# Single Particle Spectrum of the Two Dimensional Electron Gas

by

Oliver Eugene Dial III

Submitted to the Department of Physics
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2007

Author . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Department of Physics
March 13, 2007

Certified by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Raymond C. Ashoori
Professor
Thesis Supervisor

Accepted by . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . .
Thomas J. Greytak
Professor, Associate Department Head for Education

# Single Particle Spectrum of the Two Dimensional Electron Gas

by

Oliver Eugene Dial III

Submitted to the Department of Physics
on March 13, 2007, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

## Abstract

Accurate spectroscopy has driven advances in chemistry, materials science, and physics. However, despite their importance in the study of highly correlated systems, two-dimensional systems (2DES) have proven difficult to probe spectroscopically. Typical energy scales are on the order of a millielectron volt (meV), requiring high resolution, while correlated states of interest, such as those found in the integer and fractional quantum Hall effect, are destroyed by excessive electron heating. Approaches based on tunneling have been hampered by problems such as ohmic heating and low in-plane conductivity, while optical approaches probe long-wavelength excitations which can be difficult to interpret.

Here we present a refined spectroscopic technique, time domain capacitance spectroscopy (TDCS), with which we measure the single particle density of states (DOS) of a 2DES with temperature-limited resolution. In TDCS, sharp voltage pulses disequilibrate a metallic contact from a nearby 2DES, inducing a tunnel current. We detect this current by monitoring the image charge of the tunneled electrons on a distant electrode. No ohmic contact to the 2DES is required. The technique works when the 2DES is empty or has vanishing in-plane conductivity, as frequently occurs in studying the quantum Hall effect.

Using TDCS, we perform unprecedentedly high resolution measurements of the DOS of a cold 2DES in GaAs over a range from 15 meV above to 15 meV below the Fermi surface. We provide the first direct measurements of the width of the single-particle exchange gap and single particle lifetimes in the quantum Hall system. At higher energies, we observe the splitting of highly excited Landau levels by spin polarization at the Fermi surface, demonstrating that the high energy spectrum reflects the low temperature ground state in these highly correlated systems. These measurements bring to light the difficult to reach and beautiful structure present far from the Fermi surface.

Thesis Supervisor: Raymond C. Ashoori
Title: Professor

# Acknowledgments

There is no way I could list everyone that has helped me in my time at MIT or even everyone that has contributed to the work that has gone into this thesis. Thank you all. First and foremost, I'd like to thank my advisor Ray Ashoori – without him, his original vision of the pulsing experiment, his patience (or perseverance!), and his willingness to let me explore many different projects, the beautiful results presented in this thesis would never have come to be. His continuous excitement about science and ability to communicate that wonder through intuitive, physical models of complex phenomena make him a great teacher both in the classroom and in the lab. His willingness to try new measurements and new ideas in the lab have made every day as exciting as the first day I set foot in MIT. I'd also like to thank Gary Steele for being available to bounce ideas off, and teaching me more about physics than I'd ever have learned on my own. The various members of our group have been invaluable, both in developing experimental techniques and helping when things "go wrong," and in challenging each others ideas to help refine them. Aleem Siddiqui, in particular, helped keep the pulsing experiment intact and operational and helped me learn how to run it, while Michael Jura was instrumental in the construction of much of the precision electronics that drives it. The larger community in building 13, including the Greytak and Kastner groups, creates a wonderful place to work where there is always somebody around to discuss research with, or from whom to borrow a tool or fitting.

On a more personal level, I'd like to thank those who have supported me through the past years, especially my (very patient) wife Miranda (who also happens to be the first person to proofread this from cover to cover: thank you!). When things have gone well with my experiments, she's shared in my enthusiasm; more importantly, when things have gone badly, she's been a source of reassurance and encouragement. I'd also like to thank my parents Gene and Joyce, and my sister Leah. Pursuing a PhD is a long endeavor, and they've helped me continue my life outside of lab during this time. I'd also like to thank Corey Reed, Peter Monaghan, Jay Kane, Tula, Ophie, Sami, and the THDA for making my life much richer in my off-hours.

# Contents

# List of Figures

# Chapter 1

# Introduction

Accurate spectroscopy has frequently transformed our understanding of key scientific problems. From the development of quantum mechanics to describe the hydrogen spectrum to the use of emission spectra to determine the makeup of distant stars, many of the great advances of the last hundred years have resulted from asking the simple question: at what energies can this system be excited?

It is then no coincidence that many of our theoretical frameworks for describing the world around us, from a classical Hamiltonian to the most complicated Feynman diagram, provide as their most direct result the excitation spectrum of a system.

However, the two-dimensional electron gas (2DEG), which in the last 40 years has displayed a rich variety of phases and phenomena[1], has proven difficult to probe spectroscopically. Extensive measurements of its conductivity[2], magnetic moment[3], compressibility[4], and even its interaction with nuclear spins[5] have built up an impressive theoretical and experimental understanding of its rich features. However, these measurements only tell us about excitations of the 2DEG within a few $k_b$T, typically a few tens of microvolts, of the Fermi surface. The entire range of energies from the most tightly bound electrons tens of millielectron volts (meV) below the Fermi energy, to the breaking point where the 2DEG becomes three dimensional hundreds of meV above the Fermi surface remains essentially unexplored. Our basic question, "at what energies can this system be excited?” goes largely unanswered. There are many reasons for this.

Approaches based on optical excitations have difficulty dealing with the small energy scales, on the order of a meV or less, that characterize the 2DEG. Raman-like techniques[6] can still be used, but the wavelength of light at useful energies is enormously long compared to the characteristic length scales of a 2DEG, making it difficult to generate excitations that carry interesting information out of the system.

Tunneling measurements, as discussed in detail in the next chapter, can measure at which energies electrons can enter or leave a system. However, in steady state measurements, electrons must both enter and leave the system. This is tricky to arrange with a 2D system, but there are two possibilities. The current could flow out through the plane of the 2DEG, but this is only possible if the 2DEG has a high conductivity. However, the conductivity of the 2DEG at high magnetic field falls to zero in quantum Hall states. Alternatively, one can have electrons enter the 2DEG from below its plane, and exit above. Because there is no lateral current flow in this configuration, the vanishing conductivity of the 2DEG is not an issue. However, there are now two separate tunneling events; one when the electron enters the 2DEG, and a second when the electron leaves. Trying to separate the effects of these two events without changing the 2DEG’s density by allowing excess electrons to accumulate has proven difficult. Furthermore, the electrons relax on a timescale of picoseconds once they have entered the 2DEG; if they are injected at high energies, unacceptable heating can occur.

In this thesis, we present a tunneling measurement technique that allows us to measure the tunneling spectrum of a 2DEG with temperature-limited resolution over a 30 meV range of energies yet still maintain good control over the density of the 2DEG. We open up a new direction of exploration in this rich system: energy from the Fermi surface. We do this by moving away from steady-state measurements; this allows us to tunnel electrons into the 2DEG without providing a path for the current to exit. Not only does this remove the confounding artifacts discussed above, but it permits us to inject electrons at arbitrarily high energies while performing our measurement before significant heating has occurred.

The spectra that we measure show in a beautiful and concise way the physics of the integer quantum Hall effect in an electron gas. Using them, we provide simple and direct measurements of the exchange-enhanced spin gap and the finite lifetime for excited electrons. We see, for the first time, the behavior of states far from the Fermi energy as we vary the electron density. We answer the basic question, "at what energies can this system be excited?"

## 1.1 Organization

In the next chapter, we introduce the relationship between the single-particle spectrum and many-body theories in somewhat more detail and review 2D systems and the quantum Hall effect with special attention to energy spectra. In Chapter 3, we introduce our measurement technique, time domain capacitance spectroscopy (TDCS), and briefly discuss its capabilities and limitations. Enough detail is given to understand our results, but important details of the calibration and implementation are left to later for the sake of clarity. In Chapter 4 we provide copies of most of our data printed together at the same scale for easy comparison. We then discuss, in chapters 5-8, some of the physical phenomena we observe in these spectra.

In chapters 9-12, we revisit TDCS in much more detail, with a specific emphasis on calibration of the axes of our spectra and on systematic errors. Finally, in the appendices, we discuss the details of our simulations and experimental setup.

# Chapter 2

# Background

The physics of a single quantum-mechanical particle in isolation is well understood and typically dealt with easily. However, when two or more particles are allowed to interact, they can form correlations between each other which have no analogue in classical mechanics. As the number of particles is increased, the number of possible configurations of the system quickly grows unmanageably large.

To approach these difficult problems, we describe the system in terms of the available excitations away from its ground state. Some of these excitations are long-lived. Because these long-lived excitations give near-singular contributions to the response of the system to external perturbations, they dominate the observable physics[7].

In a complex system, many different types of excitations are possible. One of the simplest is the quasiparticle, which describes the system with an extra particle added and the screening response of the many body ground state. In the case of an electron gas, we can imagine a quasiparticle as consisting of an extra electron plus a screening cloud of virtual plasmons that create a localized pocket of positive charge around it[8]. Because we have attached this screening cloud, the quasiparticles of a system will have different properties from the bare particles, including different dispersion relations and energy spectra. Although our system may have strong interactions, the residual interactions between two of these

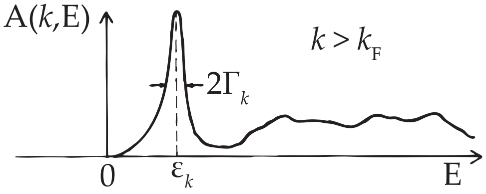


Figure 2-1: **Simple spectral function** (Schematic) probability $A_k$ of finding an electron injected with wavevector $k$ at an energy E. Adapted from [7]. The quasiparticle is identified as the peak at $\varepsilon$, with lifetime $\Gamma$. This spectral function is that for an electron, at a wavevector $k > k_f$.

quasiparticles (electrons + screening clouds in the case of the electron gas) tend to be weak. This allows us to use these quasiparticles to describe our complex, interacting system in terms of simple single particle physics [9, Chapter 0].

Other long-lived excitations can describe, for example, neutral collective modes of the system. One well known example for an electron gas would be the plasmons we have already mentioned (a propagating charge density wave). However, a wide variety of other collective excitations exist.

The Green's function provides a powerful tool for locating and describing these quasiparticles. It allows us to ask questions like, "if I excite this system from its ground state by injecting an electron at a wavevector $k$, what is the probability of finding the new N+1 particle system at an energy E?" That is, if $|0, \mathrm{N}\rangle$ is the ground state of the N particle system, $|n, \mathrm{N}+1\rangle$ are the exact eigenstates of the N+1 particle system with energies $\mathrm{E}_{n,\mathrm{N}+1}$, and $c_k$ is the single-particle creation operator with wavevector $k$,

$$\mathrm{A}(k, \mathrm{E}) = \sum_n \langle n, \mathrm{N}+1|c_k|0, \mathrm{N}\rangle \, \delta(\mathrm{E} - \mathrm{E}_{n,\mathrm{N}+1}) \tag{2.1}$$

This probability distribution is called a spectral function, and denoted by $\mathrm{A}(k, \mathrm{E})$. Although the full probability distribution gives the most accurate description of our many-body system, it tends to be dominated by a number of approximately Lorentzian peaks. We can locate the peaks in this probability by $\mathrm{A}(k, \mathrm{E})$, and identify them as quasiparticles; see Figure 2-1. Furthermore, we can identify

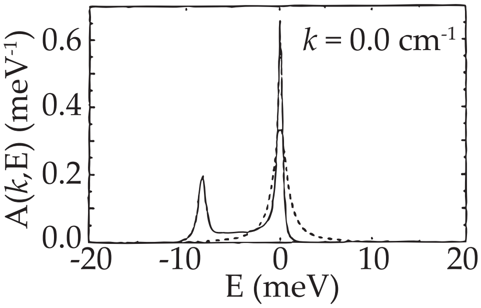


Figure 2-2: **Spectral function for holes with multiple quasiparticles:** A calculated spectral function for a two-dimensional electron gas in GaAs, adapted from [11]. The lower (more negative) energy peak is a plasmaron, while the higher (closer to zero) energy peak is a conventional quasiparticle. This spectral function is that of a hole with $k = 0$.

the half-widths $\Gamma_k$ of these peaks as the quasiparticle lifetimes, as discussed in Chapter 7.[†] We can then extract the properties of these quasiparticles from the behavior of the peaks in the spectral functions; for example, the dispersion of the energy $\varepsilon_k$ of the peak with $k$ vector can be used to compute an effective mass.

Notice that we are neglecting the long, smooth background of excitations in the probability distribution. This background is made up of more complicated excitations, such as creating an electron plus some distribution of plasmons or phonons. Typically, the coupling between the injected electron and these other excitations are weak. These complex excitations will then carry very little spectral weight because there they are almost orthogonal to the single electron at wavevector $k$.

In some cases there can be strong coupling which mixes the electronic state with these collective excitations. If this occurs, there may be additional quasiparticle peaks; see Figure 2-2. In particular, this has been predicted to occur in electron gases when there is resonant coupling between plasmons and electrons [12]. This quasiparticle, consisting of an electron screened by real plasmons[8], is called a “plasmaron” to distinguish it from the more conventional quasiparticle state.[†]

---

[†] More typically, the quasiparticles are identified using the poles in the self-energy function; see references [9] and [10].

[†] This name is sometimes also applied to a strongly coupled plasmon and phonon, particularly

Almost any measurable property of an interacting system can be calculated approximately using quasiparticles[14, 15], or more precisely using full Green's functions and spectral functions. However, some measurements can give direct access to these spectral functions and thereby the energies and lifetimes of quasiparticles in a system. These measurements provide a powerful probe as they can be directly compared to the most basic results in our many-body theories.

Conceptually, these measurements work by attempting to inject (or eject) an electron with a specific energy $\mathrm{E}$ and wavevector $k$. The spectral function $\mathrm{A}(k, \mathrm{E})$ is then the probability that this attempt will succeed. If we perform many attempts and examine the number that succeed, this number will be proportional to $\mathrm{A}(k, \mathrm{E})$. Because the spectral function is strongly peaked at the quasiparticle energies, when we inject an electron at an energy where the spectral function is peaked, we interpret this as having created that quasiparticle.

## 2.1 Observing Quasiparticles: The Single Particle Density of States

Perhaps the simplest measurement that can access these spectra in detail is x-ray photoemission spectroscopy (XPS)[16]. A high energy x-ray of known wavelength is used to excite electrons out of the material being studied into the vacuum. These electrons are then focused into a spectrometer and their energies are measured. The difference between the energy of the x-ray and the energy of the electron gives the energy needed to create a quasi-hole in the material being studied. In these simplest photoemission measurements, the $k$-vector of the quasi-hole left behind is not measured. These measurements average across all possible $k$-vectors, giving an electron yield as a function of energy proportional to

$$g(\mathrm{E}) \propto \int d^{\mathrm{N}}k\mathrm{A}(k, \mathrm{E}) \tag{2.2}$$

---

within the body of literature regarding electron energy loss spectroscopy. Care must be taken to identify which excitation is being referred to; compare, for example, references [12] and [13].

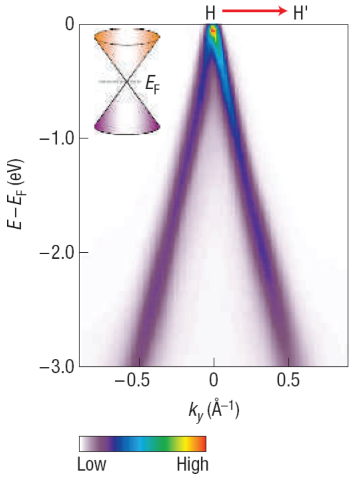


Figure 2-3: **Sample ARPES spectrum** adapted from [18], showing the k-vector dispersion of the valence band in graphite. The horizontal axis is wavevector, the vertical axis is energy, and the color is the electron yield. Note that only quasi-hole states are detected; the conduction band is missing. Also notice the coarse energy scale; the full range is several eV.

where N is the dimensionality of the system. $g(\mathrm{E})$ is referred to as the single-particle density of states, and can be thought of the density of single-electron like excitations of the system. We will refer to the single-particle density of states as the density of states (DOS) and will be more specific when referring to other densities of states.

In semiconductor systems, ultraviolet light is typically more appropriate to the energy range of the conduction band, leading to ultraviolet photoemission spectroscopy (UPS). Angularly resolved photoemission spectroscopy (ARPES) builds on this technique by measuring both the energy and the momentum of the emitted electron, allowing it to directly measure the $k$-dependent hole spectral function[17]. This allows the direct imaging of semiconductor band structure, as illustrated in Figure 2-3.

While these techniques are powerful, they have several limitations. Because they need to eject electrons, they can only investigate filled states. That is, they probe quasi-hole states, not quasi-electron. (Other techniques, including inverse photoemission spectroscopy, allow probing of only the empty states. It is

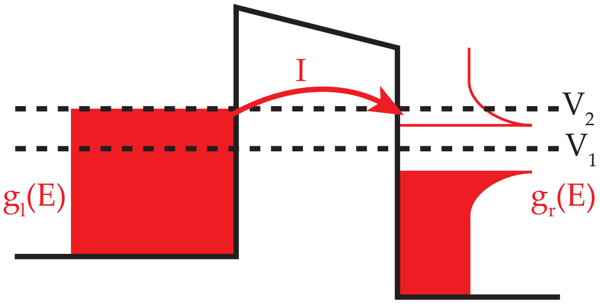


Figure 2-4: **Schematic of a tunneling measurement** between a metal and a superconductor. $g_l$ is the DOS of the left electrode, and $g_r$ is the DOS of the right electrode.

possible to measure the entire density of states at the expense of combining several techniques; however, at present the energy resolution of inverse photoemission spectroscopy is significantly worse than that of ARPES[19]) Because the electron must escape the material being studied, photoemission spectroscopy cannot be used to study structures buried deep within a sample where the highest mobility heterostructures currently available are located. Finally, while the energy range and resolution are excellent for studying band structure, they are presently inadequate for the $\mu$eV energy scales that dominate correlated physics in semiconductor structures.[†]

An alternate approach to measuring these spectral functions is through tunneling measurements. Here, two electrodes are separated by a barrier sufficiently thin to allow electron tunneling, but thick enough that the electrodes do not interact with each other appreciably; see Figure 2-4. If there is a potential difference $V_1 - V_2$ between the electrodes, we can calculate the tunnel current using Fermi's golden rule;

$$I \propto \int_{V_1}^{V_2} g_l(v) g_r(v) \left| M(v) \right|^2 dv \tag{2.3}$$

where $M(v)$ is a matrix element that describes the coupling between the elec-

---

[†] However, the energy resolution of ARPES is improving rapidly as spectrally pure photon sources become available, with sub-meV resolutions becoming possible today[20].

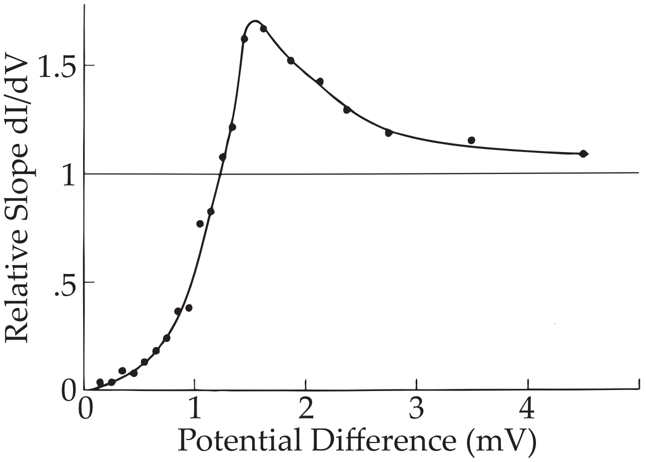


Figure 2-5: **Tunneling measurement of the superconducting gap** adapted from [21]. The vertical axis is dI/dV, proportional to the single particle density of states $g(v)$. The strong suppression of the density of states near zero energy is evidence for the existence of the quasiparticle gap.

trodes; methods for calculating $M$ are discussed in Chapter 10. Intuitively, this means the contribution to the tunnel current by electrons of energy $v$ is just proportional to the density of electron states in the left electrode at $v$ times the density of hole states in the right electrode at the same energy.

In a typical tunneling measurement, the tunnel barrier and the left electrode are selected to be uninteresting in the sense that they have neither energy nor wavevector dependence; in this case, we can pull them out of the integral;

$$\mathrm{I} \propto g_l \, |M|^2 \int_{\mathrm{V}_1}^{\mathrm{V}_2} g_r(v) \tag{2.4}$$

Differentiating then gives

$$\frac{\partial \mathrm{I}}{\partial \mathrm{V}} \propto g_r(v) \tag{2.5}$$

Tunneling measurements have been key in understanding superconductors, providing the first direct experimental confirmation of the existence of the quasiparticle gap, as illustrated in Figure 2-5[21]. Despite the fact that tunneling measurements cannot directly measure the $k$ dependence of the spectral function, the existence of a gap gives a distinctive, strong feature in the DOS. In general, the DOS is most useful as a probe when there are gaps or peaks in the DOS; a spectrum like that shown in Figure 2-3 would be fairly uniform and uninteresting as a function of energy without the wavevector dependence.

In more general circumstances, there can be rich quasiparticle structure in the $k$-dependent spectra which is unobservable using tunneling because there are no forbidden ranges of energies; that is, the momentum-averaged DOS can be featureless.

Notice that the single particle density of states, as defined above, is a distinct quantity from the thermodynamic density of states, $d\mu/d\mathrm{N}$ where $\mu$ is the chemical potential. The single particle density of states tells us about the energy of excitations of the system away from its ground state, while the thermodynamic density of states tells us how much the ground state energy changes when an additional particle is added to the system. These can look very different; for example, in superconductors the thermodynamic density of states is flat while the tunneling density of states is gaped.

## 2.2 Semiconductors and MOSFETs

Semiconductor systems have proven to be powerful environments for studying the interactions in many-particle systems. At low electron densities the Coulomb interaction between electrons is poorly screened, so interactions strongly modify electrical properties; this is in sharp contrast to metals, where electron-electron interactions are heavily screened. There are a wide variety of mature techniques for probing the electrical properties of semiconductors, making sophisticated measurements possible. Finally, there is expertise in creating highly pure semiconductors, reducing the role of impurities that might mask many-body effects.

In three-dimensional semiconductors, electrons[†] are placed in the system by doping the semiconductor, intentionally introducing charged impurities to the

---

[†] One can substitute electrons for holes and valence band for conduction band in this discussion; we concentrate on electrons for simplicity

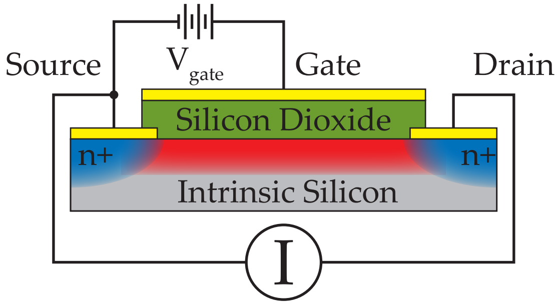


Figure 2-6: **A silicon metal oxide semiconductor field effect transistor (MOSFET)** In a MOSFET, an electric field applied by biasing a nearby gate electrode causes electrons to accumulate at the interface between a semiconductor (silicon, in this case) and an insulator (silicon dioxide). Because no donors are needed to generate these electrons, higher mobilities than those found in bulk materials can be achieved. Silicon is a popular material system for forming MOSFETs because of the relative ease and high quality of $SiO_2$ formed by thermally oxidizing the surface of silicon in an oxygen environment. Doped regions (blue) act as conductors to allow charge to flow into and out of the channel region (red).

crystal. These dopants are selected to have small ionization energies, so they will tend to give up an electron to the conduction band of the crystal while the atom left behind forms a charged impurity.[‡] Thus, there is a randomly placed impurity for electrons to scatter off for every electron in the conduction band. This scattering tends to dominate the electrical properties of three-dimensional semiconductors, making it difficult to study electron-electron interactions.

However, in fewer than three dimensions, it is possible to sidestep this problem. In devices like field-effect transistors (FETs), an electric field is generated using a distant electrode which is used to “pull” electrons up to an insulating interface, creating a sheet of charge in a semiconducting layer in the absence of donors (Figure 2-6). The higher mobilities possible in these devices allow the observation of a number of important effects, including Shubnikov-de Hass (SdH) oscillations[2].

These SdH oscillations originate with the formation of Landau levels by the electrons in the semiconductor. Classically, when a charged particle is placed

---

[‡] At low temperatures or high magnetic fields, the donors will recapture their electrons, and will form a Mott insulator unless the donor density is large enough to form an impurity band[22]

in a magnetic field B, its motion can be separated into the direction parallel to the magnetic field, in which it moves with fixed velocity, and the direction perpendicular to the magnetic field, in which it describes cyclotron orbits. For now we will neglect the velocity parallel to the magnetic field; we will return to this point later when we discuss in more detail how to construct an effectively two-dimensional system. These cyclotron orbits have a radius $R_c = v/\omega_c$, such that the Lorentz force on the particle provides the centripetal force needed to keep the particle on its circular path. Here $v$ is the particle's velocity, and $\omega_c$ is the cyclotron frequency $eB/m$, with $e$ the particle's charge, and $m$ its mass. Regardless of the particle's velocity, it completes one orbit in a time $2\pi/\omega_c$.

Semi-classically, we can argue that these cyclotron orbits will have to contain an integer $N$ number of de Broglie wavelengths; that is, $2\pi R_c = Nh/mv$. This requirement will discretize the available velocities $v_N = N\hbar\omega_c/m$, and create a discrete energy spectrum, with states only at energies $E_N = N\hbar\omega_c/2$. This qualitative argument gives a slightly incorrect constant; a full treatment gives $E_N = (N + {}^1\!/_2)\hbar\omega_c$. For any given $N$, we can have electrons with orbit centers at many different points in the space; because of this, each of these energy levels will be massively degenerate. Detailed treatment shows the degeneracy is $eB/h$, which corresponds to one electron for each quantum of magnetic flux. For each $N$ and point in space, we can also have one of two spin states, giving rise to another factor of two in the degeneracy. These massively degenerate, discrete energy levels are known as Landau levels, and they are referred to by their orbital level index $N$ (Figure 2-7a). At each orbit center, each orbital level $N$ corresponds to a different, orthogonal wavefunction; see Figure 2-7b. Each of these Landau levels gives rise to a $\delta$ function in the density of states.

When the effects of disorder are included, each of these $\delta$ function peaks in the density of states is broadened into a series of narrow peaks (Figure 2-7c). The tails of these peaks are made up of localized states, unable to carry current, while the main bodies of the peaks are made up of delocalized states; see Figure 2-7d.

Only the states near the Fermi surface of the channel of the MOSFET carry

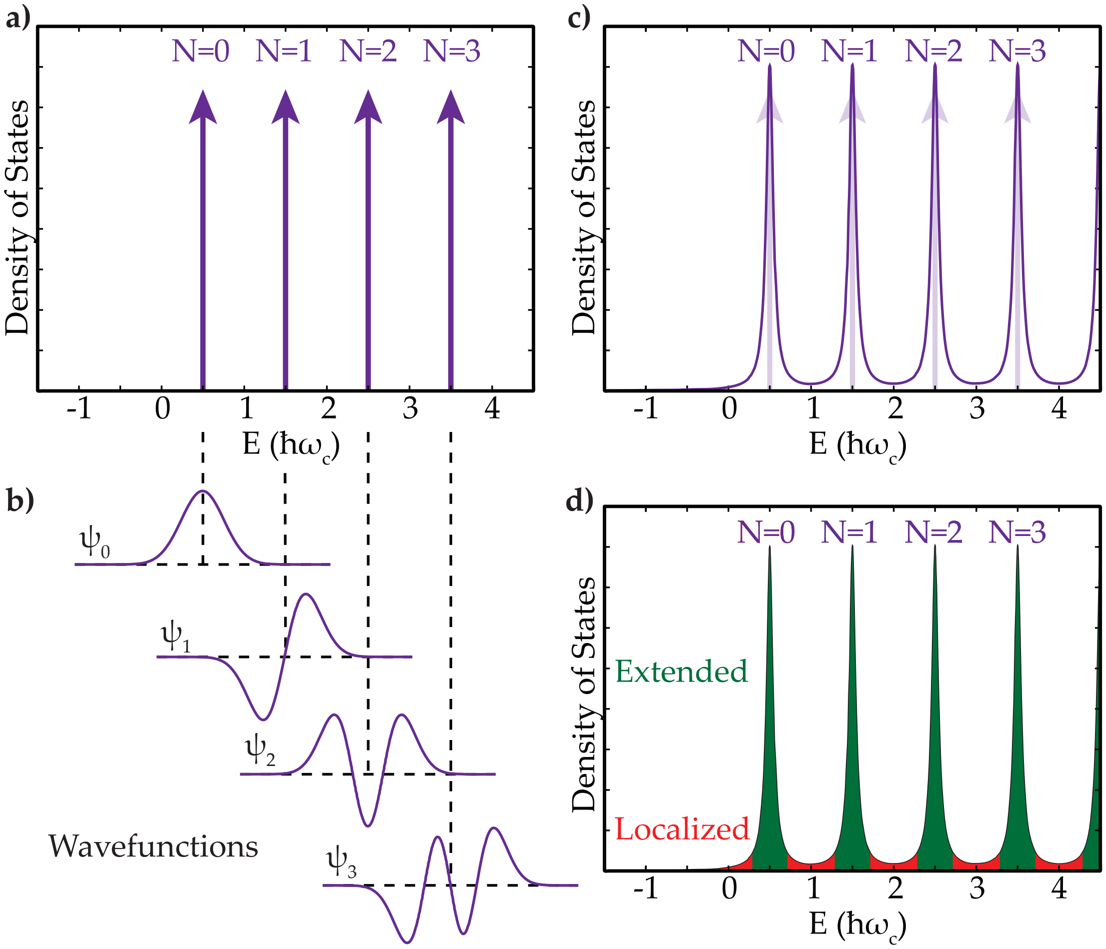


Figure 2-7: **Landau levels in an empty system: a)** The Landau spectrum for an empty system, neglecting disorder. Each evenly-spaced $\delta$-function in the density of states has the same degeneracy. **b)** The orbital wavefunctions associated with the first four Landau levels, shown in the Landau gauge. **c)** The presence of disorder broadens the $\delta$ function density of states associated with each Landau level into a narrow peak. **d)** The states that make up the middle of each peak are delocalized, and can carry charge across a sample. Those in the tails are localized, and cannot carry current.

current. If density of the MOSFET is such that the Fermi energy lies in the tails of the Landau level peaks, those states will be localized and it will have a high resistance. If, on the other hand, it falls in a region of high density of states, it will have a low resistance. This gives rise to the SdH oscillations seen in these devices.

## 2.3 Heterostructures

The interface between the semiconductor and oxide in a MOSFET is a messy place; it is the boundary between a crystal and a glass, and tends to be full of crystalline faults and charged impurities. This limits the electron mobility in these devices. However, there are semiconductors available with a wide variety of band gaps and lattice constants. By selecting two which have approximately the same lattice constant but very different bandgaps, it is possible to create barriers to electrons that do not suffer from these difficulties. This approach can be realized using molecular beam epitaxy, which allows deposition of atomically thin layers of material with precisely controlled stoichiometry. Barriers can be constructed by abruptly varying the material composition during growth, switching from one semiconductor to another.

GaAs and $\mathrm{Al_xGa_{1-x}As}$ (AlGaAs) are a popular choice in these heterostructures; the bandgap widens smoothly as the Aluminum concentration is increased in the AlGaAs, and they are almost precisely lattice matched across the entire range $x = [0, 1]$. In a simple device, a thin layer of GaAs might be surrounded by AlGaAs on both sides. Donors placed in one or both of the barriers ionize, and the electrons they give up can lower their energy by moving out of the AlGaAs and into the GaAs; see Figure 2-8a-b. Much as with the MOSFET, we can have electrons without nearby donors, giving high mobilities.

So far, we have not concerned ourselves with the electron's motion perpendicular to the oxide in the case of the MOSFETs or perpendicular to the well in the heterostructures. The electrons are confined to a thin volume along the barrier; their wavefunction perpendicular to the barrier will look like a standing

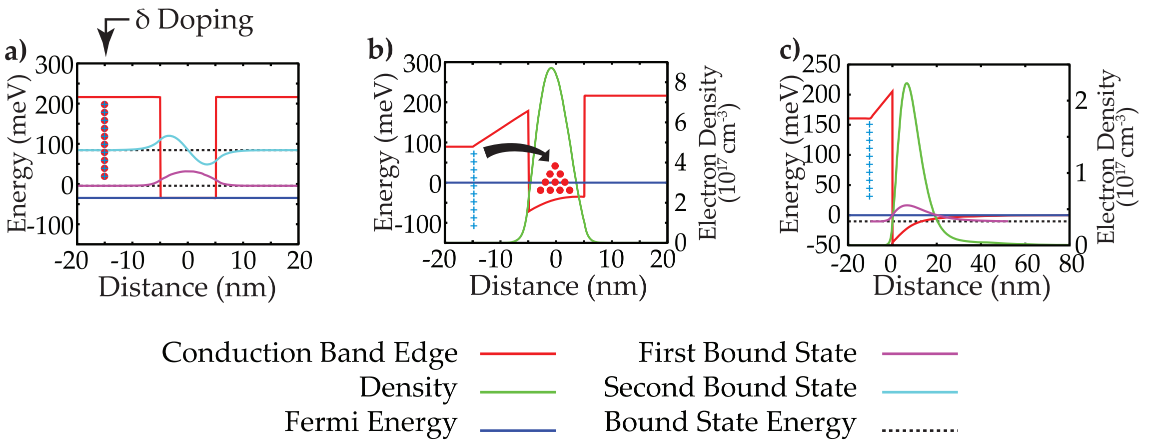


Figure 2-8: **Heterostructures and quantum wells:** **a)** A thin layer of GaAs in grown in between barriers of AlGaAs. Donors are placed inside one or both of the AlGaAs barriers. **b)** The electrons from the donors are able to lower their potential energy by moving out of the AlGaAs and into the GaAs. They create a pool of free electrons able to move freely in two dimensions, and widely separated from the donors. **c)** In single interface wells, one of the barriers is eliminated. The potential well created from the charged electrons themselves confines them to near the barrier.

wave. If the confinement in these structures is sufficiently tight, the energy separation between the lowest energy standing wave state (bound state) and the next highest can be large compared to other energy scales of interest, notably the temperature, the Fermi energy of the electrons, and Coulomb potential between nearest neighbor electrons in the 2DEG; sample bound states and energies are shown in Figure 2-8. In this case, all of the electrons will be confined in this lower bound state of the quantum well, and excitations, real and virtual, to the next bound state will be negligible. It is this inability to change the electronic wavefunction in one direction that leads us to describe these systems as two-dimensional electron gases.

One consequence of this confinement is a simplification of the energy spectrum for the case of a magnetic field applied perpendicular to the plane of the 2DEG. The electrons have no accessible degrees of freedom in the direction parallel to the magnetic field. There are now only four energy scales of note in these devices; the Coulomb energy, which we can tune through the electron density, the cyclotron energy, tunable through the magnetic field, the Zeeman energy, which we can adjust by applying fields in the plane of the quantum well, and the disorder potential. Three of the four energy scales of note are easily experimentally controllable in-situ, allowing us to easily adjust the relative importance of different effects.

