CHARACTER OF THE ELECTRONIC STATES NEAR THE METAL-INSULATOR TRANSITION IN GALLIUM MANGANESE ARSENIDE

BY

ANTHONY RICHARDELLA

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics in the Graduate College of the University of Illinois at Urbana-Champaign, 2009

Urbana, Illinois

Doctoral Committee:

Professor S. Lance Cooper, Chair
Professor Ali Yazdani, Director of Research
Professor Paul M. Goldbart
Professor Mats A. Selen
ABSTRACT

The character of electronic states near the Mott-Anderson metal-insulator transition in the ferromagnetic semiconductor Ga$_{1-x}$Mn$_x$As is studied by cross sectional scanning tunneling microscopy. 200nm thick samples grown by MBE at UCSB with doping levels from 1.5%, close the metal-insulator transition, to 5%, deep into the metallic regime, were studied at 4.2K. The thickness of the samples ensured that the electronic states were 3D in character. Strong spatial fluctuations in the local density of states, as well as corrections in the density of states at the Fermi energy due to electron-electron interactions in the presence of disorder are observed. The correlation length of the local density of states also grows significantly approaching the Fermi energy. These effects persist even in the highest doped samples, suggesting that disorder plays a profound role in the character of the electronic states and that even metallic samples are still close to the metal-insulator transition. These effects are expected to be of importance in moving beyond simple mean-field models of carrier-mediated ferromagnetism that ignore the spatial fluctuations in the density of states. Moreover, as the Mott-Anderson transition is little understood in the 3D materials where both interactions and strong disorder are equally important, these observations may provide important new insight by direct observation of the states involved. Further work is presented here comparing individual defects present in GaMnAs with their properties in a non-magnetic GaAs environment. Mn and other transition metals substituted into the first atomic layer of GaAs by STM manipulation were also studied to observe chemical trends. It is found that the effects of the symmetry breaking by the surface can account for the observed in-gap states.
ACKNOWLEDGEMENTS

In coming to the end of the long journey that has produced this dissertation, I sincerely wish to thank those who shared its path with me, in particular, my fellow graduate students Dale Kitchen and Pedram Roushan with whom I acquired the majority of the data presented here. Their help was invaluable, and learning to work together under long and stressful hours with sometimes strong differences of opinions was a one of the most valuable lessons that I will take from this experience. I also wish to thank the other members of our group: Kenjiro Gomes, Aakash Pushp, Lukas Urban, Colin Parker, Abhay Pasupathy, Jungpil Seo, Brian Zhou, Michael Vershinin, and Shashank Misra whose camaraderie I greatly enjoyed. I also thank Abhay and Shashank for being invaluable sources of knowledge and for their patience in answering my many questions. Many thanks go to my advisor Ali Yazdani for providing me the opportunity to do this work, for his valuable questions and insights, and for continually pushing to look deeper into what was hidden in the data. Lastly I wish to thank my mother and rest of my family for their support and particularly my wife Heidi; if I’ve been able to accomplish anything these past few years it is only because we walked through it together.

I also which to acknowledge the collaborations with Michael Flatté at the University of Iowa and Jian-Ming Tang, now at the University of New Hampshire, for the adatom work, David Awschalom and Shawn Mack at the University of California for the GaMnAs samples, Bob Cava for numerous other samples we had the opportunity to study, Joseph Strocio, and Ravin Bhatt and David Huse for valuable discussions. Funding for this work was provided by US ARO MURI DAAD19-01-1-0541 and W911NF-07-1-0125, NSF DMR-0704314 and the NSF MRSEC program through the Princeton Center for Complex Materials.
In memory of my father
TABLE OF CONTENTS

Chapter 1. Introduction.................................................................................................................. 1

Chapter 2. GaMnAs Background .................................................................................................. 3

2.1. Physics of Doped Semiconductors ....................................................................................... 5

2.1.1. Isolated Shallow and Deep Impurities ............................................................................. 5

2.1.2. Conduction as a Function of Doping Concentration ....................................................... 8

2.1.3. Many Body Effects ........................................................................................................ 13

2.2. History of Ga$_{1-x}$Mn$_x$As ............................................................................................... 15

2.3. Material Properties ............................................................................................................. 16

Chapter 3. Metal-Insulator Transitions ...................................................................................... 24

3.1. Conductivity ...................................................................................................................... 24

3.2. Gap Criteria ..................................................................................................................... 25

3.3. Anderson Localization ...................................................................................................... 26

3.3.1. Scaling Theory ............................................................................................................. 27

3.3.2. Wide Distributions ..................................................................................................... 31

3.3.3. Multifractal Wavefunctions ......................................................................................... 32

3.3.4. Order Parameter ......................................................................................................... 36

3.4. Mott-Anderson Transition and Interactions ..................................................................... 36

Chapter 4. Scanning Tunneling Microscopy .............................................................................. 39

4.1. Tunneling Theory ............................................................................................................ 42
4.2. Practical Considerations ........................................................................................................... 46

Chapter 5. Experimental Procedures ............................................................................................ 51

  5.1. Samples ..................................................................................................................................... 51
  5.1.1. GaAs Samples ..................................................................................................................... 51
  5.1.2. GaMnAs Heterostructure Samples ..................................................................................... 51
  5.2. Sample Preparation ................................................................................................................. 53
  5.3. STM Procedures ...................................................................................................................... 60

Chapter 6. Substitution of Transition Metals into the Surface of GaAs ...................................... 65

  6.1. Substitution of Adatoms .......................................................................................................... 66
  6.2. Isolated and Interacting Mn’s in the GaAs Surface ................................................................. 69
  6.3. Comparison of Transition Metal Species in the Surface ....................................................... 74
  6.4. Discussion and Model of the Effect of the Surface ............................................................... 79

Chapter 7. Identification of Defects in GaMnAs ................................................................. 86

Chapter 8. States near $E_F$ in GaMnAs .................................................................................. 97

  8.1. Experiment ............................................................................................................................... 97
  8.2. Local Density of States as a Function of Doping Concentration ........................................ 98
  8.3. Correlations ............................................................................................................................. 106
  8.4. Multifractal Analysis ............................................................................................................... 110

Chapter 9. Conclusions and Outlook ....................................................................................... 113

References ....................................................................................................................................... 116

Author’s Biography ....................................................................................................................... 120
CHAPTER 1. INTRODUCTION

This thesis is concerned with the physics of the dilute magnetic semiconductor Ga$_{1-x}$Mn$_x$As. In particular, how the states near the Fermi energy are modified by substitution of a large percentage of Mn into the parent GaAs lattice. GaMnAs is a unique compound, in that it is one of only a few semiconductors that exhibit high temperature ferromagnetism. The current record transition temperature ($T_C$), the temperature below which it becomes ferromagnetic, stands close to 185K. (1, 2) Much effort has been expended over the past 20 years since GaMnAs was first discovered to try to raise the transition temperature above room temperature, at which point it could become a technologically revolutionary material, as a material that allowed the seamless integration of conventional semiconductor electronics with electrical control of magnetic domains, common in persistent storage devices such as hard drives, would indeed be revolutionary. “Spintronics” is the name coined to generically describe devices based on such materials, and it has become a veritable buzzword in any discussion of next generation technologies beyond current silicon-based digital electronics. Current mean-field models have had much success describing the properties of GaMnAs and do predict that it should be possible to raise $T_C$ above room temperature.(3) However whether such theories accurately describe the physics at the smallest scales in this disordered material is unclear. For this reason we have done atomically-resolved scanning tunneling microscopy studies of Mn in GaAs, from the isolated limit of single Mn atoms in the surface of GaAs to a range of highly doped ferromagnetic samples near the metal-insulator transition and well into the metallic regime.

The flow of this document starts with background of the basic physics of GaMnAs and its material properties, followed by an overview of the metal-insulator transition in disordered and interacting systems. Details of the experimental procedure are given next. Results begin to be
presented in Chapter 6, starting with the manipulation of transition metal adatoms into the surface of GaAs and an analysis of the effect of the surface on the resulting states. Chapter 7 focuses on anisotropies and defect levels of individual Mn dopants and other defects observed in ferromagnetic GaMnAs and compares them other studies in non-magnetic GaAs. Chapter 8 focuses on our main results about the states near the Fermi energy that mediate ferromagnetism, finding they are strongly affected by electron-electron interactions and by disorder. The maps acquired of the spatial dependence of the local density of states may be the first direct observation of the character of states in a 3D material near the Mott-Anderson transition where both interactions and disorder are crucially important in the crossover from a metal to an insulator.

The studies of magnetic adatoms substituted into the surface of GaAs presented here were conducted in the Materials Research Laboratory at the University of Illinois prior to our research group moving to Princeton University in the fall of 2005. All subsequent studies were done in the Princeton Nanoscale Microscopy Laboratory at Princeton University.
CHAPTER 2. GaMnAs BACKGROUND

Ever since Bardeen, Battain and Shockley’s development of the transistor in the 1940’s semiconductor devices have become an increasingly vital component of modern life. Semiconductors have literally become the material of information, central to how we are able to use and process increasingly vast amounts of data. The transistor’s simple property is that a voltage applied at one place changes the current that comes out somewhere else. This ability to act as a switch, coupled with the relentless technological drive to make smaller and smaller transistors, has allowed the explosion of electronic products we are all now so familiar with. Current designs of computer microprocessors are rapidly approaching feature widths of 32 nanometers ($10^{-9}$ meters) where the channels carrying current can be measured in tens of atoms. At some point the process of shrinking feature sizes must run into hard physical laws where quantum effects begin to dominate. At the same time, semiconductors are not solely responsible for the information revolution. Magnetic media, such as the typical hard drive found in a computer, have provided a means of recording all this data in a way that is stable for many years without a constant power source. As a magnet has a north pole and south pole, information is recorded in the orientations of a huge number of tiny magnetic domains in one of these devices.

In looking for materials that could be the basis for the next generation of electronic devices it has been tempting to dream of the possibility of doing both of these functions, processing and storage, with the same material. Adding magnetic properties to a semiconductor is a particularly tempting way of doing this, which has been accomplished with a class of materials known as Dilute Magnetic Semiconductors (DMS). The DMS’s hold forth the possibility of control of both electronic and magnetic degrees of freedom within a single material. Information encoded in electrical signals could be used to directly change the orientation of persistent magnetic domains.
that retain the information encoded in them long after the electrical signals go away, and the domains could influence passage of electrical current, much like in a traditional transistor. Even more elaborate possibilities exist if the quantum nature of the electron’s inherent spin magnetic moment can be controlled and harnessed. This field falls broadly under the heading of spintronics, where the spin is utilized in a way analogous and complementary to the way the electron charge is currently used in electronics.

Among DMS’s, Ga$_{1-x}$Mn$_x$As has received the most attention due to its relatively high transition temperature and technologically relevant parent compound, GaAs. Though not as widely used as silicon, which is usually used in microprocessors, GaAs finds broad use in high frequency components such as the digital signal processors of cell phones. GaMnAs has been used to build numerous simple proof-of-concept spintronic elements, such as modification of $T_C$ by gating and electrically assisted domain switching. (4-10) Unfortunately, as with the development of any new material technology, there are serious issues which still need to be worked out. Primary among these is that $T_C$, the highest temperature at which the material will retain a persistent magnetic orientation, has remained below room temperature despite improved growth techniques and extensive theoretical efforts (2, 3, 11-16) to explain the mechanism of ferromagnetism and suggest strategies for raising $T_C$. Much of uncertainty centers around the exact nature of the electronic states near the Fermi level that are responsible for mediating the ferromagnetic behavior. In particular, the question is whether the carriers reside in valence band-like states or in a doping induced impurity band. Low temperature Scanning Tunneling Microscopy (STM) with its atomic scale spatial resolution and high energy resolution is well suited for studying this type of problem.

The following sections of this chapter will give some background of semiconductor physics, followed by a brief history of this compound and a more detailed exposition of the issues and open questions that exist to be studied.
2.1. Physics of Doped Semiconductors

Semiconductors are characterized as materials that have an energy gap for excitations of electrons that is larger than that of semimetals and smaller than those of insulators. This band gap arises from the electrons moving in the presence of the periodic atomic potentials of the crystal which produces a fully filled valence band separated in energy from a completely empty conduction band. As such, pure semiconductors are insulators. Nearly all of the useful properties of semiconductors are made possible by the ability to selectively introduce particular defects into the crystal and limit to a high degree the concentrations of undesired defects. There can be many different types of defects in a semiconductor, such as foreign atoms (often called dopants or impurities), vacancies, antisites or interstitials. They are generally classified into two categories, donors or acceptors, depending on whether they add or remove an electron from the crystal. Semiconductors can be made conductive by adding dopants that either add electrons to conduction band or remove them from the valence band (this is often described as adding holes to the valence band). The resulting material is called either “n-type” (n for the negative change of electrons) or “p-type” (for the effective positive change of holes). When doped semiconductors of opposite types are placed next to each and a voltage is applied across them, this type of heterostructure, called a diode, displays rectifying behavior. Three alternating layers form a transistor where current across the device can be turned on or off by a voltage applied to the central layer. It is stacked layers of n and p-type materials that form the basis for semiconductor electronics.

2.1.1. Isolated Shallow and Deep Impurities

Given their importance, the physics of dopants in semiconductors has been well studied. The introduction of an impurity atom in place of one of the host atoms of the crystal creates a perturbation in the potential at that site. The impurity atom is generally not isovalent with the host atom replaced, and its tendency to be a donor or an acceptor can often be determined by its
position to the right or the left of the host atom on the periodic table. The size of the electron binding energy largely determines the physics of the defect state and whether it is described as a shallow or deep dopant. A shallow dopant can be correctly characterized as a perturbation of the host electronic states, while a deep dopant is usually dominated by the atomic physics of the impurity. Shallow dopants are nearly always used when doping semiconductors because they allow the Fermi level to be moved close to one of the band edges and their small binding energy makes it relatively easy to ionize them, allowing the carrier they introduce to escape and participate in the transport process.

An example of this would be Si substituted for Ga in GaAs. GaAs is a III-V semiconductor, so named for its group three element gallium and group five element arsenic on the periodic table. Silicon, being group four, is a donor when substituted for Ga due to its extra valence electron. (Interestingly, Si can also substitute for As where it becomes an acceptor.) In the crystal the Si atom can be viewed as an extra proton and electron at a Ga site. This leads to shallow dopants being called “hydrogenic” because they can be described by a potential,

\[
U(r) = \frac{1}{4\pi\epsilon} \frac{e^2}{r}
\]

which is identical to a hydrogen atom except for using the permittivity of the host material instead of that of free space. This is valid when the radius of the state is much larger than the lattice constant. The solution to the Hamiltonian can be found perturbatively using the basis set of the bands of the material. Since the wavefunction is large compared to the lattice constant, it can be viewed in reciprocal space as consisting primarily of small wavevectors near the bottom of the conduction band, which in GaAs is at the center of the Brillouin zone. The dispersion relation near the bottom of the band can be well approximated as a parabola,
\[ E = \frac{\hbar^2 k^2}{2m^*} \]

where \( m^* \) is the effective mass of carriers in the band. This allows us to solve the problem in complete analogy to the hydrogen atom as,

\[ -\frac{\hbar^2}{2m^*} \Delta \psi - \frac{1}{4\pi \varepsilon} \frac{e^2}{r} \psi = E \psi \]

\[ E = \frac{e^4 m^*}{2(4\pi \hbar)^2} \frac{m^* / m_0}{\varepsilon / \varepsilon_0} \approx 13.6 \text{ eV} \]

\[ a_s = \frac{4\pi \hbar^2}{m^* e^2} = \frac{\varepsilon / \varepsilon_0}{m^* / m_0} a_B \]

which yields a binding energy close to 6meV and Bohr radius of over 100Å. These values are very close to the experimentally measured values for many shallow donors in GaAs. If the binding energy of a dopant can be predicted in this way, it is called an “effective mass” dopant. For acceptors the Bohr radius is much smaller and the binding energy is typically closer to 25meV due to the heavier effective hole-mass. The calculation for acceptors has to take the degeneracy of the valence band maximum into account, and a more exact calculation would also consider the contribution from other nearby bands, such as in the model by Luttinger and Kohn. The basic consequence remains, however, that the wave function of a shallow donor or acceptor can be well described using only the band structure of the energetically nearby bands of the host.

This is not the case for deep impurity states. A state can be characterized as deep if it is located far enough away from the band edge that it will not be thermally ionized at room temperature. The radius of the wave function of deep impurities becomes shorter the further the state is from the band edge and is typically close to a few lattice constants. Thus the potential in the Hamiltonian of a deep state can be described as an atomic potential perturbed by the crystal...
field of the host in which it sits; it does not see a periodic potential. GaAs has a zinc-blende crystal structure with tetrahedral ($T_d$) symmetry. Transition metal impurities are typically deep acceptors in GaAs with electrically active d-levels. In the presence of the tetrahedral crystal field, group theory can be used to predict, based on symmetry, which levels will hybridize with host states and which will not. The fivefold degenerate d-orbitals split into a doubly degenerate “e” symmetric pair ($d_{x^2-y^2}$, $d_{z^2}$) and a “$t_2$” symmetric triplet ($d_{xy}$, $d_{xz}$, $d_{yz}$). The $t_2$ symmetric states are oriented along the host bond directions, and thus will hybridize with the host, while the e symmetric ones point between bonds, and do not hybridize with the host. Thus, depending on the symmetry of the orbital that forms the acceptor state, it may have host-like contributions or it may have none. This will be discussed in detail in Chapter 6 to explain the wave functions observed of transition metals that were substituted into the surface of GaAs using the STM.

Finally, it is not unusual for deep defects to have multiple defect levels and for the degeneracies of the levels to make them prone to Jahn-Teller lattice relaxations or even rebonding, where the physical bonds with neighboring atoms change when the defect transitions between different charge states. Since they are spatially localized, they are broad in k-space, leading to them being very efficient recombination centers for electrons and holes because they relax the momentum conservation requirement. Except in special cases, deep levels are more of a nuisance than a help and can in most cases severely degrade the properties of devices.

**2.1.2. Conduction as a Function of Doping Concentration**

As the concentration of a species of dopant is increased, it begins to have more and more effect on the conducting properties of the semiconductor. The conduction is usually exponentially dependent on temperature, except at the highest concentrations where the material becomes essentially a metal. Upon increasing the concentration from a single isolated impurity, one begins to find a distribution of defect states centered in energy around the isolated impurity energy. As the doping is random, some impurities will be closer than others, and the local environment is
different at each location. The width of the distribution of the energies of the states is often referred to as an “impurity band” in analogy to the band of allowed energies in a crystal. The term is imprecise however because it can describe everything from the distribution of localized states at low concentrations to a quasi-continuous metallic-like distribution at high concentrations, if such a configuration is possible in the material. Also, since the impurities do not form an ordered lattice, momentum is not a good quantum number. Thus the term band should be understood here only as the range of energies and not imply metallic states with well defined crystal momenta. Nevertheless, an analogy to a tight-binding calculation can be used: as the concentration of impurities increases, the distance between them decreases and overlap of their wave functions increases, giving rise to a band of states of finite width. Such a band, arising from sites with one electron each in a crystal, is a half filled metal owing to the spin degeneracy. However at low concentrations the band is very narrow, and each impurity site is hardly modified from its ideal isolated configuration which resembles a hydrogen atom for shallow impurities. Adding a second electron to the site requires an energy associated with the coulomb repulsion between the electrons at the site. This situation is described by the Hubbard model, where there exists a lower band centered around $E_0$ and an upper band around $E_0 + U$. Each band is singly degenerate and the lower one is completely filled. The system is an insulator at zero temperature until the width of the bands increases to the point where the upper and lower bands overlap and the system undergoes a Mott metal-insulator transition. This type of transition will be described in more detail in Chapter 3.

