Cu z^2 (produced with the z -axis along the Cu-O-Cu direction) degrees of freedom and they are delocalized across the unit cell. The orbital in real space is shown in Fig. 2.8. This orbital shares the same symmetry as the O p orbital at the center of the unit cell. However, it cannot be associated with the O site *per se*. There are only three such orbitals per unit cell, not three per O site, so it is more appropriate to associate the orbital with the unit cell as a whole. This makes sense when examining the orbital in real space as well. It is spread across the unit cell and not strongly centralized on the O site.

In order to evaluate the success of the orbital basis using site-centered occupied orbitals with regard to providing an explanation of the optical spectrum in terms of orbital channels, I calculated the optical response due to the Wannier orbitals resulting from this projection. In performing this calculation I found that, in order to converge the results with the summation over $\mathbf{R}$ and $\mathbf{R}'$ in Eq. (1.128), I had to use the expression in Eq. (1.134) in which I carried out the sum analytically. This is illustrated in Fig. 2.9 for one class of particle-hole channels, where the hole is left in the Cu(1) z^2 orbital and the particle is in one of the orbitals shown in Fig. 2.8. I focused on only one particular class of channels, because the computational demand to study convergence including all of the relevant channels was too great. Decent convergence

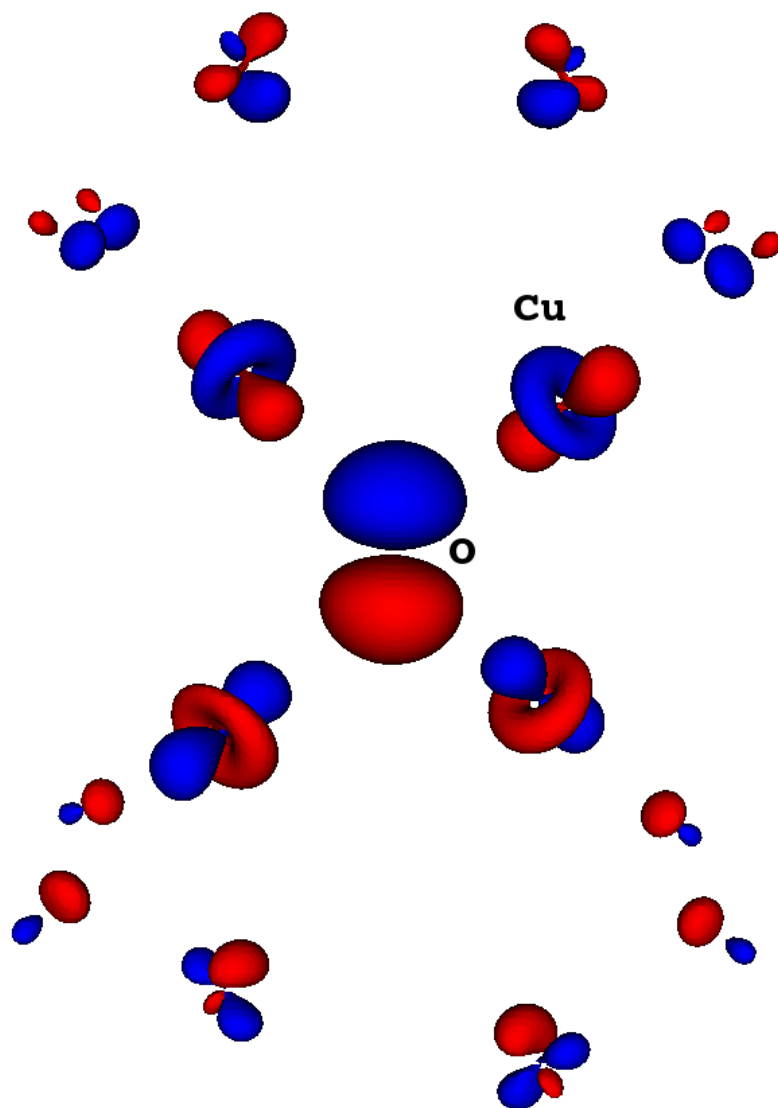


Figure 2.8: Cu_2O Cu z^2 /O p hybridized orbital. I produced this orbital using a linear combination of local orbitals, but the orbital with the highest weight in the trial linear combination is the O p orbital at the center of the unit cell. The projection onto a primarily O p orbital means that the Wannier orbital displays p -like symmetry, as is clear from the figure. For example, there is a four-fold inversion axis going through the oxygen site and one can see that under this operation this orbital changes sign, as does the p orbital. It is also apparent that there is a strong contribution from the Cu z^2 orbital. This may help to explain the strong coupling that the occupied z^2 orbital has with the unoccupied states. Though the primary weight in the projection is the O p orbital, the Cu z^2 orbital has roughly equal weight in the final orbital. This reflects the fact that the principle determinant of the final shape of the orbital rather than the details of the trial linear combination.

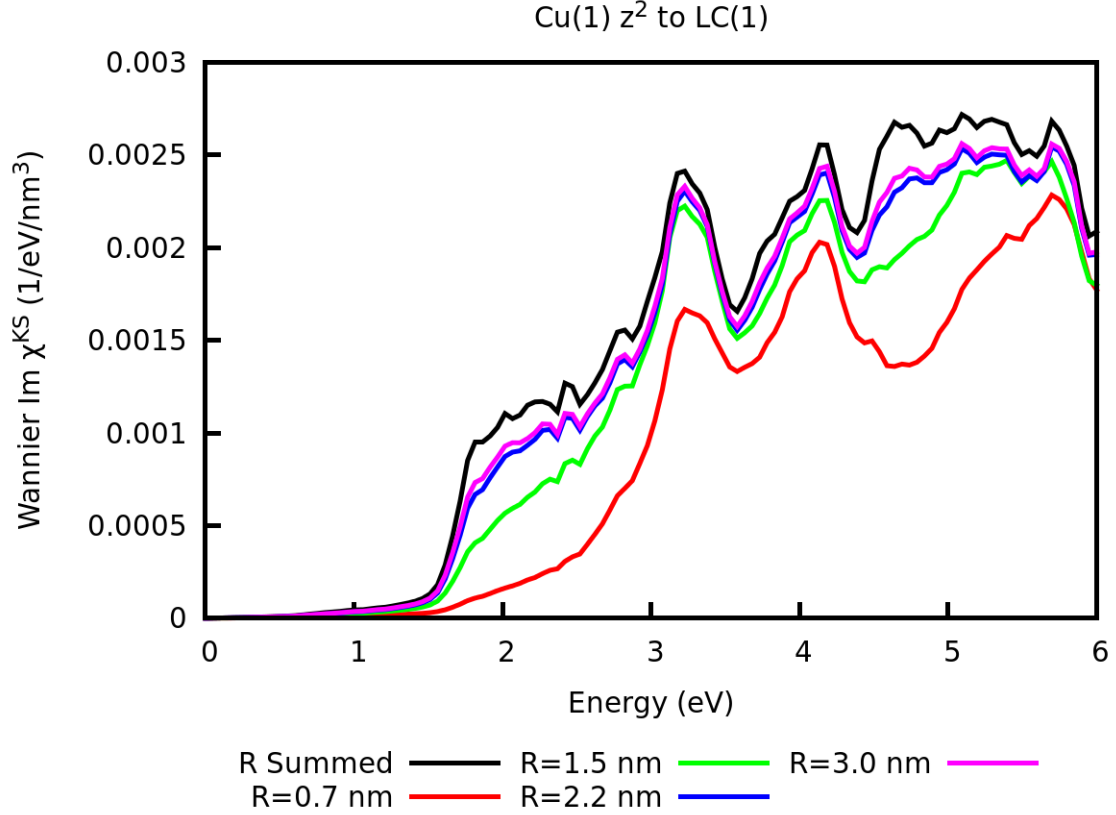


Figure 2.9: Cu_2O convergence of Wannier channels with $\mathbf{R}$ summation. Shown is the convergence of the imaginary part of χ^{KS} in the optical limit with summation in lattice vectors $\mathbf{R}$ for one class of channels. For each curve, I have only included particle-holes in unit cells whose origins are separated by a vector with magnitude less than the designated value. This gives a measure of the size of the particle-holes that make up the spectrum. I have examined only a subset of channels with the Cu z^2 orbital as the origin and Cu $z^2/\text{O } p$ hybridized orbital as the destination. This was due to the numerical cost of calculating the Wannier response for the largest separations. The black curve gives the result from analytic summation over $\mathbf{R}$ vectors. Clearly, particle-holes separated by at least one nanometer must be included in order to accurately represent the spectrum, with significant improvements up to two nanometers.

is not achieved until $|\mathbf{R}|$ is between 2 and 3 nm. Summing $\mathbf{R}$ analytically is much less computationally demanding and I am able to include all channels in the analysis.

The ability to analytically sum particle-hole channels over lattice vectors $\mathbf{R}$ allows me to decompose the optical spectrum in terms of classes of channels. The results are shown in Fig. 2.10. I have decomposed the spectrum into classes of channels differentiated by the orbital hole (i.e. the formerly occupied orbital) and including all particle orbitals (i.e. the formerly empty orbital). These classes each make qualitatively different contributions to the total spectrum. By far the largest contribution is made by the particle-hole transitions with a Cu z^2 hole. The spectrum due to these channels make up roughly half the intensity of the two main peaks, at ≈ 3.4 and 4.2 eV, whereas the transitions involving the yz hole make up roughly a sixth of the total. The spectrum involving the xy hole make up even a smaller contribution, almost nothing in the first main peak and roughly the same as a yz channels in the second.

It is not only the intensity of the spectrum for the different classes, but the lineshape that serves to distinguish them from each other. The spectrum due to the z^2 hole shows clear, sharp peaks that are very similar qualitatively to those in the total spectrum. The z^2 spectrum looks roughly like a scaled down version of the total spectrum. In contrast, the spectrum due to the yz hole displays some sharpness in the lineshape, but the peaks are not as clearly separated as for the z^2 spectrum. The spectrum due to the xy hole contains no sharp features at all. Since the lineshape of the spectrum due to the z^2 hole, in addition to its intensity, stands out compared to the other channels, it is clear that the z^2 orbital is the critical hole for the optical spectrum. However, there is considerable room for improvement, as the channels involving the z^2 orbital can only account for about half of the total signal.

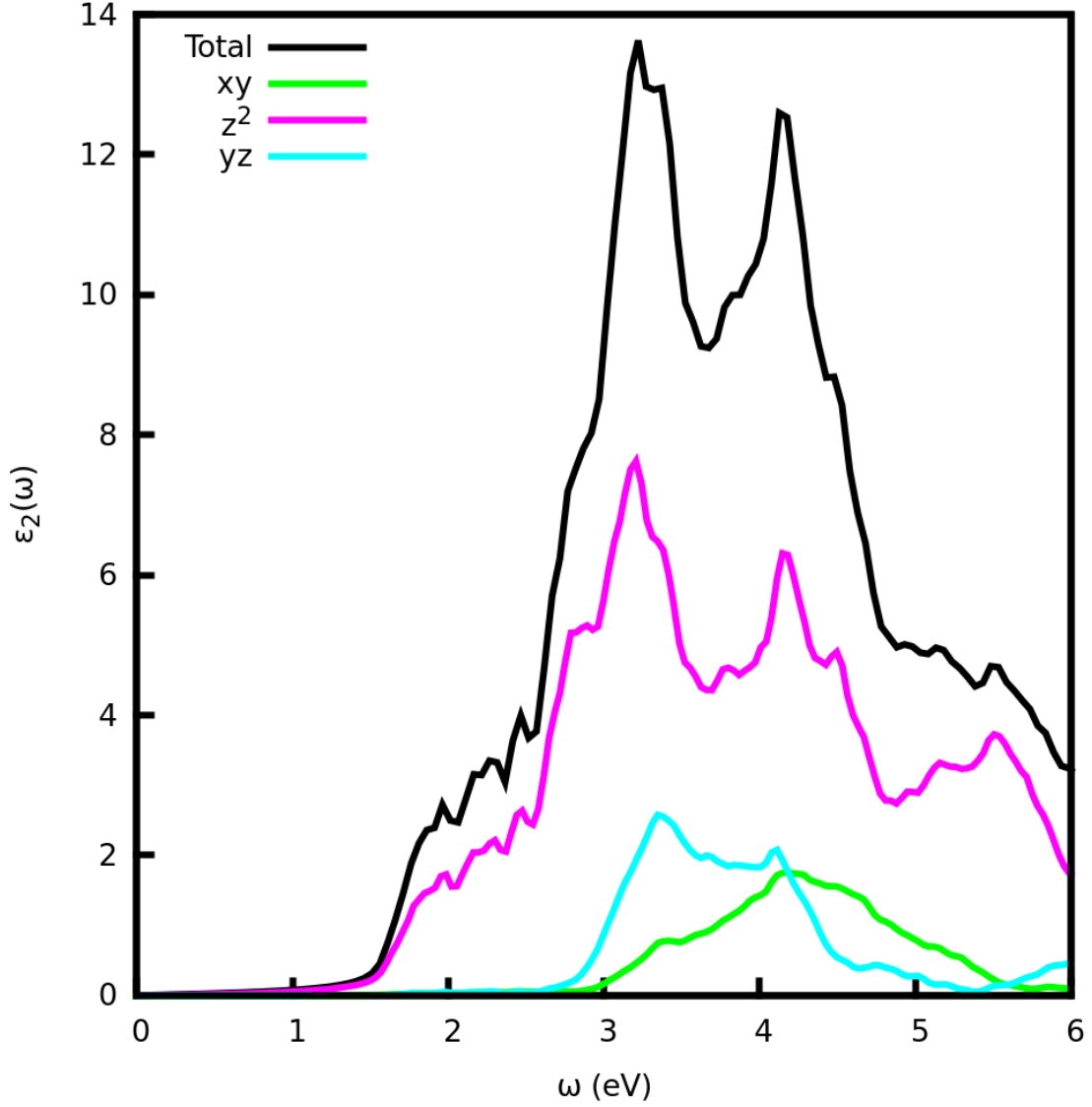


Figure 2.10: Cu_2O optical response decomposed using single-site projected Wannier orbitals. In all of the decompositions of the optical spectrum, I have performed the sum over lattice vectors $\mathbf{R}$ separating the unit cells of the particle and hole analytically. Shown is a Wannier decomposition with respect to the occupied orbitals using a set that includes five occupied, single-site projected d orbitals per Cu site. The class of orbital particle-hole channels that have the z^2 as the hole make the largest contribution, but only account for about half the total signal. The spectrum due to the xz and x^2-y^2 orbitals is related by symmetry to the ones for the yz and xy orbitals, respectively, and not shown.

2.3 Analysis of the optical spectrum involving molecular z^2 orbitals

I have shown in the discussion thus far that using single-site projected trial orbitals and calculating the response due to these already offers significant insight, particularly the importance of the z^2 orbital to local chemistry, symmetry and the optical response. However, there are clues that there is room for further optimization. A key observation that points to how such an optimization can be achieved is that the z^2 orbital resulting from this projection displays a striking two-peak structure in the DOS. Since the energy locations of features in the optical signal are primarily controlled by the energy difference between the particle and hole states, one would expect that the states in the lower peak in the DOS would not participate in the low-energy optical excitations. Thus, it becomes clear that if I were to isolate the upper from the lower peak and construct one independent orbital per Cu site for the states in each peak, I could include only those Bloch states that are important for producing the optical signal in the construction of a refined basis of Wannier orbitals. Once this has been done, I can calculate the contribution to particle-hole transitions involving just one orbital per site with a DOS at the top of the occupied band.

Vital to isolating an independent Wannier orbital representing the top peak in the z^2 DOS, as well as projecting onto the delocalized empty states, is the projection onto linear combinations of local orbitals resting on neighboring sites as described in 1.4. In particular, I must define the $\mathcal{U}_{\nu\mathbf{R}\nu'\mathbf{R}'}$ coefficients in Eq. (1.114). In that section I presented these as elements of a unitary matrix, but in practice I am allowed considerably more leeway, because after the projection the orbitals are orthonormalized. In general the $\mathcal{U}_{\nu\mathbf{R}\nu'\mathbf{R}'}$ can be elements of a rectangular matrix and the precise values can be quite small even if the resulting hybridization is significant. The shape of the final orbital in real-space is determined by the band range we are projecting into and the symmetry of the trial functions coupled with the local chemistry embodied in the DFT solution of the ground-state. The most

important things to keep in mind when deciding on the $\mathcal{U}_{\nu\mathbf{R}\nu'\mathbf{R}'}$ is that they should not unnecessarily break the natural symmetry of the orbital and they should include all the necessary orbital content for a successful projection.

In the case of Cu_2O , the primary function of projecting onto a linear combination of local orbitals for the occupied orbitals is to allow for large contributions from neighboring sites to the Wannier orbital I am trying to construct. This allows me the freedom to optimize my orbitals using a straightforward heuristic developed as part of this study. Namely, I seek to eliminate as much as possible hybridization in energy space, especially among the d orbitals. By doing so the hybridization is pushed into real space, manifesting as large, offsite contributions to the orbital that embodies important aspects of the local chemistry, in particular cation-cation bonding. By pursuing this strategy, I not only find an optimally simple reproduction of the optical spectrum, but I disentangle the d orbitals. The elimination of hybridization in energy space is achieved by selection of the energy windows used in the projection.

The key to elimination of hybridization in energy space is the recognition of the dual bonding and anti-bonding nature of the z^2 orbital, as illustrated in Fig. 2.11. Selection of energy windows for projection is driven by the two peak structure of the DOS of the z^2 , as shown in Fig. 2.11a. By using energy windows from -5 to -3.5 eV and from -2 eV to the Fermi level, i.e. windows that separately encompass these split states, I can incorporate the chemical nature of this splitting as hybridizations with d orbitals on neighboring sites. Plotting the resulting orbitals in real space makes it clear that these orbitals embody cation-cation anti-bonding (Fig. 2.11c) and bonding (Fig. 2.11d) states.

Important to note is the appearance of a hybridization peak in the DOS for the z^2 anti-bonding orbital shown in Fig. 2.11a. This peak was not present in the DOS for the single-site projected z^2 orbital but appears once I construct the anti-bonding orbital. The construction of an anti-bonding orbital where the anti-bonding state is with the neighboring Cu atoms requires strong, off-site hybridization with other d orbitals. The additional peak at ≈ -1.5 eV in the DOS for the z^2 anti-bonding orbital

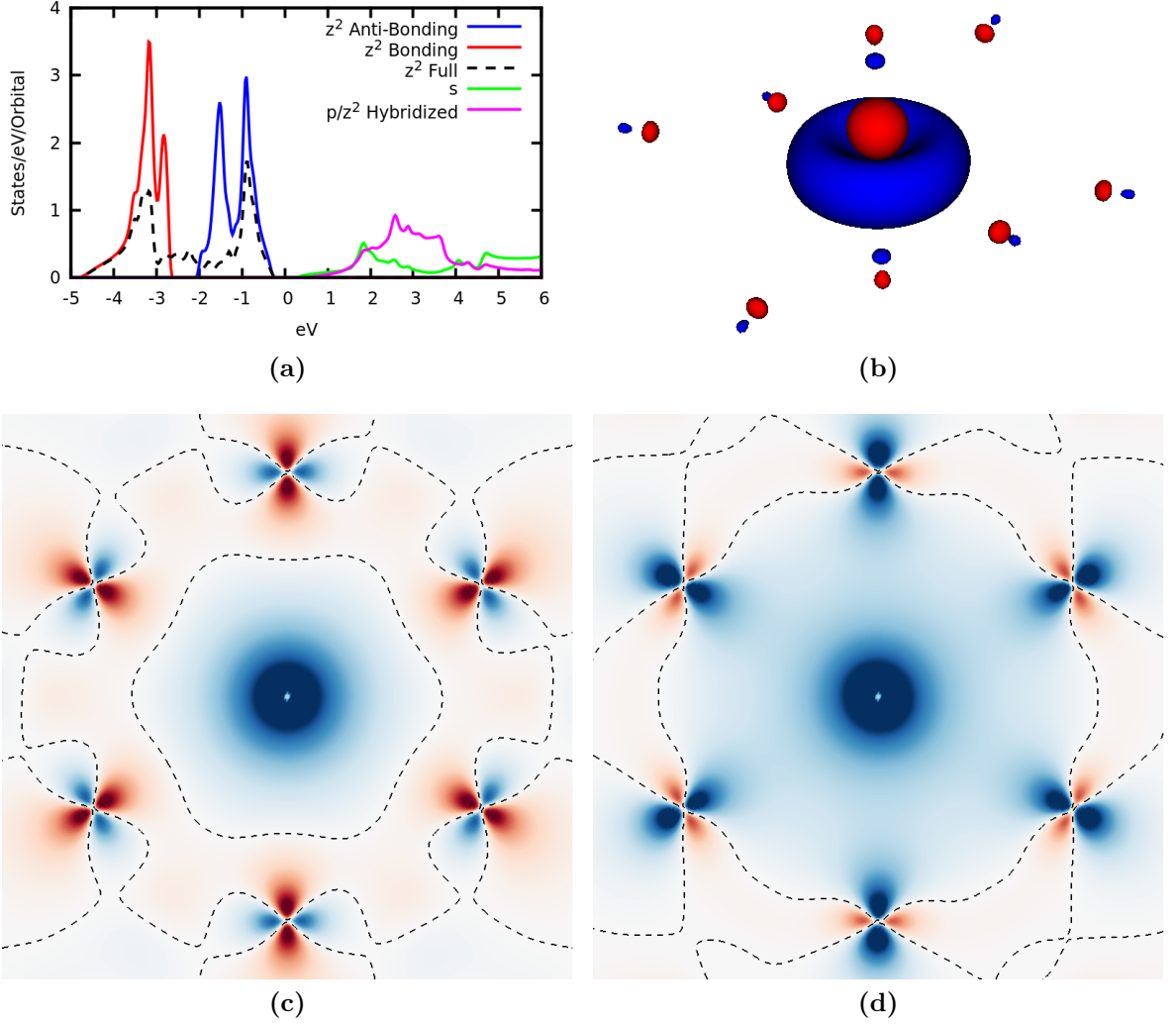


Figure 2.11: The Cu_2O z^2 bonding and anti-bonding orbitals. The critical hole for the particle-hole excitations making up the low-energy optical spectrum is a molecular orbital with z^2 symmetry in an anti-bonding configuration with d orbitals on neighboring Cu sites, forming a hexagonal ring belonging to a second Cu-O network relative to the central Cu site. This orbital is shown plotted in real space in (c) and the corresponding bonding orbital in (d). The conventionally constructed z^2 orbital, shown in (b), exists in a superposition of states involving both bonding and anti-bonding character. This superposition is clearly visible in the DOS, shown in (a). Additional peaks in the DOS of the bonding and anti-bonding orbitals, not present in the superimposed z^2 orbital, are due to the off-site hybridizations with other d orbitals, which are visible in (c) and (d). Inclusion of these additional d - d hybridizations is critical to fully representing the low-energy optical spectrum.

reflects the inclusion of these hybridizations. When a set of orbitals for the occupied states is constructed using only localized d orbitals, this peak is Cu yz (or equivalently Cu xz) character. Thus, the additional peak is due to hybridizations with other d orbitals.

In elucidating the role of hybridization in the z^2 anti-bonding orbital, it is useful to compare with the previous level of optimization, which I described as a narrowing of the band range from a window that included p and d orbitals to one of only d orbitals, the reduction of the band range resulted in hybridizations with p orbitals from neighboring O sites being rendered in real space, as shown in Fig. 2.5. In the case of the optimized bonding/anti-bonding orbitals, the hybridizations are with Cu d orbitals on neighboring sites in addition to O p orbitals. But in both the former and this latter, more refined case, the inclusion of hybridizations in real space is intimately tied to the local chemistry of Cu₂O. In the former case, including hybridizations which are O p -like resulted in the bonding between Cu and O being manifested in the shape in real space of the hybridized d -like orbitals. In the latter case, the inclusion of d hybridization along with the reduction of the band range resulted in the bonding between Cu's being embodied in d -like hybridization tails on neighboring Cu sites.

I stress that this manifestation of local chemistry was not the result of any bias I imposed on the construction of the orbitals. Instead, the bonding chemistry of the z^2 anti-bonding orbital is embodied in the Kohn-Sham eigenfunctions which encode the physics of the DFT groundstate. The linear combination of local orbitals that I used in the projection resulting in the z^2 anti-bonding orbital was necessary to allow the compression of the energy window used in the projection and the separation of the orbital from its fellow orbitals in energy space.

Now that I have been able to construct the z^2 anti-bonding orbital it is necessary to calculate the optical response due to transitions from this orbital to the unoccupied orbitals in order to evaluate the level of optimization with respect to the spectrum. The results are shown in Fig. 2.12. Again, in order to calculate the response due to these orbitals, I had to use the expression in Eq. (1.134) to sum over lattice vectors

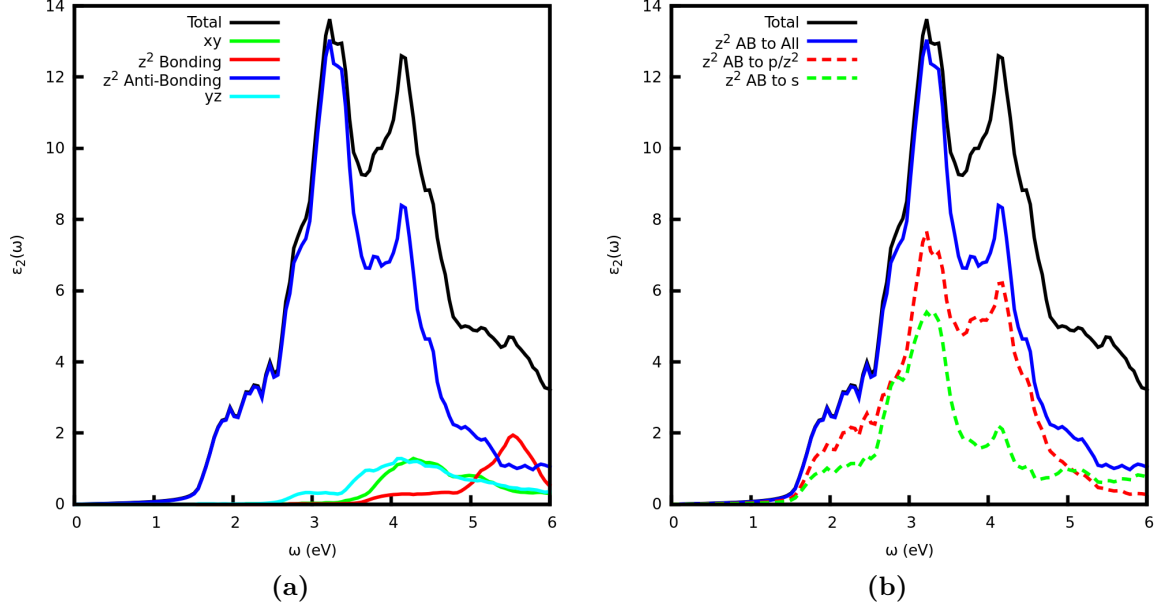


Figure 2.12: Cu_2O optical response decomposed using molecular Wannier orbitals. In all of the decompositions of the optical spectrum, I have performed the sum over lattice vectors $\mathbf{R}$ separating the unit cells of the particle and hole analytically. In (a) is a Wannier decomposition with respect to the occupied orbitals using a set in which the z^2 orbital is split into independent bonding and anti-bonding parts. Here, the class of channels involving the anti-bonding z^2 orbital can completely account for the lowest energy spectrum and displays the distinguishing sharp features of the lineshape. Once I have identified the critical hole, I can investigate the role of classes of channels broken down by particle orbitals. In (b) is a Wannier decomposition focusing on the contribution of the anti-bonding, z^2 , occupied orbital decomposed according to the types of unoccupied orbitals. There are two classes of channels that make up the signal, one involving a Cu 4s particle state and the other to a Cu $z^2/\text{O } p$ hybridized particle state.

analytically. The result is shown in Fig. 2.12a. Clearly, the low-energy line-shape, particularly the sharp peaks, is due entirely to the z^2 anti-bonding orbital. The Wannier basis with this orbital and its z^2 bonding partner is superior to the spectral decomposition using site-centered orbitals shown in Fig. 2.10. Thus, separating the bonding and anti-bonding parts of the z^2 orbital into independent orbitals is optimal with regard to the criterion that I am using here. Namely that the orbital basis gives the simplest description of the low energy optics.

The fact that the sharp line shape of the dominant peaks is fully explained by an occupied z^2 anti-bonding orbital sheds light on the sharpness in the lineshape due to the occupied yz orbital, shown in Fig. 2.10. This sharpness was due to hybridization with the z^2 orbital. When I included off-site hybridizations with the other d orbitals to produce the anti-bonding z^2 -orbital, all of the sharpness in the features became the product of transitions involving only that orbital. Thus, I find that the appearance of sharpness in the spectrum due to the yz orbital when my set of orbitals included only a localized z^2 orbital was a feature of hybridization with the z^2 orbital, not a feature of the yz orbital itself. This is further confirmed by the appearance of a peak due to off-site hybridizations in the DOS due to the anti-bonding z^2 orbital, as discussed above.

One subtlety in the difference between the spectrum due to the superimposed z^2 orbital in Figure 2.10 and the independent z^2 anti-bonding orbital in Figure 2.12a deserves mentioning. Namely, it is essential that the anti-bonding and bonding orbitals be separated in energy space as much as possible from each other and the rest of the d orbitals. One may think that the mathematics of χ^{KS} —namely that the location of a feature is controlled by the energy differences between the participating orbitals (see Eq. (1.98))—would mean that the anti-bonding part of a conventionally constructed, superimposed z^2 orbital would contribute to the spectrum in the same way as an independent anti-bonding orbital. This is not the case: the independent anti-bonding orbital is the essential hole in transitions that saturate the low-energy spectrum. In contrast, the superimposed z^2 orbital only accounts for about half of the

spectrum. This highlights the connection between optimization with respect to optics and disentanglement of orbitals, which will become more important in the discussion of the construction of downfolded Hamiltonians in Chapter 4.

The discussion thus far has been restricted to analysis of the occupied states. In order to develop a complete picture of the optical excitations, it is necessary to address the unoccupied states as well. The unoccupied orbitals consist partly of four orbitals per unit cell of Cu $4s$ character and three orbitals per unit cell with O p symmetry, which are equally of O p and Cu z^2 character. This gives two fundamentally different types of unoccupied orbitals and it is enlightening to further decompose the optical spectrum according to transitions involving these two types separately. The results in Fig. 2.12b offer a number of insights.

First, the decomposition of the signal arising from the optimized z^2 anti-bonding occupied orbital according to the unoccupied orbitals reveals two classes of transitions. One class involves the Cu $4s$ unoccupied orbital and exhibits two main features: a shoulder at 2.8 eV and the largest peak at 3.3 eV. The other class involves a O p /Cu z^2 hybridized orbital with O p symmetry shown in Fig. 2.8. This class also exhibits two main features: peaks at 3.3 and 4.2 eV. This shows that the differing symmetry and chemistry that serves to distinguish the $4s$ and p/z^2 hybridized orbitals also results in a qualitatively different spectrum that also can be used to classify them.

Second, the prominent two-peak structure of the lineshape is already embodied when considering only transitions involving the z^2 anti-bonding and the unoccupied p/z^2 hybridized orbitals. The full saturation of the optical signal for the first main peak is achieved with the addition of a strong contribution from transitions involving the z^2 anti-bonding and Cu $4s$ orbitals. Thus, it makes sense to point to the transitions involving the p/z^2 hybridized orbitals as those primarily responsible for the prominent two-peak structure in the optical spectrum. These peaks are those that are best described in RPA indicating that this is perhaps a good approximation particularly for particle-holes involving the z^2 anti-bonding and p/z^2 hybridized orbitals.

Third, the small shoulder at ≈ 2.8 eV, which is sharper in the experiment, is also due the z^2 anti-bonding to $4s$ transitions. The small peak just above 4 eV in the z^2 anti-bonding to Cu $4s$ spectrum is likely due to hybridization between the two types of unoccupied orbitals. Since this peak is on the edge of the exciton series that has been measured at lowest energies, it seems reasonable that the primary orbital degrees of freedom participating in the exciton series are the z^2 anti-bonding and $4s$ orbitals. One may speculate that these orbitals are those that must be specially treated in order to improve the optical gap and exciton series over the results from TDDFT in the RPA with GGA band structure.

It is not only the orbital character of the particle-hole transitions that I am able to determine. Since the basis set of Wannier orbitals that we are using to decompose the optical spectrum is natural for describing physics in real-space, we are able to explicitly determine the spatial range of the particle-hole transitions that contribute to the spectrum. In the case of Cu_2O I found that transitions involving particle-hole pairs separated by more than a nanometer contributed significantly to the optical signal. This is due to the delocalized nature of the orbitals involved, the possibility of long-wavelength coupling for optical transitions and the three-dimensionality of the transitions means that the number of transitions grow as $\sim R^2$. The large size of the particle-hole means that in order to calculate the spectra shown in Fig. 2.12, I summed analytically over all distances between particle-hole pairs.

The appropriateness of calculating the response in the Wannier basis by analytically summing over lattice vectors leads to further insight. The expression in Eq. (1.134) represents a return to a momentum space picture in which the optical transitions can be viewed as vertical, interband transitions, especially those involving pockets in the Brillouin zone. Thus, with the decomposition shown in Fig. 2.12b in hand, I would like to return to k -space and some key insights into Eqs. (1.134), (1.133) and (1.135) allows me to do this. The $a_{\nu j}(\mathbf{k})$ Wannier coefficients appear four times each for the origin and destination orbitals, heavily weighting each contribution in the first Brillouin zone. Also, the $a_{\nu j}(\mathbf{k})$ for origin and destination

are multiplied so both must be relatively large for a particular vertical transition to contribute significantly. Thus, I can deduce what parts of the first Brillouin zone are making the largest contributions to a feature produced by a particular class of channels.

The momentum space analysis of the spectrum is summarized in Fig. 2.13. For the class of transitions involving the Cu 4s hole, shown in Fig. 2.13b, a pocket on the line from X to M and one around the M point produce the 2.8 eV and 3.3 eV features, respectively. This region of the Brillouin zone is excluded for transitions involving the O p /Cu z^2 hybridized orbitals because the $a_{\nu j}(\mathbf{k})$ coefficients for these orbitals are small along the X-M line, as visible in Fig. 2.13c. Remarkably, the 3.3 eV peak for the class of orbitals involving the p/z^2 hybridized orbitals has significant contributions from completely different regions: a pocket around Γ and one around R. A pocket on the line X-R makes the significant contribution to the 4.2 eV peak. Thus I have found that one peak in the optical spectrum at 3.3 eV can be viewed as two distinct peaks arising from different classes of Wannier orbital transitions and each class's main contribution is associated with completely different parts of the Brillouin zone.

Though I have gained insight by returning to a momentum space picture, there are significant advantages to beginning from a local, Wannier picture rather than considering an interband picture *a priori*. First, I am able to establish using the Wannier picture that the optical transitions are very delocalized in real-space which makes a picture where I consider transitions that are localized in momentum-space a natural one. Second, there is still a vital role played by matrix-elements, as seen strikingly in the out-sized contribution from the z^2 anti-bonding orbitals to the optical response. Third, once I know which subset of orbitals are responsible for the optical features, it makes unambiguous the assignment of which pockets in the Brillouin zone are important to the optical response.

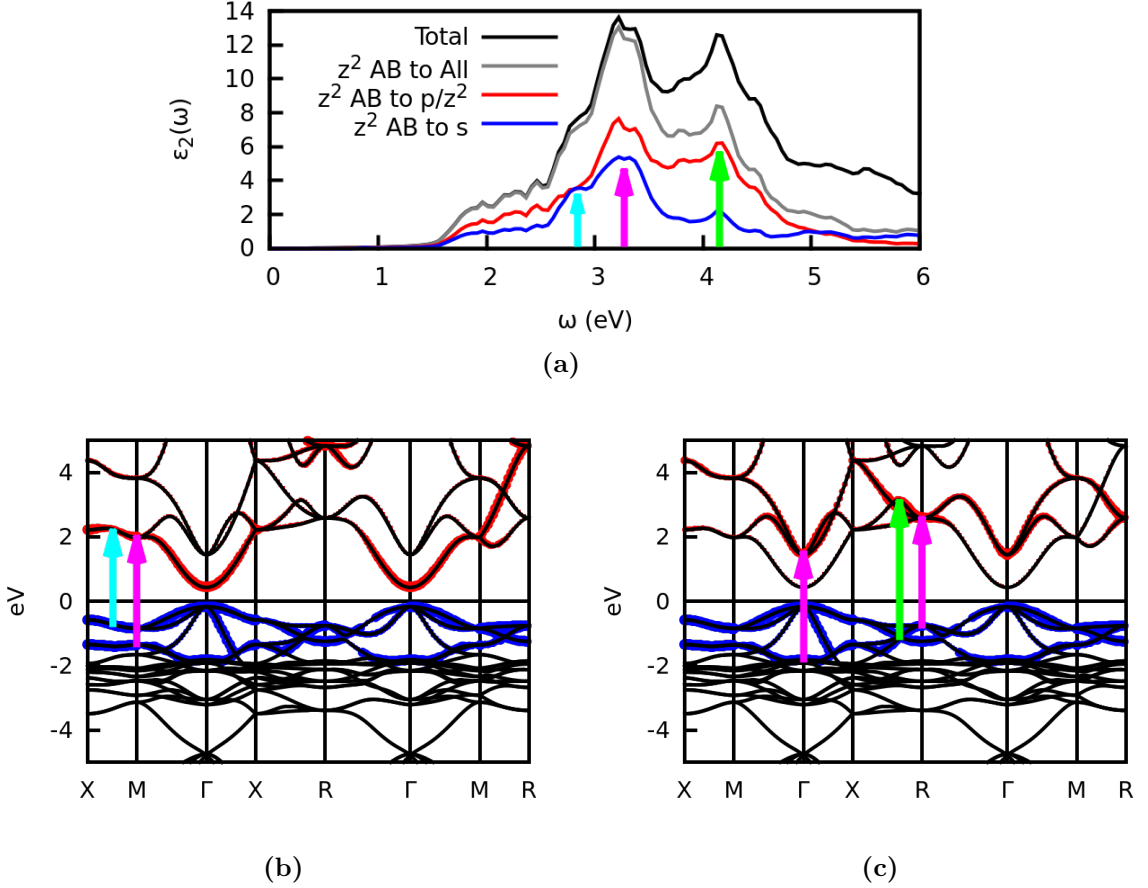


Figure 2.13: A momentum space picture of the optical spectrum of Cu_2O . Since the long-ranged nature of the particle-holes that make up the optical spectrum required me to analytically sum over distance between particle and hole, it becomes natural to return from a real space picture to a momentum space one. The decomposition of the spectrum into classes of Wannier channels aides in this endeavor. In the optical limit, a particular channel can only contribute if there is finite orbital content in both the occupied and empty bands at a particular $\mathbf{k}$. At the same time, a pair of bands at a particular $\mathbf{k}$ can only have contributed to a peak in the partially decomposed spectrum if their energy difference equals the energy location of the peak. The Cu z^2 anti-bonding to $4s$ class of channels has two features at ≈ 2.9 eV and 3.2 eV, as visible in the blue curve in (a). These features can be identified with the interband transitions marked with the cyan and magenta arrows in (b) by identifying the orbital content of the bands. Likewise, the Cu z^2 anti-bonding to Cu $z^2/\text{O } p$ hybridized class of channels has two features at ≈ 3.2 eV and 4.1 eV which can be identified with interband transitions marked by the magenta and green arrows in (c). One interesting thing to note is that interband transitions from the z^2 anti-bonding to $4s$ class and transitions from the z^2 anti-bonding to z^2/p hybridized class come from completely different regions of the Brillouin zone, though for the feature at ≈ 3.2 eV they happen to contribute to the same peak.

2.4 Implications for other complex transition metal materials

The analysis that I have been able to perform on the orbital nature of the neutral excitations of Cu_2O is not particular to this compound and can be extended to other complex transition-metal compounds. As I will demonstrate in the subsequent chapters, the accurate inclusion of local chemistry and symmetry, will have important implications for the study of the impact of orbital degrees of freedom in the iron-based superconductors and other complex, correlated compounds. In addition, the ability to calculate experimentally observable quantities from first-principles allows me to draw a direct, unbiased correspondence with experiment. Though, the methods I have used here for Cu_2O use a formulation of TDDFT in a local basis of Wannier orbitals and the challenge of applying DFT and TDDFT to correlated systems will face my local orbital methods as well, in Chapter 3, I have found that a combination of using LDA+ U in the ground state calculation can allow access to strongly correlated systems.

Thus, in this study of the impact local orbitals on the neutral excitations of Cu_2O utilizing the calculation of the TDDFT density response in a Wannier basis, I have found that this uncorrelated copper-oxide serves as a paradigm for analysis of the orbital dynamics of complex transition-metal compounds. The accurate representation of the bonding-anti-bonding chemical nature of the Cu z^2 orbital, via the wise choice of local symmetry axis and construction of Wannier orbitals, allowed me to assign the experimentally measured peaks in the optical spectrum as being intimately tied to the local chemistry and symmetry of the compound. This will serve to guide investigations into the dynamics of orbital degrees of freedom in correlated materials presented in Chapters 3 and 4.

Chapter 3

Local correlations and nanometer particle-holes in the graphene-like optical response of FeTe

In Chapter 2, I developed a method for finding an optimal set of Wannier orbitals by finding the orbitals that provided the description of the low energy optical spectrum that offered the optimal insight into the role of local chemistry and symmetry in the orbital particle-hole excitations. The primary goal in that endeavor was not to simply explain the optics of Cu_2O , but rather to test a method that could be extended to correlated compounds of greater current interest in the condensed matter physics community than the relatively simple semiconductor Cu_2O . In order to demonstrate the extensibility of the method to just such a broader class of compounds, I have chosen the iron-based superconducting parent compound FeTe.

However, as I will show in this chapter, FeTe differs from Cu_2O in that the optical spectrum is dramatically impacted by strong, onsite correlations. Whereas the optical response calculated using the GGA groundstate disagrees *qualitatively* with experiment, including the effects of onsite interactions beyond GGA via GGA+ U agrees *quantitatively* with experimental measurement of the optical conductivity in a

broad frequency range covering the mid-infrared. Examination of the band structure reveals that the effect of these correlations are *orbitally selective* in agreement with photoemission experiments. The renormalization of the band structure gives rise to a *graphene-like* optical response due to the combination of linearly dispersing energy differences between quasi-two-dimensional bands. Using response calculated in the Wannier basis, I will go beyond the momentum space interpretation of the optical spectrum and elucidate the real-space character of the excitations. Analysis of the orbitals shows the importance of *long-ranged hybridization tails* in producing the optical excitations. These long-ranged tails are also important to consider when determining the compactness of the orbitals, which is crucial for explaining the origin of orbital selectivity on the correlations. The non-locality present in FeTe raises questions about how optical properties are calculated from models and provides a clue as to how to improve the calculations that I present in this chapter.

The unexpected discovery of superconductivity in the iron arsenides has touched off an enormous effort in the investigation of related iron-based compounds [147]. The iron-chalcogenide compounds in the series $\text{FeSe}_{1-x}\text{Te}_x$ are structurally similar to the iron-pnictides, with a square lattice of Fe and tetrahedrally coordinated anions [148]. FeTe is on the end point of the this iron chalcogenide series, which can exhibit superconductivity at temperatures of up to 15K at ambient pressure or up to 37K under pressure [149, 150]. This makes FeTe one of the broad family of iron-based superconducting compounds, which includes the iron-pnictides. Thus, FeTe is just the kind of compound to which I would like to apply the method I developed using Cu_2O in Chapter 2. Investigating FeTe provides an inroad to studying the applicability of the optimization method I have developed to all of the iron-based superconducting compounds and complex transition-metal compounds in general.

However, using FeTe as a prototype for the iron-based superconducting compounds as a class is not necessarily straightforward, because FeTe has been found to differ starkly from the pnictides. While the pnictides generally display collinear spin-density wave order, with ordering vector along (100), FeTe displays a bicollinear ordering in a

zig-zag pattern [30]. While bands near the Fermi level as measured by photoemission are modestly renormalized in relation to DFT calculations in the pnictides [151], in FeTe the bands are strongly renormalized, indicating strong correlations [26]. *Ab initio* CRPA calculations of the screened interaction parameters among the d-orbital degrees of freedom showing larger onsite interactions compared to the pnictides also provide theoretical support for strong correlations in the chalcogenides [106]. Finally, and bearing directly on this work, the behavior of the infrared optical conductivity is dramatically different. In contrast to the pnictides [151], the mid-infrared response is a featureless plateau, with a value close to the universal conductivity in graphene [20, 21]. In addition, while the Drude weight decreases upon crossing the Néel temperature in the pnictides, it anomalously increases in FeTe.

As I will show in this chapter, strong correlations in FeTe have a bearing on the calculation on the optics, which makes using FeTe as a test system for optimizing the orbitals more challenging than was the case for Cu_2O . In that system, the TDDFT (RPA) calculation using the GGA groundstate already agreed well with experiment. Thus, I was able to find an optimal set of orbitals in direct comparison with experiment. Finding an adequate agreement between theory and experiment for FeTe requires handling strong, onsite correlations. So before finding an optimal set of orbitals and constructing low-energy models for FeTe I must first provide a description of the electronic structure and optical spectrum that is as complete as possible. This is what I will present in this chapter.

In addition to validating the use of FeTe as a test system for optimization by providing an explanation that is in terms of interband transitions, the formalism of response in a Wannier local orbital basis allows novel insight into the nature of the particle-hole pairs that give rise to the optical spectrum. Response in the Wannier basis allows for real-space resolution of the particle-hole transitions. particle-hole pairs separated by $\sim 1\text{-}2$ nm and below are shown to make up the majority mid-infrared response. The far infrared peak in $\epsilon_2(\omega)$ is dominated by particle holes ~ 50 nm in size, highlighting the presence of ultra-long-ranged physics. This

extreme range of the particle-hole transitions makes insights from momentum space enlightening.

I will begin my elucidation of the electronic structure of FeTe in Section 3.1 by providing an overview of the experimental and theoretical work in the literature. Though this section is useful in order to understand the broader context into which this work sits, it contains many details that bear only indirectly on the results I present and the reader may wish to skip the section on first reading. Then in Section 3.2, I will begin to present my results by showing that inclusion of the effects of strong correlations due to onsite interactions is essential to provide a good description of the mid-infrared optical spectrum. The electronic structure produced using the GGA+ U method predicts that the effect of correlations is orbitally selective with the xy orbital in particular being strongly effected and the xz/yz orbitals being weakly effected. In Section 3.3, I will explain the origin of the flat optical conductivity in FeTe in terms of a complex of approximately linearly dispersing, two-dimensional bands spanning the Fermi level. Remarkably, the topology of these bands provide a mechanism for the flat conductivity that is essentially the same as that for graphene that can explain the measured optical conductivity of graphite. In Section 3.4, I will use the calculation of the optical response in the Wannier basis to provide insight into the nature of the charge fluctuations that produce the optical spectrum, in particular their long-ranged nature. Finally, in Section 3.5, I will use all of the analysis I have performed to suggest improvements to correct deviations from experiment in the far-infrared.

3.1 Experimental and theoretical overview

This section contains a review of some of the experimental and theoretical work on FeTe and the related iron-based compounds. The picture of FeTe that emerges from the experimental evidence is that it is a strongly correlated, bad metal. In addition, the evidence from ARPES shows that these correlations are orbitally selective with the xy orbital being the most correlated. The effects of strong correlation in the electronic

structure produce an optical spectrum that is qualitatively distinct from the pnictides. The optical conductivity of the pnictides is characterized by a prominent Drude peak and a clear interband transition at between 0.6 and 0.9 eV. In contrast, FeTe does not have either and the spectrum is flat and featureless.

These strong correlations plausibly have a significant role to play in the unique magnetic and structural properties of FeTe, which produce a bicollinear magnetic order and monoclinic structural transition as opposed to the antiferromagnetic order and orthorhombic transition in the pnictides. Theoretical evidence from LDA+DMFT points to FeTe as well as the other iron-based compounds as Hund metals, a picture which can predict the trends in magnetic moment, correlation and orbital selectivity. Though I do not explicitly use the Hund metal concept in this work in this chapter and Chapter 4, I reach many of the same conclusions.