## 2.4 Quasiparticles and 2D Systems

With the discretization of the spectrum by the combination of vertical confinement and the application of a large magnetic field, the non-interacting density of states is split into a series of discrete peaks, as shown in Figure 2-7. Because of this splitting, there is the potential to be able to observe quasiparticle energies and lifetimes in the spectrum, even without momentum resolution. However, the spectrum shown in Figure 2-7 is that for non-interacting electrons. Because we expect the quasiparticles to have similar properties to the bare electrons, it is useful to assume the interacting spectrum will be similar and to qualitatively discuss what changes we expect to see as a function of density. This will provide a ready-made nomenclature for labeling quasiparticle states in the interacting electron gas. This does not result in an exact description of the quasiparticle spectrum but gives us a qualitative idea of what to expect.

Conventionally, the density of a two-dimensional electron gas is described in terms of the filling fraction $\nu = \sigma/(eN_D)$, where $N_D$ is the Landau level degeneracy. Because of the additional spin degree of freedom, the filling fraction is $\nu = 2$ when the first orbital level is filled, $\nu = 4$ when the second level is filled, etc. We will use $\nu$ exclusively to refer to the density of the electron gas.

At $\nu = 0$, there are no electrons in the 2DEG, so the DOS is just that of noninteracting particles (Figure 2-9a). We expect there to be a series of disorder-broadened peaks corresponding to putting an electron into each orbital Landau level, $N = 0, 1, 2, ....$ Because the quantum well is empty, all of these lie above the Fermi energy. Neglecting the Zeeman splitting, each of these levels carries

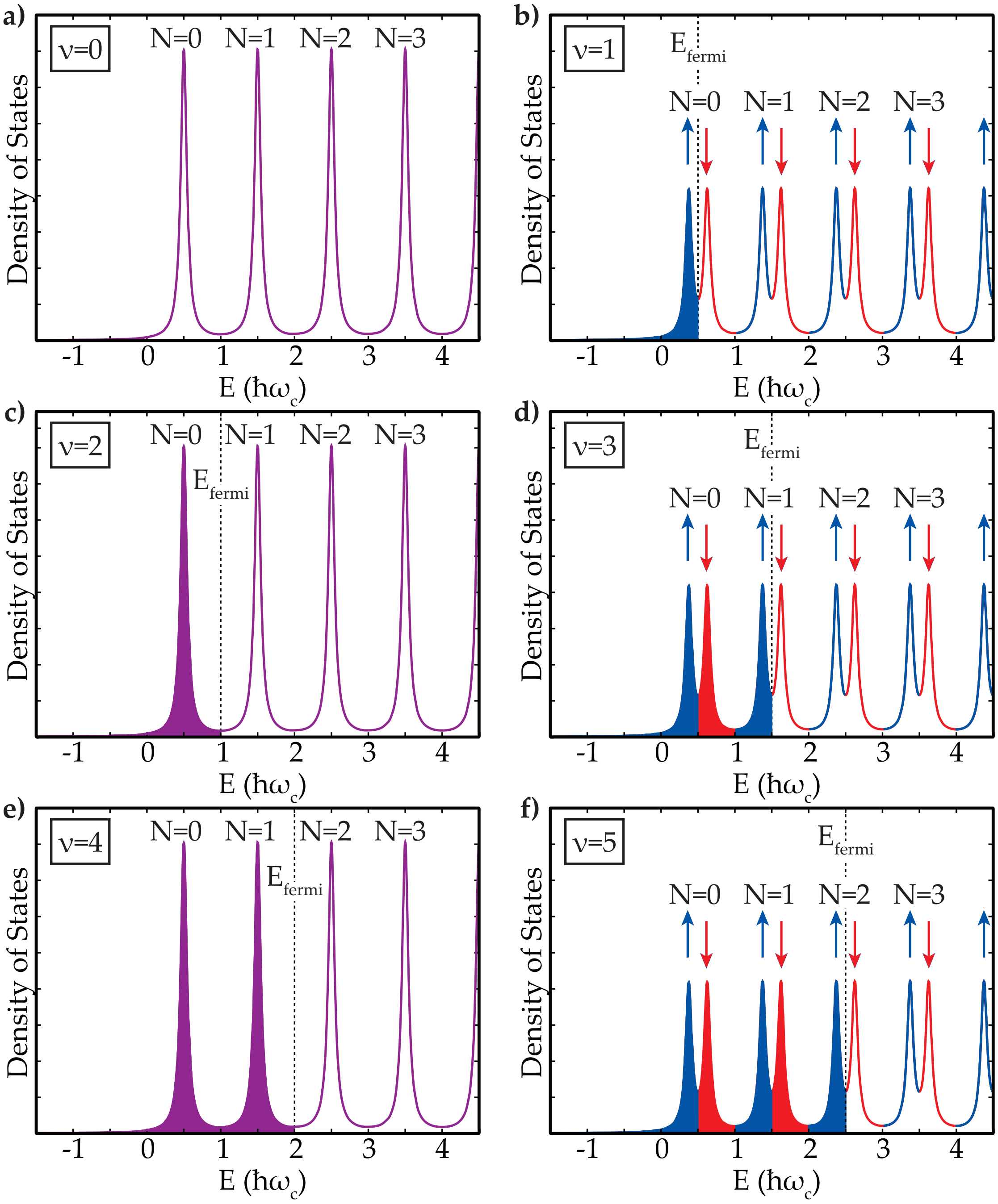


Figure 2-9: **Cartoons of the density of states of an interacting electron gas** at filling fractions $\nu = 0$ to 5.

degeneracy $2N_D$ because of spin and orbit center.

However, at $\nu = 1$ (Figure 2-9b), a more complicated picture emerges. We have put in enough charge to fill one spin-state of the $N = 0$ Landau level. These states are degenerate before we turn on electron-electron interactions, but even weak interactions can be expected to lift this degeneracy. The dominant effect is the formation of an exchange-enhanced spin gap. If two electrons in the quantum well have the same spin wavefunction, they must have an antisymmetric orbital wavefunction. This antisymmetry tends to keep the electrons apart, minimizing their Coulomb interaction. If the overall Coulomb energy of the electrons is removed, as by a uniform static positive background charge as considered in the jellium approximation, this results in a slightly attractive interaction between electrons of like spins. As we fill the quantum well, this attractive interaction causes the electrons to spin-align.[†] The Zeeman energy provides a slight symmetry breaking, and sets a preferred direction for the electron spins that we will label spin up. As we fill the spin up states, the electrons become closer together, increasing their net attraction, and the spin up states are pulled down in energy relative to the spin down. At $\nu = 1$, the spin up states of $N = 0$ are completely filled and lie below the Fermi energy, while the spin down states are completely empty and lie above the Fermi energy. The higher orbital Landau levels are similarly split; although their wavefunctions are orthogonal with those in $N = 0$, the exchange matrix elements do not vanish.

If we raise the density to $\nu = 2$ (Figure 2-9c), the spin up and spin down states become symmetric again, and the spin gaps vanish. Aside from a possible change in the quasiparticle effective mass and broadening due to finite lifetimes, we expect the spectrum to look similar to $\nu = 0$ again.

We can continue raising the density, filling successive orbital Landau levels, and see similar splitting and recombining of the orbital Landau levels. In each of the DOS spectra, it is possible to identify the electron density, and assign a unique Landau level index and spin to each of the peaks, without reference to

---

[†] This can also be understood as a special case of Hund's rule filling

any information outside of the spectrum. The spectra are almost self-calibrating, which will make them easier to understand in the more complicated scenarios we will encounter in real samples.

# Chapter 3

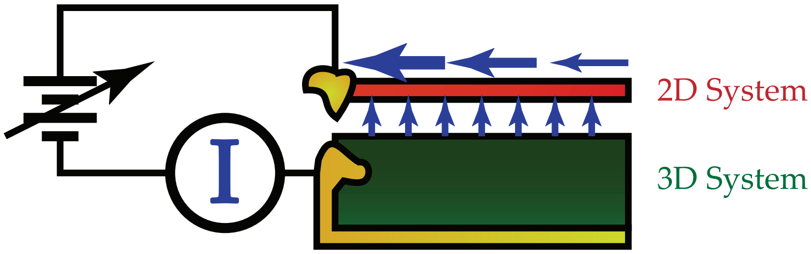
# Time Domain Capacitance Spectroscopy

Tunneling spectroscopy of two-dimensional systems poses unique difficulties. In order to probe the bulk states of the two-dimensional system, it is necessary to tunnel into the middle of the sample, not the edge where the potential barrier that forms the edge will modify the physics. Because of this, the tunneling current must be injected perpendicular to the plane of the 2D system.

In principle, the measurement could be quite simple. One can grow a 2D system in close proximity to a 3D one, with a thin tunneling barrier in between. After making separate contact to the 2D and 3D systems, one could connect a voltage source and a current meter, and simply proceed to take I-V curves or, with a bit more sophistication, use a lock-in amplifier to take dI/dV curves (Figure 3-1a).

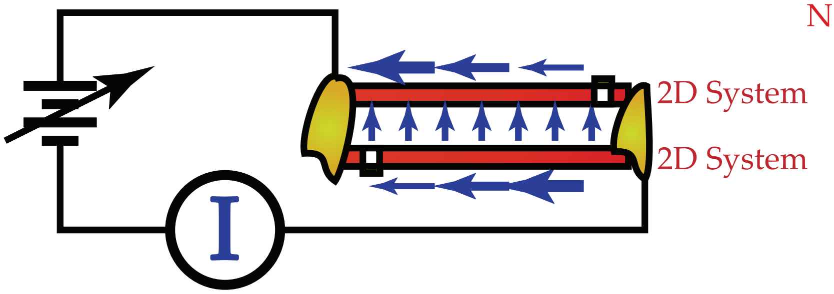
There are two basic problems with this approach. One is technological; in practice, the tunneling barrier is around 100 Å thick. Even the thinnest semiconductor ohmic contacts penetrate several thousand angstroms, and the contacts typically have poor repeatability on depth. It is virtually impossible to contact the 2D system without also shorting through to the 3D layer beneath it. This problem can be circumvented by clever sample design. For example, tunneling between two 2D systems, as opposed to between a 2D and a 3D system,

**a)** “Simple” Approach



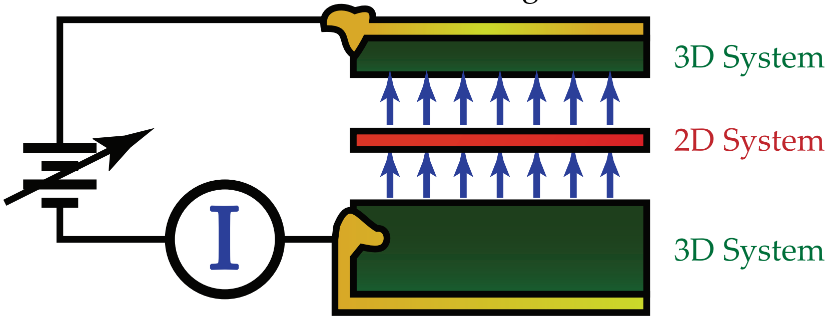
Difficult Fabrication
In-Plane Currents

**b)** Bilayer Approach



In-Plane Currents
No Simple Electrode

**c)** Double Barrier Resonant Tunneling



No In-Plane Currents
Density Control
Heating

Figure 3-1: Different possible designs for a measurement to determine the single particle density of states.

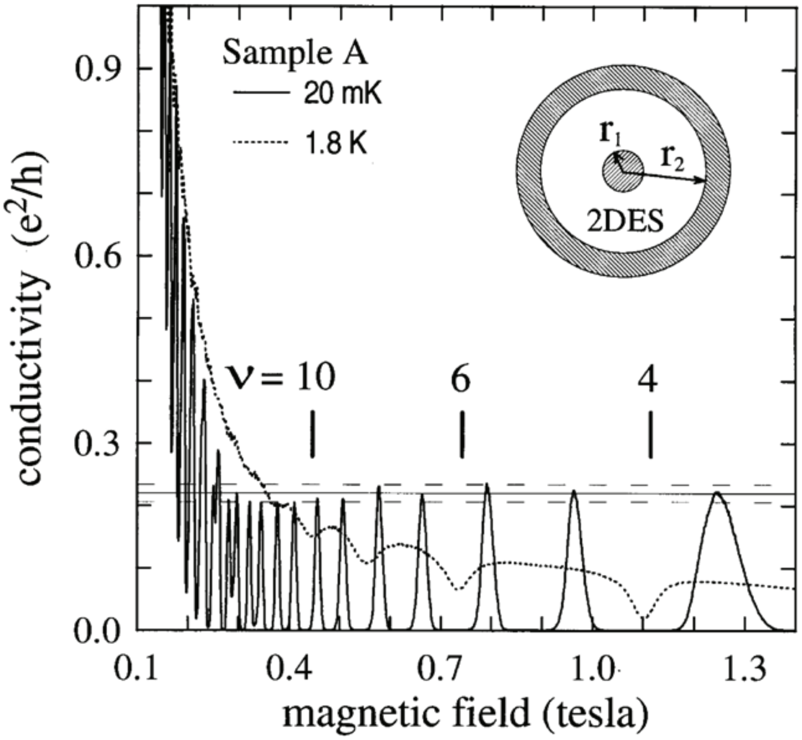


Figure 3-2: Solid trace: Conductivity of a two-dimensional electron gas at 20 mK (Rokhinson, Su, and Goldman, 1995)[24].

can be studied by using selective gating to break the short circuit (Figure 3-1b)[23]. However, in so doing one has replaced the "simple" 3D system with a complex 2D one. While the physics of tunneling between two correlated 2D systems in these quantum Hall bilayers is rich, it does not provide a simple probe of the single particle density of states.

The second problem is more fundamental; once charges have tunneled into the 2D system, they need to flow in the plane to a distant contact in order to exit. If the 2D system has any electrical resistance to in-plane current flow, this means a voltage will be dropped across its plane. Because of this voltage drop, electrons tunneling into the 2D system at different locations will tunnel in at different energies, blurring the energy scale of the measurement. This problem becomes more difficult as the injection energy grows, increasing the tunnel current.

Unfortunately, when studying the quantum Hall system, the in-plane conductivity of samples is always low and is frequently zero (Figure 3-2). This has limited approaches which use a local contact to the 2D to low injection energies and filling fractions, such as $\nu = {}^{1}/_{2}$, where the conductivity of the 2D is comparatively large.

Alternatively, one can construct a measurement where the current exits the 2D system perpendicular to the plane, on the opposite side from where it tunneled in; a broad class of measurements including double barrier resonant tunneling and resonant tunneling diodes (RTDs) fall into this category (Figure 3-1c). Because these devices are translationally symmetric in the plane of the 2D, they induce no in-plane currents. This allows them to continue to measure when the in-plane conductivity of the 2D system drops to zero and at high energies. These devices have been used to study 2DEGs in the past[25, 26]. However, varying the injected electron energy in these devices involves applying a voltage across the entire structure. This voltage in turn gates the 2D system, so it is impossible to vary the injected electron energy without also varying the 2D density. Furthermore, in these structures the tunneling barriers form a resistive voltage divider that controls the chemical potential of the 2D system. As the tunnel current fluctuates, so does the chemical potential of the 2D. The resulting density fluctuations in the 2D are difficult to measure or remove.

Ultimately, all of these techniques suffer from a final problem that limits their applicability to small injection energies: heating. Quasiparticle lifetimes at even small energies away from the Fermi surface are on the order of a few picoseconds, so a tunneled electron or hole essentially instantaneously transfers its energy to the 2D system as heat. At low temperatures, electron-phonon coupling vanishes, so dissipating this heat is a slow process.

## 3.1 Constructing a Tunnel Capacitor

In order to circumvent these difficulties we have introduced time domain capacitance spectroscopy (TDCS) [27, 28]. In TDCS, the 2D system to be studied is sandwiched inside of a tunnel capacitor (Figure 3-3). One plate of the capacitor is separated from the 2D system by a thin tunneling barrier; we will refer to this

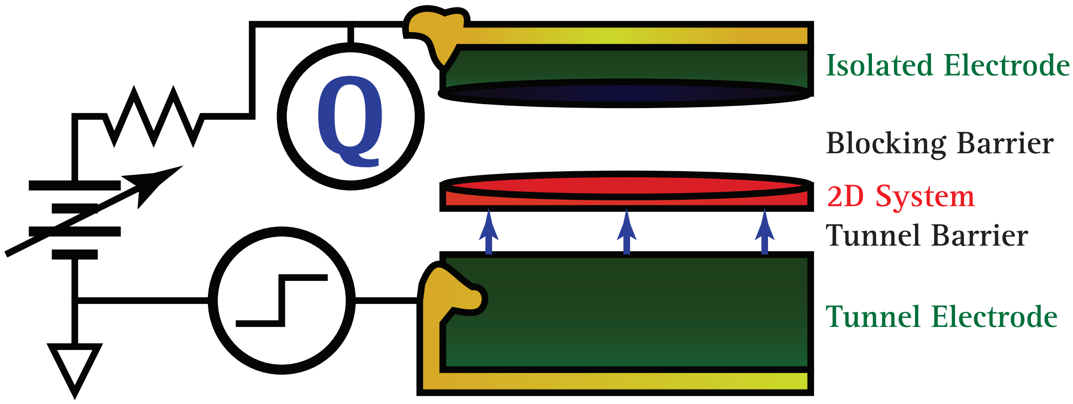


Figure 3-3: A simplified schematic of time domain capacitance spectroscopy.

plate as the tunnel electrode. The other plate is separated by a much thicker, insulating, blocking barrier. We will refer to this plate as the isolated electrode. The tunnel voltage is capacitively coupled to the 2DEG across the tunnel capacitor, while the tunnel current is measuring by monitoring its image charge on the isolated electrode. Because both the excitation voltage and charge detection are capacitively coupled, there is no need for a direct contact to the 2D system; this leads to very simple fabrication.

Like the RTD, this device has translational symmetry; we induce no in-plane currents in the 2DEG. TDCS continues to function, even when the 2DEG is fully depleted or insulating.

We believe the TDCS technique can be easily extended to many material systems, however, the results in this thesis are restricted to two-dimensional electron gases in GaAs. The maturity of the sample growth and preparation technology in this material system make it relatively easy to work with, while the outstanding mobilities available make it ideal for many-body physics.

Within this material system, we construct our tunnel capacitor using a heterostructure, with $\mathrm{Al}_x\mathrm{Ga}_{(1-x)}\mathrm{As}$ tunneling and blocking barriers (Figure 3-4). We use n+ doped 3D regions of GaAs as our isolated and tunnel electrodes. The band diagram in the inset is qualitatively accurate; we have no intentional donors in the blocking barrier. Such donors are a major source of scattering within quantum

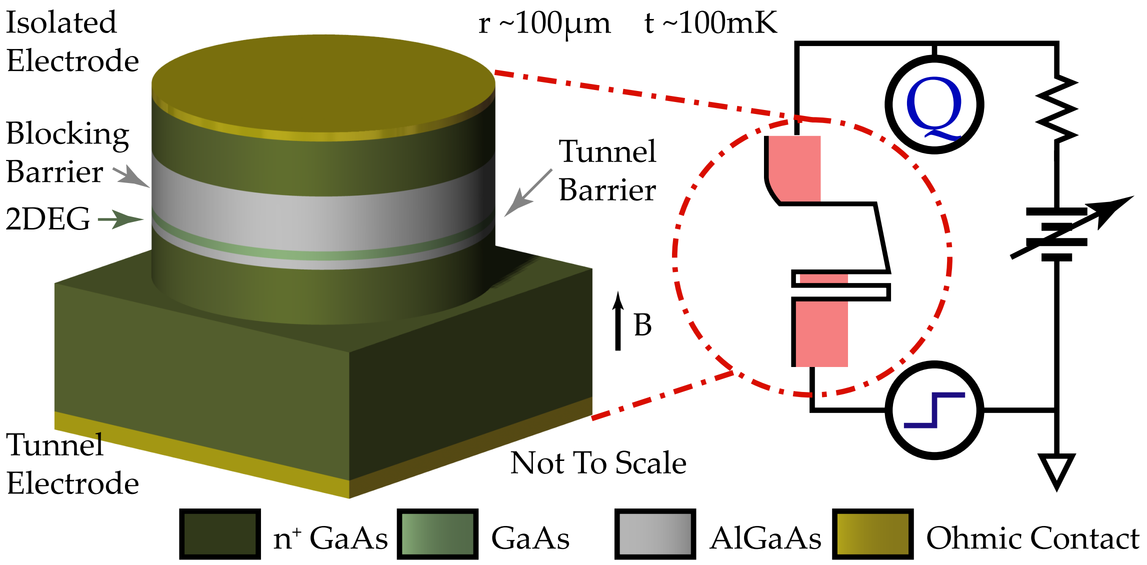


Figure 3-4: A simplified TDCS sample in GaAs.

wells; unlike the donors in the electrodes, there are no free electrons to screen their potential fluctuations[29, 30, 31]. In addition, our measurement is very slow, so we need to minimize any sources of drift within the sample. In $\delta$ doped samples, donors within the barrier often capture electrons out of the 2DEG and reionize on long time scales in response to applied voltages and temperature changes, as they tend to be close to the Fermi energy[32].

Electrical contact is made to the tunneling electrode by means of a planar ohmic contact to the back side of the substrate, while the isolated electrode is contacted from the front side of the sample. We then etch the sample (Section D.4) to isolate a section of 2DEG approximately 100 $\mu$m in radius. Our measurement averages across this entire 100 $\mu$m disc; it is intrinsically a bulk measurement. There are sparsely distributed defects in the substrate that short the tunnel and isolated electrodes together; without the isolation step, these would destroy our tunnel capacitors. Typically, we find around 90% of our devices are not shorted. This "mesa isolation" step also reduces the capacitance of our devices to around 6 pF; this is important because we need to be able to apply sharp voltage steps across the tunnel capacitor using a 50 Ω line, so we need to keep the RC time formed by the series combination of the sample capacitance and a 50 Ω source short.

## 3.2 Time Domain Capacitance Spectroscopy

Once we have our tunnel capacitor, we can begin to measure I-V characteristics for tunneling into and out of the 2DEG. We begin by applying a DC bias to set the 2DEG density by field effect gating; we can vary the density continuously from 0 to approximately $4 \times 10^{11} cm^{-2}$. Starting with the sample fully equilibrated, we apply a sudden voltage step to the tunnel electrode, disequilibrating it from the 2DEG (Figure 3-5a,b). At the first instant after the step is applied, the tunnel and blocking barriers form a capacitive voltage divider. The voltage between the 2DEG and the tunnel electrode, $\delta$V, is just a geometric lever arm times the total step height (Figure 3-6). We monitor the charge as a function of time on the isolated electrode. At the instant the voltage step is applied, a large current pulse from the geometric capacitance of the tunnel capacitor (the capacitance the structure would have in the absence of the 2DEG) results in a step in the charge on our charge sensor. This is followed by a long, slow rise from the image charge of electrons tunneling from the tunnel electrode onto the 2DEG. The charge step from the geometric capacitance is typically five times larger than the relaxation signal (this ratio is set by the geometric lever arm). In order to avoid saturating our amplifier, we subtract it by applying a pulse of the opposite sign to a standard capacitor $C_{std}$, also connected to the charge amplifier.

Because of the superposition principle, we know the amount of image charge Q must be proportional to the total amount of charge that has tunneled, so $d\mathrm{Q}/dt$ will be directly proportional to the tunnel current as a function of time between the tunnel electrode and the 2DEG.[†] At the very first instant after the step is applied, we know the voltage on the 2DEG, $\delta$V, and can measure $\mathrm{I} = d\mathrm{Q}/dt$;

---

[†] Because our device is two terminal and charge-neutral, we could in fact apply the pulse and measure the charge on whichever electrode is more convenient with only a sign change in the pulse amplitude. Typically, we will measure the charge on whichever electrode has been etched in the mesa isolation step to minimize shunt capacitance on our charge sensor.

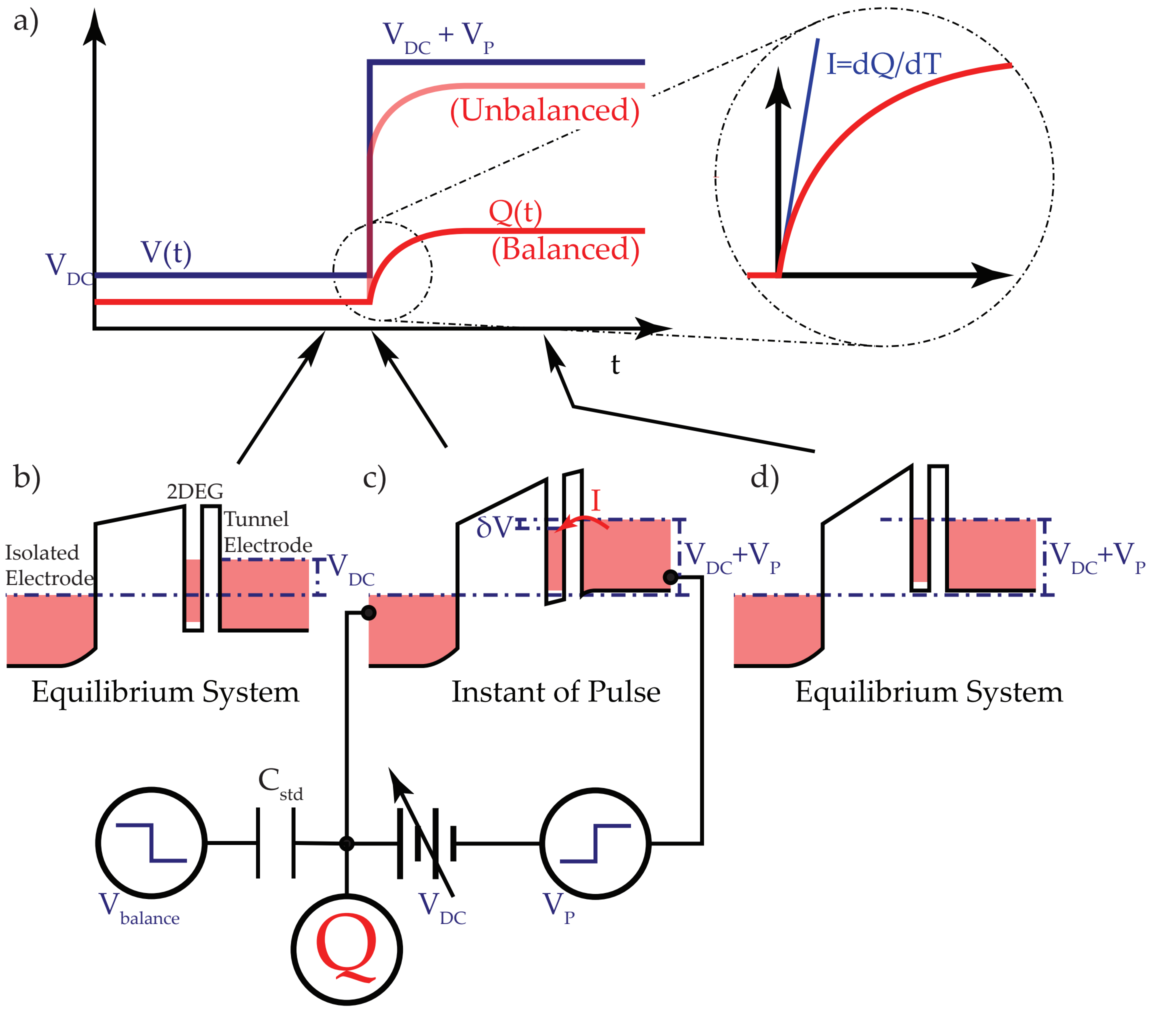


Figure 3-5: **Applying a single step**: A single step applied across a TDCS sample. **a)** Charge and voltage as a function of time for a single step applied across the sample. **b)** The sample starts with the 2DEG and tunnel electrode fully equilibrated. **c)** We suddenly apply a voltage step (blue line in **a**) to the tunnel electrode, dis-equilibrating it and the 2DEG. We monitor the current flow by measuring the image charge on the isolated electrode (red line in **a**). **d)** After a long period of time has elapsed, the 2DEG reequilibrates at a new density.

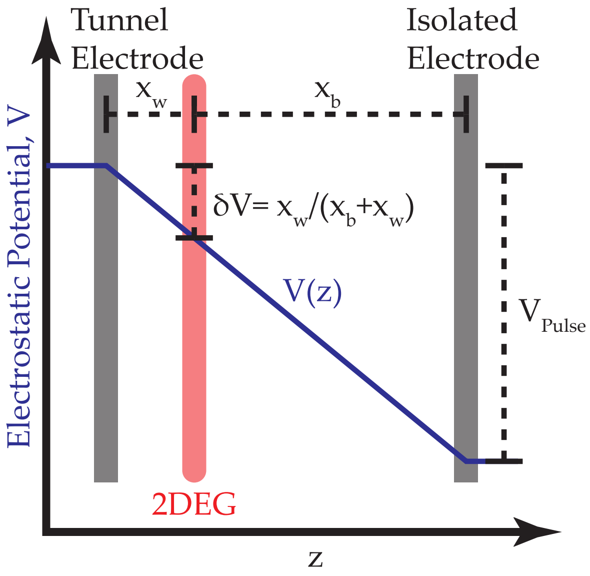


Figure 3-6: **The geometric lever arm**: At short times, the voltage between the 2DEG and the tunnel electrode is given by the applied step height times a geometric lever arm.

together, these give us a single point on an I-V curve. A very long time after the voltage step has been applied, the sample would reequilibrate at a new density; because we only need $dQ/dt$, we typically truncate the voltage step into a brief pulse just long enough to perform this measurement.

We can repeat this measurement while varying the pulse height to trace out a single I-V curve. Because we can vary both the sign and the magnitude of our voltage pulse, we can examine both the ejection and injection of electrons, measuring the DOS for both electrons and holes in the 2DEG.

This leads us to our procedure for measuring I-V characteristics using TDCS (Figure 3-7). We apply a single pulse across the sample while measuring the charge on the isolated electrode (Figure 3-7a). Because the bandwidth of our measurement is large and the charges involved are small, the thermal Johnson noise of our room temperature amplifiers dominates this measurement. To overcome this, we average together the response from on the order of 250,000 pulses, giving a clean measurement of charge as a function of time (Figure 3-7b). We fit the slope of this curve in the first instant (typically within the first 100 ns)

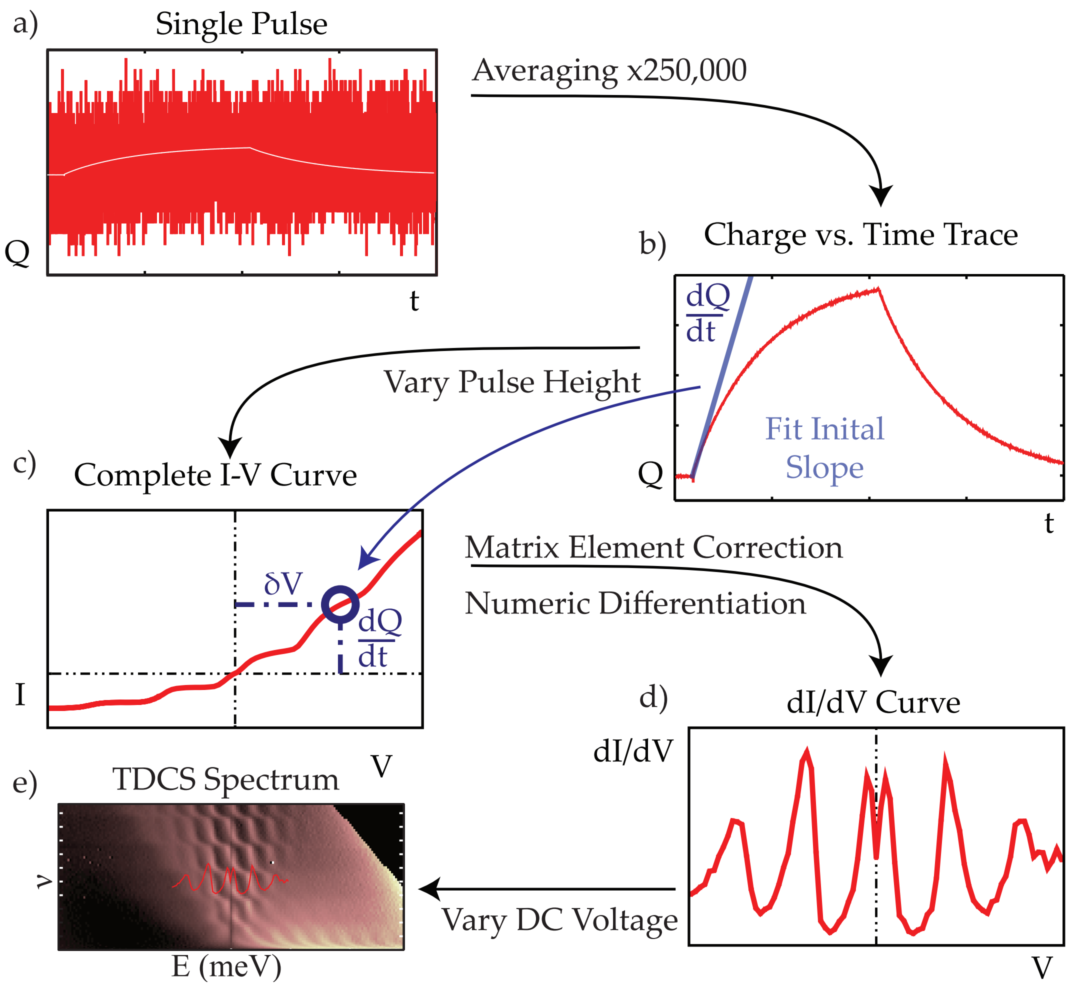


Figure 3-7: **Measuring an I-V characteristic with TDCS**: **a)** A single pulse is applied across the sample. Because of the small magnitude of the signal, we only measure noise. **b)** The response from ∼250,000 pulses is averaged together, giving a charge versus time trace. The initial slope is fit to extract the tunneling current, which, together with the known pulse height, gives us a single point on an I-V characteristic. **c)** The pulse sign and magnitude are varied to give a complete I-V characteristic. **d)** The I-V curve is corrected for matrix element variations (see Section 3.4) and numerically differentiated, giving the single particle density of states. **e)** The DC bias across the tunnel capacitor is varied to change the 2D density and the process is repeated to build up a color scale plot.

after the pulse is applied, which together with the height of the voltage pulse $V_p$ times the geometric lever arm $\lambda$ gives us a single point on an I-V curve. Because the $V_p\lambda$ gives us the voltage difference between the Fermi energy of the 2DEG and the tunneling electrode, this technique will automatically reference all energies to the Fermi energy; a zero-height pulse corresponds to tunneling at the Fermi energy. We then change the pulse height slightly and repeat the procedure, averaging together another 250,000 pulses. This gives us a second point on the I-V curve. Repeating this with several hundred pulse heights allows us to build up a complete I-V curve (Figure 3-7c). We then apply an approximate correction for the variation of the tunneling matrix element (see Chapter 10) and numerically differentiate this curve, giving the single particle density of states (Figure 3-7d).

In the absence of any context, it is difficult to interpret this single DOS curve. We provide context by varying the DC bias across the sample, which varies the carrier density in the quantum well, and repeating the entire process. Combining several hundred such traces, we can construct a color-scale plot of the DOS, with electron density as one axis, energy from the Fermi energy as the second, and intensity giving the single particle density of states (Figure 3-7e). We use bright colors to represent high DOS regions and dark colors to represent low DOS regions. Typically we will present the spectra rotated 90° from their presentation in Figure 3-7e in order to put high energy states at the top of the figure and low energy states at the bottom.