A distinction can be drawn between light doping and heavy doping by whether the material displays finite conductivity at zero temperature, i.e., whether it has crossed the metal insulator transition. A very rough measure of this is given by $N_d^g \geq 1$, that the density is high enough that the wave functions overlap significantly in space. Semiconductors close to this limit are typically called degenerately doped. One practical consequence of high concentrations of
carriers is that screening becomes important. As in a metal, the carriers will redistribute around fluctuations in the potential, tending to cancel them out. This is usually treated with Thomas-Fermi linear screening theory, which for a point potential, like that at the site of a dopant, gives the screened coulomb or Yukawa potential,

\[
\frac{e}{4\pi\epsilon r} e^{-r/r_0}
\]

where \( r_0 \) is the Thomas-Fermi screening radius. As the doping increases, the screening radius decreases, lowering the binding energy of the impurities. Eventually the attractive Coulomb potential is sufficiently screened out that the impurity binding goes to zero and the carriers are free. As we will see, Mn in GaAs sits between the cases of a shallow hydrogenic and deep impurity and has significant short range terms in its defect state Hamiltonian that cannot be screened out. This pushes the concentration at which the metal-insulator transition occurs at in GaMnAs much higher than where it typically occurs at in p-type GaAs.

Testing whether the conductivity is finite close to zero temperature requires experiments in a dilution refrigerator. Much can be learned, however, from the form of the conductivity at finite temperatures as well. There are four distinct regions that can be identified in the resistance vs. temperature curve for a semiconductor.\(^{(18)}\) The first is intrinsic and the rest are extrinsic (in the sense that they depend on impurities in the material). They are:

A. At high temperatures there is the regime of intrinsic conduction due to thermal excitation of carriers across the energy gap \( E_g \) between the conduction and valence bands of the material. This is called “intrinsic” because it does not require the existence of impurities within the material. Because \( E_g \) is typically large, the Boltzmann dependence of the carrier concentration ensures that this mechanism is only important at high temperatures.
B. If the material contains impurities whose defect energy \( E_0 \) is small compared to the temperature, then these will be fully ionized and the carrier concentration is independent of the temperature in this range. The resistance of the sample may decrease with decreasing temperature due to lowered phonon scattering.

C. When the temperature is comparable with \( E_0 \), decreasing the temperature results in a decrease in conductance called “freezing out”, as the carriers become less likely to be thermally ionized and thus become frozen on defect sites.

D. When the temperature is lowered even further, the carriers can no longer access the bands of the material and conduction must occur by carriers jumping between unoccupied defect sites. If there is no compensation of the material (i.e. no acceptor defects in a material doped with donor defects), then there are no unoccupied sites and no conduction can occur. Thus, at least some small degree of compensation is necessary for conduction at low temperature.

A metallic sample should not demonstrate freeze out, as the majority of the carriers are not localized to any particular site. The resistivity of a semiconductor, below the temperature range of intrinsic conduction, can be approximated as,

\[
\rho^{-1} = \rho_1^{-1} e^{-\varepsilon_1/kT} + \rho_2^{-1} e^{-\varepsilon_2/kT} + \rho_3^{-1} e^{-\varepsilon_3/kT}
\]

where there are three separate conductance mechanisms, each with its own resistivity and energy dependence. The first, denoted by the subscript “1”, is excitation of the carriers into the semiconductor bands. The activation energy \( \varepsilon_1 \) is close to the isolated impurity ionization energy \( E_0 \). This is the dominant term in region C above. Disorder plays an important role in the remaining cases, as the impurities are not arranged spatially on a regular lattice. The second term is only observed in cases of low compensation, near the Mott metal insulator transition in the intermediate temperature regime between band and hopping conduction. It corresponds to
exciting carriers into doubly occupied defect states, analogous to the upper Hubbard band. This mechanism has some theoretical complications, due to its dependence on the position of the Fermi level and the width of this upper impurity band, which makes it difficult to confidently ascribe features in the conductivity curve to it. The last term is the hopping conduction term, corresponding to carriers hopping from occupied to vacant impurity sites. As was stated above, at least some sites need to be vacant for this to occur, requiring a degree of compensation. Disorder, whether due to coulomb fields or structural disorder of impurity sites, will cause the energy of the defect level to vary between impurities. The factor containing the activation energy \( \varepsilon_3 \) comes from the fact that for the carrier to jump between sites with different energies it must absorb or emit a phonon, a process which has an exponential dependence on temperature. As the impurity concentration is increased from low numbers, \( \varepsilon_3 \) can increase due to the increased disorder but eventually decreases at higher concentrations as the impurity wave functions overlap more and screening becomes important. Indeed when one reaches the metal insulator transition, this activation energy disappears entirely. \( \varepsilon_3 \) can be taken to be independent of temperature except at the lowest temperature scales. If this is assumed, then the average hopping length is on the order of the average distance between impurities. When the temperature is lowered into the regime of variable range hopping, as discussed below, the activation energy becomes a function of temperature as well.

The conductivity factor in the hopping term is also a function of concentration. For low concentrations, hopping is nearly impossible due to exponential dependence with distance of the wavefunction overlap between two impurities. This leads \( \rho \) to be modeled as,

\[
\rho_3 = \rho_{03} e^{f(N_D)}
\]

where \( f(N_D) \) is a power law function of the concentration of dopant impurities.
The compensation also plays an important role in both factors. $\rho_0$ is likely to be a function of the degree of compensation because the position of the Fermi level within the impurity band will lead to either more or fewer sites that can take part in hopping. The activation energy likewise will affected because the probability of finding sites close in energy will be increased if there are more states near the Fermi level than if there are fewer. In general, one can expect that a moderate amount of compensation which moves the Fermi level toward the maximum density of states of the impurity band will help increase the conduction.

At very low temperatures it cannot be assumed that hopping will only occur between neighboring sites. Rather it may be more advantageous to jump to far away sites whose energy is close to the Fermi energy. This regime is called variable range hopping (VRH) because the hopping distance becomes longer as the temperature is decreased. In this case, the activation energy $\epsilon_3$ can no longer be assumed to be constant as a function of temperature (in fact, it decreases with decreasing temperature). Mott showed that, if the density of states in a small range near the Fermi level can be assumed to be constant, the resistivity in this regime will be given by,

$$\rho(T) = \rho_0 e^{\left[\frac{T_0}{T}\right]^{1/2}}$$

where $T_0 = \frac{\beta}{k_0 \mu a^2}$

where $a$ is the radius within which carriers are localized and $\beta$ is a numerical constant that can be estimated using percolation theory. This gives the well-known $T^{1/4}$ exponential temperature dependence of VRH.

### 2.1.3. Many Body Effects

On the insulating side of the transition, it was shown by Efros and Shklovskii that the long-range nature of the coulomb interaction between impurity sites results in a gap in the density of
states at the Fermi energy.(19) Simply, each impurity has a localized state. For an electron to move, it must jump from its current occupied site to an unoccupied site, leaving a hole behind. The interaction energy of this process is \( \frac{e^2}{4\pi \epsilon r_{ij}} \) which means that states separated by small energies must be very far apart, and thus the density of states must be low. At zero energy (\( E_F \)) there are no states, and the screening length of the system diverges. They then showed that this results in a parabolic gap in the density of states, \( g(E) \propto E^2 \). The half width of the gap in energy in three dimensions is,

\[
\Delta = \frac{e^3 g_0^{1/2}}{(4\pi \epsilon)^{3/2}}
\]

and is determined by the value of the unperturbed density of states \( g_0 \). Beyond an energy of \( \Delta \) above or below \( E_F \) the density of states is not affected. They also showed that in the regime of variable range hopping this would change the temperature dependence from Mott’s \( \frac{1}{4} \) power to a \( \frac{1}{2} \) power given by,

\[
\rho(T) = \rho_0 e^{\left[\frac{T_1}{T}\right]^{1/2}}
\]

\[
T_1 = \frac{\beta e^2}{4\pi \epsilon a}
\]

where again \( a \) is the radius within which carriers are localized and \( \beta \) is a numerical constant.

On the other side of the transition, Altshuler and Aronov showed that, even in a weakly disordered metal, the exchange part of electron-electron interaction results in a \( \sqrt{E} \) dip in the density of states but that the minimum is finite.(20) As the system becomes more metallic, the disorder should be increasingly screened out and the size of the effect should diminish. The crossover between these two regimes was studied by M. Lee who found that, close to the MIT, the functional shape of the gap could only be determined at small energies before a regime with a
common functional form was entered at energies further from $E_F$.\(^{(21)}\) In three dimensions, this leads to a $1/\sqrt{T}$ increase in the resistance at low temperatures. The temperature dependence of these effects in different dimensions, and of quantum localization described in Chapter 3, are reviewed by Lee and Ramakrishnan in Ref (22).

Lastly, carrier-carrier and carrier-dopant (ion) interactions at high doping concentrations also cause a narrowing of the bulk bandgap, sometimes called “bandgap renormalization.” This is an effect of energy lowering due to screening. Data showing the size of the effect in n and p-type GaAs, as well as a discussion of its relation to interpretation of Burstein-Moss (band filling) effect in optical absorption measurements, are given in Ref (23).

### 2.2. History of $\text{Ga}_{1-x}\text{Mn}_x\text{As}$

Discovery of the III-V dilute magnetic semiconductors was preceded by work on II-VI materials, such as $\text{Cd}_{1-x}\text{Mn}_x\text{Te}$. Various materials could display ferromagnetic or antiferromagnetic order and spin glass behavior. Co-doping was necessary to introduce carriers due to Mn being isovalent with the cations it substitutes. Carrier concentration dependent ferromagnetism was first displayed in $\text{PbSnMnTe}$ and was interpreted in terms of an RKKY-like interaction between carriers and local Mn moments.\(^{(24)}\) The observed $T_C$ was close to 4K.

The first InMnAs films were grown in 1989 at IBM’s T. J. Watson Research Center using molecular beam epitaxy (MBE).\(^{(25)}\) Melt grown Mn doped InAs and GaAs had been made much earlier, but due to the low solubility of Mn, only concentrations in the $1e18$ to $1e19$ cm$^3$ range were possible, where the material is still a paramagnetic insulator. From these it was known that Mn substitutes for Ga and is a moderately deep acceptor with a binding energy of about 113meV. Reduced MBE growth temperatures were used to grow films of the material under non-equilibrium conditions. In these conditions the substitutional Mn is metastable and much higher concentrations could be achieved. These first films were n-type instead of p-type due to a large
number of compensating defects, and higher growth temperatures resulted in metallic MnAs precipitates. Nevertheless, improvements growing these samples came quickly, and p-type samples with $T_C$’s of 110K were made within a decade. Much of this early work was pioneered by Hideo Ohno’s group at Tohoku University in Japan.

A mean-field theory of carrier mediated ferromagnetism based on a simplification of the RKKY-model called the Zener model was introduced by Tomasz Dietl in 2000. It predicted that $T_C$’s above room temperature were possible if Mn concentrations above 10% could be reached. Shortly after this it was shown that post-growth annealing of the samples at temperatures slightly below the growth temperature lead to much improved $T_C$’s. This was explained as thermal diffusion of compensating Mn interstitials to the surface of the material which allowed transition temperatures as high as 159K. In specially made delta-doped samples, where a single monolayer of Mn-doped GaAs is grown in a 2D hole accumulation layer, $T_C$’s as high as 250K have been claimed. Current efforts in “bulk” GaMnAs have pushed the doping concentration as high as 12.5% with a current record $T_C$ of 185K. This falls significantly short of room temperature and has prompted a rethinking of mechanisms responsible for ferromagnetism. Pictures stressing the importance of an impurity band have become increasingly popular after it was suggested that optical absorption measurements strongly supported that case. However, this has been contested on the grounds of temperature dependent conductivity, described above, where a freeze out of carriers excited into the valence band could not be identified. This is a fundamental question that will be addressed in Chapter 8.

### 2.3. Material Properties

Low temperature MBE growth of high quality GaMnAs films is usually done at a temperature close to 200° C. This allows the incorporation of high concentrations on Mn into the crystall without phase segregation into MnAs clusters. As with low temperature grown GaAs, the low temperature growth is prone to the creation of arsenic antisite defects (As sitting in a Ga site).
Figure 2.1: Schematic of the band structure of the direct band gap semiconductor GaAs giving the room temperature gap values. The zero temperature gap is $E_g \approx 1.52$ eV. The conduction band is singly degenerate and s-like while the p-like valence band consists of degenerate heavy $|J=3/2, J_z=\pm 3/2>$ and light hole bands $|J=3/2, J_z=\pm 1/2>$. There is also a spin split off band $|J=1/2, J_z=\pm 1/2>$ lower in energy because of the strong spin-orbit coupling present in the material. ("Band structure and carrier concentration of Gallium Arsenide (GaAs)" Sept 12th 2009 [http://www.ioffe.ru/SVA/NSM/Semicond/GaAs/bandstr.html])
Additionally, a large fraction of the Mn is known to end up at interstitial positions in the lattice, especially at Mn concentrations above a few percent. Both As antisites and Mn interstitials are double donor defects, which compensate the holes introduced by two substitutional Mn acceptors. As the Currie temperature is strongly dependent on the hole concentration, such defects are especially problematic. The sensitivity of the material quality to the parameters used during the MBE growth is discussed in (2).

Post-growth annealing can remove some of the interstitial Mn by causing them to diffuse to the surface. The energy barrier for interstitial diffusion is found to be about 0.7eV, while the energy for other processes such as $\text{Mn}_i + \text{Ga}_i \rightarrow \text{Mn}_3 \text{Ga}_i$ is significantly larger. Additional evidence for this mechanism is that thinner films tend to need to be annealed for shorter periods of time to reach their optimal $T_C$ due to the shorter diffusion path to the surface. The properties of the films are usually monitored during annealing because other processes that eventually begin to lower $T_C$ win out as the rate of out-diffusion of interstitials slows. Thinner films tend to have higher $T_C$’s in general, and it is unclear if it is a more 2D electronic structure that is beneficial or whether growth of thicker layers tends to produce more defects.

The Mn impurity state is an acceptor with an ionization energy of 113meV, several times higher than that of an effective mass acceptor (~25meV). The band structure of GaAs, shown in Figure 2.1, is typical of a zincblende structure III-V semiconductor. It is a direct band gap material with a 4s derived conduction band and a 4p derived degenerate valence band. The degeneracy of the valence band is partially lifted by spin-orbit coupling into a doubly degenerate valance band (heavy and light holes) and a spin split-off band. The 0K band gap is 1.52 eV and the spin split-off band is 0.35 eV below the valence band maximum. Mn ($4s^23d^5$) substitutes for the valence $3^+ \text{Ga}$. From ESR measurements, it was shown that the Mn configuration in GaAs is $\text{Mn}^{2+}3d^5$ plus a weakly bound host-like hole. The total core spin is then $S = 5/2$ from Hund’s rules. The total angular momentum of the bound hole state is $j = 3/2$ from the heavy hole band,
Tetrahedral Crystal Field Splitting

Figure 2.2: Crystal field splitting of atomic d-level in tetrahedral $T_d$ symmetry. Such splitting occurs for the Mn 3d level in zinc-blende GaAs. Interaction with the atoms of the crystal pushes the d-orbitals higher in energy. Three of the orbitals point toward atoms of lattice, raising their energy more than the two that are oriented between lattice sites. This splits the five-fold degeneracy of the d-level into a lower energy “e” symmetric doublet and higher energy $t_2$ symmetric triplet. The $t_2$ symmetric orbitals hybridize with the host while the e symmetric states do not.
and the total ground state spin of the neutral acceptor is \( J = 1 \) from the antiferromagnetic coupling between the Mn core spin and hole.

The origin of the antiferromagnetic coupling is essentially a result of level repulsion between the Mn d-levels and valence band states with the same spin orientation. A simplified model of this is explained in Figure 2.3 and discussed in more detail in Chapter 6. The acceptor state can be understood from the hybridization of the Mn d-orbitals with the GaAs host. The Mn 3d levels are crystal field split by the tetrahedral symmetry of the zincblende crystal. Two levels become \( E \) symmetric (\( d_{x^2-y^2} \)-like) while the other three become \( T_2 \) symmetric (\( d_{xy} \)-like). The \( E \) symmetric levels do not overlap significantly with the tetrahedral \( sp^3 \) orbitals of the nearest neighbor As atoms, while the \( T_2 \) levels do and hybridize with them. The in-gap acceptor level comes from the antibonding states of these hybridized Mn \( T_2 \) orbitals.(32, 33) Essentially this hybridization pushes a state from the valence band with the same spin orientation as the d-orbital into the gap.

Because the Mn is an acceptor, providing one less electron than a native Ga atom, this results in the removal of an electron from the crystal with the same spin alignment as the Mn core-spin and thus an antiferromagnetically aligned hole. The antibonding, and not bonding, nature of the acceptor state gives it a significant host-like character, in common with effective mass acceptors. However, it is important to remember that the binding energy is much larger than an effective mass acceptor and that much of it originates from the p-d hybridization between the Mn d states and the p-like As states that form the valence band edge. Being short ranged, this interaction will not be screened out and pushes the metal-insulator to higher doping concentrations, adding significantly to the disorder in the system.(30)

A review of the material properties of GaMnAs is given in (34). The magnetization as a function of temperature follows the Brillouin function for a homogenous ferromagnet in optimally annealed samples, though significant deviations can occur in samples with large numbers of compensating defects, as seen in many early samples. Coercive fields measured in M
Figure 2.3: Origin of the Mn acceptor state and its antiferromagnetic coupling to the hole. a) The half filled Mn d shell is Hund’s rule split into filled spin up states and empty spin down states. The top of the GaAs valence band is composed mostly of As 4p orbitals. b) The Mn d-orbitals ($t_2$ symmetric orbitals, the $e$ symmetric are non-interacting) hybridize with the GaAs states. The filled d-states push states with the same spin from the valence band into the gap. This hole state is equivalent to a missing spin up electron in the band so the effective coupling between the Mn and hole is antiferromagnetic. c) Adding more Mn’s creates a spin splitting of the valence band into spin polarized majority and minority bands.
vs. H. loops are small and on the order of a few hundred Oe. The easy axis of the film depends on the strain induced by the substrate it is grown on. For tensile strain, such as when grown on GaInAs, the easy axis is out of plane [001], while it is in plane for compressive strain, such as when grown on GaAs. When grown under compressive strain, the in plane anisotropy shows biaxial easy axes along the [100] and [010] directions at low temperature, which shifts above \( T_c/2 \) to a uniaxial [110] easy axis.\(^{(35)}\) This is not expected from the cubic zinc-blende crystal structure. It was further shown that in this higher temperature regime the easy axis can change for [110] to [1\(\bar{1}0\)] at lower hole concentrations and that this could be explained within the Zener model with a slight trigonal distortion.\(^{(36)}\) The cause of such a distortion is still unknown. A nice demonstration of the effect that strain can have on the Mn wave function was done by Yakunin et al. using STM to study an Mn located in proximity to an InAs quantum dot in a GaAs matrix.\(^{(37)}\)

Conductivities are usually on the order of a few hundred \( \Omega^{-1} \) cm\(^{-1} \) at room temperature. The resistivity as a function of temperature generally decreases as the temperature is lowered, except in the vicinity of \( T_c \), where it displays a pronounced hump, and near zero temperature, where it also rises. The hump is usually explained as being due to an increase in scattering caused by spin fluctuations near \( T_c \).\(^{(1)}\) The increase at the lowest temperatures is related to variable range hopping on the insulating side and electron-electron interactions and localization due to disorder on the metallic side.\(^{(38-41)}\) Correlation gaps in the density of states caused by electron-electron interactions have been observed in planar tunnel junction measurements.\(^{(42, 43)}\) It is speculated that these correlation effects are responsible for the large increases seen in tunneling anisotropic magnetoresistance (TAMR) signals that are seen below 4K in such devices.\(^{(42, 44, 45)}\)