Though much of what I present in this section bears directly on the results that I present subsequently in this chapter and Chapter 4, some is only indirectly related and provides context. The novel results that I will present in this chapter are focused on the electronic structure of FeTe. So, the experimental data from optical spectroscopy and ARPES will be especially useful in evaluating my calculations and building a picture of the electronic structure and I will present these first. This information on the electronic structure will also be important to understanding the content of Chapter 4. After that, I will present the experimental data regarding the magnetic and structural phases. Though not strictly relevant for the current chapter, understanding the unique phases of FeTe in relation to the iron pnictides will be important in Chapter 4. Finally, I will present some relevant, complementary theoretical results from LDA+DMFT.

First, I will discuss the data from optical spectroscopy. Here FeTe is unique, not exhibiting an obvious peak in the mid-infrared optical conductivity as is characteristic in the pnictides [151]. Instead, FeTe exhibits a flat, featureless optical conductivity up to at least 7000 cm^{-1} [20, 21]. Unexpectedly, the value of this plateau is close to that in graphene. This is seemingly the only property linking the two materials,

though I will present an underlying similarity in Section 3.3. In addition, though the pnictides show an obvious Drude tail in the far-infrared, the Drude peak in FeTe appears to be very small and some recent experimental investigations have concluded that there is no Drude response [22]. This is consistent with a view of FeTe as a “bad metal” [152–155]. These stark contrasts with the pnictides raises the question of why FeTe is so different and this requires an explanation of the optical conductivity.

Such an explanation has been challenging, since the featurelessness of the spectrum has led to contradictory empirical models for the interband contribution to the optical conductivity, with no consensus on strength or even the existence of a mid-infrared interband transitions. I argue, both on general grounds and based on the results I present in the subsequent sections, that this is largely due to the manner in which the spectrum is modeled. The dominant empirical model is the Drude-Lorentz model [156], which uses Lorentzians centered at zero (the Drude contribution) or at finite frequencies. These latter Lorentzians are used to model interband transitions and are constructed in such a way that they produce the correct symmetry on the frequency axis. That is they are guaranteed to go to zero at zero frequency.

The only physical reasoning behind this model is that a Lorentzian line shape is one that would be produced by an excitation with an energy $\hbar\omega_0$ and lifetime $1/\gamma$. Of course, in a solid state system interband transitions are not associated with a single frequency, but a range of frequencies because the bands involved in the transitions disperse across the first Brillouin zone. For transitions between flat bands, the Lorentzian may give a good fit. But for interband lineshapes that deviate significantly from Lorentzian the Drude-Lorentz model cannot be expected to give sensible results.

That is not to say that a “good fit” cannot be found using the Drude-Lorentz model. Given enough Lorentzians, any spectrum can be fit and the smoother the spectrum the easier that fitting will be. For a flat featureless spectrum like the one for FeTe, one could expect that a fit to the data could be found. Of course, that is precisely what has been done, but obtaining a fit requires the introduction of elements

into the model that have no sound physical basis. Instead, the model is rationalized *post hoc* and the result is an unphysical view of the optical excitations.

Both empirical models for the optical conductivity contain a very broad “Drude” contribution [20, 21] in addition to a narrow one that resembles a more conventional Drude peak. The only way for there to be such dramatically different Drude contributions there would have to be two types of electrons at the Fermi level. Conveniently, there is an electron pocket and a hole pocket that can be used to classify the electrons. But one must explain what could be coupling to the electronic degrees of freedom that would cause such a large width in energy. Here I find a number of problems with the introduction of this second Drude response.

One is based on firm theoretical grounds. The intraband response represented by the real part of the optical conductivity for a purely periodic electronic system should properly be a delta function [157]. The observed broadening of the delta function into a sharp peak with a tail stretching to finite frequencies away from zero must be due to scattering off of photons and impurities as well as the finite resolution of spectroscopic experiments. In other words, any intrinsic broadening of the Drude peak in the optical conductivity must be due to coupling to degrees of freedom *external* to the electronic subsystem. Though it is possible that this scattering rate can vary by band and across the Fermi surface, the second Drude peak in FeTe has a scattering rate with an associated energy in the hundreds of meV range.

It seems unlikely that magnetic fluctuations or phonons could be responsible and even if they were, why would they effect one pocket so dramatically differently relative to the other? In my estimation, the optical data alone is insufficient evidence to support such a claim and I must conclude that the second Drude peak is introduced for the sole purpose of providing an empirical fit to the observed data and does not indicate any new type of intraband physics. As such, I believe an alternative explanation is required, which I will present later.

One issue that I will not treat theoretically, but bares mentioning as an indication of the unconventional nature of FeTe is the behavior of the Drude weight as

temperature is decreased below the magnetic ordering transition. For the pnictides, for which a spin-density-wave picture for the magnetic order appears to be proper, the Drude weight decreases significantly as temperature moves below the Néel temperature [151]. This is due to the splitting of the bands due to symmetry breaking depleting the number of states near the Fermi level. The electrons at the Fermi level “fall” into the SDW state, decreasing the Drude weight. Thus, a conventional band picture can explain the behavior with the Drude weight when crossing the Néel temperature. However, FeTe displays the opposite behavior. It becomes a better metal in the magnetically ordered state when compared to the paramagnetic state [152]. This connects the uniqueness of the magnetism in FeTe with the odd behavior of the Drude weight and special nature specific to FeTe becomes apparent.

The electronic structure of FeTe as measured by angle-resolved photoemission spectroscopy (ARPES) is also in many ways unique among the iron-based superconducting parent compounds. In the paramagnetic phase the features of the measured spectral function are broad with no quasiparticle peaks crossing the Fermi level [22–26]. Though there is spectral weight at the Fermi level, this is attributed to the existence of a pseudogap like that found in the cuprates rather than a band crossing as in a conventional metal, confirming the designation of FeTe as a bad metal in the paramagnetic state [22, 24]. A pseudogap picture of the electrons at the Fermi level is subtly different from the picture of electrons exhibiting a Drude response with a large scattering rate. Any Drude response would be due to a quasiparticle band crossing the Fermi level. For a pseudogap, though there are electron states at the Fermi level, they are due to broadening of bands away from the Fermi level and there are no quasiparticle crossings.

In the magnetic state, the electronic structure at the Fermi level changes qualitatively. Small but visible quasiparticle peaks appear and bands are clearly crossing the Fermi level [22–24, 158]. This supports the observations of the increase of the Drude peak in the optics and the increase of the resistivity below the transition

temperature. Spectral weight is concentrated around the Γ and M points [22–26, 158] in agreement with my calculations, with weight observed around the X point in contradiction to theory [22] appearing unconfirmed. The peculiar appearance of quasiparticle peaks and metallic behavior below the transition raises intriguing questions on the interplay between electronic structure and optics. In contrast to the pnictides [151], no nesting in the Fermi surface is observed and no gap indicative of a spin-density wave [24, 158]. Thus, a weak coupling mechanism for the magnetic ordering transition is ruled out. Instead, strong coupling to spin fluctuations has been proposed to explain several aspects of the electronic structure.

Zhang et al. [22] have proposed that in the paramagnetic phase coupling to spin fluctuations leads to localization of otherwise itinerant charge carriers and that the suppression of spin fluctuations upon the opening of a spin gap [152, 159] allows these carriers to move more freely. The proposition of strong coupling to a bosonic mode such as spin fluctuations can also account for the broad spectra, low quasiparticle weight and existence of a pseudogap [24, 25]. Though Liu et al. [23] have proposed an explanation of the “peak-dip-hump” line shape in the antiferromagnetic magnetic phase as characteristic of strong coupling to phonons and the small quasiparticle peak as being due to coherent polaron motion in analogy to the manganites [160, 161]. Ieki et al. [25] have suggested that strong local correlations could also be pushing the system to a Mott phase, opening a pseudogap and broadening the spectral lines.

Though there is agreement about spectral weight near the Fermi level being concentrated in pockets around the Γ and M points, there is disagreement about the number of bands or features participating in the pockets in the paramagnetic phase. This is likely due to the broadness of the line shapes causing ambiguity in the data [25]. Zhang et al. [22] assign two features at Γ and two at M, while Liu et al. [23] and Lin et al. [24] two at Γ and one at M. Ambolode et al. [26] focused on the Γ point and assigned three bands after manipulating the data by taking second derivatives with respect to energy and then momentum. In contrast, Ieki et al. [25] state an inability to track any band dispersions for $\text{Fe}_{1.03}\text{Te}$. This disagreement and ambiguity

means that so far no concrete assignment of bands can be performed. Perhaps it is incorrect to even think of quasiparticle-like bands at all since the incoherence of the spectra in the paramagnetic phase makes the picture of bands being renormalized by interactions close to meaningless [24].

Another peculiar aspect of the iron-chalcogenides that is of particular interest in the context of this dissertation is the apparent orbital selectivity of the electron correlations, highlighting the vital multi-orbital nature of these compounds. Tamai et al. [162] first observed hints of this behavior in the need for several mass renormalization factors for the various Fermi surface sheets in $\text{Fe}_{1.1}\text{Se}_{0.42}\text{Te}_{0.58}$. Yi et al. [163] later identified orbital dependent renormalization in doped iron-selenide. The same orbital selective behavior observed for other members of the iron-chalcogenide series is also seen in the endpoint FeTe, both in the level of coherence [24] and mass renormalization [26]. The consensus among all of these studies is that the xy orbital (with axes aligned with the Fe square lattice) is the most strongly correlated. Considering the sensitive dependence of correlation with chalcogen height [26] and the dependence of the shape of the orbitals on this height, the iron-chalcogenides display a subtle interplay between electronic structure, local crystal environment and chemistry and correlation.

The magnetic properties of the iron pnictides is qualitatively different from those of the chalcogenides and FeTe in particular [164, 165]. All of the pnictide parent compounds (excluding LiFeAs which does not order) order in the Fe plane in the C -type structure, that is ferromagnetic along one direction of the square lattice and antiferromagnetic along the other. This ordering pattern is well explained for all of the pnictides by a magnetic instability in the nonmagnetic phase due to nesting of the hole and electron Fermi surfaces, though the predicted value of the moments are significantly larger than the measured values. Along with the magnetic order there is an orthorhombic structural distortion with the moments and antiferromagnetic ordering along the long axis.

In contrast to the pnictides, the Fe_{1+x}Te family of compounds display a number of unusual magnetic and structural properties. The phase diagram is very sensitive to the amount of excess Fe (i.e. the value of x [148, 152, 166].) Since purely stoichiometric FeTe has thus far not been produced due to the difficulty in growing such crystals, the most relevant experimental data for my purposes are those for the smallest available x [155]. All samples with $x < 10\%$ have similar enough properties to consider them analogs and provide the best available approximation to stoichiometric FeTe [152]. Fe_{1+x}Te for small x is alternatively described as a “semiconductor”, “semimetal” or “bad metal” above 70K due to the weak temperature dependence of the resistivity and its negative slope with increasing T [152–155]. This anomalous dependence of the resistivity has been rationalized as arising from scattering from magnetic fluctuations using an empirical model for the resistivity [152].

Fe_{1+x}Te for small x is a paramagnet above 70K with large local moments ($\approx 4\mu_B$) [148] and broad, gapless magnetic fluctuations at an incommensurate wave vector near $(0.5, 0.0, 0.5)$ [152]. For excess Fe at concentrations $x > 10\%$, there is an incommensurate spiral magnetic order associated with these fluctuations [30, 152, 166]. But for low excess Fe, there is an abrupt shift of the magnetic scattering peak from incommensurate to commensurate position upon moving below a first-order phase transition at 70K [159]. For the smallest values of x , this magnetic transition coincides with a monoclinic structural distortion, which is described as “locking-in” the commensurate, bicollinear ordering [159]. Since the structural distortion and magnetic ordering are locked together, occurring at the same temperature, [152] has posited a strong magneto-elastic coupling.

There has been some success in theoretical explanation of many of the features of electronic structure and magnetic properties of the pnictides and chalcogenides. Optical conductivity calculated in TDDFT using similar methods as in this dissertation have found decent agreement with experiment in e.g. LaFePO [167], though the calculations do not reveal the lowest energy feature. This is a common deficiency of DFT based calculations of the optical conductivity [168]. The magnetic structure of

the pnictides has also been extensively studied in DFT with prediction of a SDW order associated with nesting of the Fermi surface [169–171], though Lee, Yin and Ku [172] have disputed the nesting picture in favor of one based on the local chemistry of the d orbitals as revealed by Wannier functions.

However, the analysis based on Kohn-Sham DFT band structure that reveals a magnetic instability due to nesting incorrectly predicts the magnetic order of the chalcogenides [173] and over estimates the moments of the pnictides by a factor of ~ 2 [169, 171, 174]. Using a combination of DFT and dynamical mean-field theory, Yin, Haule and Kotliar have succeeded reproducing the behavior of the optical conductivity across the magnetic ordering temperature, overcoming the deficiency of the DFT-based result in the magnetic state [175]. In addition, the same authors have been able to develop a unified picture of magnetism in the iron pnictides and chalcogenides [176]. In that work they found that iron chalcogenides are significantly more correlated than the pnictides with greater orbital differentiation in the degree of correlation (with the xy orbital being the most strongly correlated), with FeTe being the most extreme, and the degree of correlation correlates with the size of the magnetic moment. This is consistent with all of the experimental studies I have reviewed in this section.

The theoretical concept that can explain this trend across compounds is that of the Hund metal [31, 32]. In a Hund metal the strong onsite exchange interaction (Hund’s coupling) tends to favor a high spin state, as in the free atom. When the electron configuration per site is one away from half filling as in the d^6 configuration of the iron pnictides and chalcogenides, increasing Hund’s coupling can both decrease the coherence of quasiparticles and increase the critical direct onsite interaction (Hubbard U) at which the transition to a Mott insulator occurs [177]. This is due to an orbital blocking mechanism associated with the high spin state that suppresses inter-orbital hopping and restricts the freedom of the electrons, slowing them down and enhancing the effective mass [176].

Lanatà et al. [178] were able to provide further theoretical evidence for this mechanism applied specifically to the iron chalcogenides. Yin, Haule, and Kotliar

also investigated the optical conductivity of FeTe using DFT+DMFT, finding good agreement with experiment [179]. However, the success of this investigation must be questioned because the experimental data to which the authors were comparing [180] does not show the flat optical conductivity observed in other measurements [20, 21] and the theory shows a qualitative deviation from this flat spectrum. I take the latter measurements as authoritative in this dissertation and I consider the flat conductivity to be reality.

The large body of experimental and theoretical work on FeTe has converged on the compound being strongly correlated with the effects of onsite interactions in particular being of vital importance. However, a comprehensive, *ab initio* study of the experimentally measured, graphene-like optical spectrum is lacking. In the next section, I will perform just such an investigation using the GGA and GGA+ U groundstates, a single-site projected Wannier basis and the optical response. This will confirm many of the aspects of the electronic structure that have been deduced so far and will provide novel insight into the graphene-like nature of the optical response.

3.2 Vital importance of local correlations

In order to clarify the physical interpretation of the optical spectrum, I have performed *ab initio* calculations in the framework of TDDFT and TDDFT based mean-field response using GGA+ U for the groundstate. As I will show in this section, the GGA groundstate gives a qualitatively incorrect description of the optical spectrum and so the electronic structure must be improved. In particular, a dramatic, orbitally selective role of local correlations on the ground-state electronic structure leads to renormalization of the mid-infrared optical conductivity by a factor of two and brings the calculation into quantitative agreement with experiment.

Figure 3.1 displays the effects of local correlations on the optical spectrum. Calculation of the GGA dielectric function and optical conductivity shows a peak at ≈ 0.7 eV in the mid-infrared, as shown by the grey curve in Figs. 3.1a and 3.1b,

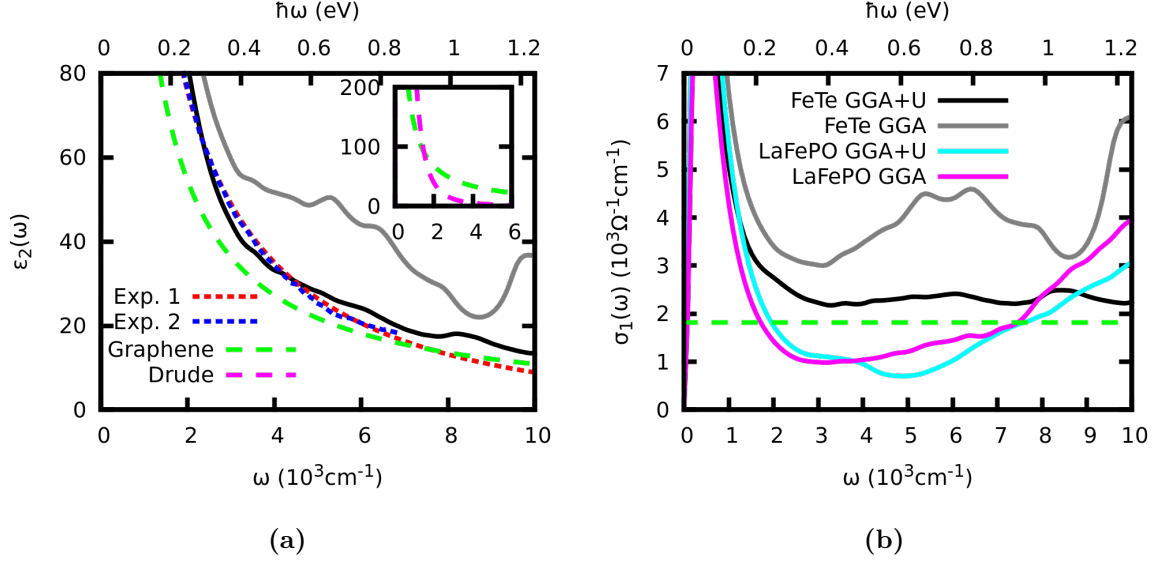


Figure 3.1: The role of onsite interactions in reproducing the measured optical spectrum of FeTe. Inclusion of correlations due to onsite interactions beyond GGA in the form of GGA+ U is critical to achieving agreement with experiment. The calculated $\epsilon_2(\omega)$ along with the measured values is shown in (a). The GGA result shown in grey is greater by a factor of 2 over a large portion of the spectrum versus experiment. The GGA+ U result shown in black achieves quantitative agreement with experiment. In particular, it reproduces the $1/\omega$ behavior of both the experiment for FeTe and the inferred behavior of Graphene, shown in green dashes. This is qualitatively different from behavior of the Drude tail, shown in magenta in the inset. Thus, the $1/\omega$ behavior is a feature of interband rather than intraband physics. The optical conductivity is shown in (b). The GGA shows a clear peak centered at $\approx 6,000 \text{ cm}^{-1}$ whereas the GGA+ U calculation is flat down to $\approx 2,500 \text{ cm}^{-1}$. I propose that the deviation from the flat behavior is due to inadequate treatment of correlations of bands at the Γ point. The impact of including onsite correlations beyond GGA is qualitatively different for LaFePO. The GGA spectrum in magenta and the GGA+ U spectrum in cyan are quantitatively, but not qualitatively different. I believe this is a general feature of the iron pnictides.

respectively. At first glance, this peak seems to correspond to the mid-infrared peaks observed across the pnictides at locations ranging from ≈ 0.6 to 0.9 eV [151]. These peaks are generally found in DFT calculation, though shifted higher in energy [151]. This shift is attributed to many-body renormalization of the band structure beyond those included in GGA. However, experimental measurements of the optical conductivity of FeTe are qualitatively different, not showing any mid-infrared peak in the optical conductivity [20, 21]. Instead, the response is virtually constant in that region. In this case, the GGA contribution not only displays a non-existent peak, but the overall intensity is larger by a factor of two. Though in GGA calculations also give a higher intensity in e.g. LaFePO, the error is significantly smaller [167].

Clearly, there is a need to dramatically improve the calculated optics. Since the spectral weight in the mid-infrared is too high, it is reasonable to think that there are too many states in the vicinity of the Fermi level that can participate in interband transitions and some bands must be removed from the Fermi level. This can be accomplished through GGA+ U , which will tend to increase the binding energy of bands due to occupied orbitals and move bands due to unoccupied orbitals higher in energy. Of course, GGA+ U will only have an effect if sufficient correlations beyond those captured in GGA are present.

Though I have presented considerable experimental evidence that strong correlations are present in FeTe in Section 3.1 and this would justify the use of GGA+ U , there is sufficient evidence from *ab initio* theory to indicate that the GGA+ U should have a significant effect. In particular, narrow bandwidth of the xy orbital (see Figs. 3.2a and 4.1) and the larger value of intra orbital U for the xy orbital from CRPA calculations (see Table 4.3 and Ref.[106]). Since the ratio of U to the bandwidth is greater than one, one can expect that the xy orbital in particular will be affected by the inclusion of local correlations using GGA+ U . Since this orbital dominates at the Fermi level, its removal from the Fermi level will tend to reduce the spectral weight of the mid-infrared.

Thus, I have performed a GGA+ U calculation of the groundstate and calculated the optical response. I used $U = 3.0$ eV and $J = 0.6$ eV. These values are in line with the average of those values present in the $U_{\nu\nu'}$ and $J_{\nu\nu'}$ matrices (See Table 4.3). The imaginary part of the dielectric function is shown in Fig. 3.1. The experimental measurements were performed by Hancock et al. [20] “using a combination of reflectivity and ellipsometry techniques” and Dai et al. [21] via reflectivity. I have extracted the imaginary part of the dielectric function from the former from the displayed optical conductivity and from the latter from the model fit of the data that the authors present. In Fig. 3.1 both are represented by dotted lines and my calculation using GGA+ U is represented by the black line. As is apparent, the calculation closely matches Experiment 2 [20] from $\approx 2,000$ cm^{-1} to $7,000$ cm^{-1} where the experimental data stops. This agreement also holds for Experiment 1 [21] and beyond to the edge of the figure at $10,000$ cm^{-1} . However, as the experiment used a model fit intended to conform to lower frequencies, it is perhaps unwise to trust the details of the red dotted curve between $7,000$ cm^{-1} and $10,000$ cm^{-1} . Clearly, the effect of including correlation effects of the onsite interaction beyond that included in GGA is essential. Whereas the GGA fails to agree even *qualitatively* with experiment (grey line in Fig. 3.1), the GGA+ U agrees *quantitatively*.

The dramatic effect of treating onsite interactions within the GGA+ U method indicates that there are strong onsite correlations in FeTe. The localized d orbital degrees of freedom are insufficiently screened by the rest of the degrees of freedom. In particular, GGA+ U treats both the self-interaction correction arising from the Pauli exclusion principle and non-local exchange whereas GGA can only approximately include these effects. The strongly correlated nature of FeTe indicated by my calculations is corroborated by the experimental evidence reviewed in Section 3.1. Though the GGA+ U calculation is evidence for strong onsite correlations, it does not exclude the possibility of additional off-site correlations. In addition, there are some hints that the current calculation could be improved by inclusion of correlations from interactions between the d orbitals of different sites, as I will present later.

The effect of inclusion of onsite correlations in GGA+ U for FeTe is in sharp contrast to the effect for the iron-pnictides. I will use LaFePO as an example. I calculated the GGA+ U groundstate for LaFePO using $U = 2.7$ eV to reflect the smaller value of onsite interactions as calculated by CRPA [106]. As shown in Fig. 3.1b, the difference between GGA and GGA+ U on the calculated optical spectrum of LaFePO is quantitative rather than qualitative. There are some minor differences between the GGA and GGA+ U , but the overall line shape of the spectrum is unchanged. This is in contrast to FeTe where an entire optical peak disappeared upon inclusion of onsite correlations.

Remarkably, the result of choosing a U in line with CRPA calculations is a *quantitative* agreement with experiment for FeTe and has a minimal effect on LaFePO as would be expected since GGA produces an optical conductivity much closer to experiment [167]. This is compelling evidence that CRPA provides a good guideline for the selection of interaction parameters. This raises an interesting question. The CRPA parameters do not stop with the onsite interactions. There are additional offsite interactions present that have not been treated here beyond their implicit treatment in GGA. If CRPA provides a good guide for the onsite interactions, why not the offsite ones? I will not provide an answer here, but I believe that these interactions must at least be investigated before they are dismissed.

The effect of U on the optical spectrum alone does not give any information on the orbital selectivity of the effect of correlations. In order to investigate the effect on the orbital degrees of freedom, it is vital to examine the orbitally resolved band structure. In agreement with the expectation from the optics calculation as to the difference of the effect of U between FeTe and LaFePO, the effect of performing a GGA+ U calculation on the band structure is qualitatively different between the two compounds, as shown in Fig. 3.2. On one hand, the xy and z^2 orbital in FeTe are shifted downwards in energy by more than an eV upon application of the U , completely removing the former from the Fermi level. On the other hand, the effect of GGA+ U on LaFePO is not characterized by dramatic, orbitally dependent shifts,

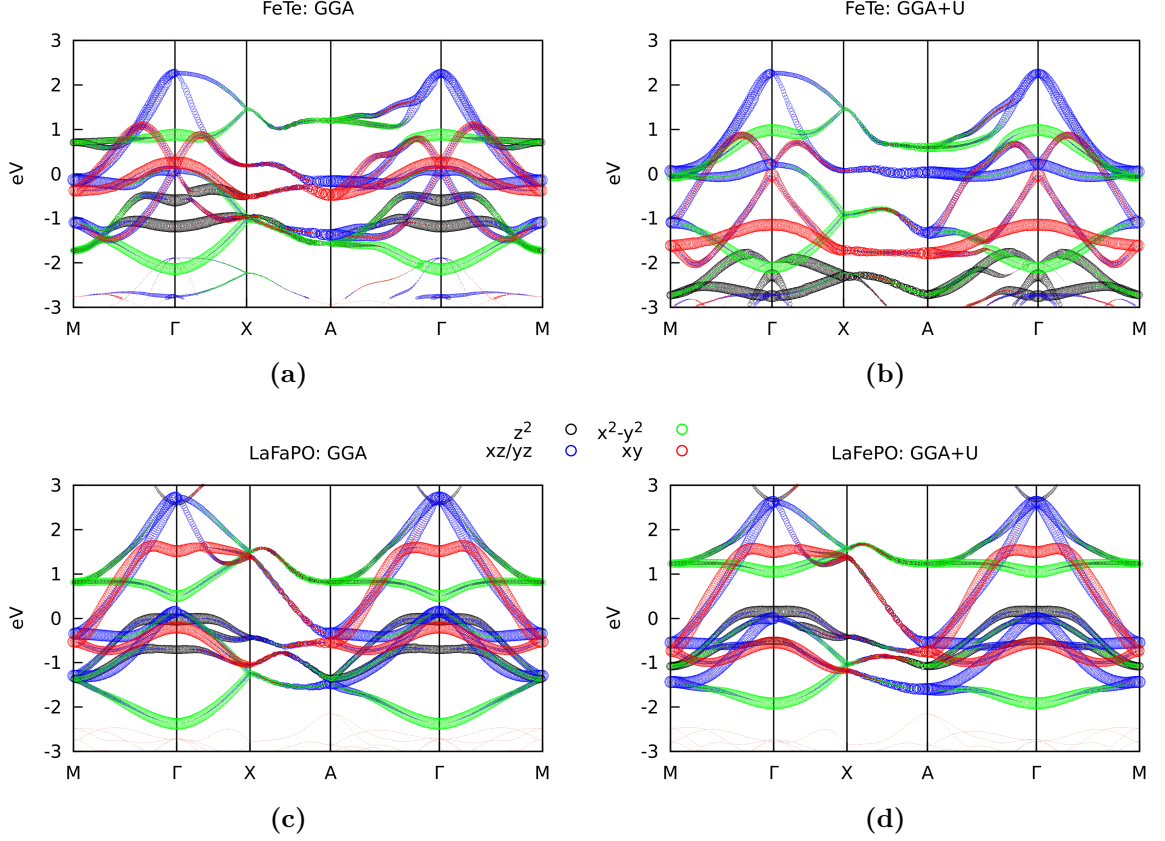


Figure 3.2: GGA and GGA+ U band structures for FeTe and LaFePO. In order to compare the effect of inclusion of correlations due to onsite interactions via GGA+ U on FeTe versus the pnictides, I have taken LaFePO as a representative member. For FeTe I have used a value of 3 eV for U and for LaFePO a value of 2.7 eV to reflect the lower value of the onsite interactions of the latter in CRPA calculations. However, even with a value of U of 3 eV the picture does not change qualitatively. In FeTe, the difference between the GGA band structure in (a) and GGA+ U band structure in (b) is a dramatic shift downward of the xy and z^2 character bands by ≈ 1.5 eV in the latter. Meanwhile the x^2-y^2 and xz/yz character bands are largely unchanged. This is a manifestation of orbital selectivity in the degree of correlation of the various orbital degrees of freedom. In LaFePO, the effect of the additional correlations due to U is qualitatively different from FeTe. There are no dramatic, orbitally dependent shifts. Rather, the occupied bandwidth narrows slightly, by ≈ 0.5 eV.

but an overall narrowing of the band width. This accounts for the differing effects on the optical spectrum of GGA+ U : removal of an orbital from the Fermi level removes a large number of transitions which had produced a peak in the GGA calculation. With no such removal taking place in the case of LaFePO, the effect on the optical spectrum is far less dramatic.

However, I must point out that the GGA+ U method assumes that all of the intra-orbital interactions are the same, which is not indicated by the CRPA calculations, which allows for more general types of interactions. In fact, it is precisely the difference in the values of the intra-orbital interactions which indicates orbital selectivity. So how is it that only the xy and z^2 orbitals are strongly effected? This has to do with the degree of localization of the orbitals, in particular the localization inside of the muffin-tin. I have calculated the integral of the density of the Wannier orbitals inside of the muffin-tin as a percentage of the total density, which is one. I found that the value of this integral for the xy orbital is 74%, for the z^2 orbital it is 79%, for the x^2-y^2 it is 66%, and for the xz and yz orbitals it is 69%.

Clearly the orbitals can be classified according to the percent of the orbital inside the muffin-tin and this coincides exactly with the classification according to the effect of the U . The xy and z^2 orbitals are the most localized and most strongly effected; the xz , yz and x^2-y^2 are less localized and least strongly effected. The effected orbitals are pulled to lower energies, away from the Fermi level because they are mostly occupied. The effect on the optical conductivity is in line with expectation: the overall intensity is reduced due to states being removed from the vicinity of the Fermi level (compare Figs. 3.2a and 3.2b).

In addition to the remarkable agreement of the GGA+ U calculation with the measured optical conductivity, I find that the flat, occupied band starting at M and extending toward Γ that has been observed in photoemission [23, 24] is also present in the Kohn-Sham band structure shown in Fig. 3.2b (cf. also Fig. 3.9a). Though it is common to compare GGA or GGA+ U band structure with ARPES data, a few caveats must be made. The band structure for, say, GGA only has strict theoretical

meaning within TDDFT linear response with reference to particle-hole excitations accessible via the density response function. There is no strong, theoretical way to draw a connection between Kohn-Sham bands and photoemission, a one-particle probe. Though any comparison is empirical and not strictly *ab initio*, it is still suggestive that I am able to produce a subtle, low-energy feature in the ARPES once I allow for onsite correlations.

However, I also found that the presence of this flat band was sensitive to the value of U to a much greater degree than the optical conductivity, as summarized in Fig. 3.3. In Fig. 3.3a, I have shown the optical conductivity calculated with GGA+ U groundstates with several values of U . For values between 2.7 and 3.3 eV the results compare well with experiments, especially compared to the GGA. But the band structure for $U=3.3$ eV has an occupied band along M to Γ that is not flat as it is for $U=3.0$ eV, as shown in Fig. 3.3c. Thus, the band structure is considerably more sensitive to the values of U , but since the optical response involves an integral over the Brillouin zone it is less sensitive. Some of the discrepancies are averaged out. I also note that I find a reasonable estimate of the CRPA values for interactions to be within 10%. Thus, I do not believe the CRPA alone could distinguish between $U=3.0$ eV and $U=3.3$ eV. Instead, I have chosen $U=3.0$ eV so that the flat band from M to Γ agrees with ARPES data.

Thus, it is clear from the results that I have presented in this section that the local correlations due to onsite interaction have a dramatic, orbitally selective effect on the electronic structure. Furthermore, these correlations beyond GGA must be included to achieve agreement with experiment. The resulting electronic structure and the particularly large effect of correlation on the xy orbital are confirmed by the experimental and theoretical evidence I reviewed in Section 3.1. However, I have not so far provided any explanation for the flat line shape of the conductivity. This deserves further analysis because of its resemblance to graphene, a seemingly unrelated compound. In the next section I will provide evidence that this correspondence goes far deeper than mere coincidence.

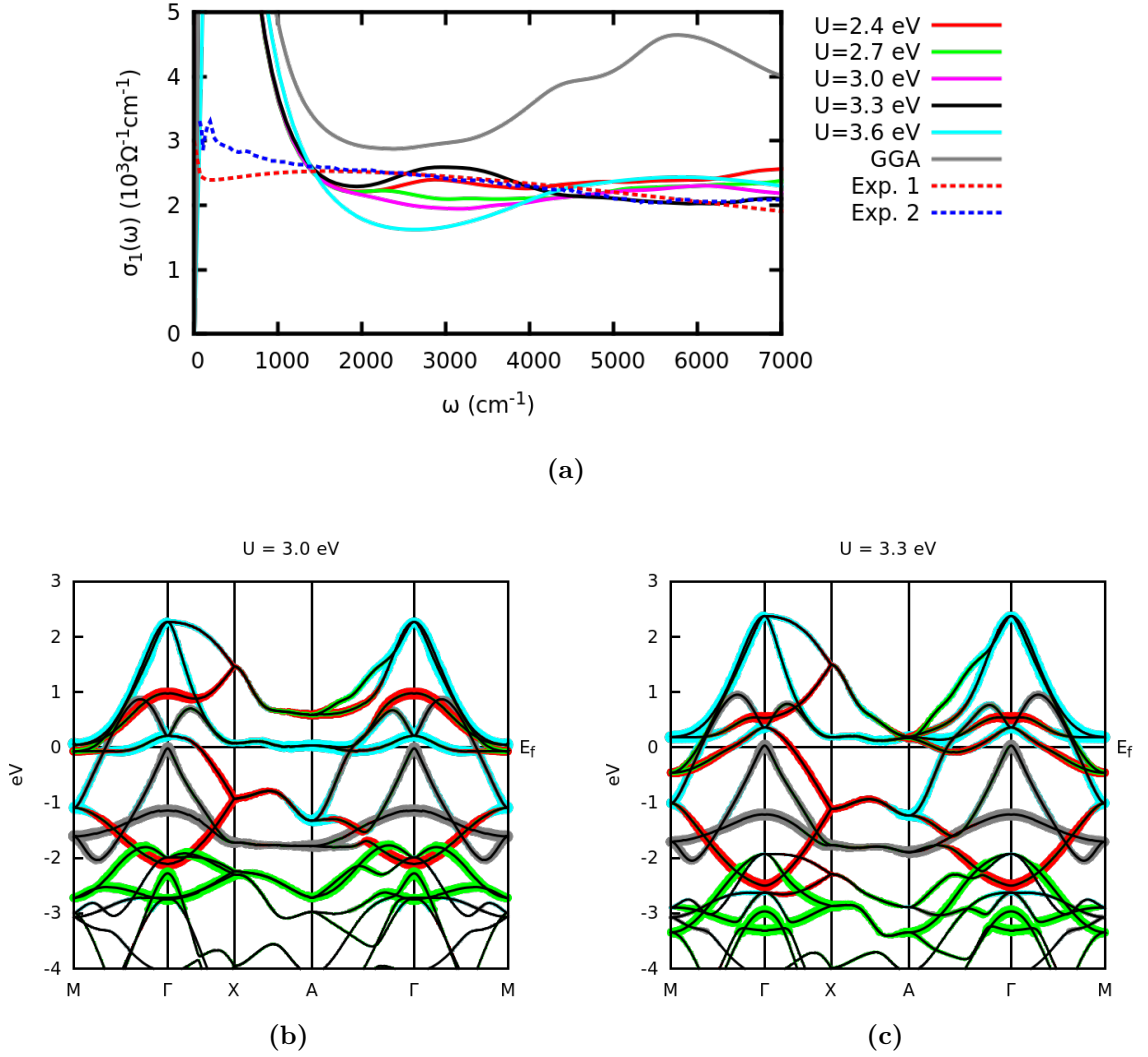


Figure 3.3: Dependence of the electronic structure on U . Though I chose $U=3.0$ eV for my GGA+ U calculations of the optical response based on CRPA calculations, based on the optical conductivity I could have chosen a range of values. Shown in (a) is the optical conductivity calculated from various GGA+ U groundstates with differing values of U . Clearly, a roughly equally good fit to the experimental data could be achieved for U values from 2.7 eV to 3.3 eV. The band structure, however is more sensitive. In (b) is the band structure for $U=3.0$ eV, which in particular displays the flat band from M to Γ that is observed in photoemission [23, 24]. When U is raised to 3.3 eV, the band structure for which is shown in (c), this band is no longer flat. Also, there is some band reordering around Γ and A. In light of the experimental data from photoemission, the choice of $U=3.0$ eV is superior.

3.3 Mechanism for the graphene-like conductivity

One remarkable feature of the spectrum shown in Fig. 3.1 is its resemblance both qualitatively and quantitatively with the measured spectrum of graphite and the proposed spectrum of graphene, which can be deduced via the Slonczewski-Weiss-McClure model [27–29]. This raises the intriguing question of whether this is a coincidence or if there is an underlying connection between the electronic structure of the two otherwise very different materials. There are two features that largely define the low-energy electronic structure of graphene: the “Dirac cone” at the K and K’ points that produces a linear dispersion of the occupied and empty bands, and the two-dimensionality of the this cone [181]. Within the RPA, these two features conspire to produce a JDOS that is proportional to the frequency. In the approximation that the matrix elements of the momentum operator in the Bloch basis are held constant, this JDOS produces an $\epsilon_2(\omega)$ that is inversely proportional to frequency, as is the case in FeTe.

The first thing to note is that the JDOS in the RPA only depends on the energy differences between the occupied and empty single-particle energy eigenvalues. It is not essential that the bands form a Dirac cone, but that there are linearly dispersing energy differences between the bands. The second thing to note is that the requirement of two-dimensionality is placed on the *bands*, not on the material as a whole. Though graphene is explicitly two-dimensional, the important feature of the band structure is that the bands are non-dispersive in one of the three spatial directions. Thus, even if a material is three-dimensional, if the relevant bands disperse only two dimensionally then the dimensionality requirement is fulfilled.

FeTe’s band structure shares the relevant topology with graphene in that there are two-dimensional bands with linearly dispersing energy differences. This is clear in Fig. 3.4. One member of the band complex that gives rise to the flat mid-infrared conductivity is the flat band running from M to about 2/3 of the way to Γ . This is the same flat band that is observed in ARPES [23, 24], which provides evidence of

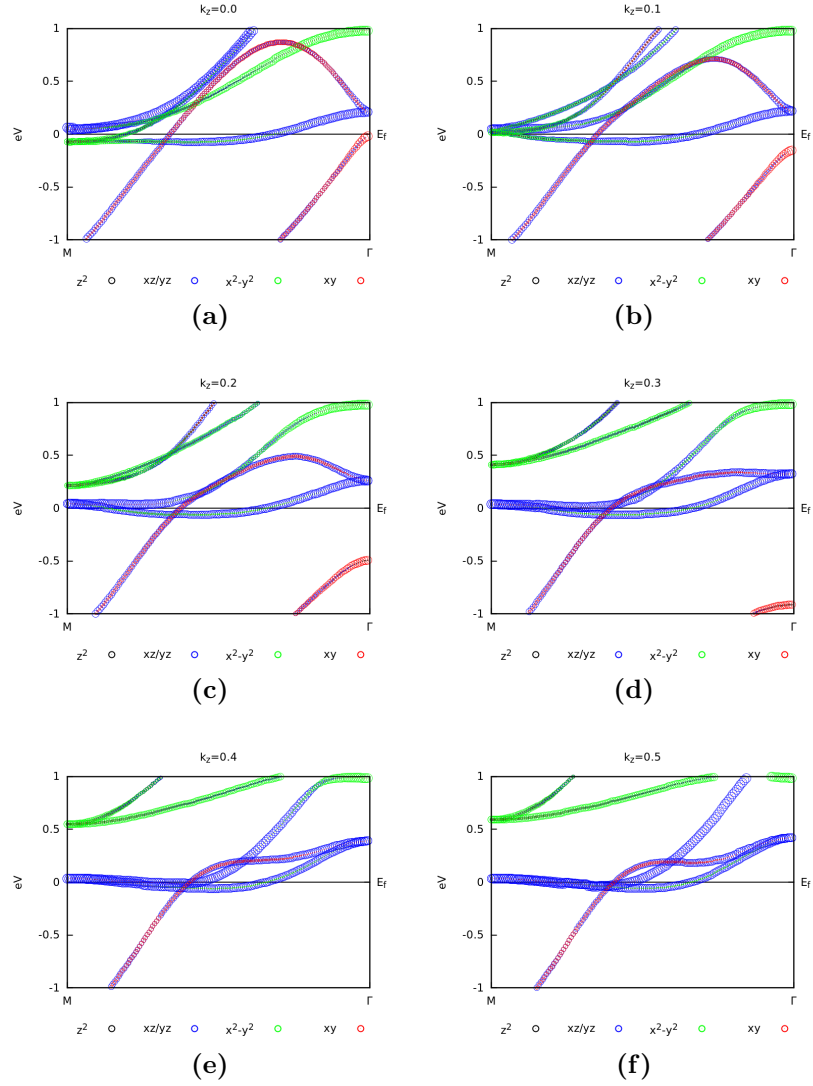


Figure 3.4: Two-dimensionality in the band structure for FeTe. Though FeTe is on the whole a three-dimensional material, there are some bands that are roughly two-dimensional. In (a) is the low-energy band structure from M to Γ . There are several linearly dispersing bands which are critical for the graphene-like optical response of FeTe. But in addition to linear dispersion, two-dimensionality is required. To gain information on the dimensionality, I have plotted the band structure for $\mathbf{k}$ points above the M to Γ line for various heights given by k_z , shown in (b) to (f). The flat band at the Fermi level, a linearly dispersing occupied band running from M to $\approx 40\%$ of the way to Γ where it crosses the Fermi level, and a linearly dispersing unoccupied band starting $\approx 50\%$ of the way to Γ at the Fermi level make up a two-dimensional band complex. This is discernible from their lack of dispersion along k_z . Since these are two-dimensional bands with linearly dispersing *energy differences* they can produce a graphene-like response.

the real-world existence of this feature in the band structure. The other members of the two-dimensional band complex are a pair of roughly linearly dispersing bands. One is linear and almost perfectly two-dimensional while it remains occupied. This band can be seen as a combination of xy and xz/yz character.

This band also begins as linearly dispersing when it becomes unoccupied, but disperses with k_z . However, the other linearly dispersing band takes over, as is already visible in Fig. 3.4b. This second band starts as x^2-y^2 and changes character to xz/yz with increasing k_z . The transition between the two characters is visible in Figs. 3.4c and 3.4d. However, it remains largely linearly dispersing. Important also is that this band complex is the only one near the Fermi level for much of the way along k_z meaning there are no intruders that would ruin the resulting flat conductivity. Thus, FeTe shares essential band topology with graphene that can explain the resemblance of the two systems' optical response, at least in the mid-infrared.

Though the topology of the low energy band structure can explain the shape of the optical response of FeTe being similar to graphene, it is insufficient to determine the magnitude of that response. In graphene, the “universal” value of the conductivity being a constant determined from fundamental constants results algebraically from a cancellation of material dependent parameters (hoppings in the tight-binding model) arising in the JDOS and the optical matrix elements [29]. With regards to the JDOS these parameters arise from their determination of the slope of the linearly dispersing bands, the Fermi velocity. In the matrix elements these parameters arise from their contribution to the single electron wavefunctions that are integrated over. The empirical fact that the optical conductivity in FeTe is so close to the universal value may be evidence that there is a deeper connection between these two contributions than mere coincidence that is material independent. However, this remains an open question and I do not offer any answer here.

Thus, it is possible to find a mechanism for the graphene-like line shape of the optical spectrum from arguments based purely on the band structure of the

Bloch states. The Wannier orbital basis was not essential to the discussion given the GGA+ U band structure, which incorporates orbital selectivity, as I showed in Section 3.2. However, using the optical response calculated in a Wannier basis, I can offer greater insight into the real space and orbital nature of the mid-infrared spectrum. This will have an impact on the picture of charge fluctuations particular to FeTe and will reveal the role of long length-scales, in contrast to the local nature of the correlations that produce the groundstate.

3.4 Insight into the mid-infrared using Wannier response

Thus far, I have been able to discuss the optical response largely without having the assistance of decomposition into Wannier particle hole channels. However, the remarkable agreement in the mid-infrared invites further interpretation and insight into the optical response beyond that gleaned by arguments based purely in momentum space. In particular, I can ascertain the size of the particle-holes that make up the mid-infrared spectrum and determine the classes of orbital channels involved. This is important because it indicates the types of charge fluctuations present in the system, which will have bearing on how FeTe and related compounds should be treated in a many-body calculation.

First, I will discuss the spatial size of the particle-holes. The reproduction of the mid-infrared spectrum requires that I include particle-holes separated by distances of more than a nanometer. This is apparent in Fig. 3.5. In the optical conductivity, shown in Fig. 3.5a, including particle-holes within ≈ 1 nm reproduces a majority of the interband spectrum. An visualization of a representative particle-hole is shown in Fig. 3.6. One thing to note is that the orbitals appear to have no overlap and so one may ask how there is any transition between them. The answer is that, in addition to the two Wannier functions, there is an $\mathbf{x}$ in the integral for the optical

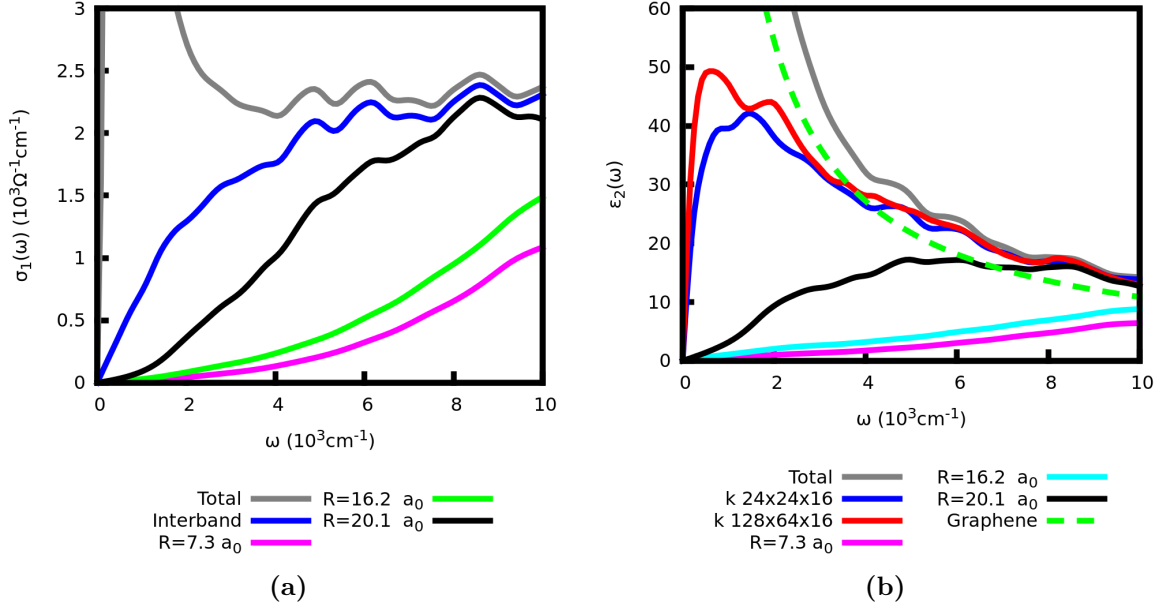


Figure 3.5: Spatially resolved contributions to the optical spectrum of FeTe. Using response in a Wannier basis, I can determine the size of the particle-holes that contribute to the optical spectrum. In (a) is just such a decomposition for the optical conductivity. A reasonable reproduction of the optical signal is not achieved until particle and hole are separated by more than 20 Bohr radii, equivalent to ≈ 1 nm. The mid-infrared is adequately reproduced with a macro-crystal size of $24 \times 24 \times 16$ unit cells. This is roughly equivalent to particle and hole separated by ≈ 5 nm. Thus, the difference between the blue and black curves in (a) is due to particle-holes with sizes between ≈ 1 and 5 nm. This difference is more dramatic in $\epsilon_2(\omega)$, which has a peak in the far-infrared, as shown in (b). Particle-holes less than 1 nm in size cannot reproduce this peak and only those up to 5 nm in size can adequately reproduce the peak. In fact, full convergence of the peak is not achieved until a macro-crystal with a size of $128 \times 64 \times 16$ unit cells is used. This reflects large, anisotropic particle-holes that are ≈ 24 nm large along the long dimension. These ultra-long-ranged particle-hole transitions are obscured in the optical conductivity because $\sigma_1(\omega) \propto \omega \epsilon_2(\omega)$.