## 3.3 Getting Oriented on a TDCS Spectrum

One of the more striking features of these spectra is that once we have applied a magnetic field, there are enough distinctive features to roughly determine both the energy and density scale without any additional information. In Figure 3-8, the vertical ($y$) axis is energy; $E = 0$ at the center corresponds to tunneling at the Fermi energy. Moving upwards corresponds to injecting electrons at increasingly large energies, while moving downwards corresponds to ejecting electrons from

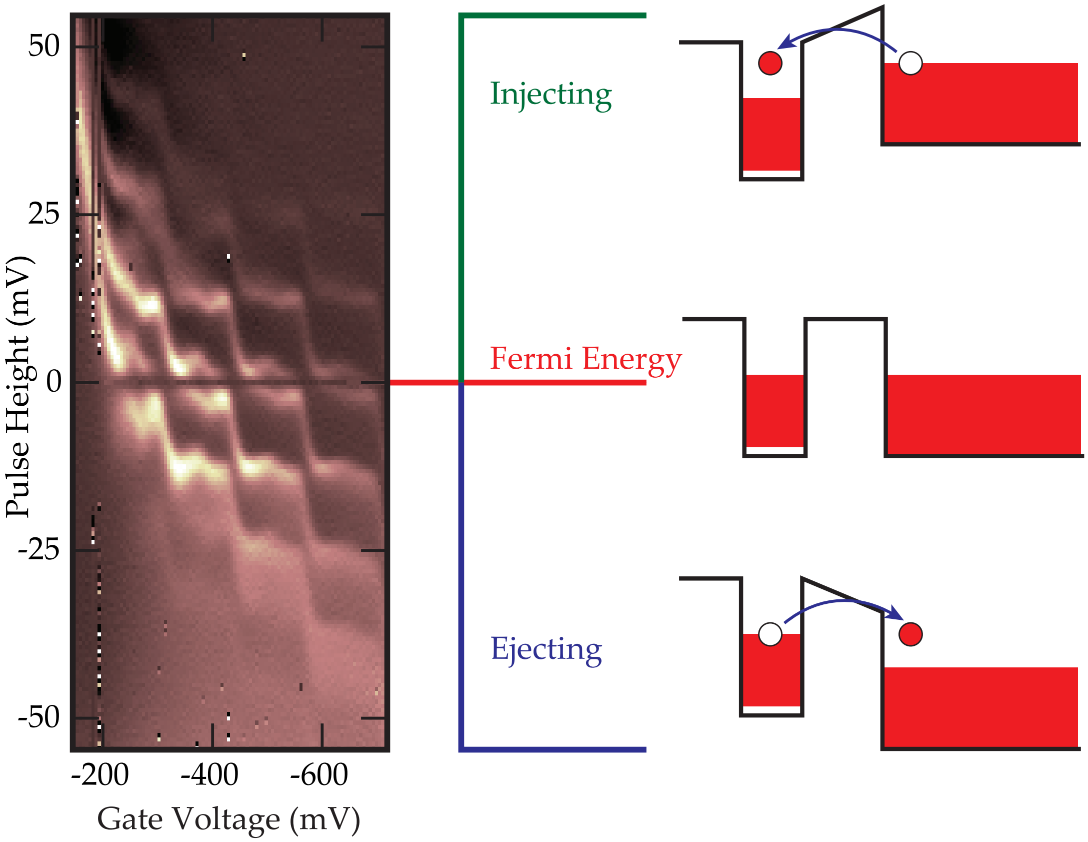


Figure 3-8: **The energy axis**: Interpreting the energy axis; at zero pulse height, we are examining equilibrium tunneling. Increasingly positive pulses inject electrons into the 2DEG, while increasingly negative pulses eject electrons out of the 2DEG

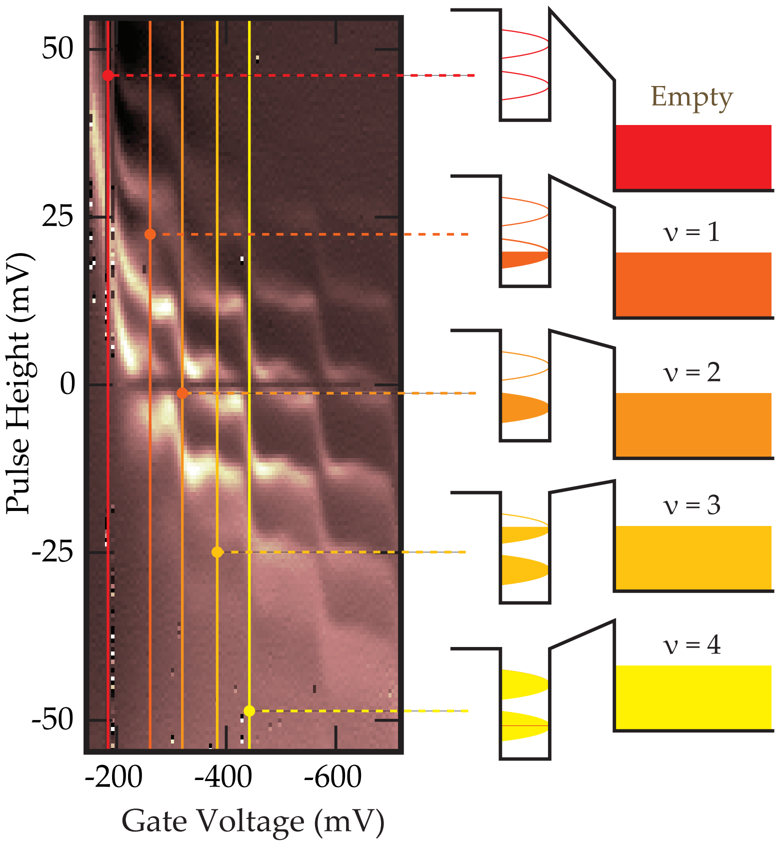


Figure 3-9: **The density axis**: Interpreting the density axis; at the left edge of the plot, the quantum well is fully depleted. As we move right, we begin to lower the lowest bound state energy in the quantum well with the applied DC field, and pull electrons into the quantum well. Each bright stripe corresponds to a single orbital Landau level, so we can count up the density by counting the number of Landau levels that have crossed the Fermi surface.

increasingly deep in the quantum well. Because the injection or ejection energy is linearly proportional to the pulse height with a fixed scale factor, we can measure relative energies on this plot by simply comparing vertical distances. We always pulse away from equilibrium, measuring all energies relative to the Fermi energy.

The horizontal ($x$) axis is the applied DC voltage, with the far left of the figure corresponding to depleting voltages large enough to completely empty the well, and the far right of the plot corresponding to a large enhancing voltage. The structure as a whole acts like a field effect transistor, and the electron density in the well is roughly proportional to the applied DC voltage.

Each bright stripe in the figure is a single orbital Landau level. As we apply a negative voltage to the blocking electrode, we push more electrons into the quantum well, and pull the Landau levels downward towards the Fermi energy. Each time a Landau level reaches the Fermi energy, we must fill it completely before the chemical potential changes appreciably; this results in a flat plateau in the plot. Once the Landau level is filled completely, the next Landau level must be pulled down to the Fermi energy before we can put the next electron into the well, resulting in a sharp downward jump in the spectra. This alternation between pinning levels at the Fermi surface and the chemical potential jumps results in the staircase appearance of the spectrum[33, 34]. As shown in Figure 3-9, we can determine the density in the 2DEG simply by counting the number of Landau levels that have crossed the Fermi surface.

## 3.4 Adding Value: Calibrating the Axes

One of the chief advantages of TDCS is that we can make the calibration on all three axes of the figure quantitative. The simplest axis to calibrate absolutely is the energy axis. While in principle we could determine the geometric lever arm from the as-designed device geometry, the spectra provide a convenient way to calibrate it experimentally. When the well is completely depleted, we are exploring the spectrum of a non-interacting electron gas; there are no electrons for the injected electrons to interact with. Because of this, the spacing between the orbital Landau levels is just given by $\hbar\omega_c$. We use the measured spacing between the Landau levels to calibrate the geometric lever arm, and find $\lambda = 0.267$ for PF-6-19-98-1, the wafer used for the bulk of the measurements in this thesis. This is in excellent agreement with the calculated lever arm for this structure. The result of this energy calibration is shown in Figure 3-10a (compare to Figure 3-8). As we apply biases to the structure, the electrodes deplete somewhat and the $z$-wavefunction of the 2DEG in the quantum well gets pulled from side to side, resulting in some small variation of the geometric lever arm. We estimate the resulting energy error to be less than 1%; see Chapter 12 for details.

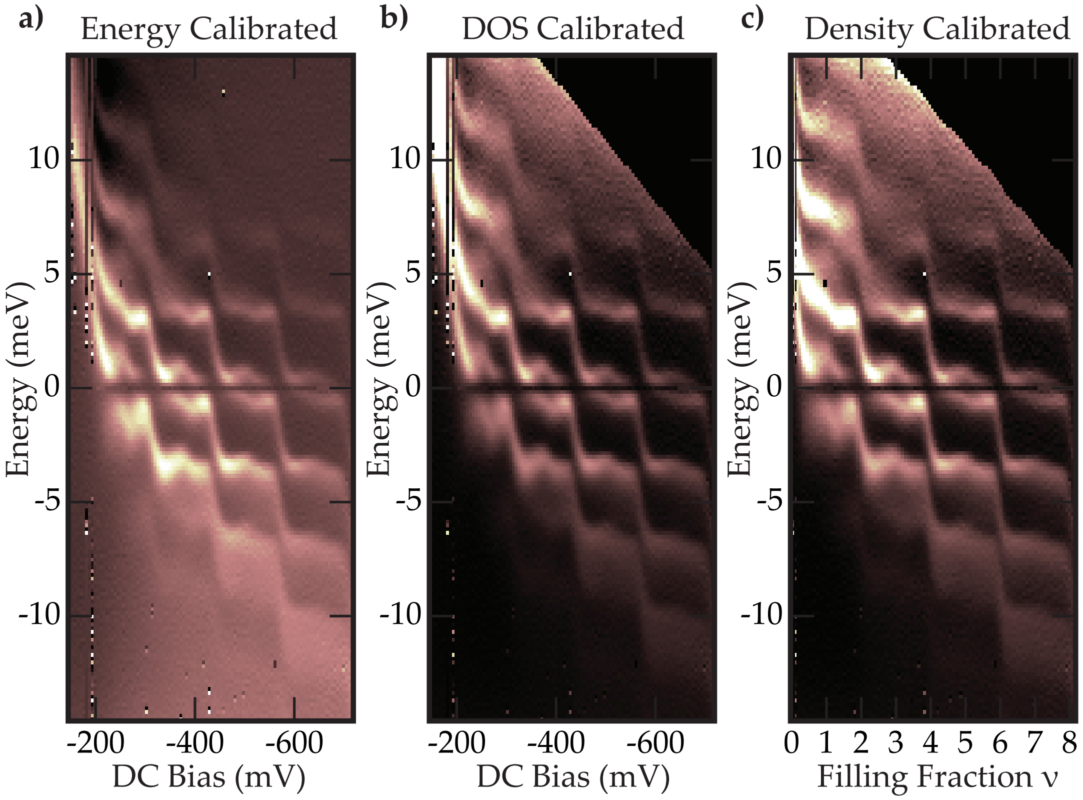


Figure 3-10: **Calibrating the axes**: **a)** TDCS Spectrum taken at 2 Tesla, with the (trivially) calibrated energy scale on the *y* axis. **b)** The same TDCS spectrum from **a**, with the tunneling matrix element correction applied to the color scale. **c)** The same TDCS spectrum as **a** and **b**, with the *x* axis distorted to show density rather than gate voltage, along with the matrix element correction.

Examining Figure 3-10a carefully, we notice that while the total number of states in each Landau level is identical, The Landau levels near the Fermi surface and slightly below it are quite bright, while those near the top-right corner of each plot are substantially dimmer. This reveals the error in one of our assumptions in deriving Equation 2.5; the tunneling matrix element and 3D density of states are not actually fixed.

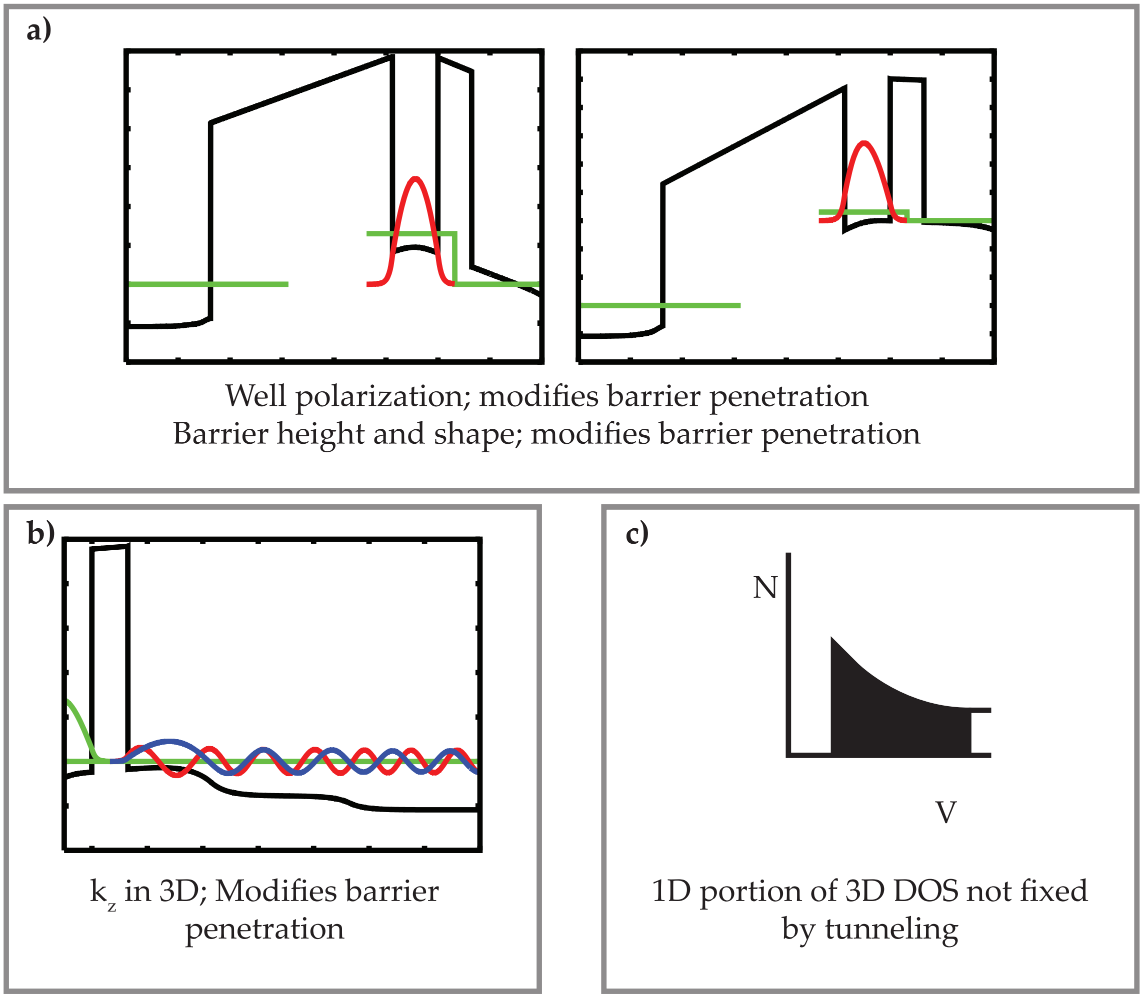
There are a number of sources for the variation of each of these; the same polarization of the quantum well that changed our geometric lever arm pulls the $z$-wavefunction either into or away from the tunneling barrier, modifying the tunneling matrix element. This electric field also changes the effective barrier height and shape (Figure 3-11a). The matrix element also varies because electrons in the 3D electrode with short wavelengths in the $z$ direction penetrate deeper into the tunneling barrier. Finally, unlike the 2D density of states at zero field, the 3D density of states is not flat.[†] We develop a detailed analytic model of this matrix element and density dependence in Chapter 10, and use it to convert our dI/dV spectra to single particle density of states (see Figure 3-10b). In the top-right corner of each plot, we are attempting to tunnel electrons into the 2DEG from below the 3D band edge, so the 3D density of states vanishes and our matrix element correction diverges. We substitute a black band in this region. This cutoff limits the energy range of our measurements to the 3D Fermi energy.[‡]

This correction is *approximate*, not exact. However, because it is smoothly varying, it neither changes the position nor width of features in our spectra; we use it to normalize contrast across the image and thus simplify fitting and visual inspection. Therefore, this “matrix element correction” should be thought of as an aesthetic decision with a physical motivation, and heights and spectral weights should not be quantitatively compared when they are separated by large distances in the spectra.

---

[†] Actually, $k_x$ and $k_y$ are fixed by momentum conservation, so the density of states that enters into Fermi’s golden rule is a 1D density of states.

[‡] We can eject electrons at arbitrarily high energies, so, more precisely, our injection range is limited by the 3D Fermi energy, and our ejection range by the 2D Fermi energy. The second sub-band of the quantum well provides a way to sidestep this restriction; see Section 8.5



Well polarization; modifies barrier penetration
Barrier height and shape; modifies barrier penetration

$k_z$ in 3D; Modifies barrier penetration

1D portion of 3D DOS not fixed by tunneling

Figure 3-11: **Tunneling matrix element corrections**: Origins of the corrections to the tunneling matrix element. **a)** The applied electric field polarizes the quantum well, pulling the bound state closer to and farther from the tunneling barrier. This modifies the barrier penetration. In addition, the shape and height of the tunneling barrier vary with the applied field. **b)** Electrons leaving the 3D with higher momentum towards the barrier penetrate more easily than those with lower momentum towards the barrier. **c)** The density of states in the 3D is not flat, so there are differing numbers of electrons available to tunnel at differing energies. Transverse momentum is conserved, so the remaining degree of freedom, $k_z$, gives a 1D density of states.

Finally, because our structure is sandwiched inside of a capacitor, it is possible to keep track of the amount of charge entering and leaving the quantum well. This allows us to distort the $x$ axis of our plot to transform it from applied DC bias to charge density in the quantum well. This is shown in Figure 3-10c. The DC bias is almost directly proportional to the density. However, near the chemical potential jumps between Landau levels and the region where the quantum well depletes a finite range of DC bias is used to pull down the quantum well bound state energy relative to the 3D electrode rather than to charge the quantum well. In these regions, we increase the DC bias without also increasing the density. Thus, when we transform to density rather than DC bias on the $x$ axis, we find somewhat steeper steps between Landau levels. This calibration method and the errors intrinsic to it are discussed in Chapter 11.

## 3.5 Signal Generation and Detection Concerns

Although in principle TDCS is a very simple technique, there are a number of experimental difficulties in practice. The voltage at which we are tunneling (the voltage across the sample) is set by the height of our pulse in the first moments after the pulse is applied. This means that to keep the voltage noise small, we need to have spectacularly flat pulses. This problem is compounded when we differentiate our I-V curve to get dI/dV; errors in V get moved up to the numerator during this step. Commercial pulse generators are available that provide pulses with very sharp transitions, but typically not very flat tops. Furthermore, they generally only provide coarsely tunable ( 1%) pulse amplitudes. We circumvent these problems through the use of a piece of custom hardware that flattens the top of the pulses and modulates their amplitude; this is discussed in Appendix A.

Once a sharp and flat pulse is generated, it needs to reach the sample itself. If we want to measure at energies up to 15 meV (our current typical range) with temperature limited resolution (about 10 $\mu$eV), we need our pulses to be flat to

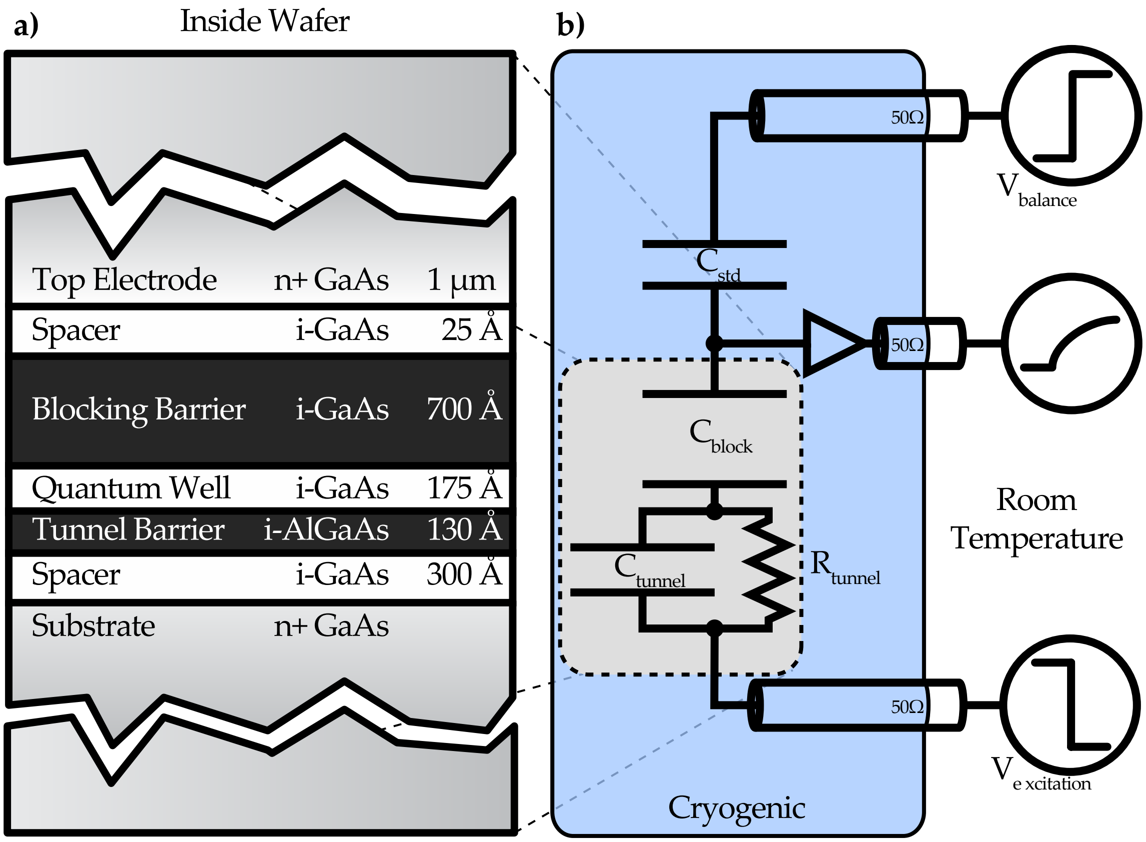


Figure 3-12: **Cryogenic cabling**: A schematic of the experimental setup showing the locations of the amplifiers and 50 Ω cabling

about 1 part in 1000. This means a reflection of -60 dB in one of our cables is enough to limit our resolution. That is, it will directly change the value of the excitation by enough to swap the order of two successive points on the measured I-V curve. In practice, we need to fit a rise of less than $^1/_{100}{}^{th}$ of the step height immediately after the pulse, so we are sensitive to ringing at less than -80 dB, although under many circumstances if the ringing is fixed we can ignore it at this level. This means the utmost care must be taken at every point in the experimental setup to limit and to attenuate reflections. This includes the presence of 50 Ω cables leading directly to the sample in the mixing chamber of our dilution refrigerator (Figure 3-12). While this substantially increases the base temperature of our fridge, it is necessary for the experiment. Some details of the measures taken can be found in Appendix A.

Finally, in order to measure the slope of our charging pulse in the first instants after it is applied, we need an amplifier able to drive this miniscule signal up to room temperature. To this end, we have designed a HEMT based amplifier with a 50 Ω output, a voltage gain of almost unity, and 10 $\mu$W of power dissipation. Although the voltage gain is non-existent, the power gain is substantial; it matches a 10 pF source (typically a few kΩ to a few MΩ at the frequencies of interest) to a 50 Ω line. The design of this amplifier is detailed in Section A.4.

## 3.6 Balancing

As introduced above, we balance the large geometric capacitance of the sample in order to remove an inconveniently large charge step immediately after applying the pulse. There is also a more pressing need for balancing the charge. The "balance point" at the input to our amplifier is a high impedance point ($R > 100$ MΩ, $C \sim 6$ pF). When the geometric capacitance of our sample dumps a large chunk of charge onto the balance point, the voltage there jumps. The series combination of the sample and the shunt capacitance of the balance point forms, at short times, a capacitive voltage divider, reducing the voltage across the sample. Applying the balancing pulse keeps the balance point of the bridge a virtual ground. This not only holds the amplifier at a fixed bias point, minimizing gain drifts, but also guarantees the full amplitude of the applied pulse is applied across the tunnel capacitor. Because the quality of the balance feeds back into the pulse height across the tunnel capacitor, errors in the balance result in noise on the measured dI/dV trace. We balance by fitting a straight line through the trace a few tens of nanoseconds after the pulse, and adjusting the pulse height on the standard capacitor until the line passes through zero at the moment the pulse is applied; this allows us to reject any transients at the moment the pulse is applied. As a rule of thumb, we try to guarantee that the error in this "short-time" balance corresponds to a voltage change of less than 1% of the voltage resolution of the dI/dV trace we are taking; at that level, it contributes a DOS noise of less than

1%.

## 3.7 The Coulomb Gap

At large magnetic fields, we find there is a gap in the density of states at the Fermi energy. This gap is known as the Coulomb gap, and has been discussed extensively elsewhere[35, 36, 27, 37, 38, 28, 39]. It is closely related to the linear Coulomb gap in disordered, localized 2D systems [40]. As suggested by the fact that it is pinned to the Fermi energy, the Coloumb gap reflects a many-body effect. When a tunneling electron enters the 2DEG, the surrounding electrons need to move away for the system to reach its new ground state. In a strong magnetic field, they are curved back towards the tunneled electron by the Lorentz force, preventing them from doing so (Figure 3-13a). This gives rise to a gap in the density of states when tunneling in at low energies (Figure 3-13b,c). The gap manifests itself in our spectra as a wide, dark band centered about $E = 0$. The gap becomes broader and deeper as the magnetic field is increased.

## 3.8 Measuring the Initial Slope

We have said that we need to measure the "initial slope" of the charge versus time trace. Clearly, we need to fit the signal over some range of times to measure the slope; we will refer to this as our "measurement window." There is a tradeoff inherent in the amount of time we include in this measurement window. On one hand, if we include more time in the window, we will be able to fit the slope more accurately. This will allow us to make the same measurement with less averaging. On the other hand, if we include too much time in our measurement window, the chemical potential difference between the tunnel electrode and the 2DEG will relax noticeably, and we will blur our spectrum.

A practical criterion for how much blurring is too much is that the energy relaxation that occurs during the measurement window should be less than the

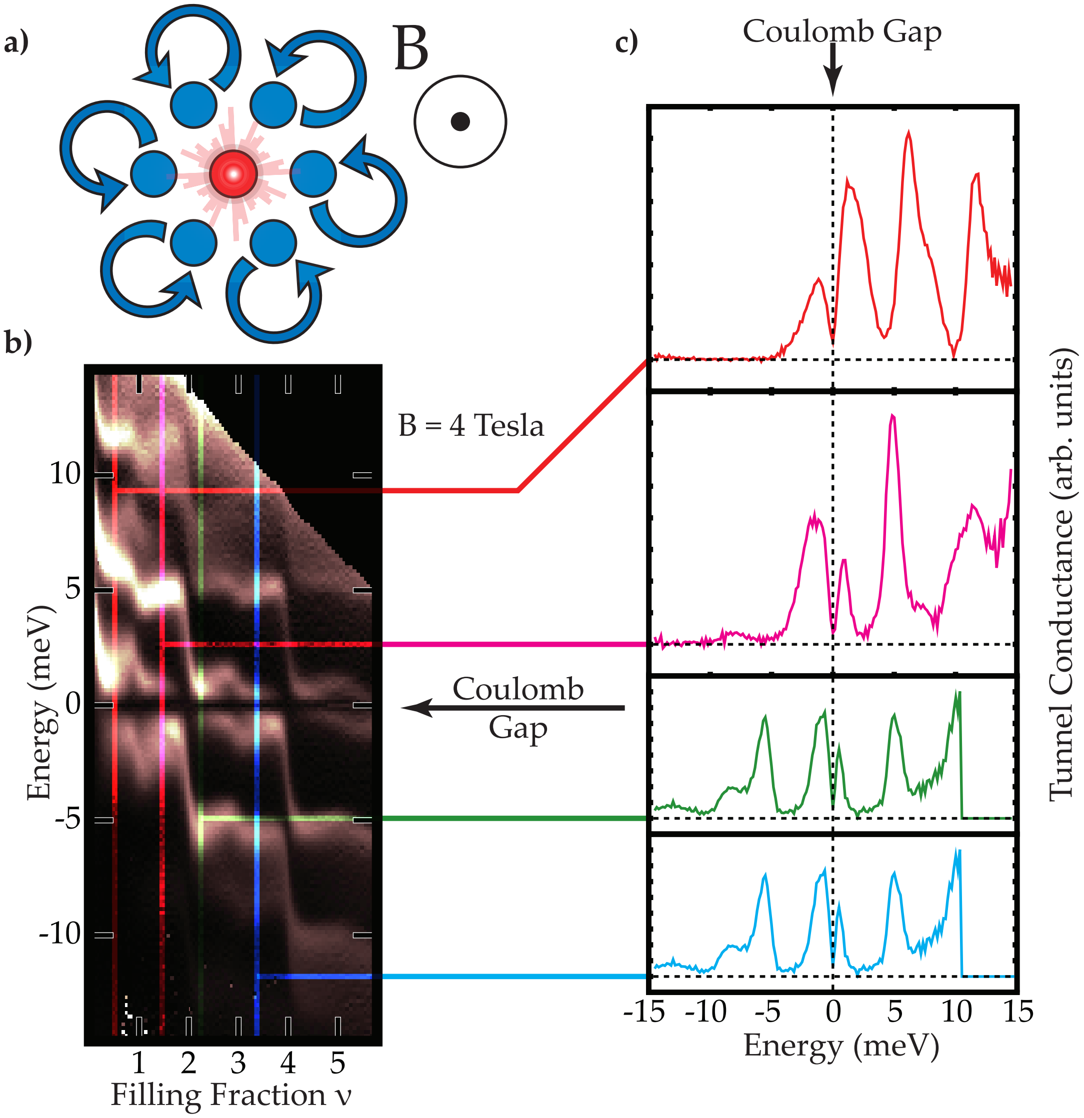


Figure 3-13: **The linear Coulomb gap a)** A tunneled electron (red) tries to repel the surrounding electrons to reduce its Coulomb energy, but in a magnetic field the Lorentz force causes its neighbors to spiral back in. **b)** A TDCS spectra taken at 4 Tesla magnetic field, showing the Coulomb gap. **c)** Several line cuts taken through **b**, all with the same vertical and horizontal scales, showing the Coulomb gap at the origin.

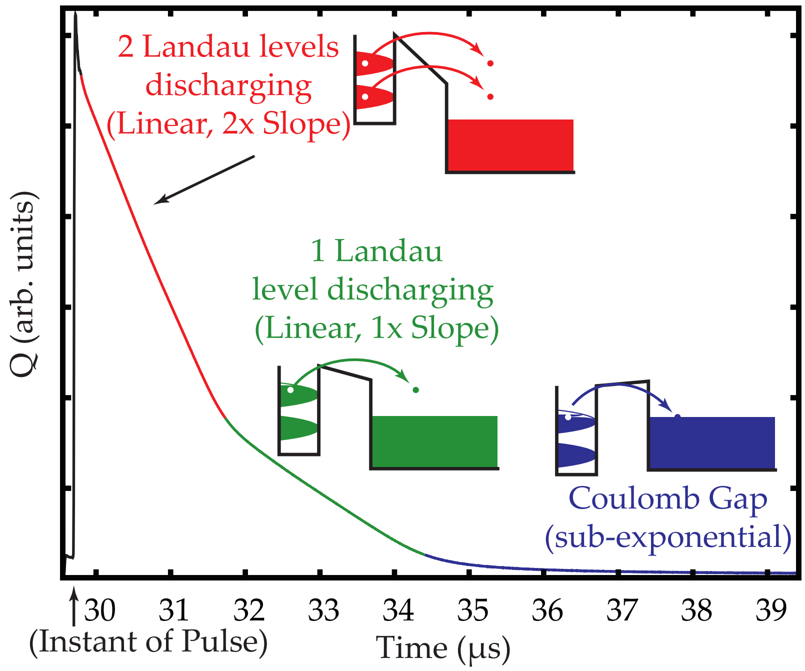


Figure 3-14: **Non-linear decay at 2 Tesla**: Charge response to a 34 mV pulse at 2 Tesla applied field, taken on a sample with a thin (fast) tunnel barrier. As the 2DEG and the tunnel electrode equilibrate, fewer Landau levels are available to tunnel into, resulting in a distinctive piecewise-linear relaxation curve.

energy step between adjacent measurements. For a temperature limited spectrum taken out to an energy of 15 meV, this means that only the first part in one thousand of the re-equilibration may be allowed to occur during the measurement window. That is, the energy relaxation must be limited to less than 1 part in 1000 of the total change that would occur if the structure were allowed to relax completely. For example, at zero field, the DOS is nearly constant with energy, so the relaxation is exponential in time. If the time constant of the relaxation is around a millisecond, we need to fit the initial slope in the first microsecond. As we raise the magnetic field, the relaxation time constant for small pulses will grow because of the Coulomb gap. However, the DOS is no longer constant, and the relaxation is no longer exponential. In fact, the relaxation becomes essentially piecewise linear at large excitations (Figure 3-14), with the slope proportional to the number of Landau levels able to discharge. As we turn up the magnetic field and increase the degeneracy of the Landau levels, the peaks in the DOS become large leading to larger currents than those at zero field for some excitation voltages. Because of this, we need to use a very short measurement window despite the long time constant for equilibration at small excitations.

Our measurement window is typically 50-300 nanoseconds long, depending on the thickness of the tunnel barrier of the sample. There is some tradeoff between the initial rise time of the pulse and the quality of it at long times (see Appendix A), so we only use the shortest windows when absolutely necessary. Our sample has a relaxation rate at zero field of a few milliseconds, so this is a very conservative range. Immediately after the pulse is applied, it takes a few tens of nanoseconds for the ringing from reflections and the signal generator to die down, so we actually fit a time period starting 20-75 nanoseconds after the pulse is applied, and ending 50-300 nanoseconds after the pulse is applied.

As can be seen in Figure 3-14, in only using the initial slope of the Q-t trace, we throw out an enormous amount of information; in principle, the entire decay tells us about the tunneling density of states for varying amount of disequilibrium between the 2DEG and the tunnel barrier. One could thus extract a complete tunneling I-V curve by looking at the detailed response to a single pulse (averaged appropriately). There are several reasons why we do not actually do this. First, once the 2DEG and the tunnel electrode have begun to equilibrate, calculating the potential difference between them requires a detailed and exact knowledge of the geometry of the sample. As we discuss in Chapters 9-11, the positions of the 2DEG and the electrodes are dynamic quantities that change in response to the AC and DC signals we apply to the structure, making it virtually impossible to know the potential difference between the 2DEG and the electrode with any precision after a significant amount of relaxation has occurred. Second, as we inject electrons into the 2DEG at energies far from the Fermi surface, we heat it up. Immediately after the pulse is applied, we are tunneling into a cold electronic system. We then inject highly excited electrons which give up their energy to the electrons already present in the 2DEG in a few picoseconds (see Chapter 7), heating it. By limiting ourselves to the first few electrons to tunnel in, we minimize the amount of heat delivered to the 2DEG during our measurement. We are able to see effects of the low-energy ground state of the system even in the high-energy DOS.

We can reassure ourselves that our measurement interval is adequately short by doubling it and ensuring the spectra do not change. This can be done without repeating the measurement since we acquire far more of the charge vs. time trace than we actually use. If the measurement interval is too long, features at high energies tend to move farther from the Fermi energy, as some relaxation of the applied pulse is occurring during the measurement. In addition, sharp features blur, while regions of low conductivity stretch, emphasizing dark regions in the spectra.