The mean free path can be extracted from the low temperature conductivity and is found to be a few nanometers, slightly longer of the average Mn-Mn spacing.\(^{(46)}\) The Zener model assumes that the Fermi wavelength is longer than the inter-Mn spacing, meaning that k\(_{\parallel}\) is close to one. Mott made a general argument, known as the Ioffe-Regel criterion, that a material must be close
the metal-insulator transition when $k_F l \approx 1$ because a state at $E_F$ can hardly be described as a plane-wave if it is scattered every wavelength. Although he was wrong about the details of how the transition takes place, the criterion serves as a useful rule of thumb. The low value implies that transport in GaMnAs is highly diffusive and that quantum localization effects could be important. Based on this, it has been suggested that GaMnAs is always close the metal-insulator transition even at dopings $> 5\%$, which is usually described as metallic.\(^{(16)}\) It is often claimed, however, that in optimally annealed samples with high $T_c$’s the disorder is lowered to the point where it is qualitatively unimportant. It is clear from mesoscopic measurements of universal conductance fluctuations that the phase coherence lengths at low temperatures can be surprisingly long for a ferromagnetic material (up to 100nm at 100 mK) and the predominant dephasing mechanism is electron-electron interaction which gives a weak $L_\phi \propto 1/\sqrt{T}$ temperature dependence in one dimensional wires.\(^{(39, 47, 48)}\) Magnetic field dependent weak localization effects have also been confirmed, though the sign of the effect, whether it is weak localization or weak anti-localization, is in dispute.\(^{(39, 40)}\) First attempts at understanding this theoretically have been made as well.\(^{(49)}\) Thus it is reasonable to assume that electron-electron interactions and quantum interference effects due to disorder play an important role in the physics of GaMnAs and should have consequences for the character of states observed near the Fermi level.
CHAPTER 3. METAL-INSULATOR TRANSITIONS

3.1. Conductivity

At the beginning of a discussion of how a material goes from being a metal to an insulator, it is perhaps good to discuss some basic definitions. A metal is often described as a material that obeys Ohm’s law, $V = IR$, which describes how the application of a voltage across the material results in the flow of an electrical current. A current is defined as the rate of transport of charge carriers from one location to another. A current is inherently a non-equilibrium process. When we discuss conduction in a material, we almost always mean processes that occur very close to equilibrium, such that the properties of the ground state of the system can be used to describe the physics. Ohm’s law is in fact an application of linear response theory; when a system is lightly perturbed from equilibrium, it responds linearly. The current density induced by an electric field in a homogeneous system can be written as,

$$J_{\alpha}(q, \omega) = \sum_{\beta} \sigma_{\alpha\beta}(q, \omega) E_{\beta}(q, \omega)$$

where the conductivity tensor $\sigma_{\alpha\beta}$ is a complex function of the many-body ground state of the system $\psi$, given by the Kubo formula,

$$\sigma_{\alpha\beta}(q, \omega) = \frac{1}{\omega} \int_{0}^{\infty} dt \ e^{i\omega t} \langle \psi | [j_{\alpha}^{\dagger}(q, t), j_{\beta}(q, 0)] | \psi \rangle + \frac{n_{e} e^{2}}{m \omega} i \delta_{\alpha\beta}$$

The conductivity tensor can be viewed as the factor in the first order term of a Taylor expansion of the current calculated around zero field. It contains a retarded two-particle Green’s function which can be written as a two-particle correlation function describing, for instance, an electron-hole excitation. The excitation raises an electron above the Fermi level and leaves a hole
in the Fermi sea behind. Both particles are changed and the conduction of the system is a function both of the applied external electric field and of the internal field created by the excitation. Thus, it is important to realize that conduction is a (very complicated) two-particle process, described by a two particle Green’s function. In the theories of the metal-insulator transition that follow, these are often formulated as single-electron problems, which is equivalent to saying that the motion of the electron and hole is completely uncorrelated and that the problem is equivalent to adding (or creating) an extra electron in the system and describing its motion alone. This is often sufficient to describe the physics, but it should be remembered that in some cases there can be important consequences to conduction not being solely defined by single particle quantities like the density of states.

3.2. Gap Criteria

Why some materials are insulators in essentially due to interactions, either between electrons and the ions of the lattice or because of electron-electron interactions (Mott insulators). The simplest examples are insulators that arise from band theory, where there is a forbidden region in energy between a completely filled valence band and a completely empty conduction band. This gap in the single-particle density of states results from the electrons moving in the presence of the periodic potential created by the ions of the lattice. From Bloch’s theorem it can be shown that there are certain energy ranges where no solutions exist to the relation between the wave vector of the electron wave function and the periodicity of the crystal. These energy ranges are called gaps, where no electronic states exist in a homogeneous crystal. At finite temperature, if the gap in energy is small, as in semiconductors, there can be some conduction due to thermal excitation of carriers across the gap, but at zero temperature the material will be an insulator. In contrast, if the bands overlap in energy, the material may not be a particularly good conductor, as in a semimetal, but it will still be a conductor at zero temperature. In the case of Mott insulators, a gap opens up
for single particle excitations because of electron-electron interactions, and in the absence of pairing (superconductivity, see warning above) the material will also be an insulator.

### 3.3. Anderson Localization

In 1958 Anderson proposed another way in which a material could become an insulator: if it contains a sufficient amount of disorder.(51) Considering what happens in an impurity band, such as in a randomly doped semiconductor, he proposed a model where each site has a random potential ($\varepsilon_j$) taken from a uniform distribution of width $W$ and where electrons can hop between neighboring sites with energy $t$. The Anderson Hamiltonian can be written as,

$$
H = \sum_{\langle jk \rangle} t_{jk} | j \rangle \langle k | + \sum_j \varepsilon_j | j \rangle \langle j |$

$$
\varepsilon_j \in \left[ -\frac{W}{2}, \frac{W}{2} \right]
$$

This model often does not allow analytical solutions, instead Anderson used the criteria of whether the probability of finding a carrier at its initial site would decrease substantially over time, in which case it was delocalized and could diffuse in space, or whether it stayed finite, in which case it was localized. For a sufficiently large ratio of $W$ to $t$, all states become localized. Anderson localization is understood as interference phenomena where multiple coherent scattering paths constructively interfere to localize a particle within a region. (Weak localization can be seen as the precursor effect to this in the metallic regime where constructive backscattering decreases the conductivity.)

In the transition from localized to extended states, Mott argued for the existence of a mobility edge at a certain energy $E_C$. States below it in energy are exponentially localized with a form,

$$
|\psi(r, r_0; E)|^2 \propto e^{-|r-r_0|/\xi(E)}
$$
within a characteristic distance $\xi$, called the localization length, while states above are delocalized with a form like a plane wave. A metal-insulator transition can occur by varying the disorder, and thus the mobility edge, or since the Anderson model is single electron model, changing the electron number to shift the Fermi energy above $E_C$. Unlike a band insulator, there is no gap in the single-particle density of states. Instead the notion of the gap is replaced by the distance of $E_F$ from the mobility edge.

The Anderson transition is a quantum phase transition (QPT) between an insulating phase with localized wave functions and a conductive phase with delocalized character. A change in the nature of the ground state in a quantum phase transition results from a change in relative influence of non-commuting operators in the Hamiltonian as a function of changing some parameter, such as the energy. The metal-insulator transition can be viewed as a competition between the kinetic energy of the carriers which prefers to spread them out as much as possible and the potential energy that is minimized by confining them to a localized region. The transition is only well defined at zero temperature. At finite temperature a crossover region exists near the transition point where the true ground state of the system cannot be distinguished, as shown in Figure 3.1.

### 3.3.1. Scaling Theory

Whether the transition occurs abruptly or continuously was an early point of contention. In formulating the concept of the mobility edge, Mott argued that plane-wave-like extended states at high energy (short wavelength) above the mobility edge would scatter from impurities as expected, but that as the energy is lowered, once the wavelength becomes longer than the mean free path, the state must abruptly become localized. (52) Known as the Ioffe-Regel criterion, $k_F l \approx 1$ is a good measure of when the regime in which quantum interference effects should be important, but the transition is not now believed to be sudden. The work of Abrahams, Anderson, Licciardello and Ramakrishnan in 1979 showed that it could be cast in the scaling theory.
Figure 3.1: Diagram of the Metal-Insulator quantum phase transition. At zero temperature the transition is well defined at a critical value of some parameter $\lambda$, such as the amount of disorder. At finite temperature a crossover region exists where whether the material is an insulator or metal cannot be distinguished. The energy scale of the transition determines how low in temperature an observation must be done at before a distinction can be made. For the Anderson metal-insulator transition, $\lambda_c$ is $E_c$, the mobility edge.
language of a continuous phase transition. Around the same time Wegner created an effective field theory called the non-linear $\sigma$ model (NL$\sigma$M) that allowed a proper renormalization group (RG) analysis of the transition around $d=2$ dimensions. For a metallic sample of side length $L$ the conductance is related to the conductivity by $G(L) = \sigma L^{d-2}$ in $d$ dimensions, while in an insulating sample it should follow the exponential falloff of the wave function, $G \propto e^{-L/\xi}$. Using the dimensionless conductance $g = \frac{G}{(e^2/\hbar)}$ the renormalization group flow equation can be defined,

$$\beta(g) = \frac{d}{d \ln L} \ln g$$

which states how $g$ will change as the system as the system size is increased. If $\beta$ is positive, the conductance will increase, and if it is negative, the conductance will decrease, tending towards an insulator. For large values of disorder where the conductance obeys the exponential form, $\beta$ is negative. From $G(L) = \sigma L^{d-2}$ it is clear that only in dimensions greater than 2 can $\beta$ be positive and the material be a metal. For dimensions $\leq 2$ any small amount of disorder, no matter how small, will localize the entire system. Thus $d=2$ is the lower critical dimension of the problem. In 3D, connecting the two asymptotic limits, there will be some critical value $g^*$ where the metal-insulator transition occurs where $\beta=0$ and the conductance is independent of $L$. This implies that some length scale has diverged and the system has become scale independent. This length scale is identified as the localization length on the insulating side of the transition, $\xi$, and the correlation length on the metallic side. Near the transition point the metal insulator transition should obey the theory of critical phenomena with properties that diverge according to power laws with critical exponents. With respect to the mobility edge, examples for the localization length and conductivity are,
## Critical Exponents of the Non-interacting Anderson Metal Insulator Transition in 3D

<table>
<thead>
<tr>
<th>Exponent</th>
<th>Exponent of</th>
<th>Value</th>
<th>Notes</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\nu)</td>
<td>Correlation length</td>
<td>1.57±0.02</td>
<td>Orthogonal class</td>
<td>(55)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1.43±0.06</td>
<td>Unitary class</td>
<td>(56)</td>
</tr>
<tr>
<td>(s)</td>
<td>Conductivity</td>
<td>(s=(d-2)\nu)</td>
<td>Wegner scaling law</td>
<td>(57, 58)</td>
</tr>
<tr>
<td>(\eta)</td>
<td>Correlation</td>
<td>1.7</td>
<td>(\eta=d-D_2)</td>
<td>(58)</td>
</tr>
</tbody>
</table>

### Multifractal measures of critical wavefunctions in 3D

<table>
<thead>
<tr>
<th>Value</th>
<th>Notes</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(D_2)</td>
<td>1.3±0.05</td>
<td>(58)</td>
</tr>
<tr>
<td>(\alpha_0)</td>
<td>4.03±0.05</td>
<td>Close to (\alpha_0 = d+\epsilon = 4) predicted by 2+\epsilon expansion</td>
</tr>
</tbody>
</table>

Table 3.1: Selected critical exponents in 3D and multifractal measures of critical wavefunctions. Many of these values are found from numerical simulations because of a lack of an analytical solution in 3D.
\[ \xi \propto (E - E_C)^{-\nu} \quad E < E_C \]

\[ \sigma = (E - E_C)^{s} \quad E > E_C \]

As common to critical phenomena, the exponents are often related by scaling relations such as the Wegner scaling law \( s = v(d - 2) \) which states that the exponents of the conductivity and localization length are the same in three dimensions.\(^{(57)}\) The values of the exponents are often the same between widely different materials which is called “universality”. The universality class of the transition in a material is determined by the symmetries of the Hamiltonian, irrespective of the microscopic details. For the Anderson transition, the three most common classes are orthogonal, unitary and symplectic, corresponding to time reversal invariance, broken time reversal symmetry, and broken spin rotation symmetry, though a large number of other classes have been found as well.\(^{(58)}\) In 3D, the exponents found by numerical simulations for the orthogonal and unitary classes are generally similar and are listed in Table 3.1.

### 3.3.2. Wide Distributions

The scaling theory of localization is called a one-parameter theory because only the dimensionless conductance, \( g \), is used to predict whether the material tends towards being an insulator or conductor. Really what is meant by \( g \) is some average value \( \langle g \rangle \), ensemble averaged over multiple realizations of disorder. It was shown by Altshuler \textit{et al.} that close to the transition the fluctuations in the conductance become as large as the conductance itself, and it was argued that the conductance must be described by its entire distribution of possible values, which can take a lognormal form with asymptotically long tails.\(^{(59)}\) Further, it was shown by Lerner that local observables such as the local density of states near the transition can take the same lognormal distribution.\(^{(60)}\) A similar shaped distribution of the LDOS was also shown by Mirlin and Fyodorov.\(^{(61)}\) Such wide non-gaussian distributions that are non-self-averaging imply that all the moments of the distribution and not just the mean need to be considered and that one-
parameter scaling might be an inadequate description. However, it was shown by Shapiro that if a representative value of the conductance could be found that is not influenced by the far tails of the distribution, the one-parameter scaling can be cast in terms of that typical value.\( (62, 63) \) For a lognormal distribution this is naturally the median, given by the geometric mean, \( e^{\ln g} \).

### 3.3.3. Multifractal Wavefunctions

In addition to a lognormal distribution of the local density of states, it was shown that the wave functions near the transition must have a multifractal spatial structure. Wegner calculated inverse participation ratios of eigenfunctions near the transition (which map back to the two-particle Green’s function describing conduction described at the beginning of this chapter) as a way of describing their change in character from localized to extended states.\( (64) \) Consider the inverse participation number (sometimes called the inverse participation ratio),

\[
P = \int d^d r |\psi(r)|^4
\]

which is the integral of the fourth power of the wave function over a sample of volume \( L^d \), where \( d \) is the dimension. The number of sites goes as the volume so if the wave function is extended, such that it is spread out over all sites, \( P = 1/N = L^{-d} \) and goes to 0 in the limit that \( L \to \infty \). Conversely, in the limit that the wave function is localized to a single site, \( P = 1 \). A participation ratio can then be defined as,

\[
p = \left( L^d P \right)^{-1} \sim (a/L)^d
\]

which gives the fraction of the total volume that the wavefunction occupies. (To understand this, consider if the volume of the wave function was approximated as a box of length of “a”, then that would give the last term in the formula above, though this would grossly mischaracterize what the wave function’s spatial dependence near criticality really looks like.) Aoki reasoned that, since
For localized states, at the mobility edge the first extended states must occupy an infinitesimal fraction of the volume and thus be filamentary. Furthermore, being at the critical point there should be no length scales, and therefore the structure should be a self-similar fractal with a dimension \( d^* < d \). This interpretation is too simplistic, however. Near the transition, as discussed, observables have broad distributions. Wegner had in fact calculated a whole set of generalized inverse participation numbers,

\[
P_q = \int d^d r |\psi(r)|^{2q}
\]

which correspond to different moments of the spatial probability distribution of the wave function. Castellani and Peliti interpreted these as showing that the critical wave function is, in fact, a multifractal. The wave function is described by a whole set of fractal dimensions

\[
P_q \propto L^{-\tau_q}
\]

\[
\tau_q = D_q (q - 1)
\]

given by the generalized dimensions \( D_q \), each of which is a critical exponent.

Multifractals are usually described by a multifractal spectrum \( f(\alpha) \) which is the Legendre transform of \( \tau_q \),

\[
f(\alpha) = a \tau(q), \quad a = \frac{d\tau(q)}{dq}
\]

It can be interpreted as follows: Consider a discretized picture of the probability of finding an electron at each site in the lattice \( r_i \) given by \( |\psi(r_i)|^2 = p_i \). The generalized inverse participation number can then be written as,
\[ P_q = \sum_{i}^{N} p_i^q = N \sum_{p} P(p)p^q = L^d \int_{0}^{1} P(p)p^q \, dp \]

where \( P(p) \) is the probability distribution of the site probabilities, i.e., how likely a particular probability is. In the limit that \( L \to \infty \) the integral can be evaluated by the method of steepest descent (a.k.a. saddle point approximation) by assuming that, for each \( q \), a particular \( p = p^* \) dominates. The product of the remaining terms scales with \( L \) by definition as \( \tau_q \),

\[ P_q \sim L^d P(p^*)p^{*q} \quad p^* \propto L^{-\alpha}, \quad L^d P(p^*) \propto L^f \]

\[ \tau(q) = aq - f(\alpha) \]

giving the result above. From \( P(p^*) \propto L^f / L^d \), the number of sites where the wave function intensity falls off locally as \( |\psi(r_i)|^2 \propto L^{-\alpha} \) scales with the power \( f(\alpha) \). Therefore \( f(\alpha) \) can be understood as the fractal dimension of the set composed of those sites. The multifractal is a collection of spatially interpenetrating sets where each set is the group of different sites with the same \( \alpha \) and where each set has a different fractal dimension. The \( f(\alpha) \) curve is concave with a maximum at \( f(\alpha_0) = d \) where \( \alpha_0 = \alpha(q=0) \geq d \). \( \alpha_0 \) describes the scaling of the distribution of typical sites. If \( \tau(q) \) is a linear function of \( q \), then the object is a simple fractal described by a single dimension \( D \). This is the case of a good metal where \( \alpha_0 = D = d \) and \( f(\alpha) \) collapses to a single point at \( f(d) = d \). Some values for multifractal wavefunctions in 3D are listed in Table 3.1, and qualitative plots of typical \( \tau(q) \) and \( f(\alpha) \) curves are shown in Figure 3.2. An efficient method for calculating the \( f(\alpha) \) curve was suggested by Chhabra and Jensen.(67) Many of the properties of multifractals can be interpreted as being analogous to concepts in thermodynamics such as entropy or free energy. For a discussion of this see (68).
Figure 3.2: Qualitative plots of the multifractal measures $\tau(q)$ and $f(\alpha)$. a) $\tau(q)$ describes the scaling of the $q^{th}$ moments of the distribution. If $\tau(q)$ is not a straight line, like the dashed line shown, then there is a continuum of fractal dimensions in the object being measured. b) The distribution of values is described by the $f(\alpha)$ curve called the multifractal spectrum. $\alpha_0$ is the value of $\alpha$ where $f(\alpha)$ is maximal and corresponds to $q=0$. Parts of the curve to the left of $\alpha_0$ are dominated by places where $|\psi|^2$ is larger than the typical value and to the right by places where it is smaller than typical. Due to fact that $f(\alpha)$ is calculated at all points in the space, $D_0=d$. 
### 3.3.4. Order Parameter

Phase transitions are usually accompanied by an order parameter. In the ferromagnetic transition the order parameter is the magnetization, which is 1 if the system is fully polarized and zero in the unordered phase, i.e., above the Curie temperature. It is a natural question to ask what is the order parameter of the Anderson metal-insulator transition? The average local density of states is non-critical at the transition; it does not vanish at the mobility edge. Unlike a band-insulator or Mott insulator, there is not a gap in the single particle density of states that closes at the transition. Building on previous work within the effective field theoretical non-linear $\sigma$ model that showed that an order parameter function could be defined with different analytical properties in the localized and extended phases, Mirlin and Fyoderov showed that this function is related to the distribution of the one-site Green’s functions and therefore the distribution of the local density of states. Janssen has suggested that the typical LDOS does vanish upon the transition to the localized phase and, as the typical conductance could be considered as the proper scaling variable in the one-parameter scaling theory, so the typical local density of states could be considered as the order parameter. Similar considerations led Dobrosavljević to incorporate the typical density of states into a mean-field “typical medium theory” that qualitatively reproduces most of the expected features of the Anderson transition. Because the theory only considers single particle quantities, it easily lends itself to the study of the Anderson transition in the presence of Mott-like electron-electron correlations via dynamic mean field theory (DMFT) as has been done. It should be noted as well that Wigner’s inverse participation ratio also vanishes at the transition and so could also be used as an order parameter, although the LDOS is a more physically meaningful quantity.