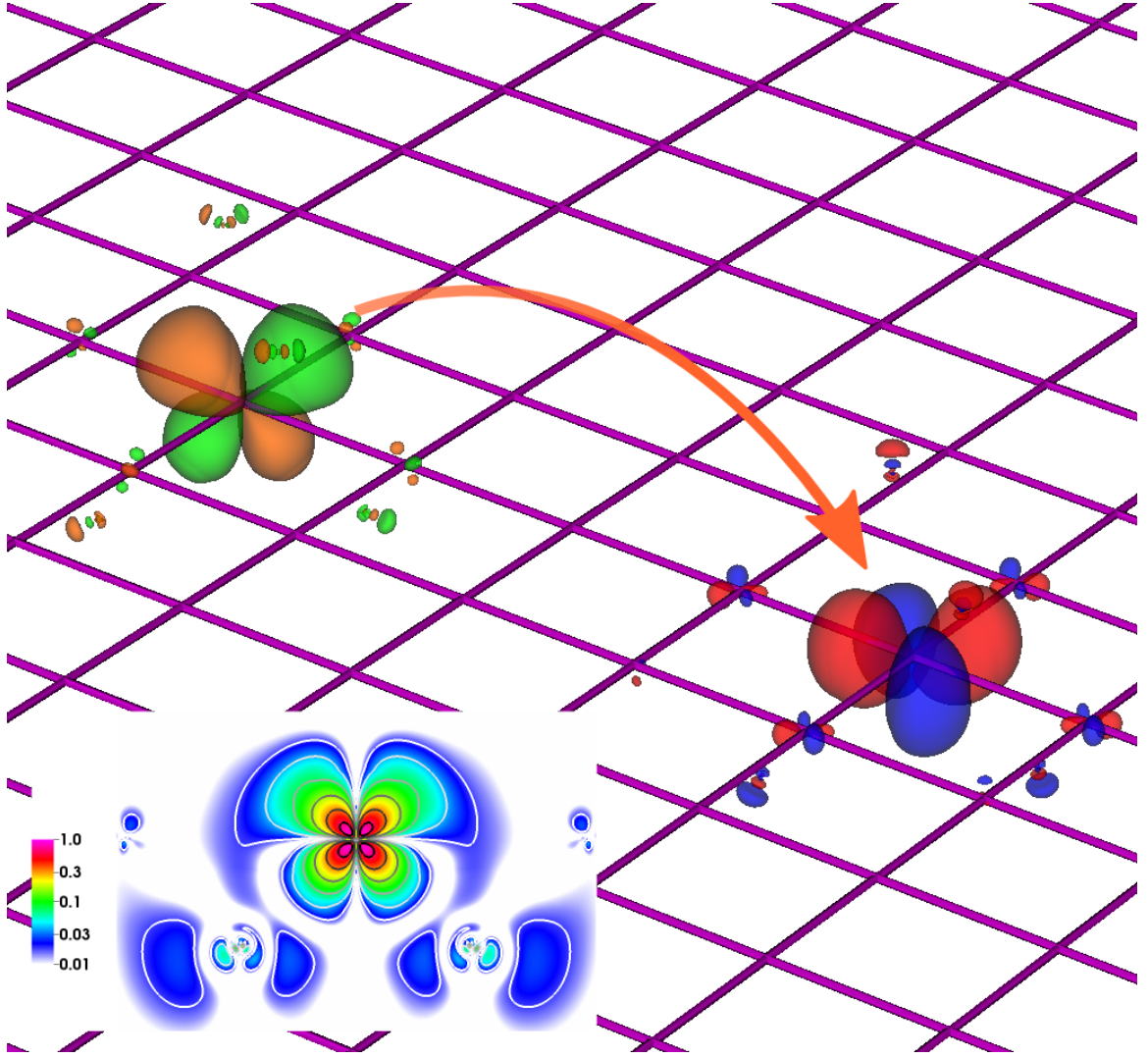


Figure 3.6: Real-space visualization of a typical optical transition in FeTe. Shown is a real space visualization a particle-hole transitions which is representative of the many transitions that make up the mid-infrared interband spectrum of FeTe. The hole is an xz/yz orbital and the particle is an xy orbital. They are separated by ≈ 1 nm, which is a minimal separation to begin to achieve a decent reproduction of the optical conductivity. Notice that each is quite different from the atomic-like local orbitals that were used as trial functions to construct these Wannier orbitals. In particular, the parity symmetry of an atomic-like d orbital is broken, as shown for the xz/yz orbital in the inset. This is also visible for the xy orbital. This is due to the presence of the Te atoms, which make it so that the Fe atom is not on an inversion center. This parity symmetry breaking allows for onsite optical $d-d$ transitions, though these have a small weight in the total spectrum.

matrix element. Around the 1 nm separation, the tails of the Wannier orbitals drop off slowly enough to produce a finite integral. This will be come important again later.

Also, I note as a curiosity that onsite transitions are included. This is despite the transition being between two d orbitals in the optical limit! But onsite d - d transitions are forbidden in the optical limit when the orbitals are parity eigenstates. In FeTe, the Fe atom does not sit on an inversion center of the crystal. Thus, the Wannier orbitals do not have to respect parity. This is plainly visible in the inset to Fig. 3.6. The presence of the Te atoms breaks inversion symmetry and the result is a finite (in fact leading, see inset to Fig. 3.7) optical matrix element.

Though looking at the conductivity leads one to think that particle-holes separated by ≈ 1 nm and within are enough to qualitatively describe the optical spectrum, looking at the dielectric function paints a different picture. As shown in Fig. 3.5b, there is a far-infrared peak in $\epsilon_2(\omega)$ that is missed by the particle-holes of sizes within ≈ 1 nm. It turns out that this is because in the far-infrared region, the particle-holes are on the order of the size of the macrocrystal. This makes direct investigation of the role of the various Wannier orbital channels beyond our current computational abilities. However, for such long-ranged particle-holes I can use the expression in Eq. (1.134) to sum classes of Wannier channels over the entire macrocrystal.

The reason for such long-ranged particle-holes is that the chemistry induced hybridization tails drop off as $\sim 1/R$ combined with the increased number of transitions at longer distances. This is clearly illustrated in Fig. 3.7, where I have multiplied a Wannier function by x and shown the result at a maximum distance from the center of about 25 nm. With a macrocrystal of 24x24x16 unit cells, represented by the black dashed box in Fig. 3.7, I am able to reproduce the peak in epsilon shown in Fig. 3.5b. This corresponds to particle-holes about 4 nm in size. There is a significant drop off in the Wannier function beyond this range, but as the red curve in Fig. 3.5b shows, full convergence of the far-infrared peak requires particle-holes separated by

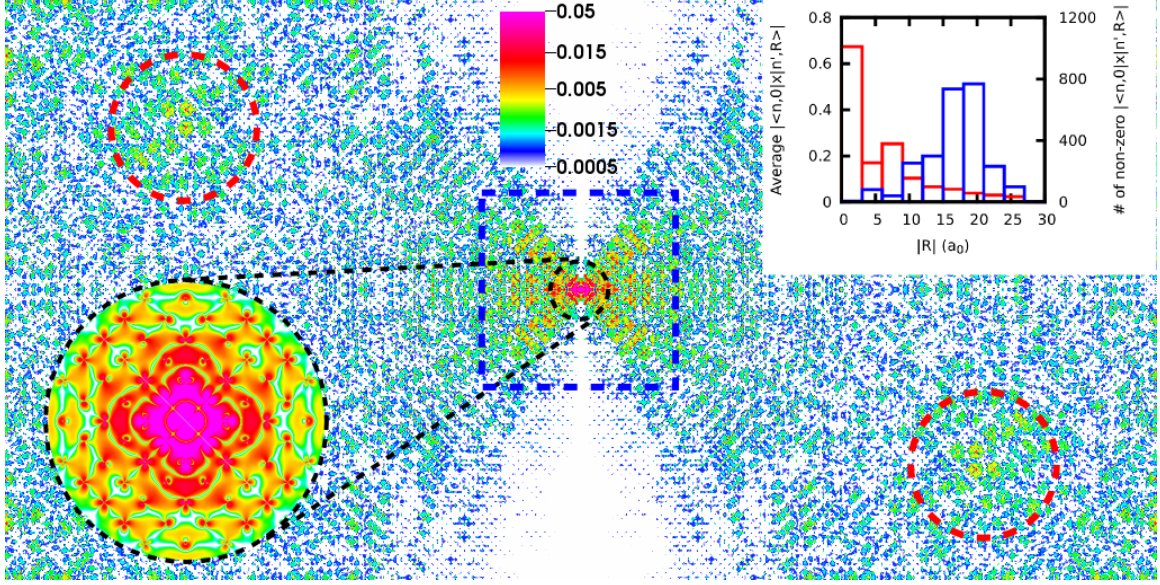


Figure 3.7: Visualization of the long-ranged nature of the particle-hole transitions in FeTe. Shown is a real-space map of the absolute value of the Wannier xy orbital multiplied by the coordinate x on a log scale with color values shown in the legend. When discussing optical transitions, it is appropriate to consider this quantity because it represents one half of the integrand that takes part in calculating the Wannier matrix element for the response in the optical limit. In particular, if this quantity is small at a location, then the matrix element between the central orbital and an orbital at that location will be small. Several length scales are shown, which represent different, significant contributions to the optical spectrum. One is the length scale of 1 nm, shown by the black dashed circle with a zoomed image as an inset. Particle-holes with one orbital in the center and the other in this circle represent the minimum required to adequately describe the mid-infrared spectrum. Full description of the mid-infrared requires particle-holes across the entire macro-crystal of size $24 \times 24 \times 16$ unit cells, marked by the blue dashed square. Part of the reason for the importance of long-ranged transition, even as their value decreases with distance, is that the number of such transitions increases with distance to compensate, as shown in the histogram in the inset. Though the onsite transitions makes the largest single contribution, there are very few of them. In contrast, the transitions between ≈ 15 and 20 Bohr radii are about ten times smaller in value, but there are on the order of 100 times more of them. The longest-ranged optical transitions are at ≈ 20 nm and contribute to the far-infrared peak in $\epsilon_2(\omega)$. These are due to islands of relatively large value of the Wannier function multiplied by x highlighted by the red dashed circles.

tens of nanometers! This is due to islands of relatively large tails in the Wannier function, highlighted by the red dashed circles in Fig. 3.7.

The extreme range of particle-holes in effect returns me to a momentum-space picture of the optical response. The goal in this case is to identify the regions in the first Brillouin zone that produce the interband transition and the orbital character of the bands involved. I find that a pocket that is centered on the M point and extends along the Γ -M direction makes the dominant contribution to the far-infrared peak in the dielectric function. A remarkably flat, occupied band hugging the Fermi level is predominantly xy (in my coordinate system) and z^2 at the M point and switches to xz/yz as it approaches the Γ point. There is a complex of dispersing xy , x^2-y^2 and xz/yz bands along Γ -M that must provide the empty states for the transition.

In the mid-infrared region, where particle-holes are separated by up to ~ 1 nm, it is possible to investigate the role of various classes of Wannier orbital channels in producing the interband spectrum in the optical conductivity. I find three classes of channels that constitute the interband transitions up to $\approx 10,000$ cm^{-1} , each making up a roughly equal portion of the spectral weight, as shown in Fig. 3.8. These are 1) transitions among the xz/yz orbitals, 2) transitions between the xy orbital and the xz/yz orbitals, and 3) transitions between the x^2-y^2 orbital and the xz/yz orbitals. Thus, the z^2 orbital plays virtually no role and the dominant transitions are from one type of orbital to another rather than from one type of orbital to the same type on a different site. Only the xz/yz orbitals exhibit relevant transitions that take e.g. an xz/yz orbital on one site to an xz/yz orbital several sites away. Notably, there are virtually no transitions within or between the xy and x^2-y^2 orbitals.

One must be cautious about assigning one orbital as a hole and another as a particle as all of the orbitals are partially occupied and partially empty. All of the orbitals could be both particle and hole. This is true near the Fermi level, in the energy region relevant for the mid-infrared optics. As visible in Fig. 3.4, there are regions in the Brillouin for each orbital in which they are either occupied or empty. However, I can make a couple of approximate statements. First, in the energy region

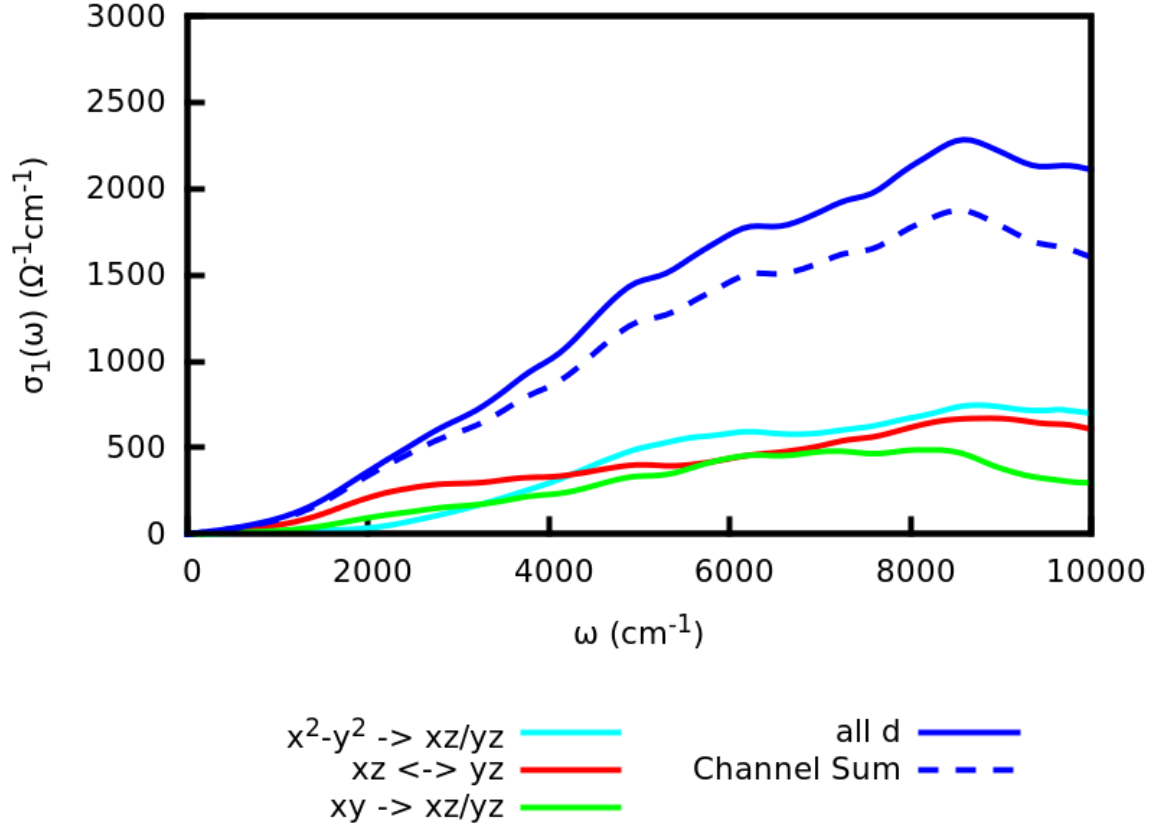


Figure 3.8: Decomposition of the mid-infrared spectrum of FeTe according to classes of Wannier channels. I am able to use response in the Wannier basis to decompose the contribution to the mid-infrared spectrum due to particle-holes up to 1 nm in size into classes of channels. All of the channels have as the particle an xz/yz orbital. The classes separate according to the hole of the transition with three primary contributors: x^2-y^2 , xz/yz and xy . These are represented by the cyan, red and green lines and each class makes an approximately equal contribution to the spectrum. The sum of all of these classes is shown by the blue dashed curve and clearly these three classes constitute the primary contribution to the mid-infrared spectrum.

of interest, the x^2-y^2 orbital is mostly unoccupied, with the only occupied region very near the Γ point, as visible in Fig. 3.4a. Thus I can say that the x^2-y^2 orbital is a particle in the transitions making up the mid-infrared. Second, most of the xy content in the relevant two-dimensional complex of bands is in the occupied, linearly dispersing band. Thus, I can say that the xy orbitals is mostly a hole in the transitions that make up the mid-infrared. The xz/yz orbitals must be taken as play the roles of both particle and hole when taken as a whole throughout the Brillouin zone.

Thus, I am able to use the calculation of the optical response in the Wannier basis to establish the orbital nature of the particle hole content and establish the size of the particle-holes that must be considered. I find that there are several length-scales that are progressively necessary to describe more and more of the spectrum. Particle-holes on the order of 1 nm are required to describe the mid-infrared, several nanometers to describe the far-infrared peak in $\epsilon_2(\omega)$ and tens of nanometers to fully converge this peak. This importance of long-ranged charge fluctuations is in contrast to the local correlations that are important in the groundstate. This shows the necessary coexistence of disparate length scales and I argue that both of these scales must be considered in a complete theoretical description of FeTe and related compounds. Even in the current context of optical spectra, there is a need for a more complete theoretical treatment as the reproduction of the far-infrared does not agree with experiment. I will turn to this issue in the next section.

3.5 Improving the far-infrared

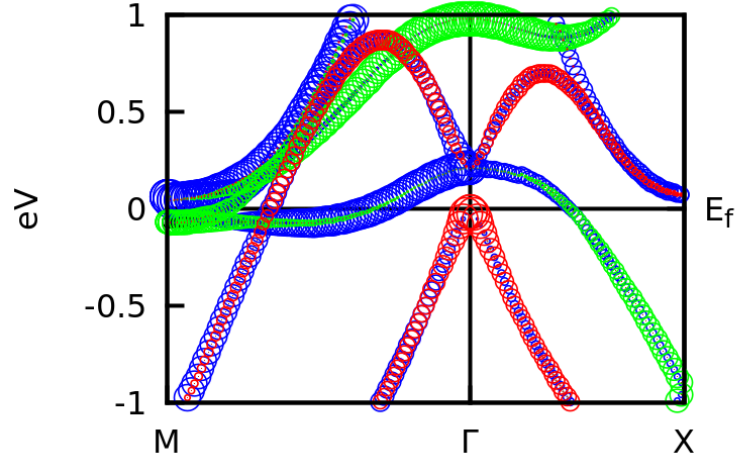
Though I am able to achieve agreement to experiment down to $\approx 2,000 \text{ cm}^{-1}$, it is clear that in Fig. 3.5 the interband response begins to deviate from a $1/\omega$ behavior of ϵ_2 at $\approx 4,000 \text{ cm}^{-1}$. Thus, I am helped somewhat between 2,000 and 4,000 cm^{-1} by the tail of the Drude, intraband response. This is too large in the calculation when compared to experiment, which also accounts for the dramatic deviation from experiment below 2,000 cm^{-1} . So, if the intraband response is reduced to come in

line with experiment, the interband contribution to ϵ_2 must track the $1/\omega$ trend to lower frequencies. This will entail an enhancement of the interband peak in epsilon shown in Fig. 3.5b. Thus, any proposal as to how to improve the current calculation must have as simultaneous consequences the reduction of the Drude weight and the increase of the interband response at frequencies below $4,000 \text{ cm}^{-1}$.

I believe the most likely place where the calculation could be improved is in the vicinity of the Γ point. Figure 3.9 shows the topology of the bands of FeTe at the Γ point in contrast to that of LaFePO. As is apparent in Fig. 3.9a, there is a gap between the occupied and unoccupied states around the Γ point. This means that the interband contribution to the optical spectral weight for low energies will be reduced relative to if there were no gap. If the lowest two unoccupied bands which become degenerate at the Γ point could be shifted down in energy and the lowest unoccupied band that continues to become the flat band at M could be flattened, then there would be both more interband contribution to the low-energy optics and the effective mass at the Fermi level would be increased, decreasing the Drude weight.

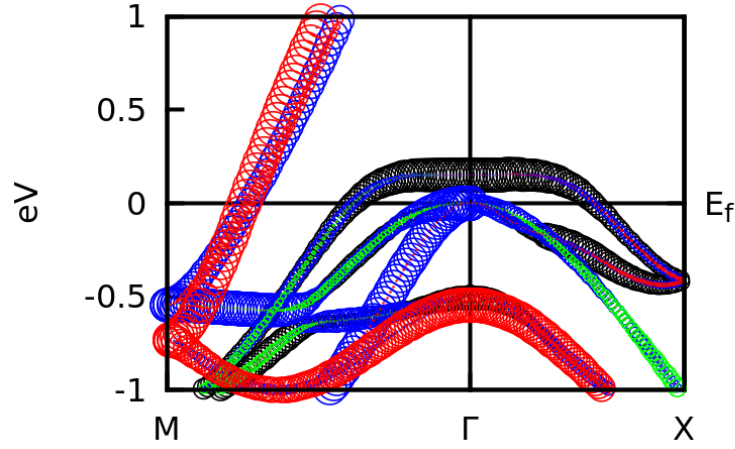
However, in order to reproduce a graphene like conductivity, the bands participating must also be two-dimensional. As is visible in Fig. 3.4, the V-shaped occupied and unoccupied bands at the Γ point in Fig. 3.9a strongly disperse along the k_z direction indicating that they are three-dimensional. This may represent an over estimation of the three-dimensionality of the bands as the method I used to include correlations, the GGA+ U , assumes spherically symmetric interactions when FeTe is in reality anisotropic due to the local chemistry involved with the placement of the Te atoms and the layered crystal structure. A more realistic treatment of the interactions may result in increased confinement of electrons to the Fe-Te layers, reducing interlayer hopping and hence making the bands more two-dimensional.

There is some support from photoemission data for this proposal on what the features of an improved electronic structure may be. Though photoemission is blind to the unoccupied states, Lin et al. [24] observed an *occupied* band around around the Γ point that qualitatively resembles the lowest *unoccupied* in my calculation.



(a)

z^2 ○ x^2-y^2 ○
 xz/yz ○ xy ○



(b)

Figure 3.9: Low-energy GGA+ U band structure for FeTe and LaFePO. By comparing (a) and (b), it is clear that the low-energy band structure is dramatically different between FeTe and LaFePO, a representative of the pnictides. In particular, there is no evidence for band complexes with linearly dispersing energy differences in (b), excluding a graphene-like optical response. This difference between the band structures can explain the qualitative differences in the optical spectrum of FeTe versus the pnictides.

In addition, this band appears to be degenerate with a lower band at the Γ point, which may correspond to the upside-down V-shaped band in my calculation. This is supported by the fact that this lower band in experiment is quite incoherent indicating strong correlations and the corresponding band in my calculation is heavily *xy* (the correlated orbital in my calculation and in experiment) near the Γ point (see Fig. 3.9a). If the correspondences between the experimental bands and my calculated bands are correct, then this indicates that in the real system the band above the gap at the Γ point in my calculation is shifted downward in energy until it touches the upside-down V-shaped band. This is precisely what I have proposed is needed for the interband contribution to the optical response to stretch to lower energies.

Given the importance of beyond nanometer scale charge fluctuations in producing the far-infrared spectrum in my calculation, I believe any further theoretical work to improve the far-infrared would have to be able to handle long length scales and restriction to physics at the level of nearest neighbors would be insufficient. The ability to conduct such a theoretical investigation that also includes many-body physics beyond GGA+ U will be assisted by the formulation of an *ab initio* microscopic model of FeTe. I will turn to the construction of just such a model in the next chapter.

Chapter 4

Construction of an optimized Hamiltonian for FeTe

Since the many-body methods that are often used become rapidly more numerically demanding with increasing degrees of freedom, a primary goal in constructing low-energy models is finding a minimal number of degrees of freedom. The question of a minimal model for the iron-based superconductors, of which I will be using FeTe as a representative, has been anything but clear. An early model for the pnictides was a two-orbital Hubbard model with one Fe atom per unit cell [182–184]. However, as pointed out by Graser, Maier, Hirschfeld and Scalapino [185], the two-orbital model fails to describe the experimentally observed Fermi surface. Thus, it was determined that a five-orbital model was minimal [186–188], though three [189, 190] and four [191, 192] orbital models have been utilized and, going the other way, ten [193] and sixteen [194, 195] orbital models that include the more realistic two Fe atoms per unit cell have been studied. Determining the fewest degrees of freedom necessary for a given set of physical observables and constructing a minimal model is vital to the study of the iron-based superconductors and multiorbital systems in general.

In this chapter, I will present a method for optimizing the construction of an *ab initio* model that naturally leads to the minimal number of orbital degrees of freedom per site. This is done by systematically eliminating chemistry-induced hybridization among the d orbitals leading to a “target” subspace whose single-particle bands are disentangled from the “rest,” i.e. the degrees of freedom outside of the target. In both FeTe and LaFeAsO, the construction of a minimal target necessitates the classification of the d orbitals into “ t_{2g} ” and “ e_g ” subspaces and splitting the e_g orbitals into bonding and anti-bonding orbitals reflecting *cation-cation bonding chemistry*. Remarkably, the e_g bonding and anti-bonding orbitals in the optimal solution correspond directly to the experimentally observed differences in the structural and magnetic phase transitions of FeTe and the iron pnictides, of which LaFeAsO serves as a representative. Using CRPA, I will calculate the interaction parameters for the “ t_{2g} model” and compare it to the “ d - p model” and “ d model” that have been previously studied [106]. The trend in the interaction parameters with decreasing target subspace size is that of increasing importance of Hund’s coupling J relative to the Hubbard U , decreasing spherical symmetry and increasing orbital differentiation. I will also discuss the importance of long-ranged physics in both the single-particle hopping and two-body interactions that will have implications for the future development of solvers for the *ab initio* models.

Though I will be presenting an optimization procedure that focuses on using the groundstate electronic structure as a guide in the form of eliminating d - d hybridization, it is important to place this approach in the context of the broader work I have presented in this dissertation. In Chapter 2, I presented a method for optimizing the choice of gauge used in the construction of the Wannier orbital basis with respect to optical spectroscopy, with Cu_2O providing the testbed. However, performing the optimization in that way necessitated the splitting of the z^2 orbital into molecular bonding and anti-bonding orbitals to facilitate the elimination of hybridization between the z^2 orbital degrees of freedom and the other d orbitals (see Section 2.3 and particularly Fig. 2.11.) Thus, optimization through the elimination

of hybridization and optimization through achieving an accurate description of the optical spectrum represent two sides of the same coin. The benefit of optimization with respect to groundstate properties is that it does not require the repeated calculation of χ^{KS} which represents savings in both computer and human time.

In light of the close kinship between the optimization procedure I present in this chapter and the one with respect to optics, it is fruitful to discuss briefly the relationship between achieving a minimal description of the optical spectrum and the utility of the minimal model that is produced. The method that I have developed in Chapter 2 is well suited to provide answers as to the minimal number of degrees of freedom needed in an eventual model. One can think of the optical spectrum calculated in the Wannier basis as the result for the model in the RPA before any many-body corrections are made to the groundstate. In this sense, finding a set of orbitals that can most simply describe the optical spectrum is the same as finding the minimal number of orbital degrees of freedom in the model that is required to yield the same spectrum.

The minimum number of orbitals required to describe the low-energy RPA optics can certainly set an educated lower bound on the required number of orbitals. If a set of orbitals cannot adequately describe the RPA spectrum calculated using all necessary degrees of freedom, as in a calculation in the Bloch basis, then it will be insufficient to describe other quantities at the same energy scale, no matter what many-body corrections are used. A poor starting point will yield poor results. However, even degrees of freedom that are relevant at higher energies in the RPA optical spectrum could be relevant to the low-energy optics after many-body corrections are introduced. For example, they may be important in virtual transitions in a diagrammatic approach. Thus, more orbital degrees of freedom may be required in a model than are required to reproduce the RPA spectrum. Of course, the precise answer will depend on what observables are to be calculated.

I will begin the discussion with a description of the single-site projection in Section 4.1. This construction of the Wannier orbitals uses methods widely used [2],

but I will show that these methods are inadequate for producing an optimal set of orbitals. In a manner similar to that described in Section 2.3, incorporating the strong, offsite hybridizations inherent to the cation-cation bonding chemistry apparent from examining the single-site projection requires the use of the extension of the projection method that I developed as part of this dissertation (described in Section 1.4) which uses Eq. (1.112) to construct molecular orbitals. I will present this construction in Section 4.2.

Constructing the molecular orbitals will involve choosing between two different ordering patterns for the e_g orbitals. I will show that the optimization process I have developed results in a striking correspondence to the monoclinic structural and bi-collinear magnetic phase transition present in FeTe and is able to distinguish this from the orthorhombic and antiferromagnetic transition present in LaFeAsO [166, 196]. Showing this required me to develop an algorithm to produce a $\mathbf{k}$ -dependent energy window for LaFeAsO, which I will also describe in Section 4.2.

In Section 4.3 I will give a detailed description of the orbitals that make up the optimal Wannier basis. There I will highlight the crucial role of eliminating hybridization between t_{2g} and e_g in producing a disentangled t_{2g} model. Eliminating hybridization will also be shown to be critical to achieving a description of the low-energy optical features that only involves the t_{2g} orbitals.

In Section 4.4, I will present the first-principles-derived parameters for the t_{2g} model and compare them to the d and $d-p$ models. There I will examine the trends which reveal that the interaction parameters of t_{2g} model displays increased importance of Hund’s coupling and increased orbital differentiation. These observations are in line with the view of FeTe as a Hund metal [31, 32] and the experimental evidence for orbitally selective correlations [24, 26]. In addition to examining the onsite interactions, I will show that long ranged hopping on the order of several nanometers must be included to correctly reproduce the GGA band structure.

4.1 Single-site projection

The minimal number of orbital degrees of freedom will depend on the chemistry, which will determine the shape and energy location of the orbitals. FeTe is a metal, so we are primarily concerned with the degrees of freedom near the Fermi level. In order to start to understand the chemistry of the orbital degrees of freedom in FeTe, I will construct an initial set of Wannier orbitals in a single-site projection scheme. This is the projection method described in Section 1.4 which uses Eq. (1.106). The first step in performing this construction is the determination of the coordinate system.

The crystal structure of FeTe consists of layers of Fe atoms in a two-dimensional square lattice with Te atoms above and below the Fe plane, tetrahedrally coordinated with the Fe atoms. The strong anisotropy of this structure makes the direction normal to the Fe plane the logical choice for that of the local z -axis. With the z -axis fixed, there remains two reasonable choices for the orientation of the x - and y -axes. In one option, the x - and y -axes are pointing towards the Fe next-nearest neighbors, placing maxima of the xz and yz orbitals towards the Te atoms. I choose this option so that the Fe-Te bonding chemistry will be clearly manifest in the xz and yz orbitals. This is similar to the choice of local z -axis that I made for Cu_2O that placed the z^2 orbital such that its maxima were pointed towards the O neighbors.

The other option is to place these axes towards the Fe nearest neighbors, effectually utilized elsewhere [172] and the more logical from the stand point of the Fe square lattice. As this convention is used widely in the designation of the orbitals in the iron-based compounds in general. I will name the orbitals according to this scheme. Thus, the x^2-y^2 will have maximal density towards the Fe nearest neighbors and the xy towards the Fe next-nearest neighbors. The z^2 orbital does not depend on this choice of x - and y -axes. The xz and yz orbitals form an irreducible representation of the site symmetries, are thus degenerate and must always be considered together. Thus I call them the xz/yz orbitals. As a set, they do not depend on the choice of x - and y -axes. The local coordinate system for Fe(2) is rotated by 180° with respect

to that in Fe(1) to reflect the glide plane symmetry so that the z -axis for Fe(2) is in the opposite direction as that for Fe(1).

With the local coordinate systems chosen, it is now necessary to choose the energy window for projection. Since I would like to consider almost exclusively the Fe d orbitals and not have to concern myself with the Te p orbitals I will project the p orbitals in one window and the d orbitals in a different window. Ideally, these windows would not overlap, but there is a band crossing between the predominantly p and predominantly d bands so that the top of the p band lies above the bottom of the d band. Thus I must choose overlapping windows, with the p window running from -6 eV to $\approx$ -0.5 eV and the d window running $\approx$ -2.5 eV to 2.5 eV (see Fig. 4.1). Fortunately, the chemistry of FeTe is such that there is very little hybridization between p and d orbitals so that the resulting Wannier functions are in approximately disjoint subspaces of the Bloch states. That is, the d orbitals constructed from overlapping windows can be thought of as independent from the p orbitals, as desired.

The DOS for the resulting projection is shown in Fig. 4.1. The Wannier orbitals can be classified into two qualitatively different types. The xy and xz/yz orbitals dominate in a $\approx$ 1 eV window about the Fermi level. These I will call collectively the t_{2g} orbitals. The x^2-y^2 and z^2 are depressed at the Fermi level and instead dominate away from the Fermi level. These I will call the e_g orbitals. I stress that these are simply names in this context reflecting that these orbitals are true t_{2g} and e_g orbitals in the cubic environment. FeTe has a tetragonal structure, so there is no t_{2g} irreducible representation and the e_g irreducible representation is different, being spanned by the xz/yz orbitals. Still, since the orbitals classify in this way in terms of their chemical nature, as evinced in the DOS, they are useful names.

One reason for choosing FeTe in particular is that this classification in terms of t_{2g} and e_g orbitals is especially clean with respect to energy location in the DOS. In FeTe, the t_{2g} orbitals are especially concentrated at the Fermi level and the e_g orbitals are so depleted in this region that the DOS due to these orbitals resolve into clear peaks above and below the Fermi level. This splitting is similar to the two-peak structure

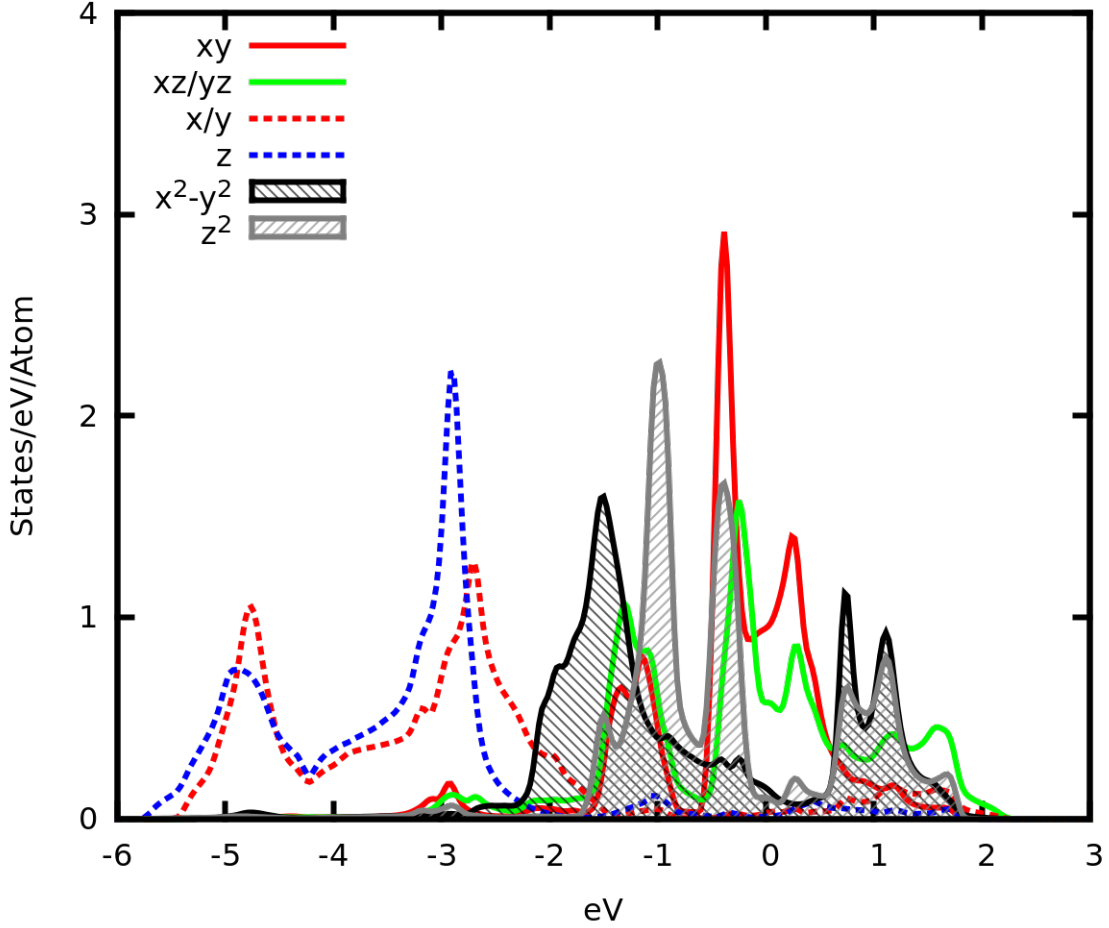


Figure 4.1: FeTe single-site projected DOS. Shown is the Wannier decomposed DOS for orbitals constructed using the single-site projection scheme. I projected Te p orbitals and Fe d orbitals into two different energy windows. The energy window for the p orbitals runs from -6 eV to ≈ -0.5 eV, while the window for the d orbitals runs from ≈ -2.5 eV to 2.5 eV. The windows must overlap because there is a band crossing between the predominantly p band and predominantly d band. However, as there is little chemical hybridization between p and d orbitals in FeTe, there is little hybridization in the DOS resulting from these overlapping windows. The xy and xz/yz orbitals (which I call the t_{2g} orbitals in an abuse of nomenclature, but in line with a convention in the literature) dominate at the Fermi level, while the x^2-y^2 and z^2 (which I call the e_g orbitals) have a pronounced depletion of content about the Fermi level. This is more evident for the x^2-y^2 , which in real space has maximal density pointed toward the Fe nearest neighbors. This is reminiscent of the z^2 orbital in Cu_2O and the cause is the same: a bonding/anti-bonding chemistry for the e_g orbitals.

of the z^2 DOS in Cu_2O , which I have shown there to have resulted in a route toward optimization of the Wannier orbital set. Such a structure is also present in other parent compounds of the Fe-based superconductors, but I found it to be especially clear in FeTe, making it a clean test case.

Within each group, t_{2g} and e_g , there is considerable hybridization between the orbitals that would make it difficult and likely counter-productive to further break them into subgroups. This is apparent in the DOS shown in Fig. 4.1 by looking at matching line shapes in the peaks. Between the t_{2g} orbitals this is clear both in the main peak at the Fermi level and the smaller peak at ≈ -1.3 eV, where the line shapes of the xy and xz/yz orbitals closely resemble each other. Between the e_g orbitals the hybridization is more pronounced at the double peak at ≈ 1 eV. This hybridization makes it appropriate to consider the orbitals in each group together, as a unit rather than individual orbitals.

Though the d orbitals are nicely classified into t_{2g} and e_g , this is not to say that these orbitals are completely separate from each other. In fact, there is considerable hybridization between the two groups, particularly apparent in the peaks at ≈ -1.1 and -0.4 eV, though there is hybridization throughout. This means that there is a qualitative classification in terms of chemistry between the t_{2g} and e_g orbitals, but the single-site projection results in orbitals that are heavily entangled. Since the t_{2g} orbitals dominate near the Fermi level, it is those that are most likely to dominate the optical signal and those that would be most desirable for forming a model with a minimal number of degrees of freedom. However, both simple explanation of the optics and construction of a t_{2g} -only model would require greater disentanglement.

4.2 Bonding/anti-bonding projection and orbital ordering

The route to optimization is furnished by the work I have done for Cu_2O in Chapter 2. There, cation-cation bonding produced a two-peak structure in the Cu z^2 DOS. By splitting the z^2 into independent bonding and anti-bonding orbitals, I was able to disentangle the critical hole for the optical excitations from the other d orbitals. The same two-peak structure in the DOS exists for the e_g orbitals in FeTe. Thus I will first attempt to split the window for the e_g orbitals into two parts. One running from ≈ -2.5 to 0.0 eV and the other from 0.0 to 2.5 eV while projecting the t_{2g} orbitals in the same window as for the single-site projection.

In order to place independent e_g orbitals in the two separate windows, it is necessary to form a linear combination of local orbitals using Eq. (1.114). This is to allow for strong offsite hybridization to become manifest in real space due to the bonding and anti-bonding nature of the resulting orbitals. I placed contributions from the same combination of t_{2g} orbitals on each nearest-neighbor Fe site. There will be one orbital per Fe atom per energy window. Using the same combination on each site ensures that the symmetry of the final orbital will respect the crystal symmetries and using t_{2g} orbitals will aid in allowing the e_g orbitals to have contributions from bands that are primarily t_{2g} in the single-site projection. This latter point is important as “room” must be made for the e_g orbitals in the more restricted energy window.

The resulting DOS is shown in Fig. 4.2. The two e_g orbitals I call the e_g bonding orbitals and e_g anti-bonding orbitals to reflect their chemical nature. The e_g bonding orbitals are almost entirely occupied while the e_g anti-bonding orbitals are mostly unoccupied. By virtue of this much of the hybridization between the e_g in the single-site projection is removed, particularly for the double peak at ≈ 1 eV. This allows me to separate the e_g orbitals into two groups, which was impossible for the single-site projection. Again, as with the classification of the d orbitals into t_{2g} and e_g , the

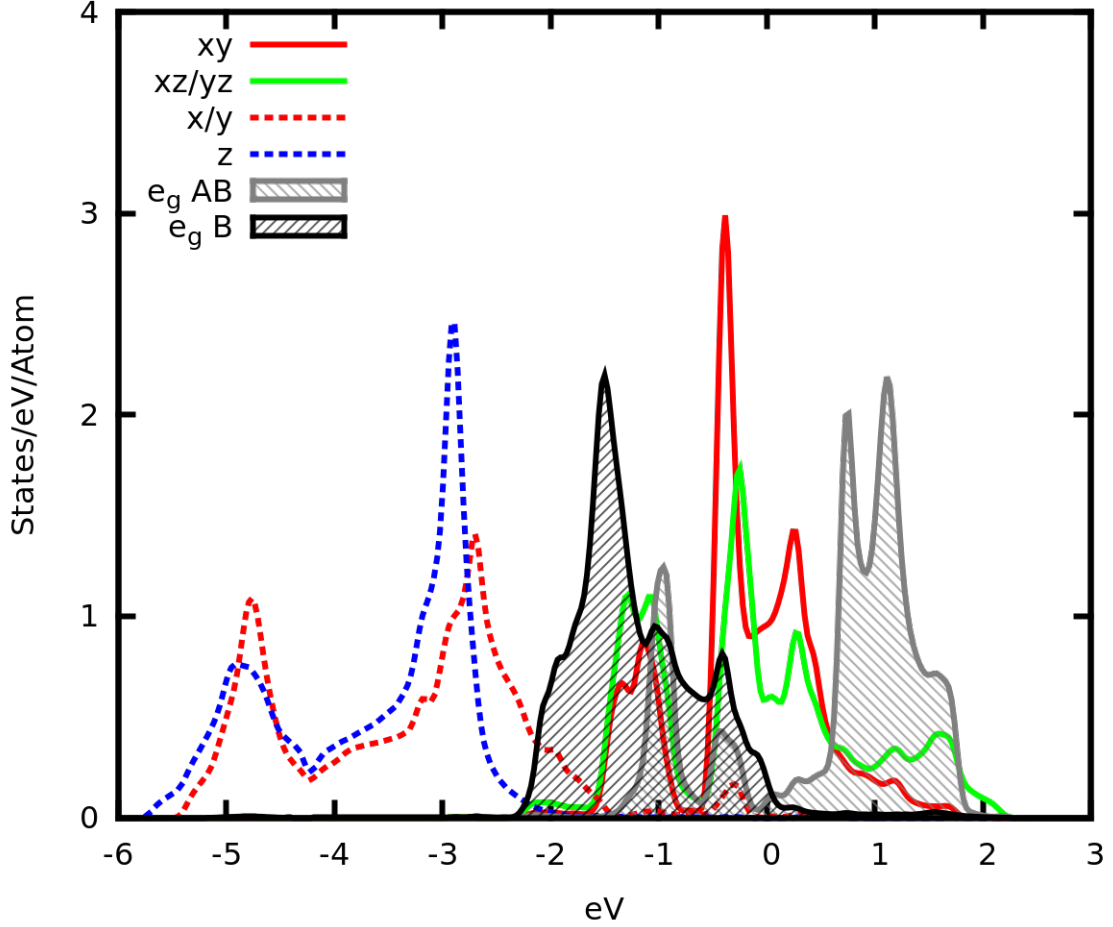


Figure 4.2: FeTe DOS using a projection with e_g bonding and anti-bonding orbitals. Shown is the Wannier decomposed DOS where I have split the e_g orbitals into bonding and anti-bonding orbitals. The energy windows for the p and t_{2g} orbitals is the same as in the single site projection, but the energy window for the e_g bonding orbital runs from ≈ -2.5 eV to 0.0 eV and that for the e_g anti-bonding orbital runs from 0.0 eV to 2.5 eV. As a trial orbital for both I used a linear combination of local orbitals with symmetric contributions from each of the four neighboring Fe sites. The result are bonding and anti-bonding orbitals that are mostly occupied and empty, respectively. However, there is still significant hybridization between t_{2g} and e_g orbitals, as visible in the peaks at 1.0 eV and 0.5 eV. This means that if one wants to construct a model for only the t_{2g} orbitals, this projection is not suitable.

classification of the e_g orbitals into bonding and anti-bonding is one based on the differing chemical nature of the orbitals. From this perspective, the bonding/anti-bonding classification is more desirable than the single e_g group, because it reveals the chemistry of the compound.

Though I have been successful in placing one set of e_g orbitals below the Fermi level and one largely above the Fermi level, there remains nearly the same level of hybridization between t_{2g} and e_g orbitals. This is particularly visible in the peaks at ≈ -1.1 and -0.4 eV. There is also still some remaining hybridization between the e_g orbitals at ≈ -0.4 eV. This latter location is particularly important for optimizing the orbitals with respect to the optical spectrum and for creating a minimal low-energy model. It is quite close to the Fermi level and one can expect, especially in light of my work on Cu_2O in Chapter 2, that the e_g orbitals will make considerable contributions to the low-energy optics and would have to be included in a minimal model. Thus, I would like to further optimize the orbitals. However, this requires a closer examination of the e_g orbitals, and in particular the e_g bonding orbitals.

I am able to produce the projection with bonding and anti-bonding e_g orbitals whose DOS is shown in Fig. 4.2 even if the central part of the linear combination to produce the orbitals contains only the x^2-y^2 trial orbital, with no contribution from the z^2 trial orbital. Of course, in the final Wannier orbital basis, what was the z^2 orbital in the single-site projection must go somewhere and it implicitly ends up in the hybridization tails on neighboring sites. However, the apparent disappearance of the z^2 orbital is not very physically pleasing. And as a practical matter, exclusion of the z^2 trial orbital from the linear combinations does not allow any further optimization of the orbitals.

Instead, I will use an onsite linear combination of the x^2-y^2 and z^2 trial orbital so that the z^2 trial orbital is explicitly included in the projection of the Wannier functions. There are two inequivalent choices for this linear combination. If I fix Fe(1) to have the combination $|x^2 - y^2\rangle + |z^2\rangle$ up to a normalization, then I can choose the combination on Fe(2) to be $|x^2 - y^2\rangle + |z^2\rangle$ or $|x^2 - y^2\rangle - |z^2\rangle$. The first

choice for Fe(2) will produce an orbital that is rotated 90° with respect to that on Fe(1), whereas the second choice will produce an orbital that is unrotated. The first ordering pattern I refer to as “zigzag” ordering and the second as “parallel” ordering. It turns out that in FeTe, I am only able to optimize the orbital basis further by choosing zigzag ordering. This reflects a preference for this ordering pattern of the bonding and anti-bonding orbitals that is *intrinsic* to FeTe and not imposed by me.