## 3.9 Discharge Pulses

After a single pulse in which we inject (or eject) some charge, we need to wait for the sample to re-equilibrate (Figure 3-15). Charge returning to the 2DEG after we remove our pulse sees roughly the same spectra as the charge exiting does[†], so any gap near the Fermi energy will tend to slow re-equilibration of the 2DEG and the 3D electrode. Because of the linear Coulomb gap, as we raise the magnetic field, this relaxation takes an increasingly large amount of time. At even moderate magnetic fields ($\gtrsim$ 2 Tesla), the time it takes for the sample to relax dominates the acquisition time.

We can help matters somewhat by, instead of applying a long step as shown in Figures 3-14 and 3-15, removing the AC voltage as soon as our measurement window has elapsed. Assuming the pulse is already short enough that the DOS is relatively flat across the amount of relaxation that occurs, the amount of charge that tunnels into the 2DEG is linearly proportional to the pulse length. When we wait for the sample to relax, assuming the amount of charge we allowed to tunnel in was small, the relaxation will be exponential with a time constant set

---

[†] The 2DEG has been heated slightly by the injected charge, which may modify the spectra.

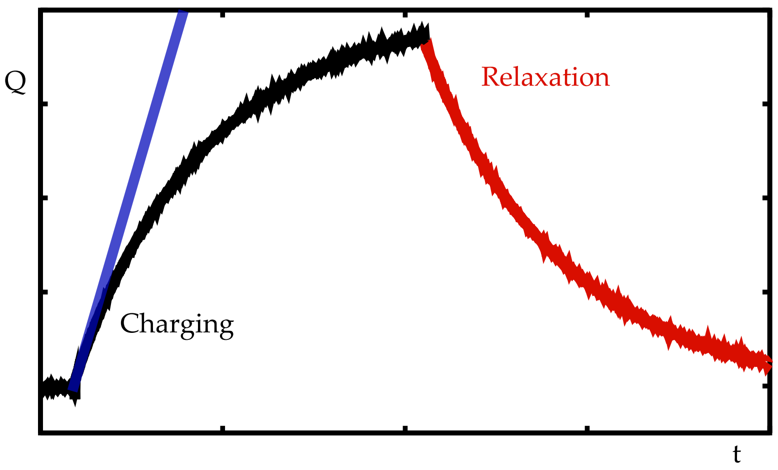


Figure 3-15: **Relaxation after a pulse is applied** After application of a single pulse, the sample must be allowed to re-equilibrate before the next pulse can be applied

by the small density of states at the bottom of the Coulomb gap. Thus, an $n$-fold reduction in the pulse width reduces the length of time we need to wait for relaxation by $\tau \ln n$ where $\tau$ is the time constant of the exponential decay. Reducing the pulse width reduces the relaxation time, but only logarithmically.[‡]

Fortunately, we can resolve this issue. It is important to realize that this long relaxation time is *not* an intrinsic property of the 2DEG. The time constant is given by $\tau = \mathrm{RC}$, where R is the (large) tunnel resistance at equilibrium, and C is the capacitance of the 2DEG to the outside world (C is dominated by the capacitance between the 2DEG and the tunnel electrode, which is about four times larger than the geometric capacitance of our structure as a whole). If we could develop the technology to perform our measurement on a structure with a much thinner tunnel barrier, and thus a much smaller R, we could reduce this relaxation time. This solution does not scale well with magnetic field. As we raise the magnetic field, we need to make our measurement window shorter because the peaks in the DOS become larger. At the same time, we would like to make our tunnel

---

[‡] There are other benefits to reducing the pulse width; most notably, we keep the overall magnitude of the charge signal smaller, allowing us to use more gain on our charge amplifier without overloading our detectors, and also it allows us to use smaller coupling capacitors after the amplifier.

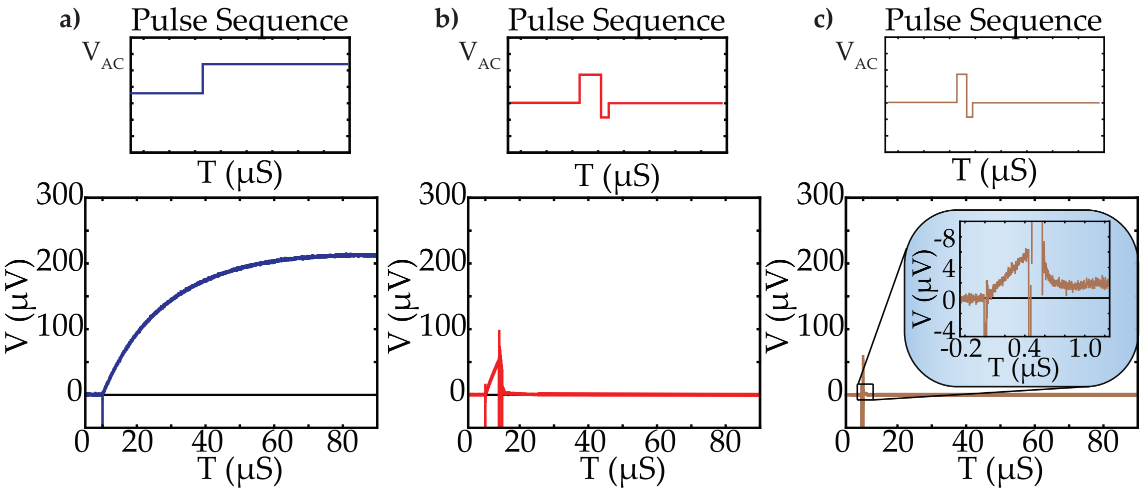


Figure 3-16: **Discharge pulse sequence**: **a)** Charge versus time response for a voltage step applied to the sample, showing the long, slow natural relaxation. **b)** Charge versus time response to the same step, cut short and immediately followed by a pulse of the opposite polarity to discharge the 2DEG. **c)** As **b**, but the pulse and discharge pulse are shortened to a length typical of the actual measurement.

barrier thinner to reduce the relaxation time, once again forcing us to make the measurement window even shorter. We are already pushing our capabilities on the width of the measurement window, so this solution is untenable.

However, this line of thought suggests an answer. If despite the tunnel barrier we could force the correct amount of charge back into the 2DEG, then we would only need to wait for internal relaxation to return the 2DEG to its ground state before applying the next pulse. The resistance of our tunnel barrier is only large at small biases, so all we need to do is arrange to discharge the 2DEG at large bias instead of near equilibrium. We do this by applying a second "discharge pulse" of the opposite polarity immediately after the pulse we use to measure the tunnel current. We then carefully tune this pulse to return the 2DEG to its original density. Figure 3-16 shows the general idea. Figure 3-16a shows the step response of the 2DEG, including the long, slow decay caused by the Coulomb gap. In Figure 3-16b, we apply a short pulse of the opposite polarity immediately after the first pulse, rapidly forcing the charge out of the 2DEG. Note that we can tell our discharge pulse is returning the structure to the correct density because the charge is not changing as a function of time long after the discharge pulse. In Figure 3-16c, we show the same measurement, with the pulse's widths reduced to values comparable to what we use in collecting our spectra. Note that both the signal amplitude and the total measurement time are much smaller, allowing us to use a higher gain and a higher repetition rate in our measurement.

Our criteria for the discharge pulse width is that the measured charge be constant as a function of time ($dQ/dt = 0$) long after the discharge pulse is done, not that the total charge on the 2DEG be the same before and after the measurement. The reason for this is that we repeat the measurement as part of a steady stream of several million equally spaced pulses. Regardless of whether or not our discharge pulse has forced the sample back to equilibrium, the structure will reach a steady state quickly (quite probably one where the 2DEG and the electrode never equilibrate). Examination of the tunnel current long after the discharge pulse is the only means we have to determine if equilibration has occurred. As we increase the magnetic field and the Coulomb gap becomes deep, we need to measure this tunnel current over a long time with enormous precision because relatively large disequilibria between the 2DEG and the electrode will result in small currents.

We now have to decide what about the discharge pulse to tune to force $dQ/dt$ to zero, returning the structure to equilibrium. As we already need excellent amplitude control over our pulses to acquire our spectra, it would seem natural to tune the amplitude of the discharge pulse. As we vary the amplitude of the discharge pulse, the tunnel current in response to it will change. Since we are only trying to remove a small amount of charge, we can approximate the tunnel current as constant as a function of time. We will thus remove an amount of charge $Q = I(V_{\mathrm{discharge}})t$ where $t$ is the width of the discharge pulse, and $V_{\mathrm{discharge}}$ is its height. The I-V curve for the discharge pulse will be almost identical to the I-V curve for the charging pulse. A sample of such a curve is shown in Figure 3-17. Because of the plateaus created by the gaps between Landau levels, for wide variations of pulse height, the tunnel current, and hence the charge removed by

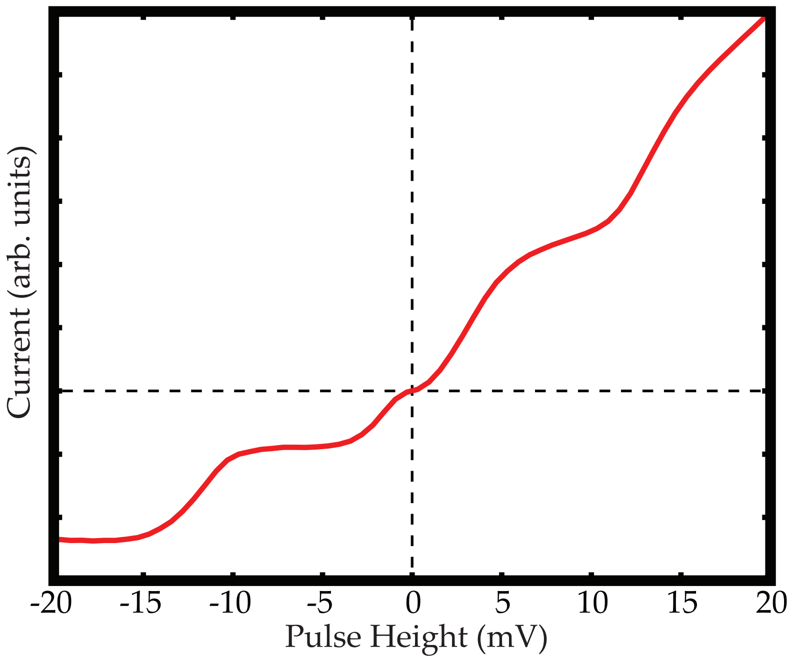


Figure 3-17: **Sample I-V curve**: Because of the highly non-linear density of states, this I-V curve, taken at 2 Tesla, has a series of plateaus separated by sharp jumps.

the discharge pulse, varies only slightly. In other regions, the current, and thus the charge, jumps suddenly as the pulse brings the next Landau level into play. While we are tuning the discharge pulses we do not yet have access to this I-V curve, so we must use a nonlinear optimization method, such as the secant method[41, Chapter 9]. These techniques do not deal well with the almost step-function nature of the I-V curve; they tend to change the pulse height by far too much when they are sampling one of the flat regions. Even if we find the correct pulse height to discharge the 2DEG, if it occurs in one of the quick steps between Landau levels our discharge pulse will tend to be sensitive to very small amounts of noise and drift.

Instead, we pick a fixed height for the discharge pulse, and vary its length. This way, the amount of charge we extract is almost exactly linear in our control variable (time). The step-like nature of the I-V curve helps us; usually the discharge pulse will fall on one of the plateaus, rendering us insensitive to small amounts of drift and noise in the 2D chemical potential and the discharge pulse height.

We still must simultaneously adjust the balancing pulse for the measurement pulse and the width of the discharge pulse. Fortunately, the ratio of the excitation pulse height to the balancing pulse height is the ratio of the geometric capacitance of our sample to the capacitance of our standard capacitor. Because this quantity is almost fixed, we can generate a good initial guess for our balancing algorithm. Similarly, we can use the width of discharge pulses from similar excitations to generate a good guess for the width of discharge pulse to use. We are now left with an optimization problem where we are trying to zero two parameters (the discontinuity at the start of the measurement pulse, $\Delta$Q and the slope long after the pulses have been applied, $m_{\text{discharge}}$) using two controls; the height of the balancing pulse sent to the standard capacitor, $\text{V}_{\text{balance}}$ and the width of the discharge pulse $w$. With a good initial guess, we can linearize this, and represent the response of the system with a 2x2 matrix $\mathcal{A}$ and a two element offset vector $\vec{o}$;

$$\begin{bmatrix} \Delta\text{Q} \\ m_{\text{discharge}} \end{bmatrix} = \mathcal{A} \begin{bmatrix} \text{V}_{\text{balance}} \\ w \end{bmatrix} + \vec{o} \tag{3.1}$$

At low fields, we can frequently go further and neglect the off-diagonal elements of $\mathcal{A}$, and even at intermediate fields we can usually neglect the impact of the balancing pulse on the discharge of the sample $\mathcal{A}_{21}$. We can then take 3 (or 4, depending on how many elements of $\mathcal{A}$ we set to zero) traces, systematically varying $\text{V}_{\text{balance}}$ and $w$ away from our initial guess by small steps to generate an initial guess for $\mathcal{A}$. For the case where we neglect the off diagonal elements of $\mathcal{A}$, the discharge pulse and balance pulse heights separate, and we apply the secant method to them individually. Otherwise, we can either assume $\mathcal{A}$ is fixed and update $\vec{o}$ with each added observation to zero in on the correct balance and discharge point, or we can try to update both $\mathcal{A}$ and $\vec{o}$ using a solver appropriate for over-determined systems, such as a least-squares fitter. In general, multi-dimensional nonlinear zero finding is as much art as science, and algorithmically determining the balance pulse height and the discharge pulse width is no exception.

Once we have successfully taken a trace using discharge pulses, we still need to guarantee that we have allowed sufficient time for the 2DEG to relax internally, as well as discharge. We do so by increasing the amount of time between pulses, and ensuring the spectrum does not change. We need to measure the charge versus time trace for a long time (tens of microseconds) to ensure no current is flowing after the pulses; even with our fastest pulse repetition rate there is a long delay between pulses. Thus, we have never seen a change in our spectrum upon increasing the period of the measurement. However, we do sometimes see the current through the device fluctuate on a timescale of microseconds immediately after the discharge pulse. We believe these fluctuations reflect chemical potential fluctuations in the quantum well as it cools.

## 3.10 Heating

As we saw earlier, one advantage to only using the initial slope of the charge versus time trace is that we minimize heating effects. We are essentially using a very short duty cycle measurement to allow our 2DEG to have time to cool. Given that this short duty cycle provides TDCS with one of its chief advantages, the ability to measure at high energies with negligible heating, it is natural to ask if we could similarly use a short duty cycle with a RTD or a quantum Hall bilayer.

In designing a short duty cycle experiment with a RTD, one needs to ensure that the applied pulse is long enough to establish a well-defined tunneling voltage between the 2D and 3D systems. In practice, this means the RC circuit formed by the tunneling barrier must be allowed to relax (Figure 3-18). However, this very relaxation requires that a substantial amount of charge tunnel into the system. In other words, establishment of the steady-state situation required to perform a traditional RTD-style measurement requires injection of a large quantity of charge prior to the measurement, heating the 2D system. Similar arguments hold for the quantum Hall bilayers, except that the charging is now spatially inhomogeneous

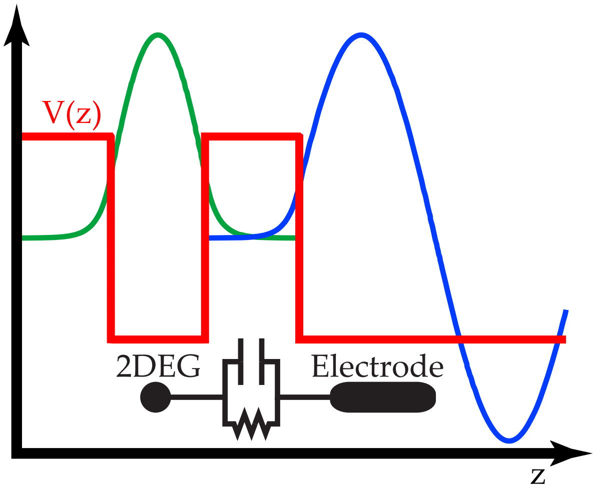


Figure 3-18: The RC circuit formed by a tunnel barrier.

as well as time varying.

## 3.11 Summary

TDCS allows us to study both electron and hole quasiparticle spectra with temperature limited resolution. Because of its pulsed nature, we can tunnel into cold electron systems at energies of tens of millivolts, equivalent to hundreds of degrees Kelvin. It provides a beautiful and intuitive probe of the physics of two-dimensional electron systems. The sample geometry allows concrete and quantitative calibration of the energy and density axes, while analytic corrections for the tunneling matrix elements and 3D density of states allows us to understand and partially compensate for differences between dI/dV and the quasiparticle density of states.

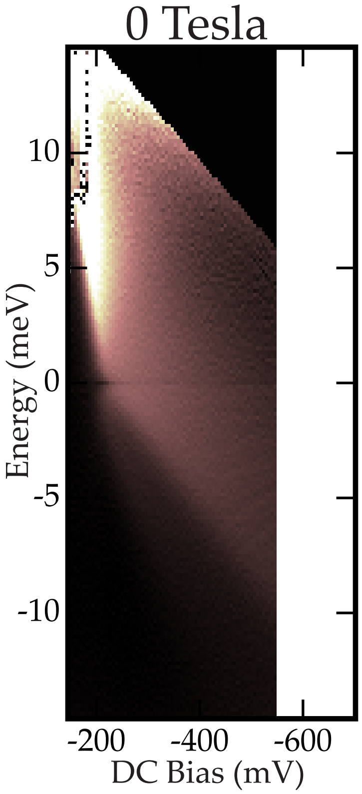
# Chapter 4

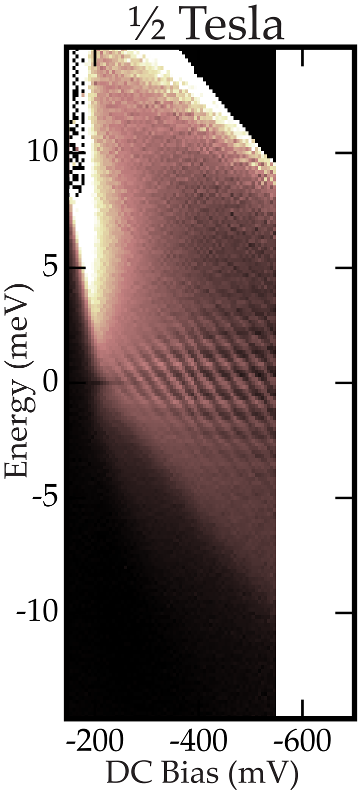
# Spectra

Figure 4-1: **TDCS Spectra** (following pages)
The main results of our measurement are tunneling density of states spectra at fields from zero to four Tesla. While the next several chapters analyze in detail the features in these spectra, they are collected here for ease of comparison. They are all plotted with the same X and Y scales, although the contrast has been adjusted individually for each. Because density calibration data is only available for some of these spectra, they are shown here as a function of DC bias.

Left: At zero field, the step in the density of states associated with filling the well is clearly visible. The plasmaron mode forms a general haze below the bottom of the band. The logarithmic gap in the density of states is visible at low densities and zero energy.

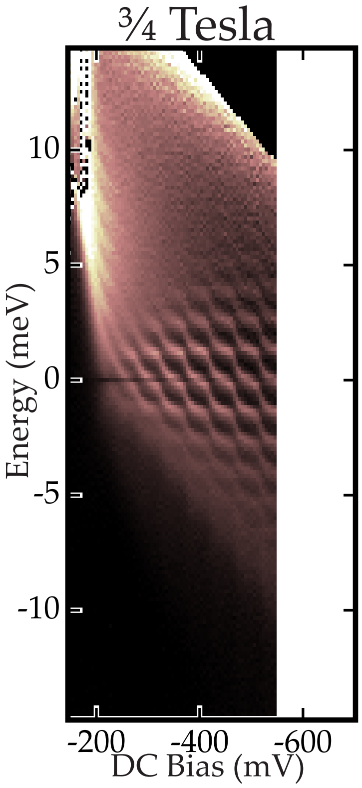
Right: At ½ Tesla, the ripples in the density of states from Landau levels become clear. Lifetime broadening blurs away the Landau levels at high energies and low densities.

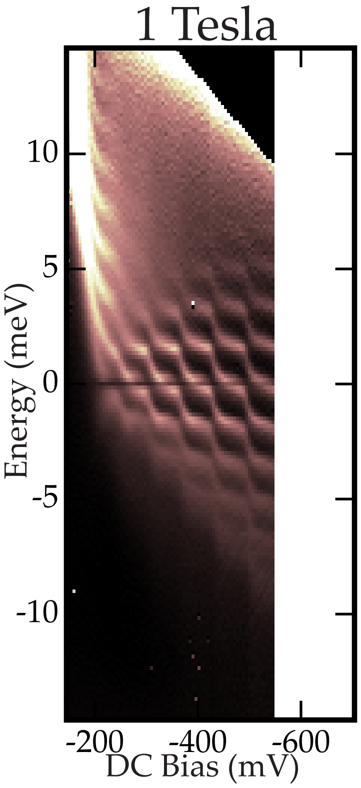

0 Tesla


½ Tesla

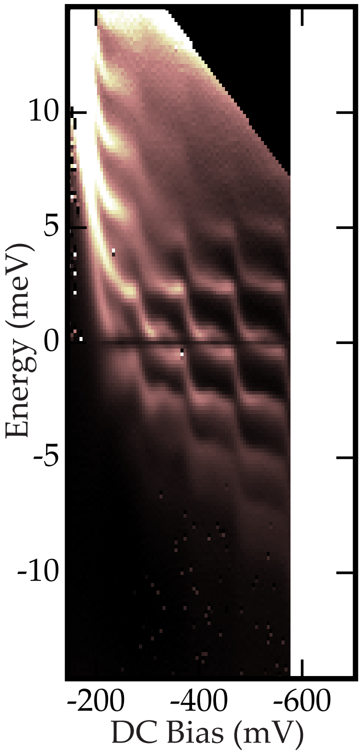
Left: At ¾ Tesla, the ripples in the density of states from Landau levels become more clear, especially at zero density, and a slight stair-step pattern from pinning of the Landau levels to the Fermi energy becomes visible.

Right: At 1 Tesla, the Landau levels become distinct and a slight exchange splitting is visible.

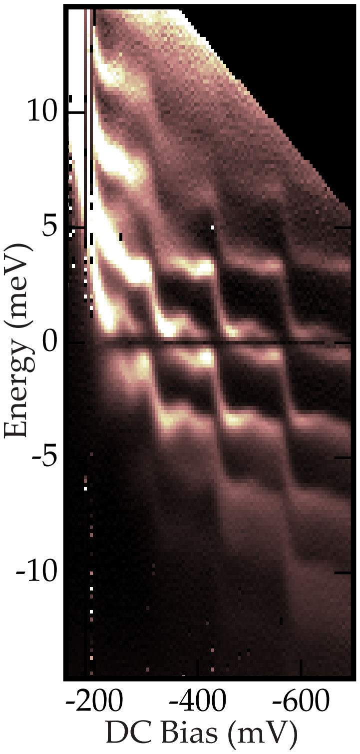

¾ Tesla


1 Tesla

1½ Tesla



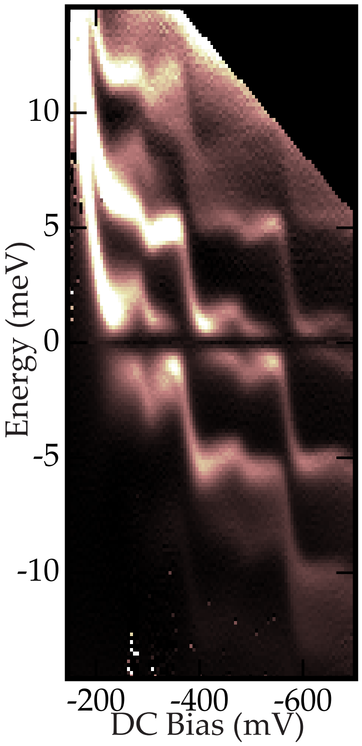
2 Tesla



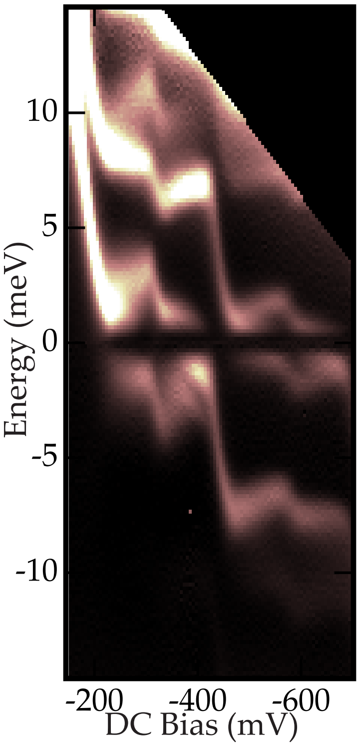
Left: At 1½ Tesla, the exchange gaps become much more evident as Landau levels cross the Fermi energy. Some indirect exchange splitting is evident in N=0.

Right: At 2 Tesla, the indirect exchange splitting becomes prominent both above and below the Fermi energy.

3 Tesla



4 Tesla



Left: At 3T, the indirect splittings continue to grow more pronounced. An unexplained sash begins to grow in N=0 between ν=1 and 2.

Right: At 4T, the sash becomes pronounced, and begins to emerge from the region between ν=0 and 1.

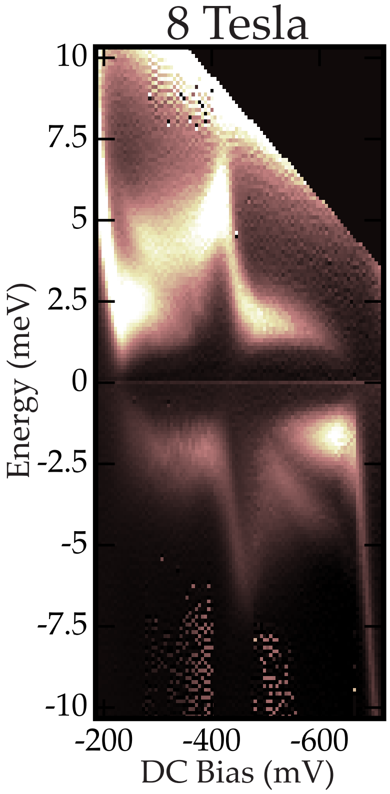


Figure 4-2: **8 Tesla TDCS Spectrum**, on a different scale from the prior plots. The jaggedness visible in the lower spin state of N=0 between filling factors $\nu = 0$ and $\nu = 1$ are a measurement artifact from poor discharge pulses, as are the speckled regions at the extreme edges of the energy range. The "sash" is clearly developed on both sides of $\nu = 1$.

# Chapter 5

# Exchange Splitting

As each Landau level crosses the Fermi surface in our spectra, we see the exchange gap open up, splitting it in two. This splitting reaches a maximum at half-filling, with one spin state completely filled. As we continue to raise the density, this exchange enhanced spin-splitting collapses, and the Landau level becomes a single peak again. We will, by careful analysis of our spectra, examine the quantitative dependence of the exchange splitting on magnetic field and filling fraction and develop an understanding of the detailed structure of the exchange splitting as Landau levels cross the Fermi energy.

## 5.1 Energy Scales

At even-integer $\nu$, the cyclotron energy, $\hbar\omega_c$ characterizes the separation between the orbital Landau levels. At odd-integer $\nu$ the Coulomb interaction between two nearest neighbor electrons in the same Landau level, $e^2/(\epsilon l_b) \propto \sqrt{\mathrm{B}}$, characterizes the exchange enhanced spin-splitting. The Zeeman energy, $g\mu_b\mathrm{B}$, is negligible by comparison in GaAs. These quantities are plotted in Figure 5-1 for electrons in GaAs.

Over the bulk of the magnetic field range accessible in our experiment, the exchange energy is larger than or comparable to the cyclotron energy. We will see that the cyclotron splitting will limit the exchange splitting, and at low fields will

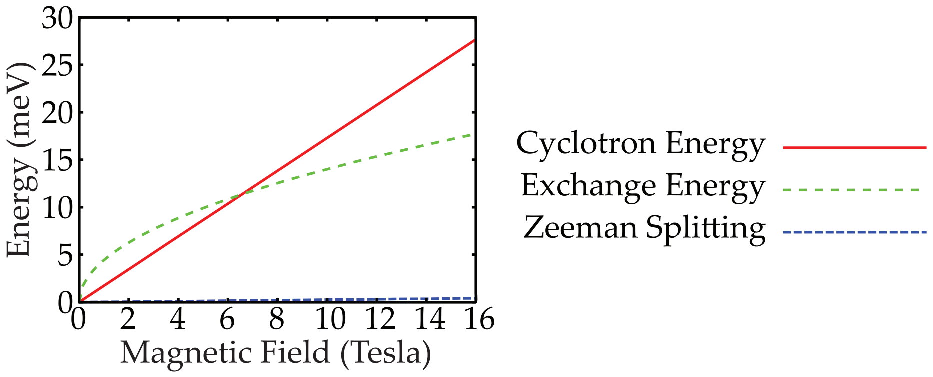


Figure 5-1: **Magnetic energy scales in GaAs**: The magnetic field dependence of three important energy scales at high magnetic fields in GaAs: the cyclotron energy $\hbar\omega_c = 1.73$ meV/T × B (solid red), the inter-electron Coulomb energy ("Exchange Energy") $e^2/(\epsilon l_b) = 4.43\ \text{meV}/\text{T}^{1/2} \times \sqrt{\text{B}}$ (dashed green), and the Zeeman splitting $g\mu_b\text{B} = 0.026$ meV/T × B (dotted blue).

give rise to a B-linear magnetic field dependence of the exchange splitting.

## 5.2 Prior Measurements

The width of the exchange gap, at its maximum, has been measured in many different ways before. Perhaps the simplest way to measure the width of the exchange gap is to examine the depth of the $\sigma_{xx}$ conductivity minima in the integer quantum Hall effect as a function of temperature[42, 43, 44, 45, 46]. At temperatures much less than the exchange splitting, this conductivity is small and dominated by hopping. As the temperature becomes comparable to the exchange splitting, large $k$-vector thermally generated excitons with the electron and hole in different spin states or different Landau levels (Figure 5-2) and a seperation proportional to $k \times l_b^2$, begin to seperate and contribute to the conductivity.

These thermal-activation based measurements consistently measure smaller exchange gaps than any of the other techniques we will examine. There are a number of possible reasons for this. First, it is unclear that the activation energy required to allow current to flow is the same as the width of the single particle exchange gap[48, 49]. Secondly, these experiments measure the width of the gap

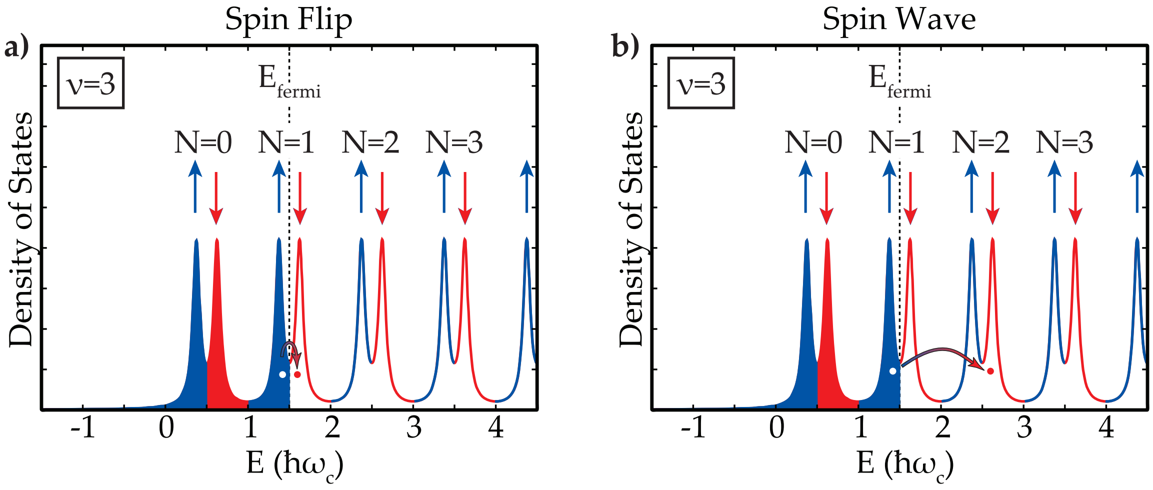


Figure 5-2: **The formation of excitons** gives rise to a temperature dependence in the conductivity minima at integer filling fractions. This temperature dependence can be used to learn about the width of the gaps between Landau levels. We use the names in common use in inelastic scattering measurements of quantum wells for different "magnetic" excitons[47]. **a)** A spin flip mode, in which an electron has its spin flipped but keeps the same orbital quantum number. These excitations dominate the temperature dependence of the conductivity. **b)** A spin wave, in which the electron flips its spin and is promoted to the next orbital Landau level. These modes are important in extracting exchange energies from long-wavelength collective excitations.

by convolving it with a Fermi function; because of this, they require a detailed a-priori model of the shape of the Landau level peaks in order to account for the narrowing of the gap because of the Landau level width. Finally, the width of the exchange gap depends on temperature; as thermally activated carriers begin to populate the upper spin-band, they narrow the gap. The measurements look only at the very low temperature tail of the conductivity, so the last effect may not be important, but this low temperature tail is also changed the most by finite Landau level widths.

A second approach that has been used is a variant on photoemission spectroscopy, discussed in Chapter 2. Photoemission spectroscopy is typically difficult to apply in 2DEGs as they are buried deep within a semiconductor and electrons are unable to escape. Kukushkin et al. overcame this problem by placing a layer of acceptors very close to the 2DEG; excited electrons recombine with these acceptors, and emit a photon[50]. This indirect luminescence technique effectively moves the electron detector into the sample. Line-widths are wider than the exchange splitting using this technique, but by manipulating selection rules, only one of the two spin states can be made visible. The exchange splitting can then be extracted by examining the deviation of the observed energy from B-linear as it is swept through the Fermi energy by varying the magnetic field[51]. The exchange gap is only published at one filling fraction and one magnetic field for this technique; this gap is somewhat larger than that found by thermal activation. Interactions between the emitted electron and the localized acceptor complicate and distort spectra obtained by this technique.

Inelastic light scattering[52] also has access to magnetic excitons. Like Raman scattering, inelastic light scattering measures the energy and momentum carried off by phonon-like excitations created during virtual excitations of carriers in a semiconductor. Because of the relatively long wavelength of light of relevant energies compared to length scales in GaAs, in the absence of impurities, only long wavelength modes can be excited. As described by Kohn's theorem, the long wavelength excitation spectrum of an ideal 2D system cannot depend on interactions[53, 54, 55]. This can be overcome in several ways. One can rely on residual disorder to break momentum conservation[6]. One can, by heroic measures, prepare a structure with low enough density to make $l_b$ comparable to the wavelength of the light used ($n \sim 10^9\text{cm}^{-2}$) [56, 57]. Alternatively, with the realization that excitations that combine a spin-flip with an orbital change can depend on electron-electron interactions, one can measure long wavelength "spin wave" modes[47]. Although all of these measurements give a quantity which depends on electron-electron interactions and can be expressed, in Hartree-Fock, in terms of the exchange integral, none of them directly measure the single particle exchange splitting. Nonetheless, this elegant technique complements tunneling measurements in its capabilities. While tunneling measures localized, charged excitations, inelastic scattering observes long-wavelength, neutral collective excitations.