### 3.4. Mott-Anderson Transition and Interactions

In real materials it is undoubtedly true that electron-electron correlation effects are important to the physics of disorder induced metal insulator transitions. Such metal-insulator transitions in...
the presence of interactions are called Mott-Anderson transitions. A solid theoretical understanding of them in strongly disordered systems has still not been developed. In general, theoretical models including interactions are treated within Landau’s Fermi liquid theory. As discussed in Chapter 2, disorder leads to a Coulomb gap in insulating materials and a $\sqrt{E}$ shaped cusp in the density of states in disordered metals.\textsuperscript{(19, 20)} In the limit of weak disorder, interactions of this form and weak localization effects have been introduced as phenomenological corrections to the semi-classical Drude formula, in general agreement with experimental evidence.\textsuperscript{(22)} Finkel’stein and others have studied electron-electron interactions in the $d=2+\varepsilon$ expansion of the non-linear $\sigma$ model. In high magnetic fields or systems with magnetic impurities, the DOS is believed to be critical, but the properties in the general non-magnetic case are not as well understood.\textsuperscript{(73, 74)} The implications for the 3D case, where $\varepsilon\rightarrow1$, are somewhat unclear and in general the exponents predicted by this method have not been reproduced in experiments. Because $2+\varepsilon$ expansions are around the lower critical dimension, where any small amount of disorder will cause localization, these can be viewed as probing the weak disorder limit, though arguments can be made for the wider applicability of the results. A review of these types of RG methods and the universality classes that can arise is given in \textsuperscript{(74)}. Another method are dynamic mean field theory approaches based on Dobrosavljević’s typical medium theory, which have allowed a qualitative phase diagram to be drawn.\textsuperscript{(71)} In particular, a two fluid behavior was predicted with sites near the median of the disorder distribution undergoing Mott localization and sites at larger disorder potentials being Anderson localized.\textsuperscript{(72)} It is still too early to assess the correctness of these models, however.

At finite temperatures interactions show up by inelastic collisions that randomize the phase of the electron and destroy the interference pattern. This is known to cut off weak localization effects as the temperature is increased. It can be modeled as a temperature dependent phase coherence length. At scales beyond this length, the multifractal nature of the wave function is
lost. The most important inelastic collisions that occur at finite temperature are with phonons, which can transfer enough energy to an electron to hop between localized states, giving rise to Mott’s variable range hopping conduction. It has been asked whether electron-electron interactions could provide a similar thermal bath that would allow hopping conduction. The answer to this appears to be “no”, however as hopping conduction will cause finite conductivity at finite temperature in a material that is otherwise an insulator at zero temperature, it appears there must be a metal-insulator transition at some non-zero temperature $T_c$ where hopping conduction freezes out, which may be influenced by the electron-electron interaction.\cite{75, 76}

Experimental studies of the transition in 3D have generally been carried out using doped semiconductors. The values found are expected to differ from the values found from non-interacting models because the coulomb interactions are important in real systems. Studies from the early 1980’s using $n$-type phosphorus doped silicon (Si:P) found the scaling exponent of the conductivity $s=0.5$.\cite{77, 78} (Recall that in 3D, $s=v$.) Other studies with semiconductors with significant compensation yielded values closer to $s\approx1$ for both Ge and Si.\cite{79, 80} This became known as the exponents puzzle and brought up the question of whether compensation was necessary for the higher value. However, it was proposed that the critical region in Si:P is very narrow and that the value of $0.5$ was due to numerical errors in fitting the conductivity.\cite{81} More recent work on uncompensated Si:P convincingly showed $s=1.0\pm0.1$ and also measured the dynamical critical exponent as a function of temperature as $z=2.94\pm0.3$.\cite{82} Recently, a larger value of $s=1.6$ was found for $p$-type Si:B, possibly suggesting that it possibly belongs to a different universality class because of stronger spin-orbit scattering, or that the measurements were not taken at sufficiently low temperature.\cite{74, 83} A more extensive review of the experimental literature can be found in \cite{74}.
CHAPTER 4. SCANNING TUNNELING MICROSCOPY

The observation of individual atoms at surfaces was accomplished as early as the 1950’s, preceding the development of the Scanning Tunneling Microscope (STM) by many decades. Two of the earliest methods were developed by Erwin Müller: the Field Emission Microscope (FEM) and the Field Ion Microscope (FIM), which both involved the application of kilovolt electric fields to a sample terminating in a sharp point and the observation of either field emitted electrons or ionized helium atoms that are accelerated from the tip onto a phosphorescent screen. Clearly however, these methods can only be applied to metals capable of withstanding the application of very high fields. For most materials a different approach is needed.

A precursor to the STM was developed by Russel Young in 1971 called the topogafiner. It also used kilovolt electric fields to field emit electrons from a sharp tip but, now the tip was attached to several piezoelectric ceramics that could be used to raster the tip across the surface of a material to be studied that the electrons were incident upon. The field emission current was regulated by a feedback circuit than moved the tip closer to the surface to increase the current or further away to decrease it. The resolution observed was not impressive however, on the order of tens of 10Å vertically and thousands of Angstroms laterally. Young was able to show that if the tip was brought very close to the surface and the voltage was reduced to less than the work function of the material, a deviation from the Fowler-Nordheim equation for the field emission current to a linear current vs. voltage regime took place, indicating that he had entered the metal-vacuum-metal tunneling regime. Unfortunately, insufficient vibration isolation prevented him from being able to carry this work further.

In 1981 Gerd Binnig and Heinrich Rohrer, working at IBM in Zürich, Switzerland, produced the first working STM, solving the vibration problem using a complex magnetic levitation
Two year later they demonstrated the power of the new technique by showing the 7x7 reconstruction of the atoms at the silicon (111) surface in real space for the first time. The STM was the first nondestructive probe that could show real space atomic structure with Angstrom resolution. This opened up a whole new field of study for surfaces and point defects in conducting materials. Even the limitation of conducting materials was removed several years later with the invention of the atomic force microscope (AFM) in 1986 by Binnig, Quate and Gerber. The field of scanned probe techniques has only continued to grow since then, providing a rich wealth of tools today that can be used to study everything from magnetic structure to dynamic processes occurring at surfaces. For this reason, Binnig and Rohrer shared a well-deserved Nobel Prize for their invention in 1986.

There are three primary modes in which STM operates. The first is topography mode where a feedback loop maintains a constant tunnel current as the tip is rastered across the sample surface. If there is a feature on surface, such as an atomic step or an area of high electron density, the feedback loop adjusts the height of the tip above the surface to maintain the setpoint current. In this way an image of topographic details of the surface of the surface can be acquired. The second is spectroscopy mode which, as explained in the following section, allows the measurement of the electronic density of states as a function of both energy and position with Angstrom resolution. STM spectroscopy provides much more flexibility than planar tunnel junction methods by allowing the states associated with individual defects to be studied, instead of averaging over macroscopic areas of the sample. This, especially when combined with angle resolved photoemission spectroscopy (ARPES) measurements, can provide a powerful set of tools for characterizing the electronic properties of a material. The final mode, atom manipulation, is arguably the most fun. The STM tip can be used to create a tunable bond to adatoms, which on some surfaces that allows the tip to move the adatoms from one location to another to create nanoscopic structures. One of the most famous examples of this is shown in Figure 4.1, the
Figure 4.1: Sequence of images showing the creation of a quantum corral using the STM to manipulate iron adatoms on a copper (111) surface. The rings inside the corral are standing wave solutions of the electron wave function which is confined within the potential barrier created by the adatoms. Image originally created by IBM Corporation.
quantum coral, where the wave-like nature of electrons confined to a small area by a “coral” of adatoms can clearly be seen. On other surfaces, such as will be shown on GaAs (110), the process is less controllable but can still be used to create interacting configurations of single atomic defects whose properties can then be studied at will. (87)

The rest of this chapter will cover the theory of the STM tunneling process and the practical considerations that must be addressed in order to do high quality STM studies in general and on semiconductors specifically. Issues, such as the sensitivity to vibrations, will be explained and solutions, which fortunately are (usually) simpler than Binnig and Rohrer’s original magnetic levitation scheme will be detailed.

4.1. Tunneling Theory

In principal, the operation of an STM is simple. The sample to be studied must be conductive. A metal tip, which is electrically biased with respect to the sample, is brought within about a nanometer of the surface, and a current is measured from the tunneling of electrons across the “forbidden” vacuum barrier between the tip and sample. From the Schrödinger equation

\[
\left( \frac{-\hbar}{2m} \frac{d^2}{dz^2} + V(z) \right) \psi(z) = E \psi(z)
\]

in one dimension, a wave function of a state in the forbidden region obeys \( \Delta^2 \psi = \kappa^2 \psi \), where \( \kappa = \sqrt{\frac{2m(V(z) - E)}{\hbar^2}} \). For the case of a square barrier of height \( V \), the probability of observing an electron at a distance \( z \) within the barrier is

\[
|\psi(z)|^2 = |\psi(0)|^2 e^{-2\kappa z}
\]

where \( \psi(0) \) is the amplitude of the wave function at the sample surface. This gives a simple explanation of why the tunnel current is an exponential function of the tip-sample distance and why STM is so sensitive to vibrations that add unwanted modulations to that distance.
For a more realistic analysis, the tunneling process can be understood as a time dependent process of electrons transferring from eigenstates of the tip to eigenstates of the sample or vice versa. From Fermi’s golden rule we find that the transition probability is,

$$\frac{2\pi}{\hbar} |M_{\mu\nu}|^2 \delta(E_v - E_\mu)$$

where $M_{\mu\nu}$ is the matrix element between the states. It is worth noting that the delta function requires conservation of energy and that, in the absence of coupling to vibrational modes and the like, the tunneling is elastic; the energy of the initial and final states must be the same. Since the current will depend on the number of states participating, it is convenient to use the local density of states (LDOS) defined as,

$$\rho(\vec{r}, E) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \sum_{E_n=E-\epsilon/2}^{E_n=E+\epsilon/2} |\psi_n(\vec{r})|^2$$

which is the number of states per unit volume per unit energy. The tunnel current, taking the two spin degrees of freedom into account and taking $k_B T$ small enough to regard the occupancy of states described by the Fermi-Dirac distribution as a step function, is then

$$I = \frac{4\pi e V}{\hbar} \int_0^{eV} \rho_s(\vec{r}, E_F - eV + \epsilon) \rho_T(E_F + \epsilon) |M|^2 d\epsilon$$

where $\rho_s$ and $\rho_T$ are the density of states of the sample and tip respectively. If our tip is a normal free-electron like metal then it has a constant density of states and can be taken out of the integral.

The task now is to evaluate the matrix element $M$, which after some manipulation can be written as a Bardeen integral over some surface separating the states as...
\[
M_{\mu\nu} = -\frac{\hbar}{2m} \int \left( \chi^*_v \nabla \psi_\mu - \psi_\mu \nabla \chi^*_v \right) \cdot dS
\]

where \( \chi \) and \( \psi \) are the wavefunctions of the tip and sample respectively. It was shown by Tersoff and Hamann that, if the eigenstates of the tip that participate in the tunneling consist only of spherical s-wave wave functions, then \( |M|^2 \propto |\psi(r_0)|^2 \). That is, the matrix element is proportional to the exponentially suppressed magnitude of the sample wave function at the position of the tip.\(^{(88)}\) Alternatively, the tunneling probability can be calculated from the barrier height between the tip and sample using the work functions of the materials and the WKB approximation. In either case a term that is exponential in the tip-sample distance is found. In general, it is best to take the matrix element as an unknown that is independent of energy in the range of interest. This allows the matrix element and the density of state of the tip to be treated as constants and the tunnel current is then given by

\[
I \propto \int_0^{E_F} \rho_s(\vec{r}, E_F + \varepsilon) d\varepsilon
\]

where \( V \) is the applied bias voltage. The derivative of the tunnel current is therefore directly proportional to the density of states of the sample. The correctness of this expression has been well proven with the density of states of niobium, as measured by STM fitting the BCS form exactly.\(^{(89)}\)

On a flat surface, therefore, STM topographs can be viewed as maps of the total integrated density of states of the sample between \( E_F \) and the bias voltage. In regions where the local density of states (LDOS) is lower, the tip moves in, and where it is higher, the tip moves back. It must be recognized however that real surfaces are not perfectly flat. Atoms may be at different distances from the tip, either because of an atomic step where a new monolayer of material abruptly begins,
where a foreign adatom is adsorbed to the surface, or where the atoms of the surface are buckled to minimize their energy, etc… Binnig and Rohrer’s work on the Si(111) 7x7 reconstruction is an example of the observation of the physical movement of the locations of the atoms. This structural information is convolved with the electronic information, and care must be taken in disentangling them. When taking topography, the tip typically rasters across the sample at a rate of about 100 nm/second. To keep the tip from making mechanical contact with the sample, the feedback loop must have sufficient bandwidth to adjust the height of the tip quickly enough to maintain the setpoint tunnel current. If the sample surface has large structural changes in height, the scan rate may need to be lowered.

Spectroscopy of the local density of states of a sample can be acquired with the help of a lock-in amplifier. To measure the DOS as a function of energy, the following procedure is used. The feedback loop is turned off so that the tip height remains constant, a small voltage oscillation is added by the lock-in to the bias voltage allowing us to measure $\frac{dI}{dV} \propto \rho_s(eV)$ and the voltage is ramped across the range of energies we wish to measure. The voltage is then returned to the original bias, the feedback loop is turned back on, and the tip can be moved to the next location. This can be done on a grid of many points to map out the density of states as both a function of position and energy. The quantity measured is not the absolute density of states, it is only proportional to it. Nevertheless, this allows one to acquire a host of useful information, such as the energies that defect states exist at inside the gap of a semiconductor, the energy dependence of quasi-particle interference patterns, or how the energy of the pairing gap changes spatially in an inhomogenous superconductor. Lastly, it should be noted that, since the tip height is set independently at each location, it is unavoidable that the spectra at each location are normalized differently. Care must be taken to choose a setup condition where the topographic height
differences between locations are minimal. Usually this means choosing a bias that is high enough that the integrated DOS is much larger that the fluctuations in the DOS between locations.

4.2. Practical Considerations

Binnig et al.’s original demonstration of STM tunneling was done under high vacuum and showed that the apparent tunnel barrier measured changed over the course of minutes due to the accumulation of a layers of contamination on the surface. Though STM has been demonstrated to work in liquids and even air, the modern standard is ultra high vacuum where pressures are at or less than 1e-10 Torr. At these pressures, samples that have been cleaned by repeated cycles of sputtering and annealing or by cleaving the sample to expose a clean face of material can be expected to stay mostly free of contamination for days or longer, depending on the reactivity of the sample.

Cyrogenic temperatures are another step that is typically taken to achieve high quality data. In the derivation of the tunneling current equation above, the Fermi-Dirac distribution function was taken to be a step function because of this. All the measurements reported in this dissertation where acquired at liquid helium temperatures (~4K). The higher the temperature, the more the energy resolution of spectral measurements will be thermally broadened. Even at 4K this effect still exists, but is on the order of about 1 mV. For temperature dependent measurements such as the studies of high temperature superconductors which are done in our research group, these effects need to be explicitly accounted for. The STM tip also tends to be more stable in low temperature studies, due to the suppression of thermal excitations which can cause microscopic reconfigurations of the atoms forming the end of the tip. An additional benefit of low temperatures is the cryopumping effect that the cold parts of the machine near the microscope have. The already low concentration of atoms in the vacuum is lowed further as they stick to the cold surfaces without sufficient thermal energy to escape again. This allows samples in our STM to stay virtually clean of contamination for weeks or longer.
Figure 4.2: The low temperature STM (LTSTM) in Prof. Yazdani’s lab at Princeton University. Former member of our research group Dale Kitchen (below) can be seen working on the tubeline where the microscope head is located. In normal operation the helium cryostat would be mounted to the system where he is standing. Two of the four air legs that provide isolation of the system from vibrations in the floor can be seen. Even acoustic waves can be a noise source that is visible in the tunnel current. For that reason the walls of the room are acoustic shielding panels. The entire room seen here can be floated on massive set of air legs shown in the next figure, and an additional room is built around the outside of this one.
Finally, noise of all sources must be suppressed as much as possible. The most important sources are electrical, vibrational and acoustic. Typical tunnel currents we use are on the order of 100 pA or $1 \times 10^{-10}$ Amperes. At these levels electrical pickup on the tip line and noise introduced onto the system ground can be a real problem. All care must be taken to shield and filter lines supplying signals to the microscope as much as possible. This is especially true for lines leading to the piezoelectrics responsible for moving or scanning the sample where submillivolt noise can lead to detectable noise in the tunnel current. Adding low band pass “pi” or similar filters is a good way of dealing with this. Noise on the system ground can also occur by accidentally creating ground loops where two parts of the system or two pieces of measurement equipment accidentally sit at slightly different potentials due to having two or more paths to ground. As much as possible, all equipment should be powered from the same source, and preferably the same receptacle using a common ground wire.

The importance of vibration isolation has already been alluded to as having ultimately doomed the success of the topogafiner. Binnig and Rohrer’s magnetic levitation approach has now largely been replaced by the use of pneumatic isolators (also called air legs) or active vibration damping systems. Ideally the microscope is also made as rigid as possible so that its subparts move together. If there is one failing of the current Besocke type design(90) of our STM, it is that the sample rests on top of the microscope head by the force of gravity and thus is especially sensitive to vertical vibrations. Vibration isolation of the system is currently achieved in Princeton by several stages, some of which are visible in Figure 4.2. The first is that the microscope head is decoupled from the rest of the system by a set of bellows that allow the assembly to which it is attached to swing freely. Secondly, a set of air legs support the optical table upon which that the system is built. Finally, the floor of the room is made of a massive 30 ton concrete plinth that can be floated on six additional pneumatic isolators. The large mass helps lower the resonance frequency to about 1Hz and, all other frequencies above that should be
strongly damped. Acoustic isolation is achieved by a room built around the system, made of acoustic isolation panels. These panels strongly damp acoustic waves above about 50Hz. For frequencies below these, there is a heavy ciderblock walled enclosure built around the entire structure.

In practice this design has not worked as well as anticipated. One problem that was identified was placing the acoustic room on top of the plinth where the large surface area of the room allowed acoustic noise and air currents to drive the plinth. There are three such rooms in Princeton, and in one case, the acoustic room has been moved off the plinth, which appears to have improved the situation.
Figure 4.3: Picture of the author with standing next to one of six air-legs (red cylinder at the top right) that can be used to float the entire room containing the STM seen in the previous figure. The floor of the room (ceiling in this picture) is a 30 ton concrete slab. The large mass helps make the resonance frequency of this vibration damping system less than 1Hz.
CHAPTER 5. EXPERIMENTAL PROCEDURES

5.1. Samples

5.1.1. GaAs Samples

For the transition metal adatom studies described in later chapters, both degenerately doped n-type and p-type commercially bought GaAs wafers were used. The n-type was doped with silicon at concentrations near $1 \times 10^8 \text{ cm}^{-3}$, while the p-type was zinc doped at higher concentrations between $5 \times 10^8$ and $1 \times 10^9 \text{ cm}^{-3}$ because of the higher concentrations needed for metallic conduction at low temperatures in p-type samples due to the larger effective mass of the valence band. Most of the GaAs wafer used were bought from Wafer Technology (www.wafertech.co.uk). Typically, 2 inch diameter wafers were bought, which could be divided into fifty or more samples of sufficient size to be cleaved in-situ in the STM chamber. Thicker samples provided a larger surface area for the STM tip to approach and so were generally preferred. Samples as large as 2mm by 2mm rods were found to yield large atomically flat areas after cleaving.