The origin for this preference can be deduced from the band structure, as shown in Fig. 4.3. The orbital content of the bands for the e_g bonding and anti-bonding orbitals is largely similar for the two ordering patterns, as visible in Figs. 4.3a and 4.3c. However, there is a crucial difference at the Γ point. There are two bands that are ≈ 0.4 eV and ≈ 1.0 eV below the Fermi level, highlighted by green circles. For zigzag ordering, the lower band has more e_g bonding content, while for parallel ordering the upper band has more content. This means that for parallel ordering, the top of the bonding orbital’s energy window cannot be compressed past ≈ 0.4 eV and the hole pocket ends up with remaining entangled e_g content. On the other hand, for zigzag ordering the window can be compressed much lower and the hole pocket is cleared of e_g content.

I can apply the same analysis to another iron-based superconducting parent compound for comparison. In LaFeAsO, the equivalent of the two bands that I identified for FeTe are at ≈ 0.2 eV and ≈ 0.5 eV below the Fermi level. For the parallel ordering, the e_g bonding content is predominant in the lower band, as shown in Fig. 4.3d. In exactly the same way as the orbital content of this band predicted zigzag ordering in FeTe, it predicts a preference for parallel ordering in LaFeAsO. For zigzag ordering, I again find a reversal from FeTe: in LaFeAsO the upper band has more e_g bonding content where for FeTe it was the lower band that had more content, as shown in Fig. 4.3b. Thus the Wannier projection can distinguish between zigzag ordering for FeTe and parallel ordering for LaFeAsO.

This is very suggestive, because this corresponds to the magnetic and structural phase transitions present in these materials, as I have described in Section 3.1. Below

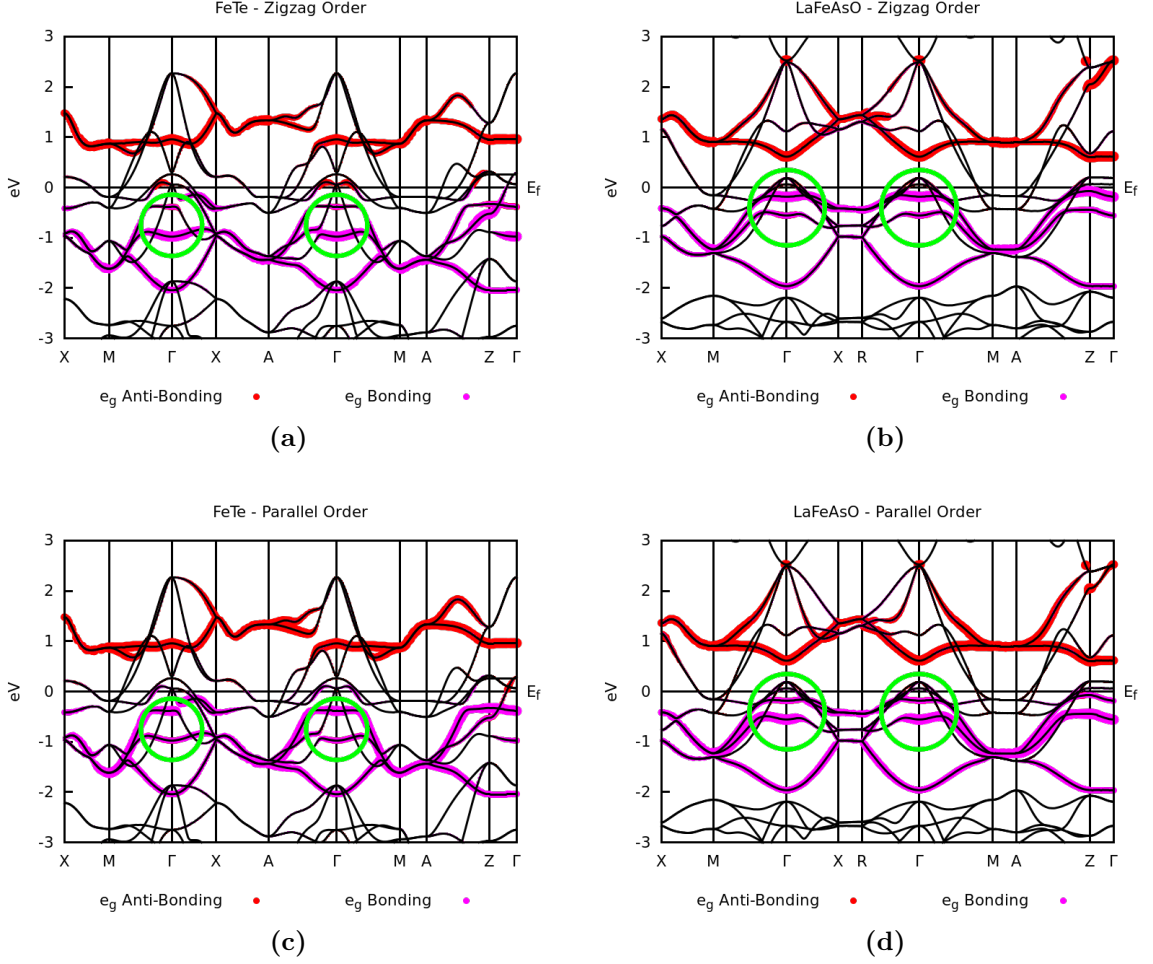


Figure 4.3: Band structure for zigzag versus parallel ordering of the e_g orbitals in FeTe and LaFePO. By examining the orbital content in the band structure of the e_g bonding orbital I am able to determine a preferred ordering pattern for each compound. The orbital content for the e_g bonding and anti-bonding orbitals is largely the same across the band structure. However, there are critical differences in the hole pocket at Γ . These are in particular lower and higher energy bands in FeTe at ≈ -1.0 eV and ≈ -0.4 eV, respectively. The corresponding higher and lower energy bands in LaFePO are at ≈ -0.5 eV and ≈ -0.2 eV respectively. These differences are marked by green circles. The orbital content for the zigzag configuration for FeTe in (a) shows more content in a lower energy band relative to the zigzag configuration for LaFePO in (b) which has more content in the higher energy band. The situation is flipped for the parallel configuration. For FeTe shown in (c), there is more e_g bonding content in the higher energy band and for LaFePO shown in (d), there is more content in the lower energy band. Therefore, energetically the zigzag configuration is preferred for FeTe and the parallel for LaFePO.

the Néel temperature, FeTe exhibits a bi-collinear, zigzag ordering pattern of its magnetic moments and a monoclinic distortion that results in a compression of the Fe square lattice along one of the square diagonals. On the other hand, LaFeAsO and the iron-pnictides broadly exhibit an antiferromagnetic order in one direction along the edge of the square and ferromagnetic along the other with an orthorhombic distortion that compresses the square lattice along one of the edges. The zigzag ordering with bonding along alternating directions and the parallel ordering with bonding along one direction points directly to these differing structural distortions. This is despite the fact that there is no distortion in my calculation.

Thus, there is an apparent tendency in the non-optimized projection with e_g bonding and anti-bonding orbitals to prefer one ordering pattern over the other. But the question remains whether the orbital optimization procedure can predict the difference between FeTe and e.g. LaFeAsO. Namely, if the zigzag ordering pattern in FeTe reflects the monoclinic distortion, then a parallel ordering pattern should be predicted for LaFeAsO, reflecting the orthorhombic distortion. I find that this is indeed the case, though LaFeAsO and the pnictides in general pose a special challenge. I have illustrated the effects of optimization in both compounds by focusing on the hole pocket at the Γ point in Figs. 4.4 and 4.5.

In FeTe, I was able to use a $\mathbf{k}$ -independent energy window and perform a binary search algorithm by hand to find the optimal energy window. For the bonding orbital I keep the lower bound on the energy window fixed at ≈ -2.5 eV, but I lower the upper bound by some amount. If the Wannier functions are properly orthonormalized, i.e. there are no linear dependencies in the set, then I lower the upper bound again. However, if the projection fails due to linear dependencies *anywhere* in the first Brillouin zone, meaning that I cannot find the $S_{\nu\nu',\sigma}(\mathbf{k})$ matrix in Eq. (1.106) for some $\mathbf{k}$, then I raise the upper bound by half the amount I lowered it. I proceed in this way resulting in a binary search that converges on the optimal window. I do the same for both the upper and lower bounds of the xy and xz/yz windows to further aid in disentangling these orbitals from the e_g orbitals.

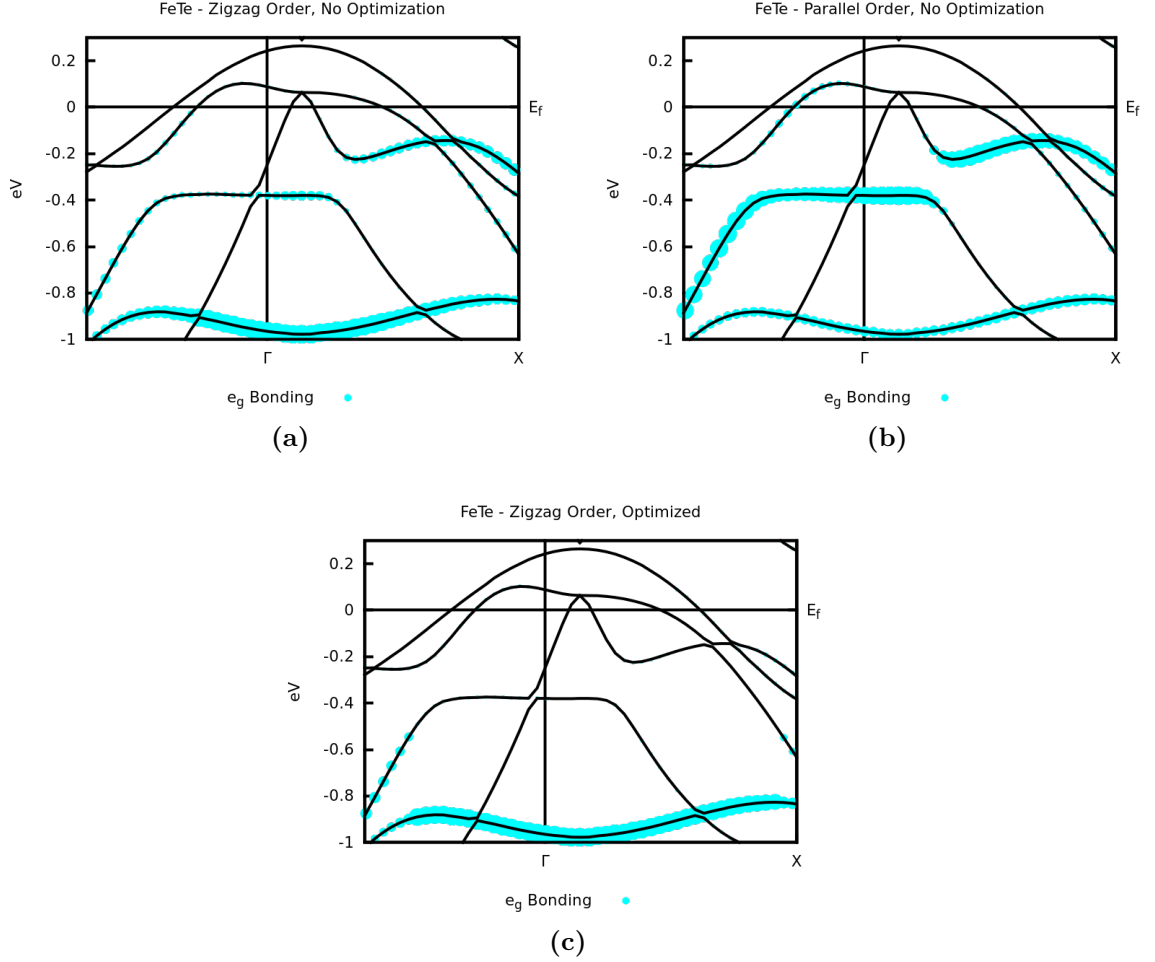


Figure 4.4: Optimization of the hole pocket in FeTe. Shown is the detailed structure of the hole pocket at the Γ point. Visible in (a) and (b) is the differing concentration of orbital content for zigzag and parallel ordering of the e_g bonding orbital, respectively. Since the orbital content for zigzag ordering is lower in energy and the goal of optimization is to disentangle the t_{2g} orbitals, which lie close to the Fermi level, from the e_g orbitals, applying the optimization scheme to the zigzag ordered bonding orbitals is appropriate. Optimization proceeds by moving the upper limit of the energy window used to project the e_g bonding orbitals from the Fermi level, where it lies for the projections shown in (a) and (b), to ≈ -0.5 eV, where it lies in (c). The result is that the higher energy e_g content around ≈ -0.2 eV is removed, disentangling the e_g orbitals from the hole pocket. Though this calculation was performed for a structure appropriate for the paramagnetic, undistorted phase, the preference for a zigzag ordering pattern corresponds to a monoclinic structural distortion observed in FeTe.

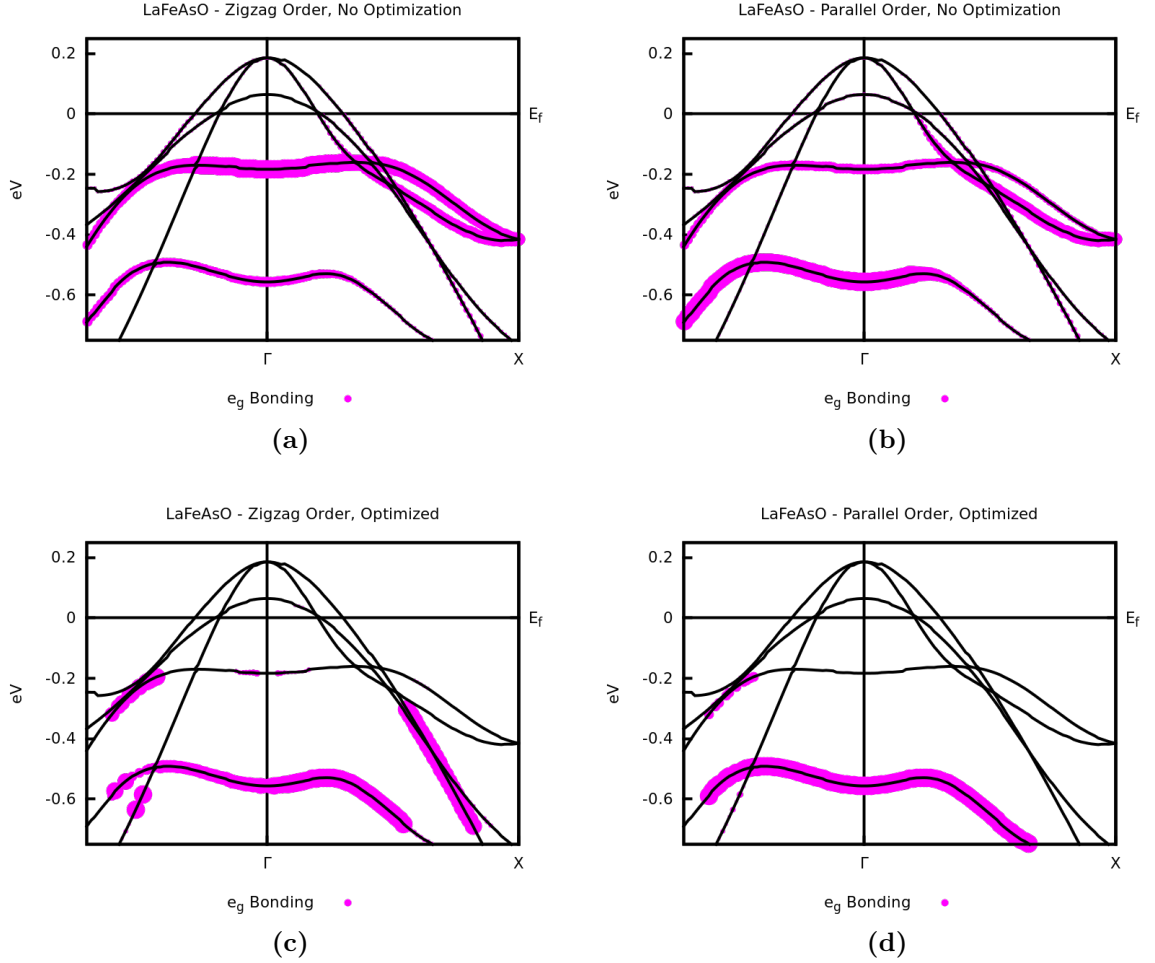


Figure 4.5: Optimization of the hole pocket in LaFeAsO. Shown is the detailed structure of the hole pocket at the Γ point. Visible in (a) and (b) is the differing concentration of orbital content for zigzag and parallel ordering of the e_g bonding orbital, respectively. The energy locations of the content concentration for the two orderings is reversed with respect to FeTe. In LaFeAsO, the parallel ordering pattern shows a greater concentration of orbital content at lower energies, here at ≈ -0.5 eV, while the zigzag ordering pattern shows greater concentration at higher energies, here at ≈ -0.2 eV. Thus, the parallel ordering should be preferred for optimization. In contrast to FeTe, for which the optimization of the energy windows used for projection could be performed independently of $\mathbf{k}$, in LaFeAsO I found it necessary to implement a $\mathbf{k}$ -dependent optimization of the energy window. This is because the e_g content lies higher in energy than in FeTe. The results for zigzag and parallel ordering are shown in (c) and (d), respectively. As expected, the optimization is superior for the parallel ordering, with greater removal of e_g content between -0.2 and -0.4 eV. Though this calculation was performed for a structure appropriate for the paramagnetic, undistorted phase, the preference for a parallel ordering pattern corresponds to a orthorhombic structural distortion observed in LaFeAsO.

If I were to insist on parallel ordering of the e_g orbitals in FeTe, then the orbital content in the hole pocket at Γ prevents me from optimizing the energy window in this way. As shown in Fig. 4.4b, the heavy concentration of the e_g bonding orbitals in bands at ≈ -0.2 and 0.4 eV prevents me from lowering the energy window appreciably and the “optimized” orbitals for parallel ordering are virtually identical to those which are not optimized. On the other hand, for zigzag ordering, the orbital content is heavily concentrated at a band at ≈ -1.0 eV, meaning that I am able to considerably lower the upper bound on the energy window. The result of optimization is shown in Fig. 4.4c. I am able to lower the energy window to ≈ -0.5 eV resulting in the bonding orbital being almost entirely confined to ≈ -1.0 eV and below. I cannot compress the window all the way to the top of this band because of bands elsewhere in the Brillouin zone with slightly higher energy and heavy e_g bonding orbital content. The resulting, zigzag ordered orbitals are shown in real space in Fig. 4.6a.

In LaFeAsO, the binary search optimization procedure using $\mathbf{k}$ -independent energy windows does not work and I am unable to compress the energy window virtually at all for either ordering pattern. I found that the solution was to implement the binary search algorithm to find the optimal window computationally so that it could be performed independently at each point in the Brillouin zone. This in effect produced a $\mathbf{k}$ -dependent energy window. In this way I am able to optimize the orbitals in LaFeAsO. I can then perform the optimization procedure for both ordering patterns and see which one gives a superior result.

In Figs. 4.5a and 4.5b, I show the non-optimized orbital content in the hole pocket at Γ for the zigzag and parallel ordering respectively. The increased content in the lower band for parallel ordering is clearly visible. Upon optimization by implementing $\mathbf{k}$ -dependent energy windows, much e_g content is removed from the pocket in both cases, as shown in Figs. 4.5c and 4.5d. But more is removed for the parallel ordering, as is to be expected from the naturally increased content of the orbitals in that ordering pattern in a lower energy band. Optimization is easier for parallel ordered orbitals because they naturally like to sit at lower energies. They are thus also

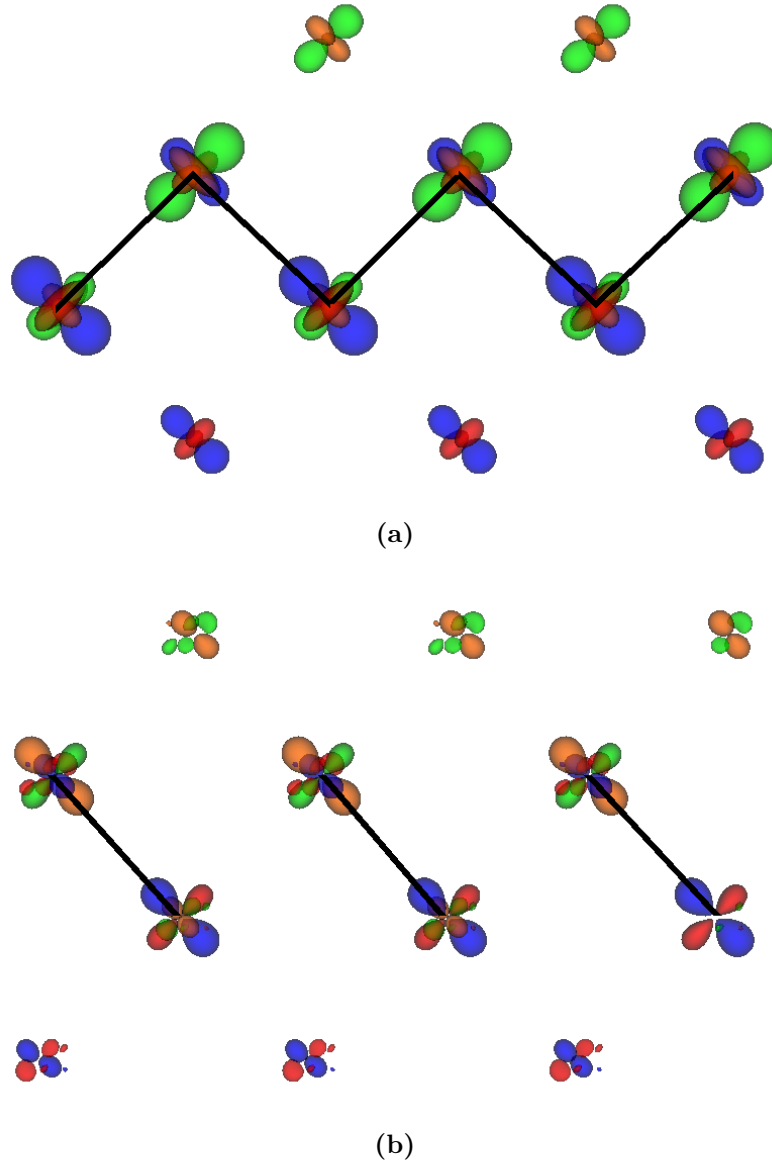


Figure 4.6: Ordering pattern for the e_g bonding orbitals in FeTe and LaFeAsO. In (a) and (b) are several e_g bonding orbitals shown in real space for FeTe and LaFeAsO, respectively. These orbitals were produced after optimization of the energy windows. The orbitals for Fe(1) are shown in blue and red denoting positive and negative and for Fe(2) in green and orange denoting positive and negative, respectively. The zigzag pattern in (a) is characterized by a rotation of the orbitals on the Fe(2) sites by $\pi/2$ with respect to the orbitals on the Fe(1) sites. The resulting zigzag pattern is denoted by black lines to guide the eye. The parallel pattern in (b) is characterized by a lack of rotation, so that the orbitals going from site to site are simply translations of each other.

energetically preferred. The resulting parallel ordered, optimized orbitals are shown in Fig. 4.6b.

Though it is perhaps intuitive that the ordering of bonding orbitals in the tetragonal phase of FeTe and LaFeAsO would lead to the structural distortions that are observed in those materials, there is no obvious connection to the magnetic phase and the ordering pattern there. In all likelihood, it should be the t_{2g} orbitals that control the magnetic ordering since they dominate at the Fermi level. However, a connection has been proposed between the structural phase transition and the magnetic transition [166, 196]. Thus, if the ordering in the e_g orbitals drives the structural transition and this in turn drives the magnetic transition, then this would be a connection between the ordered orbitals and the magnetic structure. Also, the ordered solution for the e_g orbitals are reflected in the details of the tails of the t_{2g} orbitals as a result of the requirement that the orbitals be orthogonal to each other. The importance of such tails to magnetism in the pnictides has been shown in the theoretical work of Lee et al. [172] and I suggest that a similar mechanism could be at work in FeTe.

4.3 The optimal Wannier basis

Via the correct choice of ordering pattern for the e_g bonding and anti-bonding orbitals and subsequent optimization of the energy windows for all the orbitals using the binary search algorithm, I am able to accomplish elimination of hybridization between the e_g and t_{2g} orbitals. The band structure and DOS for the resulting Wannier orbitals are shown in Fig. 4.7. In Fig. 4.7a, it is clear that each group of orbitals occupy different subspaces of Bloch bands. The e_g bonding orbitals lie in two bands between ≈ -2.2 and -1.0 eV that become degenerate along the X-M, X-A and M-A directions. There are two bands because there are two Fe atoms in the unit cell and one bonding orbital per atom. The e_g anti-bonding orbitals lie in two bands between ≈ 0.7 and 2.5 eV with the same degeneracies as the bands that hold the bonding

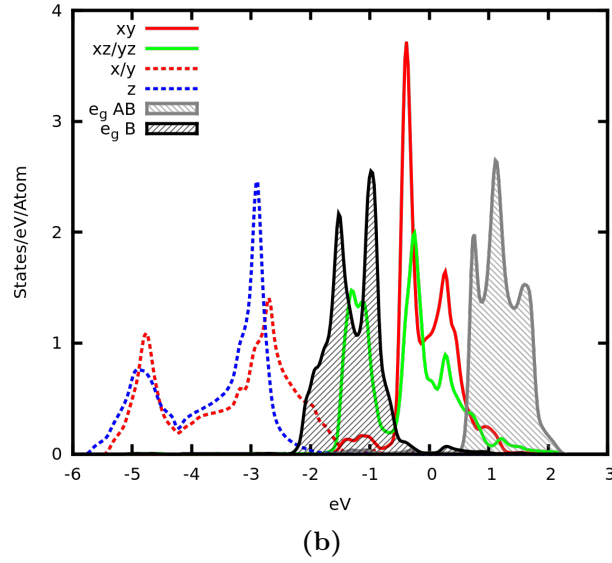
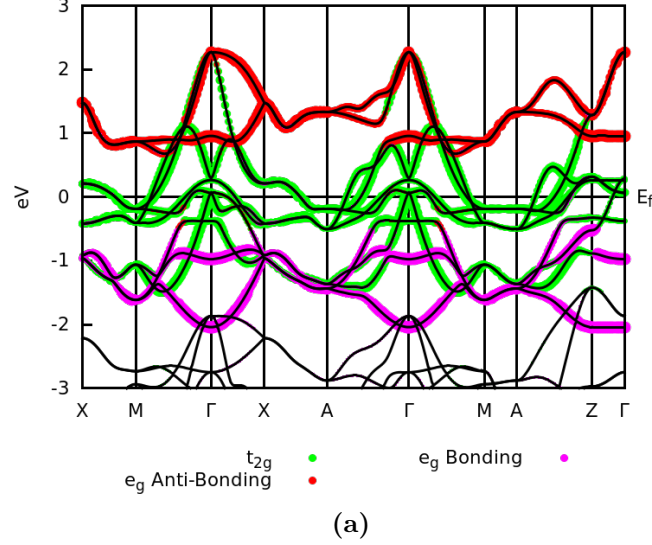


Figure 4.7: Band structure and DOS of the optimal Wannier basis for FeTe. In (a) is the band structure with orbital character upon optimization of the energy windows in order to eliminate as much as possible hybridization between e_g and t_{2g} orbitals. The energy window for the e_g bonding orbitals runs from ≈ -2.2 to -0.5 eV and the resulting orbitals are confined almost entirely to a complex of two bands between ≈ -2.2 to -1.0 eV. The energy window for the e_g anti-bonding orbitals runs from ≈ 0.7 to 2.5 eV with the resulting orbitals almost entirely confined to two bands in that window. The t_{2g} orbital primarily sit in the remaining bands, dominate around the Fermi level and are almost completely disentangled from the e_g orbitals. In (b) I display the DOS for the same set of orbitals. The bonding and anti-bonding orbitals are made to be almost entirely occupied and empty, respectively. The t_{2g} band between ≈ -1.0 to 1.0 eV is seen to dominate the Fermi level. The xy orbital in particular is heavily confined to a narrow band about the Fermi level.

orbitals. That the band complexes for the bonding and anti-bonding orbitals have degeneracies in the same places in the first Brillouin zone is expected as these orbitals have the same symmetry properties.

The t_{2g} orbitals rest in bands that are mostly concentrated in a 2 eV window centered on the Fermi level, but these bands extend almost the entire width of the d band. This means that the bands that contain the t_{2g} orbitals cross those that contain the e_g orbitals, but these are generally unavoided crossings and so do not represent hybridization. Most crucially, the t_{2g} orbitals completely dominate the hole pocket at Γ and the electron pocket at M, which are critical aspects of the electronic structure for explaining low energy phenomena like superconductivity.

The elimination of hybridization can also be monitored by comparing the DOS for the non-optimized orbitals in Fig. 4.2 and the DOS for the optimized orbitals in Fig. 4.7b. Additional peaks appear in the DOS of the e_g orbitals, the most striking of which at ≈ -1.0 eV. This peak contained clear evidence of hybridization between t_{2g} and e_g in the DOS for the non-optimized orbitals and the transfer of this peak entirely to the e_g bonding orbital is a clear indication of the elimination of hybridization. A similar situation occurred for the z^2 bonding and anti-bonding orbitals in Cu_2O , as I discussed in Chapter 2. In both cases, this is a reflection of the hybridization moving from energy space into real space. The transfer of weight in a peak from t_{2g} to e_g upon optimization reflects the increased t_{2g} content in the tails of the e_g orbitals in real space.

The chemical nature of the e_g orbitals is most immediately ascertained by examining them in real space. The orbitals in real space in a cut through the x - y -plane are shown in Fig. 4.8. The axes of the Fe square lattice are rotated 45° with respect to the horizontal and vertical axes of the figures. Whether the orbital is bonding or anti-bonding is obvious when looking at the structure of the nodes, which are shown in the black dashed line. The anti-bonding orbital in Fig. 4.8a shows a clear node between the central Fe site and the nearest neighbor Fe sites. The locations of the nodes correspond to a change in sign from negative (red) on the Fe neighbor in

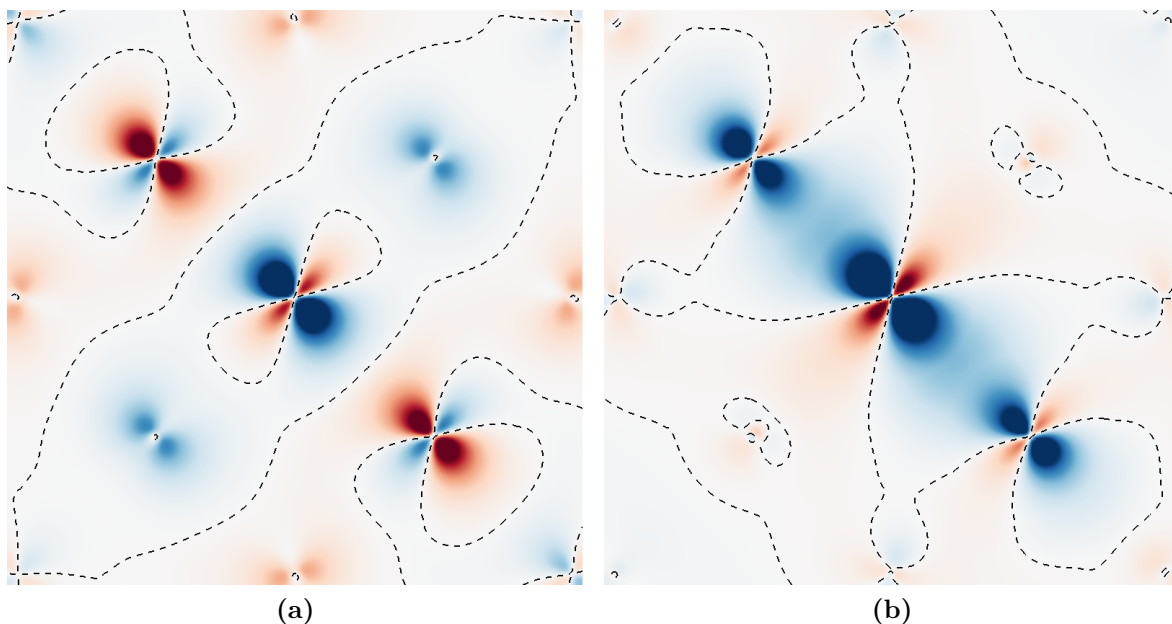


Figure 4.8: Bonding and anti-bonding e_g orbitals for FeTe. The orbitally resolved DOS of FeTe displays a splitting of the e_g orbitals into peaks above and below the Fermi level, similar to the two-peak structure of the z^2 DOS in Cu_2O . By splitting these orbitals into independent e_g anti-bonding, shown in (a) and e_g bonding, shown in (b), I am able to achieve set of Wannier orbitals that provide an optimal description of the optics. The directional preference of the orbitals that causes greater concentration on two of the Fe neighbors is not the result of a bias in the linear combination of local orbitals used as a trial molecular orbital for the projection. Rather, I used an unbiased projection that treated all neighbors the same and the directional preference is a result of the chemistry of FeTe as reflected in the KS band structure.

the upper left to positive (blue) on the central Fe and back to negative for the Fe neighbor in the bottom right. In contrast, the bonding orbital in Fig. 4.8b does not display any such node and the sign of the orbital going through the center and the Fe neighbors in the upper left and lower right does not change, remaining positive.

The directionality of the orbitals is also strikingly apparent. Both bonding and anti-bonding orbitals are strongly concentrated on the Fe nearest neighbor sites in the upper left and lower right, but have very small absolute value on the neighbors in the upper right and lower left. The orbitals shown are those associated for Fe(1). Those for Fe(2) look exactly the same, only rotated by 90° , reflecting the zigzag ordering of the orbitals. It is also possible to project bonding and anti-bonding orbitals that do not display any directionality or ordering. These would have the same absolute value located on all four of the Fe nearest neighbor sites. However, these orbitals cannot be optimized. Thus, the orbitals shown in Fig. 4.8 are unique once optimization with respect to optics is chosen as a criterion.

The t_{2g} orbitals for the optimized Wannier basis are shown in Fig. 4.9 in an isosurface plot. By comparing to the e_g orbitals in Fig. 4.8, one can clearly see that the chemical nature of the two groups of orbitals are qualitatively distinct. Whereas the e_g orbitals display very strong offsite hybridizations caused by their nature as bonding and anti-bonding molecular orbitals, the t_{2g} orbitals are almost entirely contained within a short range of the central Fe site. These orbitals are not participating in any cation-cation bonding.

However, that is not to say that the offsite hybridizations that are present are not important. The requirement that these orbitals be orthogonal to the e_g orbitals means that the chemical nature of the e_g orbitals are manifested somewhat in the t_{2g} orbitals. There are significant tails on the Fe next-nearest neighbors in the case of the xy orbital in Fig. 4.9a and on both Fe nearest and next-nearest neighbors in the case of the xz/yz orbital in Fig. 4.9b. These are the tails that play such an important role in determining the magnetic structure in the theory of Lee et al. [172].

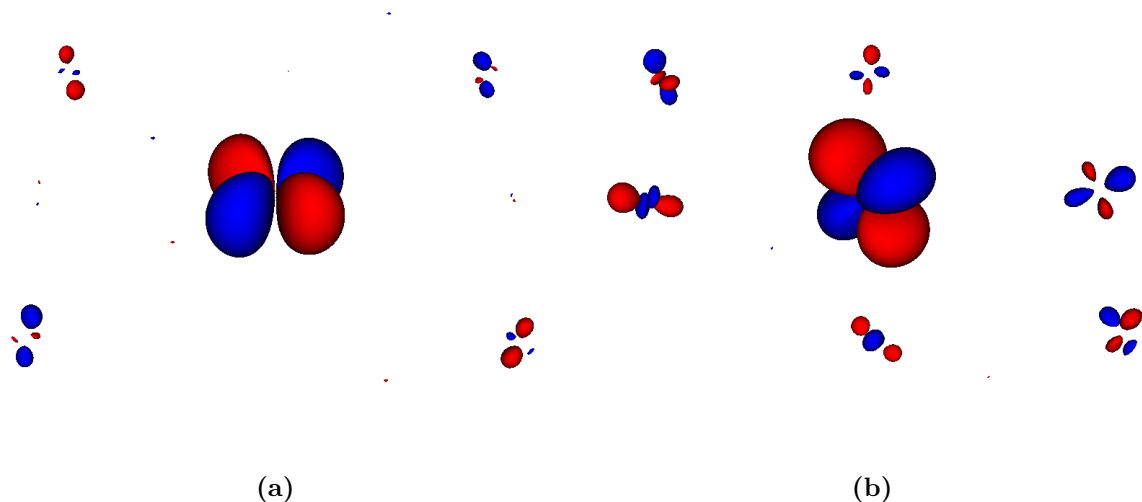


Figure 4.9: t_{2g} orbitals for FeTe. Shown are the t_{2g} orbitals in FeTe: in (a) the xy orbital and in (b) an xz/yz orbital. Both orbitals display a qualitatively different chemistry from that found in the e_g orbitals. There is no significant bonding chemistry apparent and though there is non-negligible off-site hybridization, it is not to the degree of the e_g orbitals. Thus, chemically the t_{2g} and e_g orbitals are distinct subsets. The xy orbital displays the least off-site hybridization and is the most localized. This corresponds to it being a more strongly correlated orbital than the xz/yz orbitals, which display a far larger degree of off-site hybridization and delocalization. The coordinate system that I used to construct the orbitals had the x - and y -axes pointed along the Fe next-nearest neighbor direction so that the density of the xz/yz orbitals would have maxima pointed towards the Te atoms. Thus, the xz and yz trial orbitals used to construct the orbitals are rotated 45° from those in the coordinate system in which the x - and y -axes point along the Fe nearest neighbor direction. The nomenclature I use is for the latter coordinate system to align with the most popular convention in the literature, so that the density of the xy orbital shown in (a) has maxima pointed towards the Fe next-nearest neighbors.

In addition, though they are not visible in Fig. 4.9, there are longer ranged tails that decay slowly with increasing distance. I showed in Chapter 3 that it is precisely these tails that give rise to the long-ranged transitions that make up the optical spectrum (see Fig. 3.7). So, though the t_{2g} orbitals do not display as dramatic of a departure from the d orbitals used as trial orbitals that give rise to the Wannier orbitals, the fact that the t_{2g} orbitals display offsite hybridizations due to the crystal environment is critical to the physics of FeTe and the iron-based compounds in general. I believe that this will be general feature for complex transition-metal compounds.

In the preceding discussion, I optimized with the goal of finding the Wannier basis that provided the simplest description of the optical spectrum, but I did so without presenting any spectra decomposed into Wannier particle-hole channels. Instead, I focused on removing hybridization and disentangling the t_{2g} orbitals from the e_g orbitals. This was based on the insights that the t_{2g} orbitals dominated near the Fermi level even in the single-site projection, that the single-particle states near the Fermi level will dominate in the low-energy optical excitations and, as we learned in the case study of Cu_2O in Chapter 2, elimination of hybridization between the subclasses of orbitals is essential to achieving an optimal description of the optical spectrum. Thus, even without calculating the response in the Wannier basis, it was possible for me to optimize the Wannier orbitals.

Even so, I would be remiss if I did not confirm this heuristic approach to the optimization of the orbital set. To this end, I have calculated the response due to particle-hole channels involving the t_{2g} orbitals. However, there is a complication to calculating the response using the Wannier basis that was irrelevant for Cu_2O . One significant difference between the orbitals basis in Cu_2O and that in FeTe is that in the former all of the orbitals were either fully occupied or fully empty in contrast to the latter in which some of the orbitals are partially occupied. This presents a challenge for decomposition of the optical signal into orbital channels because matrix elements that make no contribution to a calculation in the Bloch basis, because they

would represent particle-particle or hole-hole transitions, can make a contribution to a particular Wannier channel. These spurious channels tend to overwhelm the true channels in any decomposition that naively extends the approach used in Cu_2O . The solution is to find a form for χ^{KS} that excludes these spurious channels from the matrix elements. Just such a form is given by using the quantities given in Eqs. (1.142) and (1.143). I stress that these equations refer only to χ^{KS} , which is appropriate in the optical limit and in the absence of crystal local fields. It is an open question what approach should be used for finite $\mathbf{q}$.

With the expressions appropriate for a metallic system, I am able to calculate the contribution to the optical spectrum from transitions within the t_{2g} subspace in each of the three projections that I have presented: the single-site projection, the projection with the e_g orbitals split into independent bonding and anti-bonding orbitals and the projection in which I optimized the energy windows. I examined the contributions from particle-holes separated by all distances by analytically summing over lattice vectors, i.e. using the expression in Eq. (1.134). The resulting spectra are shown in Fig. 4.10. The first optical feature is a peak centered at ≈ 0.7 eV. Of course, this peak does not reflect the experimental reality, as discussed in detail in Chapter 3. However, it is the lowest energy feature in the GGA reference groundstate that I am using to form a low-energy model and thus is the most relevant part of the spectrum.

The t_{2g} orbitals in the single-site projection make a significant contribution to this peak, as shown by the blue curve in Fig. 4.10. This is in line with the expectation gleaned from examining the DOS (see Fig. 4.1). It can account for roughly half of the total intensity of the peak. This is exactly like the result for Cu_2O shown in Fig. 2.10. There the transitions involving the single-site projected z^2 were also able to account for roughly half of the total intensity of the low-energy features. The conclusion for FeTe is the same for Cu_2O as well. The t_{2g} orbitals are clearly the critical degrees of freedom for explaining the low-energy optics, but the orbital basis is not optimized

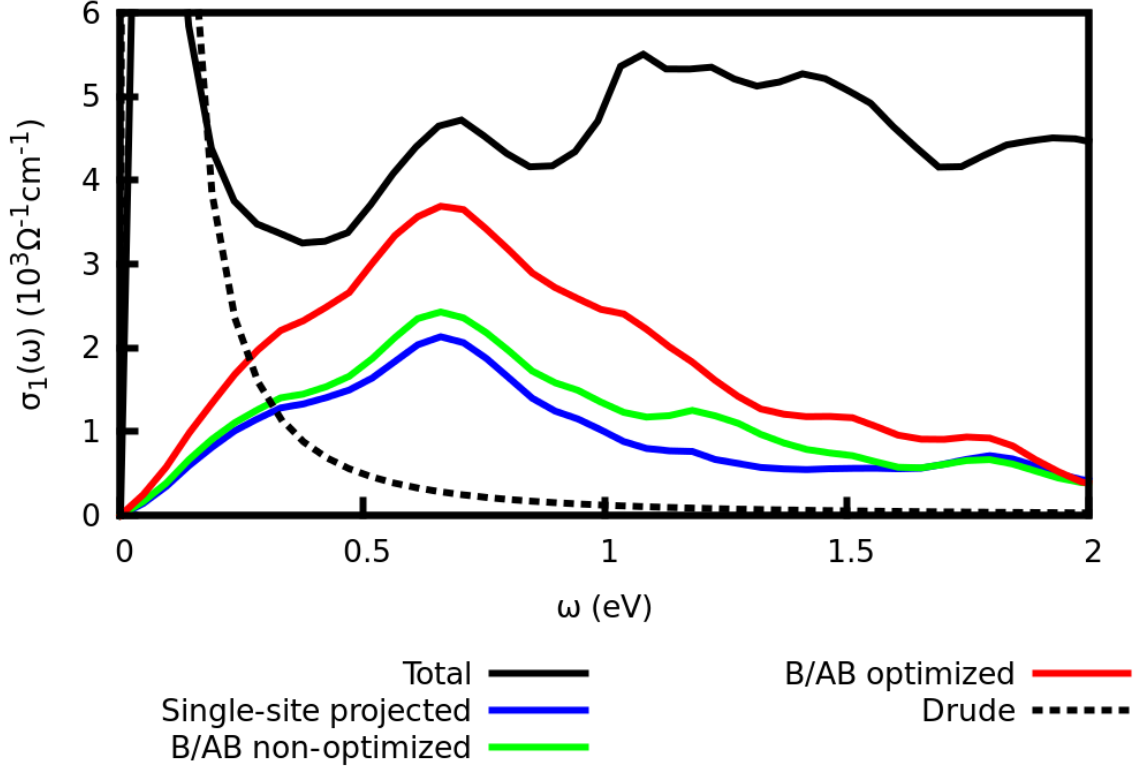


Figure 4.10: Optimization of the Wannier basis in FeTe monitored by optics. Exactly as I have done for Cu_2O , I am able to evaluate the degree of optimization of the orbital basis by observing which construction gives the simplest picture of the low-energy optical spectrum. Shown is the total optical conductivity calculated in the Bloch basis in black along with the contribution from the t_{2g} orbitals calculated using the Wannier basis for various constructions using the GGA groundstate. In all response calculations in the Wannier basis, I summed over lattice vectors analytically. Just as in Cu_2O , the orbitals constructed using the conventional, single-site projected approach, shown in blue, accounts for about half of the total weight of the first prominent peak. Simply splitting the e_g orbitals into bonding and anti-bonding orbitals without optimizing the energy windows results in the green curve, which is only a marginal improvement over the single-site projection. This is because there is still significant hybridization between t_{2g} and e_g orbitals, especially near the Fermi level. When this hybridization is removed by optimizing the energy windows to confine the e_g orbitals away from the Fermi level, the resulting spectrum shown in the red curve shows a marked improvement, almost doubling the weight due to transitions among the t_{2g} orbitals.

and does not yet represent the appropriate one to use to gain the simplest description for the optical spectrum.

The first step towards optimization that I performed was to split the e_g orbitals into independent bonding and anti-bonding orbitals. However, when this construction is checked against the ≈ 0.7 eV peak in the optical conductivity, I do not find much improvement over the orbitals constructed using single-site projection, as shown by the green curve in Fig. 4.10. The t_{2g} orbitals in both cases account for about half of the peak. There is still significant hybridization between the t_{2g} and bonding/anti-bonding orbitals and this accounts for the lack of improvement.

By removing the hybridization and disentangling the t_{2g} orbitals from the e_g orbitals, I am able to find that the description of the ≈ 0.7 eV peak is greatly improved, as shown by the red curve in Fig. 4.10. Though the splitting of the e_g orbitals into independent bonding and anti-bonding parts was necessary to this construction, it was not sufficient to achieving an optimal description of the optical spectrum. Only by eliminating the hybridization, especially the e_g content near the Fermi level, am I able to find the optimal set of Wannier orbitals. Also, that the optical calculation using these orbitals confirms this conclusion means that I can optimize orbitals by targeting hybridization without the need to calculate the optical response along the way. This will be of considerable practical utility in extending this method to other compounds.

The minimal description of the optical spectrum achieved using the optimized t_{2g} orbitals is analogous to the role played by the z^2 anti-bonding orbital in Cu_2O which I presented in Chapter 2. However, there is an important difference in using optics as a guide to optimization in Cu_2O and in FeTe. In the presentation of Cu_2O , the TDDFT (RPA) calculation of the optical spectrum agreed very well with experiment. Since the calculation corresponded closely to reality, I could be confident that the underlying chemistry and dynamics of the orbital degrees of freedom reflected some truth about the material.

In contrast, The insufficiency of optical spectra based on GGA band structure is starkly apparent in the case of FeTe. I showed decisively in Chapter 3 that the GGA electronic structure produces a qualitatively incorrect optical spectrum. There the GGA+ U proved sufficient to include the necessary many-body effects on the band structure and bring the calculation in line with experiment. This raises the question of whether it was appropriate to use the GGA groundstate electronic structure to construct the optimal basis in the first place. Why not then use the GGA+ U band structure as a starting point? This may be possible, but this raises an issue of double-counting. If I use the GGA+ U band structure, I have already included significant many-body effects beyond GGA and I must have some way of subtracting the effect of these contributions in a subsequent many-body calculation.

Instead, it is simpler and more consistent to use the GGA band structure as a starting point. The model that I produce from this reference electronic structure will eventually be used to include many-body effects that should correct the discrepancies between the calculation and experiment. In this light, the work in Chapter 3 could be seen as an analog for what would be done with a model constructed from the GGA groundstate. The GGA+ U would be akin to a Hartree-Fock solution to the model constructed with Wannier orbitals where only the onsite interactions are included. The construction of an *ab initio* model has considerably more flexibility than GGA+ U though. In particular, many-body corrections such as the GW and the ability to investigate the role of offsite and long-ranged interactions are in reach.