Finally, capacitance measurements of tunnel capacitors, similar or identical to our TDCS structure, can be used to measure the chemical potential jump as the density is swept across the exchange gap[33, 48, 58]. The swept-density nature of this measurement is worrisome. The single particle exchange gap is a measure of the amount of energy it takes to add suddenly a single electron in an empty Landau level above the Fermi surface. The chemical potential jump, on the other hand, measures how much the energy it takes to add a single electron to an N particle system once the N+1 particle system has been allowed to relax completely varies as one slowly tunes the density from the center of one Landau level to the center of the next. One of these is a function of the single particle density of states, the other of the tunneling density of states. There is no guarantee that these densities of states are the same; for example, in the superconductors discussed in Chapter 2, the single particle density of states is gapped, while the thermodynamic density of states is continuous.

A second concern with the capacitance measurements is an experimental difficulty. These measurements rather indirectly measure the chemical potential. First, one calculates the amount of charge that has accumulated in the quantum well by comparing the capacitance of the structure at low and high frequencies; we examine this process in some detail in Chapter 11. Because of approximations made in this process, one expects errors on the order of 5% in the calculated densities. Second, one calculates the total electric field in the structure and then, using the extracted density, the amount screened by the 2DEG. Multiplying the remaining electric field by the distance between the center of the bound state of the quantum well and the edge of the tunneling electrode gives the chemical potential difference between the quantum well and the electrode. This subtraction process involves subtracting two very large electric fields to get a small one; only around 10% of the electric field penetrates the quantum well. Even after the electric field has been extracted, extraordinarily precise knowledge of the distance between the quantum well and the tunnel barrier is necessary, even though these structures move with applied bias. Finally, one needs to correct for the change in the bound state energy of the quantum well as the DC bias across it is varied in order to remove this component of the variation in the chemical potential. All told, these difficulties make extraction of the size of the chemical potential stepts with any precision impractical.

To circumvent these difficulties, the authors experimentally calibrate the total contribution of all these factors using one or both of the known Landau level degeneracy and the splitting between orbital Landau levels at even-integer $\nu$. This makes the (justifiable) assumption the distance and field corrections involved in calculating the DOS are approximately fixed across small ranges of density. However, it also assumes prior knowledge of the orbital splittings; the authors assume they are $\hbar\omega_c$. While this is true at zero density, it is not exactly true at nonzero density. As we will examine in Section 8.3, the exchange interaction modifies splittings at even-integer $\nu$. Thus, although the ratios of the chemical potential jumps at even and odd-integer $\nu$ extracted by capacitance measurements can be precise, their absolute magnitudes may be incorrect.

## 5.3 Measurements using TDCS

Because we have a precisely calibrated energy scale and can directly measure the exchange splittings as a function of B and $\nu$ in our spectra, we can easily extract the magnitude of the exchange splittings at their maximum values. Although we can see the exchange-split Landau levels separating and recombining in our spectra, as the exchange splitting shrinks it becomes confused with the Coloumb gap. As the filling fraction is raised from an even-integer $\nu$ to odd-integer $\nu + 1$, the gap at the Fermi surface changes from being dominated by the Coulomb gap to the exchange gap. As the density is raised further, back to even-integer $\nu + 2$, the gap at the Fermi surface becomes dominated by the Coulomb gap once more. While some theoretical work has been done on the evolution of the combined Coulomb and exchange gap at the Fermi surface, it is only applicable at very large densities and filling fractions[59]. We do not achieve these conditions in our experiments. Without a good model for the line-shape of our peaks when the Coulomb gap and exchange splitting are both important, we restrict our quantitative analysis of the width of the exchange splitting to exactly odd-integer $\nu$ where this confusion is minimized.

For each of the spectra in Chapter 4 at fields greater than or equal to 1 Tesla, we fit the location and width of each Landau level at each density. With the exception of levels that cross the Fermi surface, and as such have a triangular Coulomb gap cut out of them, we find that superpositions of Lorentzians fit our spectra well. At each odd-integer filling, we extract the exchange enhanced spin-splitting $E_J$ as the energy difference between the lower and upper spin state of the orbital Landau level at the Fermi energy (see Figure 5-3).

Figure 5-4a plots the set of measured splittings as a function of B and $\nu$. We will develop an empirical formula to describe the magnitudes of these splittings. At each magnetic field, there are measurements at several filling fractions, while at each filling fraction there are measurements at several magnetic fields. This overlap ensures that we can separately extract the dependence of the exchange

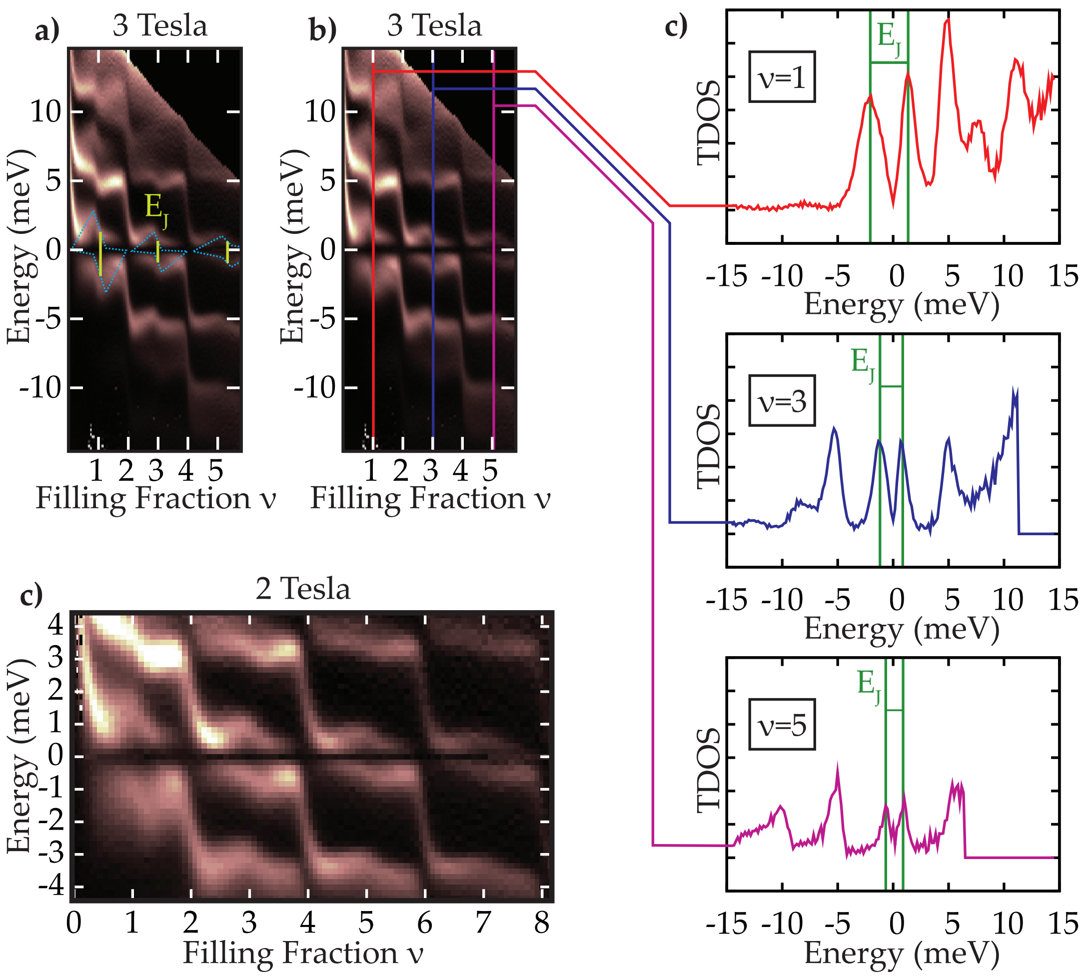


Figure 5-3: **Measuring the exchange splitting at 3 T: a)** As each Landau level crosses the Fermi surface, it splits apart as the exchange splitting grows. The high and low energy spin states jump downward at the chemical potential step at odd-integer filling while maintaining the same separation, and then the exchange splitting shrinks again, allowing the levels to recombine. **b)** We measure the exchange splitting at odd-integer filling by extracting individual columns from our spectra, as in **c**, and fitting the centers of the Landau levels using Lorentzians (not shown). The centers are illustrated using vertical green lines in the line cuts. The exchange splitting is the distance between the peak centers, shown as light-green vertical lines in **a**. **c)** TDCS spectra at 2 Tesla, showing a wide variation of the exchange enhanced spin-splittings at the Fermi surface.

enhanced spin-splitting on both filling fraction and magnetic field. We begin by looking at the exchange splitting at fixed filling fraction as a function of magnetic field. This means that as we raise the magnetic field, we increase the density of the sample to keep the fractional occupation of the Landau levels fixed.† As shown in Figure 5-4b, we find that the exchange splitting at fixed filling fraction $E_J$ is proportional to magnetic field. This is initially surprising in light of the $\sqrt{B}$ dependence of the exchange energy scale. However, as we saw in Figure 5-1, the cyclotron energy is smaller than the exchange energy scale in the range of magnetic fields discussed here.

We can think about the observed behavior in terms of screening of the electron-electron interactions or in terms of avoided crossings and Landau level mixing. In terms of screening, at high magnetic fields the electron gas at integer $\nu$ is unable to screen the Coulomb potential that gives rise to the exchange splitting because doing so would require high energy excitations to the next orbital Landau level. At lower magnetic fields screening becomes energetically favorable as the Coulomb energy is larger than the orbital energy associated with these excitations. Calculations based on simple screening models have successfully modeled a B-linear change in the cyclotron gap caused by the exchange interaction in GaN/AlGaN devices, as measured by thermal activation[60].

In terms of level mixing, we can imagine what the Landau energy spectrum would look like if we sat at fixed filling fraction, and slowly turned on the exchange interaction; physically, we could imagine placing a metal electrode near the 2DEG to screen the electron's potential, then moving it away slowly. For the moment, we ignore any splitting of Landau levels away from the Fermi surface. As the exchange interaction becomes stronger, the exchange-enhanced spin splitting of the Landau level at the Fermi surface becomes large enough that it begins to cross the next higher and lower orbital Landau levels, as shown in Figure 5-5a. However, there is no special symmetry that prevents mixing of

---

[†] By contrast, the bulk of the measurements described in Section 5.2 are performed at fixed density, not fixed filling fraction.

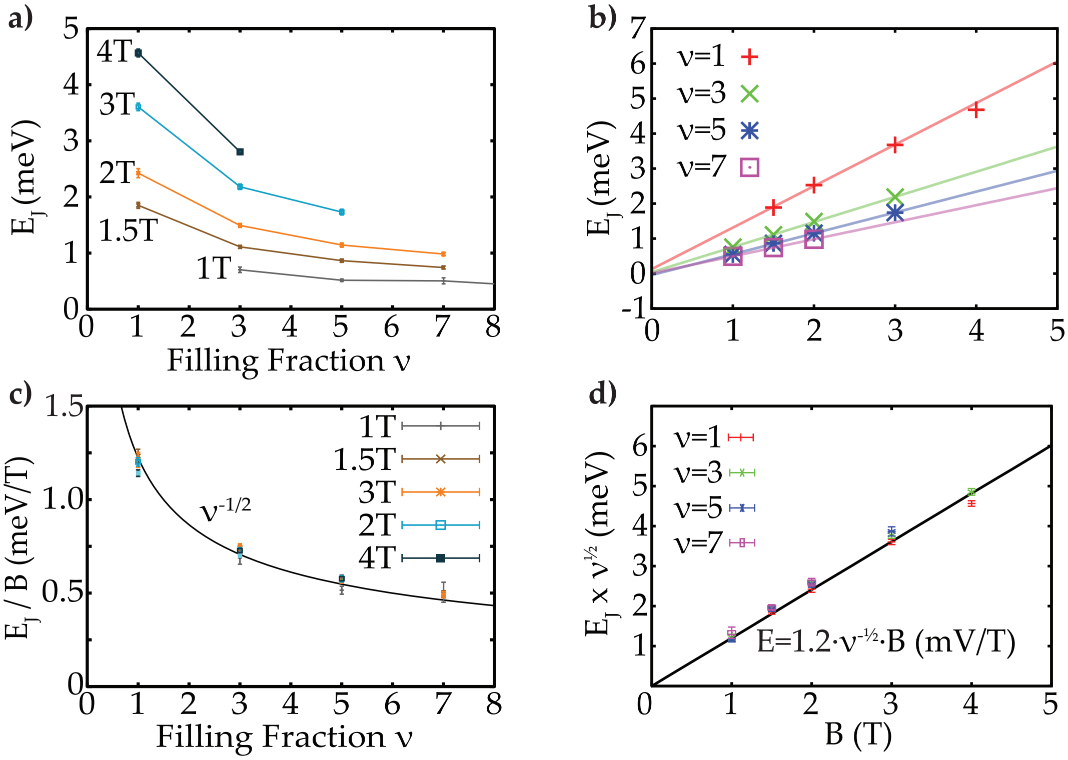


Figure 5-4: **Exchange-enhanced spin splittings** measured at a variety of magnetic fields and densities. **a)** Exchange enhanced spin-splittings as a function of filling fraction at different magnetic fields. There is substantial overlap between our data sets; we measure the splitting at at least 3 magnetic fields for each $1 \leq \nu \leq 7$. **b)** Exchange splitting as a function of magnetic field for different filling fractions. Note the clear linear dependence of the exchange splitting on magnetic field. **c)** Plotting the exchange splitting as a function of filling fraction, we can collapse the data onto a single curve by scaling out the magnetic field dependence. We find this curve fits $\nu^{-1/2}$. **d)** To check our $\nu^{-1/2}$ dependence for invariance as a function of magnetic field, we plot the exchange splitting versus magnetic field, scaling out the $\nu$ dependence. Although some deviation begins to show at 4 Tesla applied field, we find excellent agreement at lower fields.

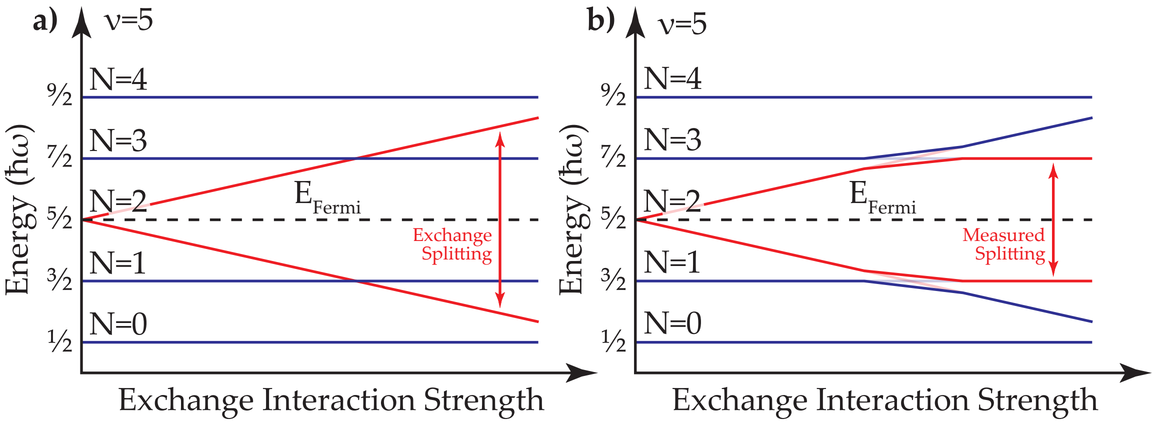


Figure 5-5: **Landau level mixing** illustrated by hypothetically turning off the exchange interaction, then slowly turning it on. **a)** If we neglect any level mixing, as the exchange interaction reaches full strength, the spin split Landau level at the Fermi energy crosses the next higher and lower orbital Landau levels. **b)** Level repulsion between the different orbital Landau levels causes an avoided crossing, limiting the apparent exchange splitting to a value comparable to the cyclotron energy.

different orbital Landau levels by the Coulomb interaction. There will be an avoided crossing, which in this simplistic picture limits the exchange splitting to a value on the order of the cyclotron energy; see Figure 5-5b. It is worth noting that within this simple model, on the far side of the avoided crossing, the nature of the orbital wavefunctions has changed. For small exchange splittings, the states just above and below the Fermi energy have orbital wavefunctions that look like $N = 2$. For large splittings (the far right of Figure 5-5b), the state just above the Fermi energy has an orbital wavefunction that looks like the non-interacting $N = 3$ state, while the state just below the Fermi energy has an orbital wavefunction that looks like the $N = 1$ state. Aside from some small modification of the tunneling matrix element, this wavefunction switch should not have observable consequences in TDCS. However, it may be possible to detect it in the magnetoplasmon dispersion curve, as measured by inelastic scattering spectroscopy.

Detailed calculations for specific devices[61] have shown B-linear exchange splittings, and a general calculation only valid for $\nu = 1$ gives $\mathrm{E}_J = \hbar\omega_c/2$[62].

Our observed splitting at $\nu = 1$ of 0.7 $\hbar\omega_c$ is somewhat larger that this value.

Having examined our B-linear exchange gap at small magnetic fields, we now turn to the filling fraction dependence. We can divide each observed splitting by the magnetic field at which it was observed, leaving us with only the filling fraction dependence; see Figure 5-4c. We find a clear $\nu^{-1/2}$ dependence once the magnetic field dependence has been removed. Finally, we check for deviations from this dependence with applied field by replotting our splittings as a function of magnetic field, with the $\nu^{-1/2}$ dependence scaled out (Figure 5-4d). The resulting data fall onto a straight line with remarkable precision; however, we see that at 4 Tesla it begins to deviate slightly. We attribute this deviation to the beginnings of the breakdown of our level mixing argument. The Coulomb energy scale is becoming comparable to the cyclotron energy, and we begin to see a crossover from our B-linear dependence to an expected (not yet observed) $\sqrt{B}$ dependence of the exchange gap.

Overall, it is not surprising that the exchange gap shrinks as we increase the filling fraction. We know that at extremely high densities (not attained here), our results must be identical to those for a non-interacting gas, so the exchange splitting must vanish at large filling fraction. Within the screening picture, at large filling fractions there are more electrons available to screen the Coulomb interaction, so the exchange splitting will shrink. In addition, as we move to higher orbital Landau levels, the wavefunctions become more spread out spatially and have more oscillations, so the magnitude of the exchange matrix element begins to shrink[63].

However, this result directly contradicts thermal activation data on the size of the exchange gap, where no filling fraction dependence is observed[42]. Given the relatively complex origin of the transport phenomena in quantum Hall systems, it is hard to interpret this difference. Capacitive measurements of the thermodynamic exchange gap do show some filling fraction dependence [48], but beyond questions about accuracy, these measurements are only available at relatively high fields and low filling fractions.

We fit the slope of the points in Figure 5-4d, forcing the fit through zero at zero field (doing so does not change the slope by of the fit by more than our statistical error). Combining this with our $\nu$ dependence, we find

$$\begin{aligned} E_J &= (1.2 \pm 0.03\text{mV/Tesla}) \cdot \nu^{-1/2} B & (5.1) \\ &= (0.70 \pm 0.02)\hbar\omega_c \cdot \nu^{-1/2} & (5.2) \end{aligned}$$

To date, no theory predicts this dependence of the splitting on filling fraction and field.

One remaining potentially contentious question regards our ability to distinguish between the Coulomb gap and the exchange splitting. There are theoretical predictions that, even at densities much higher than ours and magnetic fields much lower than ours, there is a range of filling fractions in the vicinity of odd-integer filling fractions where the exchange interaction sets the width of the gap at the Fermi surface[59, 64]. Beyond this, we observe that the gap at odd-integer filling is substantially larger than Coulomb gap we observe at filling fractions just greater than even-integer $\nu$, where there is a Landau level lying at the Fermi surface. The Coulomb gap does not have a strong filling fraction dependence, suggesting the exchange gap dominates the Coulomb gap near odd-integer $\nu$.

## 5.4 Evolution of the Exchange Splitting with Filling Fraction

Beyond understanding the magnitude of the exchange gap at its maximum, we would like to understand the way the exchange splitting grows and shrinks as the system evolves from a paramagnet at, for example, $\nu = 0$ to a ferromagnet at $\nu = 1$ and back to a paramagnet again at $\nu = 2$. Because the Coulomb gap partially obscures the evolution of the exchange enhanced spin-splitting, we will have to extrapolate somewhat near even-integer filling. We will refer to the evolution of the positions of the spin up and spin down states within the orbital Landau level that straddles the Fermi energy as the “shape” of the exchange gap, as outlined with dashed-blue lines in Figure 5-3a.

This shape is generated by interplay between disorder, pinning of the quantum well chemical potential to the Landau level energies, and the exchange interaction. In order to approach it, we will start with a simple, noninteracting system, and imagine turning on each of these one at a time until we arrive at an accurate description of our system. We will begin by imagining varying the filling fraction of a well filled with uncharged electrons; see Figure 5-6a. Because there are no interactions and no Coulomb energy, the energy of the $N = 0$ Landau level does not change as we vary the filling fraction.

We can now turn on the exchange interaction, while neglecting the capacitive energy associated with the accumulation of charge in the quantum well. We can imagine this as being in the jellium approximation, where a uniform static positive background is spread throughout the well canceling the net charge from the electrons while still allowing them to interact with each other. We will then vary the electron density by varying the charge density of this background. In this case, as we begin to fill the $N = 0$ orbital state by putting in the first electron, the spin up and spin down states are degenerate. Neglecting the Zeeman energy, there is perfect symmetry between these states. For the sake of concreteness, we will label the spin of this first electron as spin down.[†] However, when we put in a second electron, it can interact with the first. The attractive exchange interaction reduces the amount of energy it takes to inject a spin down electron, while the energy it takes to inject a spin up electron remains fixed. As we continue to fill the Landau level, the increasing density of electrons to interact with continues to pull the spin down level to lower energies. As discussed in the next section, in simple models of the exchange splitting this interaction energy is directly proportional to the Landau level occupation, so we draw the energy of the spin down level as decreasing linearly with filling fraction (Figure 5-6b). Once the spin down state

---

[†] In reality, the Zeeman energy will break this degeneracy; however, the Zeeman energy is too small for us to measure directly in the face of the exchange interaction.

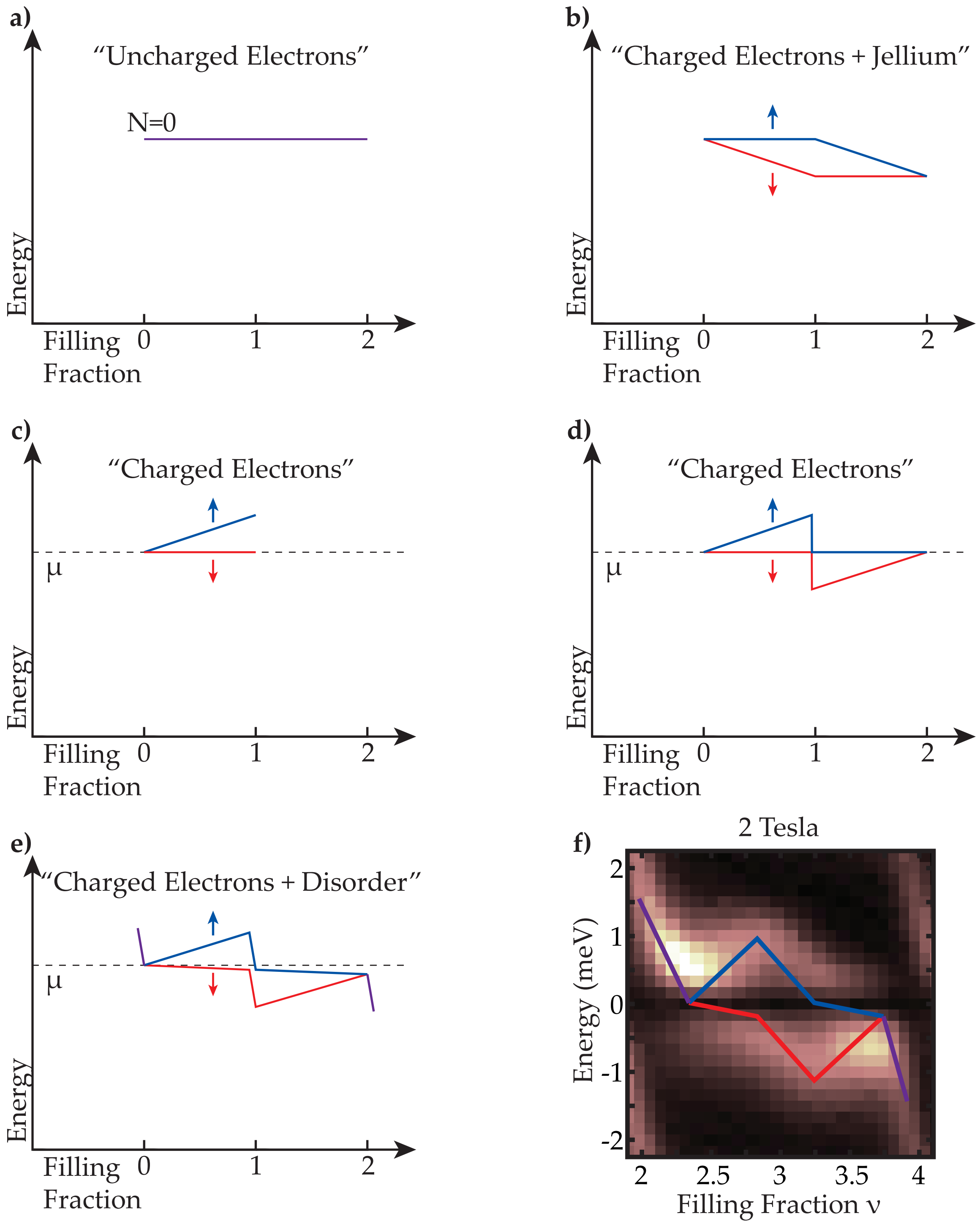


Figure 5-6: **The shape of the exchange gap**, built up by considering a simplified system of uncharged electrons, and adding effects one at a time. **a)** Non-interacting electrons **b)** Interacting electrons with uniformly charged background **c,d)** Charged electrons in tunnel capacitor **e)** Charged electrons with disorder **f)** Actual data at 2 Tesla for comparison.

has been completely filled with electrons at $\nu = 1$, we are forced to begin putting electrons into the higher energy spin up state. Just as with the spin down state, the exchange interaction between the tunneling electron and those already in the spin up state pull the spin up state down in energy as we fill it, until at $\nu = 2$ we restore the symmetry between the spin up and spin down electrons and the levels become degenerate again.

In reality, we do not control our electron density using a uniform positive background of charge; we gate it by using a distant electrode to vary the electron density. This gate pulls the bound state energy of the quantum well down relative to the tunneling electrode, which pulls with it the energies of all of the Landau levels. When the chemical potential lies inside of a Landau level, it is pinned there by the high density of states; we need to fill the level completely before we can pull the next Landau level down to the chemical potential. The electrostatics of this pinning are discussed in somewhat more detail in Chapter 11 and the associated references. Because of this pinning, as we fill the spin down states between $\nu = 0$ and $\nu = 1$, the energy of the spin down states are pinned to the chemical potential. This pinning forces the spin up states to move upward in energy; see Figure 5-6c. Once we have filled the spin down states completely, in order to put the next electron into the system, we have to pull the spin up states down to the chemical potential (Figure 5-6d); this results in the chemical potential jumps between Landau levels. Finally, as we fill the spin up states in between $\nu = 1$ and $\nu = 2$, the pinning of those states to the chemical potential pulls the spin down states up in energy until they reach the chemical potential. This combination of exchange splitting and pinning gives us a distinctive bow-tie shape.

Of course, our real sample has some disorder. This broadens our levels so the density of states in between Landau levels is not exactly zero. Because of this, our chemical potential jumps do not occur instantaneously, but rather over a small range of densities, slanting the edges of the chemical potential jumps somewhat (Figure 5-6e).

This simple model captures the basic shape that we see repeatedly, each time an orbital Landau level crosses the Fermi surface (Figure 5-6f).

## 5.5 The Exchange Splitting and Disorder

The cartoon picture we have developed qualitatively explains the evolution of the exchange splitting in our data. However, we vastly simplified the role of disorder in this picture. In reality, because of the finite width of the Landau levels, we do not fill one spin state up completely, and then begin filling the other; we fill them both simultaneously. When the exchange splitting is small, we may put nearly equal numbers of electrons into both spin states; in this case, the exchange splitting may grow very slowly at first, and then suddenly "open up" once it is comparable to the disorder induced level width.

We can construct a very simple model for the exchange splitting after Ando and Uemura [65];

$$\mathrm{E}_J \propto \left(n_\uparrow - n_\downarrow\right) \tag{5.3}$$

where $n_\uparrow$ and $n_\downarrow$ are the occupations of the spin up and spin-down sub-bands, respectively. In the case that the Landau levels are infinitely narrow, this will give us our earlier cartoon model. If the Landau levels are broadened, we will have to calculate self-consistently the populations of the levels and the exchange splitting.

Based on our observed lineshapes, we represent the Landau levels as Lorentzians of half-width half-maximum $\Gamma$, where $\Gamma$ is the disorder-induced width of the Landau levels. For simplicity, we assume $\Gamma$ is not a function of filling fraction. We compute the occupation of each Landau level as

$$\mathrm{N}_i^{\uparrow,\downarrow} = \frac{\mathrm{N}_D}{\pi}\left(\arctan\left(\frac{\mu - \mathrm{E}_i^{\uparrow,\downarrow}}{\Gamma}\right) + \frac{\pi}{2}\right) \tag{5.4}$$

where $\mathrm{N}_D$ is the degeneracy of a spin-split Landau level, $\mathrm{E}_i^\uparrow$ and $\mathrm{E}_i^\downarrow$ are the energies of the centers of the $i$th orbital Landau level's spin up and spin down states, and $\mu$ is the chemical potential. Given $N_i^{\uparrow,\downarrow}$, we compute the energies of the centers of the Landau levels as

$$E_i^{\uparrow,\downarrow} = \hbar\omega_c(i + 1/2) - J_{ii}N_i^{\uparrow,\downarrow} \tag{5.5}$$

where $J_{ii}$ gives the maximum possible magnitude of the exchange splitting of the $i$th Landau level. We self-consistently solve Equation 5.4 and 5.5 using a non-linear root finder. In order to generate a simple simulated map of the DOS, we calculate the occupations of the Landau levels as a function of chemical potential, and compute a simplified DOS spectrum for each chemical potential as

$$D(\epsilon) = \sum_i \left( \frac{N_i^{\uparrow}}{\pi} \frac{\Gamma}{\left(\epsilon - E_i^{\uparrow}\right)^2 + \Gamma^2} + \frac{N_i^{\downarrow}}{\pi} \frac{\Gamma}{\left(\epsilon - E_i^{\downarrow}\right)^2 + \Gamma^2} \right) \tag{5.6}$$

In constructing this DOS, corresponding to placing a Lorentzian at the center of each of our Landau levels, we have assumed the tunneling density of states is identical to the thermodynamic density of states, neglecting both the finite quasiparticle lifetime and any renormalization of the cyclotron energy as a function of density or quasiparticle energy. However, this simple model is still adequate to illustrate that a simple occupation-dependent exchange splitting survives in the presence of disorder, and does not introduce additional features not seen in our spectra.

On doing so, a difficulty arises with the attractive nature of the exchange interaction. The total density, $N_{total}$, as a function of the chemical potential is multi-valued; this is a consequence of the regions of negative compressibility of the electron gas[66, 67]. This does not actually result in bistability of the density in real samples; we have not included the capacitive energy of the structure in our calculation of the chemical potential. Because we only measure energies relative to the chemical potential in our experiment, we have no direct access to this capacitive energy; neglecting it has no measurable consequences. One solution to this difficulty is to perform the calculation of the spectral function as

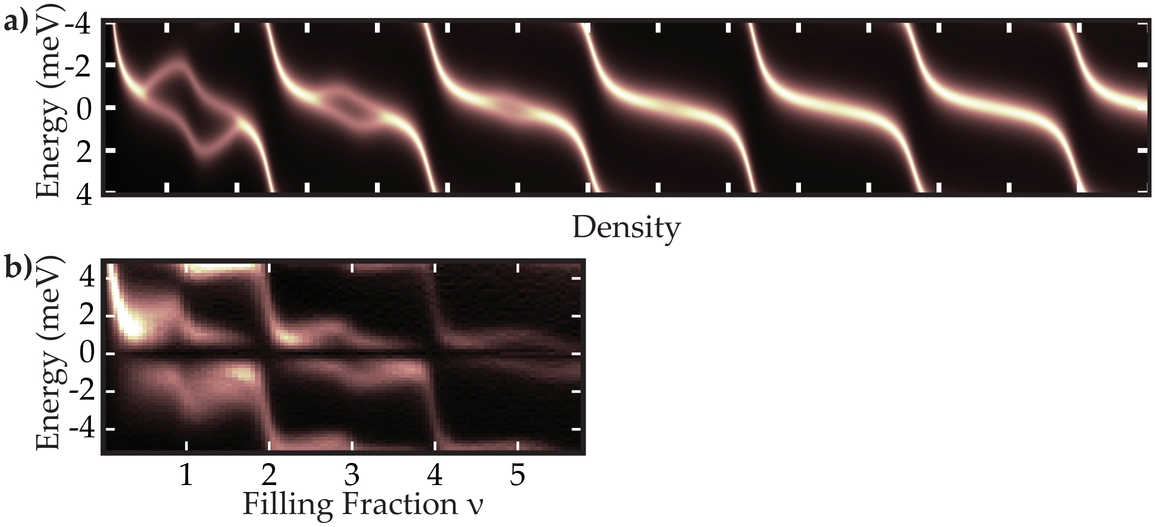


Figure 5-7: **a) Calculated exchange DOS at 3 Tesla** using Equation 5.6. Values of $J_{ii}$ are extracted from our experiment. **b)** Measured DOS spectrum at 3 Tesla. Note the good agreement on the overall shapes formed by the spin-split Landau levels, neglecting the Coulomb gap.

a function of density rather than chemical potential[63]. However, we eventually intend to combine this density of states with our numeric models of the structure (Appendix C); it would be difficult to rework this large body of code in terms of the density rather than the chemical potential.

Rather than do this, we use an approach motivated by the capacitive energy we neglected earlier. We transform from the real energies and chemical potentials to variables than include a "fake" capacitive energy, $E_i^{'\uparrow,\downarrow} = E_i^{'\uparrow,\downarrow} + CN_{total}$ and $\mu' = \mu + CN_{total}$, where $C$ is a parameter that characterizes this energy. Note that for any given solution, we have just offset our energy scale by a fixed amount, so there is a one-to-one mapping between solutions $\mu, N_i^{\uparrow,\downarrow}, E_i^{\uparrow,\downarrow}$ and solutions using our modified energies $\mu', N_i^{\uparrow,\downarrow}, E_i^{'\uparrow,\downarrow}$. The advantage of working in terms of the fake energies is that if $C$ is large enough, there is only one solution to our equations for any given $\mu'$. In fact, in the limit where $C$ goes to infinity, working in terms of $\mu'$ is equivalent to working in terms of density. In Figure 5-7, we show the calculated DOS spectrum at 3 Tesla, with a single $\Gamma$ taken from our experimental data near the Fermi energy and $J_{ii}$ taken from our experimental splittings.