5.1.2. GaMnAs Heterostructure Samples

The Ga$_{1-x}$Mn$_x$As samples used were grown by molecular beam epitaxy (MBE) in Prof. Awschalom’s lab at UC Santa Barbara. In this process layers of GaMnAs are grown on a GaAs substrate one atomic layer at a time by having calibrated fluxes in atoms incident on the substrate from pure gallium, arsenic and manganese sources. To study the samples with the STM, they are cleaved in the vacuum, exposing a clean cross-section of the sample to which the STM tip is brought within tunneling range. The process of bringing the tip over the fraction of the cross-section that consists of GaMnAs is described later. Samples with 1%, 1.5%, 3% and 5% Mn concentrations were grown for our research group. These consisted of a 500μm thick, 2 inch diameter, degenerately doped n-type GaAs substrate followed by a 150nm thick MBE grown
1e18 cm$^3$ silicon doped n-type buffer, a 150nm thick MBE grown 2e18 cm$^3$ beryllium doped p-type buffer and finally 200nm of Ga$_{1-x}$Mn$_x$As, where ‘x’ is the doping fraction of Mn used. No capping layer was used. The GaMnAs layer was purposely chosen to be fairly thick in order to make sure its electronic properties were substantially 3D in character and also to aid in locating the layer with the STM. Usually such samples are grown on insulating substrates, but in order to be able to do low temperature STM, metallic substrates had to be used because of the near certainty that the STM tip would first come into tunneling range over the much thicker substrate than the GaMnAs when first approaching the sample. Unfortunately, this also made conductivity measurements of the GaMnAs films impossible.

The samples were grown in a “non-rotated” geometry. Typically, the substrate is rotated during growth so that an even flux of atoms is distributed uniformly across the sample. In the non-rotated geometry used to grow these samples, the Ga and As sources are at opposite angles to surface normal of the substrate, resulting in a higher concentration of Ga atoms incident at one end of the sample and higher As concentration at the other. The Mn source is incident normal to the surface so that the Mn flux should be uniform across the surface. This results in a gradient of growth conditions across the sample with Ga rich growth at one end, As rich growth at the other, and a stoichiometric region somewhere in-between, where the flux of incident atoms was balanced to yield the best material properties. This can be better than growth in the rotated geometry, which is equivalent to growing the entire sample at some composition along the gradient, which may or may not be ideal. Both As rich and Ga rich growth tend to lead to the formation of compensating defects that lower the Curie temperature. The added complication to this procedure is that the region with the best material properties must be identified after growth.

In order to do this, the samples were diced into pieces of sufficient size to perform STM measurements on, and the pieces from the region with the highest $T_C$ were identified by superconducting quantum interference device (SQUID) measurements that were performed
independently at both UCSB and Princeton University. The T_c’s measured for the samples were 27K for the 1.5%, 43K for the 3% and 70K for the 5%. Stoichiometric pieces of the 5% samples were also post-annealed to yield T_c’s as high as 100K, though STM measurements were only successfully performed on a sample that was annealed for only for a few hours to yield a TC of 86K. The annealing was done by placing the samples GaMnAs side up on a clean glass slide on a hot plate in air. Temperatures used were between 195 to 210° C. Due to the thickness of the GaMnAs layer, substantially higher T_c’s would have required much longer annealing times of up to weeks.

5.2. Sample Preparation

The first step in preparing a semiconductor sample was dicing a piece of sufficient size from the full wafer. Samples were usually about 10mm long by 2mm by some thickness up to 2mm. For GaAs the crystalline directions are marked by flats on the wafers according to either the US SEMI or EJ standards. The wafers we used followed the EJ standard where the plane of the major flat is the (-1-10) direction. The natural cleavage planes of GaAs are along the {110} directions. Beyond this relatively little had to be done to prepare GaAs samples for use in the STM. They could be mounted, scribed, placed in the chamber and cleaved.

This was not the case for the GaMnAs samples. Due to the fact that the GaMnAs film was the top layer of the wafer, it was necessary to get an ordered atomically flat surface all the way to the front edge of the sample when cleaving it cross-sectionally in the STM chamber. It is well known that the thicker the sample is, the more likely it is that it will cleave poorly near the front edge, resulting either in a disordered surface or a myriad of closely spaced atomic steps. To avoid this we thinned our samples to less than 100 µm in thickness using an Allied High Tech MultiPrep™ polisher. The samples were mounted GaMnAs side down on a parallel polishing fixture using Crystal Bond 509 and were then thinned on the Multiprep using 6 µm diamond lapping films. After thinning, the samples were unmounted by soaking in acetone. In order to not let a film of
dissolved crystalbond dry on the samples, they were soaked in several baths of clean acetone followed by several baths of isopropanol. The samples were then removed from the bath, GaMnAs side up, and blown with compressed air to ideally cause the isopropanol to delaminate off the surface of sample before the isopropanol could evaporate on the surface. At this thickness the samples are very fragile and must be handled with extreme care to avoid accidentally breaking them.

The samples could then be mounted in the STM sample holder, which consisted of two flat-faced clamp pieces which tightened together with screws to hold the wafer and which then sat inside the larger sample holder assembly. Handling the samples was done with carbon tipped tweezers which would not scratch their surface. To cushion the wafer and to provide good electrical contact, a v-shaped piece of 1mil thick gold film was placed around the half of the sample that sat inside the clamp before tightening the screws. Without this, imperfections in the faces of the clamp pieces tended to cause the samples to break. The cleavage plane of the sample should be as close to parallel with the top face of the clamp as possible. One way to do this was to assemble the clamp on a flat surface and insert the sample into it vertically so that the flat surface aligned them before tightening the screws. A short scribe mark was made with a diamond scribe one edge of the front face of the sample about a half of a millimeter above the clamp piece to form the place where the sample would cleave. This could be done by resting the shank of the scribe on the clamp piece while making the scribing motion in order to get a straight scribe mark. The scribe mark did not need to be very deep. In fact, sometimes it could only be seen under a 2x microscope, and could be very short, sometimes only consisting of a v shaped notch taken out of the edge of the wafer. No consistent pattern could be observed of which type of scribe marks yielded better cleaves than another.

After this the sample was pumped down to ultra high vacuum (UHV) in the load lock and transferred into the chamber. It was found that if the GaMnAs samples were cleaved at this point,
Figure 5.1: GaMnAs sample mounted in the sample holder clamp pieces. A short scribe mark would be placed at one edge of the front face of the wafer. In the chamber the sample would be cleaved by pushing down on the part of the sample extending out of the clamp pieces, causing it to cleave at the scribe mark. A pair of carbon tipped tweezers used to handle the sample while mounting it can be seen in the background.
contamination from absorbates on the front face of the wafer (most likely water) would diffuse over the edge onto the freshly cleaved surface covering a micron or more of the surface with one or more self-assembled of layers of contamination, as shown in Figure 5.2. For a GaAs sample this was not a problem because of the large usable surface away from the edges, but for the GaMnAs samples it would completely cover the heterostructure. This required that the samples be heated in vacuum prior to cleaving to evaporate the adsorbates from off the surfaces that had been exposed to the atmosphere. We tried this in two ways. The first was a heater in the sample holder that consisted of a resistive filament that was used to heat the sample holder for several hours. This method was not as reliable as might have been desired. The temperature of the sample was not known and only part of the sample inside the clamp piece really got heated. The other end that was sticking out probably didn’t warm up much due to its large surface area. This resulted in cleaves where sometimes all or part of surface was still covered in contamination. Baking the sample in the loadlock for over a day at a temperature of ~125C seemed to be a more consistent solution. In both cases the sample could be placed in the SQUID after the run completed, and the Tc of the sample would have increased by a degree or two due to the light annealing effect of the heating.

It is common lore that heating these type samples for long periods of time anneals the scribe mark, leading to worse cleaves. We were not able to verify this trend. It appeared that the thickness of the sample played a much more important role, as well as the length of the sample and the way in which the force was applied during the cleave. We cleaved our samples using a flat face of a manual wobble stick to apply force very close to the end of the sample that was sticking out from the sample holder. Ideally the sample would cleave with a negligible amount of force applied. The length of sample sticking out from the clamp pieces was usually 5mm or more. Attempts to cleave samples with less material sticking out invariably resulted in bad cleaves. This wasted of a large amount of material but appeared to be necessary, possibly by allowing the strain
Figure 5.2: STM topography of contamination at the edge of the wafer caused by insufficient degassing of
the sample prior to cleaving. The first image is 5μm across taken at -2.5V and the second is 500Å taken at 2V.
The composition of the contamination is unknown but it is speculated that it is water. It is clear that it tends to
form an ordered structure on the GaAs (110) surface.
Figure 5.3 Example of the multitude of atomic steps that make up the surface after a bad cleave (5000Å x 5000 Å area, acquired at -2.5V).
Figure 5.4: Optical microscope image of the sample shown in the previous figure. The curved lines running across the surface are characteristic of a bad cleave. A small monocular was used to observe the sample through a viewport on the chamber after cleaving. If any such lines could be observed on the cleaved surface, except in the immediate vicinity of the scribe mark, the sample was generally not inserted into the STM.
to spread out more evenly across the wafer. The other factor which seemed to be important was alignment of the normal of the surface with the direction in which force was applied. If the sample had some angle with respect to the force, the cleaves were also bad, resulting in a large number of steps sloping across the resulting surface (Figure 5.3). Though areas near the center of the substrate might be large atomically flat planes, the edge in this case would be an “infinite” staircase of closely spaced atomic steps.

A rule of thumb is that if steps can be seen optically on the surface after the cleave, it is likely that the edge will not have cleaved well (Figure 5.4). For this reason we used a small high magnification monocular to look at the surface through a viewport of the chamber after each cleave. If a number of steps away from the vicinity of the scribe mark could be seen, then the sample was not inserted into the STM and another sample was prepared instead.

5.3. STM Procedures

The STM microscope head is located at the bottom of a tubeline hanging from to the chamber above. The tubeline sits inside an exchange gas can within the liquid helium cryostat, and the microscope is at temperature of approximately 4.2K. The tubeline swings freely inside the exchange gas can, mounted on bellows below the chamber, and the rest of system is mounted on air legs inside an acoustic room for vibrational and acoustic isolation. A diagram of the system is shown in Figure 5.5. Additionally, the entire room housing the microscope can be floated. Since the tunnel current is exponentially sensitive to noise sources that influence the tip-sample distance, these kinds of measures are necessary to obtain high quality, low noise data. This design does not allow for access to the microscope head, so the STM tip cannot be changed in-situ. Before each sample, therefore, a clean metal sample (either copper or silver (111)) is prepared by cycles of sputtering and annealing and is placed in the microscope to check the quality of the tip. The tip is field emitted to the metal sample to clean the end of the tip, and then the tip’s sharpness and ability to replicate the known density of states spectra of the metal are checked. Once the
spectral properties and the stability of the arrangement of atoms at the end of the tip are confirmed, the metal sample is removed, and the system is ready for the next sample to be studied. Once a new sample is cleaved to expose a clean surface, a long vertical manipulator is used to take the sample down the tubeline and place it in the STM microscope, where it is typically left for 12 to 18 hours to cool. If studies involving adatoms are to be done, the sample holder can be flipped over so that the face of the sample is facing the chamber above, where one of three sources can be used to evaporate sub-monolayers of metal atoms onto the cold surface. The sample is then flipped over to face the STM scan-head again.

To begin STM studies the tip then needs to be brought to within a distance of about 1 nanometer from the sample, without coming into contact with the sample, in order for a tunneling current to be established. Our STM is uses a Besocke type walker that consists of three outer piezoelectric ceramic legs and an inner piezo scanner on which the tip is mounted. High voltage can be applied to electrodes on the four quadrants of each piezo, causing them to bend in the x or y direction or extend or contract on the z direction. The sample plugs into a round piece we call “the ramp” which consists of three sloped surfaces which each rest on top of one of the legs. To bring the tip within range of the sample, the sample is walked by applying series of high voltage pulses to each of the piezo legs in a method known as “pizza mode.” The legs are extended in z while being contorted in x and y to cause the sample to rotate slightly. The legs are then contracted quickly in z while returning to the center position in x and y. This causes the sample to fall and the legs to catch it at a slightly higher or lower position on the ramps. In this way the sample can be moved towards or away from the tip.

The drawback to this scheme is that the sample inevitably shifts in the x and y directions while walking so that it is impossible to align the tip with the center of the sample a priori before beginning the approach. It is fortunate therefore that our machine allows the detection of a significant change in the capacitive response of the tip circuit when the sample is brought into
Figure 5.5: Schematic diagram of the components of the LTSTM. The liquid helium cryostat is shown in blue and the microscope head located near the bottom of the cryostat is labeled.
proximity with the tip. Since the sample sticks out about a half millimeter past the clamp pieces, it is possible to walk the sample in x or y and see the capacitance increase and then decrease as the sample transits over the location of the tip. Using this method the tip can be realigned as the sample is brought closer and a successful approach to the center of the sample can be achieved.

For most samples, acquiring data could now begin, but for the GaMnAs heterostructure samples it was necessary to bring the tip to the very edge of the sample where the heterostructure was located. The capacitance was again a useful measure for doing this. It proved impossible to walk laterally and not crash the tip without walking some distance away in z. Practically this meant approaching, checking the area to see if the heterostructure was there, if not, walking out, walking laterally and reapproaching. It was extremely helpful that we could scan an area up to several microns across with each approach but this process could still require between 50 and 100 approaches before the heterostructure was located. Each approach carried a risk of a tip crash, and walking off the end of the wafer also carried a risk of crashing the side of the tip into the sample. The rate at which the capacitance dropped as the tip was moved towards the edge of the sample provided a useful measure of whether it was close to the edge or not, and allowed for far fewer approaches to have to be made. Once near the edge, it was not uncommon to see many atomic steps for up to tens of microns away from the heterostructure. Staircases of steps were usually a sign that the edge would not be good, but occasionally it would turn out that the steps would stop at the heterostructure and the GaMnAs layer would be perfectly flat.

Acquisition of data and control of the STM was done using an RHK SPM-1000 control electronics box which was managed with the RHK XPMPro program on a Windows PC. The tunnel current was measured with a Femto LCA-1K-5G current preamplifier with $5 \times 10^9$ V/A gain. Tunneling measurements were typically done with a feedback loop tunnel current of 100 to 200 pA and a tunnel junction resistance of several gigaohms. Density of states differential conductance ($dI/dV$) measurements were done using a Perkin-Elmer 7265 digital lockin amplifier.
with an oscillation frequency of 757Hz. Typical oscillation amplitudes were 10 mV or less.

Particular care must be taken to get the phase of the lockin correct so that zero tunneling current results in zero dI/dV signal. In studies of the distributions of values of the local density of states, an incorrect phase adds a constant offset to the distribution that is impossible to correct for later.
CHAPTER 6. SUBSTITUTION OF TRANSITION METALS INTO THE SURFACE OF GaAs

Cleaving GaMnAs heterostructures, as described in Chapter 5, to produce atomically flat surfaces is a difficult task. Cleaving wafers of GaAs, however, is relatively easy, and given the ability of our STM system to deposit sub-monolayers of isolated transition metal adatoms onto the cleaved surface at cryogenic temperatures, experiments of this sort were easy to undertake. A variety of effects have been seen by STM with magnetic adatoms on surfaces ranging from observation of the Kondo effect, to spin-flip excitations to interactions between dimmers or even longer chains of adatoms created by STM manipulation.\(^1\) The original aim of the experiment was to see if interactions between closely spaced Mn adatoms on the surface could be observed, potentially mediated through the states of the substrate, similar to how the carriers mediate ferromagnetism in GaMnAs. The details would be different, but the observation of an interaction would clearly be interesting. It came as a surprise, therefore, when we realized that in the process of trying to manipulate the adatoms to create closely spaced pairs, we were actually causing the adatoms to exchange places with atoms in the lattice. This allowed the creation of custom-made arrangements of pairs, trimers, or even larger numbers of Mn atoms in the GaAs surface along various crystallographic orientations that could be studied for interactions. The observation of interactions between pairs could help explain the coupling between Mn’s in GaMnAs, as long as the effect of the surface is not too great.

In addition to Mn, we tried depositing Fe and Co onto the GaAs surface and found that for all it was possible to induce the transition metal adatom to substitute with a Ga atom in the surface,

ejecting the Ga atom to the surface. We were also able to identify Zn dopant atoms that happened to be in the surface of the Zn-doped p-type GaAs wafers used. This allowed the observation of both magnetic and non-magnetic transition metal acceptor states in the GaAs (110) surface. The resulting in-gap defect states have similar spatial symmetries (shown in Figure 6.5) that are generally reproduced by tight binding calculations of Mn in bulk GaAs(32), pointing to a large influence of the GaAs host on the states. However, from the splitting of the acceptor level of isolated Fe and Co at the surface, it can be shown that the surface does play an important role. This was also recently suggested for Mn near to the surface by tight binding calculations that take the surface explicitly into account.(95, 96) It is shown in section 6.4 that the similarities originate from the antibonding nature of the acceptor states which arises from the hybridization of the impurity d-levels with the host. A second, deeper, in-gap state observed for Fe and Co can be explained by the effect of symmetry breaking of the surface on the degeneracy of the acceptor level. The work presented in this chapter can be found in Ref’s (87, 97).

6.1. Substitution of Adatoms

The experiments were performed near 4 K in ultra-high-vacuum. Wafers of p-type GaAs, doped at 10^{19} Zn atoms cm^{-3}, were cleaved in situ to expose a (110) surface. The Zn dopant concentration was measured in STM topographs of the GaAs (110) surfaces prior to depositing foreign atoms. In addition to Zn below the surface, a defect in the surface matching the expected concentration of Zn dopants was observed, which we identified as Zn in the surface layer. This added an additional effective mass-like, non-magnetic transition metal acceptor in the same configuration to the other magnetic atoms studied. We evaporated small concentrations of Mn, Fe, or Co atoms (~0.5% monolayer) from in situ sources onto the cold GaAs surface. (The process of depositing the adatoms onto the surface is described in section 5.3.) The various species could be distinguished between samples in filled state topography. Previous STM studies of metal atoms on GaAs (110) deposited enough atoms to form clusters(98-101), in contrast,
Figure 6.1: Two substitutions of cobalt adatoms into the lattice showing cobalt can de-substitute and come back to the surface. (a) Several cobalt adatoms that appear as depressions on the lattice. A subsurface Zn dopant can be seen in the upper left. (b) A voltage pulse from the STM tip, as described in the text, causes a cobalt atom in the lower right to substitute into the lattice, kicking a gallium atom to the surface. (c) The gallium from the first site is moved out of view, and a second cobalt is substituted at the center. (d,e) The gallium adatom from the second site is manipulated to the location of the first site. (f) A voltage pulse causes the gallium to go in and the cobalt comes back to the surface. (All images are 120Å x 90Å at -1.5V)
adsorbate atoms were clearly identifiable as isolated atoms on the surface, and no Fermi level pinning was observed.