In addition to establishing the optimized t_{2g} orbitals as providing a superior description of the optics, by calculating the response in a local orbital basis I am able to establish the range of separation between particle and hole in the transitions that make a significant contribution to the optical signal. The results for FeTe corresponding to the GGA groundstate are shown in Fig. 4.11. In both Cu₂O and FeTe, I find that transitions involving particles and holes separated by up to a few nanometers must be included in order to saturate the total spectral weight. This may be unsurprising when considering a semiconductor with delocalized empty states,

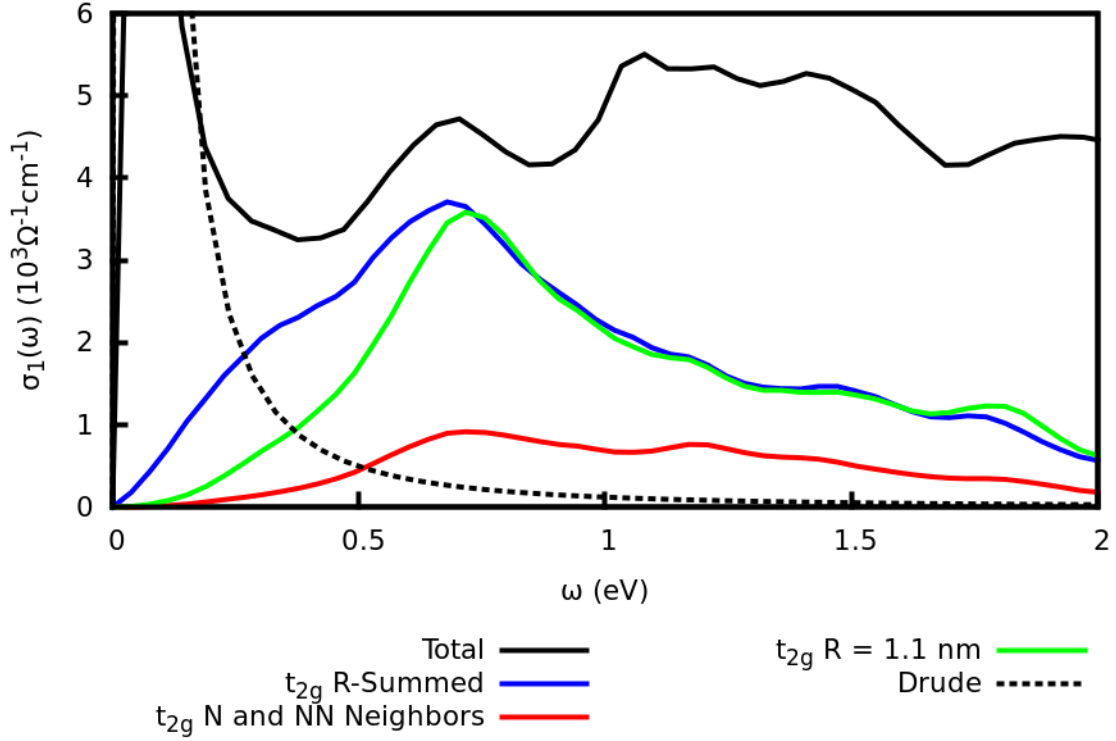


Figure 4.11: The importance of long-ranged particle-hole pairs on the optical spectrum of FeTe. Using the calculation of response in the Wannier basis, I am able to establish the size of the particle-holes that generate the low-energy optical peak in the spectrum calculated using the GGA groundstate. Just as in Cu_2O , using only particle-holes separated by nearest or next-nearest neighbors is insufficient to account for the low-energy optics. The spectrum due to particle-hole transitions with separation of nearest and next-nearest neighbors among the t_{2g} using the optimized projection is shown in red. Clearly, the result cannot even qualitatively describe the low-energy peak. In order to get good agreement on the high energy side of the peak, particle-holes separated by ≈ 1 nm must be included, as shown by the green curve. However, even then, the low energy side of the peak is not well reproduced and only by analytically summing over all separations is the peak properly described, indicating relevant charge fluctuations at a beyond-nanometer length scale. Though the optical spectrum derived from the GGA groundstate does not correspond well to experiment, these long-ranged charge fluctuations will be in any microscopic model using the GGA as a reference system and so should properly be considered.

such as Cu_2O . However, considering that FeTe and the other Fe-based compounds are frequently modeled in an explicitly local picture, with short-ranged hopping and onsite interactions, my results provide a stark reminder of the vital role of long-ranged as well as local physics in these compounds. In light of the need to include such long-ranged particle-hole transitions, I have presented results in which the sum over distance between particle and hole has been performed analytically.

4.4 Comparison of *ab initio* models

Since I have optimized the orbital set, I now have the freedom to calculate the interaction parameters using the CRPA method for a minimal model containing only the t_{2g} orbitals in addition to the d model, which contains only the Fe d orbitals in a site-centered projection, and the d - p model, which contains both Fe d and Te p orbitals. Both the d and d - p models have been investigated elsewhere using CRPA [106], but I will present my own results here to discuss the trends in the onsite interactions as we proceed from d - p , to d , to t_{2g} models. I contend that the t_{2g} model is appropriate to consider both because of the general interest in such a minimal model [189, 190] and because the optimized t_{2g} orbitals can account for the lowest energy optical feature and low-energy band structure.

I will first consider the onsite direct and exchange interaction matrices for the d - p model, $U_{\nu\nu'}$ and $J_{\nu\nu'}$ respectively, which are shown in Tables 4.1 and 4.2. The Fe d matrices are in the former and the Te p matrices are in the latter. The intra-orbital $U_{\nu\nu}$ is approximately the same for all orbitals for both Fe d and Te p with $U_{\nu\nu} \approx 6.0$ eV for the Fe d orbitals and $U_{\nu\nu} \approx 3.8$ eV for the Te p orbitals. This would be expected if both the orbitals and the constrained screened interaction W^c obeyed approximate spherical symmetry. Certainly, since the energy window for the d - p model is the largest the orbitals will be the closest to obeying spherical symmetry.

However, spherical symmetry also implies that the inter-orbital direct interaction $U_{\nu\nu'} = U'$ for all $\nu \neq \nu'$ and the exchange interaction $J_{\nu\nu'} = J$ for all $\nu \neq \nu'$,

Table 4.1: Fe d onsite direct (U) and exchange (J) Coulomb parameters in eV for the d - p model.

$U_{\nu\nu'}$	xy	yz	xz	x^2-y^2	z^2	$J_{\nu\nu'}$	xy	yz	xz	x^2-y^2	z^2
xy	6.1	5.0	5.0	5.6	4.8	xy	***	0.5	0.5	0.2	0.7
yz	5.0	6.0	5.0	4.9	5.4	yz	0.5	***	0.5	0.5	0.3
xz	5.0	5.0	6.0	4.9	5.4	xz	0.5	0.5	***	0.5	0.3
x^2-y^2	5.6	4.9	4.9	5.9	4.7	x^2-y^2	0.2	0.5	0.5	***	0.6
z^2	4.8	5.4	5.4	4.7	6.0	z^2	0.7	0.3	0.3	0.6	***

Table 4.2: Te p onsite direct (U) and exchange (J) Coulomb parameters in eV for the d - p model.

$U_{\nu\nu'}$	x	y	z	$J_{\nu\nu'}$	x	y	z
x	3.8	3.0	3.0	x	***	0.4	0.4
y	3.0	3.8	3.0	y	0.4	***	0.4
z	3.0	3.0	3.6	z	0.4	0.4	***

in addition to $U_{\nu\nu} = U$. Though this is the case for the Te p orbitals, with $U_{\nu\nu} = 3.0$ eV and $J_{\nu\nu} = 0.4$ eV, the inter-orbital interactions for the Fe d orbitals range from 4.7 to 5.6 eV and the exchange interactions range from 0.2 to 0.7 eV. Thus, even in the d - p model where spherical symmetry would be the most respected, there is considerable asphericity. This is important for dynamical mean-field theory methods as they most often assume that the interactions obey spherical symmetry [32].

In order to perform the calculation of both the d and t_{2g} models, I utilized the expressions for response in the Wannier basis, extending our spectroscopic decomposition methods to the CRPA downfolding procedure. This is done by splitting the polarization into that due to Wannier channels involving only target orbitals with the remainder comprised of the rest of the Wannier channels. I do this by first constructing a Wannier basis that includes all d and p orbitals. These degrees of freedom as a whole are disentangled from the rest. The rest due to all of the bands outside this d - p subspace constitutes one contribution to the polarization entering into the constrained screened interaction (see Eq. (1.119)).

Table 4.3: Onsite direct (U) and exchange (J) Coulomb parameters in eV for the d model.

$U_{\nu\nu'}$	xy	yz	xz	x^2-y^2	z^2	$J_{\nu\nu'}$	xy	yz	xz	x^2-y^2	z^2
xy	3.6	2.3	2.3	2.9	2.1	xy	***	0.5	0.5	0.3	0.7
yz	2.3	3.1	2.0	2.2	2.5	yz	0.5	***	0.5	0.5	0.3
xz	2.3	2.0	3.1	2.2	2.5	xz	0.5	0.5	***	0.5	0.3
x^2-y^2	2.9	2.2	2.2	3.3	2.0	x^2-y^2	0.3	0.5	0.5	***	0.6
z^2	2.1	2.5	2.5	2.0	3.2	z^2	0.7	0.3	0.3	0.6	***

For the contribution due to orbitals other than the t_{2g} orbitals, I calculate the KS response resolved as a sum over Wannier indices given by Eq. (1.128), but I exclude λ and λ' that have all indices in the t_{2g} subspace. This constitutes the second contribution to the polarization that enters into Eq. (1.119). However, I stress that the t_{2g} orbitals are almost entirely entangled already, so the interaction parameters calculated by disentanglement of the t_{2g} bands [108] should not differ dramatically. For the d model, I use the exact same procedure only the λ and λ' indices that contain all d orbitals are excluded. I must do this because of the small remaining hybridization between p and d at the bottom of the d band.

The interaction matrices for the d model are shown in Table 4.3. Since the p orbitals are now included as screening channels, the elements of the $U_{\nu\nu'}$ matrix are significantly reduced. The average intra-orbital interaction is ≈ 3.3 eV versus ≈ 6.0 eV when the p orbitals are not included in the screening. However, the exchange interactions are virtually unchanged. In general, screening has far less of an effect on exchange parameters when compared to direct parameters. The asphericity of the direct interactions has also increased, with the intra-orbital interaction ranging from 3.1 eV for the xz and yz orbitals to 3.6 eV for the xy orbital. This is to be expected as a narrower energy window will drive the orbital further away from atomic-like due to increased hybridization in real space. Since the intra-orbital interaction for the xy orbital is the largest, we start to see the important orbital differentiation in FeTe in the interactions.

Table 4.4: Onsite direct (U) and exchange (J) Coulomb parameters in eV for the t_{2g} model.

$U_{\nu\nu'}$	xy	yz	xz	$J_{\nu\nu'}$	xy	yz	xz
xy	1.1	0.4	0.4	xy	***	0.3	0.3
yz	0.4	0.6	0.3	yz	0.3	***	0.2
xz	0.4	0.3	0.6	xz	0.3	0.2	***

However, the orbital differentiation becomes more dramatic for the t_{2g} model, the interaction matrices for which are shown in Table 4.4. Here, the intra-orbital interaction for the xy orbital is 1.1 eV versus 0.6 eV for the xz and yz orbitals, thus differing by nearly a factor of two. Thus, the t_{2g} model exhibits orbital differentiation more explicitly, though it must exist implicitly in the d and $d-p$ models. The e_g orbitals also provide a significant amount of screening, reducing the direct interactions by about a factor of 3 and, in contrast to the effect of the Te p orbitals going from the $d-p$ to d model, reducing the exchange interactions by about a factor of 2. Still, in going from the $d-p$ to t_{2g} model the exchange interactions have overall gained in significance when compared to the direct interactions. This may be seen as consistent with the proposed picture of FeTe as a Hund metal [176].

I have not presented the particulars of the hoppings for the microscopic models for which I have described the interactions. This is partially because I have found that achieving the full bandwidth of the optimized t_{2g} orbitals requires the inclusion of hoppings at a distance that is on the order with the size of the macro-crystal, as is apparent in Fig. 4.12. There it is clear that long-ranged hopping on the order of a few nanometers is critical to faithfully reproducing the GGA band structure. Thus presenting, for example, only the nearest or next-nearest neighbor hoppings would be deceptive as to the importance of these over longer-ranged hoppings. Instead, I would suggest using the full non-interacting reference Green's function in a future solver for the models. I comment also that these long-ranged hoppings are consistent with the requirement that optical transitions involving distances between particle and hole of over a nanometer be included in the calculation of the optical response, as indicated

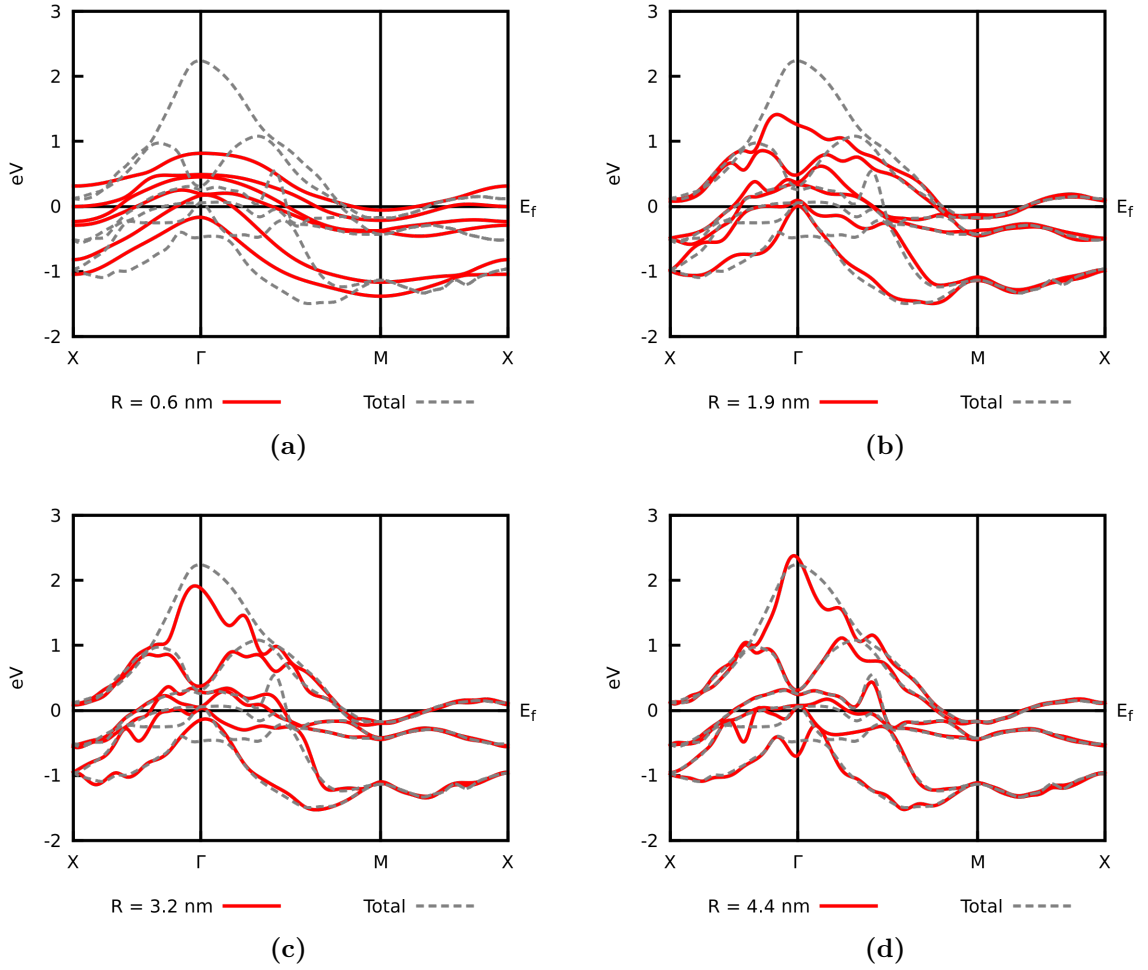


Figure 4.12: Evolution of the band structure with hopping for the t_{2g} model. The hopping parameters define the single-particle part of the Hamiltonian to which the interactions are added. These parameters are usually assumed to fall off rapidly with distance and when they are presented only the values for a few shells of neighbors are given. However, I find that when these parameters are calculated *ab initio* using the Wannier basis using Eq. (1.124), long-ranged hopping must be included. Shown are the band structures generated when the hopping is cut off between an orbital at the origin and orbitals at a radius greater than 0.6 nm in (a), 1.9 nm in (b), 3.2 nm in (c) and 4.4 nm in (d) together with the band structure generated when all hopping is included. In all plots I calculated the eigenvalues only at $\mathbf{k}$ -points in the original $24 \times 24 \times 16$ mesh and interpolated the points between with a cubic spline. Clearly, quantitative agreement is not even achieved with the largest value of the cutoff radius and even for qualitative agreement hopping for distances larger than 2 nm must be included. The largest cutoff radius is close to the boundary of the macro-crystal and the errors generated when the hopping is cutoff here likely means that a denser $\mathbf{k}$ -point mesh would be preferred.

in Fig. 4.11. Both of these facts point to the itinerant and long-ranged nature of the t_{2g} -like subspace.

Here, I feel that it is necessary to make a few comments regarding the convergence of the CRPA calculations. Given the large numerical costs of the calculations, the results I have presented are at the frontier of convergence. Thus, there is a non-negligible margin of error associated with the values I have presented. There are three important convergence parameters that must be considered. I will discuss each in detail below, but here I will state that I believe all the parameters that I have presented to be converged to within a few percent. Since the level systematic error involved in the CRPA approximation is expected to be $\sim 10\%$, I consider this to be a sufficient level of convergence.

The first important convergence parameter is the maximum reciprocal lattice vector, $G_{\max}$, for which $W_{\mathbf{G}\mathbf{G}'}^c$ is calculated. For inter-orbital direct interactions and exchange interactions, the calculation converges relatively quickly, with $\approx 1,000$ $\mathbf{G}$ vectors, corresponding to $G_{\max} \approx 4.5 a_0^{-1}$ for FeTe, being sufficient to converge to a level of less than 1%. The offsite interactions converge even more rapidly. However, the intra-orbital $U_{\nu\nu}$ tends to converge much more slowly, with $> 3,000$ $\mathbf{G}$ vectors being required to converge the d orbitals to within a few percent. Before full convergence $U_{\nu\nu}$ is a monotonically increasing function of $G_{\max}$, meaning the fully converged values are somewhat higher than values I have presented. The increase of $U_{\nu\nu}$ with $G_{\max}$ has an intuitive physical meaning. The interaction of an orbital with itself will have a large contribution from short range screening. Small values of $G_{\max}$ represent more long-ranged screening, which will always be more effective than screening at short range. As length scales become smaller the screened interaction will approach the bare interaction. Thus, since increasing $G_{\max}$ corresponds to a decreasing length scale, the significant contribution from short-ranged screening in $U_{\nu\nu}$ will increase, raising the total value. For inter-orbital interactions, and even more for offsite interactions, short-ranged screening is less important and so convergence with $G_{\max}$ is faster.

The second important convergence parameter is the number of interband transitions that are included in the calculation of $\chi^{\text{KS},r}$. This is represented by the energy window containing the band that are paired to enumerate interband transitions. For FeTe, this parameter was easier to converge than G_{max} , with a energy window ≈ 50 eV above and below the Fermi level being sufficient to converge to within a few percent. However, FeTe is a small system, containing only four atoms per unit cell. For a larger system, there will be a greater number of bands for the same energy window. If there were 8 atoms per unit cell the number of bands in the same energy window would roughly double. Since interband transitions are enumerated over pairs of bands, the number of transitions would quadruple. Clearly, larger systems will represent a challenge for future developments of CRPA. The trend with increasing band energy window for the interactions is the opposite of that with G_{max} : the interactions tend to decrease with increasing energy window. This is also intuitive, since more interband transitions contribute more screening channels, which will suppress the interactions.

The last convergence parameter is the number of $\mathbf{k}$ points used to sample the first Brillouin zone. This remains the parameter I have investigated the least. Intuitively, the onsite parameters should converge rapidly with the size of the $\mathbf{k}$ -mesh since this corresponds to the size of the macro-crystal, i.e. the integration volume. The integrand for onsite parameters will drop off the most rapidly with distance, so only a small integration volume is required. I have found that the difference between a $6 \times 6 \times 4$ and $9 \times 9 \times 6$ mesh to be within a few percent, though ideally I would like to look at larger $\mathbf{k}$ -meshes in order to establish a trend. Unfortunately, as of the time of writing, this remains too numerically difficult to do since the problem scales like the number of $\mathbf{k}$ points squared. Offsite parameters require a larger $\mathbf{k}$ -mesh since the orbitals are now separated by some distance. I have found this to be numerically true in preliminary investigations and this remains a topic for future research.

Also, though I have only presented the onsite interactions in detail, one can show that the offsite direct interactions decrease like $1/R$ as distance between sites become large. This behavior has also be confirmed numerically [105]. This is due to the

constrained, static Kohn-Sham response used in CRPA vanishing like q^2 in the optical limit. A $1/R$ tail in the interactions represents a different manifestation of long-ranged physics than that in the hoppings and optical matrix elements and any solver for the Hamiltonian we present should be able to handle these interactions. One could perhaps argue that a small value for this $1/R$ tail relative to the onsite interaction would justify ignoring it. However, the nearest-neighbor interaction for the optimized t_{2g} -like orbitals is 0.2 eV, which is on the same order as the onsite interactions.

Finally, I would like to point out that, in contrast to the orbitals used in the construction of many models, the t_{2g} Wannier orbitals still depart from atomic-like orbitals, as visible in Fig. 4.9. Lobes directed towards the Te sites of the xz and yz orbitals are reduced in size, all the orbitals display significant tails on neighboring Fe sites and they do not drop to zero with increasing distance from the center as rapidly as atomic-like orbitals. The details of the orbitals in the neighborhood of neighboring sites account for the long-ranged hoppings and optical matrix elements discussed above and these details can also have significant bearing on magnetism. These and the previously discussed characteristics of the optimized Hamiltonian for the *ab initio* electronic structure and should be considered in any calculation using the Hamiltonian.

The procedure I demonstrate here provides a route to *ab initio* models that have a minimal number of degrees of freedom, which is of vital importance as the difficulty of any many-body calculation increases dramatically as the number of degrees of freedom increases. However, I note that the value for the nearest-neighbor, direct interactions are ≈ 0.2 eV, comparable to the onsite parameters presented in Table 4.4. It is likely that methods that are able to go beyond Hubbard-like models will be required to fully investigate the models that I present here. I suggest that these models are potentially well suited to Green's function methods such as the *GW*. Such methods can potentially treat a model with considerably more dense $\mathbf{k}$ point meshes (large macro-crystal size) and long-ranged interactions than the methods that have been

developed for use with models such as the Hubbard model that consider only local interactions and small system sizes.

Conclusion

In this dissertation, I have used *ab initio* methods together with the Wannier local orbital basis to study the role of orbital degrees of freedom on spectroscopy. The ability to calculate the contributions to optical spectra in particular due to Wannier orbital particle-hole excitations has allowed me to develop a method to find an optimized construction of the Wannier basis. This optimization has as its criterion the simplest description of the optical spectrum in terms of Wannier particle-hole channels. This allows for an unbiased construction of orbitals tied to a vitally important, ubiquitously experimentally measured physically observable. This means that the orbitals are validated by the real world physics that they can describe rather than the physics that one wishes to produce. With the Wannier basis constructed, I was able to calculate the parameters to be used in models, the form of which is driven by the realistic, material dependent local symmetry and chemistry as well as long-ranged solid-state phenomena. This theoretical approach relies on the unification of several techniques and theoretical tools: Kohn-Sham DFT, TDDFT linear response, the Wannier transformation from the Bloch states and CRPA downfolding.

The eigenvalues and eigenfunctions that define Kohn-Sham groundstate provide a way to calculate particle-hole excitation spectra via TDDFT. Using linear density response theory within TDDFT I have been able to calculate particle-hole excitation spectra, with particular success in the case of the optical spectrum of Cu_2O . By using the GGA+ U groundstate for FeTe, I have also found good agreement with experiment in a linear response method that can be seen to be closely related to strict TDDFT. By

focusing on the optical spectrum as measured by ellipsometry and reflectivity, I have been able to investigate the role of local orbital degrees of freedom on the particle-hole excitation spectrum without the complicating effects of long-ranged screening by examining the Kohn-Sham density response as opposed to the true density response.

The transformation from the Kohn-Sham Bloch states to the Wannier functions provides a way to define localized functions that still contain all of the information on chemistry and symmetry embodied in the Kohn-Sham groundstate. This is because it is possible to perform the transformation such that it is unitary. I have used the Wannier basis to study the chemistry of the groundstate by combining analysis of the functions in real space and analysis of their DOS in energy space. I also presented an expression for the Kohn-Sham density response as a sum over pairs of Wannier functions which represent particle-holes in real space. Using this expression I have been able to decompose the optical spectrum into contributions from different classes of channels. This allowed me to optimize the gauge for the Wannier transformation to yield a basis that provided the simplest description of the optical signal.

Given a Wannier basis and using TDDFT linear response, it is possible to calculate screened interaction parameters for a target subspace using the CRPA method. I have applied CRPA to a target space defined by the optimized Wannier basis in order to construct a minimal model for the low-energy physics. I have found that the optimization of the Wannier basis with respect to optics amounts to the elimination of hybridization between subclasses of orbitals defined by their chemical properties. This is precisely what is required to disentangle the target subspace from the rest of the degrees of freedom so that I am able to downfold to a well defined target. Thus there is a close correspondence between optimizing orbitals with respect to optical spectroscopy and downfolding to minimal models.

Cu₂O provided the first test case for the development of my optimization method. I was aided by the remarkable agreement that the TDDFT (RPA) spectrum using the GGA groundstate had with ellipsometry experiments. This allowed me to investigate the impact of local orbital degrees of freedom on the optical spectrum with the

assurance that the conclusions were tightly bound to reality. I found that the best choices for the parameters that enter into the Wannier transformation were those that reflected the local symmetry and chemistry. By constructing a Wannier basis in this way, I was able to identify a single z^2 orbital per Cu site as the critical hole for the low-energy (few eV) optical spectrum. Finding an optimal basis which could simply describe the optics required the construction of molecular orbitals that embodied the cation-cation bonding chemistry which stabilizes the structure of Cu_2O . Using this set, I find that particle-hole transitions involving a z^2 anti-bonding orbital could account for the entire low-energy spectrum. In addition, I had to consider nanometer-sized particle-holes in order to represent the spectrum. I developed a way to analytically sum over lattice vectors in order to account for these large particle-holes.

Since Cu_2O is a relatively weakly correlated semiconductor, it is not the type of system that would be most appropriate for constructing a low-energy model using Wannier local orbitals. However, FeTe is a parent compound for a class of iron-based superconductors and is thought to be strongly correlated. So, it is a system for which models based on local orbitals is appropriate. However, precisely because of correlation effects on the groundstate, the GGA electronic structure yields a qualitatively incorrect optical spectrum. Thus, I found it necessary to conduct a further investigation into the optical response. I found that inclusion of correlations due to onsite interactions via GGA+ U brought the calculated spectrum into qualitative agreement with experiment. Remarkably, the effect of these correlations was orbitally selective, with the xy and z^2 orbitals being strongly effected and the xz , yz and x^2-y^2 being hardly effected. Also, as in Cu_2O , I found that nanometer-sized particle-holes were required to reproduce the spectrum. In addition, I found that the effect of orbitally selective renormalization of the band structure was the emergence of a two-dimensional band topology that gives rise to a graphene-like optical response.

I then used the insight gained from the study of Cu_2O that the key to optimization was the elimination of hybridization in energy space, transferring it into real space,

to optimize the orbitals in FeTe with respect to the GGA band structure. I chose to use the GGA groundstate since this would be the appropriate non-interacting reference system for a many-body calculation. I found that optimization of the Wannier basis required splitting the z^2 and x^2-y^2 orbitals into bonding/anti-bonding orbitals, recapitulating the theme of cation-cation bonding that was so important for Cu_2O . Combination of the z^2 and x^2-y^2 orbitals and the creation of bonding/anti-bonding orbitals required that I consider two ordering patterns for the orbitals, zigzag and parallel. I found that the optimization scheme I developed was able to select one ordering pattern over the other and moreover this preferred pattern corresponded to the observed structural and magnetic phase transition in both FeTe and LaFeAsO.

Through optimization I was able to produce a minimal, three-orbital model for the xy , xz and yz orbitals that could reproduce the calculated, low-energy optical spectrum. I found evidence for strong orbital selectivity in the onsite interactions, with the xy orbital having twice the intra-orbital interaction strength as the xz and yz orbitals. I also found that screening reduced the direct interactions from those in the d - p and d models much more than the exchange interactions. I argue that this corresponds to the view of FeTe as a Hund metal. Though the onsite interactions are clearly important, I also found that long-ranged physics, especially in considering the single-particle hopping, is a crucial aspect to the model.

Throughout this dissertation, I have found that access to a Wannier basis that is optimized to give maximal insight into the role of local orbital degrees of freedom on spectroscopy is crucial to understanding the role of local chemistry and symmetry on the two-particle excitation spectrum. Of particular importance is the interplay of physics on the atomic length scale, including local correlations in the case of FeTe, and physics on the nanometer length scale, which is critical to understanding the optical excitations. I have been able to reveal a complex relationship between the atom and the solid that may have broad implications on the theoretical treatment and intuitive understanding of complex transition metal compounds.

Bibliography

- [1] Wannier, G. H. The Structure of Electronic Excitation Levels in Insulating Crystals. *Phys. Rev.* **52**, 191–197 (1937). [2](#), [45](#)
- [2] Anisimov, V. I. *et al.* Full orbital calculation scheme for materials with strongly correlated electrons. *Phys. Rev. B* **71**, 125119 (2005). [5](#), [7](#), [45](#), [49](#), [154](#)
- [3] Marzari, N., Mostofi, A. A., Yates, J. R., Souza, I. & Vanderbilt, D. Maximally localized Wannier functions: Theory and applications. *Rev. Mod. Phys.* **84**, 1419–1475 (2012). [2](#), [46](#), [79](#)
- [4] Runge, E. & Gross, E. K. U. Density-Functional Theory for Time-Dependent Systems. *Phys. Rev. Lett.* **52**, 997–1000 (1984). [2](#), [5](#), [34](#)
- [5] Ullrich, C. A. *Time-dependent density-functional theory: concepts and applications* (Oxford University Press, New York, 2012). [2](#), [5](#), [13](#), [43](#), [242](#)
- [6] Hohenberg, P. & Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **136**, B864–B871 (1964). [2](#), [15](#), [16](#), [18](#)
- [7] Kohn, W. & Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **140**, A1133–A1138 (1965). [2](#), [4](#), [15](#), [18](#), [232](#)
- [8] Springer, M. & Aryasetiawan, F. Frequency-dependent screened interaction in Ni within the random-phase approximation. *Phys. Rev. B* **57**, 4364–4368 (1998). [3](#), [56](#)
- [9] Kotani, T. Ab initio random-phase-approximation calculation of the frequency-dependent effective interaction between 3d electrons: Ni, Fe, and MnO. *Journal of Physics: Condensed Matter* **12**, 2413 (2000).
- [10] Aryasetiawan, F. *et al.* Frequency-dependent local interactions and low-energy effective models from electronic structure calculations. *Phys. Rev. B* **70**, 195104 (2004). [3](#), [56](#), [58](#)

- [11] Ambrosch-Draxl, C. Augmented Planewave Methods. *Physica Scripta* **2004**, 48 (2004). [4](#), [26](#), [30](#)
- [12] Singh, D. & Nordstrom, L. *Planewaves, Pseudopotentials, and the LAPW Method* (Springer US, New York, 2006), 2nd edn. [5](#), [26](#), [27](#), [32](#)
- [13] Exciting Plus FP-LAPW Code. <https://github.com/toxa81/exciting-plus>. [5](#), [31](#), [32](#), [48](#), [53](#), [63](#), [241](#)
- [14] Gross, E. K. U. & Kohn, W. Local density-functional theory of frequency-dependent linear response. *Phys. Rev. Lett.* **55**, 2850–2852 (1985). [5](#), [37](#)
- [15] Anisimov, V. I., Aryasetiawan, F. & Lichtenstein, A. I. First-principles calculations of the electronic structure and spectra of strongly correlated systems: the LDA+U method. *Journal of Physics: Condensed Matter* **9**, 767 (1997). [6](#), [74](#)
- [16] Zuo, J. M., Kim, M., O’Keeffe, M. & Spence, J. C. H. Direct observation of d-orbital holes and Cu–Cu bonding in Cu₂O. *Nature* **401**, 49–52 (1999). [6](#), [8](#), [81](#), [83](#), [84](#), [97](#), [99](#)
- [17] Ito, T., Kawashima, T., Yamaguchi, H., Masumi, T. & Adachi, S. Optical properties of Cu₂O studied by spectroscopic ellipsometry. *Journal of the Physical Society of Japan* **67**, 2125–2131 (1998). [7](#), [86](#), [87](#)
- [18] Haidu, F. *et al.* Dielectric function and magneto-optical Voigt constant of Cu₂O: A combined spectroscopic ellipsometry and polar magneto-optical Kerr spectroscopy study. *Phys. Rev. B* **84**, 195203 (2011).
- [19] Meyer, B. K. *et al.* The Physics of Copper Oxide (Cu₂O). In Svensson, B. G., Pearton, S. J. & Jagadish, C. (eds.) *Oxide semiconductors*, 201 – 226 (Academic Press, San Diego, 2013). [7](#), [81](#), [85](#), [86](#)

- [20] Hancock, J. N. *et al.* Strong coupling to magnetic fluctuations in the charge dynamics of iron-based superconductors. *Phys. Rev. B* **82**, 014523 (2010). [8](#), [9](#), [119](#), [121](#), [123](#), [129](#), [131](#), [132](#)
- [21] Dai, Y. M. *et al.* Spectral weight transfer in strongly correlated $\text{Fe}_{1.03}\text{Te}$. *Phys. Rev. B* **90**, 121114 (2014). [8](#), [9](#), [119](#), [121](#), [123](#), [129](#), [131](#), [132](#)
- [22] Zhang, Y. *et al.* Strong correlations and spin-density-wave phase induced by a massive spectral weight redistribution in $\alpha\text{-Fe}_{1.06}\text{Te}$. *Phys. Rev. B* **82**, 165113 (2010). [8](#), [122](#), [124](#), [125](#)
- [23] Liu, Z. K. *et al.* Measurement of Coherent Polarons in the Strongly Coupled Antiferromagnetically Ordered Iron-Chalcogenide $\text{Fe}_{1.02}\text{Te}$ using Angle-Resolved Photoemission Spectroscopy. *Phys. Rev. Lett.* **110**, 037003 (2013). [125](#), [135](#), [137](#), [138](#)
- [24] Lin, P.-H. *et al.* Nature of the Bad Metallic Behavior of $\text{Fe}_{1.06}\text{Te}$ Inferred from Its Evolution in the Magnetic State. *Phys. Rev. Lett.* **111**, 217002 (2013). [124](#), [125](#), [126](#), [135](#), [137](#), [138](#), [149](#), [155](#)
- [25] Ieki, E. *et al.* Evolution from incoherent to coherent electronic states and its implications for superconductivity in $\text{FeTe}_{1-x}\text{Se}_x$. *Phys. Rev. B* **89**, 140506 (2014). [125](#)
- [26] Ambolode, L. C. C. *et al.* Dependence of electron correlation strength in $\text{Fe}_{1+y}\text{Te}_{1-x}\text{Se}_x$ on Se content. *Phys. Rev. B* **92**, 035104 (2015). [8](#), [119](#), [124](#), [125](#), [126](#), [155](#)
- [27] McClure, J. W. Band Structure of Graphite and de Haas-van Alphen Effect. *Phys. Rev.* **108**, 612–618 (1957). [9](#), [138](#)
- [28] Slonczewski, J. C. & Weiss, P. R. Band Structure of Graphite. *Phys. Rev.* **109**, 272–279 (1958).

- [29] Kuzmenko, A. B., van Heumen, E., Carbone, F. & van der Marel, D. Universal Optical Conductance of Graphite. *Phys. Rev. Lett.* **100**, 117401 (2008). [9](#), [138](#), [140](#)
- [30] Bao, W. *et al.* Tunable $(\delta\pi, \delta\pi)$ -Type Antiferromagnetic Order in α -Fe(Te,Se) Superconductors. *Phys. Rev. Lett.* **102**, 247001 (2009). [10](#), [119](#), [127](#)
- [31] Haule, K. & Kotliar, G. Coherenceincoherence crossover in the normal state of iron oxypnictides and importance of Hund’s rule coupling. *New Journal of Physics* **11**, 025021 (2009). [11](#), [128](#), [155](#)
- [32] Georges, A., de’ Medici, L. & Mravlje, J. Strong Correlations from Hunds Coupling. *Annual Review of Condensed Matter Physics* **4**, 137–178 (2013). [11](#), [128](#), [155](#), [183](#)
- [33] Neudert, R. *et al.* Manifestation of Spin-Charge Separation in the Dynamic Dielectric Response of One-Dimensional Sr_2CuO_3 . *Phys. Rev. Lett.* **81**, 657–660 (1998). [11](#), [220](#), [227](#)
- [34] Moskvina, A. S. *et al.* Evidence for Two Types of Low-Energy Charge Transfer Excitations in Sr_2CuO_3 . *Phys. Rev. Lett.* **91**, 037001 (2003). [11](#), [220](#), [226](#)
- [35] Thomas, L. H. The calculation of atomic fields. *Mathematical Proceedings of the Cambridge Philosophical Society* **23**, 542–548 (1927). [15](#)
- [36] Fermi, E. Eine statistische methode zur bestimmung einiger eigenschaften des atoms und ihre anwendung auf die theorie des periodischen systems der elemente. *Zeitschrift für Physik* **48**, 73–79 (1928). [15](#)
- [37] Rayleigh, J. In finding the correction for the open end of an organ-pipe. *Phil. Trans.* **161**, 77 (1870). [16](#)
- [38] Ritz, W. Über eine neue Methode zur Lösung gewisser Variationsprobleme der mathematischen Physik. *J. reine angew. Math.* **135**, 1–61 (1908). [16](#)

- [39] Levy, M. Universal variational functionals of electron densities, first-order density matrices, and natural spin-orbitals and solution of the v-representability problem. *Proceedings of the National Academy of Sciences* **76**, 6062–6065 (1979). [17](#)
- [40] Lieb, E. H. Density functionals for Coulomb systems. *International Journal of Quantum Chemistry* **24**, 243–277 (1983). [17](#)
- [41] Kohn, W., Savin, A. & Ullrich, C. A. Hohenberg-Kohn theory including spin magnetism and magnetic fields. *International Journal of Quantum Chemistry* **100**, 20–21 (2004). [23](#)
- [42] Capelle, K., Ullrich, C. A. & Vignale, G. Degenerate ground states and nonunique potentials: Breakdown and restoration of density functionals. *Phys. Rev. A* **76**, 012508 (2007). [23](#), [234](#), [235](#)
- [43] Sakurai, J. J. *Modern Quantum Mechanics* (Addison Wesley, 1993), revised edn. [24](#)
- [44] Ashcroft, N. W. & Mermin, N. D. *Solid State Physics*. Science: Physics (Saunders College, 1976). [24](#)
- [45] Phillips, J. C. & Kleinman, L. New Method for Calculating Wave Functions in Crystals and Molecules. *Phys. Rev.* **116**, 287–294 (1959). [26](#)
- [46] Hamann, D. R., Schlüter, M. & Chiang, C. Norm-Conserving Pseudopotentials. *Phys. Rev. Lett.* **43**, 1494–1497 (1979).
- [47] Bachelet, G. B., Hamann, D. R. & Schlüter, M. Pseudopotentials that work: From H to Pu. *Phys. Rev. B* **26**, 4199–4228 (1982).
- [48] Kleinman, L. & Bylander, D. M. Efficacious Form for Model Pseudopotentials. *Phys. Rev. Lett.* **48**, 1425–1428 (1982).

- [49] Vanderbilt, D. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B* **41**, 7892–7895 (1990).
- [50] Pickett, W. E. Pseudopotential methods in condensed matter applications. *Computer Physics Reports* **9**, 115 – 197 (1989). [26](#)
- [51] Skriver, H. L. *The LMTO Method: Muffin-Tin Orbitals and Electronic Structure*, vol. 41 of *Springer Series in Solid-State Sciences* (Springer-Verlag, Berlin, 1984). [26](#)
- [52] Methfessel, M., Rodriguez, C. O. & Andersen, O. K. Fast full-potential calculations with a converged basis of atom-centered linear muffin-tin orbitals: Structural and dynamic properties of silicon. *Phys. Rev. B* **40**, 2009–2012 (1989).
- [53] Methfessel, M. Elastic constants and phonon frequencies of Si calculated by a fast full-potential linear-muffin-tin-orbital method. *Phys. Rev. B* **38**, 1537–1540 (1988).
- [54] Wills, J. M., Eriksson, O., Alouani, M. & Price, D. L. *Electronic Structure and Physical Properties of Solids: The Uses of the LMTO Method Lectures of a Workshop Held at Mont Saint Odile, France, October 2–5, 1998*, chap. Full-Potential LMTO Total Energy and Force Calculations, 148–167 (Springer Berlin Heidelberg, Berlin, Heidelberg, 2000). [26](#)
- [55] Slater, J. C. & Koster, G. F. Simplified LCAO Method for the Periodic Potential Problem. *Phys. Rev.* **94**, 1498–1524 (1954). [26](#)
- [56] Evarestov, R. A. *Quantum Chemistry of Solids: The LCAO First Principles Treatment of Crystals*, vol. 153 of *Springer Series in Solid-State Sciences* (Springer-Verlag, Berlin, 2007).
- [57] Goringe, C., Bowler, D. & Hernandez, E. Tight-binding modelling of materials. *Reports on Progress in Physics* **60**, 1447–1512 (1997). [26](#)

- [58] Soler, J. M. *et al.* The SIESTA method for ab initio order-N materials simulation. *Journal of Physics: Condensed Matter* **14**, 2745 (2002). [26](#)
- [59] Andersen, O. K. Linear methods in band theory. *Phys. Rev. B* **12**, 3060–3083 (1975). [26](#), [27](#), [30](#)
- [60] Koelling, D. D. & Arbman, G. O. Use of energy derivative of the radial solution in an augmented plane wave method: application to copper. *Journal of Physics F: Metal Physics* **5**, 2041 (1975). [26](#), [27](#)
- [61] Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **50**, 17953–17979 (1994). [26](#)
- [62] Kresse, G. & Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **59**, 1758–1775 (1999). [26](#)
- [63] Slater, J. C. Wave Functions in a Periodic Potential. *Phys. Rev.* **51**, 846–851 (1937). [27](#)
- [64] Slater, J. Energy band calculations by the augmented plane wave method. vol. 1 of *Advances in Quantum Chemistry*, 35 – 58 (Academic Press, 1964). [27](#)
- [65] Sjöstedt, E., Nordström, L. & Singh, D. An alternative way of linearizing the augmented plane-wave method. *Solid State Communications* **114**, 15 – 20 (2000). [30](#)
- [66] Li, C., Freeman, A. J., Jansen, H. J. F. & Fu, C. L. Magnetic anisotropy in low-dimensional ferromagnetic systems: Fe monolayers on Ag(001), Au(001), and Pd(001) substrates. *Phys. Rev. B* **42**, 5433–5442 (1990). [31](#)
- [67] Schlosser, H. & Marcus, P. M. Composite Wave Variational Method for Solution of the Energy-Band Problem in Solids. *Phys. Rev.* **131**, 2529–2546 (1963). [32](#)

- [68] van Leeuwen, R. Mapping from Densities to Potentials in Time-Dependent Density-Functional Theory. *Phys. Rev. Lett.* **82**, 3863–3866 (1999). [35](#), [242](#), [247](#)
- [69] van Leeuwen, R. Causality and symmetry in time-dependent density-functional theory. *Phys. Rev. Lett.* **80**, 1280–1283 (1998). [36](#)
- [70] Kubo, R. Statistical-Mechanical Theory of Irreversible Processes. I. General Theory and Simple Applications to Magnetic and Conduction Problems. *Journal of the Physical Society of Japan* **12**, 570–586 (1957). [36](#)
- [71] Fetter, A. L. & Walecka, J. D. *Quantum Theory of Many-Particle Systems* (Dover Publications, Mineola, New York, 2003). [36](#), [38](#), [229](#)
- [72] Ando, T. Inter-subband optical absorption in space-charge layers on semiconductor surfaces. *Zeitschrift für Physik B Condensed Matter* **26**, 263–272 (1977). [37](#)
- [73] Ando, T. Inter-subband optical transitions in a surface space-charge layer. *Solid State Communications* **21**, 133 – 136 (1977). [37](#)
- [74] Zangwill, A. & Soven, P. Resonant Photoemission in Barium and Cerium. *Phys. Rev. Lett.* **45**, 204–207 (1980). [37](#)
- [75] Petersilka, M., Gossmann, U. J. & Gross, E. K. U. Excitation Energies from Time-Dependent Density-Functional Theory. *Phys. Rev. Lett.* **76**, 1212–1215 (1996). [37](#)
- [76] Slater, J. C. The Theory of Ferromagnetism: Lowest Energy Levels. *Phys. Rev.* **52**, 198–214 (1937). [45](#)
- [77] Slater, J. C. Electrons in Perturbed Periodic Lattices. *Phys. Rev.* **76**, 1592–1601 (1949). [45](#)

- [78] Löwdin, P. On the nonorthogonality problem connected with the use of atomic wave functions in the theory of molecules and crystals. *The Journal of Chemical Physics* **18**, 365–375 (1950). [45](#), [47](#), [49](#)
- [79] Kohn, W. Analytic Properties of Bloch Waves and Wannier Functions. *Phys. Rev.* **115**, 809–821 (1959). [45](#)
- [80] Des Cloizeaux, J. Orthogonal Orbitals and Generalized Wannier Functions. *Phys. Rev.* **129**, 554–566 (1963). [46](#), [79](#)
- [81] Marzari, N. & Vanderbilt, D. Maximally localized generalized Wannier functions for composite energy bands. *Phys. Rev. B* **56**, 12847–12865 (1997). [46](#), [79](#)
- [82] Posternak, M., Baldereschi, A., Massidda, S. & Marzari, N. Maximally localized Wannier functions in antiferromagnetic MnO within the FLAPW formalism. *Phys. Rev. B* **65**, 184422 (2002). [46](#), [79](#)
- [83] Freimuth, F., Mokrousov, Y., Wortmann, D., Heinze, S. & Blügel, S. Maximally localized Wannier functions within the FLAPW formalism. *Phys. Rev. B* **78**, 035120 (2008).
- [84] Kuneš, J. *et al.* Wien2wannier: From linearized augmented plane waves to maximally localized wannier functions. *Computer Physics Communications* **181**, 1888 – 1895 (2010). [46](#)
- [85] Stephan, U., Martin, R. M. & Drabold, D. A. Extended-range computation of Wannier-like functions in amorphous semiconductors. *Phys. Rev. B* **62**, 6885–6888 (2000). [46](#)
- [86] Ku, W., Rosner, H., Pickett, W. E. & Scalettar, R. T. Insulating Ferromagnetism in $\text{La}_4\text{Ba}_2\text{Cu}_2\text{O}_{10}$: An *Ab Initio* Wannier Function Analysis. *Phys. Rev. Lett.* **89**, 167204 (2002). [46](#), [79](#)