Comparing to actual data in Figure 5-7b, we see that not only does the realistic inclusion of disorder not invalidate our cartoonish model of exchange, but it also brings the evolution of the splittings into even better agreement with our experimental results. Because the shapes formed depend only on the ratio between the level width $\Gamma$, taken here to be 500 $\mu$eV, and the exchange energy $J_{ii}$, the array of shapes in Figure 5-7 can be taken to be representative of the different shapes we should expect to see in our data at all magnetic fields. This value of $\gamma$ is taken to be roughly the minimum observed Landau level width, roughly approximating our peak-to-peak disorder. Comparing these shapes to those in Chapter 4, we see excellent agreement in the general shape, ignoring the Coulomb gap.

## 5.6 Summary

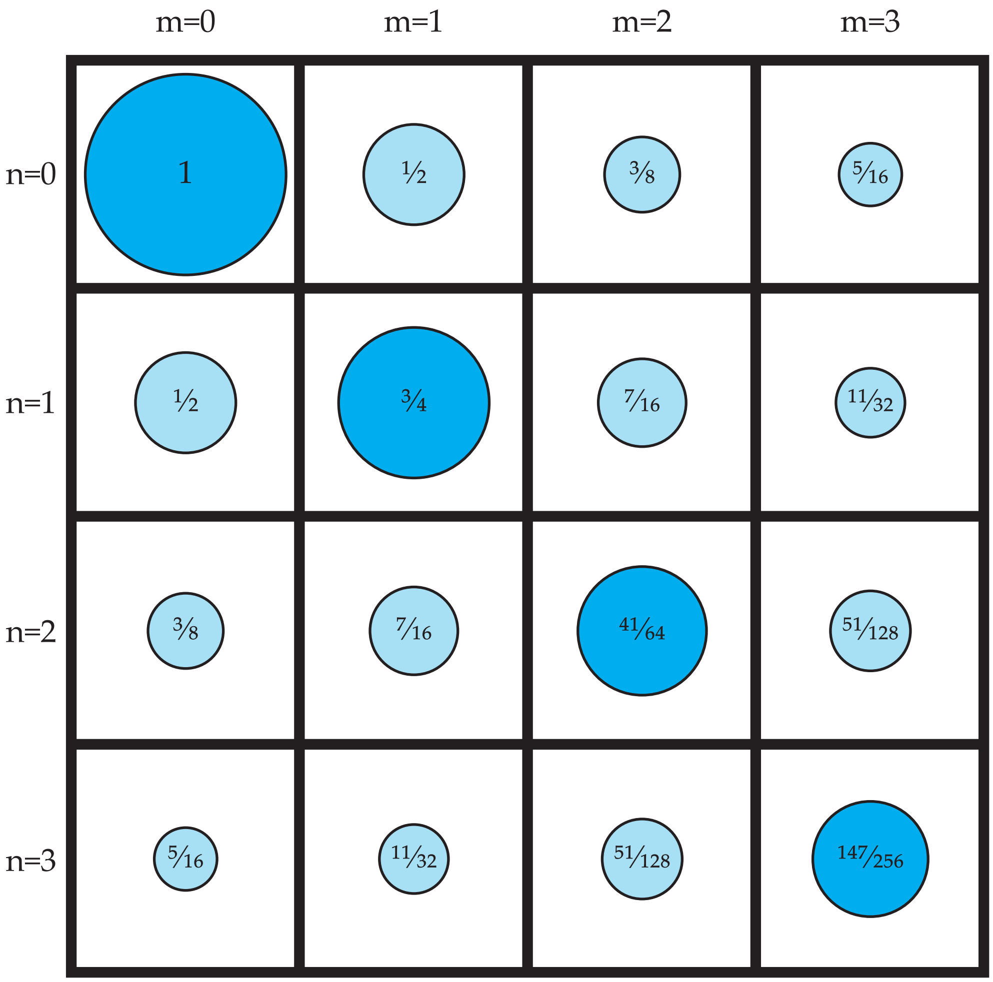
Using our DOS spectra, we directly extract the exchange enhanced spin-splitting at odd-integer $\nu$ at fields up to 4 Tesla. We develop an empirical formula to describe this splitting and briefly discuss its general form. The evolution of the splittings as the filling fraction is varied originates with the interplay between disorder, pinning of the Landau levels to the chemical potential, and the exchange splitting. We describe a simple model that explains the shape in our spectra and use a more detailed numerical model to examine its robustness in the face of disorder.

# Chapter 6

# Indirect Exchange Splittings

In the previous chapter, we examined the way the exchange splitting splits the orbital Landau level closest to the chemical potential. If we look at our higher field spectra (see Figure 4-1) in the vicinity of odd-integer filling fractions, we notice a peculiar feature in the empty states in the Landau levels above the Fermi energy and in the filled states in the Landau levels below the Fermi energy. They split apart in a way similar to the Landau level at the Fermi surface. There is no net spin polarization in either of these Landau levels; they are completely empty and completely filled, respectively. The states in these levels are orthogonal to those that we are filling and emptying as we vary the density. Naively we might not expect to see these completely empty and completely filled levels split as we vary the filling fraction. However, the exchange splitting involves matrix elements of the Coulomb interaction $e^2/(r_1 - r_2)$; these matrix elements do not vanish between Landau levels with different orbital indices. In Figure 6-1, we illustrate the magnitude of these exchange matrix elements relative to that for two electrons of like spin in the $N = 0$ orbital state, as calculated in [63]. We denote the exchange matrix element between an electron in $N = m$ and an electron in $N = n$ by $\mathrm{X}_{mn}$. The diagonal elements of this X matrix, $\mathrm{X}_{ii}$ give the usual exchange matrix elements between two electrons in the same orbital Landau level.

We will refer to this splitting as the "indirect exchange splitting" in order to distinguish it from the more familiar splitting at the Fermi surface. Indirect refers



|  | m=0 | m=1 | m=2 | m=3 |
|---|---|---|---|---|
| n=0 | 1 | 1⁄2 | 3⁄8 | 5⁄16 |
| n=1 | 1⁄2 | 3⁄4 | 7⁄16 | 11⁄32 |
| n=2 | 3⁄8 | 7⁄16 | 41⁄64 | 51⁄128 |
| n=3 | 5⁄16 | 11⁄32 | 51⁄128 | 147⁄256 |

Figure 6-1: **Relative magnitudes of exchange integrals**, as calculated in [63]. The sizes of the circles depict the magnitude of the exchange matrix element between an electron in $N = m$ and $N = n$, normalized to the interaction between two electrons in $N = 0$. Note that the matrix element is always largest between electrons in the same orbital Landau level, where the wavefunction overlap is the largest, and becomes smaller as $N$ increases and the orbital wavefunctions become spread out.

to the fact that it is a splitting induced by electrons in one orbital state on electrons in a different orbital state. Our spectra are, to the best of our knowledge, the first experimental observations of these indirect splittings. Much as with the direct exchange splittings, we can try to understand the shape of these splittings and their relative magnitudes.

## 6.1 A New Probe of Many-Body Ground States

In Figure 6-2, we examine the indirect splittings at 3 Tesla in some detail. As shown in **b** and **e**, when the 2DEG is at even-integer filling, all of the orbital Landau levels, including those above and below the Fermi surface, appear as a single, well defined peak. As we begin to fill one spin state of a given orbital level, the filled Landau levels above (**c**) and below (**f**) the Fermi surface begin to split apart due to the indirect exchange splitting, and continue to grow farther apart as we approach odd-integer filling.

The exchange enhanced spin-gap is a relatively delicate, low temperature phenomenon. The gap at the Fermi surface is typically on the order of a few hundred microvolts, or a few Kelvin in temperature units. If the temperature is raised to on the order of the exchange splitting, thermal occupation of the upper and lower spin states will equalize their populations and collapse the exchange gap. In injecting electrons or holes into excited orbital states, we are probing the 2DEG at energies on the order of ten millivolts, equivalent to one hundred Kelvin. The observation of this indirect splitting shows that the high energy spectra can reflect the low temperature many-body ground state of the electron liquid. This high energy spectrum constitutes a new and unexplored probe of the nature of the many-body ground state of the 2DEG.

Injecting electrons at energies equivalent to 100 Kelvin into a low temperature system is not equivalent to measuring equilibrium properties in a 100 Kelvin system. The persistence of these indirect exchange gaps at high energies is further evidence of the non-invasive nature of TDCS; if we were substantially heating the

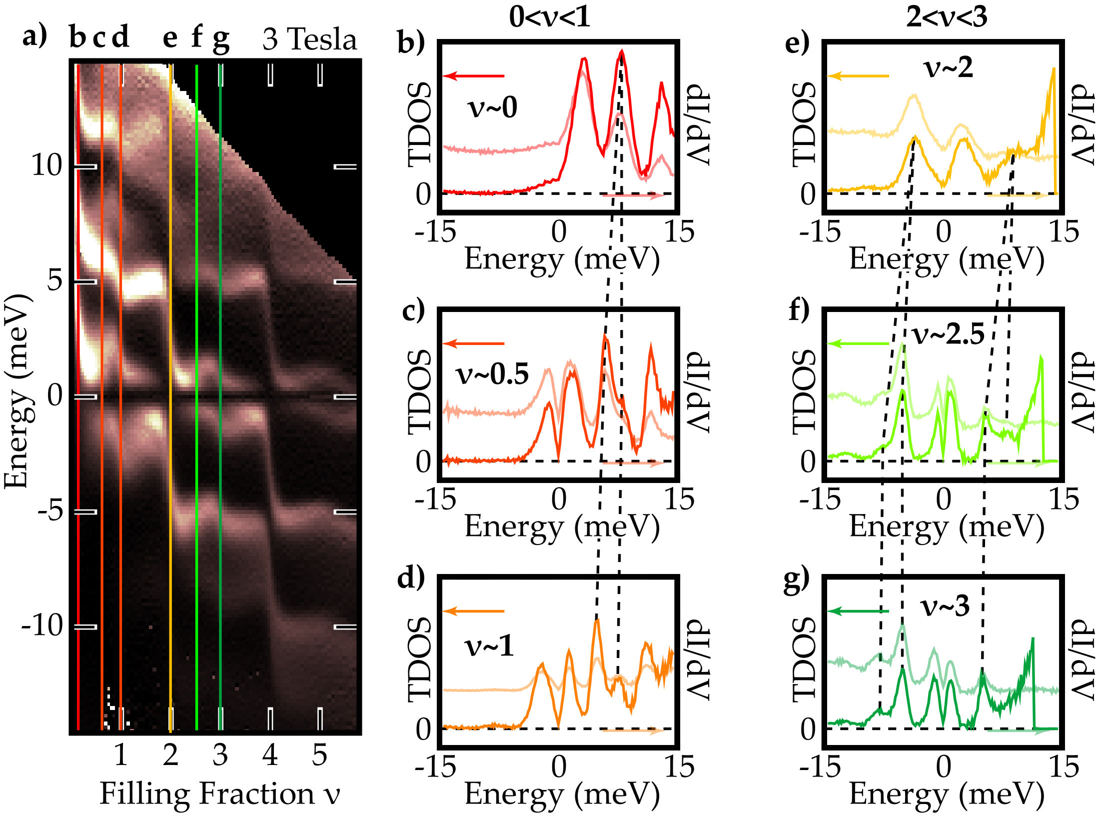


Figure 6-2: **Indirect exchange at 3 Tesla: a)** 3 Tesla TDCS spectrum, showing both direct exchange at the Fermi surface, and indirect exchange in Landau levels above and below the Fermi surface. **b-g)** Line cuts from **a** showing indirect exchange splittings. Exchange-split peaks far from the Fermi energy are indicated with dashed black lines. Data without matrix element correction is included in the line cuts as a lighter line. When the indirect exchange splitting occurs near the region where the correction becomes singular, the splitting can be clearer before the correction is applied.

electron gas, these gaps would collapse. Because we perform our measurement at the first instants of our applied pulse, before a significant amount of heating has occurred, we probe a cold electron gas, even at the highest energies at which we measure.

We consistently find that the spin-state closest to the Fermi surface is brighter and sharper than that farther away in these indirect-split levels. For example, continuing to call the lower energy spin-state "spin down" and the higher energy spin state "spin up", we see that at filling fractions $3 < \nu < 4$, the spin up state of $N = 0$ and the spin down state of $N = 2$ are bright and easy to see. By comparison, the spin down state of $N = 0$ and the spin up state of $N = 2$ are dim and broad features; on the line-cuts through the spectra, they often contribute a small shoulder rather than a separate well defined peak. In many places, it is not possible to resolve these states at all. The difference appears to involve both a broadening of the farther state and a reduction of its spectral weight. At present, the reason for this disparity between the spin state closer to the Fermi surface and that farther away is not understood. It prevents us from accurately assessing the magnitude of the indirect exchange splitting at its maximum. Because of this, we will restrict ourselves to qualitative discussions of the size of the indirect exchange splitting, unlike our detailed fitting of the direct splitting.

## 6.2 Evolution of the Indirect Exchange Splitting with Filling Fraction

Just as with the direct splitting, we can develop a physical model for the unusual shape the indirect exchange splitting takes. The situation is somewhat more complicated now; we now have two energy scales: the magnitude of the direct exchange gap and that of the indirect exchange gap. We will consider two rough cases in parallel; the indirect splitting can be smaller than the direct (as illustrated in the top row of Figure 6-3), or larger (as illustrated in the bottom

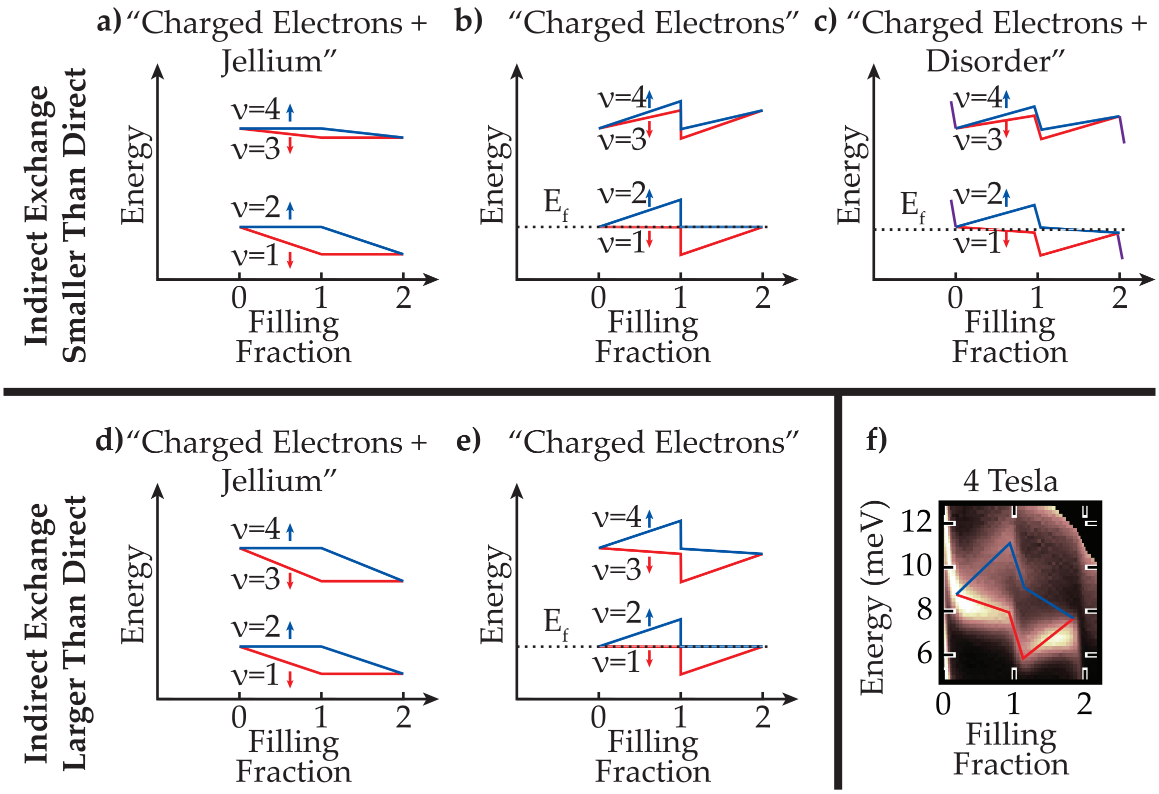


Figure 6-3: **The shape of the indirect exchange gap**, built up by considering a simplified system of uncharged electrons, and adding effects one at a time. We consider both the case where the indirect gap above the Fermi surface is smaller than the direct gap (**a**-**c**), and the case where it is larger (**d**,**e**). **a)** Interacting electrons with uniformly charged background **b)** Charged electrons in tunnel capacitor **c)** Charged electrons with disorder **d)** Interacting electrons with uniformly charged background **e)** Charged electrons in tunnel capacitor **f)** Actual data at 4 Tesla for comparison.

row of Figure 6-3). In panels **a** and **d**, we show the energies of two adjacent orbital Landau levels as we proceed from even to odd and back to even integer filling fractions. Without loss of generality, we take these to be $\nu = 0, 1, 2$, and consider injecting electrons into an orbital level above the Fermi energy. The shape generated by injecting holes into a level below the Fermi energy is similar.

At zero density, both the $N = 0$ and the $N = 1$ Landau levels are spin degenerate. As before, we begin to fill the lower spin state, spin down, of $N = 0$ by increasing the DC bias. If we now tunnel an electron into the spin down state of $N = 0$, it feels an attractive exchange interaction to the electrons already in the well. This lowers the energy of the spin down $N = 0$ state on our spectrum. If we consider tunneling an electron into the *empty* spin-down state of $N = 1$, there are no other electrons in $N = 1$ with which it can interact. However, it has a weak attraction to the electrons already in $N = 0$; referring to Figure 6-1, the matrix element is half the size of that for an electron tunneling into $N = 0$. This weak attraction pulls down the energy of the spin down state of $N = 1$, just as it pulls down the spin down state of $N = 0$. As we continue filling the 2DEG, the $N = 0$ states and $N = 1$ states continue to evolve similarly, and in the same fashion as discussed in Section 5.4. From the viewpoint of the bare exchange matrix elements, we only expect the situation in panel **a** to be observed; panel **d** apparently shows more attraction between an electron in $N = 1$ and $N = 0$ than between two electrons in $N = 0$.

As before, to include the effects of Landau level pinning we distort the diagrams so that the Landau level being filled is always at the same energy. This is illustrated in panels **b** and **e**. In the case where the indirect splitting is smaller than the direct splitting, this will cause the $N = 1$ spin down state to apparently move upwards in energy as we fill the well from $\nu = 0$ to $\nu = 1$ (panel **b**). This is a manifestation of the negative compressibility of the quantum well in this region; neglecting the external capacitive energy of the 2DEG, the chemical potential of the well drops as we raise its density. This is characterized by the tendency of the well to over-screen in these regions; it will pull in more charge than it needs to screen applied electric fields, and the capacitance of our tunnel capacitor will exceed what would be observed if the 2DEG were replaced by a perfect metal[66, 49]. If the indirect splitting were larger than the direct splitting, as in **e**, the spin-down state of the $N = 1$ orbital level will move downward in energy.

Because of disorder, the Landau levels will have finite width. As we raise the density, the centers of the Landau levels move downwards as we fill the lowest energy states in them first and then fill the higher energy states. This is illustrated in **c**. In extreme cases, this slope may be large enough that the $N = 1$ states will both move downward in energy, even when the indirect splitting is smaller than the direct. This is the case near $\nu = 1$ at lower magnetic fields; see Figure 4-1. Putting together all of these effects, we find a bow-tie shape for the Landau level nearest the Fermi surface, while the Landau level farther away can form either a bow-tie or a wave-like shape. We find examples of both of these shapes in our data.

## 6.3 The Indirect Exchange Splitting and Disorder

When we add disorder to the indirect exchange splitting, we have interplay between three energy scales: the direct exchange splitting, the indirect exchange splitting, and the disorder amplitude. The simple cartoon developed above includes only a rough approximation of the disorder. We can see if our simple model of Section 5.5 can reproduce the range of shapes we see in our actual data by extending it to include these indirect exchange splittings. We do so by continuing to use our experimentally observed exchange-enhanced spin splittings near the Fermi surface, but using the ratio of the elements of $X_{mn}$ to approximate the indirect exchange splittings. we will replace Equation 5.5 with

$$E_i^{\uparrow,\downarrow} = \hbar\omega_c(i + 1/2) - \sum_j J_{ii}\frac{X_{ij}}{X_{ii}}N_j^{\uparrow,\downarrow} \tag{6.1}$$

This form for the energy dependence is, aside from the use of the experimentally extracted on-diagonal exchange splittings, that derived by MacDonald et al. in the absence of Landau level mixing[63].

The resulting spectrum is illustrated in Figure 6-4. We find the variety of indirect exchange splitting shapes in our simple model (**a**) mirror those we see in our data (**b**). However, substantial and qualitative differences are visible.

Because the off-diagonal elements of $X_{mn}$ are smaller than the on-diagonal elements in the same row and column, in the simulation the indirect splittings are always smaller than the splittings at the Fermi surface. This is in general true

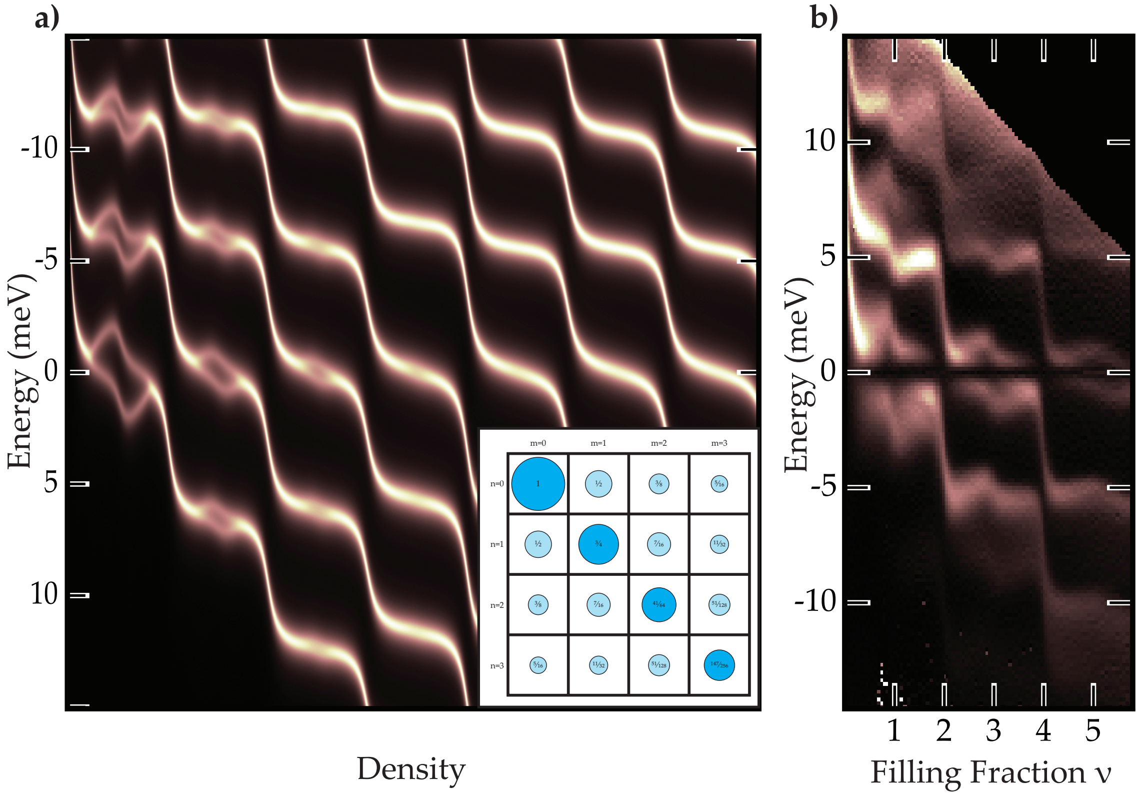


Figure 6-4: **a) Calculated exchange DOS at 3 Tesla** using Equation 5.6. Values of $J_{ii}$ are extracted from our experiment, while the ratio between $J_{ii}$ and $J_{ij}$ are taken from those in [63] (Figure 6-1). **b)** Measured DOS spectrum at 3 Tesla. Note the good agreement on the overall shapes, and the disagreement on the magnitude of the indirect exchange splitting. Inset; a graphical representation of the relative sizes of the exchange matrix elements, as in Figure 6-1, for reference.

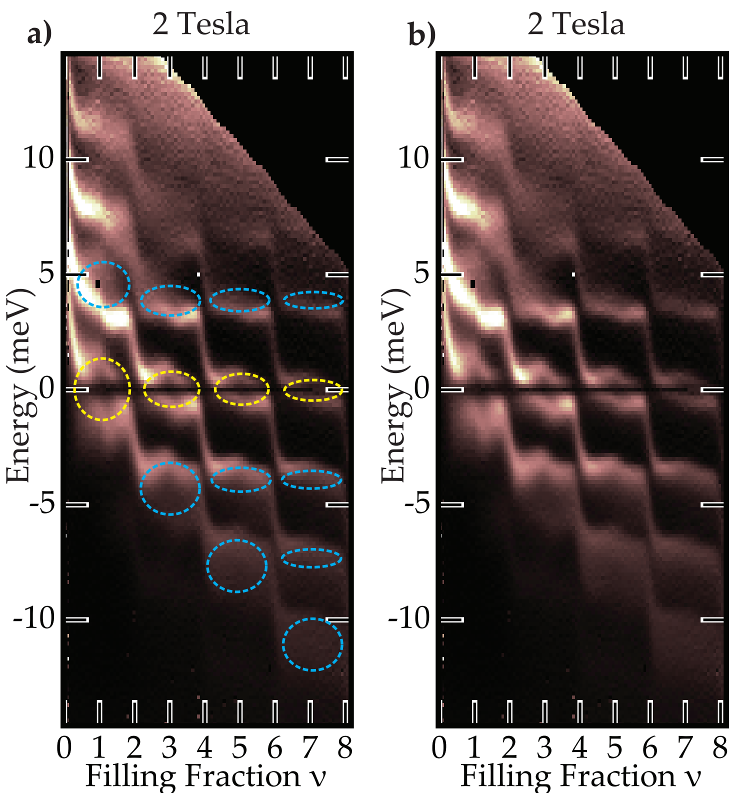


Figure 6-5: **Indirect Exchange Splittings at 2 Tesla**, with the splittings emphasized by blue ovals. In regions where one spin state is not easily visible, the ovals have been placed assuming the splitting to be symmetric. Yellow ellipses mark the direct exchange splitting at the Fermi energy. **b)** has the ovals removed for clarity.

in our spectra; see Figure 6-5. However, at any given density, the $N = 0$ Landau level is always split by as much as or more than every other level in the spectrum.

For most orbital levels, the nearby energy spectrum is symmetric; there are Landau levels that lie above and below, separated by (approximately) $\hbar\omega_c$. However, for the $N = 0$ Landau level, there is no lower level. This special position in the spectrum gives it markedly different properties.

Avoided crosses limit the indirect exchange gaps just as they limit the exchange-enhanced spin gap at the Fermi surface. However, the $N = 0$ state's unique position in the spectrum means that there is no lower Landau level for it to mix with, as shown in Figure 6-6. Its exchange splitting can grow larger than any

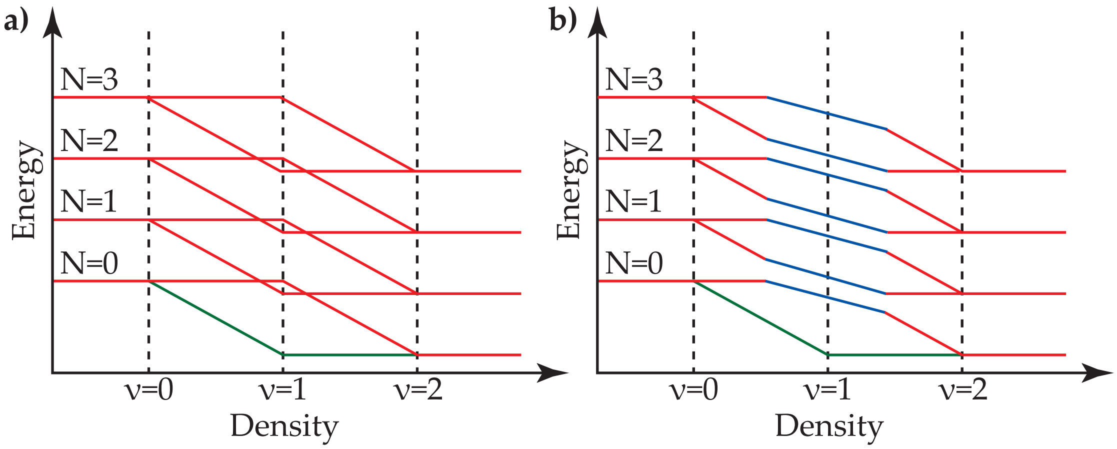


Figure 6-6: **Special role of the $N = 0$ Landau level: a)** neglecting Landau level mixing, the exchange interaction causes the energies of Landau levels with different orbital indices to cross near odd-integer filling. Energies shown neglect the capacitive energy of charging the well, as per the jellium panels of Figure 6-3. **b)** While avoided crossings limit the indirect exchange splittings of levels with orbital indices $N > 0$, $N = 0$ has no lower lying level to mix with, and is able to grow without bound.

other in the spectrum can.[†]

## 6.4 Alternative Approaches to Level Mixing

Some success has been had in treating Landau level mixing at intermediate fields by replacing the usual Coulomb interaction with the screened interaction present in a zero field 2D electron gas[60, 68, 62]. Conceptually, this deals with the Coulomb interaction first and then approximately adds Landau quantization on top. This approach neglects the orbital energy required to introduce a charge density change in a quantum Hall system relative to the Coulomb interaction strength, an approximation appropriate in the low-field limit. We choose to speak in terms of mixing Landau levels instead both because of the intuitive insight it can give, especially in explaining the limiting the splittings to the order of $\hbar\omega_c$, and because in those terms we are able to discuss special cases such as the $N = 0$

---

[†] The spin down state of the $N = 0$ level also has a fairly narrow line-width and high spectral weight; this is suggestive that the asymmetry of the peaks in the spin-split Landau levels may be related to mixing between Landau levels

Landau level.

The framework for including Landau level mixing has been laid out and used in the calculation of thermodynamic and electrical properties of two-dimensional electron gases[61, 63]. However, there are no published calculations of the detailed high energy spectra of an interacting 2DEG, nor are there analytic calculations of splittings in this intermediate field range.

## 6.5 Summary

Spin-polarized Landau levels at the Fermi surface induce exchange splittings in empty orbital states above the Fermi surface and filled orbital states below the Fermi surface. These "indirect" exchange splittings have been observed, for the first time, using time domain capacitance spectroscopy. The spectrum of these highly excited states reflects correlations formed in the many-body ground state of the electron liquid. With the ability to measure this high energy spectrum comes a unique new probe of the many-body ground state.

We develop an understanding of the general shape and magnitude of the indirect exchange splitting, and provide a qualitative interpretation of the special role of $N = 0$ in the spectrum. A quantitative measurement of these splittings is hampered by an unexplained imbalance between the spectral weight in the two branches of the splitting; the state closest to the Fermi energy is sharper and carries more spectral weight than the state farther away.

# Chapter 7

# Quasiparticle Lifetimes

For the past three chapters, we have discussed quasiparticle energies in the integer quantum Hall effect; we have asked, and answered the question “at what energies can we insert and remove electrons from our system, creating quasi-electrons and quasi-holes?” However, that is only part of the information contained in our spectra; another important question is “given that we have injected a quasi-electron or quasi-hole, how long does it live in the electron gas?” These quasiparticles are not true eigenstates of the many-body system; their lifetimes are finite due to electron-electron interactions.

These finite lifetimes lead to observable broadenings of the line-widths in our spectra. There are also contributions to the broadening from the disorder present in our samples. Our task, then, is to extract the line-widths from our spectra, to separate the contributions from disorder and from the quasiparticle lifetime, and to understand the origins of both sources of broadening.

## 7.1 Lifetimes and Energy Broadening

Consider a state $|\psi\rangle$ which is “almost” an eigenstate of a system with energy $\hbar\omega_0$, such that it has a long lifetime $\Gamma^{-1} \gg \omega_0^{-1}$. That is, if $|\Psi(t)\rangle|_{t=0} = |\psi\rangle$ is a solution to Schrödinger's equation $\mathcal{H}\left|\Psi(t)\right\rangle = i\hbar\frac{\partial\left|\Psi(t)\right\rangle}{\partial t}$,

$$\langle\Psi(t)|\psi\rangle = e^{i\omega_0 t}e^{-\Gamma t}\ ,\ t > 0 \tag{7.1}$$

In calculating the tunneling rates, we need to know what the overlap M(E) of $|\Psi(t)\rangle$ is with a state $|\epsilon\rangle$ of definite energy E due to the contribution from $|\psi\rangle$;

$$\begin{aligned} \mathrm{M(E)} &= \int_0^\infty e^{i(\omega_0 - \mathrm{E}/\hbar)t}e^{-\Gamma t}\,\langle\epsilon|\psi\rangle && (7.2)\\ &\propto \frac{1}{\Gamma + i(\omega_0 - \mathrm{E}/\hbar)} && (7.3)\end{aligned}$$

Our differential tunnel current is proportional to $|\mathrm{M(E)}|^2$,

$$|\mathrm{M(E)}|^2 \propto \frac{1}{\Gamma^2 + (\omega_0 - \mathrm{E}/\hbar)^2} \tag{7.4}$$

Thus, if we consider the contribution to our spectra of a state of energy $\hbar\omega_0$ and lifetime $\Gamma$, we expect to see a Lorentzian peak in our spectrum with a half-width half-maximum (HWHM) of $\Gamma$. Note that our definition of $\Gamma$ in terms of the amplitude decay rate, rather than the probability decay rate, is consistent with that typically used in many-body theories, and the $\Gamma$ shown in Figure 2-1.

## 7.2 Electron-Electron Interactions and Finite Lifetimes

Figure 7-1 shows a TDCS spectrum taken at low field. If we look at the zero and low density region of the spectrum, we can see clear Landau levels all the way from the Fermi energy to 15 meV (dashed red box in Figure 7-1b). However, as we begin to raise the density in the quantum well, we see the high energy levels broaden and disappear near $\nu \sim 0.5$ (right edge of red box). As we continue to raise the density, these high energy levels reappear as they approach the Fermi surface (dashed blue box in Figure 7-1b). The overall high DOS in the

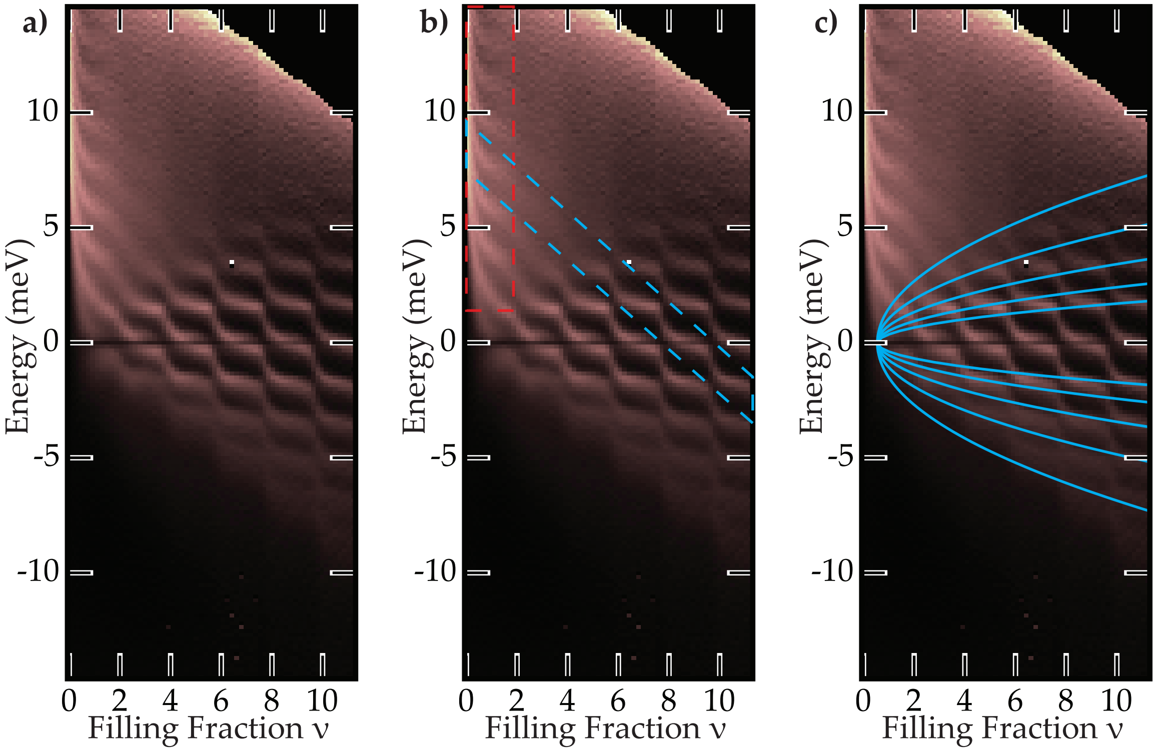


Figure 7-1: **Line-widths and lifetimes at 1 Tesla: a)** TDCS Spectrum taken at 1 Tesla, provided for reference. **b)** At zero density and low density, clear Landau levels are visible from the Fermi surface all the way up to 15 meV (dashed red box). However, as the density is raised, the high energy levels disappear. They become visible again as they approach the Fermi surface (dashed blue box). **c)** Imaginary contours of constant visibility for the Landau levels are roughly parabolic in shape.