The STM atomic manipulation technique involved placing the STM tip above the adsorbate atom and sweeping the tunneling bias from negative to positive voltage across the GaAs energy gap with the feedback loop off, tunneling energetic electrons into the adatom. In doing so, we found that Mn, Fe and Ga adatoms could be made to move randomly on the surface within a half circle centered about the [001] direction, while Co atoms would move oppositely for an unknown reason. Using higher current and voltages, the transition metal atoms could be made to substitute into Ga sites in the first layer of the (110) surface, ejecting the Ga atom to the surface. The energetics of the substitution process varied between species but were generally similar. Using a setup condition of 200pA at -1.5V, the tunnel current would drop to zero as the bias voltage was swept through the gap, begin to rise sharply near +1V and then drop suddenly a few hundred mV later when the substitution occurred, providing a rough qualitative measure of the activation barrier required to be overcome to occupy the cation site. In the case of Co we were even able to get a Ga atom ejected from one substitution to resubstitute with another Co previously placed in the lattice at another site. The sequence of this is shown in Figure 6.1. We found that only Co could be induced to swap with a Ga adatom and return to the Co adatom configuration again, suggesting that it was less energetically favorable for it to stay in the surface. After substitution, the displaced Ga atom often remained loosely bound to the newly created defect site, proving more difficult to remove for Co than for Fe or Mn. It has been shown that the same STM manipulation is possible for Mn adatoms on InAs.(102) The substitution process is likely similar to the ejection of cation atoms to the surface that can occur during the initial growth of transition metal overlayers on III-V surfaces at room temperature.(103, 104) Our observation that the adatoms at low temperature sit on the surface and do not substitute with atoms of the substrate until we manipulate them is potentially of interest for the deposition of magnetic films on GaAs
for spin injection in spintronic devices. Recent work on the deposition of Fe films on GaAs found that the intermixing of species and formation of compounds at the interface can be significantly reduced by low temperature deposition, which is potentially consistent with our observations. (105, 106)

6.2. Isolated and Interacting Mn’s in the GaAs Surface

Initial studies focused on Mn adatoms because of their obvious connection to GaMnAs. Once a single Mn adatom was substituted into the lattice, careful voltage pulses with the tip could be used to try to move another Mn close to it and substitute it into the lattice to create a pair. This process was difficult to control, and trying to create a pair of a specific orientation and distance required many attempts. In addition to the controlled substitution mechanism, it was found that taking topographies at positive bias would incorporate most of the Mn within the scan area into the lattice. In doing so, large numbers of Mn could be substituted into an area, as shown in Figure 6.2(a, b), creating numerous pairs, triplets or larger numbers of closely spaced impurities. Depending on the density of Mn adatoms, occasionally this method was useful for creating a large number of features and hunting for interesting configurations of isolated pairs. Under the influence of the STM tip, Mn in the surface sometimes switched to a neighboring Ga lattice site. In the process of creating or studying the pairs, it was found that they seemed to have a tendency to want to form or reconfigure to be along the <001> direction. Once a pair was in this orientation it tended to be more stable, with one Mn very occasionally moving over one lattice site to create a <211> pair but never returning to an in-row <110> configuration.

The defect state created by the Mn in the surface was studied with dI/dV spectroscopy. It consists of a single in-gap state that was wide in energy centered at 0.85V (see Figure 6.2(c)). The width of the state is not uncommon for in-gap states measured by STM, but it does mean that any fine multiplet splittings of the acceptor state would not be able to be observed. The acceptor state of Mn in bulk is triply degenerate, and it was assumed that this was the reason they could not be
Figure 6.2: (a) Impurity states of many Mn’s substituted into the surface of GaAs. Typically, lower densities were used to avoid complications from too many nearby defects, but very high densities like this could be created with many pairs and clusters. A subsurface Zn can be seen near the center of the image. (200Å x 200Å, 1.5V). (b) Valence band filled states topography. Each Mn causes an enhancement on the two As atoms in the surface to which it is bound. (-1.5V). (c) dI/dV density of states of an isolated Mn impurity showing it causes a single state in the gap with an energy of about 0.85V.
Figure 6.3: (a) In-gap topography of the states of a Mn-Mn pair 8Å apart along the <110> direction (1.1V, 40Å x 40Å). (b) Valence band topography of the pair (-1.5V). (c) dI/dV density of states showing there are two in-gap states associated with the pair. (d) Spatially resolved energy map of the lower state at 0.91V showing an antibonding-like arrangement. (e) Higher energy bonding-like state at 1.35V.
observed. This will be discussed more in the next section, where in the context of studying other species of transition metals, it became clear that the surface was responsible for strongly breaking this degeneracy. When a pair of Mn’s are created, the single in-gap state either splits or does not change, depending on the orientation of the atoms in the pair. The splitting is most strong for pairs along the in row <110> direction. Such a pair 8 Å apart is shown in Figure 6.3 with the splitting in energy and the spatial dependence of the two states. Surprisingly, there is no observable splitting of the states for pairs across the rows in the <001> direction. The splitting observed for pairs in various orientations is displayed in Figure 6.4.

For states of two interacting magnetic atoms in a gap, one way this can be explained is by their relative spin alignment. If the states are ferromagnetically aligned they will split into bonding and antibonding states, while if they are antiferromagnetic then the states can stay degenerate and no splitting will be observed. (107) This implies that Mn-Mn pairs along the <110> direction are ferromagnetically coupled and those along <001> are antiferromagnetic. Similar strong directional anisotropies in the coupling constant J(r) had been predicted by ab initio density functional calculations which questioned whether such large anisotropies could be compatible with an RKKY-like interaction. (14) Tight-binding calculations of Mn pairs in bulk also produced directional dependences similar to those observed. (87) Such a large anisotropy would suggest that if the material could be grown to maximize the number of Mn’s along the <110> directions while minimizing those closely spaced along <001>, then the ferromagnetic properties of the material could be optimized. It would also suggest that frustration is an important component of the physics. Subsequent STM studies of Mn doped GaAs (x << 1%) and tight binding calculations that take the surface explicitly into account have shown that the symmetry breaking of the surface and the strain in the crystal from the GaAs surface reconstruction do have a large effect on the Mn states near the surface. (95, 96, 108-110) This may mean that the large splittings are more a property of the surface than representative of the bulk
Figure 6.4: (a-d) Images of the acceptor levels of closely spaced pairs of Mn impurities in the surface. Crystallographic orientation between the Mn’s is labeled in the upper right. Images are 40 Å x 40 Å acquired at 1.5V. (e) Comparison of the observed splitting in energy of acceptor level to those from tight binding calculations of Mn pairs in bulk. The splitting is the delta in energy between the upper and lower states such as shown in Figure 6.3. If the splitting is due to hybridization of the wave functions, pairs along the <100> direction are likely to be antiferromagnetically aligned due to the lack of splitting. In general, it would imply that the magnetic coupling is highly anisotropic.
properties. The effects of the surface on the individual acceptor state are considered in more detail in the next section.

6.3. Comparison of Transition Metal Species in the Surface

In GaAs, Zn forms a shallow acceptor whose valence band-like hole state is well described by effective mass theory. Fe and Co are deep acceptors that are expected to have a localized state that is largely determined by their atomic central cell potential. Mn sits between these cases as a fairly deep acceptor that binds a hole in an Mn$^{2+}$3$d^5$-hole configuration. In this sense it is like Zn, but its binding energy is larger because there is hybridization between the d levels and the host which, importantly, provides a means of coupling to the spin of the core state. The strength and influence of this hybridization on the nature of the hole state has been a source of debate. The effective mass limit is clearly compatible with the RKKY model. Others found the hybridization to be large enough to give the hole significant impurity-like d character causing highly anisotropic exchange interactions between localized Mn impurity states as mentioned. The two limits will in general predict different anisotropies in both the defect state wave functions and the magnetic interactions. Insight into the nature and character of the Mn acceptor state can be gained by comparing it to the acceptor states of both shallow and deep transition metal impurities. To do this we used the STM manipulation technique to substitute Fe and Co in addition to Mn into a GaAs host. While we expect the surface to have significant effects, it is not clear a priori which would be the most important. As it turns out, it appears that the surface significantly changes the picture from in bulk.

Topographies of Mn, Zn, Co, and Fe in the surface layer are displayed in Figure 6.5. The topography maps the acceptor states by applying the bias voltage near the conduction band edge so that the image is the sum of all states in the gap above the Fermi energy. An anisotropic star-shaped structure is apparent for each state with $C_s$ symmetry similar to that observed for subsurface acceptors in GaAs. The wave function of Mn in bulk GaAs that is predicted
Figure 6.5: Topography of in-gap states (40Å x 40Å) and insets of valence band states (30Å by 30Å) of the four transition metal species studied in the (110) surface of GaAs. Topography in insets has been multiplied by two to enhance the contrast. (a) Zn (1.6V, inset: -1.4V), (b) Mn (1.6V, inset: -1.3V), (c) Fe (1.5V, inset: -1V) (d) Co (1.5V, inset: -1.3V)
Figure 6.6: dI/dV measurements of an Mn and Zn acceptor; each showing a broad in-gap resonance. The Zn resonance peaks near 0.65±0.05 V and the Mn resonance peaks near 0.85±0.03 V. The minor peak near 1.3V changes with different tips and is believed to be a tip induced shadow of the main peak. (b) dI/dV measurements near an Fe acceptor, with the tip between head and arm features (see inset), bring out two states near 0.87±0.05 V and 1.52±0.05 V. (c) Two dI/dV measurements taken near a Co acceptor also show two distinct states. With the tip centered over the Co/Ga lower energy resonance is seen near 0.92±0.05 V, and with the tip over the arm feature a higher energy peak near 1.15±0.05 V.
by tight-binding calculations viewed in the plane of the Mn site resembles our data well.\(^{(32, 87)}\)

It is known that the depth of the dopant beneath the surface affects the symmetry of the acceptor wave function observed by STM, with Mn going from \(C_2\) to \(C_s\) symmetry the nearer to the surface it is, possibly as a result of surface-related strain.\(^{(108)}\) The wave functions in Figure 6.5 can be viewed as the extreme limit of that trend. Recent tight-binding work that attempted to take the surface into account is consistent with this trend as well.\(^{(95, 96)}\) The wave functions can be described as a central lobe with arms along the [1\(\bar{T}\)0] directions and legs along [1\(\bar{T}\)1]. The dominant half of the subsurface Mn bowtie wave function is associated with the central lobe, which expands in the [00\(\bar{T}\)] direction when located in layers beneath the surface, while the weaker half arises from the leg features. The arms are not observed in subsurface Mn in GaAs, but a similar feature has been observed for Cd in GaP.\(^{(116)}\) The arm and leg features of the Zn impurity are weaker than Mn, suggestive of the triangular wave function observed for subsurface Zn dopants.\(^{(114)}\) Whereas the triangular shape of subsurface Zn dopants observed by STM has been associated with multiple resonances that are observed throughout the band gap,\(^{(117-119)}\) the spatial structure of a Zn in the surface layer is clearly identified with a distinct single state in the dI/dV density of states spectra. The observed wave function is measured as a wide in-gap state in spectroscopy peaked near 0.65 eV for the surface Zn acceptor and at 0.85 eV for Mn, shown in Figure 6.6(a).

While the shape of the Mn and Zn acceptors in the topographs directly correlate with single, strong resonant levels, Fe and Co acceptors have more complicated spectral signatures. The Fe and Co states show increased spectral weight on the arm and leg features. Spectroscopy shows that both Fe and Co have a low energy state close to the Mn level at 0.87 and 0.92 eV, and a higher energy state at 1.5 and 1.15 eV respectively, shown in Figure 6.6(b) and (c). Spatial maps of the energy-resolved density of states in Figure 6.7 show that the shape of the lower energy state is similar to that of Mn. The higher energy state is responsible for the greatly enhanced arm-
Figure 6.7: Differential conductance maps of Fe and Co acceptors (50 Å x 50 Å). Energy maps near the Fe$_{Ga}$ show (a) the lower resonance state mapped at 0.88 V (inset: 30 Å by 30 Å topography taken at 1V better shows spatial extent of the state), and (b) the higher energy resonance mapped at 1.50 V. Energy maps of an isolated Co$_{Ga}$ (c) near the lower resonance peak energy of 0.85 V and (d) near the higher energy resonance of 1.15 V with a setup voltage of 1.05 V. The Fe$_{Ga}$ setup voltage was 1.45 V which created a larger shadow effect in (a) than in (c) since the setup voltage for the Fe$_{Ga}$ map was nearer its high energy resonance. Both (b) and (d) exaggerate the depression in the center because the tip is further from the surface due to the tunneling contribution from the lower energy resonance.
like features. The low contrast at the location of the dopant in the maps of the higher energy states in Figure 6.7 is somewhat exaggerated due to the lower energy state causing the STM tip to move out, decreasing its sensitivity. Nevertheless, the maps clearly show that spatial dependence is quite different for the two states.

The filled state topography of the defects (insets Figure 6.5) shows an enhancement on the two in-row [1\bar{1}1] nearest-neighbor As sites, with some minor variation between the species. The lack of any depression in the valence band topography of the p-type GaAs around the impurities shows they are not positively charged, and thus the states in the gap are acceptor-like. At biases close the Fermi energy we observe an apparent lattice distortion involving the nearest neighbor As atoms for all species except Co. This is most clearly seen with Mn, where at low bias one As appears to buckle upward and another downward. Electrons tunneling inelastically with sufficient energy can cause the distortion to switch to the other equivalent configuration, with the down As now buckling upward and vice versa as seen in (b) and (c) of Figure 6.8. Larger biases cause the switching to occur more rapidly, resulting in noise in the topography around the feature, similar to that observed due to switching behavior in a variety of other systems with STM.(120, 121) At large biases the switching occurs quickly enough that only the average is observed(122), as in the insets of Figure 6.5. Coupling to phonons can provide an efficient means of dissipating energy during the capture and recombination process of carriers trapped by deep levels.(123) A similar multiphonon emission process could explain the large magnitude of the dI/dV signal when tunneling into these impurities deep in the gap as well as explain their width in energy.

6.4. Discussion and Model of the Effect of the Surface

While the similarity of the ground state of all four species of acceptor may at first be surprising, it follows from the well-known effect of symmetry on the hybridization of a substitutional impurity with the host. The defect states can be seen as arising from the interaction of the atomic states of the isolated impurity with the dangling bond states of an ideal vacancy in
Figure 6.8: (a) Valence band topography of two Mn$_{Ga}$ sites showing an enhancement of the adjacent arsenic atoms. (50Å x 50Å, -1.5V) (b) Low bias topography of the same area shows an enhancement on only one arsenic and a depression at the site of the other. Scan proceeds from right to left. A glitch occurs ¾ of the way through the measurement (-0.25V). (c) Result of glitch in (b), feature has switched to the mirror configuration. (d) Composite image made by combining (c) and (d) showing how the high bias image is a mixture of the two configurations. (e) As the bias is increased the switching rate increases, appearing as noise in the topography. (-0.5V) (f-h) Noise patterns were observed around (f) Fe$_{Ga}$ (45Å by 45Å, -0.3V) and (h) Zn$_{Ga}$ (32Å by 32Å, -0.5V) but not (g) Co$_{Ga}$ (32Å by 32Å, -0.3V).
the lattice (Figure 6.9). Clearly, if the vacancy is filled with the host atom, the hybridization forms a bonding state in the valence band and an antibonding one in the conduction band. It is the difference in the hybridization that occurs with an impurity that gives rise to the defect states.\(^{124}\) The case of Zn is the easiest to understand: the Zn d levels are deeply bound below the valence band edge and are non-interacting, and the Zn valence orbitals are comparable to but higher than those of Ga. This gives rise to a state with primarily host-like bonding character that isn’t pushed all the way down into the valence band and an antibonding state with impurity character above the conduction band edge.

As shown in Figure 6.9, the case when the d levels are interacting is more involved and is well summarized by Mahadevan and Zunger.\(^{112}\) The s and p levels that make up the sp\(^3\) hybrids of the vacancy transform under zinc-blende T\(_d\) symmetry according to A\(_1\) and T\(_2\) irreducible representations. The a\(_1\) state is a singly degenerate s-like state, while the t\(_2\) is triply degenerate p-like. Before hybridization, the t\(_2\) levels of the Ga vacancy are near the valence band edge. Under the same symmetry, the d-levels of the transition metal are crystal field split into an e symmetric doublet and t\(_2\) symmetric triplet. The t\(_2\) states of the impurity and vacancy will strongly hybridize into bonding and antibonding states while the e states do not. Mahadevan and Zunger calculate that the impurity d levels of Mn, Fe and Co lie energetically beneath the vacancy level. Therefore, the t\(_2\) derived in-gap states are antibonding and predominantly host-like, though the stronger the hybridization the more mixed character they will have. This is the origin of the Mn\(^{2+}\)3d\(^5\) + hole configuration, the hole is host-like and, though antibonding instead of bonding, it shares the same t\(_2\) (p-like) symmetry as the valence band. The core spin is preserved in the deep impurity-like e and bonding t\(_2\) states that fill in a high spin configuration according to Hund’s rules.

With this in mind, a few conclusions can be drawn. First, the anisotropy of the ground state for all four acceptors is similar, implying they are all states of the same (t\(_2\)) symmetry. It should
be noted that the character of acceptor state will depend sensitively on the order of levels in energy and be determined by the lowest unoccupied state. The similarity between them is most simply explained if the bonding d levels fill according to their atomic configuration. Second, the spatial extent of the Mn, Fe and Co ground state is the same, and that of Zn is smaller. This is a result of the antibonding nature of the Mn, Fe and Co states which causes the wave function to be more delocalized onto the As dangling bonds than the bonding nature of the Zn state. This result is somewhat counterintuitive, because as the shallowest acceptor the Zn wave function is expected to be the most delocalized. It arises from the fact that STM doesn’t sample the wave function within the material but rather the part that is exponentially decaying into the vacuum at the location of the tip. Since the acceptors at the surface have no layers above them for the wave function to spread out in, instead the STM only images the core of the wave function at the impurity site. Third, a partially filled triply degenerate t\(_2\) level could be expected to be susceptible to Jahn-Teller distortions as are observed for some configuration terms of Fe and Co in III-V materials, potentially explaining Figure 6.8.\(^{(125, 126)}\)

In our analysis thus far, however, we have ignored the effect of the surface and surface reconstruction which lowers the symmetry from T\(_d\) all the way to C\(_s\) (\((1\bar{1}0)\) mirror plane). This will split the t\(_2\) states into functions of parity, creating two a’ (even) levels and one a’’ (odd) level (Figure 6.9). While it is hard to estimate the resulting energies of these levels, \textit{ab initio} calculations of the unrelaxed surface Ga vacancy in GaP found the t\(_2\) level to split into one a’ level in the valence band and the remaining a’ and a’’ levels to be nearly degenerate above the valence band maximum.\(^{(127)}\) This implies the possibility of two states in the gap, one having even symmetry and other odd. The comparison to the spatial structure of the two states of Fe and Co is immediately obvious, the lower energy state being symmetric with respect to the (1\bar{1}0) plane, and the higher energy state being consistent with the wave function squared of an odd state with respect to the mirror plane. For the higher state, it is unlikely that the doubly ionized state is
Figure 6.9: (a) Schematic of hybridization of an atom with an ideal gallium vacancy (V_{Ga}) leading to a shallow acceptor level such as for Zn. The valence levels of the impurity are higher in energy than the host atom, resulting in a state just above the band edge that shares the valence band’s bonding character. (b) Hybridization of levels giving rise to the Mn acceptor following the model of Ref 5. In contrast to (a), the acceptor is antibonding in character. States are labeled +(-) for majority(minority) spin. Colors represent the initial state the hybridized Mn_{Ga} level most resembles in character. (c) Possible level ordering that would give rise to a symmetric acceptor state for a Mn under the influence of the symmetry breaking of the surface, as described below. (d) Splitting of a t_{2} state by the surface. Removal of one neighboring As atom lowers the T_{d} symmetry to C_{3v}; the rest of the missing atoms at the surface and the GaAs surface reconstruction lower the symmetry further to C_{s} ((110) mirror plane). For there to be a symmetric (a’) acceptor level at the surface there can be at most three electrons in the vacancy-like t_{2} level. For Fe and Co, this would require that their additional electrons fill the impurity-like e and t_{2} levels according to their atomic configuration.
being observed, but rather that the electron is tunneling into an unoccupied excited state. In Figure 6.9(c) one possible level order that allows for two unoccupied states of different symmetry is shown. This implies that the d-level configuration of Fe and Co is $3d^6$ and $3d^7$ respectively, as they would have as isolated atoms.