- [87] Lu, W. C., Wang, C. Z., Chan, T. L., Ruedenberg, K. & Ho, K. M. Representation of electronic structures in crystals in terms of highly localized quasiautomatic minimal basis orbitals. *Phys. Rev. B* **70**, 041101 (2004).
- [88] Qian, X. *et al.* Quasiautomatic orbitals for *ab initio* tight-binding analysis. *Phys. Rev. B* **78**, 245112 (2008). [46](#)
- [89] Sporkmann, B. & Bross, H. Calculation of Wannier functions for fcc transition metals by Fourier transformation of Bloch functions. *Phys. Rev. B* **49**, 10869–10876 (1994). [46](#)
- [90] Sporkmann, B. & Bross, H. Calculation of Wannier functions for zinc-blende-type semiconductors. *Journal of Physics: Condensed Matter* **9**, 5593 (1997). URL <http://stacks.iop.org/0953-8984/9/i=26/a=008>.
- [91] Smirnov, V. P. & Usvyat, D. E. Variational method for the generation of localized Wannier functions on the basis of Bloch functions. *Phys. Rev. B* **64**, 245108 (2001). [46](#), [79](#)
- [92] Solovyev, I. V. Combining DFT and many-body methods to understand correlated materials. *Journal of Physics: Condensed Matter* **20**, 293201 (2008). [56](#)
- [93] Gunnarsson, O. Calculation of parameters in model Hamiltonians. *Phys. Rev. B* **41**, 514–518 (1990). [56](#)
- [94] Gunnarsson, O., Andersen, O. K., Jepsen, O. & Zaanen, J. Density-functional calculation of the parameters in the Anderson model: Application to Mn in CdTe. *Phys. Rev. B* **39**, 1708–1722 (1989).
- [95] Anisimov, V. I. & Gunnarsson, O. Density-functional calculation of effective Coulomb interactions in metals. *Phys. Rev. B* **43**, 7570–7574 (1991). [56](#), [74](#)

- [96] Imada, M. & Miyake, T. Electronic Structure Calculation by First Principles for Strongly Correlated Electron Systems. *Journal of the Physical Society of Japan* **79**, 112001 (2010). [56](#)
- [97] Honerkamp, C. Effective interactions in multiband systems from constrained summations. *Phys. Rev. B* **85**, 195129 (2012). [56](#)
- [98] Negele, J. W. & Orland, H. *Quantum Many-Particle Systems* (Westview Press, Boulder, CO, 1998). [57](#)
- [99] Şaşıoğlu, E., Friedrich, C. & Blügel, S. Effective Coulomb interaction in transition metals from constrained random-phase approximation. *Phys. Rev. B* **83**, 121101 (2011). [57](#)
- [100] Aryasetiawan, F., Karlsson, K., Jepsen, O. & Schönberger, U. Calculations of Hubbard U from first-principles. *Phys. Rev. B* **74**, 125106 (2006). [57](#)
- [101] Miyake, T. & Aryasetiawan, F. Screened Coulomb interaction in the maximally localized Wannier basis. *Phys. Rev. B* **77**, 085122 (2008). [57](#)
- [102] Solovyev, I. V. & Imada, M. Screening of Coulomb interactions in transition metals. *Phys. Rev. B* **71**, 045103 (2005). [57](#)
- [103] Arita, R., Kuneš, J., Kozhevnikov, A. V., Eguluz, A. G. & Imada, M. *Ab initio* Studies on the Interplay between Spin-Orbit Interaction and Coulomb Correlation in Sr_2IrO_4 and Ba_2IrO_4 . *Phys. Rev. Lett.* **108**, 086403 (2012). [57](#)
- [104] Wehling, T. O. *et al.* Strength of Effective Coulomb Interactions in Graphene and Graphite. *Phys. Rev. Lett.* **106**, 236805 (2011). [57](#)
- [105] Nakamura, K., Arita, R. & Imada, M. Ab initio Derivation of Low-Energy Model for Iron-Based Superconductors LaFeAsO and LaFePO . *Journal of the Physical Society of Japan* **77**, 093711 (2008). [57](#), [188](#)

- [106] Miyake, T., Nakamura, K., Arita, R. & Imada, M. Comparison of ab initio low-energy models for LaFePO, LaFeAsO, BaFe₂As₂, LiFeAs, FeSe, and FeTe: electron correlation and covalency. *J. Phys. Soc. Japan* **79**, 044705 (2010). [119](#), [131](#), [133](#), [153](#), [182](#)
- [107] Aichhorn, M., Biermann, S., Miyake, T., Georges, A. & Imada, M. Theoretical evidence for strong correlations and incoherent metallic state in FeSe. *Phys. Rev. B* **82**, 064504 (2010). [57](#)
- [108] Miyake, T., Aryasetiawan, F. & Imada, M. *Ab initio* procedure for constructing effective models of correlated materials with entangled band structure. *Phys. Rev. B* **80**, 155134 (2009). [58](#), [184](#)
- [109] Nakamura, K., Yoshimoto, Y., Nohara, Y. & Imada, M. Ab initio Low-Dimensional Physics Opened Up by Dimensional Downfolding: Application to LaFeAsO. *Journal of the Physical Society of Japan* **79**, 123708 (2010). [70](#)
- [110] Nomura, Y. *et al.* Effective on-site interaction for dynamical mean-field theory. *Phys. Rev. B* **86**, 085117 (2012). [70](#), [71](#)
- [111] Zaanen, J., Sawatzky, G. A. & Allen, J. W. Band gaps and electronic structure of transition-metal compounds. *Phys. Rev. Lett.* **55**, 418–421 (1985). [74](#)
- [112] Terakura, K., Oguchi, T., Williams, A. R. & Kübler, J. Band theory of insulating transition-metal monoxides: Band-structure calculations. *Phys. Rev. B* **30**, 4734–4747 (1984). [74](#)
- [113] Sawatzky, G. A. & Allen, J. W. Magnitude and Origin of the Band Gap in NiO. *Phys. Rev. Lett.* **53**, 2339–2342 (1984). [74](#)
- [114] Anisimov, V. I., Zaanen, J. & Andersen, O. K. Band theory and Mott insulators: Hubbard U instead of Stoner I . *Phys. Rev. B* **44**, 943–954 (1991). [75](#)

- [115] Anisimov, V. I., Solovyev, I. V., Korotin, M. A., Czyżyk, M. T. & Sawatzky, G. A. Density-functional theory and NiO photoemission spectra. *Phys. Rev. B* **48**, 16929–16934 (1993). [76](#), [77](#)
- [116] Perdew, J. P., Parr, R. G., Levy, M. & Balduz, J. L. Density-functional theory for fractional particle number: Derivative discontinuities of the energy. *Phys. Rev. Lett.* **49**, 1691–1694 (1982). [76](#), [77](#), [234](#)
- [117] Czyżyk, M. T. & Sawatzky, G. A. Local-density functional and on-site correlations: The electronic structure of La_2CuO_4 and LaCuO_3 . *Phys. Rev. B* **49**, 14211–14228 (1994). [76](#)
- [118] Petukhov, A. G., Mazin, I. I., Chioncel, L. & Lichtenstein, A. I. Correlated metals and the LDA + u method. *Phys. Rev. B* **67**, 153106 (2003). [77](#)
- [119] Souza, I., Marzari, N. & Vanderbilt, D. Maximally localized Wannier functions for entangled energy bands. *Phys. Rev. B* **65**, 035109 (2001). [79](#)
- [120] Ku, W., Berlijn, T. & Lee, C.-C. Unfolding First-Principles Band Structures. *Phys. Rev. Lett.* **104**, 216401 (2010). [79](#)
- [121] Bruneval, F. *et al.* Exchange and correlation effects in electronic excitations of Cu_2O . *Phys. Rev. Lett.* **97**, 267601 (2006). [81](#), [85](#), [88](#)
- [122] Ghijsen, J. *et al.* Electronic structure of Cu_2O and CuO . *Phys. Rev. B* **38**, 11322–11330 (1988). [81](#), [84](#)
- [123] Hu, J. P. *et al.* On-site interband excitations in resonant inelastic x-ray scattering from Cu_2O . *Phys. Rev. B* **77**, 155115 (2008). [81](#), [84](#), [85](#)
- [124] Filippetti, A. & Fiorentini, V. Coexistence of ionic and metallic bonding in noble-metal oxides. *Phys. Rev. B* **72**, 035128 (2005). [81](#), [83](#)

- [125] Kirfel, A. & Eichhorn, K. Accurate structure analysis with synchrotron radiation. The electron density in Al_2O_3 and Cu_2O . *Acta Crystallographica Section A* **46**, 271–284 (1990). [81](#)
- [126] Hafner, S. S. & Nagel, S. The electric field gradient at the position of copper in Cu_2O and electronic charge density analysis by means of k-factors. *Physics and Chemistry of Minerals* **9**, 19–22 (1983). [81](#)
- [127] O’Keeffe, M. Madelung Constants for the C3 and C9 Structures. *J. Chem. Phys.* **38**, 3035–3035 (1963). [83](#)
- [128] Dahl, J. & Switendick, A. Energy bands in cuprous oxide. *Journal of Physics and Chemistry of Solids* **27**, 931 – 942 (1966). [84](#)
- [129] Kleinman, L. & Mednick, K. Self-consistent energy bands of Cu_2O . *Phys. Rev. B* **21**, 1549–1553 (1980).
- [130] Robertson, J. Electronic structure and x-ray near-edge core spectra of Cu_2O . *Phys. Rev. B* **28**, 3378–3385 (1983).
- [131] Marksteiner, P., Blaha, P. & Schwarz, K. Electronic structure and binding mechanism of Cu_2O . *Zeitschrift für Physik B Condensed Matter* **64**, 119–127 (1986). [84](#)
- [132] Elliott, R. J. Intensity of Optical Absorption by Excitons. *Phys. Rev.* **108**, 1384–1389 (1957). [84](#), [85](#)
- [133] Elliott, R. J. Symmetry of Excitons in Cu_2O . *Phys. Rev.* **124**, 340–345 (1961).
- [134] Balkanski, M., Petroff, Y. & Trivich, D. Optical properties of cuprous oxide in the ultra-violet. *Solid State Communications* **5**, 85 – 88 (1967).
- [135] Forman, R., Brower, W. & Parker, H. Phonons and the green exciton series in cuprous oxide, Cu_2O . *Physics Letters A* **36**, 395 – 396 (1971).

- [136] Fröhlich, D., Kenklies, R., Uihlein, C. & Schwab, C. Assignment of the Even-Parity Excitons in Cu_2O . *Phys. Rev. Lett.* **43**, 1260–1263 (1979).
- [137] Baumeister, P. W. Optical Absorption of Cuprous Oxide. *Phys. Rev.* **121**, 359–362 (1961).
- [138] Brahms, S., Nikitine, S. & Dahl, J. On the band structure and the absorption spectrum of Cu_2O . *Physics Letters* **22**, 31 – 33 (1966). [84](#), [85](#), [86](#)
- [139] Nie, X., Wei, S.-H. & Zhang, S. B. First-principles study of transparent *p*-type conductive SrCu_2O_2 and related compounds. *Phys. Rev. B* **65**, 075111 (2002). [85](#)
- [140] Tran, F. & Blaha, P. Implementation of screened hybrid functionals based on the Yukawa potential within the LAPW basis set. *Phys. Rev. B* **83**, 235118 (2011). URL <http://link.aps.org/doi/10.1103/PhysRevB.83.235118>. [85](#)
- [141] Isseroff, L. Y. & Carter, E. A. Importance of reference Hamiltonians containing exact exchange for accurate one-shot GW calculations of Cu_2O . *Phys. Rev. B* **85**, 235142 (2012). [85](#)
- [142] Kotani, T., van Schilfgaarde, M. & Faleev, S. V. Quasiparticle self-consistent GW method: A basis for the independent-particle approximation. *Phys. Rev. B* **76**, 165106 (2007). [85](#), [88](#)
- [143] van Schilfgaarde, M., Kotani, T. & Faleev, S. Quasiparticle Self-Consistent GW Theory. *Phys. Rev. Lett.* **96**, 226402 (2006). [85](#)
- [144] Bruneval, F. *Exchange and Correlation in the Electronic Structure of Solids, from Silicon to Cuprous Oxide: GW Approximation and beyond*. Ph.D. thesis, École Polytechnique (2005). [87](#), [88](#)
- [145] Lany, S. Band-structure calculations for the 3d transition metal oxides in GW. *Phys. Rev. B* **87**, 085112 (2013). [86](#), [87](#)

- [146] Baym, G. & Kadanoff, L. P. Conservation Laws and Correlation Functions. *Phys. Rev.* **124**, 287–299 (1961). [88](#), [256](#), [257](#)
- [147] Stewart, G. R. Superconductivity in iron compounds. *Rev. Mod. Phys.* **83**, 1589–1652 (2011). [118](#)
- [148] Li, S. *et al.* First-order magnetic and structural phase transitions in $\text{Fe}_{1+y}\text{Se}_x\text{Te}_{1-x}$. *Phys. Rev. B* **79**, 054503 (2009). [118](#), [127](#)
- [149] Hsu, F.-C. *et al.* Superconductivity in the PbO-type structure α -FeSe. *Proceedings of the National Academy of Sciences* **105**, 14262–14264 (2008). [118](#)
- [150] Margadonna, S. *et al.* Pressure evolution of the low-temperature crystal structure and bonding of the superconductor FeSe ($T_c = 37$ K). *Phys. Rev. B* **80**, 064506 (2009). [118](#)
- [151] Charnukha, A. Optical conductivity of iron-based superconductors. *J. Phys.: Condens. Matter* **26**, 253203 (2014). [119](#), [121](#), [124](#), [125](#), [131](#)
- [152] Rodriguez, E. E. *et al.* Magnetic and structural properties near the Lifshitz point in Fe_{1+x}Te . *Phys. Rev. B* **88**, 165110 (2013). [122](#), [124](#), [125](#), [127](#)
- [153] Liu, T. J. *et al.* Charge-carrier localization induced by excess Fe in the superconductor $\text{Fe}_{1+y}\text{Te}_{1-x}\text{Se}_x$. *Phys. Rev. B* **80**, 174509 (2009).
- [154] Chen, G. F. *et al.* Electronic properties of single-crystalline $\text{Fe}_{1.05}\text{Te}$ and $\text{Fe}_{1.03}\text{Se}_{0.30}\text{Te}_{0.70}$. *Phys. Rev. B* **79**, 140509 (2009).
- [155] Hu, R., Bozin, E. S., Warren, J. B. & Petrovic, C. Superconductivity, magnetism, and stoichiometry of single crystals of $\text{Fe}_{1+y}(\text{Te}_{1-x}\text{S}_x)_z$. *Phys. Rev. B* **80**, 214514 (2009). [122](#), [127](#)
- [156] Dressel, M. & Grüner, G. *Electrodynamics of Solids: Optical Properties of Electrons in Matter* (Cambridge University Press, Cambridge, 2002). [122](#)

- [157] Millis, A. J. Optical conductivity and correlated electron physics. In Baeriswyl, D. & Degiorgi, L. (eds.) *Strong interactions in low dimensions*, vol. 25 of *Physics and Chemistry of Materials with Low-Dimensional Structures*, 195–235 (Springer Netherlands, Dordrecht, Netherlands, 2004). [123](#)
- [158] Xia, Y. *et al.* Fermi Surface Topology and Low-Lying Quasiparticle Dynamics of Parent $\text{Fe}_{1+x}\text{Te/Se}$ Superconductor. *Phys. Rev. Lett.* **103**, 037002 (2009). [124](#), [125](#)
- [159] Parshall, D. *et al.* Competition between commensurate and incommensurate magnetic ordering in Fe_{1+y}Te . *Phys. Rev. B* **85**, 140515 (2012). [125](#), [127](#)
- [160] Mannella, N. *et al.* Polaron coherence condensation as the mechanism for colossal magnetoresistance in layered manganites. *Phys. Rev. B* **76**, 233102 (2007). [125](#)
- [161] de Jong, S. *et al.* Quasiparticles and anomalous temperature dependence of the low-lying states in the colossal magnetoresistant oxide $\text{La}_{2-2x}\text{Sr}_{1+2x}\text{Mn}_2\text{O}_7$ ($x = 0.36$) from angle-resolved photoemission. *Phys. Rev. B* **76**, 235117 (2007). [125](#)
- [162] Tamai, A. *et al.* Strong Electron Correlations in the Normal State of the Iron-Based $\text{FeSe}_{0.42}\text{Te}_{0.58}$ Superconductor Observed by Angle-Resolved Photoemission Spectroscopy. *Phys. Rev. Lett.* **104**, 097002 (2010). [126](#)
- [163] Yi, M. *et al.* Observation of Temperature-Induced Crossover to an Orbital-Selective Mott Phase in $\text{A}_x\text{Fe}_{2-y}\text{Se}_2$ ($\text{A}=\text{K, Rb}$) Superconductors. *Phys. Rev. Lett.* **110**, 067003 (2013). [126](#)
- [164] Dai, P., Hu, J. & Dagotto, E. Magnetism and its microscopic origin in iron-based high-temperature superconductors. *Nature Physics* **8**, 709–718 (2012). [126](#)

- [165] Dai, P. Antiferromagnetic order and spin dynamics in iron-based superconductors. *Rev. Mod. Phys.* **87**, 855–896 (2015). [126](#)
- [166] Fobes, D. *et al.* Ferro-Orbital Ordering Transition in Iron Telluride Fe_{1+y}Te . *Phys. Rev. Lett.* **112**, 187202 (2014). [127](#), [155](#), [170](#)
- [167] Qazilbash, M. M. *et al.* Electronic correlations in the iron pnictides. *Nature Physics* **5**, 647–650 (2009). [127](#), [131](#), [133](#)
- [168] Charnukha, A. *et al.* Superconductivity-induced optical anomaly in an iron arsenide. *Nat. Commun.* **2**, 219 (2011). [127](#)
- [169] Yin, Z. P. & Pickett, W. E. Antiphase magnetic boundaries in iron-based superconductors: A first-principles density-functional theory study. *Phys. Rev. B* **80**, 144522 (2009). [128](#)
- [170] Dong, J. *et al.* Competing orders and spin-density-wave instability in $\text{La}(\text{O } 1x \text{ F } x)\text{FeAs}$. *EPL (Europhysics Letters)* **83**, 27006 (2008).
- [171] Mazin, I. I. & Johannes, M. D. A key role for unusual spin dynamics in ferropnictides. *Nature Physics* **5**, 141–145 (2009). [128](#)
- [172] Lee, C.-C., Yin, W.-G. & Ku, W. Ferro-Orbital Order and Strong Magnetic Anisotropy in the Parent Compounds of Iron-Pnictide Superconductors. *Phys. Rev. Lett.* **103**, 267001 (2009). [128](#), [156](#), [170](#), [174](#)
- [173] Subedi, A., Zhang, L., Singh, D. J. & Du, M. H. Density functional study of FeS, FeSe, and FeTe: Electronic structure, magnetism, phonons, and superconductivity. *Phys. Rev. B* **78**, 134514 (2008). [128](#)
- [174] Cao, C., Hirschfeld, P. J. & Cheng, H.-P. Proximity of antiferromagnetism and superconductivity in $\text{LaFeAsO}_{1-x}\text{F}_x$: Effective Hamiltonian from *ab initio* studies. *Phys. Rev. B* **77**, 220506 (2008). [128](#)

- [175] Yin, Z. P., Haule, K. & Kotliar, G. Magnetism and charge dynamics in iron pnictides. *Nature Phys.* **7**, 294–297 (2011). [128](#)
- [176] Yin, Z. P., Haule, K. & Kotliar, G. Kinetic frustration and the nature of the magnetic and paramagnetic states in iron pnictides and iron chalcogenides. *Nature Materials* **10**, 932–935 (2011). [128](#), [185](#)
- [177] de’ Medici, L., Mravlje, J. & Georges, A. Janus-Faced Influence of Hund’s Rule Coupling in Strongly Correlated Materials. *Phys. Rev. Lett.* **107**, 256401 (2011). [128](#)
- [178] Lanatà, N. *et al.* Orbital selectivity in Hund’s metals: The iron chalcogenides. *Phys. Rev. B* **87**, 045122 (2013). [128](#)
- [179] Yin, Z. P., Haule, K. & Kotliar, G. Fractional power-law behavior and its origin in iron-chalcogenide and ruthenate superconductors: Insights from first-principles calculations. *Phys. Rev. B* **86**, 195141 (2012). [129](#)
- [180] Mirri, C. *et al.* Optical conductivity of $\text{FeTe}_{1-x}\text{Se}_x$. *Superconductor Science and Technology* **25**, 045002 (2012). [129](#)
- [181] Wallace, P. R. The Band Theory of Graphite. *Phys. Rev.* **71**, 622–634 (1947). [138](#)
- [182] Raghu, S., Qi, X.-L., Liu, C.-X., Scalapino, D. J. & Zhang, S.-C. Minimal two-band model of the superconducting iron oxypnictides. *Phys. Rev. B* **77**, 220503 (2008). [152](#)
- [183] Yao, Z.-J., Li, J.-X. & Wang, Z. D. Spin fluctuations, interband coupling and unconventional pairing in iron-based superconductors. *New Journal of Physics* **11**, 025009 (2009).

- [184] Das, T. & Balatsky, A. V. Stripes, spin resonance, and nodeless d -wave pairing symmetry in Fe₂Se₂-based layered superconductors. *Phys. Rev. B* **84**, 014521 (2011). [152](#)
- [185] Graser, S., Maier, T. A., Hirschfeld, P. J. & Scalapino, D. J. Near-degeneracy of several pairing channels in multiorbital models for the Fe pnictides. *New Journal of Physics* **11**, 025016 (2009). [152](#)
- [186] Kuroki, K. *et al.* Unconventional Pairing Originating from the Disconnected Fermi Surfaces of Superconducting LaFeAsO_{1-x}F_x. *Phys. Rev. Lett.* **101**, 087004 (2008). [152](#)
- [187] Kontani, H. & Onari, S. Orbital-Fluctuation-Mediated Superconductivity in Iron Pnictides: Analysis of the Five-Orbital Hubbard-Holstein Model. *Phys. Rev. Lett.* **104**, 157001 (2010).
- [188] Ikeda, H., Arita, R. & Kuneš, J. Phase diagram and gap anisotropy in iron-pnictide superconductors. *Phys. Rev. B* **81**, 054502 (2010). [152](#)
- [189] Lee, P. A. & Wen, X.-G. Spin-triplet p -wave pairing in a three-orbital model for iron pnictide superconductors. *Phys. Rev. B* **78**, 144517 (2008). [152](#), [182](#)
- [190] Daghofer, M., Nicholson, A., Moreo, A. & Dagotto, E. Three orbital model for the iron-based superconductors. *Phys. Rev. B* **81**, 014511 (2010). [152](#), [182](#)
- [191] Korshunov, M. M. & Eremin, I. Theory of magnetic excitations in iron-based layered superconductors. *Phys. Rev. B* **78**, 140509 (2008). [152](#)
- [192] Yu, R. *et al.* Magnetic and metallic state at intermediate Hubbard U coupling in multiorbital models for undoped iron pnictides. *Phys. Rev. B* **79**, 104510 (2009). [152](#)

- [193] Saito, T., Onari, S. & Kontani, H. Emergence of fully gapped s_{++} -wave and nodal d -wave states mediated by orbital and spin fluctuations in a ten-orbital model of KFe_2Se_2 . *Phys. Rev. B* **83**, 140512 (2011). [152](#)
- [194] Yanagi, Y., Yamakawa, Y. & no, Y. Electronic States and Pairing Symmetry in the Two-Dimensional 16 Band dp Model for Iron-Based Superconductor. *Journal of the Physical Society of Japan* **77**, 149–150 (2008). [152](#)
- [195] Yanagi, Y., Yamakawa, Y. & no, Y. Magnetism and Superconductivity in the Two-Dimensional 16-Band d p Model for Iron-Based Superconductors. *Journal of the Physical Society of Japan* **77**, 123701 (2008). [152](#)
- [196] Martinelli, A. *et al.* From antiferromagnetism to superconductivity in $\text{Fe}_{1+y}\text{Te}_{1-x}\text{Se}_x$ ($0 \leq x \leq 0.20$): Neutron powder diffraction analysis. *Phys. Rev. B* **81**, 094115 (2010). [155](#), [170](#)
- [197] Schlappa, J. *et al.* Spin-orbital separation in the quasi-one-dimensional Mott insulator Sr_2CuO_3 . *Nature* **485**, 82–85 (2012). [220](#), [228](#)
- [198] Kim, K. W. & Gu, G. D. Optical excitations in Sr_2CuO_3 . *Phys. Rev. B* **79**, 085121 (2009). [220](#), [223](#)
- [199] Gunnarsson, O. & Lundqvist, B. I. Exchange and correlation in atoms, molecules, and solids by the spin-density-functional formalism. *Phys. Rev. B* **13**, 4274–4298 (1976). [229](#)
- [200] von Barth, U. & Hedin, L. A local exchange-correlation potential for the spin polarized case. i. *Journal of Physics C: Solid State Physics* **5**, 1629 (1972). [232](#), [234](#)
- [201] Capelle, K. & Vignale, G. Nonuniqueness of the Potentials of Spin-Density-Functional Theory. *Phys. Rev. Lett.* **86**, 5546–5549 (2001). [233](#), [234](#)

- [202] Eschrig, H. & Pickett, W. Density functional theory of magnetic systems revisited. *Solid State Communications* **118**, 123 – 127 (2001). [233](#), [234](#)
- [203] Capelle, K., Vignale, G. & Ullrich, C. A. Spin gaps and spin-flip energies in density-functional theory. *Phys. Rev. B* **81**, 125114 (2010). [234](#)
- [204] Kohn, W. Density functional theory: Fundamentals and applications. In Bassani, F., Funi, F. & Tosi, M. P. (eds.) *Highlights of Condensed Matter Theory*, 1–15 (North-Holland, Amsterdam, 1985). [234](#)
- [205] Dreizler, R. M. & Gross, E. K. U. *Density Functional Theory: An Approach to the Quantum Many-Body Problem* (Springer-Verlag, Berlin, 1990). [234](#)
- [206] Friedrich, C., Schindlmayr, A., Blügel, S. & Kotani, T. Elimination of the linearization error in GW calculations based on the linearized augmented-plane-wave method. *Phys. Rev. B* **74**, 045104 (2006). [236](#)
- [207] Mattheiss, L. F. & Hamann, D. R. Linear augmented-plane-wave calculation of the structural properties of bulk Cr, Mo, and W. *Phys. Rev. B* **33**, 823–840 (1986). [237](#), [238](#)
- [208] Singh, D. & Krakauer, H. H -point phonon in molybdenum: Superlinearized augmented-plane-wave calculations. *Phys. Rev. B* **43**, 1441–1445 (1991). [237](#), [238](#)
- [209] Singh, D. Ground-state properties of lanthanum: Treatment of extended-core states. *Phys. Rev. B* **43**, 6388–6392 (1991). [238](#)
- [210] Giuliani, G. & Vignale, G. *Quantum Theory of the Electron Liquid* (Cambridge University Press, Cambridge, 2005). [242](#)
- [211] Vilk, Y. M., Chen, L. & Tremblay, A.-M. S. Theory of spin and charge fluctuations in the Hubbard model. *Phys. Rev. B* **49**, 13267–13270 (1994). [256](#)

- [212] Vilk, Y. M. & Tremblay, A.-M. S. Non-Perturbative Many-Body Approach to the Hubbard Model and Single-Particle Pseudogap. *J. Phys. I France* **7**, 1309–1368 (1997). [256](#)
- [213] Miyahara, H., Arita, R. & Ikeda, H. Development of a two-particle self-consistent method for multiorbital systems and its application to unconventional superconductors. *Phys. Rev. B* **87**, 045113 (2013). [256](#)
- [214] Baym, G. Self-Consistent Approximations in Many-Body Systems. *Phys. Rev.* **127**, 1391–1401 (1962). [256](#)

Appendix

Appendix A

An orbitally driven collective excitation in Sr_2CuO_3

The cuprates are one of the most widely studied classes of correlated materials, an interest that began in earnest with the discovery of high-temperature superconductivity in many members of the class. Though this focused on those materials with stacked two-dimensional Cu-O planes, the same CuO_4 building blocks can form one-dimensional chains. The prototype of this subclass is Sr_2CuO_3 , which has been studied as a close analog to one-dimensional versions of the Hubbard and Heisenberg models. The role of the local orbital degrees of freedom on the two-particle excitation spectrum has been a crucial line of research for understanding Sr_2CuO_3 , particularly in connection to spin-charge and more recently spin-orbital [197] separation. In this appendix, I present preliminary results of calculations using the same methods of TDDFT linear response in the Wannier basis used in Chapters 2, 3 and 4. The results show that the first optical peak, measured experimentally [198], can be explained as arising from the transition of a particle in the Cu x^2-y^2 /O p hybridized orbital to the neighboring site. This optical excitation drives a collective mode (plasmon), which has been measured in EELS [33, 34] and is well reproduced by the calculation in the Bloch basis.

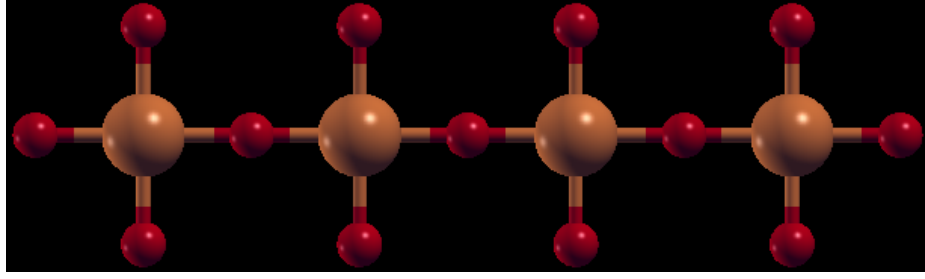


Figure A.1: One-dimensional chain in Sr_2CuO_3 . The relevant structural unit for describing the low-energy electronic structure is the one-dimensional Cu-O chain. This chain consists of corner-sharing CuO_4 plaquettes with the x -direction pointed along the chain and the z -direction normal to the plaquette.

Sr_2CuO_3 is one of a class of strongly anisotropic cuprates that contain structural units that are effectively one-dimensional. This structural unit for Sr_2CuO_3 is shown in Fig. A.1. The basic unit for this chain, and the cuprates in general, is the CuO_4 plaquette: a single Cu atom surrounded on four sides by O atoms forming a square. In Sr_2CuO_3 , the plaquettes are corner-sharing, meaning that each plaquette shares an O atom with each of its two neighbors. The two unshared O plus the two shared, each contributing “half” an atom combine with the central Cu to form the CuO_3 formula unit. In this view of the solid, the Sr atoms are “spectators,” each contributing two electrons to the O atoms. The remaining two electrons to fill the shells of all three O atoms are donated by the Cu atom.

Thus, in a simple ionic view of the solid, the Cu atom would have nine electrons in its d shell, leaving one hole. If all of the d orbitals were roughly degenerate, then one would expect weak correlations, because the d band would be far from half filling. However, the crystal field due to the electrostatic potential of the O atoms exert a strong influence on the local electronic structure of the Cu atom. The square of negatively charged O atoms serve to raise the energy of the x^2-y^2 orbital relative to the other d orbitals because an electron in the x^2-y^2 orbital will have maximal charge density in the direction of the O atoms. In simple band theory, this orbital would be partially occupied and Sr_2CuO_3 would be a metal.

However, Sr_2CuO_3 is an insulator because the “band” due to the x^2-y^2 orbital is narrow and the screened onsite interaction strong enough to effect a metal-insulator transition. For this reason, LSDA or GGA approximations to Kohn-Sham DFT give a predictably poor description of the electronic structure and the LDA+ U (or GGA+ U , as I have used) is required to sufficiently open the gap. In the low-temperature insulating phase, Sr_2CuO_3 is an antiferromagnet, with the magnetic moment of on the Cu atoms alternating along the one-dimensional chain.

It is possible to project Wannier functions for only the d orbitals and produce a good representation of the electronic structure. Each Cu site has nine occupied orbitals and one unoccupied orbital per site, but it is useful for later numerical convenience in the calculation of the response to only project orbitals for one spin component. Since Sr_2CuO_3 is an antiferromagnet, the orbitals for one spin will be same as that for the opposite spin on the neighboring Cu sites. If only one spin component is considered, then the occupancy of the d orbitals on the Cu atoms alternates along the chain. In the designation that I will be using here, Cu(1) has four occupied and one unoccupied d orbital, which is x^2-y^2 , and Cu(2) had five occupied d orbitals.

The details of the electronic structure of the x^2-y^2 orbital is visible in the band structure, which I present in Fig. A.2a. The unoccupied x^2-y^2 orbitals, one for each spin, or alternatively one for each Cu atom, sit in a pair of almost entirely degenerate flat bands just above the gap. The occupied x^2-y^2 orbitals sit in a another pair of almost entirely degenerate bands, though there is some slight mixture with the other d bands, with avoided crossings between Γ and X and Γ and S. Still, for the most part the occupied x^2-y^2 orbitals occupy their own subspace of bands. As may be expected, there is a pronounced dispersion in these directions while the bands in the other directions are relatively flat, reflecting the one-dimensional nature of the material. The unoccupied x^2-y^2 orbital lies entirely in two flat, largely degenerate bands. The occupied x^2-y^2 orbital also mostly lies in two bands, though there are band crossings with bands that have other d -orbital content.

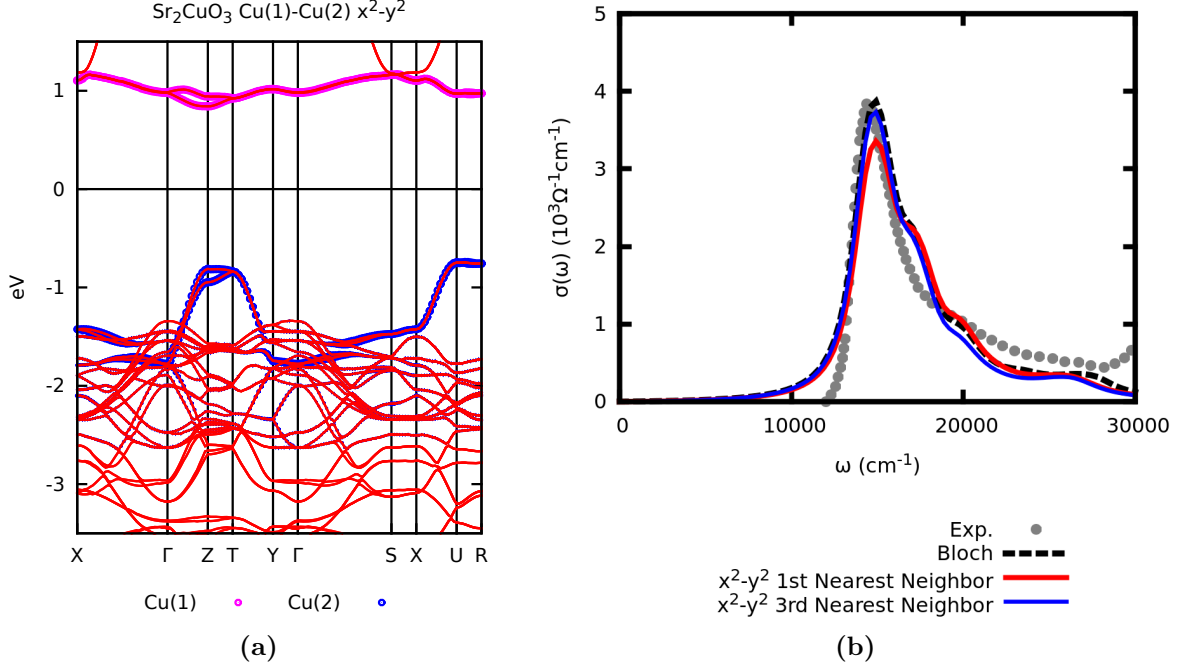


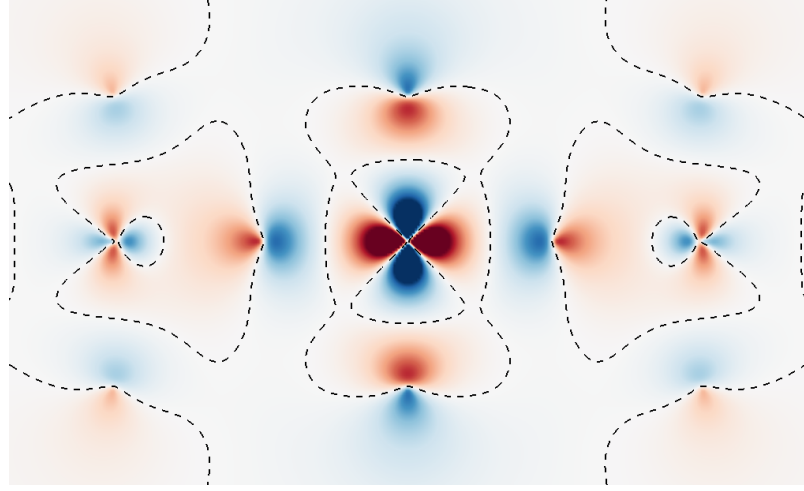
Figure A.2: Band structure and contributions of Wannier channels to $\sigma_1(\omega)$ in Sr_2CuO_3 . The amount of orbital content of the x^2-y^2 Wannier states for Cu(1) and Cu(2) is shown with magenta and blue symbols, respectively in (a). The Γ -Z, T-Y and X-U directions are in the same direction as the chain (the z -direction). Since the optical peak in (b) is at ≈ 2 eV, we must focus on the flat sections of the occupied band structure at ≈ -1 eV as providing the hole states for the particle hole transition. The spectrum calculated using the Bloch basis is represented by the black dashed curve and agrees well with optical measurements from Ref. [198], shown in gray dots. The peak can be almost entirely explained by the x^2-y^2 transitioning to the neighboring Cu site, as shown by the red curve. There is a small ($\approx 10\%$) contribution from a transition to the site three Cu atoms down the chain. Transitions to the second nearest neighbor Cu site are excluded because these would involve a spin-flip. The other d orbitals play almost no role.

Figure A.2b shows the optical peak calculated using the Bloch basis in the black dashed line along with two decompositions into Wannier particle-hole channels. The red curve is made up of a single type of transition: a particle moving from the occupied x^2-y^2 orbital on one Cu site to the unoccupied orbital in the neighboring Cu site, leaving a hole behind. There are two such transitions for each site, one to the right and one to the left. Clearly, these transitions account for almost the entire optical peak. Virtually all of the remaining $\sim 10\%$ of the peak height is due to the same type of transition, only to the third-nearest neighbor site. Transitions to the second-nearest neighbor are excluded because they would include a spin flip, which is not allowed for a charge fluctuation.

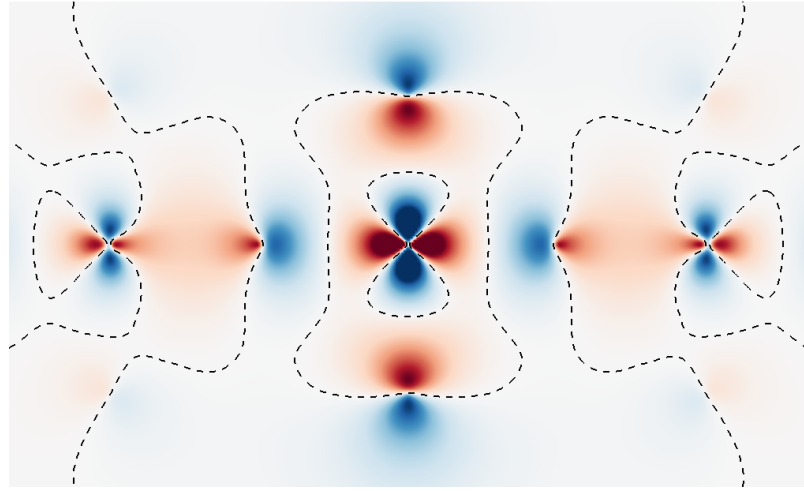
It is useful to examine the orbitals themselves in real space. The x^2-y^2 orbitals for a single spin projection are shown in Fig. A.3. The most noticeable non-trivial aspect of these orbitals is that they exhibit significant hybridization with the neighboring p orbitals. In both the unoccupied orbitals, the central x^2-y^2 core is in an anti-bonding configuration with the neighboring x and y orbitals. It is this chemistry that places the orbital above the other d orbitals, which do not participate in bonding or anti-bonding with the O p orbitals.

There are two primary differences between the occupied and unoccupied x^2-y^2 orbitals. First, the x hybridization tails on the neighboring O atoms along the chain are in an anti-bonding configuration with the x^2-y^2 hybridization tail on the neighboring Cu atom along the chain for the unoccupied orbital shown in Fig. A.3a. However, the same hybridization tails are in a bonding configuration for the occupied orbital shown in Fig. A.3b. This chemistry can explain why one orbital is occupied and the other empty: the bonding configuration will always have lower energy. Second, the unoccupied orbital is more localized than the occupied orbital, with more weight in the central x^2-y^2 core than in the hybridization tails.

The polarization process that produces the peak in the imaginary part of $\epsilon(\omega)$ modulates the real part to produce a minimum in the dielectric function which manifests as a peak in the loss function. This is the hallmark of a collective mode and



(a)



(b)

Figure A.3: Real space plots of the x^2-y^2 Wannier functions in Sr_2CuO_3 . The orbitals for a single spin projection are shown centered on Cu(1) in (a) and Cu(2) in (b). The black line shows the nodes (zeros) of the Wannier function, blue is positive value and red is negative value. (a) shows an unoccupied orbital and (b) show an occupied one. Both orbitals display an anti-bonding configuration between the central Cu and the neighboring O atoms. However, the occupied orbital on Cu(2) shows a bonding configuration between the along-chain oxygen hybridization tail and the tail on the neighboring Cu site along the chain while the unoccupied orbital on Cu(1) shows an anti-bonding configuration of the same. In addition, the occupied orbital has more weight in the hybridization tails and consequently less in the core of the orbital centered on Cu(2). Thus, there are a subtle differences in the chemistry of the two orbitals both in the core and in the longer-ranged hybridization tails that serves to distinguish the two orbitals.

this mode for increasing momentum transfer is shown in Fig. A.4. The peak starts out at ≈ 2.5 eV and disperses upwards in energy to ≈ 3.0 eV as the momentum transfer approaches the edge of the first Brillouin zone. The peak also damps out, almost disappearing from the loss function near the zone edge. Both of these properties are shared with the conventional plasmon in the homogeneous electron gas.

The picture of the excitation that drives the collective mode in Fig. A.4 is largely, if qualitatively, in agreement with what Neudert et al. concluded using calculations in the one-dimensional, one-band Hubbard model. This may not be surprising since the x^2-y^2 orbitals shown in Fig. A.3 correspond to degrees of freedom in the one-band system. However, one difference that may be important is that I constructed these orbitals from a groundstate that already included many-body effects of the onsite interaction that served to open the charge-transfer gap. In a later refinement on the earlier one-band calculation, Moskvin et al. [34] used a multi-band model that included O p degrees in freedom in recognition that Sr_2CuO_3 should be considered as a charge-transfer insulator like the other cuprates.

Clearly, the p orbitals are critical in forming the orbitals shown in Fig. A.3. Though the orbitals have symmetry appropriate to an x^2-y^2 orbital, they are strongly hybridized with the p orbitals. But there is a potentially relevant difference between the interpretation for the two-particle spectrum I have presented here and that derived from the multi-band model in Ref. [34]: in the calculation that produces the optical peak in Fig. A.2b, the only degrees of freedom are the x^2-y^2/p hybridized orbitals, whereas in Ref. [34] the x^2-y^2 and p orbitals are separate degrees of freedom. These then generate multiple excitations versus the single excitation that the current Wannier/TDDFT calculation indicates.

It is also possible to construct a Wannier basis that explicitly includes the p orbitals as degrees of freedom independently from the x^2-y^2 orbital and calculate the response due to these orbitals. We have done this that find that the exact same correspondence can be achieved as is shown in Fig. A.2b, but it involves several channels which include the x^2-y^2 , x and y orbitals. As these channels together reproduce the same

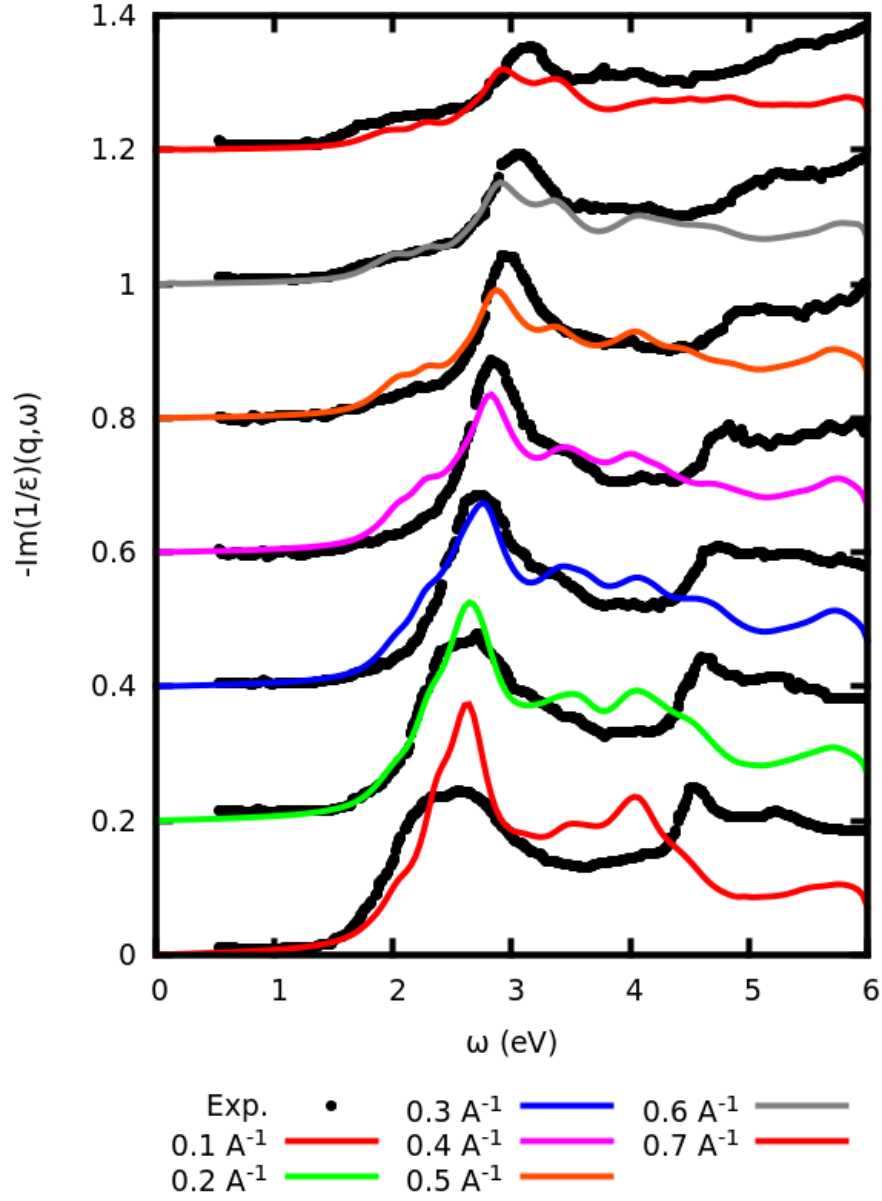


Figure A.4: Dispersion of the loss peak in Sr_2CuO_3 . Both the intensity and trend in dispersion of the calculated loss peak agree fairly well with experimental EELS measurements from Ref. [33], shown in black dots. From the decomposition of χ^{KS} into Wannier channels, I can identify the orbital particle-hole excitation that drives this collective mode. This basic unit of the collective mode is a particle in x^2-y^2 orbital transitioning to the neighboring x^2-y^2 orbital, leaving a hole behind. However, the appearance of the collective mode at small $\mathbf{q}$ relies on the long-ranged nature of the interaction and the collective mode is the cooperative motion of these basic units all along the chain.

optical peak that is produced by the orbitals shown in Fig. A.3, it is more natural to consider the basis involving only hybridized d orbitals as offering the best insight into the physics behind the excitation. This is another manifestation of the optimal Wannier basis resulting from minimizing the number of channels needed to explain the low-energy optical spectrum, exactly like that described in Chapters 2 and 4.

In contrast to the interpretations of the two-particle spectrum that involve only x^2 - y^2 and p orbitals, Schlappa et al. [197] have proposed, based on RIXS measurements, the existence of excitations that involve the other d orbitals and result in spin-orbital separation, a novel extension of spin-charge separation. In the calculation that I used to produce the Wannier response spectra in Fig. A.2b, the other d orbitals play no role at all. However, there are fundamental differences in the quantity that EELS measures and that which RIXS measures. The spectrum measured by EELS is determined by the density-response function χ , precisely what I am calculating. RIXS measures a second order process which can couple to excitations that are inaccessible to EELS, but with the disadvantage that the comparison between theory and experiment is less clear. Thus, it is an open question what the relation is between the picture that I have presented here and that of Schlappa et al.