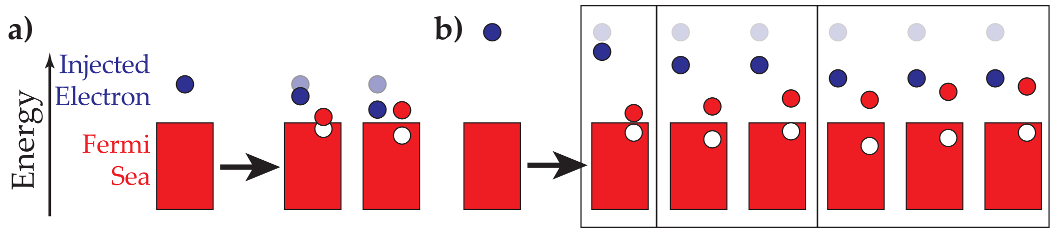


Figure 7-2: **Lifetimes of quasiparticles near the Fermi surface: a)** An injected electron can decay by scattering off electrons already in the Fermi sea. **b)** As the energy of the injected electron increases, the number of different ways it can excite electron hole pairs increases, decreasing the electron's lifetime. Both the energy lost by the injected electron (different black boxes), and the distance from the Fermi surface of the scattered electron (different cartoons within the same black box) can vary.

region where the Landau levels have disappeared shows they disappear due to broadening rather than a reduction of their amplitude; this is borne out by an examination of the high energy Landau levels at higher magnetic fields, which show a large spectral weight but increased broadening.

The strong density dependence of this broadening indicates that it originates with many-body effects. This leads us to a well-known result from Landau's Fermi-Liquid theory[14, 15, 69] that the lifetime of quasiparticles near the Fermi surface is a strong function of their energy from the Fermi surface. Conceptually, injected quasiparticles can decay by scattering off electrons in the Fermi sea. As we raise the energy of the injected electron, we increase the number of possible electron hole pairs we can excite (Figure 7-2). This increase in the number of potential decay channels decreases the electron lifetime.

This energy dependence can be quantitatively understood by examining the phase space available for electron-electron scattering[70]. At zero temperature, consider an injected electron (or more strictly speaking, quasiparticle) an energy $\epsilon$ away from the Fermi surface, where $\epsilon$ is small enough that the single-particle density of states at $\epsilon$ can be approximated with that at the Fermi energy. The electron can decay by scattering off an electron already in the Fermi sea, creating an electron-hole pair; see Figure 7-3. When the injected electron scatters, it will lose an amount of energy $\Delta$. Because of the Pauli exclusion principle, the electron

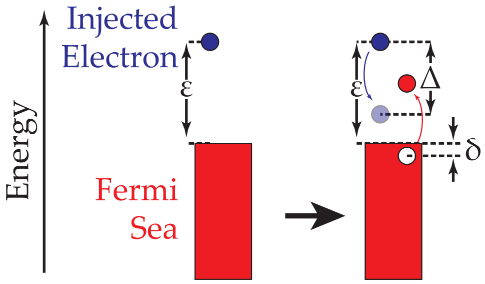


Figure 7-3: **Phase space for scattering near the Fermi surface:** An electron injected with energy $\epsilon$ can scatter, losing any amount of energy $\Delta < \epsilon$. The generated electron-hole pair will also have energy $\Delta$. Because the electron must be above the Fermi surface and the hole must be below, the hole must be generated from an energy $0 \leq \delta \leq \Delta$.

must scatter to an energy above the Fermi surface, so $0 \leq \Delta \leq \epsilon$. Similarly, the electron scattered out of the Fermi sea must be scattered into a state above the Fermi surface; it can be drawn from any energy $0 \leq \delta \leq \Delta$ beneath the Fermi surface. Thus, neglecting momentum conservation, the phase space $\Omega$ available for scattering is given by

$$\begin{aligned} \Omega &\propto \int_0^\epsilon \int_0^\Delta d\delta d\Delta && (7.5) \\ &\propto \int_0^\epsilon \Delta d\Delta && (7.6) \\ &\propto \epsilon^2 && (7.7) \end{aligned}$$

If $\epsilon$ is small enough that the variation of the scattering matrix element is negligible, the lifetime $\tau$ is simply proportional to $\Omega^{-1}$. The only remaining energy scale is $\mathrm{E_F}$, so, using dimensional analysis,

$$\frac{\hbar}{\tau} \propto \frac{\epsilon^2}{\mathrm{E_F}} = \mathrm{E_F} \left( \frac{\epsilon}{\mathrm{E_F}} \right)^2 \qquad (7.8)$$

Detailed examinations[9] confirm this result in the limit of small $\epsilon$ but rely on momentum conservation; with care, it can be derived in the absence of momentum conservation[71, Section 5.2].

At low magnetic fields, the Fermi energy is roughly proportional to the charge density, or equivalently the filling fraction $\nu$. With this approximation, curves of constant lifetime are curves of constant $\epsilon^2\nu^{-1}$, or $\nu \propto \epsilon^2$; these are parabolas oriented along the x-axis of our plots. In Figure 7-1**c** we superimpose such parabolas on our 1 Tesla data. On qualitative examination, these parabolas do seem to describe contours of constant Landau level visibility, as we would expect if electron-electron scattering were the origin of our disappearing Landau levels.

Of course, at zero energy our Landau levels have finite line-widths. Fermi liquid theory tells us that lifetimes diverge near zero energy, indicating there is a second source of broadening in our spectrum.

## 7.3 Disorder and Screening

This second source of level broadening in our structures is disorder. Charged impurities can locally modify the electrostatic potential in the quantum well, broadening the Landau levels[72]. Fluctuations in the width of the quantum well and variations in the Aluminum concentration in our $\mathrm{Al_xGa_{1-x}As}$ barriers also lead to variation in the confinement energy in our quantum well. In either case, the energy of the bottom of the 2D band will vary as a function of position. For simplicity of language, we can combine these effects into an effective electrostatic potential.

If the thermodynamic density of states in the quantum well is large, the local electron density in the quantum well can change in response to long length-scale disorder. These density fluctuations will tend to reduce the magnitude of the disorder potential (Figure 7-4a). If the density of states is infinite, the 2DEG is able to screen the disorder perfectly and the potential is constant across the entire 2D plane; this is the familiar case of a metal sheet placed in a non-uniform electrostatic potential.

As we sweep the density in the quantum well, the density of states fluctuates. When the chemical potential falls inside of a Landau level, we can locally add

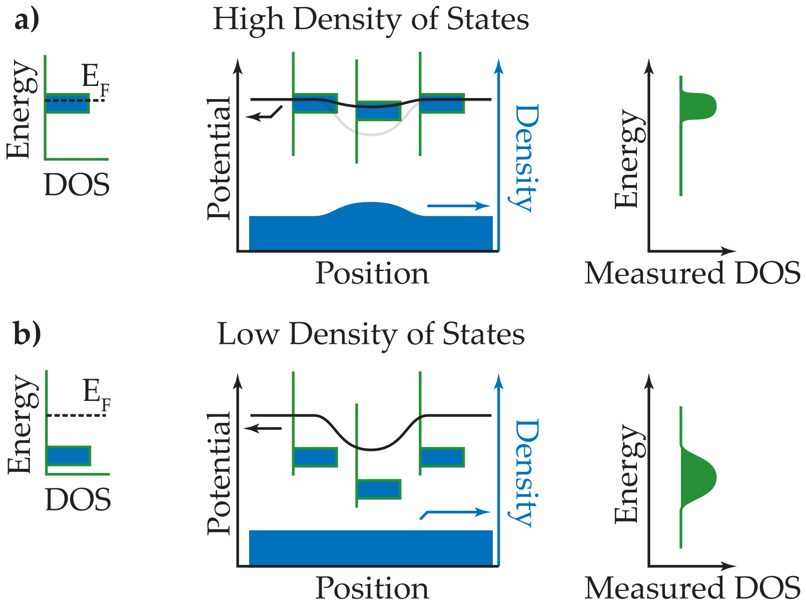


Figure 7-4: **Screening of disorder by the 2DEG: a)** When the density of states at the Fermi surface is large, potential fluctuations induce changes in the carrier density, which reduce the magnitude of the potential variation. The observed DOS spectrum shows little broadening. **b)** When the density of states at the Fermi surface is small, the 2DEG is unable to screen potential fluctuations. The observed DOS spectrum is heavily broadened.

electrons to the quantum well to screen disorder potentials. We expect in these regions to observe narrower line-widths due to disorder. On the other hand, when the chemical potential is between two Landau levels, it would take a potential fluctuation on the order of $\hbar\omega_c$ to locally add electrons. Screening of disorder by the quantum well will be negligible, and the observed line-widths will reflect the unscreened disorder in the quantum well (Figure 7-4b).

## 7.4 Extracting Low-Field Lifetimes

Starting with the 1 Tesla spectrum, we can fit the position and width of each distinguishable Landau level; see Figure 7-5a. We are left with an irregularly spaced mesh of line-widths that are difficult to visualize or compare with theory. We simplify the presentation of this data by fitting a fifth order polynomial through each column of our data, and using these fits to interpolate a continuous scalar field of broadenings. We then plot contours of constant broadening using this field (Figure 7-5b).

If we look along the Fermi surface at $\mathrm{E} = 0$, we see the broadening oscillates. The levels are very broad at even-integer $\nu$, and narrower elsewhere. This is indicative of disorder broadening, as discussed above. At even-integer $\nu$, the 2DEG is incompressible and unable to screen disorder. At odd-integer $\nu$, the density of states is high (the exchange splitting is smaller than the level width at these fields), and disorder is screened effectively.

If we want to find regions where lifetime broadening dominates our spectra, we should go to odd-integer $\nu$ where the disorder broadening is minimized, and move out to higher energies where we expect shorter lifetimes.

Detailed calculations of the lifetime due to electron-electron scattering in a 2D system at zero temperature have been performed by Chaplik[73, 74];

$$\frac{\hbar}{\tau_{ee}}(\mathrm{E}) \simeq \frac{E_f}{4\pi}\left(\frac{\mathrm{E}}{E_f}\right)^2\left[\ln\left(\frac{E_f}{\mathrm{E}}\right)+\frac{1}{2}+\ln\left(\frac{2q_{TF}}{k_f}\right)\right] \tag{7.9}$$

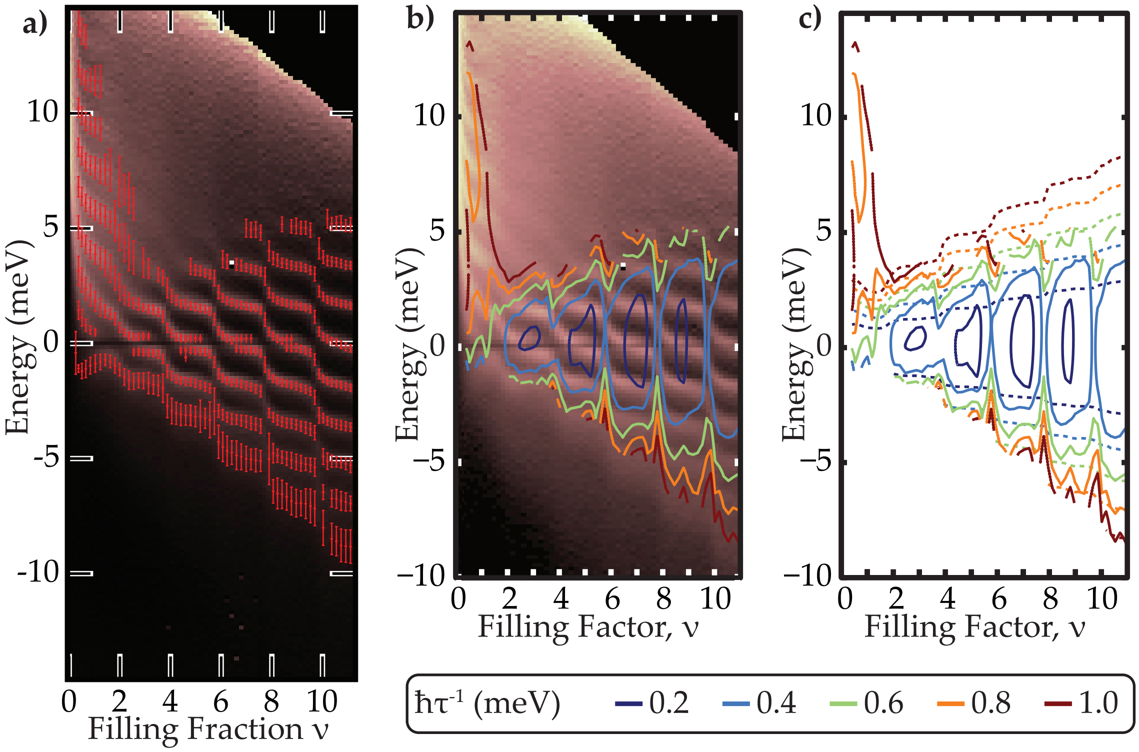


Figure 7-5: **Fitting lifetimes at 1 Tesla: a)** 1 Tesla TDCS spectrum with fitting results superimposed. Each marked point corresponds to the center of a Landau level, while the error-bars denote the FWHM of the fit. In **b**, we interpolate this irregularly spaced grid vertically across our data using 5th order polynomials, and plot contours of constant line-width. **c)** Comparison of extrapolated contours of constant line-width to Equation 7.9 (dashed lines).

where $q_{TF}$ is the Thomas-Fermi screening wave vector $2me^2\hbar^{-2}$. This is a zero field result that is only valid near the Fermi energy. We expect that at adequately small fields this result will hold, and the half-widths of our levels will be given by $\hbar/\tau_{ee} + \Gamma_0$, where $\Gamma_0$ is the broadening due to disorder. We approximate the Fermi energy by the energy of the most tightly bound Landau level, and the quasiparticle mass by the bulk electron mass in GaAs, $m^* = 0.067m_0$; this leaves no adjustable parameters in determining $\tau_{ee}$. In Figure 7-5**c**, we plot this result against our contours of constant line-width. At high energies and odd-integer $\nu$, where we expect lifetime broadening to dominate our disorder-induced line-widths $\Gamma_0$, we find excellent quantitative agreement between our measured widths and this first principles result. The useful range of comparison covers the range of $r_s = \frac{q_{TF}}{\sqrt{2}k_f}$ from around 1.5 to 4.5, testing the logarithmic dependence on $r_s$.

Perhaps the most remarkable feature of this measurement is that we find this agreement holds out to energies comparable to the Fermi energy, well outside the expected range of validity of this result. This measurement shows experimentally that even in the presence of strong Coulomb interactions, quasiparticles with well defined energies and long lifetimes exist over the full range of energies from $0 \lesssim \mathrm{E} \lesssim 2\mathrm{E}_F$.

The lifetime due to electron-electron scattering has been probed experimentally before. In transition metals, it contributes a region of $\mathrm{T}^2$ temperature dependence of the resistivity [70]. In two-dimensional systems, the phase decoherence length has been measured as a function of energy using an electron interferometer[75]. Unlike our measurement, this was done at zero field, allowing more precise comparisons with Equation 7.9. These measurements of the phase decoherence length cover a single density and a narrow range of energies (up to approximately 25% of the Fermi energy). Zero-field double well tunneling experiments have been used to probe the temperature dependence of the electron-electron scattering lifetime[76], which takes a similar form to that of the energy dependence[77]. These measurements are over a narrow range of energies and densities, and there is a large unexplained multiplicative factor between the results and theory. Our measurements represent the first available at high energies, and probe the quasiparticle lifetime directly, rather than the phase or momentum scattering time. They also provide the first tests of the logarithmic corrections to Equation 7.9.

## 7.5 Choice of Dataset

There is an unavoidable compromise in our choice of the 1 Tesla data set for measuring line-widths. Although we would like to measure lifetimes at zero field for easy comparison with currently available theories, the zero field spectrum is relatively featureless. We need to apply a magnetic field to discretize the spectrum, providing sharp features whose line-width we can measure. We need these features to be a distance comparable to the line-width apart in order to resolve them. As we increase the magnetic field, we are able to measure shorter lifetimes, and our fitting accuracy increases as our peaks become better separated. However, these increased fields move us further out of the domain of Equation 7.9, and we sample the lifetime at more widely spaced intervals as the energy spacing between the Landau levels increases. Finally, the emergence of direct and indirect exchange-enhanced spin gaps at higher magnetic fields makes it difficult to extract accurate widths; in regions where the gap is not wide enough to resolve it appears as extra broadening in our levels.

We have several datasets available to choose between to measure the electron-electron lifetimes, shown in Figure 7-6. At $^1/_2$ Tesla, we find that although we are able to see several Landau levels, their separation is much smaller than their line-width. The DOS does not come close to zero in between the Landau levels, even near the Fermi surface.

At $^3/_4$ Tesla, the dip between the Landau levels near the Fermi surface becomes pronounced, dropping to approximately 50% of the peak value, but we are still not able to resolve levels very far from the Fermi surface.

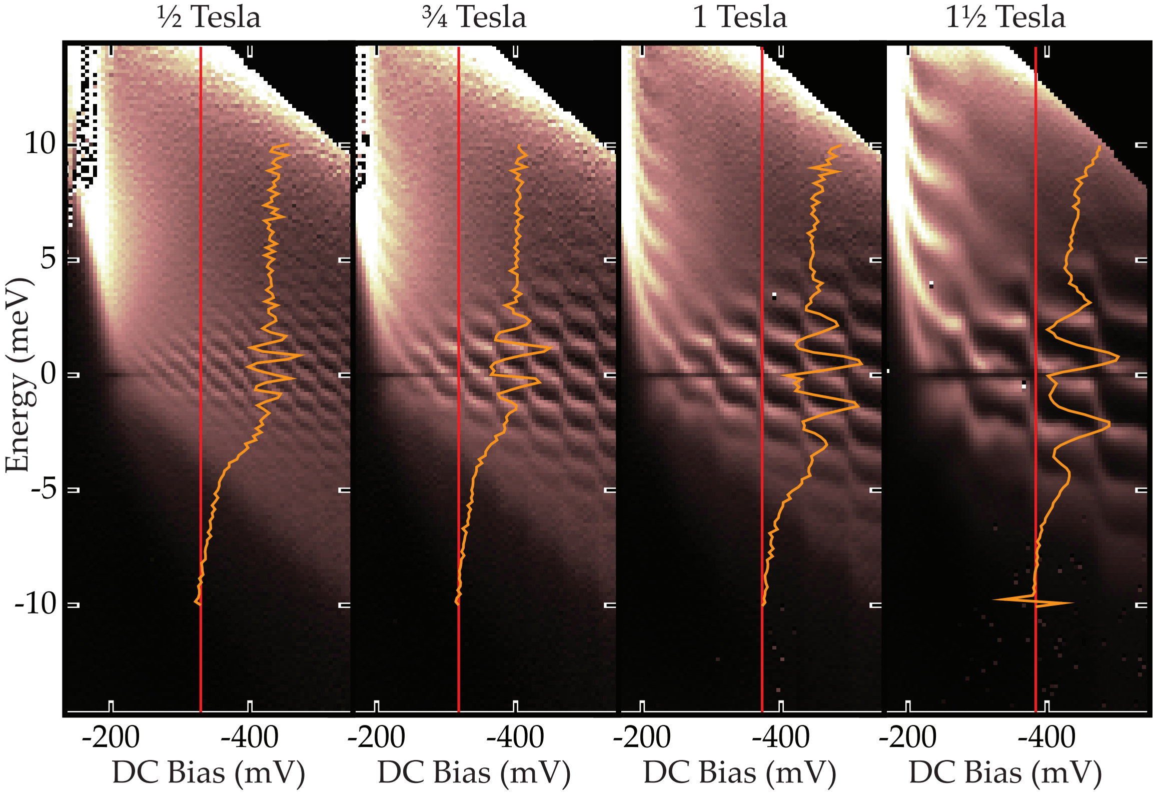


Figure 7-6: **Choice of spectrum for lifetime fits**: The 1 Tesla spectrum was selected as a compromise for fitting the quasiparticle lifetime. A low magnetic field, closely approximating zero-field theories, allows tightly spaced lifetime measurements in energy. Higher magnetic fields give good separation between adjacent peaks, allowing accurate fitting, but exchange splitting and the Coulomb gap can modify line-widths. Representative line cuts near even-integer filling are shown for each spectrum. The red line is the zero of the line cut, and is drawn at the point in the spectrum from which the line cut was taken.

By 1 Tesla, we are able to resolve Landau levels at a large range of energies around the Fermi surface, and the DOS is dropping to near zero in between the levels at the Fermi surface. Some exchange splitting is becoming visible near the Fermi surface, but we can separate the levels there with the aid of the Coulomb gap.

At $1\,{}^{1}/_{2}$ Tesla, the exchange splitting is becoming large enough to be a serious problem, with the indirect exchange splitting becoming noticeable in the excited states that are important for our lifetime fits.

On this basis, we choose to use the 1 Tesla data for our lifetime fits. In the future, with more data acquisition, we could use sweeps at fields near 1 Tesla to provide lifetime data in the gaps between Landau levels in our current spectra. It may also be possible to relate the peak-to-valley visibility of the Landau levels to their width in a simple and robust way, allowing us to use much lower field data for lifetime fitting. Any such approach would rely on the quantitative accuracy of our matrix element correction, which at this point would lead to systematic errors in the measured lifetimes.

## 7.6 Summary

Although the resolution of our spectra is only limited by temperature and acquisition time, the Landau levels we observe are substantially broadened. At energies near the Fermi surface, disorder dominates the broadening, and we see clear oscillations in the observed level width as the 2DEG transitions between compressible and incompressible regions. At higher energies, finite lifetimes due to electron-electron scattering dominate. We find good agreement with no adjustable parameters between our observed line-width and theoretical predictions, even substantially outside of the strict domain of applicability of those theories. Quasiparticles in the 2D electron gas remain a useful construct even at energies comparable to $E_F$ away from the Fermi surface.

When we move to very low density, there are no electrons off which the injected electrons can scatter. Here, we recover the physics of non-interacting electrons, and are able to resolve Landau levels at the highest energies we measure. As we transition to finite densities, we see Fermi liquid theory "turn on". This happens somewhat suddenly near $\nu \sim 1/2$. It is unclear why this occurs at a finite density rather than at zero density.

# Chapter 8

# New Features

Even with the current capabilities of TDCS, much remains to be explored in the 2DEG. There are several unidentified and tentatively identified features in the spectra that need to be examined with more care. There are also simple ways to extend the measurement that we expect to reveal new phenomena. We explore these opportunities here.

## 8.1 Plasmarons

As we discussed in the introduction, in 2D systems there is predicted to be a second type of quasiparticle, a plasmaron, much more tightly bound than a normal quasi-hole[12]. To our knowledge, this quasiparticle has not been directly observed despite extensive efforts to identify it in X-ray spectra[80, 81]. Efforts to identify plasmarons in 3D may have been in vain. There are suggestions that in 3D the plasmarons result from approximations used in solving Dyson's equation[82, 83], and in reality the excitations are too heavily damped to observe. However, these excitations are expected to be real in 2D systems[11]. Furthermore, they are expected to contribute a distinctive peak or sharp shoulder of substantial spectral weight[84, 79], suggesting they should be easily visible in tunneling experiments in 2DEGs. There is, however, substantial disagreement as to the exact structure in the DOS[12, 11, 84, 79, 78].

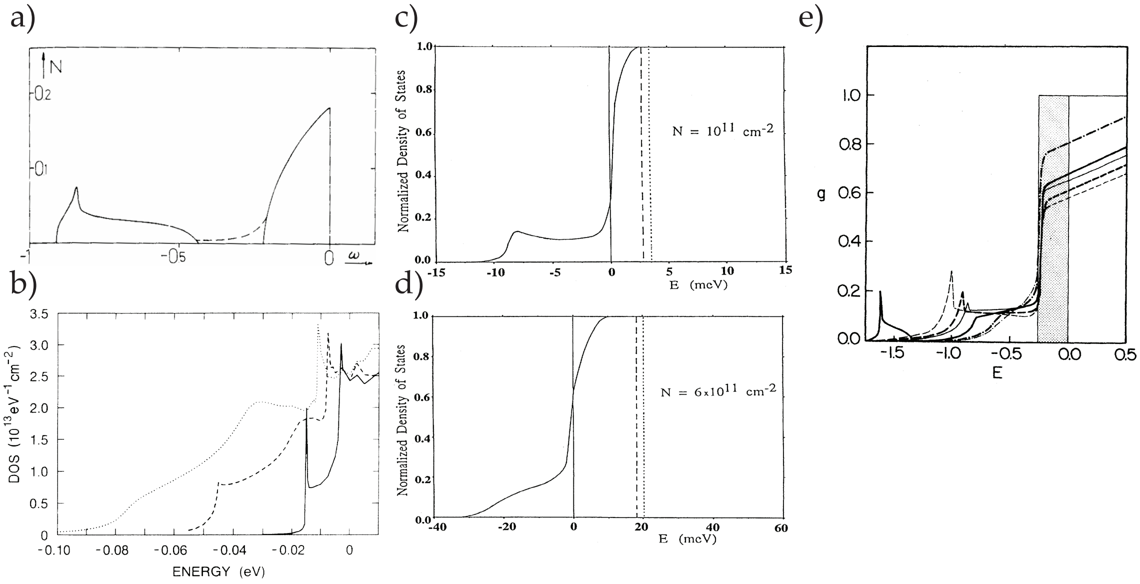


Figure 8-1: **Calculated DOS** at zero magnetic field, in different approximations. **a)** Hedin [12], 3D electron gas diagram. The peak/shoulder structure below zero energy originates with the plasmaron modes. **b)** von Allmen[78], 2D electron gas, $n = 2.2,\ 7.1,\ 12 \times 10^{11}$ cm$^{-2}$ moving from right to left. **c)** Guillemot and Clérot[11], 2D electron gas, $n = 10^{11}$ cm$^{-2}$. The Fermi energy is given by the vertical dashed line. **d)** Guillemot and Clérot[11], 2D electron gas, $n = 6 \times 10^{11}$ cm$^{-2}$. The Fermi energy is given by the vertical dashed line. **e)** Jalabert and das Sarma [79], 2D electron gas, $n = 10^{11}$ cm$^{-2}$ for the dashed lines, $2 \times 10^{11}$ cm$^{-2}$ for the solid, and $10^{12}$ cm$^{-2}$ for the dash-dotted lines. Bold lines include phonon coupling. Energy is in units of four times the Fermi energy, with the shaded region corresponding to filled states for non-interacting electrons. The same basic shoulder structure is present in all the spectra, at an energy a few times the Fermi energy.

However, the basic prediction is consistent. There is expected to be a sharp shoulder or peak in the DOS substantially below the most tightly bound conventional quasiparticle, owing to simultaneous generation of a strongly bound plasmon and quasi-hole (Figure 8-1). Even at zero field, these quasiparticles appear as a peak in the DOS as they have a peaked contribution at long $k$-vectors. This band occurs at an energy roughly proportional to $E_F$, and is a few times $E_F$ below the Fermi surface.

Examining our 0 T data carefully (Figure 8-2), we see a shoulder of states below the 2D band bottom that seems consistent with these predictions, and tentatively identify it as due to plasmarons. This shoulder becomes peak-like at zero density, but is never well resolved from the tail of the main band. We can emphasize this feature by examining the derivative of our data, d(DOS)/dE, which will show a peak each time there is a step in the data. We encounter the same difficulties as we have in taking dI/dV. In this case, we overcome them by low pass filtering the data on a column-by-column basis, convolving it with a $\sigma = 500\ \mu$eV Gaussian, although this solution is less than ideal.

At higher magnetic fields, this shoulder sharpens into a well defined peak roughly $\hbar\omega_c$ below the lowest Landau level. Although there are no predictions we are aware of for the energy spectrum of magnetic plasmarons, this seems consistent with the zero field result. At high fields, the magnetoplasmon is broken into bands associated with excitations between Landau levels with different orbital indices. In particular, at long wavelengths the plasmon energies tend towards $\hbar\omega_c$ as predicted by Kohn's theorem[55, 85]. If we think of the magnetic plasmaron as an $N = 0$ quasiparticle bound to a long-wavelength magnetoplasmon, we expect it to be about $\hbar\omega_c$ below the $N = 0$ Landau level. The plasmarons appear and disappear as the filling fraction is varied, typically appearing strongest at even-integer filling. This may be because of suppression of the lower energy spin-flip plasmon modes at these filling fractions. These modes tend to the Zeeman energy at long wavelengths, and we would not resolve a plasmaron mode separated from the lowest Landau level by only Zeeman energy.

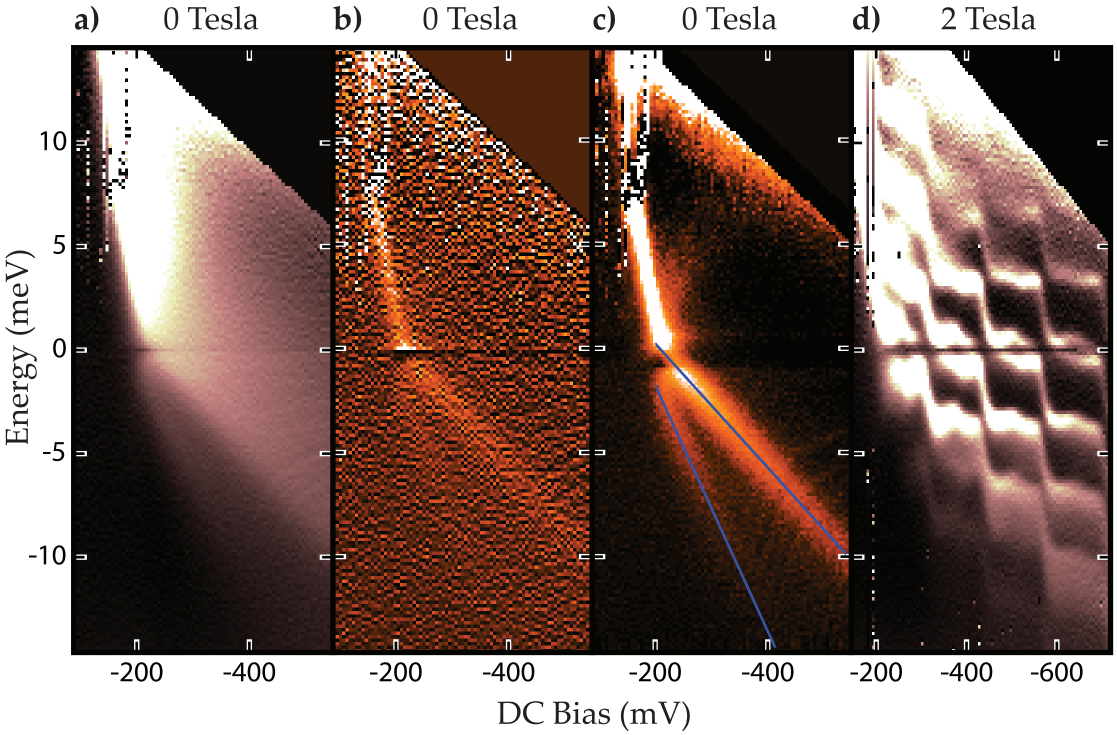


Figure 8-2: **Plasmarons** at zero field and 2 Tesla. **a)** Zero field TDCS spectra, with the contrast adjusted to show clearly the plasmaron band below the main quasiparticle band. Moving from the bottom of the figure towards zero energy, there are two steps in brightness, the first corresponding to the plasmarons, and the second to the conventional quasiparticles. **b)** $\frac{\partial \mathrm{DOS}}{\partial \mathrm{E}}$, where the steepest points in the two steps of **a** appear as peaks. The numeric differentiation introduces high frequency noise. **c)** $\frac{\partial \mathrm{DOS}}{\partial \mathrm{E}}$, with each column low-pass filtered in energy. The steps associated with the quasiparticles and the plasmarons become peaks. Blue lines illustrate the band edges. The plasmaron line is drawn at $-(2|\mathrm{E}_{\mathrm{quasiparticle}}| + 2)meV$, where $\mathrm{E}_{\mathrm{quasiparticle}}$ is the normal quasiparticle energy. Some quantization artifacts from the measurement are visible in the bottom-right corner. **d)** Plasmarons at 2 Tesla. The plasmarons are visible as an additional faint line approximately $\hbar\omega_c$ below the lowest lying Landau level.