Finally, we attempt to address the energies of the states. The order of the acceptors in bulk is Zn (31 meV), Mn (113 meV), Co (160 meV) and Fe ($\sim$0.5 eV). At the surface we observe the same trend with the exception that the ground state of Fe is below Co. The shift to higher energy is likely due both to tip induced band-bending and the energy of the acceptor states being affected by being at the surface. If the Ga vacancy dangling bond levels are higher at the surface than in bulk, this would lead directly to a deeper acceptor state. It is worth noting also that charge transfer from cation to anion sites due to the surface reconstruction is known to increase the binding energy of the surface Ga 3d levels by $\sim$0.3 eV. If this applies to the transition metal impurities, it would be expected to push the d-states deeper into the valence band lessening the p-d hybridization, and pinning the acceptor energies closer to the Ga vacancy dangling bond energy, resulting in the confluence of similar energies for the first states of Mn, Fe and Co. The energy splitting of the states may also be correlated with the amount of lattice distortion observed for each impurity, with Co having no observable distortion and the smallest energy splitting and Fe having more distortion and splitting. As Mn had the most distortion, the upper energy level would be pushed into the valence band. This would explain the lack of an observation of a second level for Mn which would be expected from the preceding analysis of the origin of the acceptor states.

We have demonstrated an atomic scale technique to study single transition metal impurities in the surface of a III-V semiconductor, successfully substituting single Mn, Fe, and Co atoms into the GaAs (110) surface, as well as identifying native Zn dopants in the same layer. The similarity of the anisotropic wavefunctions measured for all these acceptors shows the importance of the
host in determining the shape of these states, even for deep non-effective mass acceptors. The character of these states can be seen to follow directly from simple symmetry arguments and the hybridization of the impurity and host states. The interactions of pairs of closely spaced Mn’s in the surface must be affected by these changes from bulk. Nevertheless, the observations described here are a unique example of the interaction of magnetic impurities with a semiconductor surface, and the demonstration of the ability to dope a material one atom at a time could be an extremely useful capability for the custom creation of nanoscopically engineered materials in the future.
CHAPTER 7. IDENTIFICATION OF DEFECTS IN GaMnAs

Defects are what make semiconductors useful. By controlled addition of specific defects to pure semiconductors, they can be made to display a huge array of useful properties and novel physics. One specific example of this is the substitution of the transition metal Mn into GaAs, which can result in the creation of a ferromagnetic semiconductor. The emergence of ferromagnetism occurs when Mn is substituted for more than about 1% of the Ga in the non-magnetic GaAs lattice. This unique property, arising at such dilute concentrations, is therefore strongly dependent on the interaction of the Mn defect states and the properties of the bulk crystal, which makes the characterization of defect state of critical importance to the understanding of the underlying physics. One tool for doing this is scanning tunneling microscopy (STM), which can provide an atomic-scale view of the electronic properties of individual defects. While there have been many STM studies of Mn defects in lightly doped paramagnetic samples (< 1%), there have not been any that focused on individual defects in ferromagnetic samples.

Studies in the paramagnetic regime have focused on the anisotropy of the acceptor wave function and, in particular, whether it is determined primarily by the band structure or by the local potentials of the Mn itself.\textsuperscript{(115)} This has provided an important confirmation of our understanding of the nature of the acceptor state and insight into the types of interactions that are possible that give rise to ferromagnetism. Further studies looking at whether the wave function is modified in the presence of other nearby Mn’s, trying to anticipate the effects possible at much higher dopings, found no discernable changes.\textsuperscript{(131)} In the process of cataloging these properties, it has been shown that the shape of the impurity wave function is strongly dependent on the proximity of the Mn to the (110) surface in both GaAs and the closely related compound.
InAs.\cite{109,110} This effect was first ascribed to strain in the crystal from the relaxation of the atoms at the surface by Loth and was later studied and confirmed by tight-binding calculations.\cite{95,108,132} The warping of the Mn wave function in the presence of strong strain fields had already been demonstrated.\cite{133} Near the surface, it is seen as an increasing asymmetry in the weight of the wave function across the $\{110\}$ mirror plane. At the surface the effect is most pronounced, where the reduced symmetry can result in splittings of the acceptor level that are quite large.\cite{134} Such depth dependence has allowed a cataloging of the wave function of isolated Mn impurities in each layer for up to 10 or more layers beneath the surface, which can be used for comparison with Mn in other environments.

A fundamental question for the properties of GaMnAs is whether the Mn wave functions are the same in the ferromagnetic state as they are in the paramagnetic regime. Several factors might be expected to cause possible changes. Interactions between impurities should be more pronounced at concentrations beyond the metal-insulator transition where $N_{\text{Mn}}a_b^3 \approx 1$. Screening of the Coulomb potentials, due to the high hole concentration, should be larger. Also tight-binding calculations of Mn in bulk GaAs predicted that the strong spin-orbit effect couples the shape of the wave function to the direction of the Mn core spin.\cite{135} It was recently predicted that this effect should be present for Mn near the surface and be observable by STM.\cite{132} Since in the ferromagnetic state the spins should be preferentially aligned instead of thermally randomized, this is another possible mechanism by which changes might be seen in the ferromagnetic state, compared with in the very dilute non-magnetic environment. Instead we observe the wave function of Mn viewed from either the $(110)$ or $(\bar{1}0\bar{1})$ planes to be very similar to the isolated limit. We do however find a discrepancy in the identification of Mn one layer beneath the surface from that given in Ref \cite{110}.

A sample of $\text{Ga}_{0.985}\text{Mn}_{0.015}\text{As}$ with a $T_C$ of 30K was grown by MBE. A thick 200nm layer of the GaMnAs layer was grown to aid finding it with the STM, as well as to provide large step-free
Figure 7.1: (a) Topography of the in-gap states at the (110) surface acquired at 1.52V. Mn impurities in various layers labeled by layer, where 1 means the surface layer. Distinguishing individual defect wave functions can be difficult because of the large number of overlapping states. As antisites are labeled “A”. (b) Determination of the depth beneath the surface is aided by the valence band states where Mn in even layers cause an enhancement centered on a single As in the surface, while those in odd subsurface layers cause a more diffuse enhancement. The distinctive enhancement of two As atoms in a row by Mn in the surface is also clearly visible. For the antisites, the positively charge of the donor-type defect bends the bands locally, causing depressions in the topography. Topography was acquired at -1V.
areas, as the cleaving process often left stepped regions near the front edge of the wafer.

Degenerately doped n-type substrates were used to allow the STM to tunnel into the sample while attempting to find the MBE grown layers. The STM measurements where performed using a home built ultra-high vacuum (UHV) STM at a temperature of 4.2K. A mechanically sharpened iridium tip that was cleaned by field emission and checked to give reproducible spectral measurements on a clean Au(111) surface before each measurement was used. The samples were cleaved in situ at room temperature to expose either the (110) or (11̄0) face of the material before being transferred within a few minutes into the cold microscope where they were zero-field cooled.

STM topography acquired at a voltage just below the conduction band edge maps out the unoccupied states that exist in the band gap of the material above the Fermi energy. Such topography of the states in GaMnAs viewed at the (110) surface is shown in Figure 7.1(a). A number of Mn related states and a few As antisites can be observed. Defect states of several Mn impurities in the surface layer are visible which are consistent with those observed in previous work where Mn adatoms were substituted into the surface of GaAs using STM manipulation.(87) As can be seen from the two Mn states labeled as being in layer 4 near the center of the image, the wave functions observed vary somewhat due to the local environment around them. Given the high density of Mn impurities, the states often overlap spatially, making the identification of individual wave functions difficult. In particular, Mn in layers closer to the surface often obscure those further down such that it is hard to get a clear view of them further down than the fourth or fifth layer. The distribution of Mn locations appears to be consistent with random doping and no tendency for the Mn to cluster near each other could be discerned. For those Mn that happen to be in locations where they are relatively well isolated, the wave functions are very similar those observed in for isolated Mn impurities.(109, 110) The only discrepancy we find is that what we identify as Mn in the second layer (first layer beneath the surface) was identified as a pair of Mn
in the surface layer in Ref (110). In fact, such a pair of Mn in the surface can be seen in the lower left hand corner of Figure 7.1(a) which is clearly distinct from the features marked in layer two near the opposite corner. This also means that the features Ref (110) labeled as being in layer 2 are actually in layer 3 and so on. Further support for this can be found in topography taken of the valence band states. Because of the directions of the bonds, a dopant in an even layer causes an enhancement in the valence band centered on a single As atom at the surface while those in odd layers do not, as can be seen in Figure 7.1(b). (114) We believe this classification is also more consistent with the identifications made for Mn in InAs. (109)

The wave functions were also observed when the sample was cleaved on the perpendicular (11̄0) plane as shown in Figure 7.2. The shape of the wave function is inverted along the [001] direction compared to the (110) surface, reflecting the change in the bond directions at this surface, but is otherwise identical, similar to observations of the deep acceptor Cd in GaP. (116) At the low temperatures GaMnAs films display biaxial easy axes along the [100] and [010] directions which undergo a second order transition to uniaxial at higher temperature along either the [110] or [11̄0] direction. (35) At higher hole concentrations [110] is the easy axis and [11̄0] is hard. (36) In looking for whether the Mn wave function is dependent on the direction of the core spin, comparing STM data of the ferromagnetic state, where the spin should have a definite preferential alignment, with previous non-magnetic studies or observing the wave function at different surfaces might be expected to yield some effect. However, a comparison with the very-dilute non-magnetic regime does not reveal discernable differences. It should be noted that, because of the influence of the local environment on the Mn acceptors we observe, a quantitative comparison is difficult, and if the changes are small we may not be able to observe them.

The energy dependence of the states as a function of the depth of the Mn beneath the surface is plotted in Figure 7.3. As shown for donor impurities near the GaAs surface, we see the binding energy of the defect state increases the closer the impurity is located to the surface. (136) The
Figure 7.2: Acceptor state wave functions of Mn measured at the (110) surface in comparison to the (110) surface of Figure 7.1. (a) Mn in the surface layer, (b) second layer (first subsurface), (c) third layer and (d) fourth layer. The shapes of the wave functions observed at both surfaces are very similar, not showing any discernable dependence on the surface studied. The [001] direction is rotated in the opposite direction from in Figure 7.1 for easier comparison of the wave functions.
Figure 7.3: STM dI/dV measurements of the density of states of Mn in various layers showing that the acceptor level energy decreases as the Mn is located further beneath the surface. Mn in the surface layer causes a very deep state near 0.85V. Mn in the second layer still has a very deep acceptor state close to 0.4V and possibly a second state at higher energy. In the third layer, the acceptor state is observed much closer in energy to that of an isolated Mn acceptor in bulk GaAs. The acceptor energies observed for a Mn in any of these layers can vary by approximately ±50meV depending on the local environment around it.
energies observed are qualitatively in agreement with a recent tight-binding calculation of Mn near the surface which found that the binding energies are large for Mn in the surface and second layer and drop in energy to slightly above the isolated acceptor energy of 113meV by the third layer (second subsurface).\textsuperscript{(132)} The energy for Mn in the surface is close to the previously reported value of 0.85V for Mn substituted into the surface of GaAs.\textsuperscript{(87)} The value is not appreciably different for samples with Mn dopings as high as 5%, where the higher hole concentration might be expected to better screen the band bending caused by the electric field of the tip. Additionally, ionization rings can often be seen at large negative biases typically < -0.3V for Mn in the second and third layers, which is consistent with states in the gap becoming occupied when they are pulled below E\textsubscript{F} by the tip bending the bands downward into depletion at the surface. For these reasons and the observed energy of the antisite defect, discussed later, we believe we are close to the flat band condition near zero bias.

We note that the energies observed for Mn in the same layer typically vary by about ±50meV due to differences in the local environment, such as proximity to other Mn’s or to compensating defects. Such effects are expected due to the disorder in the system. The distribution of energies for Mn in the third and deeper layers might be expected to reflect the “band” of impurity energies. However, the dip in the density of states at E\textsubscript{F} that can be seen in the spectra of the third layer Mn in Figure 7.2 can be observed with varying strength at all locations in the sample. Given that this gap always occurs around E\textsubscript{F}, we attribute it to a correlation gap caused by electron-electron interactions, as is commonly observed in GaMnAs.\textsuperscript{(41, 42)} For this reason the peak in the density of states may not accurately reflect acceptor energy, and the width of the impurity band is likely to extend down to the valence band.

Finally, we identify compensating donor defects that degrade the ferromagnetic properties of the material by reducing the hole concentration. Figure 7.4 shows the wave functions and density of states of arsenic antisite defects in several layers near the surface that are created due to the
Figure 7.4: (a-c) Topography of the wave function of As antisites in the 2\textsuperscript{nd} to 4\textsuperscript{th} layers of the (110) surface showing the satellite peaks of the wave function as previously observed. (d) dI/dV spectra showing a deep state in the gap near 0.85V and other states near the conduction band edge. The spectra is suppressed for the third layer antisite because the STM tip was not exactly centered on the maximum of the wave function intensity, but the same peaks can be discerned.
low temperature MBE growth. Mn interstitials are another important defect but, even in more highly doped samples, we were not able to identify them. The As$_{Ga}$ wave functions were first observed by Feenstra et al. in low temperature grown GaAs.\(^{137}\) His observations were made in samples where the Fermi energy lay above the midgap defect state. It is nice to see that the same result is obtained when tunneling into the unoccupied state as when tunneling out of the occupied state. The antisite in the second layer shown in Figure 7.4(a) is surprisingly well reproduced by the \textit{ab initio} calculation of Capaz et al.\(^{138}\) As they predicted, the antisite in the surface has a significantly different character, much like a single dangling bond, and we do not observe it. In the spectra the mid gap level is observed \~100mV higher in energy than the accepted value of 0.75eV for the EL2 defect level. Capaz found the donor level is not strongly changed from the bulk value in layers below the surface. This could imply that there is only an upward tip-induced band bending of about this amount for energies near mid-gap, consistent with our previous reasoning. It is interesting that we don’t see the double donor level in the spectra, which Feenstra explained by suggesting that the two lie close enough in energy that they cannot be individually resolved.\(^{137}\) From topography acquired in the gap at 1V, we can tell that the wave functions observed are due to the mid-gap state. The peaks in the density of states near the conduction band are thus likely to be due to the positive charge of the antisite bending the bands downward slightly to create localized resonances in an accumulation region at the conduction band edge.

We have observed the wave functions of Mn acceptors at low temperature in the ferromagnetic state of GaMnAs. Their shape and energy as a function of depth beneath the GaAs (110) and (110) surfaces match previously reported results at low concentrations in non-magnetic samples and are consistent with recent tight binding calculations of Mn states near the GaAs surface. We show that the defect state of Mn in the second layer has been previously misidentified, however. Though the Mn core spins should have a preferential alignment in the ferromagnetic state, we are unable to confirm predictions that the anisotropy of the wavefunction
depends on its orientation. Antisite defects arising from the low temperature growth of GaMnAs were also observed. The energy that the mid-gap antisite donor state is observed at is only moderately larger than its expected value, providing a measure of the STM tip induced band bending at energies deep inside the gap.
CHAPTER 8. STATES NEAR $E_F$ IN GaMnAs

Understanding of ferromagnetism in GaMnAs would benefit from a microscopic picture of the electronic states near the Fermi level that give rise to magnetic order. Additionally, the observation of electronic states near the metal insulator transition, subject to strong disorder and carrier interactions, is of general theoretical interest. To this end, cross sectional scanning tunneling spectroscopy (STM) studies were carried out on samples with dopings across metal insulator transition. A disordered material with a broad distribution of Mn states in the gap that overlap with the valence band is observed. Signatures of localization and electron-electron correlations are observed at all dopings, suggesting that the effects of disorder persist well into the metallic regime. The local density of states fluctuates strongly in space with a diverging correlation length near the Fermi energy. An analysis of these fluctuations is consistent with the multifractal character of wave functions that is known to occur near the non-interacting Anderson metal-insulator transition. The observation of these effects in an interacting system provides a new means of studying the Mott-Anderson transition.

8.1. Experiment

MBE was used at UCSB to grow 200 nm thick samples of Ga$_{1-x}$Mn$_x$As with x=1.5, 3 and 5% Mn concentrations in order to span the range of the metal-insulator transition. The thickness of the GaMnAs layer was chosen to aid finding it with the STM, as well as to provide large step free areas, as the cleaving process often left stepped regions near the front edge. Degenerately doped n-type substrates were used to allow the STM to tunnel into the sample while attempting to find the MBE grown layers. The conducting nature of the substrate unfortunately complicated any attempt to measure transport properties of these samples. The Ga$_{1-x}$Mn$_x$As was grown on p and n-type buffer layers on top of the substrate in a non-rotated geometry, creating a gradient of Ga to
As composition across the wafer. The samples where diced into pieces along the gradient every 2mm and the \( T_C \) of each piece was measured using SQUID magnetometry to find the stoichiometric region with the highest \( T_C \). The \( T_C \)’s of the as-grown pieces studied were for the 1.5% 30K, 3% 43K and for the 5% 70K. In addition, a 5% piece annealed for 5 hours at 200C with a \( T_C \) of 86K was also studied. The STM measurements were performed using a home built ultra-high vacuum (UHV) STM at a temperature of 4.2K. A mechanically sharpened iridium tip that was cleaned by field emission and checked on a clean Au(111) surface before each measurement was used. The sample pieces were cleaved \textit{in situ} at room temperature before being transferred into the cold microscope within a few minutes, where they were zero-field cooled.

### 8.2. Local Density of States as a Function of Doping Concentration

The electronic density of states (DOS) of a material is number of states per unit energy per unit volume. The local density of states (LDOS) is the spatially dependent measure of the DOS, that is, how the DOS varies from location to location. STM is a local probe that allows the measurement of the LDOS by placing the STM tip over various locations of the sample and measuring the tunneling conductance there. STM topography is proportional to the integrated local density of states between the Fermi energy and the bias voltage. Therefore, since in GaMnAs \( E_F \) sits near the valence band edge, acquiring topography with a bias just below the conduction band provides a simple way of observing all the defect states that exist in the gap within some area. Such topographs are shown in Figure 8.1 for samples with 1.5 and 5% doping. The large number of states that can be seen arises mostly from randomly distributed Mn sites, but also from compensating defects due to the low temperature growth, such as As antisites and Mn interstitials. As the doping is increased, the number of states increases as expected. These images give some idea of the amount of disorder that exists in GaMnAs due to the high doping levels. Nevertheless, a well-ordered lattice can be seen in the topography, and no indication of Mn
Figure 8.1: (a-c) Topography of the in-gap states acquired just below the conduction band edge at 1.5V. The number of states visible gives some idea of the amount of disorder present in the sample between Mn acceptors and compensating donor defects. Zooming in (b,c), individual Mn wave functions can be identified in various layers. As the doping is increased identification of individual defects in layers other than the first or second becomes more difficult due to the increased overlap of the wave functions.
nearest-neighbor clustering, which would lead to anti-ferromagnetic interactions, was observed.

Long range potential variations, due to the random spatial distribution of positively charged compensating defects, shift the energies of the states up and down spatially. This can be seen in the changes of the predominantly observed bond direction of the lattice in different areas as the C3 surface state, which lies in energy near the conduction band and which extends predominantly along the [001] direction, is above the bias voltage of the tip in some areas and below it in others.\textsuperscript{(139)}

The energy dependence of the states was studied using STM dI/dV spectroscopy. Figure 8.2(a,b) shows the LDOS measured along a line, starting over the p-type buffer and extending deep into the GaMnAs heterostructure for the 1.5 and 5% samples. Compared with the Be-doped GaAs buffer, the valence band edge is observed at a higher energy in GaMnAs. This could be due to a larger amount of tip induced band bending occurring over the lightly doped GaAs buffer layer, where the electron density is lower and the screening is weaker, than over the much more highly doped GaMnAs. The result is also however consistent with the shifting of the valence band edge called bandgap renormalization that occurs in highly doped GaAs due to many-body effects.\textsuperscript{(23)} The energy of valence band edge also varies in space much more, suggesting the importance of band-tail states caused by disorder and compensation, as observed in the topography. In the gap, a wide distribution of defect states is observed above the valence band due to Mn acceptors. As expected, the number of states increases with increased doping. Some states deeper in the gap can be identified with donor defects from the local depressions they cause in the valence band topography due to charge induced band bending.