Appendix B

Explicit formula for the exchange-correlation energy

Though the exchange correlation functional $E^{\text{xc}}[n]$ that appears in Kohn-Sham DFT is, in the vast majority of cases, approximated by some local functional such as LDA or GGA, there does exist an exact expression for $E^{\text{xc}}[n]$. In order to make $E^{\text{xc}}[n]$ less of a mysterious object and put it on a more concrete foundation, I will sketch the derivation here following Gunnarsson and Lundqvist [199].

I begin with the textbook expression for the interaction energy [71]:

$$U = \frac{e^2}{2} \int d^3x d^3x' \frac{\langle \hat{n}(\mathbf{x}) \hat{n}(\mathbf{x}') \rangle - \delta(\mathbf{x} - \mathbf{x}') \langle \hat{n}(\mathbf{x}) \rangle}{|\mathbf{x} - \mathbf{x}'|}. \quad (\text{B.1})$$

Here, the angled brackets denote a groundstate expectation value, though it is possible to perform this derivation for finite temperature, in which case the brackets would denote an equilibrium average. Now I will define a Hamiltonian $\hat{H}(g)$ for which the coupling constant g is taken as a variable. Not only that, I will also define an external potential $\hat{V}[n, g]$ that will depend on the coupling constant as well. It will serve to fix the true external potential $\hat{V}$ such that the groundstate of $\hat{H}(g)$ has the density

$n(\mathbf{x})$. That is

$$\hat{H}(g) = \hat{T} + \hat{V} + \hat{V}[n, g] + \frac{g}{2} \sum_{i \neq j} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|}. \quad (\text{B.2})$$

Then the total energy can be written as a function of the coupling constant via the expectation value:

$$E[n, g] = \langle \Psi[n, g] | \hat{H}(g) | \Psi[n, g] \rangle. \quad (\text{B.3})$$

Now it is possible to write the expression in Eq. (B.3) in terms of the non-interacting energy and an integration over the coupling constant:

$$E[n, g] = E[n, 0] + \int_0^g dg \left\langle \Psi[n, g] \left| \frac{d\hat{V}[n, g]}{dg} \right| \Psi[n, g] \right\rangle + \int_0^g \frac{dg}{g} U[n, g], \quad (\text{B.4})$$

where

$$U[n, g] = \left\langle \Psi[n, g] \left| \frac{g}{2} \sum_{i \neq j} \frac{1}{|\mathbf{x}_i - \mathbf{x}_j|} \right| \Psi[n, g] \right\rangle. \quad (\text{B.5})$$

Now it is actually possible to evaluate the second term in Eq. (B.4) in the case where $g = e^2$ because the expectation value of the potential can be evaluated in the same way as in Eq. (1.12). Evaluating the expectation value in this way crucially depends on the density being independent of the coupling constant, which we have guaranteed by our construction.

$$\begin{aligned} \int_0^{e^2} dg \left\langle \Psi[n, g] \left| \frac{d\hat{V}[n, g]}{dg} \right| \Psi[n, g] \right\rangle &= \int_0^{e^2} dg \int d^3x \frac{d\tilde{v}[n, g](\mathbf{x})}{dg} n(\mathbf{x}) \\ &= - \int d^3x \tilde{v}[n, 0](\mathbf{x}) n(\mathbf{x}), \end{aligned} \quad (\text{B.6})$$

because $\tilde{v}[n, e^2](\mathbf{x}) = 0$ by construction. Now, we can recognize the non-interacting kinetic energy in $E[n, 0]$ as precisely what we defined as $T^{\text{KS}}[n]$ in Eq. (1.19). Thus,

$$E[n, 0] = T^{\text{KS}}[n] + \int d^3x [v(\mathbf{x}) + \tilde{v}[n, 0](\mathbf{x})] n(\mathbf{x}). \quad (\text{B.7})$$

Putting together Eq. (1.22) with Eqs. (B.4), (B.6) and (B.7) and defining the density fluctuation operator $\hat{\tilde{n}}(\mathbf{x}) = \hat{n}(\mathbf{x}) - n(\mathbf{x})$ (which allows for absorption of the Hartree term), we can write an explicit expression for $E^{\text{xc}}[n]$:

$$E^{\text{xc}}[n] = \frac{1}{2} \int_0^{e^2} dg \, d^3x \, d^3x' \frac{\langle \Psi[n, g] | \hat{\tilde{n}}(\mathbf{x}) \hat{\tilde{n}}(\mathbf{x}') | \Psi[n, g] \rangle - \delta(\mathbf{x} - \mathbf{x}') n(\mathbf{x})}{|\mathbf{x} - \mathbf{x}'|}. \quad (\text{B.8})$$

Here we can identify the expression $\langle \Psi[n, g] | \hat{\tilde{n}}(\mathbf{x}) \hat{\tilde{n}}(\mathbf{x}') | \Psi[n, g] \rangle$ as the zero-temperature, static density-density correlation function, *evaluated for each value of the coupling constant*. Thus, if Eq. (B.8) were to be used to actually calculate the exchange-correlation energy, the many-body calculation yielding the density-density correlation function would have to be done at several values of the coupling constant. Also, the calculation depends only implicitly on the density, so it is not a method for deriving a universal functional into which a density can simply be “plugged in”. But even if Eq. (B.8) is not used to explicitly calculate $E^{\text{xc}}[n]$, it is still a useful tool theoretically.

Appendix C

A discussion of spin density-functional theory

In presentations of spin density-functional theory (SDFT), the generalization to spin is often treated as a natural extension of the spin-degenerate form with no new complications being introduced. This approach was already presented by Kohn and Sham in the original presentation of their method [7]. However, the proof of the Hohenberg-Kohn theorems as I have presented cannot be extended to systems in which an additional external field couples to the spin-density. This was first pointed out by von Barth and Hedin [200], who were able to construct a counter example to the one-to-one mapping between (nondegenerate) groundstate wavefunctions and external potentials in the spin-dependent case. In that construction, the potentials do not differ by at most a trivial constant, as I presented for the spin-independent case. Instead, they were able to construct an infinitely large class of potentials, which if small enough, would all give the same groundstate density. They were able to salvage the situation for nondegenerate groundstates by presenting a proof that bypassed the mapping between external potentials and groundstate wavefunctions and instead directly proving that there is a one-to-one mapping between external potentials and groundstate densities.

However, the complications involved with extending DFT to spin-densities did not end there. The problem was addressed decades later independently by Cappelle and Vignale [201] and Eschrig and Pickett [202] and in considerably more detail. Cappelle and Vignale were able to show that a necessary condition for two external potentials to yield the same groundstate is that there exists a linear combination of the charge and magnetization density operators for which the groundstate is an eigenfunction. From there they classified the cases for nonuniqueness of the potentials into systematic nonuniqueness and accidental nonuniqueness.

For systematically nonunique cases, where there is a gap between the groundstate and first excited state (a condition for most examples of nonuniqueness), a family of potentials that give the same groundstate wavefunction can be constructed from constants of motion. Accidentally nonunique cases will tend to be pathological examples, like the case of a ferromagnet in which the Kohn-Sham eigenfunctions of only one spin are occupied [201]. Cappelle and Vignale were able to identify specific problem areas for which the uniqueness of the potentials presents a particular problem. Luckily, from the perspective of this dissertation, none of the problem areas will be encountered in this work.

Eschrig and Pickett were able to narrow the types of groundstates for which nonuniqueness is a problem to states that, when coordinates are rotated to align with the magnetic field, are a linear combination of components that all have the same number of electrons with spin up (and hence also have the same number of electrons with spin down). In this case, the magnetic field can be increased by a small amount, changing the external potential and producing a Zeeman shift in the eigenenergies. As long as the field is not large enough to produce a crossing between ground state and an excited state, the ground state will remain the same while the external potential has changed. Recall that the external potential in spin-independent DFT could also be shifted by a constant while preserving the groundstate, but that was due to the ambiguity in choosing a zero from which to measure the energy. Once the energy zero is fixed, the ability to change the magnetic field while preserving the

groundstate remains in SDFT and is thus an example of nonuniqueness. Also, in the case of noncollinear magnetizations, the magnetic field can be changed in a spatially dependent manner and so the change is not constant as it was for the potential in spin-independent DFT.

Both Cappelle and Vignale and Eschrig and Pickett also pointed out that the energy functional is generally not differentiable with respect to the spin-up and spin-down densities. This is essential if we are to be able to take the functional derivatives that are needed in the Kohn-Sham method. The problem is intimately connected to the well-known “gap problem” in DFT when it is extended to systems with a variable number of particles [116]. There, the energy functional is not differentiable for systems with a gap when varying particle number across an integer because the chemical potential jumps discontinuously. In the case of SDFT, there are two particle numbers and if even one of the spin components is gapped, as in a half-metal, then the same problem arises [203]. This makes calculating spin-flip energies using Kohn-Sham eigenvalues in half-metals unreliable [201, 202].

Fortunately, Capelle, Ullrich and Vignale [42] were able to prove several years later that the existence of the universal, exchange correlation and Kohn-Sham kinetic energy functionals in the most general case, including the case of degenerate groundstates, which the proof of von Barth and Hedin excluded [200]. Separately, degeneracy and nonuniqueness do not pose a problem for the existence of the DFT energy functionals. For spin-independent DFT with the possibility of degenerate groundstates, one can show [204, 205] that, even though a one-to-one correspondence between groundstate densities and wavefunctions is no longer possible, any one of the groundstate wavefunctions still determines the external potential and the functional $V[n]$ still exists. Note that this functional is not necessarily invertible since several densities may give the same potential. However, this is not necessary to define the universal functional $F[n]$ by

$$F[n] \equiv E - V[n]. \quad (\text{C.1})$$

Though definition of the universal functional this way is possible, it is no longer possible to define the kinetic and interaction energy functionals $T[n]$ and $U[n]$. For SDFT without degenerate groundstates, the one-to-one correspondence between groundstate densities and wavefunctions is retained and the universal functional is defined as

$$F[n] \equiv \langle \Psi[n] | \hat{T} + \hat{U} | \Psi[n] \rangle, \quad (\text{C.2})$$

where $|\Psi[n]\rangle$ is defined as in Eq. (1.14).

However, for SDFT with degenerate groundstates, neither the mappings between groundstate densities and wavefunctions nor between groundstate wavefunctions and potentials can be defined. Instead, Capelle, Ullrich and Vignale proved two weaker theorems: the joint-degeneracy theorem and internal-energy theorem and used them to define $F[n]$ in the most general case [42]. I will state them without proof. The joint-degeneracy theorem states that if two groundstate wavefunctions of a Hamiltonian with one potential have the same density, then for a Hamiltonian with any other potential, they are either both groundstates or neither are. The internal-energy theorem states that two systems with the same groundstate density must also have the same internal energy F . The internal-energy theorem establishes that F is a functional of the density. The joint-degeneracy theorem can be used to show that this functional $F[n]$ does not depend on the potentials and is thus universal.

Appendix D

Proof of orthogonality of core orbitals in the LAPW basis

An important property of the radial functions defined by Eq. (1.56) is the expression for the inner product of two radial functions evaluated for different linearization energies. In the derivation that follows I drew on the proof of Friedrich et al. [206]. First, let us examine the inner product to two radial functions multiplied by the linearization energy. Using a textbook transformation of the radial equation Eq. (1.56), I find the relation

$$\begin{aligned} E_L \int_0^{R_{\text{MT}}} dr r^2 f_{\alpha\ell}(r; E_L) f_{\alpha\ell}(r; E_C) = & -\frac{\hbar^2}{2m} \int_0^{R_{\text{MT}}} dr r \frac{d^2}{dr^2} \left[r f_{\alpha\ell}(r; E_L) \right] f_{\alpha\ell}(r; E_C) \\ & + \frac{\hbar^2 \ell(\ell+1)}{2m} \int_0^{R_{\text{MT}}} dr f_{\alpha\ell}(r; E_L) f_{\alpha\ell}(r; E_C) \\ & + \int_0^{R_{\text{MT}}} dr r^2 v_{\sigma,\alpha 00}^{\text{KS}}(r) f_{\alpha\ell}(r; E_L) f_{\alpha\ell}(r; E_C). \end{aligned} \tag{D.1}$$

The last two terms on the right hand side are symmetric under the exchange of E_L and E_C , so I can take the following difference and only the terms with derivatives

remain:

$$\begin{aligned}
(E_L - E_C) \int_0^{R_{\text{MT}}} dr r^2 f_{\alpha\ell}(r; E_L) f_{\alpha\ell}(r; E_C) \\
= -\frac{\hbar^2}{2m} \left\{ \int_0^{R_{\text{MT}}} dr r \frac{d^2}{dr^2} [r f_{\alpha\ell}(r; E_L)] f_{\alpha\ell}(r; E_C) \right. \\
\left. - \int_0^{R_{\text{MT}}} dr r f_{\alpha\ell}(r; E_L) \frac{d^2}{dr^2} [r f_{\alpha\ell}(r; E_C)] \right\}.
\end{aligned} \tag{D.2}$$

Integration by parts and several cancellations yield the desired result:

$$\begin{aligned}
(E_L - E_C) \int_0^{R_{\text{MT}}} dr r^2 f_{\alpha\ell}(r; E_L) f_{\alpha\ell}(r; E_C) \\
= -\frac{\hbar^2 R_{\text{MT}}^2}{2m} \left\{ \frac{df_{\alpha\ell}(r; E_L)}{dr} \Big|_{R_{\text{MT}}} f_{\alpha\ell}(R_{\text{MT}}; E_C) \right. \\
\left. - f_{\alpha\ell}(R_{\text{MT}}; E_L) \frac{df_{\alpha\ell}(r; E_C)}{dr} \Big|_{R_{\text{MT}}} \right\}.
\end{aligned} \tag{D.3}$$

The crucial point is that if $f_{\alpha\ell}(R_{\text{MT}}; E_C) = 0$ and $\frac{df_{\alpha\ell}(r; E_C)}{dr} \Big|_{R_{\text{MT}}} = 0$ then $f_{\alpha\ell}(r; E_L)$ and $f_{\alpha\ell}(r; E_C)$ are orthogonal. This is precisely how core states are usually treated and are treated throughout this dissertation. Since the orbitals for the core are taken to vanish outside the muffin tin, the linearized APW functions for the valence and the local orbitals for the core are strictly orthogonal to each other. This turns out to be essential to ensure the stability of the calculation [207, 208].

This property of local orbitals is relevant to the treatment of extended semi-core states that are particularly important for early transition-metal and rare-earth elements, among others. Using Eq. (D.3) I showed that the valence and the core orbitals are orthogonal if the radial part of the core orbital and its derivative vanish at the muffin tin boundary. The problem lies in states that are not quite deep enough in the potential well of the nucleus to be contained entirely in the muffin tin, but are also significantly deeper than the valence states. In this situation, there are two problems. First, the variational freedom for APW functions with a linearization energy close to the valence energies is insufficient to represent the radial part of the

semi-core states [209]. This should be obvious since the radial part of a 4p orbital that may reside in the valence and a 3p orbital in the semi-core should have completely different radial parts (e.g. the number of nodes will differ.) Second, the semi-core and valence are no longer orthogonal. This turns out to introduce instability in the total energy when the linearization energy is varied [207, 208]. The solution introduced by Singh [209] is to add local orbitals to the basis set by forming linear combinations of the radial functions $f_{\alpha\ell}(r; E_L)$, $\frac{df_{\alpha\ell}(r; E_L)}{dr}$ and $f_{\alpha\ell}(r; E_{SC})$ and enforcing the condition that the local orbital radial function and its first derivative vanish at the muffin tin boundary. E_{SC} is chosen so that it lies in the semi-core level and can be adjusted adaptively during the self-consistent loop.

Appendix E

Explicit expressions for LAPW matrix elements

In Section 1.2, I outlined the construction of the APW+LO+lo basis set. Here, I will present some explicit expressions for the matrix elements that must be calculated in order to solve the Kohn-Sham eigenvalue problem. I focus only on the calculation of the elements of $\mathbf{H}^{\text{APW,APW}}(\mathbf{k})$ and simply note the differences between the expressions for those elements and the elements for the other blocks. Also, I will consider separately the kinetic energy and effective potential parts of $\mathbf{H}^{\text{APW,APW}}(\mathbf{k})$, which are simply added together.

First is the kinetic energy $\mathbf{T}^{\text{APW,APW}}(\mathbf{k})$. As is apparent in the expression for the Kohn-Sham kinetic energy in Eq. (1.63) the muffin-tin and interstitial contributions are simply added together. I will consider them separately. The interstitial part is given by examining Eq. (1.63) in light of the expression for the APW functions in the interstitial in Eq. (1.96):

$$T_{\mathbf{G}\mathbf{G}'}^{\text{APW,APW};\text{I}}(\mathbf{k}) = \frac{\hbar^2}{2m}(\mathbf{k} + \mathbf{G}) \cdot (\mathbf{k} + \mathbf{G}') \frac{1}{\Omega} \int_{\text{I}} d^3x e^{-i(\mathbf{G}-\mathbf{G}')\cdot\mathbf{x}}. \quad (\text{E.1})$$

The most important thing to note here is that the integral is not equal to $\delta_{\mathbf{G}\mathbf{G}'}$ as it would be for planewaves as the integral is only over the interstitial, not over all of

space. Instead, it can be viewed as the Fourier transform of a so-called “characteristic function”:

$$\Theta_{\mathbf{G}-\mathbf{G}'}^{\text{C}} = \frac{1}{\Omega} \int d^3x e^{-i(\mathbf{G}-\mathbf{G}') \cdot \mathbf{x}} \Theta^{\text{C}}(\mathbf{x}), \quad (\text{E.2})$$

where $\Theta^{\text{C}}(\mathbf{x}) = 1$ if $\mathbf{x}$ is in the interstitial and zero otherwise. For the muffin-tin contributions, the angular integrals can be computed analytically yielding an expression containing only one-dimensional radial integrals:

$$\begin{aligned} T_{\mathbf{G}\mathbf{G}'}^{\text{APW,APW;MT}}(\mathbf{k}) &= \frac{\hbar^2}{2m} \sum_{\alpha\ell m} A_{\mathbf{k}+\mathbf{G}}^{*\alpha\ell m} A_{\mathbf{k}+\mathbf{G}'}^{\alpha\ell m} \\ &\quad \times \left\{ - \int_0^{R_{\text{MT}\alpha}} dr f_{1\alpha\ell}(r) \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial f_{1\alpha\ell}}{\partial r} \right) + \ell(\ell+1) f_{1\alpha\ell}(r) \right] \right. \\ &\quad \left. + R_{\text{MT}\alpha}^2 f_{1\alpha\ell}(R_{\text{MT}\alpha}) \frac{\partial f_{1\alpha\ell}}{\partial r} \Big|_{R_{\text{MT}\alpha}} \right\}. \end{aligned} \quad (\text{E.3})$$

The matrix elements for the APW-APW block of the Kohn-Sham effective potential is also the sum of an interstitial contribution and contributions from the muffin-tin. The interstitial integral is again complicated by the fact that the integral is not over all space. Thus what appears is not just $v_{\mathbf{G}-\mathbf{G}'}^{\text{KS}}$ as written in Eq. (1.50). Instead, I must again use the characteristic function given in Eq. (E.2). For the muffin-tin contributions, I can use the expansion radial functions of the potential inside the muffin-tin sphere in spherical harmonics like that given in Eq. (1.55): $v_{\alpha\ell m}^{\text{KS}}(r)$. With these considerations it is straightforward to calculate the matrix elements of the potential:

$$\begin{aligned} V_{\mathbf{G}\mathbf{G}'}^{\text{APW,APW}}(\mathbf{k}) &= \sum_{\mathbf{G}_1} v_{\mathbf{G}-\mathbf{G}_1}^{\text{KS}} \Theta_{\mathbf{G}_1-\mathbf{G}'}^{\text{C}} \\ &\quad + \sum_{\alpha} \sum_{\substack{\ell m \\ \ell' m' \\ \ell'' m''}} A_{\mathbf{k}+\mathbf{G}}^{*\alpha\ell m} A_{\mathbf{k}+\mathbf{G}'}^{\alpha\ell' m'} \int_0^{R_{\text{MT}\alpha}} dr r^2 f_{1\alpha\ell}(r) v_{\alpha\ell'' m''}^{\text{KS}}(r) f_{1\alpha\ell'}(r) \\ &\quad \times \int d\Omega Y_{\ell m}^*(\hat{\mathbf{r}}) Y_{\ell'' m''}(\hat{\mathbf{r}}) Y_{\ell' m'}(\hat{\mathbf{r}}). \end{aligned} \quad (\text{E.4})$$

The angular integral is given by the Gaunt coefficients and they serve to couple the different angular momentum components of the APW functions inside the muffin-tin. This is what makes the APW+LO+lo method as implemented in Exciting Plus [13] a full potential method, meaning the non-spherical parts of the potential are included. This is in contrast for the expression for the kinetic energy in Eq. (E.3) which did not couple different angular momentum components.

Since the local orbitals are zero in the interstitial, the other blocks of $\mathbf{H}^{\text{KS}}(\mathbf{k})$ do not have any contribution from the interstitial. The calculation inside the muffin tins proceeds in exactly the same way as for the APW-APW block except the APW-LO block will include only one factor of $A_{\mathbf{k}+\mathbf{G}}^{\alpha\ell m}$ and the LO-LO block will not contain any of these matching coefficients. I also note that, since the APW+LO+lo basis is not orthogonal, an overlap matrix must be calculated. This will proceed as a special case of the calculation of the potential matrix elements with the replacement of the potential with the constant 1.

Appendix F

Proof of the Runge-Gross Theorem

In the proof of the Runge-Gross theorem that I present in what follows I am closely following the proof of van Leeuwen [68] and the presentations of Giuliani and Vignale [210] and Ullrich [5]. To begin, we need a differential equation for the density relating it to the potential. This will in essence relate the density at a particular time to the potential at all previous times. It is useful to prove the theorem for an unspecified initial state for this differential equation, because the proof uses a Taylor series which need only converge in the neighborhood of t_0 . If it diverges outside of this radius of convergence then a new initial state can be chosen at a time after t_0 but inside the radius of convergence and the proof can thus be analytically continued to arbitrarily large times after t_0 . This means that the differential equation will indeed cover all times without issues of convergence. With this in mind, I will only consider the proof for times inside the radius of convergence of all the Taylor series involved [68].

I first point out that the density at an infinitesimally later time is determined by the current via the continuity equation:

$$\frac{\partial}{\partial t}n(\mathbf{x}, t) = -\nabla \cdot \mathbf{j}(\mathbf{x}, t). \quad (\text{F.1})$$

This by itself is not enough though, since there is no explicit appearance of the potential. Thus we must consider the time derivative of the current given by

$$i\hbar \frac{\partial}{\partial t} \mathbf{j}(\mathbf{x}, t) = \langle \Psi(t) | [\hat{\mathbf{j}}(\mathbf{x}), \hat{H}(t)]_- | \Psi(t) \rangle. \quad (\text{F.2})$$

The first part of the commutator,

$$\hat{\mathbf{Q}}(\mathbf{x}) = \frac{1}{i\hbar} [\hat{\mathbf{j}}(\mathbf{x}), \hat{T} + \hat{W}]_-. \quad (\text{F.3})$$

does not depend explicitly on the potential and so its details need not concern us. The second part of the commutator is given by

$$\begin{aligned} & [\hat{\mathbf{j}}(\mathbf{x}), \int d^3x' v(\mathbf{x}', t) \hat{n}(\mathbf{x}')]_- \\ &= -\frac{i\hbar}{2m} \int d^3x' v(\mathbf{x}', t) \{ [\hat{\psi}^\dagger(\mathbf{x}) \nabla \hat{\psi}(\mathbf{x}), \hat{n}(\mathbf{x}')]_- - [\nabla \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}(\mathbf{x}), \hat{n}(\mathbf{x}')]_- \} \\ &= -\frac{i\hbar}{2m} \int d^3x' v(\mathbf{x}', t) \{ \hat{\psi}^\dagger(\mathbf{x}) \nabla [\hat{\psi}(\mathbf{x}), \hat{\psi}^\dagger(\mathbf{x}')]_+ \hat{\psi}(\mathbf{x}') - \hat{\psi}^\dagger(\mathbf{x}') [\hat{\psi}^\dagger(\mathbf{x}), \hat{\psi}(\mathbf{x}')]_+ \nabla \hat{\psi}(\mathbf{x}) \\ &\quad - \nabla \hat{\psi}^\dagger(\mathbf{x}) [\hat{\psi}(\mathbf{x}), \hat{\psi}^\dagger(\mathbf{x}')]_+ \hat{\psi}(\mathbf{x}') + \hat{\psi}^\dagger(\mathbf{x}') \nabla [\hat{\psi}^\dagger(\mathbf{x}), \hat{\psi}(\mathbf{x}')]_+ \hat{\psi}(\mathbf{x}') \}. \end{aligned} \quad (\text{F.4})$$

Using the result

$$\nabla [\hat{\psi}(\mathbf{x}), \hat{\psi}^\dagger(\mathbf{x}')]_+ = \nabla \delta^3(\mathbf{x} - \mathbf{x}') = -\nabla' \delta^3(\mathbf{x} - \mathbf{x}') \quad (\text{F.5})$$

and integration by parts I arrive at

$$\begin{aligned} & [\hat{\mathbf{j}}(\mathbf{x}), \int d^3x' v(\mathbf{x}', t) \hat{n}(\mathbf{x}')]_- \\ &= -\frac{i\hbar}{2m} \{ 2\nabla v(\mathbf{x}, t) \hat{n}(\mathbf{x}) + v(\mathbf{x}, t) \hat{\psi}^\dagger(\mathbf{x}) \nabla \hat{\psi}(\mathbf{x}) - v(\mathbf{x}, t) \hat{\psi}^\dagger(\mathbf{x}) \nabla \hat{\psi}(\mathbf{x}) \\ &\quad - v(\mathbf{x}, t) \nabla \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}(\mathbf{x}) + v(\mathbf{x}, t) \nabla \hat{\psi}^\dagger(\mathbf{x}) \hat{\psi}(\mathbf{x}) \}. \end{aligned} \quad (\text{F.6})$$

Thus

$$\frac{\partial}{\partial t} \mathbf{j}(\mathbf{x}, t) = -\frac{1}{m} \nabla v(\mathbf{x}, t) n(\mathbf{x}, t) + \mathbf{Q}(\mathbf{x}, t). \quad (\text{F.7})$$

The same procedure applies for the primed potential, so if we assume $n(\mathbf{x}, t)$ is the time-dependent density associated with both potentials we find that

$$\frac{\partial}{\partial t} [\mathbf{j}'(\mathbf{x}, t) - \mathbf{j}(\mathbf{x}, t)] = -\frac{1}{m} \nabla [v'(\mathbf{x}, t) - v(\mathbf{x}, t)] n(\mathbf{x}, t) + \mathbf{Q}'(\mathbf{x}, t) - \mathbf{Q}(\mathbf{x}, t). \quad (\text{F.8})$$

Since $\nabla \cdot \mathbf{j}(\mathbf{x}, t) = \nabla \cdot \mathbf{j}'(\mathbf{x}, t) = -\frac{\partial}{\partial t} n(\mathbf{x}, t)$,

$$\frac{1}{m} \nabla \cdot \{ \nabla [v'(\mathbf{x}, t) - v(\mathbf{x}, t)] n(\mathbf{x}, t) \} = \zeta(\mathbf{x}, t), \quad (\text{F.9})$$

where

$$\zeta(\mathbf{x}, t) = \nabla \cdot [\mathbf{Q}'(\mathbf{x}, t) - \mathbf{Q}(\mathbf{x}, t)]. \quad (\text{F.10})$$

Now I can replace everything in Eq. (F.9) with Taylor expansions about t_0 yielding the equation

$$\frac{1}{m} \sum_{k_1, k_2=0}^{\infty} \nabla \cdot \{ \nabla [v'_{k_1}(\mathbf{x}) - v_{k_1}(\mathbf{x})] n_{k_2}(\mathbf{x}) \} (t - t_0)^{k_1+k_2} = \sum_{k_3=0}^{\infty} \zeta_{k_3}(\mathbf{x}) (t - t_0)^{k_3}. \quad (\text{F.11})$$

Making the replacement $k_1 + k_2 = k$ and matching terms I find an expression relating the Taylor coefficients:

$$\frac{1}{m} \nabla \cdot \left\{ \sum_{k_1=0}^k \nabla [v'_{k_1}(\mathbf{x}) - v_{k_1}(\mathbf{x})] n_{k-k_1}(\mathbf{x}) \right\} = \zeta_k(\mathbf{x}). \quad (\text{F.12})$$

Using this equation I can find an expression for the k 'th Taylor coefficient for the potential in terms of the others:

$$\frac{1}{m} \nabla \cdot \{ \nabla [v'_k(\mathbf{x}) - v_k(\mathbf{x})] n_0(\mathbf{x}) \} = \zeta_k(\mathbf{x}) - \frac{1}{m} \nabla \cdot \left\{ \sum_{k_1=0}^{k-1} \nabla [v'_{k_1}(\mathbf{x}) - v_{k_1}(\mathbf{x})] n_{k-k_1}(\mathbf{x}) \right\}. \quad (\text{F.13})$$

Thus, $\nabla \cdot \{ \nabla [v'_k(\mathbf{x}) - v_k(\mathbf{x})] n_0(\mathbf{x}) \}$ depends on a quantity of the same form for $k_1 < k$. This suggests a proof by induction. I want to show that $v'_k(\mathbf{x}) - v_k(\mathbf{x}) = \text{const.}$ for all k . If I assume that this is true for all $k_1 < k$ and show that this implies that it is true for k , then it is true for all k .

First it is necessary to examine equations of the form

$$\nabla \cdot \{ f(\mathbf{x}) \nabla g(\mathbf{x}) \} = 0. \quad (\text{F.14})$$

Green's theorem implies that

$$\int d^3x f(\mathbf{x}) \nabla g(\mathbf{x}) \cdot \nabla g(\mathbf{x}) = - \int d^3x \nabla \cdot \{ f(\mathbf{x}) \nabla g(\mathbf{x}) \} g(\mathbf{x}) + \int_S [f(\mathbf{x}) g(\mathbf{x}) \nabla g(\mathbf{x})] \cdot d\mathbf{S}, \quad (\text{F.15})$$

where the last term is a surface integral over the boundary of the integration volume. Since we are assuming that Eq. (F.14) holds, the first term on the right-hand side vanishes, because the integrand is zero. If $f(\mathbf{x})$ goes to zero faster than $|g(\mathbf{x}) \nabla g(\mathbf{x})|$ increases as $|\mathbf{x}| \rightarrow \infty$ then the surface term vanishes as well. In principle, this requirement places restrictions on $n_k(\mathbf{x})$ and $v_k(\mathbf{x})$ when we will use this result, but it turns out that for all physical problems, these restrictions do not bear on the problem. So,

$$\int d^3x f(\mathbf{x}) |\nabla g(\mathbf{x})| = 0. \quad (\text{F.16})$$

If $f(\mathbf{x}) \geq 0$ for all $\mathbf{x}$ then either $f(\mathbf{x})$ must be identically zero, a trivial case that does not interest us, or $\nabla g(\mathbf{x}) = \mathbf{0}$. This means that $g(\mathbf{x}) = \text{const.}$ is the solution to Eq. (F.14).

So, back to the proof. Let us assume that $v'_{k_1}(\mathbf{x}) - v_{k_1}(\mathbf{x}) = \text{const.}$ for all $k_1 < k$. Then the second term on the right-hand side of Eq. (F.13) is zero. What about $\zeta_k(\mathbf{x})$? The operator $\hat{\mathbf{Q}}(\mathbf{x})$ does not depend on the potential at all, so $\zeta(\mathbf{x}, t)$, and thence

$\zeta_k(\mathbf{x})$, only depend on the potential via $|\Psi(t)\rangle$ and its Taylor series

$$|\Psi(t)\rangle = \sum_{k=0}^{\infty} |\Psi_k\rangle (t - t_0)^k. \quad (\text{F.17})$$

Plugging this into the Schrödinger equation and matching terms, I find

$$i\hbar k |\Psi_k\rangle = [\hat{T} + \hat{W}] |\Psi_{k-1}\rangle + \sum_{k_1=0}^{k-1} \int d^3x v_{k_1}(\mathbf{x}) \hat{n}(\mathbf{x}) |\Psi_{k-k_1-1}\rangle. \quad (\text{F.18})$$

That is, the k 'th Taylor coefficient of $|\Psi(t)\rangle$ only depends on $v_{k_1}(\mathbf{x})$ for $k_1 < k$. So if $v'_{k_1}(\mathbf{x}) - v_{k_1}(\mathbf{x}) = \text{const.}$ for all $k_1 < k$, then the following holds

$$\langle \Psi'_k | \hat{\mathbf{Q}}(\mathbf{x}) | \Psi'_k \rangle - \langle \Psi_k | \hat{\mathbf{Q}}(\mathbf{x}) | \Psi_k \rangle = \text{const.} \quad (\text{F.19})$$

Thus,

$$\zeta_k(\mathbf{x}) = \nabla \cdot \sum_{k_1=0}^k [\langle \Psi'_{k_1} | \hat{\mathbf{Q}}(\mathbf{x}) | \Psi'_{k-k_1} \rangle - \langle \Psi_{k_1} | \hat{\mathbf{Q}}(\mathbf{x}) | \Psi_{k-k_1} \rangle] = 0. \quad (\text{F.20})$$

Therefore,

$$\nabla \cdot \{ \nabla [v'_k(\mathbf{x}) - v_k(\mathbf{x})] n_0(\mathbf{x}) \} = 0, \quad (\text{F.21})$$

which I have shown is only satisfied if $v'_k(\mathbf{x}) - v_k(\mathbf{x}) = \text{const.}$, meaning that $v'_k(\mathbf{x}) - v_k(\mathbf{x}) = \text{const.}$ for all k .

If all of the Taylor coefficients of $v'(\mathbf{x}, t) - v(\mathbf{x}, t)$ are constant, then this difference must be equal to a purely time-dependent function. So if $n(\mathbf{x}, t) = n'(\mathbf{x}, t)$, then $v'(\mathbf{x}, t) = v(\mathbf{x}, t) + c(t)$. The converse is also true. If $v'(\mathbf{x}, t) = v(\mathbf{x}, t) + c(t)$, then $|\Psi(t)\rangle$ and $|\Psi'(t)\rangle$ differ by at most a purely time-dependent phase, so $n(\mathbf{x}, t) = n'(\mathbf{x}, t)$. Since this is true, the contrapositive is also true. Namely if $n(\mathbf{x}, t) \neq n'(\mathbf{x}, t)$, then $v'(\mathbf{x}, t) \neq v(\mathbf{x}, t) + c(t)$. This completes the proof of the Runge-Gross theorem.

Though the proof of the Runge-Gross theorem was considerably more complicated than the Hohenberg-Kohn theorem, the effort turns out to be well worth it because

much of the work has already been done to prove a more general theorem, the van Leeuwen theorem [68], which leads directly to a time-dependent analog of the Kohn-Sham scheme in groundstate DFT. If the interaction $\hat{W}$ as well as the external potential $\hat{V}(t)$ is different for $\hat{H}'$, then $\hat{Q}(\mathbf{x})$ will also differ. Then, rather than using Eq. (F.9) to show that $v'(\mathbf{x}, t) = v(\mathbf{x}, t) + c(t)$, it can be used to yield a solution for $v'(\mathbf{x}, t) - v(\mathbf{x}, t)$. The Hamiltonian with external potential $\hat{V}'(t)$ and interaction $\hat{W}'$ will then have the same time-dependent density as the one with $\hat{V}(t)$ and $\hat{W}$. The Runge-Gross theorem is then a particular case of the van Leeuwen theorem when $\hat{W} = \hat{W}'$. The other important particular case is that of $\hat{W}' = 0$. In that case, as long as there is a solution to Eq. (F.9) then there is an external potential for a non-interacting system that will produce the same time-dependent density as the interacting system. This is precisely what is needed to provide a foundation to the time-dependent extension of the Kohn-Sham method.

Appendix G

Derivation of the Kohn-Sham density response

The response equation in Eq. (1.94) is a simple matrix equation that is straightforward to solve. However, the equation for χ^{KS} given in Eq. (1.87) is not useful for actual computation. Instead, there is still analytical work to do. In the derivation that follows, I am following unpublished notes of Adolfo Eguiluz. First notice that the Kohn-Sham groundstate is an eigenstate of the Kohn-Sham Hamiltonian,

$$e^{-\frac{i}{\hbar}\hat{H}_{\text{KS}}t}|\Phi_{0\text{KS}}\rangle = e^{-\frac{i}{\hbar}E_0t}|\Phi_{0\text{KS}}\rangle, \quad (\text{G.1})$$

where I have made the definition $E_0 \equiv \sum_{\mathbf{k}j\sigma} \epsilon_{\mathbf{k}j\sigma}$. In addition to this groundstate, I will introduce a complete set of N-particle states $|\Phi_i\rangle$ that are made of Slater determinants of Kohn-Sham single-particle states. These will also be eigenstates:

$$\hat{H}_{\text{KS}0}|\Phi_i\rangle = E_i|\Phi_i\rangle. \quad (\text{G.2})$$

I can insert the resolution of the identity using this set of eigenstates such that, upon insertion of Eq. (1.84), one of the brackets in the commutator in Eq. (1.87) becomes

$$\begin{aligned} \langle \Phi_{0\text{KS}} | \hat{n}_{(I)}(\mathbf{x}, t) \hat{n}_{(I)}(\mathbf{x}', t') | \Phi_{0\text{KS}} \rangle &= e^{\frac{i}{\hbar} E_0(t-t')} \sum_i e^{-\frac{i}{\hbar} E_i(t-t')} \\ &\times \langle \Phi_{0\text{KS}} | \hat{n}(\mathbf{x}) | \Phi_i \rangle \langle \Phi_i | \hat{n}(\mathbf{x}') | \Phi_{0\text{KS}} \rangle \end{aligned} \quad (\text{G.3})$$

The density operator can be expanded in terms of Kohn-Sham creation and annihilation operators as

$$\hat{n}(\mathbf{x}) = \sum_{\sigma} \sum_{\substack{\mathbf{k}j \\ \mathbf{k}'j'}} \phi_{\mathbf{k}j\sigma}^*(\mathbf{x}) \phi_{\mathbf{k}'j'\sigma}(\mathbf{x}) \hat{d}_{\mathbf{k}j\sigma}^{\dagger} \hat{d}_{\mathbf{k}'j'\sigma}. \quad (\text{G.4})$$

Making such an expansion forces us to consider the matrix element $\langle \Phi_i | \hat{d}_{\mathbf{k}'j'\sigma}^{\dagger} \hat{d}_{\mathbf{k}j\sigma} | \Phi_{0\text{KS}} \rangle$. Let us define $|\Phi_{i'}\rangle = \hat{d}_{\mathbf{k}'j'\sigma}^{\dagger} \hat{d}_{\mathbf{k}j\sigma} |\Phi_{0\text{KS}}\rangle$. $|\Phi_{0\text{KS}}\rangle$ can be written as $\prod_{\mathbf{k}j\sigma}^{\text{occ}} \hat{d}_{\mathbf{k}j\sigma}^{\dagger} |0\rangle$. So, in order for $|\Phi_{i'}\rangle \neq 0$, the state $|\mathbf{k}j\sigma\rangle \equiv \hat{d}_{\mathbf{k}j\sigma}^{\dagger} |0\rangle$ must be occupied in the Kohn-Sham groundstate. Otherwise, $\hat{d}_{\mathbf{k}j\sigma}$ would commute with all of the creation operators in $|\Phi_{0\text{KS}}\rangle$ and annihilate the vacuum. Likewise, $|\mathbf{k}'j'\sigma\rangle$ must be unoccupied. Otherwise, there would be a term $(\hat{d}_{\mathbf{k}'j'\sigma}^{\dagger})^2 = 0$ in the product making up $|\Phi_{0\text{KS}}\rangle$, causing it to vanish. If $\mathbf{k} = \mathbf{k}'$ and $j = j'$, then $|\Phi_{i'}\rangle = |\Phi_{0\text{KS}}\rangle$. Since $[\hat{d}_{\mathbf{k}j\sigma}, \hat{d}_{\mathbf{k}j\sigma}^{\dagger}]_+ = 1$, the state $|\mathbf{k}j\sigma\rangle$ will be eliminated from $|\Phi_{i'}\rangle$. $\hat{d}_{\mathbf{k}'j'\sigma}^{\dagger}$ will create a particle in the state $|\mathbf{k}'j'\sigma\rangle$ in $|\Phi_{i'}\rangle$. Thus, $|\Phi_{i'}\rangle$ is a state with a hole in $|\mathbf{k}j\sigma\rangle$ and a particle in $|\mathbf{k}'j'\sigma\rangle$ with energy $E_{i'} = E_0 + \epsilon_{\mathbf{k}'j'\sigma} - \epsilon_{\mathbf{k}j\sigma}$.

Then I am left with $\langle \Phi_i | \Phi_{i'} \rangle = \delta_{ii'}$. The Kronecker delta will fix the excited states in the first bracket in Eq. (G.3) so that the particle-hole pair must match the particle-hole pair in the second bracket. With this in mind, I find that Eq. (G.3)

becomes

$$\begin{aligned}
\langle \Phi_{0\text{KS}} | \hat{n}_{(I)}(\mathbf{x}, t) \hat{n}_{(I)}(\mathbf{x}', t') | \Phi_{0\text{KS}} \rangle &= \sum_{\sigma} \sum_{\substack{\mathbf{k}j \\ \mathbf{k}'j'}} e^{-\frac{i}{\hbar}(\epsilon_{\mathbf{k}'j'\sigma} - \epsilon_{\mathbf{k}j\sigma})(t-t')} \Theta(\epsilon_{\mathbf{k}'j'\sigma} - E_F) \Theta(E_F - \epsilon_{\mathbf{k}j\sigma}) \\
&\times \phi_{\mathbf{k}j\sigma}^*(\mathbf{x}) \phi_{\mathbf{k}'j'\sigma}(\mathbf{x}) \phi_{\mathbf{k}'j'\sigma}^*(\mathbf{x}) \phi_{\mathbf{k}j\sigma}(\mathbf{x}).
\end{aligned} \tag{G.5}$$

The partner of Eq. (G.5) in the bracket of the commutator in Eq. (1.87) is the same expression with $\mathbf{x}$ and t switched with $\mathbf{x}'$ and t' . I can manipulate the dummy variables in this second term in order to factor out everything but the step functions. Then I am able to recognize that

$$\begin{aligned}
&\Theta(\epsilon_{\mathbf{k}'j'\sigma} - E_F) \Theta(E_F - \epsilon_{\mathbf{k}j\sigma}) - \Theta(\epsilon_{\mathbf{k}j\sigma} - E_F) \Theta(E_F - \epsilon_{\mathbf{k}'j'\sigma}) \\
&= \Theta(E_F - \epsilon_{\mathbf{k}j\sigma}) - \Theta(E_F - \epsilon_{\mathbf{k}'j'\sigma}) = f_{\mathbf{k}j\sigma} - f_{\mathbf{k}'j'\sigma},
\end{aligned} \tag{G.6}$$

where I have made the definition of the occupation numbers $f_{\mathbf{k}j\sigma} \equiv \Theta(E_F - \epsilon_{\mathbf{k}j\sigma})$.

Then the bracket of the commutator becomes

$$\begin{aligned}
\langle \Phi_{0\text{KS}} | [\hat{n}_{(I)}(\mathbf{x}, t), \hat{n}_{(I)}(\mathbf{x}', t')]_- | \Phi_{0\text{KS}} \rangle &= \sum_{\sigma} \sum_{\substack{\mathbf{k}j \\ \mathbf{k}'j'}} e^{-\frac{i}{\hbar}(\epsilon_{\mathbf{k}'j'\sigma} - \epsilon_{\mathbf{k}j\sigma})(t-t')} [f_{\mathbf{k}j\sigma} - f_{\mathbf{k}'j'\sigma}] \\
&\times \phi_{\mathbf{k}j\sigma}^*(\mathbf{x}) \phi_{\mathbf{k}'j'\sigma}(\mathbf{x}) \phi_{\mathbf{k}'j'\sigma}^*(\mathbf{x}) \phi_{\mathbf{k}j\sigma}(\mathbf{x}).
\end{aligned} \tag{G.7}$$

Now, I am actually interested in the Fourier transform of $\chi^{\text{KS}}(\mathbf{x}t, \mathbf{x}'t')$. In order to find the time Fourier transform, I must introduce the following expression for the step function:

$$\Theta(t - t') = -\lim_{\eta \rightarrow 0} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi i} \frac{e^{-i\omega(t-t')}}{\omega + i\eta}. \tag{G.8}$$

Then, after inserting (G.7) and (G.8) into (1.87) and shifting the frequency in the integral I find that

$$\begin{aligned} \chi^{\text{KS}}(\mathbf{x}t, \mathbf{x}'t') &= \lim_{\eta \rightarrow 0} \sum_{\sigma} \sum_{\substack{\mathbf{k}j \\ \mathbf{k}'j'}} \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega(t-t')} \phi_{\mathbf{k}j\sigma}^*(\mathbf{x}) \phi_{\mathbf{k}'j'\sigma}(\mathbf{x}) \\ &\quad \times \frac{f_{\mathbf{k}j\sigma} - f_{\mathbf{k}'j'\sigma}}{\epsilon_{\mathbf{k}j\sigma} - \epsilon_{\mathbf{k}'j'\sigma} + \hbar(\omega + i\eta)} \phi_{\mathbf{k}'j'\sigma}^*(\mathbf{x}') \phi_{\mathbf{k}j\sigma}(\mathbf{x}'). \end{aligned} \quad (\text{G.9})$$

From this I can recognize the time Fourier transform of $\chi^{\text{KS}}(\mathbf{x}t, \mathbf{x}'t')$:

$$\chi^{\text{KS}}(\mathbf{x}, \mathbf{x}'; \omega) = \lim_{\eta \rightarrow 0} \sum_{\sigma} \sum_{\substack{\mathbf{k}j \\ \mathbf{k}'j'}} \phi_{\mathbf{k}j\sigma}^*(\mathbf{x}) \phi_{\mathbf{k}'j'\sigma}(\mathbf{x}) \frac{f_{\mathbf{k}j\sigma} - f_{\mathbf{k}'j'\sigma}}{\epsilon_{\mathbf{k}j\sigma} - \epsilon_{\mathbf{k}'j'\sigma} + \hbar(\omega + i\eta)} \phi_{\mathbf{k}'j'\sigma}^*(\mathbf{x}') \phi_{\mathbf{k}j\sigma}(\mathbf{x}'). \quad (\text{G.10})$$

In addition, lattice periodicity implies $\chi^{\text{KS}}(\mathbf{x} + \mathbf{R}, \mathbf{x}' + \mathbf{R}; \omega) = \chi^{\text{KS}}(\mathbf{x}, \mathbf{x}'; \omega)$ so that $\chi^{\text{KS}}(\mathbf{x}, \mathbf{x}'; \omega)$ can be written

$$\chi^{\text{KS}}(\mathbf{x}, \mathbf{x}'; \omega) = \frac{1}{\Omega} \sum_{\mathbf{q}} \sum_{\mathbf{G}\mathbf{G}'}^{\text{BZ}} e^{i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} \chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}}(\mathbf{q}, \omega) e^{-i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{x}'}. \quad (\text{G.11})$$

Inverting this equation and using Eq. (G.10), I can find an expression for the Fourier transform for the Kohn-Sham response as a function in $\mathbf{q}$ and ω and a matrix in $\mathbf{G}$ and $\mathbf{G}'$:

$$\begin{aligned} \chi_{\mathbf{G}\mathbf{G}'}^{\text{KS}}(\mathbf{q}, \omega) &= \lim_{\eta \rightarrow 0} \frac{1}{\Omega} \sum_{\sigma} \sum_{\substack{\mathbf{k}j \\ \mathbf{k}'j'}} \langle \mathbf{k}, j, \sigma | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} | \mathbf{k}', j', \sigma \rangle \frac{f_{\mathbf{k}j\sigma} - f_{\mathbf{k}'j'\sigma}}{\epsilon_{\mathbf{k}j\sigma} - \epsilon_{\mathbf{k}'j'\sigma} + \hbar(\omega + i\eta)} \\ &\quad \times \langle \mathbf{k}', j', \sigma | e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{x}} | \mathbf{k}, j, \sigma \rangle. \end{aligned} \quad (\text{G.12})$$

I can use a change of variables and the Bloch theorem to simplify the expressions for the matrix element and eliminate one of the sums over $\mathbf{k}$. This amounts to using the

periodic symmetry of the lattice:

$$\begin{aligned}
\langle \mathbf{k}, j, \sigma | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} | \mathbf{k}', j', \sigma \rangle &= \int d^3x \phi_{\mathbf{k}j\sigma}^*(\mathbf{x}) e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} \phi_{\mathbf{k}'j'\sigma}(\mathbf{x}) \\
&= e^{-i(\mathbf{k}+\mathbf{q}-\mathbf{k}')\cdot\mathbf{R}} \langle \mathbf{k}, j, \sigma | e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{x}} | \mathbf{k}', j', \sigma \rangle \implies \mathbf{k}' = \mathbf{k} + \mathbf{q}
\end{aligned}
\tag{G.13}$$

This leads to the final result in Eq. [1.98](#).