Focusing on the states near the valence band edge, due to the variation of the LDOS between locations, it makes sense to look at the average of spectra taken over many locations of each sample to get a sense of the average DOS. These are shown in Figure 8.2(c). For all dopings, there appears to be a distribution of Mn related impurity band states in the gap. These impurity
Figure 8.2: STM $dI/dV$ spectroscopy of the local density of states in GaMnAs. (a) Spectra probing from deep inside the valence band to the conduction band edge, acquired along a line of points from the Be doped p-type buffer layer into the 1.5% Mn doped material. A number of in gap states are observed in the GaMnAs region, as explained in Chapter 7. The edge of the valence band can be seen to shift about 200meV higher in energy and become very inhomogeneous as band tail states are formed due to variations in the local potential. (b) Same measurement for the 5% sample where a higher density of in-gap states is observed. (c) Spatially averaged spectra showing a suppression of the density of states around the Fermi energy due to electron-electron interactions with an approximately $E^{1/2}$ energy dependence.
band states exist well into the metallic regime, showing no obvious trend of collapsing into the valence band at higher dopings, where screening of the acceptor binding energy would be expected to be strongest. This is likely due to the short range non-coulombic terms in the Mn acceptor Hamiltonian, such as the p-d hybridization, which cannot be screened out. At the same time, there is no region in energy of zero density of states between the VB and IB states as would be observed if the impurity band was well separated from the valence band. Instead, a dip in the DOS near the valence band edge is observed that almost goes to zero in the 1.5% sample and is non-zero at higher dopings. This dip could be interpreted as the boundary between the impurity band and valence band states, which overlap more with higher doping. This interpretation is problematic, however, because the position of the dip would be expected to shift in energy between various dopings, and there is no obvious reason for it to always occur at \( E_F \). For this reason, we ascribe the dip in the density of states to a gap caused by carrier-carrier correlations that occur in disordered materials. Such gaps have previously been observed in planar tunnel junction measurements of GaMnAs. This implies that the overlap in energy of the impurity and valence band states is masked by the correlation gap and that the overlap has begun even on the insulating side of the metal-insulator transition. In principle, therefore, there will be holes residing in the valence band if the compensation is low enough. However, as the impurity band is composed of states lifted from the valence band, the crossover is not sharp. The shape of the gap has a roughly \( E^{-1/2} \) Altshuler-Aronov dependence, though it varies in space and is sometimes closer to the \( E^2 \) dependence of an Efros-Shklovskii coulomb gap. The existence of a correlation gap may also complicate the interpretation of optical adsorption measurements with excitations occurring across the region of suppressed DOS around \( E_F \).

As states of the valence band are Bloch-like extended plane waves and impurity states in the gap are exponentially-localized, distinguishing the character of the states at \( E_F \) is important. For
instance, the crystal momentum is a good quantum number for Bloch states, while it is not for localized ones. This will have important consequences of the types of interactions that can lead to itinerant carrier ferromagnetism in the material. Figure 8.3 shows maps of density of states of a 500 Å² area of the 1.5% sample at various energies above and below E_F. Below E_F, going deeper into the valence band, the states become more uniformly distributed, while in the gap the local density of states (LDOS) is predominantly localized on individual Mn dopant sites. At E_F, the LDOS fluctuates strongly in space acquiring a patchy, filamentary character. Such patterns are expected near criticality in an Anderson disorder-driven metal-insulator transition. As the states begin to experience localization, they go from a close to normal distribution to a skewed log-normal distribution with a long tail due to sites with rare, anomalously large, values of the local density of states. (59, 60) Fits in Figure 8.3(f) show that histograms of the LDOS are well described as being lognormally distributed. Such wide, non-self averaging distributions are a hallmark of the metal-insulator transition, where the variance of quantities measured near the transition, such as the conductance, is on the order of the mean value.

As the doping is increased and the magnitude of the suppression of the density of states at E_F decreases, the spatial distribution of the local density of states still varies strongly but begins to become less patchy (Figure 8.4). Because the scales of the distributions are different for the different dopings, it is convenient to normalize them by their arithmetic mean in order to put them on a common scale where they can be compared. This has the advantage of providing a simple measure of the weight of anomalously localized sites in the distribution. The arithmetic mean is a good estimator of the typical value of a normal distribution, but not for a skewed one, where it is determined by large but unlikely values in the tail of the distribution. The normalized distribution of metallic states will therefore be peaked (i.e., have its most likely value) near a value of one, while the peak will shift closer to zero as the states become more localized. Just such a trend can
Figure 8.3: (a-e) Spatially resolved dI/dV maps of the local density of states of the 1.5% plotted on a linear color scale that was saturated above 8 standard deviations in order to improve the contrast. States go from being uniform to very patchy at the Fermi energy (0 meV) before becoming highly peaked on individual defect sites deep in the gap. (f) Histogram of the local density of states at each energy. States deep in the valence band states are close to Gaussian distributed and develop long tails approaching the Fermi energy. Solid lines are fits to a lognormal distribution showing that the states can be well described as being lognormally distributed.
Figure 8.4: (a,b) Comparison of the density of states at the Fermi energy in the 1.5% and 5% samples respectively. Images are of 500Å x 500Å areas. (c) Probability distributions of the local density of states at the Fermi energy for 1.5, 3 and 5% Mn concentrations. The distributions are divided by their arithmetic mean to allow them to be compared on a common scale. All of the distributions have long tails corresponding to large rare values that make the mean (1 on the x axis) much larger than the typical value. As the doping increases the distributions become less skewed, suggesting that the states are becoming more uniform, as seen in (b). The solid lines are lognormal fits to the data. The dI/dV signal is small near $E_F$ and measurement noise broadens the distribution resulting in some negative values.
be seen as a function of doping in Figure 8.4(c), with the 1.5% sample, which is closest to insulating phase, having its most likely value the closest to zero.

The same trend can be shown as a function of energy, with extended states in the valence band becoming more localized approaching $E_F$ and crossing into the gap. This behavior can be analyzed in more concrete terms, using the parameters which determine the shape of the log-normal distribution,

$$f(x; \mu, \sigma) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln(x) - \mu)^2}{2\sigma^2}}$$

the geometric mean, $e^\mu$, and geometric standard deviation, $e^\sigma$. The geometric mean represents the typical (median) value of the LDOS, which has been suggested to be the order parameter of the Anderson metal-insulator transition, which would go to zero in the insulating phase. The geometric standard deviation is a measure of the variance of the LDOS and thus determines the length of the tails of the distribution. In the critical regions the fluctuations are expected to become of the same order of magnitude as the typical value, and thus a larger value of the geometric standard deviation indicates states that are more localized. These values are plotted as a function of energy in Figure 8.5. An anomaly near $E_F$ is seen in both, though it is stronger in the lower doped sample, where the typical value of the LDOS decreases toward zero and the fluctuations in the distribution increase. These measures indicate that the states near $E_F$ have become more localized than the states around them, likely due to the carrier-carrier interactions that are responsible for the correlation gap. A similar analysis can be done with the position peak, the mode, of the mean divided distribution and yields the same result.

### 8.3. Correlations

Having established that the states around $E_F$ show signs of increased localization, it can be asked whether they show other signs of being close to the metal insulator transition. The non-
Figure 8.5: (a,b) Mean divided distributions of the local density of states for the 1.5% and annealed 5% samples as a function of energy from states in the valence band to deep in the gap. States in the valence band are close to normally distributed and all states of the annealed 5% are less skewed as the material becomes more metallic. States deep in the gap may conduct by hopping over localized sites, which would also give a lognormal distribution. (c,d) Plots of the typical local density of states given by the geometric mean and the geometric standard deviation. Both show deviations near $E_F$ in the energy range where electron-electron interactions cause corrections to the density of states. The typical density of states has been proposed as a possible order parameter that goes to zero at the metal-insulator transition.
interacting Anderson transition is a continuous phase transition, where a correlation length is expected to diverge approaching criticality. On the insulating side, this length scale is associated with the localization length of states. At criticality, or on length scales less than the correlation length, the correlation falls off with distance as a power law with a critical exponent. Approaching the transition from the metallic side should be expected to show the same behavior.

To calculate this, the autocorrelation of the local density of states maps was computed. The correlation as a function of distance is plotted in Figure 8.6 for the energies around $E_F$. As might have been inferred by eye from the size of the patches of high and low conductance in Figure 8.3, the correlation length grows dramatically at $E_F$. At higher doping the same divergence can be seen. The length scale is much larger than the size of individual defects, which predominately determines the length at other energies. At $E_F$, the functional form can be fit to either a power law or exponential. The exponent of the power law fit, $r^{-\eta}$, gives $\eta \approx 1.2 \pm 0.3$. Power law behavior would be expected in the critical regime but measurements over much larger areas would be necessary to show it definitively. Fitting an exponent to the doping dependence of correlation length is beyond the reach of the measurements that have been made, however a rough estimate that the energy dependence goes close to of $E^{-1}$ appears to be justified. Unlike in the non-interacting case where there exists a mobility edge between localized and extended states at a certain energy, in the interacting case the states above and below $E_F$ have to be thought of as quasiparticle excitations in terms of Fermi liquid theory. Thus the exponent of the energy dependence is related to a dynamical critical exponent of the lifetime of quasiparticle excitations. Correlation effects responsible for the correction in the density of states near $E_F$ will cause the screening length to increase. In the limit where states become localized and the Efros-Shklovskii Coulomb gap opens, the screening length diverges. The regions of low density of states at $E_F$ tend to have gaps that are more parabolic in shape, like the Coulomb gap, suggesting that these areas
Figure 8.6: (a,b) Autocorrelation of the local density of states as a function of energy for 1.5 and 5% doping showing that the correlation length appears to be diverging at $E_F$. In (a) the dashed line is a guide showing that the correlation length falls off with approximately a power of $E^{-1}$. (c) Loglog plot of the distance dependence of the correlation. The correlation can be fit to a power law at $E_F$, while energies above or below it in the valence band or in the gap cannot. In the inset, the same data is plotted on a semilog plot showing the distance dependence could also be described as exponential, though with a much longer correlation length.
that are influenced by the long range Coulomb interaction which goes as \( E \propto 1/r \). This gives a simple explanation of the \( 1/E \) correlation length dependence.

### 8.4. Multifractal Analysis

Near the Anderson MIT states are known to exhibit multifractal wave functions on scales less than the quantum phase correlation length. Such patterns are usually described by looking at the self-similarity of the object at different length scales and calculating their singularity spectrum \( f(\alpha) \), which characterizes the whole set of fractal dimensions embedded in it. Such analyses are common of wave functions generated by numerical simulations but several complications arise when applied to experimental STM data. Firstly, the data were acquired at a finite temperature of 4K which will lower the length scale over which multifractal behavior can be observed due to thermal inelastic scattering. Estimates of the phase correlation length at these temperatures, however, are on the order of a few hundred of Angstroms, as are extrapolations from the temperature dependence of coherence lengths measured in transport at lower temperatures.\(^{16, 48}\) Secondly, our measurement of the LDOS averages over an unknown number of states, such that the spatial dependence of individual wave functions cannot be observed. Nevertheless, states that are close in energy should be correlated spatially so that the measured LDOS may still carry meaningful signatures of multifractality. This analysis would not be expected to work for states deep in the gap that are exponentially localized at different sites and the values obtained at those energies should be viewed critically. Lastly, we are looking at a 2 dimensional cut through 3 dimensional wave functions. This will lower the observed fractal dimensions of the wave functions but not prohibit the observation of fractal behavior.\(^{140}\) The \( f(\alpha) \) curves shown in Fig 4c were calculated by splitting the measured the LDOS into boxes of different sizes and using the method of Chhabra and Jensen.\(^{67}\) The most important value of the singularity spectrum is \( \alpha_0 \) which is the value at which the maximum of \( f \) occurs. This gives the scaling of typical values in the distribution, and deviations from the dimension of support of the measure, in this case a 2D
plane, represent anomalous fractal scaling. As plotted in Figure 8.7, $a_0$ is close to 2 for valence band states, shows an anomaly near $E_F$, and grows for states deep in the gap where the values cannot be taken seriously. Such an analysis has been previously carried out for states of a 2D electron gas in the integer quantum hall regime, but no shift of the value of $a_0$ away from 2 was able to be observed for states near the mobility edge.\textsuperscript{(141)} The value of $a_0$ we find is still much lower than predicted for the 3D Anderson transition. The dependence of the value on lower temperature measurements with finer energy resolution will determine the confidence that can be placed on this value. In particular, the shape of the multifractal spectrum can give access to useful quantities like $D_2$, which is calculated from the second moment of the distribution, and in the non-interacting case is related to the two particle Green’s function describing conduction.

Taken together, these measures of the spatial characteristics of the states and their scaling behavior at different length scales provide strong evidence for increased localization of the states around the Fermi level due to carrier correlations in a disordered material. These states show signs of multifractal wave functions, expected for states close to the metal-insulator transition. Such wavefunctions are known to cause anomalous dispersion of carriers due to the lower effective spatial dimension they occupy, and likely increase the electron-electron interaction effects. As the material becomes more metallic these effects decrease but do not disappear. The effects of a wide distribution of the local density of states and long range correlations on the carrier-mediated exchange interactions that make GaMnAs a ferromagnet are thus far largely unexplored and may explain why room temperature ferromagnetism predicted by the mean field Zener model has not yet been achieved.\textsuperscript{(3)}
Figure 8.7: (a,b) Multifractal spectra $f(\alpha)$ at various energies for the 1.5% and annealed 5% samples. The spectrum describes the full set of scaling exponents of the LDOS distribution. Consistent with the states becoming more uniform, the spectra are narrower for the 5% sample. The value of $\alpha$ where the peak of the spectra occurs is called $\alpha_0$ and describes the scaling of typical sites. Since the maps are 2D, values greater than 2 indicate anomalous scaling. (c) Plot of $\alpha_0$ versus energy, indicating the states at $E_F$ become more localized than those around them.
CHAPTER 9. CONCLUSIONS AND OUTLOOK

The experiments detailed in this thesis cover observations of magnetic impurities in a semiconducting GaAs host from the isolated impurity limit to the highly doped metallic phase. Using the atomically resolved STM probe, local properties of the Mott-Anderson metal insulator transition in a strongly disordered 3D metal have been observed. Metal insulator transitions where localization and carrier-carrier interaction effects are equally important are representative of most real materials but are still poorly understood theoretically. Measures of the distribution function of the local density states and the correlation length show the importance of interactions in localizing states near the Fermi energy. Properties of wave functions near the Anderson transition, such as multifractality, are observed, and rudimentary estimates of critical exponents are made. While finer measurements are needed to extract precise numbers, it is shown that such quantities are accessible by microscopic observations. This provides a motivation for further studies with local probes that may finally provide a consistent theoretical understanding of the metal insulator transition in disordered interacting systems such as this.

For the material properties of Ga$_{1-x}$Mn$_x$As, the main observation is that the electronic states near the Fermi energy are highly non-uniform, with a wide distribution in the values of the local density of states from point to point. This is expected as a consequence of disorder in any system close the metal-insulator transition, and the observation that these effects persist across a wide range of dopings suggests that the material is always, in some sense, close to the transition even in samples that are nominally metallic. This leads to long range correlations that are missed in a mean field treatment of ferromagnetism, and it will be interesting to see the results of theoretical models that take them into account. The local density of states spectra show no sign of detached impurity band but rather that the impurity band is merged with the valence band with a long tail.
extending out into the gap. As such, it is best described as a disordered valence band. As observed in transport measurements, electron-electron interactions are critically important for the states near $E_F$.

Observations of various transition metals substituted into the GaAs surface by STM manipulation showed the importance of the surface symmetry breaking on the resulting electronic states. Pairs of closely spaced impurities could be created at the surface and their microscopic interactions probed, showing a highly anisotropic coupling. This result resembles theoretical predictions of anisotropic exchange couplings between Mn impurities in bulk GaMnAs, but understanding of the effect that the surface has on these couplings requires further theoretical studies. Using the ferromagnetic MBE-grown heterostructure samples, individual Mn defects were studied in deeper layers than just the surface. The acceptor state energy becomes deeper, the closer to the surface the impurity sits. Consequently, states observed near $E_F$ are located deep beneath the surface and generally cannot be identified in topography. The observations of Mn acceptors in up to the first four layers can be made in the ferromagnetic material and are very similar to those reported in paramagnetic samples at much lower dopings by other groups. As the shape of the Mn wave function is expected to depend on the orientation of its $5/2$ core spin, this result is surprising. Defects that are isolated enough that their wave functions can be distinguished from the background environment may not be strongly coupled enough to become ferromagnetically aligned. Alternatively, the strain in the crystal near the surface may quench the spin degree of freedom. Future spin polarized STM measurements could be useful to resolve this.

Looking forward, there are a number of directions for future research. Further studies of the metal insulator transition in GaMnAs done at lower temperatures, finer energy resolution, and acquired over larger areas could extract the full shape of the multifractal spectrum and determine the critical exponents of the transition. Studies in other strongly disordered semiconductors, which likely belong to other universality classes, could be done as well for comparison. One
possibility is Cd doped GaP which has a moderately deep acceptor energy, like Mn in GaAs, but where the material should be nonmagnetic with weak spin-orbit coupling, unlike GaAs.

How the character of the states near the Fermi energy in GaMnAs changes in even higher doped samples than those studied, with $T_c$’s close to the highest attainable values today, will be important to observe. Whether or not strong fluctuations in the local density of states persist in these samples will determine whether the Zener model is qualitatively correct or not in this limit. Other possible measurements are more closely tied to the magnetic properties. There is reason to believe, based on the large tunneling anisotropic magnetic resistance (TAMR) effects seen in planar tunnel junctions, that interesting effects could be seen in STM studies of GaMnAs with a vector magnet.\(^{(42)}\) One particularly easy measurement would be to observe the effect of the rotation of the easy axis of the material from a biaxial [100] direction to [110] as the sample is warmed towards $T_c$.\(^{(35)}\) This could be particularly powerful if combined with spin-polarized STM. Spin-polarized measurements of low doped samples near the metal-insulator transition with very low magnetic transition temperatures could also study the critical behavior of percolation networks of bound magnetic polarons that give rise to the ferromagnetic state.

The study of phase transitions at the atomic scale is one of the most interesting applications of local probes. Whether or not GaMnAs ever becomes technologically relevant, it turns out to be a good material for studying the physics of the metal insulator transition. As shown, STM can provide much insight into the microscopic processes at work. Surely, our understanding of them will continue to be extended in the future by studies of this and other materials using this unique tool that provides a window into a world within our own.
REFERENCES

52. K. Slevin, T. Ohtsuki, Physical Review Letters 78, 4083 (1997).
AUTHOR’S BIOGRAPHY

Anthony Richardella was born in Washington, DC on July 18, 1977. He graduated from The College of William and Mary with honors with a Bachelor of Science in Physics and a Bachelor of Science in Computer Science. For three years he worked at SRA International in Fair Lakes, Virginia working on machine learning algorithms for natural language processing. He then moved to Champaign, Illinois to pursue graduate studies in condensed matter physics. After joining the research group of Prof. Ali Yazdani, he completed a Master of Science in 2005. In 2006 Prof. Yazdani’s group moved to Princeton University in Princeton, New Jersey where a large part of the research presented in this dissertation was completed. Following the completion of his PhD he will start a post doctoral research position working with Prof. Nitin Samarth at Penn State University in State College, Pennsylvania.