Appendix H

Derivation of an expression for the density response in CRPA

In Section 1.5 I showed that the screened interaction in RPA W can be exactly reproduced in CRPA by replacement of the bare interaction v with the constrained screened interaction W^c and the density response χ with the density response in the target χ^t . However, W is not a physical observable, so this correspondence is not so useful. Instead, one would like to have an expression for an observable in terms of W^c and quantities calculated in the target. The natural quantity to look at is the density response. How exactly is χ^t related to χ ? In order to answer this question I will conduct a similar derivation to Eq. (1.118). I will drop the explicit $\mathbf{q}$ and ω arguments to save space, but they are present in exactly the same way as in

Eq. (1.118):

$$\begin{aligned}
\chi &= (\chi^{\text{KS}, \text{t}} + \chi^{\text{KS}, \text{r}}) [\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{t}} - \mathbf{v} \chi^{\text{KS}, \text{r}}]^{-1} \\
&= (\chi^{\text{KS}, \text{t}} + \chi^{\text{KS}, \text{r}}) [(\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{r}})(\mathbf{1} - [\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{r}}]^{-1} \mathbf{v} \chi^{\text{KS}, \text{t}})]^{-1} \\
&= (\chi^{\text{KS}, \text{t}} + \chi^{\text{KS}, \text{r}}) [\mathbf{1} - \mathbf{W}^{\text{c}} \chi^{\text{KS}, \text{t}}]^{-1} [\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{r}}]^{-1} \\
&= \chi^{\text{t}} [\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{r}}]^{-1} + \chi^{\text{KS}, \text{r}} [\mathbf{1} - \mathbf{W}^{\text{c}} \chi^{\text{KS}, \text{t}}]^{-1} [\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{r}}]^{-1} \quad (\text{H.1}) \\
&= \chi^{\text{t}} [\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{r}}]^{-1} + \chi^{\text{KS}, \text{r}} \mathbf{W}^{\text{c}} \chi^{\text{t}} [\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{r}}]^{-1} + \chi^{\text{KS}, \text{r}} [\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{r}}]^{-1} \\
&= [\mathbf{1} + \chi^{\text{KS}, \text{r}} \mathbf{W}^{\text{c}}] \chi^{\text{t}} [\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{r}}]^{-1} + \chi^{\text{KS}, \text{r}} [\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{r}}]^{-1} \\
&= [\mathbf{1} - \chi^{\text{KS}, \text{r}} \mathbf{v}]^{-1} \chi^{\text{t}} [\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{r}}]^{-1} + \chi^{\text{KS}, \text{r}} [\mathbf{1} - \mathbf{v} \chi^{\text{KS}, \text{r}}]^{-1},
\end{aligned}$$

where I have used the identities $[\mathbf{1} - \mathbf{W}^{\text{c}} \chi^{\text{KS}, \text{t}}]^{-1} = \mathbf{1} + \mathbf{W}^{\text{c}} \chi^{\text{t}}$ and $\mathbf{1} + \chi^{\text{KS}, \text{r}} \mathbf{W}^{\text{c}} = [\mathbf{1} - \chi^{\text{KS}, \text{r}} \mathbf{v}]^{-1}$.

If $\chi^{\text{KS}, \text{r}} = 0$, then Eq. (H.1) tells us that $\chi = \chi^{\text{t}}$. This should be expected since if there were no bands in the rest, i.e. if the target were identical to the full system, then $\chi^{\text{KS}, \text{r}} = 0$ and $\chi = \chi^{\text{t}}$ trivially. However, this is not a realistic situation. Instead, consider a target that consists of a set of bands approximately symmetrically placed across the Fermi level with bandwidth ΔE . Then for frequencies given by $\hbar\omega \lesssim \Delta E/2$, $\Im \chi^{\text{KS}, \text{r}} = 0$ because there are no interband transitions at these frequencies for the rest. However, $\Re \chi^{\text{KS}, \text{r}} \neq 0$, because there are likely interband transitions in the rest at higher frequencies that feed back into the real part of $\chi^{\text{KS}, \text{r}}$ at low frequencies due to the Kramers-Kronig relations. Thus, $\chi^{\text{KS}, \text{r}}$ will not vanish except at isolated frequencies and in general $\chi \neq \chi^{\text{t}}$.

This may seem somewhat distressing as we would like to at least be able to describe the low energy excitations using only the target degrees of freedom. I will examine the expression in Eq. (H.1) with this goal in mind by considering only the imaginary part for frequencies satisfying $\hbar\omega \lesssim \Delta E/2$. The imaginary part is the relevant one because the excitation spectrum for density fluctuations is given by $\Im \chi$. In this case, only χ^{t} has a nonzero imaginary part and $\Im \chi$ is given by (restoring the explicit $\mathbf{q}$

and ω arguments)

$$\mathfrak{Im}\chi(\mathbf{q}, \omega) = [\mathbf{1} - \chi^{\text{KS}, \text{r}}(\mathbf{q}, \omega)\mathbf{v}(\mathbf{q})]^{-1} \mathfrak{Im}\chi^{\text{t}}(\mathbf{q}, \omega) [\mathbf{1} - \mathbf{v}(\mathbf{q})\chi^{\text{KS}, \text{r}}(\mathbf{q}, \omega)]^{-1}. \quad (\text{H.2})$$

Seemingly, there is a hope that an excitation at a particular $\mathbf{q}$ and ω in $\mathfrak{Im}\chi$ will correspond to one in $\mathfrak{Im}\chi^{\text{t}}$, though it is more difficult to ascertain the roles of the matrices involving $\chi^{\text{KS}, \text{r}}$.

For ease of discussion, I will consider the approximate expression for which the matrices can instead be treated as scalars. Then Eq. (H.2) becomes

$$\mathfrak{Im}\chi(\mathbf{q}, \omega) = \frac{\mathfrak{Im}\chi^{\text{t}}(\mathbf{q}, \omega)}{(1 - v(\mathbf{q})\chi^{\text{KS}, \text{r}}(\mathbf{q}, \omega))^2}. \quad (\text{H.3})$$

The relation between $\mathfrak{Im}\chi$ and $\mathfrak{Im}\chi^{\text{t}}$ is controlled by a quantity that looks like a simple dielectric function (squared). This is the dielectric function for the polarization $\chi^{\text{KS}, \text{r}}$, which will be exactly like the polarization for an insulator, for the case I am considering. Thus, it will begin at some finite value larger than one at $\omega = 0$ and gradually decrease as ω becomes larger. So, the quantity $(1 - v(\mathbf{q})\chi^{\text{KS}, \text{r}}(\mathbf{q}, \omega))^2$ will likely decrease and modulate a peak in the spectrum calculated in the target compared to a peak for the full system, but a true peak in the spectrum calculated in the full system should have a corresponding peak in the spectrum calculated using only the target degrees of freedom and vice versa.

Appendix I

A Multi-Orbital Two-Particle Self-Consistent Method

The purpose of this appendix is to present work that I pursued to rigorously derive an extension of the two-particle self consistent method for the one-band Hubbard model presented by Vilk and Tremblay [211, 212] to the multi-band Hubbard model. Miyahara, Arita and Ikeda [213] first attempted this extension to multi-orbital systems by forming an ansatz for the renormalized vertices in a manner that resembled the original formulation of Vilk and Tremblay. Part of this appendix is to attempt a justification for this ansatz. I draw heavily on the ideas of Baym and Kadanoff [146, 214] to provide a theoretical foundation.

The equation of motion for the single-particle Green's function in the presence of an arbitrary external field U which couples to the density non-locally in space, but locally in Matsubara time may be written in terms of (Wannier) orbital indices as

$$G_{\nu\nu',\sigma\sigma'}^{-1}(\mathbf{R}\tau; \mathbf{R}'\tau') = G_{0\nu\nu',\sigma\sigma'}^{-1}(\mathbf{R}\tau; \mathbf{R}'\tau') - U_{\nu\nu',\sigma\sigma'}(\mathbf{R}, \mathbf{R}'; \tau) \delta(\tau - \tau') - \Sigma_{\nu\nu',\sigma\sigma'}(\mathbf{R}\tau; \mathbf{R}'\tau'). \quad (\text{I.1})$$

Here, ν is an orbital index that runs within a unit cell, $\mathbf{R}$ is a lattice vector and runs over unit cells, σ is a spin index and τ is Matsubara time. The purpose of the

external field U is to facilitate the derivation on an equation of motion for the two-particle response function. In an *ab initio* downfolding scheme, the Green's function in Eq. (I.1) will apply to the Wannier orbitals which span some downfolded subspace of the full Hilbert space.

The non-interacting Green's function in Eq. (I.1) is defined as

$$G_{0\,\nu\nu',\sigma\sigma'}^{-1}(\mathbf{R}\tau; \mathbf{R}'\tau') \equiv \delta_{\sigma\sigma'}\delta(\tau-\tau') \left[-\hbar \frac{\partial}{\partial \tau} \delta_{\nu\nu'} \delta_{\mathbf{R}\mathbf{R}'} - (t_{\nu\nu'}(\mathbf{R}-\mathbf{R}') - v_{\text{Hxc}\,\nu\nu'}(\mathbf{R}-\mathbf{R}') - \mu \delta_{\nu\nu'} \delta_{\mathbf{R}\mathbf{R}'} \right]. \quad (\text{I.2})$$

Here, the hopping coefficients are defined in terms of the differential operator $\frac{\hbar^2 \nabla^2}{2m} + v(\mathbf{x}) + v_{\text{Hxc}}(\mathbf{x})$ for which we can obtain eigenvalues and eigenvectors from a groundstate DFT calculation. v is the potential due to the nuclei and v_{Hxc} is the Kohn-Sham self-consistent potential. Written explicitly:

$$t_{\nu\nu'}(\mathbf{R}-\mathbf{R}') = \int d^3x w_{\nu}^*(\mathbf{x}-\mathbf{R}) \left[\frac{\hbar^2 \nabla^2}{2m} + v(\mathbf{x}) + v_{\text{Hxc}}(\mathbf{x}) \right] w_{\nu'}(\mathbf{x}-\mathbf{R}'). \quad (\text{I.3})$$

With the hopping coefficients defined this way, we must subtract off the contribution due to the self-consistent potential in the downfolded subspace. In terms of Wannier orbitals this is given by

$$v_{\text{Hxc}\,\nu\nu'}(\mathbf{R}-\mathbf{R}') = \int d^3x w_{\nu}^*(\mathbf{x}-\mathbf{R}) v_{\text{Hxc}}(\mathbf{x}) w_{\nu'}(\mathbf{x}-\mathbf{R}'). \quad (\text{I.4})$$

Also appearing in Eq. (I.1) is the self-energy Σ . The self-energy is defined in the same manner as in Ref. [146] by two equations that it must simultaneously satisfy:

$$\begin{aligned} \sum_{\sigma_1 \nu_1 \mathbf{R}_1} \int_0^{\beta \hbar} d\tau_1 \Sigma_{\nu\nu_1, \sigma\sigma_1}(\mathbf{R}\tau; \mathbf{R}_1\tau_1) G_{\nu_1\nu', \sigma_1\sigma'}(\mathbf{R}_1\tau_1; \mathbf{R}'\tau') \\ = \hbar \sum_{\sigma_1} \sum_{\nu_1 \nu_2 \nu_3} W_{\nu\nu_1 \nu_2 \nu_3, \sigma\sigma_1}^c G_{2\,\nu_3 \nu_1 \nu_2 \nu', \sigma_1 \sigma \sigma_1 \sigma'}(\mathbf{R}\tau, \mathbf{R}\tau; \mathbf{R}\tau^+, \mathbf{R}'\tau') \end{aligned} \quad (\text{I.5})$$

and

$$\begin{aligned}
\sum_{\sigma_1 \nu_1 \mathbf{R}_1} \int_0^{\beta \hbar} d\tau_1 G_{\nu \nu_1, \sigma \sigma_1}(\mathbf{R}\tau; \mathbf{R}_1 \tau_1) \Sigma_{\nu_1 \nu', \sigma_1 \sigma'}(\mathbf{R}_1 \tau_1; \mathbf{R}' \tau') \\
= \hbar \sum_{\sigma_1} \sum_{\nu_1 \nu_2 \nu_3} G_{2 \nu \nu_3 \nu_1 \nu_2, \sigma \sigma_1 \sigma' \sigma_1}(\mathbf{R}\tau, \mathbf{R}' \tau'^-; \mathbf{R}\tau', \mathbf{R}' \tau') W_{\nu_1 \nu' \nu_2 \nu_3, \sigma \sigma_1}^c.
\end{aligned} \tag{I.6}$$

Here G_2 is the two-particle Green's function and W is the static interaction which in our case will be calculated within CRPA. The spin indices allow for the definition of a Hubbard-like interaction where terms that are zero due to the Pauli exclusion principle are eliminated from the Hamiltonian. Written explicitly:

$$W_{\nu_1 \nu_2 \nu_3 \nu_4, \sigma \sigma'}^c = \begin{cases} 0 & \text{if } \sigma = \sigma' \text{ and} \\ & \nu_1 = \nu_3 \text{ or } \nu_2 = \nu_4 \\ \int d^3x d^3x' w_{\nu_1}^*(\mathbf{x}) w_{\nu_2}(\mathbf{x}) W^c(\mathbf{x}, \mathbf{x}') w_{\nu_3}^*(\mathbf{x}') w_{\nu_4}(\mathbf{x}') & \text{otherwise} \end{cases} \tag{I.7}$$

A consistent set of equations of motion can be derived by starting with an approximation to the self-energy. The approximation I will use takes the form

$$\Sigma_{\nu \nu', \sigma \sigma'}(\mathbf{R}\tau; \mathbf{R}' \tau') \approx \delta_{\sigma \sigma'} \delta_{\mathbf{R} \mathbf{R}'} \delta(\tau - \tau') \Sigma_{\nu \nu', \sigma}^{(1)}. \tag{I.8}$$

$\Sigma^{(1)}$ is defined in analogy to the standard Hartree-Fock approximation, but with an effective interaction $\tilde{W}$ which will be different from the “bare” interaction W^c :

$$\begin{aligned}
\Sigma_{\nu \nu', \sigma}^{(1)} \equiv \hbar \sum_{\nu_1 \nu_2} [\tilde{W}_{\nu \nu' \nu_1 \nu_2, \sigma \sigma} G_{\nu_2 \nu_1, \sigma \sigma}(\mathbf{R}\tau; \mathbf{R}\tau^+) + \tilde{W}_{\nu \nu' \nu_1 \nu_2, \sigma \bar{\sigma}} G_{\nu_2 \nu_1, \bar{\sigma} \bar{\sigma}}(\mathbf{R}\tau; \mathbf{R}\tau^+) \\
- \tilde{W}_{\nu \nu_2 \nu_1 \nu', \sigma \sigma} G_{\nu_2 \nu_1, \sigma \sigma}(\mathbf{R}\tau; \mathbf{R}\tau^+)].
\end{aligned} \tag{I.9}$$

We can find an equation for $\tilde{W}$ by plugging Eq. (I.8) into Eq. (I.5):

$$\begin{aligned} & \sum_{\nu_1} \Sigma_{\nu\nu_1,\sigma}^{(1)} G_{\nu_1\nu,\sigma\sigma}(\mathbf{R}\tau; \mathbf{R}\tau^+) \\ &= \hbar \sum_{\sigma_1} \sum_{\nu_1\nu_2\nu_3} W_{\nu\nu_1\nu_2\nu_3,\sigma\sigma_1}^c G_{2\nu_3\nu_1\nu_2\nu,\sigma_1\sigma\sigma_1\sigma}(\mathbf{R}\tau, \mathbf{R}\tau; \mathbf{R}\tau^+, \mathbf{R}\tau^+) \end{aligned} \quad (\text{I.10})$$

This equation can be put into a more attractive form by recognizing $G_{\nu\nu',\sigma\sigma}(\mathbf{R}\tau; \mathbf{R}\tau^+)$ as the Wannier orbital matrix elements of the average spin-density $\langle \hat{n}_{\nu'\nu,\sigma} \rangle$. There is no index for spin or lattice vector because we are concerned here only with densities for which there is no symmetry breaking. I.e. the density is SU(2) symmetric and periodic with the periodicity of the lattice. Also, $G_{2\nu_3\nu_1\nu_2\nu',\sigma_1\sigma\sigma_1\sigma}(\mathbf{R}\tau, \mathbf{R}\tau; \mathbf{R}\tau^+, \mathbf{R}\tau^+)$ can be written as double occupancies, though this must be done carefully since the creation and annihilation operators that appear in G_2 do not always commute. After some algebra Eq. (I.10) can be written as

$$\begin{aligned} & \sum_{\nu_1\nu_2\nu_3} [\tilde{W}_{\nu\nu_1\nu_2\nu_3,\sigma\sigma} \langle \hat{n}_{\nu\nu_1,\sigma} \rangle \langle \hat{n}_{\nu_2\nu_3,\sigma} \rangle + \tilde{W}_{\nu\nu_1\nu_2\nu_3,\sigma\bar{\sigma}} \langle \hat{n}_{\nu\nu_1,\sigma} \rangle \langle \hat{n}_{\nu_2\nu_3,\bar{\sigma}} \rangle - \tilde{W}_{\nu\nu_3\nu_2\nu_1,\sigma\sigma} \langle \hat{n}_{\nu'\nu_1,\sigma} \rangle \langle \hat{n}_{\nu_2\nu_3,\sigma} \rangle] \\ &= \sum_{\nu_1\nu_2\nu_3} \{ (1 - \delta_{\nu\nu_2})(1 - \delta_{\nu\nu_3}) [\delta_{\nu_2\nu_3} W_{\nu\nu_1\nu_2\nu_3,\sigma\sigma} - (1 - \delta_{\nu_1\nu_2}) W_{\nu\nu_3\nu_2\nu_1,\sigma\sigma}^c] \langle \hat{n}_{\nu'\nu_1,\sigma} \rangle \langle \hat{n}_{\nu_2\nu_3,\sigma} \rangle \\ & \quad + W_{\nu\nu_1\nu_2\nu_3,\sigma\bar{\sigma}}^c \langle \hat{n}_{\nu'\nu_1,\sigma} \rangle \langle \hat{n}_{\nu_2\nu_3,\bar{\sigma}} \rangle \}. \end{aligned} \quad (\text{I.11})$$

This is a system of linear equations for $\tilde{W}$, though it must be noted that there are potentially far fewer equations than unknowns, so it is likely that this system would be under-determined. This means that it will probably be the case that if there is a solution then there will be infinitely many solutions, thus we must have some way to pick the solution we want. One simple way to ensure that Eq. (I.11) is satisfied is to match the sums in the right and left hand sides term-by-term. These give the following conditions, which are analogous to the TPSC conditions in the one-band

case:

$$\tilde{W}_{\nu\nu_1\nu_2\nu_3,\sigma\bar{\sigma}} = W_{\nu\nu_1\nu_2\nu_3,\sigma\bar{\sigma}}^c \frac{\langle \hat{n}_{\nu'\nu_1,\sigma} \hat{n}_{\nu_2\nu_3,\bar{\sigma}} \rangle}{\langle \hat{n}_{\nu'\nu_1,\sigma} \rangle \langle \hat{n}_{\nu_2\nu_3,\bar{\sigma}} \rangle}, \quad (\text{I.12})$$

$$\tilde{W}_{\nu\nu_1\nu_2\nu_3,\sigma\sigma} = (1 - \delta_{\nu\nu_2})(1 - \delta_{\nu_1\nu_3})\delta_{\nu_2\nu_3} W_{\nu\nu_1\nu_2\nu_3,\sigma\sigma}^c \frac{\langle \hat{n}_{\nu'\nu_1,\sigma} \hat{n}_{\nu_2\nu_3,\sigma} \rangle}{\langle \hat{n}_{\nu'\nu_1,\sigma} \rangle \langle \hat{n}_{\nu_2\nu_3,\sigma} \rangle}. \quad (\text{I.13})$$

and

$$\tilde{W}_{\nu\nu_3\nu_2\nu_1,\sigma\sigma} = (1 - \delta_{\nu\nu_2})(1 - \delta_{\nu_1\nu_3})(1 - \delta_{\nu_1\nu_2}) W_{\nu\nu_3\nu_2\nu_1,\sigma\sigma}^c \frac{\langle \hat{n}_{\nu'\nu_1,\sigma} \hat{n}_{\nu_2\nu_3,\sigma} \rangle}{\langle \hat{n}_{\nu'\nu_1,\sigma} \rangle \langle \hat{n}_{\nu_2\nu_3,\sigma} \rangle}. \quad (\text{I.14})$$

I will stress here that these conditions are not the only way to satisfy Eq. (I.11). However, they are closely in line with the previous formulation of multi-orbital TPSC, so they may be preferable to some other scheme.

It remains to find equations for the double-occupancies. In principle, they would be the exact values, but in practice they will be approximate. By using the spin-density response function, derived in a way consistent with the approximation in Eq. (I.8), we can find equations for the double occupancies. I will define the spin-density response in the Wannier basis as

$$\chi_{\nu_1\nu_2\nu_3\nu_4,\sigma\sigma'}(\mathbf{R}\tau; \mathbf{R}'\tau') \equiv \frac{\delta \langle \hat{n}_{\nu_1\nu_2,\sigma}(\mathbf{R}\tau) \rangle}{\delta U_{\nu_3\nu_4,\sigma'\sigma'}(\mathbf{R}', \mathbf{R}'; \tau')}. \quad (\text{I.15})$$

There are two ways to compute the functional derivative in Eq. (I.15). One way, using the definition for the single-particle Green's function yields:

$$\chi_{\nu_1\nu_2\nu_3\nu_4,\sigma\sigma'}(\mathbf{R}\tau; \mathbf{R}'\tau') = -\frac{1}{\hbar} \left[\langle \hat{T}_\tau \{ \hat{n}_{\nu_1\nu_2,\sigma}(\mathbf{R}\tau) \hat{n}_{\nu_3\nu_4,\sigma'}(\mathbf{R}'\tau') \} \rangle - \langle \hat{n}_{\nu_1\nu_2,\sigma}(\mathbf{R}\tau) \rangle \langle \hat{n}_{\nu_3\nu_4,\sigma'}(\mathbf{R}'\tau') \rangle \right]. \quad (\text{I.16})$$

From here we can see how the response function will yield the double occupancies. By setting $\mathbf{R} = \mathbf{R}'$ and $\tau = \tau^-$ we obtain the equation:

$$\langle \hat{n}_{\nu_1\nu_2,\sigma} \hat{n}_{\nu_3\nu_4,\sigma'} \rangle = -\hbar \chi_{\nu_1\nu_2\nu_3\nu_4,\sigma\sigma'}(\mathbf{R}\tau; \mathbf{R}\tau^-) + \langle \hat{n}_{\nu_1\nu_2,\sigma} \rangle \langle \hat{n}_{\nu_3\nu_4,\sigma'} \rangle. \quad (\text{I.17})$$

It still remains to find an equation with which we can actually calculate χ . This equation can be obtained by using the equation of motion for G in Eq. (I.1) and a convenient relation, written schematically as $\delta[G^{-1}G]/\delta U = 0$. After some algebra one finds the following:

$$\begin{aligned} \frac{\delta G_{\nu_2\nu_1,\sigma\sigma}(\mathbf{R}\tau; \mathbf{R}\tau^+)}{\delta U_{\nu_3\nu_4,\sigma'\sigma'}(\mathbf{R}', \mathbf{R}'; \tau')} &= G_{\nu_2\nu_3,\sigma\sigma'}(\mathbf{R}\tau; \mathbf{R}'\tau') G_{\nu_4\nu_1,\sigma'\sigma}(\mathbf{R}'\tau'; \mathbf{R}\tau) \\ &+ \sum_{\substack{\nu_5\nu_6 \\ \nu_7\nu_8}} \sum_{\substack{\mathbf{R}_5\mathbf{R}_6 \\ \mathbf{R}_7\mathbf{R}_8}} \sum_{\substack{\sigma_5\sigma_6 \\ \sigma_7\sigma_8}} \int_0^{\beta\hbar} d\tau_5 d\tau_6 d\tau_7 d\tau_8 G_{\nu_2\nu_5,\sigma\sigma_5}(\mathbf{R}\tau; \mathbf{R}_5\tau_5) G_{\nu_6\nu_1,\sigma_6\sigma}(\mathbf{R}_6\tau_6; \mathbf{R}\tau) \\ &\times \frac{\delta \Sigma_{\nu_5\nu_6,\sigma_5\sigma_6}(\mathbf{R}_5\tau_5; \mathbf{R}_6\tau_6)}{\delta G_{\nu_8\nu_7,\sigma_8\sigma_7}(\mathbf{R}_8\tau_8; \mathbf{R}_7\tau_7)} \frac{\delta G_{\nu_8\nu_7,\sigma_8\sigma_7}(\mathbf{R}_8\tau_8; \mathbf{R}_7\tau_7)}{\delta U_{\nu_3\nu_4,\sigma'\sigma'}(\mathbf{R}', \mathbf{R}'; \tau')}. \end{aligned} \quad (\text{I.18})$$

If we assume a form for $\delta\Sigma/\delta G$ similar to the one we would obtain in standard Hartree-Fock, i.e.

$$\frac{\delta \Sigma_{\nu_5\nu_6,\sigma_5\sigma_6}(\mathbf{R}_5\tau_5; \mathbf{R}_6\tau_6)}{\delta G_{\nu_8\nu_7,\sigma_8\sigma_7}(\mathbf{R}_8\tau_8; \mathbf{R}_7\tau_7)} = \hbar \delta_{\mathbf{R}_5\mathbf{R}_6} \delta_{\mathbf{R}_5\mathbf{R}_7} \delta_{\mathbf{R}_5\mathbf{R}_8} \delta_{\sigma_5\sigma_6} \delta_{\sigma_7\sigma_8} \delta(\tau_5 - \tau_6) \delta(\tau_5 - \tau_7) \delta(\tau_5^+ - \tau_8) \Gamma_{\nu_5\nu_6\nu_7\nu_8,\sigma_5\sigma_7}, \quad (\text{I.19})$$

and make the definition

$$\chi_{\nu_1\nu_2\nu_3\nu_4,\sigma\sigma'}^0(\mathbf{R}\tau; \mathbf{R}'\tau') \equiv \hbar G_{\nu_2\nu_3,\sigma\sigma'}(\mathbf{R}\tau; \mathbf{R}'\tau') G_{\nu_4\nu_1,\sigma'\sigma}(\mathbf{R}'\tau'; \mathbf{R}\tau), \quad (\text{I.20})$$

we find an RPA-like equation for χ :

$$\begin{aligned} \chi_{\nu_1\nu_2\nu_3\nu_4,\sigma\sigma'}(\mathbf{R}\tau; \mathbf{R}'\tau') &= \chi_{\nu_1\nu_2\nu_3\nu_4,\sigma\sigma'}^0(\mathbf{R}\tau; \mathbf{R}'\tau') \\ &+ \sum_{\substack{\nu_5\nu_6 \\ \nu_7\nu_8}} \sum_{\mathbf{R}_5} \sum_{\sigma_5\sigma_7} \int_0^{\beta\hbar} d\tau_5 \chi_{\nu_1\nu_2\nu_5\nu_6,\sigma\sigma_5}^0(\mathbf{R}\tau; \mathbf{R}_5\tau_5) \Gamma_{\nu_5\nu_6\nu_7\nu_8,\sigma_5\sigma_7} \chi_{\nu_7\nu_8\nu_3\nu_4,\sigma_7\sigma'}(\mathbf{R}_5\tau_5; \mathbf{R}'\tau'). \end{aligned} \quad (\text{I.21})$$

We must still examine $\delta\Sigma/\delta G$ to see if the assumption in Eq. (I.19) is reasonable. For the approximate self-energy given by Eq. (I.8) and Eq. (I.9), we find the following:

$$\begin{aligned} \frac{\delta\Sigma_{\nu_5\nu_6,\sigma_5\sigma_6}(\mathbf{R}_5\tau_5;\mathbf{R}_6\tau_6)}{\delta G_{\nu_8\nu_7,\sigma_8\sigma_7}(\mathbf{R}_8\tau_8;\mathbf{R}_7\tau_7)} &= \hbar\delta_{\mathbf{R}_5\mathbf{R}_6}\delta_{\mathbf{R}_5\mathbf{R}_7}\delta_{\mathbf{R}_5\mathbf{R}_8}\delta_{\sigma_5\sigma_6}\delta_{\sigma_7\sigma_8}\delta(\tau_5-\tau_6)\delta(\tau_5-\tau_7)\delta(\tau_5^+-\tau_8) \\ &\times (\delta_{\sigma_5\sigma_7}\tilde{W}_{\nu_5\nu_6\nu_7\nu_8,\sigma_5\sigma_5} + \delta_{\bar{\sigma}_5\sigma_7}\tilde{W}_{\nu_5\nu_6\nu_7\nu_8,\sigma_5\bar{\sigma}_5} - \delta_{\sigma_5\sigma_7}\tilde{W}_{\nu_5\nu_8\nu_7\nu_6,\sigma_5\sigma_5}) \\ &+ \delta_{\mathbf{R}_5\mathbf{R}_6}\delta_{\sigma_5\sigma_6}\delta(\tau_5-\tau_6)\sum_{\nu_1\nu_2}\langle\hat{n}_{\nu_1\nu_2}\rangle\frac{\delta[\tilde{W}_{\nu_5\nu_6\nu_1\nu_2,\sigma_5\sigma_5} + \tilde{W}_{\nu_5\nu_6\nu_1\nu_2,\sigma_5\bar{\sigma}_5} - \tilde{W}_{\nu_5\nu_2\nu_1\nu_6,\sigma_5\sigma_5}]}{\delta G_{\nu_8\nu_7,\sigma_8\sigma_7}(\mathbf{R}_8\tau_8;\mathbf{R}_7\tau_7)}. \end{aligned} \quad (\text{I.22})$$

The problem with taking the $\delta\tilde{W}/\delta G$ functional derivatives in Eq. (I.22) is that these parameters depend on the spin-density response, as we can see in Eq. (I.11). This means that calculating these functional derivatives would require us to calculate the three-particle Green's function, which is not practical. We will have to make an ansatz for the second term in Eq. (I.22) that allows us to use Eq. (I.21). I will assume the following:

$$\begin{aligned} \sum_{\nu_1\nu_2}\langle\hat{n}_{\nu_1\nu_2}\rangle\frac{\delta[\tilde{W}_{\nu_5\nu_6\nu_1\nu_2,\sigma_5\sigma_5} + \tilde{W}_{\nu_5\nu_6\nu_1\nu_2,\sigma_5\bar{\sigma}_5} - \tilde{W}_{\nu_5\nu_2\nu_1\nu_6,\sigma_5\sigma_5}]}{\delta G_{\nu_8\nu_7,\sigma_8\sigma_7}(\mathbf{R}_8\tau_8;\mathbf{R}_7\tau_7)} \\ = \delta_{\mathbf{R}_5\mathbf{R}_7}\delta_{\mathbf{R}_5\mathbf{R}_8}\delta_{\sigma_7\sigma_8}\delta(\tau_5-\tau_7)\delta(\tau_5^+-\tau_8)\Xi_{\nu_5\nu_6\nu_7\nu_8}. \end{aligned} \quad (\text{I.23})$$

Notice that I have assumed that Ξ does not depend at all on spin. This will become important later as we attempt to decouple the spin and charge response functions.

Let us now turn our attention to the spin response function. I will define the spin operator as

$$\hat{\sigma}_{\nu\nu'}^i(\mathbf{R}\tau) = \sum_{\sigma_1\sigma_2}\hat{c}_{\nu,\sigma_1}^\dagger(\mathbf{R}\tau)\sigma_{\sigma_1\sigma_2}^i\hat{c}_{\nu',\sigma_2}(\mathbf{R}\tau), \quad i = x, y, z. \quad (\text{I.24})$$

Here $\hat{c}^\dagger$ and $\hat{c}$ are creation and annihilation operators for Wannier orbitals and σ^i is a Pauli matrix. The definition in Eq. (I.24) allows me to write the spin response as

$$\begin{aligned}\chi_{\nu_1\nu_2\nu_3\nu_4}^{\text{sp}}(\mathbf{R}\tau; \mathbf{R}'\tau') &= -\frac{1}{\hbar} [\langle \hat{T}_\tau \{ \hat{\sigma}_{\nu_1\nu_2}^z(\mathbf{R}\tau) \hat{\sigma}_{\nu_3\nu_4}^z(\mathbf{R}'\tau') \} \rangle - \langle \hat{\sigma}_{\nu_1\nu_2}^z(\mathbf{R}\tau) \rangle \langle \hat{\sigma}_{\nu_3\nu_4}^z(\mathbf{R}'\tau') \rangle] \\ &= \chi_{\nu_1\nu_2\nu_3\nu_4, \uparrow\uparrow}(\mathbf{R}\tau; \mathbf{R}'\tau') + \chi_{\nu_1\nu_2\nu_3\nu_4, \downarrow\downarrow}(\mathbf{R}\tau; \mathbf{R}'\tau') \\ &\quad - \chi_{\nu_1\nu_2\nu_3\nu_4, \uparrow\downarrow}(\mathbf{R}\tau; \mathbf{R}'\tau') - \chi_{\nu_1\nu_2\nu_3\nu_4, \downarrow\uparrow}(\mathbf{R}\tau; \mathbf{R}'\tau').\end{aligned}\tag{I.25}$$

Using this equation together with Eq. (I.21) and assuming that there is no SU(2) symmetry breaking, one can show that

$$\begin{aligned}\chi_{\nu_1\nu_2\nu_3\nu_4}^{\text{sp}}(\mathbf{R}\tau; \mathbf{R}'\tau') &= \chi_{\nu_1\nu_2\nu_3\nu_4}^{0\text{ sp}}(\mathbf{R}\tau; \mathbf{R}'\tau') \\ &\quad - \frac{1}{2} \sum_{\substack{\nu_5\nu_6 \\ \nu_7\nu_8}} \sum_{\mathbf{R}_5} \int_0^{\beta\hbar} d\tau_5 \chi_{\nu_1\nu_2\nu_5\nu_6}^{0\text{ sp}}(\mathbf{R}\tau; \mathbf{R}_5\tau_5) U_{\nu_5\nu_6\nu_7\nu_8}^{\text{sp}} \chi_{\nu_7\nu_8\nu_3\nu_4}^{\text{sp}}(\mathbf{R}_5\tau_5; \mathbf{R}'\tau'),\end{aligned}\tag{I.26}$$

where

$$U_{\nu_5\nu_6\nu_7\nu_8}^{\text{sp}} \equiv \Gamma_{\nu_5\nu_6\nu_7\nu_8, \uparrow\downarrow} - \Gamma_{\nu_5\nu_6\nu_7\nu_8, \uparrow\uparrow}.\tag{I.27}$$

Now we see the utility of using the spin response. If we take the ansatz in Eq. (I.23) as given, then the term involving $\delta\tilde{W}/\delta G$ cancels out when calculating U^{sp} . This means that U^{sp} is given by

$$U_{\nu_5\nu_6\nu_7\nu_8}^{\text{sp}} = \tilde{W}_{\nu_5\nu_6\nu_7\nu_8, \uparrow\downarrow} - \tilde{W}_{\nu_5\nu_6\nu_7\nu_8, \uparrow\uparrow} + \tilde{W}_{\nu_5\nu_8\nu_7\nu_6, \uparrow\uparrow}.\tag{I.28}$$

We are close to finding a set of equations that can be solved self-consistently. The spin response is related to the double-occupancies by

$$\chi_{\nu_1\nu_2\nu_3\nu_4}^{\text{sp}}(\mathbf{R}\tau; \mathbf{R}\tau^-) = -\frac{2}{\hbar} [\langle \hat{n}_{\nu_1\nu_2, \uparrow} \hat{n}_{\nu_3\nu_4, \uparrow} \rangle - \langle \hat{n}_{\nu_1\nu_2, \uparrow} \hat{n}_{\nu_3\nu_4, \downarrow} \rangle].\tag{I.29}$$

The problem is that we must know $\langle \hat{n}_\uparrow \hat{n}_\uparrow \rangle$ and $\langle \hat{n}_\uparrow \hat{n}_\downarrow \rangle$ independently from each other. Luckily, we can utilize a consequence of SU(2) symmetry to find $\langle \hat{n}_\uparrow \hat{n}_\downarrow \rangle$ independently.

First I will use the spin raising and lowering operators $\hat{\sigma}^\pm = \hat{\sigma}^x \pm i\hat{\sigma}^y$ to write another spin response function:

$$\begin{aligned}
\chi_{\nu_1\nu_2\nu_3\nu_4}^\pm(\mathbf{R}\tau; \mathbf{R}'\tau') &= -\frac{1}{\hbar} [\langle \hat{T}_\tau \{ \hat{\sigma}_{\nu_1\nu_2}^+(\mathbf{R}\tau) \hat{\sigma}_{\nu_3\nu_4}^-(\mathbf{R}'\tau') \} \rangle - \langle \hat{\sigma}_{\nu_1\nu_2}^+(\mathbf{R}\tau) \rangle \langle \hat{\sigma}_{\nu_3\nu_4}^-(\mathbf{R}'\tau') \rangle] \\
&= -\frac{1}{\hbar} [\langle \hat{T}_\tau \{ \hat{\sigma}_{\nu_1\nu_2}^x(\mathbf{R}\tau) \hat{\sigma}_{\nu_3\nu_4}^x(\mathbf{R}'\tau') \} \rangle - \langle \hat{\sigma}_{\nu_1\nu_2}^x(\mathbf{R}\tau) \rangle \langle \hat{\sigma}_{\nu_3\nu_4}^x(\mathbf{R}'\tau') \rangle \\
&\quad + \langle \hat{T}_\tau \{ \hat{\sigma}_{\nu_1\nu_2}^y(\mathbf{R}\tau) \hat{\sigma}_{\nu_3\nu_4}^y(\mathbf{R}'\tau') \} \rangle - \langle \hat{\sigma}_{\nu_1\nu_2}^y(\mathbf{R}\tau) \rangle \langle \hat{\sigma}_{\nu_3\nu_4}^y(\mathbf{R}'\tau') \rangle \\
&\quad + i \langle \langle \hat{T}_\tau \{ \hat{\sigma}_{\nu_1\nu_2}^y(\mathbf{R}\tau) \hat{\sigma}_{\nu_3\nu_4}^x(\mathbf{R}'\tau') \} \rangle - \langle \hat{\sigma}_{\nu_1\nu_2}^y(\mathbf{R}\tau) \rangle \langle \hat{\sigma}_{\nu_3\nu_4}^x(\mathbf{R}'\tau') \rangle \rangle \\
&\quad - i \langle \langle \hat{T}_\tau \{ \hat{\sigma}_{\nu_1\nu_2}^x(\mathbf{R}\tau) \hat{\sigma}_{\nu_3\nu_4}^y(\mathbf{R}'\tau') \} \rangle - \langle \hat{\sigma}_{\nu_1\nu_2}^x(\mathbf{R}\tau) \rangle \langle \hat{\sigma}_{\nu_3\nu_4}^y(\mathbf{R}'\tau') \rangle \rangle].
\end{aligned} \tag{I.30}$$

Due to SU(2) symmetry, we can shuffle around x , y and z without changing the averages involved. So we recognize that $\chi^\pm = 2\chi^{\text{sp}}$. Additionally,

$$\begin{aligned}
\chi_{\nu_1\nu_2\nu_3\nu_4}^\pm(\mathbf{R}\tau; \mathbf{R}'\tau') &= -\frac{4}{\hbar} \langle \hat{c}_{\nu_1,\uparrow}^\dagger \hat{c}_{\nu_2,\downarrow} \hat{c}_{\nu_3,\downarrow}^\dagger \hat{c}_{\nu_4,\uparrow} \rangle = -\frac{4}{\hbar} \langle \hat{c}_{\nu_1,\uparrow}^\dagger \hat{c}_{\nu_4,\uparrow} \hat{c}_{\nu_2,\downarrow} \hat{c}_{\nu_3,\downarrow}^\dagger \rangle \\
&= \frac{4}{\hbar} [\langle \hat{n}_{\nu_1\nu_4,\uparrow} \hat{n}_{\nu_3\nu_2,\downarrow} \rangle - \delta_{\nu_2\nu_3} \langle \hat{n}_{\nu_1\nu_4,\uparrow} \rangle].
\end{aligned} \tag{I.31}$$

So with Eq. (I.29) together with Eq. (I.31), we will be able to determine both $\langle \hat{n}_\uparrow \hat{n}_\uparrow \rangle$ and $\langle \hat{n}_\uparrow \hat{n}_\downarrow \rangle$.

So now we have a closed set of equations that can be solved self-consistently. To review, we can start out with a guess for G and χ^{sp} . Using Eqs. (I.29) and (I.31) we can find $\langle \hat{n}_\uparrow \hat{n}_\uparrow \rangle$ and $\langle \hat{n}_\uparrow \hat{n}_\downarrow \rangle$. Then we can find a $\tilde{W}$ that satisfies Eq. (I.11). With this we can solve for G self-consistently using the self-energy in Eq. (I.9). $\tilde{W}$ also yields U^{sp} from Eq. (I.28), so we can solve the RPA equation for χ^{sp} in Eq. (I.26). We will actually solve the Fourier transformed version of this equation to find $\chi_{\nu_1\nu_2\nu_3\nu_4}^{\text{sp}}(\mathbf{q}, \omega_n)$. Then by summing over $\mathbf{q}$ and ω_n we can find a new set of $\langle \hat{n}_\uparrow \hat{n}_\uparrow \rangle$ and $\langle \hat{n}_\uparrow \hat{n}_\downarrow \rangle$ and start the self-consistent cycle all over again.

What remains now is to treat the charge part of the method. Similarly to the spin response, the charge response can be written

$$\begin{aligned} \chi_{\nu_1\nu_2\nu_3\nu_4}^{\text{ch}}(\mathbf{R}\tau; \mathbf{R}'\tau') = & \chi_{\nu_1\nu_2\nu_3\nu_4, \uparrow\uparrow}(\mathbf{R}\tau; \mathbf{R}'\tau') + \chi_{\nu_1\nu_2\nu_3\nu_4, \downarrow\downarrow}(\mathbf{R}\tau; \mathbf{R}'\tau') \\ & + \chi_{\nu_1\nu_2\nu_3\nu_4, \uparrow\downarrow}(\mathbf{R}\tau; \mathbf{R}'\tau') + \chi_{\nu_1\nu_2\nu_3\nu_4, \downarrow\uparrow}(\mathbf{R}\tau; \mathbf{R}'\tau'). \end{aligned} \quad (\text{I.32})$$

This relation, together with Eq. (I.21) yields an RPA-like equation for the charge response similar to Eq. (I.26) for the spin response:

$$\begin{aligned} \chi_{\nu_1\nu_2\nu_3\nu_4}^{\text{ch}}(\mathbf{R}\tau; \mathbf{R}'\tau') = & \chi_{\nu_1\nu_2\nu_3\nu_4}^{0\text{ ch}}(\mathbf{R}\tau; \mathbf{R}'\tau') \\ & + \frac{1}{2} \sum_{\substack{\nu_5\nu_6 \\ \nu_7\nu_8}} \sum_{\mathbf{R}_5} \int_0^{\beta\hbar} d\tau_5 \chi_{\nu_1\nu_2\nu_5\nu_6}^{0\text{ ch}}(\mathbf{R}\tau; \mathbf{R}_5\tau_5) U_{\nu_5\nu_6\nu_7\nu_8}^{\text{ch}} \chi_{\nu_7\nu_8\nu_3\nu_4}^{\text{ch}}(\mathbf{R}_5\tau_5; \mathbf{R}'\tau'), \end{aligned} \quad (\text{I.33})$$

where

$$U_{\nu_5\nu_6\nu_7\nu_8}^{\text{ch}} \equiv \Gamma_{\nu_5\nu_6\nu_7\nu_8, \uparrow\downarrow} + \Gamma_{\nu_5\nu_6\nu_7\nu_8, \uparrow\uparrow}. \quad (\text{I.34})$$

The main difference at this stage between the charge and spin parts of the method is that even if we take the ansatz in Eq. (I.23) as given, the term involving $\delta\tilde{W}/\delta G$ does not cancel as it did when determining U^{sp} . This means we must find U^{ch} some other way. This can be accomplished using an equation analogous to Eq. (I.29):

$$\chi_{\nu_1\nu_2\nu_3\nu_4}^{\text{ch}}(\mathbf{R}\tau; \mathbf{R}\tau^-) = -\frac{2}{\hbar} [\langle \hat{n}_{\nu_1\nu_2, \uparrow} \hat{n}_{\nu_3\nu_4, \uparrow} \rangle + \langle \hat{n}_{\nu_1\nu_2, \uparrow} \hat{n}_{\nu_3\nu_4, \downarrow} \rangle + 2\langle \hat{n}_{\nu_1\nu_2, \uparrow} \rangle \langle \hat{n}_{\nu_3\nu_4, \uparrow} \rangle]. \quad (\text{I.35})$$

Finding U^{ch} is now a root finding problem. Now, Eq. (I.35) is a non-linear equation in terms of U^{ch} because U^{ch} itself depends on χ^{ch} . Thus, we will have the best assurances of success in finding a solution if we can pick a first guess that is close to the true answer.

We can be relatively sure that we are close to the solution for which all of the quantities I have discussed are consistent with each other if we start from a place where we know the self-consistent solution. If we are in the non-interacting limit ($W^{\text{c}} = 0$) then we know the answer. Thus, if we start with a weak interaction then we should be

close to the answer and we should be able to find a solution relatively easily. Once we have a new set of self-consistent quantities we can increase the interaction strength a little bit and be confident that we are once again close to a self-consistent solution. Thus, I propose a procedure where we gradually increase the interaction strength and at each step in the process we find the self-consistent set of quantities. I believe this should give greatest confidence in being able to stably find a solution.

Vita

Robert G. Van Wesep was born to Edward P. and Katherine Van Wesep in Grand Rapids, MI, the younger brother to Sarah L. Van Wesep. In Grand Rapids he attended Brookside Elementary School, City Middle School and Ottawa Hills High School, graduating as salutatorian in May 2004. He then entered the College of Literature, Science and the Arts at the University of Michigan in Ann Arbor under a full tuition merit scholarship. There his major field of study was Interdisciplinary Physics with minors in Mathematics and Asian Languages and Cultures (Japanese). He spent one year in the atomic physics laboratory of Prof. Georg Raithel at the University of Michigan and a summer doing computational solid state physics research under J. Woods Halley at the Minnesota Supercomputing Center at the University of Minnesota in the Twin Cities, funded by the National Science Foundation's Research Experiences for Undergraduates (REU) program. He graduated from the University of Michigan with High Distinction in May 2009. After graduation, he spent six months working for Genesia Corp. in Mitaka City, Tokyo as an optical engineer. He then returned to the United States to enter the Ph.D. program in the Department of Physics and Astronomy at the University of Tennessee in Knoxville upon being recruited under the UTK/ORNL Distinguished Graduate Fellowship Program. At first, his research was supervised by Prof. Zhenyu Zhang at UTK/ORNL, then under Dr. G. Malcolm Stocks at ORNL before starting his final dissertation work under Prof. Adolfo G. Eguiluz.