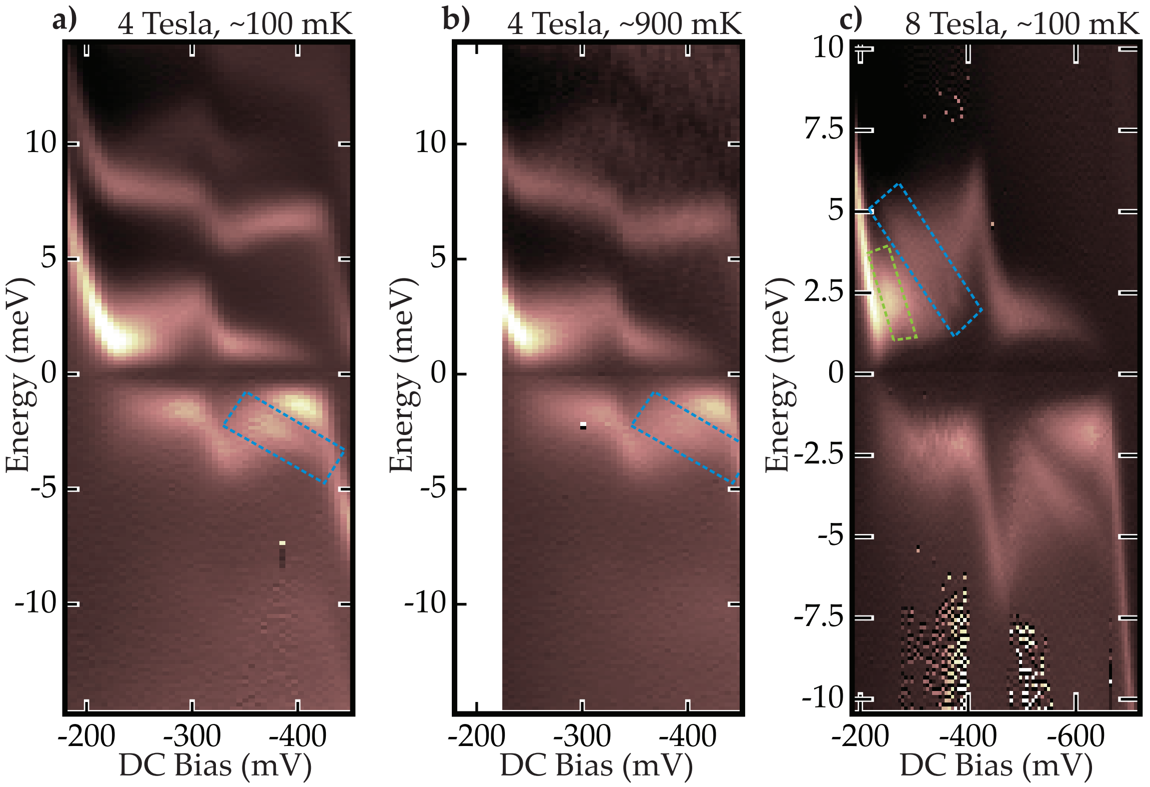


Figure 8-3: **Fine structure near $\nu = 1$: a)** A sash (dashed blue box) appears in the high field data extending down and to the right from $\nu = 1$ (-300 mV DC bias) at 100 mK. **b)** Although heavily broadened, the sash is still visible at 900 mK (dashed blue box). **c)** At 8 Tesla, the sash becomes stronger, and several similar lines (dashed blue and green boxes) appear extending upward and to the left of $\nu = 1$. Systematic measurement artifacts due to inaccuracies in the discharge pulse are visible on the hole side of $0 \leq \nu \leq 1$.

Positive identification of these bands as the plasmaron would be aided by lower noise data at zero field that could be compared more precisely to theory, lower disorder samples which might tend to sharpen the shoulder into a peak, and more extensive theoretical work on plasmarons in high magnetic fields.

## 8.2 Fine Structure Near Filling Fraction 1

As shown in Figure 8-3a, in our higher field data, extra structure begins to appear near $\nu = 1$ (-300 mV DC bias at 4 Tesla). We observe both extreme broadening of the spin-up and spin-down peaks in the vicinity of $\nu = 1$, and the appearance of an extra sash (dashed blue box) of states extending approximately linearly from roughly E $= 0$ at $\nu = 1$ downward in energy towards the center of the cyclotron gap at $\nu = 2$ (-450 mV DC bias at 4 Tesla). This feature becomes stronger at higher magnetic fields, and survives but is heavily broadened at temperatures as large as 900 mK (Figure 8-3b).

At the highest fields we measure at a second sash appears extending upward in energy from $\nu = 1$ (dashed blue box, Figure 8-3c), with approximately the same slope, towards $\nu = 0$. There may be additional lines beginning to appear at the highest fields (dashed green box). We never observe this sash around any filling fraction aside from $\nu = 1$, but it may simply not appear at higher filling fractions except at fields and densities we cannot attain in this structure.

The location of the line in the vicinity of $\nu = 1$ is suggestive of skyrmions[86], long wavelength spin waves with properties controlled by the ratio of the Zeeman to exchange energies predicted to be the low-lying excitations near $\nu = 1$ at low magnetic fields. We find the structure survives at temperatures as high as 900 mK. Prior measurements of tunnel rate structure because of the spin bottleneck associated with skyrmions found the bottleneck disappeared at around 400 mK at similar fields[27]. It is not clear if these temperature scales are directly comparable.

Application of tilted fields should indicate if these sashes are associated with skyrmions; by effectively varying the Zeeman energy relative to the exchange, tilted fields allow shifting of the skyrmion energy relative to a simple spin flip.

## 8.3 Effective Mass Renormalization

As discussed in the introduction, quasiparticles can have effective masses that are a function of both the electron density and the quasiparticle energy. In the context of our Landau spectra, these mass variations can appear as variation of the cyclotron energy at even-integer filling. While thermal activation measurements have provided accurate mass measurements at E $= 0$[87], we believe TDCS is

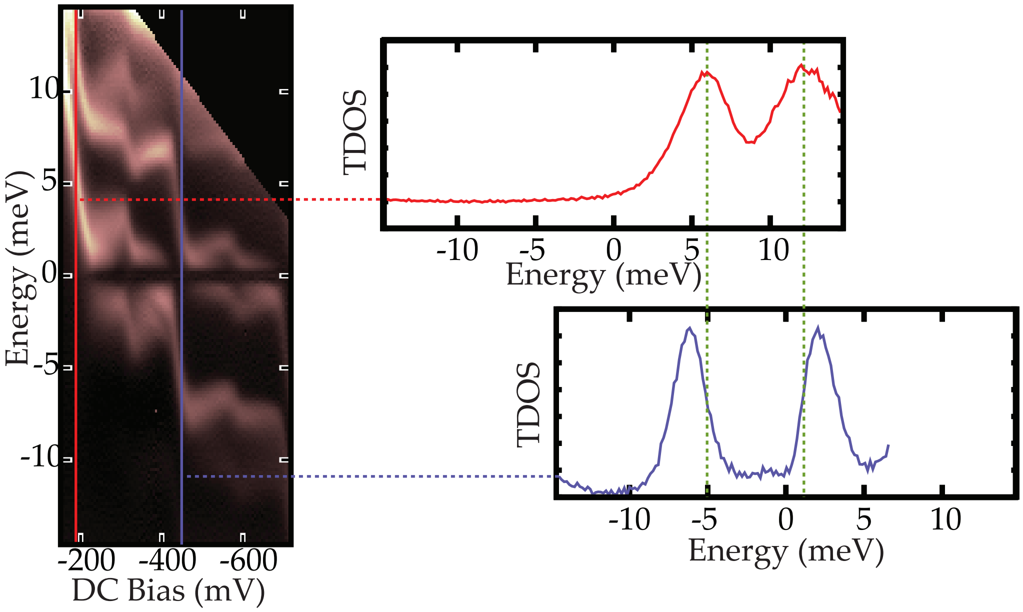


Figure 8-4: Cyclotron energy at $\nu = 2$ and $\nu = 0$ at 4 Tesla applied field, showing the enhancement of the cyclotron splitting by the exchange interaction.

unique in its ability to measure mass away from the Fermi energy.

Indeed, some evidence of this mass variation is already present at higher magnetic fields; the orbital gap at $\nu = 2$ in the 4 Tesla data is nearly 30% larger than the cyclotron energy at zero density (Figure 8-4). This can be understood either in terms of effective mass renormalization, or in terms of the exchange interaction; as a result of the direct and indirect exchange interaction between $\nu = 0$ and $\nu = 2$, both the $N = 0$ and $N = 1$ orbital levels are pulled down in energy. However, the $N = 0$ level is pulled down in energy substantially more than the $N = 1$ due to larger exchange matrix elements. This results in a gap larger than $\hbar\omega_c$ between $N = 0$ and $N = 1$ at $\nu = 2$.

However, this effect is much smaller at the low magnetic fields that would ideally be used to approximate the zero field mass renormalization. Systematic studies require higher resolution, lower noise spectra. Achieving this will require moderate improvements in our signal generation apparatus, especially in long term drift of the pulse height. A new generation of DC voltage source is in preparation to replace the one currently used to select the pulse heights. Thermal drift in the pulse shaper may also have to be addressed (see Appendix B).

## 8.4 Future Prospects: Fractional Quasiparticles

Tunneling measurements of excitations in a fractional Hall liquid have the potential to provide direct and intuitive demonstrations of hierarchy. There are predicted to be strict selection rules as to what fractionally charged states can be excited using a normal electron tunneled out of a metallic electrode[88]. These selection rules potentially provide a second channel of information in tunneling spectra; there will be information on both quasiparticle energies, and on their symmetries.

Most fractional excitations are expected to fall inside of the Coulomb gap. They may still be visible but strongly suppressed, or may be pushed outside of the gap altogether. Regardless, there may also be high energy states that reflect the fractional quantum Hall ground state. Exploring this possibility suggests two directions; moving to higher magnetic fields and lower temperatures to enter the domain of the fractional quantum Hall effect, and moving to samples with thinner tunneling barriers to ease looking for substructure in the Coulomb gap.[†] We are well positioned for both of these changes in the near future.

## 8.5 Future Prospects: Measuring the Chemical Potential

The chemical potential of a quantum well is a difficult number to measure experimentally. Electric field transparency measurements[66] and single electron transistors[89] provide direct measurements, but the structures required are not compatible with TDCS. However, accurate measurements of the chemical

---

[†] Such substructure has been observed in the vicinity of the fractional quantum Hall effect in equilibrium tunneling; see [27]. We have not yet replicated these features in non-equilibrium tunneling.

potential would ease comparison with theoretical predictions of the DOS and help understand the shape of our spectra in regions that show negative compressibility. Chemical potential measurements may also help us understand the sash near $\nu = 1$ by allowing us to measure its dispersion in absolute energy rather than relative to the Fermi energy. Finally, having measurements of both the chemical potential jump in the thermodynamic density of states and the single particle exchange-enhanced spin gap in the same sample under the same conditions would help understand the differences between these two measurements of the exchange splitting.

At zero density, the energy difference between the Fermi energy and the lowest energy state we can tunnel into is a direct measure of the chemical potential. However, as we raise the density, electron-electron interactions will tend to modify the energy of this lowest energy state, and we can no longer use it as a direct measure of the chemical potential (although it is still a useful rough estimate). Ideally, we would measure the chemical potential relative to an imaginary bound state energy for an uncharged electron in our well. Unfortunately, uncharged electrons are difficult to acquire experimentally.

There is a set of states that we expect electron-electron interactions to affect less; the second sub-band of the quantum well. For a square quantum well, the second sub-band wave functions in the z direction are odd under inversion, while those of the first sub-band are even. Because of this, the exchange matrix elements between states in the first sub-band and those in the second will vanish. This inversion symmetry will be broken once we apply DC biases to the well, but we still expect that effects of interactions between the first and second sub-bands will be small. Thus, one could imagine tracking the energy of the lowest state in the second sub-band relative to the Fermi energy, and using it trace out chemical potential changes in the quantum well.

Unfortunately, this faces some technological difficulties. The second sub-band is at a very high energy in our structure, and reaching it requires modifying our setup to generate large pulses. We can do so at the expense of increased reflections

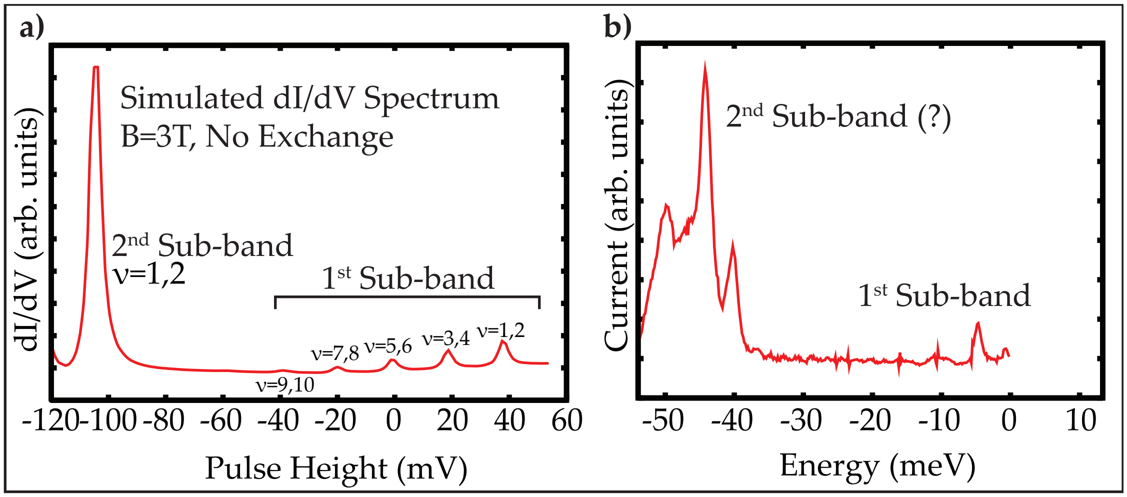


Figure 8-5: **DOS of the second sub-band** calculated (**a**) and measured (**b**). The origin of the multiple peaks in the experimental trace is unclear. Due to the large currents and pulses used, exact energies should not be trusted in the experimental trace. The lever arm in the simulated structure is $\lambda = 0.267$, so pulse heights in the simulated structure can be converted into energies in the measurement by dividing by $3.7 \approx 4$.

due to removed attenuators and increased quantization noise on our pulse height control. These introduce large systematic errors into our measurement. In addition, because of the enormous bias on the structure when we reach the second sub-band and the large penetration of the bound state wavefunction into the barriers, we expect the tunnel currents in the second subband to be approximately an order of magnitude larger than the currents in the first sub-band (Figure 8-5a). They apparently are (Figure 8-5b).† These currents require shorter measurement intervals to avoid systematic distortions of our spectra.

Nonetheless, we have several individual traces that reach the second sub-band.

---

†Because $k_x$ and $k_y$ are zero once again at the bottom of the second sub-band, the tunneling matrix element is finite even at these large energies; by comparison, tunneling out of the first sub-band has stopped due to momentum conservation.

We find the second sub-band at a somewhat higher energy than our calculations and simulations predict, but because the bound state energy is quadratic in the inverse well width, small variations in the grown structure will give large variations in the sub-band splitting.

Somewhat more disturbing, we see a multiplet of peaks at approximately the same energy, rather than a single peak. Field and density dependence may help identify these peaks. Coupling to the LO phonon mode may give at least some of the peaks; the LO phonon energy ($\approx$ 36.8 meV[84]) and the second sub-band are at a similar energy in our current structure.

We hope to grow a new structure with a wider quantum well. This will increase energy errors because of lever arm changes by making our 2DEG easier to polarize, but will also allow us to reach this second sub-band more easily and characterize it with care.

## 8.6 Future Prospects: Other Materials

Three ingredients are needed to apply TDCS to a two-dimensional system; a uniform, thin tunnel barrier; a constant thickness, highly resistive blocking barrier; and two metallic or nearly metallic electrodes. With various forms of epitaxial growth, there is the potential to extend this technique to other systems of interest, allowing examination of monolayers of high $T_c$ material at different doping densities, or the bizarre structure of the Landau levels in quasi-relativistic materials such as graphene[90, 91]. The latter would be of special interest as direct confirmation of the unusual Landau level spacing in these materials, including the presence of an $E = 0$ Landau level and easy access to both band electron and band hole states.

## 8.7 Summary

TDCS has already provided spectra that provide unique insight into the integer quantum Hall effect in a stunning fashion. In the next few months, as we continue to raise the magnetic field, lower the temperature, and increase the energy resolution of our spectra, we expect to resolve some of the open questions in our current spectra, and begin to discern features associated with the fractional quantum Hall effect.

# Chapter 9

# Polarizability of a Two-Dimensional Electron Gas

> *"Be patient, for the world is broad and wide."*
>
> I call our world Flatland, not because we call it so, but to make its nature clearer to you, my happy readers, who are privileged to live in Space.
>
> *Flatland: A Romance of Many Dimensions*, Edwin A. Abbott, 1884

While it is a convenient approximation to speak of a two-dimensional electron gas, the reality is it lives in a three-dimensional sample. The approximations we make in speaking of it as two-dimensional are very good, but sometimes the three-dimensional nature rears its ugly head; most notably in the case of this experiment through the 2DEGs polarization in response to applied voltages. Our main goal in this section will be to develop simple approximate results for this polarizability, but we will also explore several exactly soluble problems to help understand the limitations of these results.

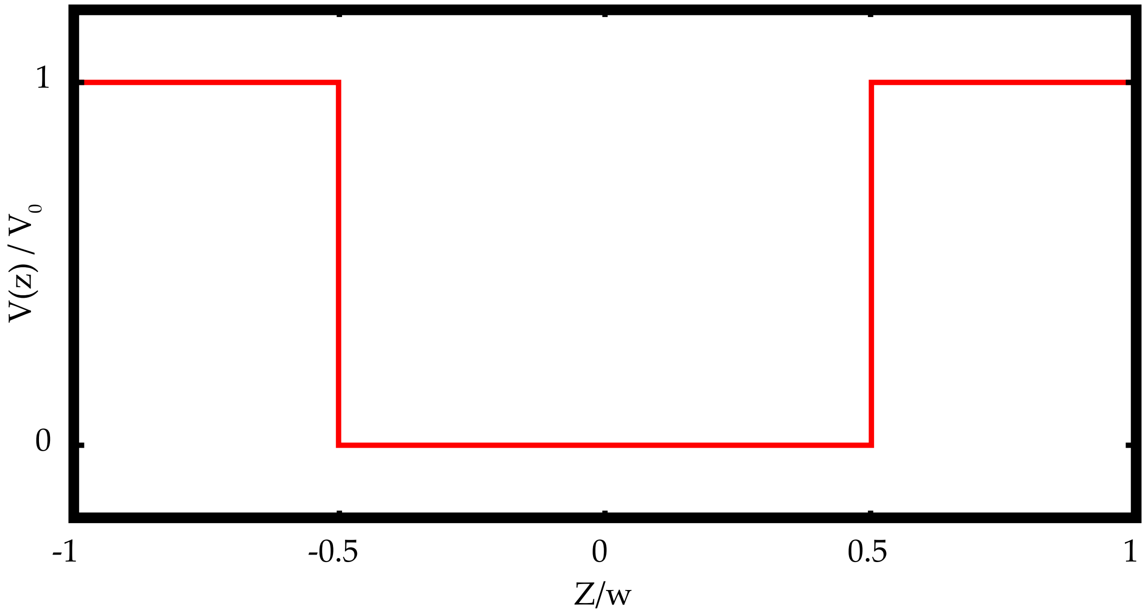


Figure 9-1: Confining potential of a simple quantum well

## 9.1 Three Dimensions into Two

We will start with a simplified model of a quantum well in which an electron is bound in a quantum well perpendicular to the $\hat{z}$ direction in a potential field

$$V(\vec{r}) = \begin{cases} V_0 & z < -w/2 \\ 0 \quad \text{if} & -w/2 \le z \le w/2 \\ V_0 & w/2 < z \end{cases} \tag{9.1}$$

where w is the width of the well (Figure 9-1).

$$\mathcal{H} = \frac{\vec{p}^2}{2m^*} + V(\vec{x}) \tag{9.2}$$

The $\hat{x}$ and $\hat{y}$ directions decouple, leaving us free to solve the $\hat{z}$ component of the problem separately.

In GaAs, for an x = .325 barrier, $V_0$ will be 251 meV. By comparison, typical Fermi energies will be a few meV; this suggests we can begin by approximating our quantum well by an infinite well. For the moment, we also neglect electro-

| Effect | Energy Scale |
|---:|:---|
| Quantum Well Barrier Height | 251 meV |
| Sub-band Energy Spacing | 55 meV (1st and 2nd sub-bands) |
| Applied Electric Field ($\lvert\vec{E}\rvert we$) | 45 meV (at $n = 3 \times 10^{11}$ cm$^{-2}$) |
| Fermi Energy | 6.2 meV (at $n = 3 \times 10^{11}$ cm$^{-2}$) |
| Coulomb Energy Scale ($\frac{e^2}{4\pi\epsilon l_b}$) | 4.4 meV (at B = 1 Tesla) |
| Cyclotron Energy | 1.7 meV (at B = 1 Tesla) |

Figure 9-2: Typical energy scales in our GaAs 2DEG

static self-consistency; we will not consider the effect of the electronic distribution on the potential in the well. We will revisit both these approximations later.

In this case, the problem reduces to a "particle in a box", and the eigenfunctions and eigenvalues of the 1D problem are given by

$$\phi_n(z) = \sqrt{\frac{2}{w}} \sin\left((z - w/2)k_n\right), \; k_n = n\pi/w \tag{9.3}$$

$$E_n = \frac{\hbar^2 k_n^2}{2m^*} \tag{9.4}$$

For a 17.5 nm thick quantum well, this gives sub-band energies of about 18 meV $\times n^2$, and the second sub-band is a full 55 meV in energy above the first. If this energy is very much larger than any other energy scale in our problem, mixing between these sub-bands will be small, and we can neglect motion in the $\hat{z}$ direction.

Figure 9-2 shows typical energy scales for our samples. While we find the Fermi, cyclotron, and Coulomb energies are safely small, the applied electric field is non-negligible.

The "Applied Electric Field" term is the electrostatic energy associated with the electric field we apply to the structure; it can be approximated as $V_{dc} \times w/W$, where $V_{dc}$ is the largest DC field we will use to gate the structure, and $W$ is the width of the entire tunnel capacitor. The finite $z$ extent of the well will allow it to polarize slightly in response to the electric field, both changing the lever arm $\lambda$ and adding an additional term to the chemical potential calculation. This gives important contributions both to the energy scale calibration (Chapter 12) and to the density calibration (Chapter 11). In addition, with a finite well this polarization will change the penetration of the wavefunction into the barriers, modifying the tunneling matrix element. We will have to consider this in the DOS calibration (Chapter 10).

Given that this polarizability of the quantum well results in corrections on all three axes of our spectra, it is essential that we examine it carefully.

## 9.2 Polarizability of an Infinite Well

With the addition of a potential term $V'(z) = e\mathrm{E}z$, we can represent our Hamiltonian inside the well as

$$\varepsilon\psi(z) = \left(\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2} + e\mathrm{E}(z)\right)\psi(z) \tag{9.5}$$

subject to the boundary condition that $\psi$ is zero at the well edges. Changing variables to the dimensionless

$$\begin{aligned} \tilde{z} &= \left(2m^* e\mathrm{E}\hbar^{-2}\right) z \equiv \alpha z & (9.6) \\ \tilde{\varepsilon} &= 2m^*\hbar^{-2}\alpha^{-2/3}\varepsilon & (9.7) \end{aligned}$$

our new differential equation is

$$\tilde{\varepsilon}\psi(\tilde{z}) = \ddot{\psi}(\tilde{z}) + (\tilde{z})\psi(\tilde{z}) \tag{9.8}$$

Which, with the final change in variables $\bar{z} = \tilde{\varepsilon} - \tilde{z}$, gives us the Airy equation

$$\bar{z}\psi(\bar{z}) = \ddot{\psi}(\bar{z}) \tag{9.9}$$

and affords the solution

$$\psi(\tilde{\varepsilon} - \tilde{z}) = c_1\mathrm{Ai}(\tilde{\varepsilon} - \tilde{z}) + c_2\mathrm{Bi}(\tilde{\varepsilon} - \tilde{z}) \tag{9.10}$$

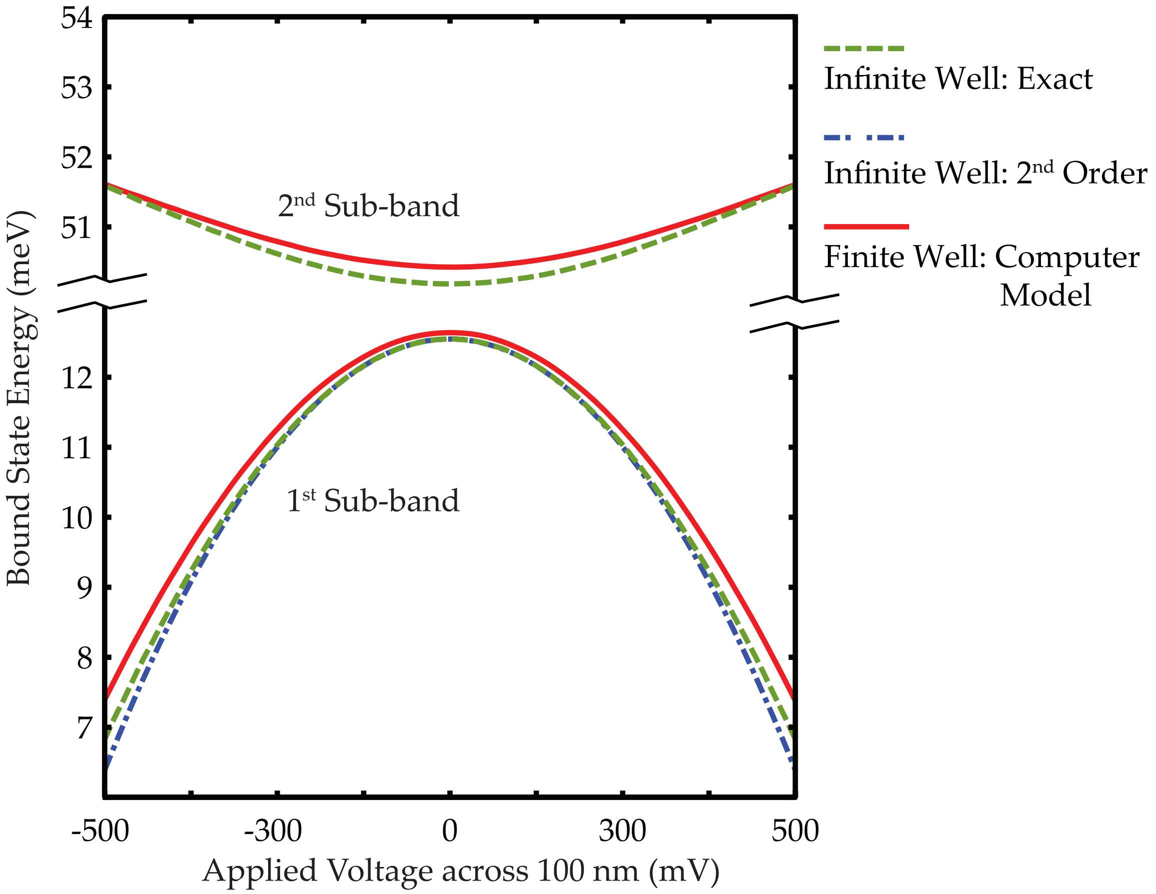


Figure 9-3: Exact eigenenergies of the first two bound states of an infinite square well as a function of applied electric field. Energies shown are for a 21.15 nm square well in GaAs. Also shown is Equation 9.12, the energy of the lowest bound state in 2nd order perturbation theory, and the energy of an approximately equivalent (see Section 9.3) 17.5 nm finite well with x = .325 AlGaAs barriers, calculated using a numeric model (Appendix C).

The boundary condition at $z = -w/2$ is trivially satisfied if we take $c_1 = \mathrm{Bi}(\tilde{\varepsilon} + \tilde{w}/2)$ and $c_2 = -\mathrm{Ai}(\tilde{\varepsilon} + \tilde{w}/2)$; we can now numerically find the zeroes of $\psi(\tilde{w}/2)$ as a function of $\tilde{\varepsilon}$ to extract the bound state energies. Note that the first order bound state energy shift is zero; we can confirm this by noting that $\langle \phi_1 | eEz | \phi_1 \rangle$ is zero; $\phi_1$ is symmetric about $z = 0$. Exact and approximate (see below) bound state energies are plotted in Figure 9-3.

In the electrostatic calculations that enter the density calibration, we also care

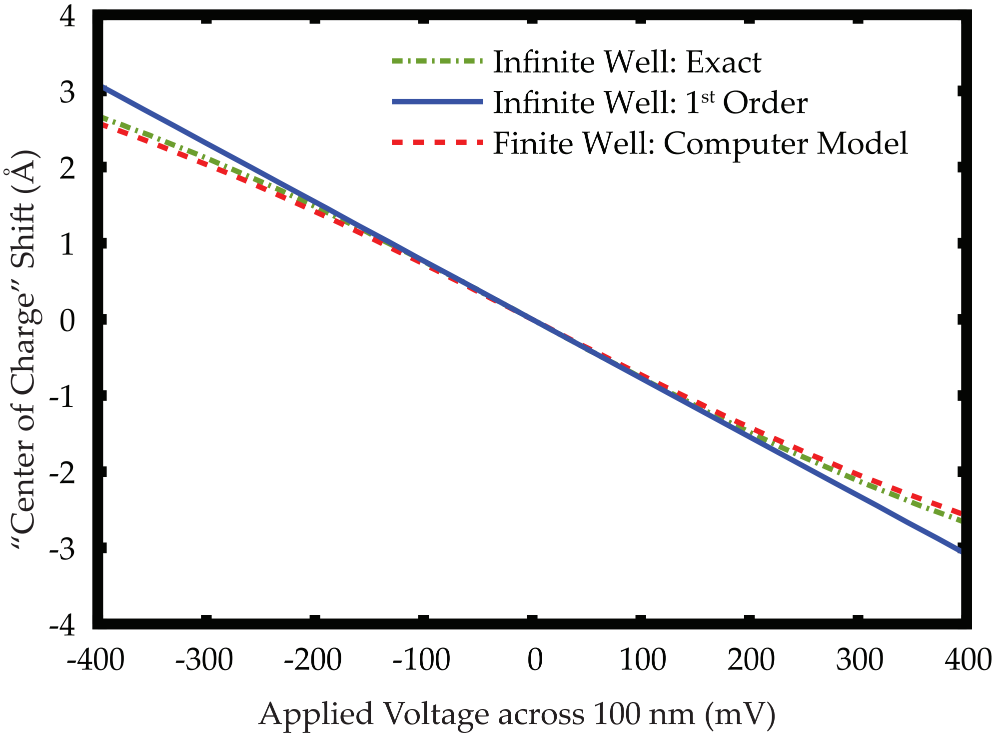


Figure 9-4: Exact polarization $\langle z \rangle$ of the first bound state of an infinite square well. Polarizations are calculated for a 21.15 nm infinite well, both exactly and in first order perturbation theory, and for an approximately equivalent 17.5 nm finite well with x = .325 barriers. The motion is approximately 7.7 Å per mV/nm applied electric field, or about 8 nm/volt in our structure.

about the center-of-charge of the wavefunction. The integral can be found in [92]:

$$\int y^2 z dz = \frac{1}{3}\left(y'y - zy'^2 + z^2y^2\right) \tag{9.11}$$

where $y$ is any linear combination of Ai and Bi. Once we have found $\tilde{\varepsilon}$, we can use this to calculate the location of the center-of-charge, as shown in Figure 9-4.

For the purposes of estimating errors and corrections in our spectra, a closed-form expression is more useful than the exact result. As noted above, if we expand this using stationary perturbation theory, the first order shift vanishes. The second order shift has terms coming from all $\phi_n$ with $n$ even, but the energy denominator grows as $n$ while the matrix element shrinks as $n^{-3}$, so the contributions from $n > 2$ are small, and we neglect them. It is worth noting that in the actual well, the states corresponding to $n > 2$ are not bound, and we expect the error associated with our approximate treatment of the finite barrier height in the well is larger than the error from neglecting these states.

Using the matrix element $\langle \phi_1 | z | \phi_2 \rangle = -\frac{16w}{9\pi^2}$ and $(E_2 - E_1)^{-1} = \frac{2m^* w^2}{3\pi^2 \hbar^2}$ we quickly arrive at the result

$$\delta E \approx \left( \frac{16wEe}{9\pi^2} \right)^2 \frac{2m^* w^2}{3\pi^2 \hbar^2} \tag{9.12}$$

Similarly, using the same approximations and matrix element the first order wavefunction change is

$$\delta\phi = -\frac{16w}{9\pi^2} \frac{2m^* w^2}{3\pi^2 \hbar^2} \mathrm{E}e|\phi_2\rangle \tag{9.13}$$

giving a first-order center of mass shift

$$\langle z \rangle = \left( \frac{16w}{9\pi^2} \right)^2 \frac{4m^* w^2}{3\pi^2 \hbar^2} \mathrm{E}e \tag{9.14}$$

Both the exact and approximate results are plotted in Figure 9-3 and Figure 9-4. Note that the in this approximation the first-order energy shift is in fact equal to $\mathrm{E}e\langle z \rangle$, as we might expect from simple electrostatic arguments.

## 9.3 Finite Well; Corrections

In our actual structure, the barriers on the quantum well are large but finite. This results in a softening of the confinement potential, making the corrections to the polarizability and bound state-energy shift slightly larger. We will approximate the electrostatic potential as flat outside of the well (Figure 9-5), and neglect the Fermi energy and variation in the bound state energy as compared to the barrier height; as we noted before, the barrier height is very large compared to these, and, because the decay constant $\kappa$ of the wavefunction in the barrier goes as

$$\sqrt{\frac{2m^*}{\hbar^2}} \sqrt{V_0 - \varepsilon} \approx \sqrt{\frac{2m^* V_0}{\hbar^2}} (1 - \varepsilon/(2V_0)) \tag{9.15}$$

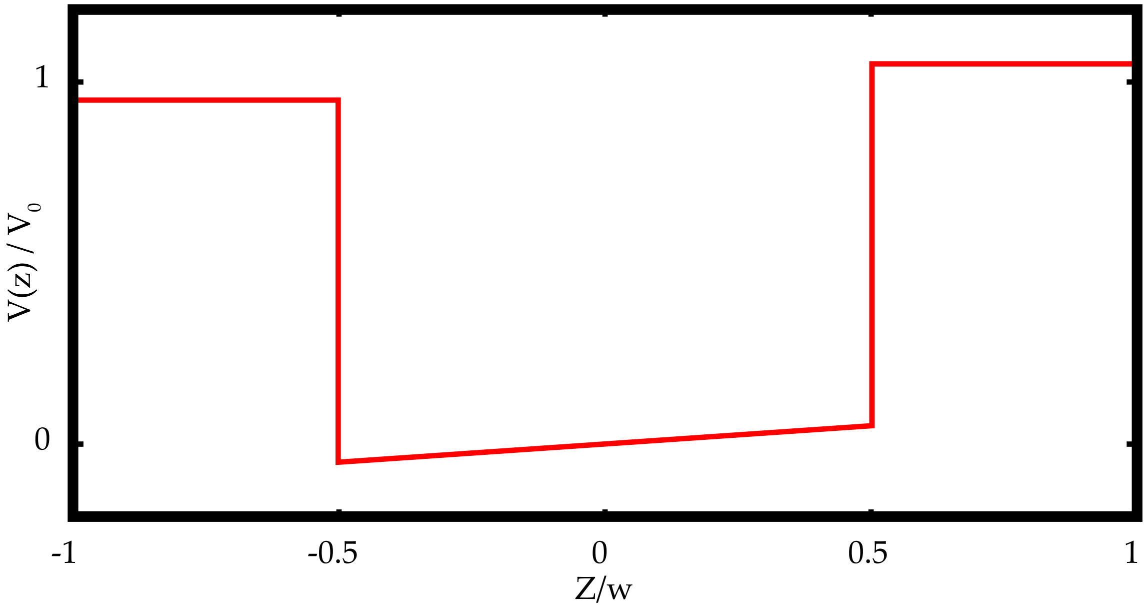


Figure 9-5: Approximate potential used for examining finite well. Shown with an (exaggerated) electric field of $0.1V_0/w$ applied.

we will at most underestimate the decay length by a few percent. (Strictly, this is only true if $\frac{e\mathrm{E}}{\kappa} \ll V_0$, but if this is not the case, we have applied enough electric field that our quantum well is beginning to field emit into the electrodes, and the sample is probably melting.)

Because the potential is finite at the interface, $\frac{1}{m^*}\frac{\partial\Psi}{\partial z}$ is continuous.[†] Matching it between the wavefunction and the decaying exponential in the barrier gives

$$\frac{1}{m^*_{\text{well}}}\frac{\partial\psi(z)}{\partial z}|_{z=\pm(w/2-\epsilon)} = \frac{1}{m^*_{\text{barrier}}}\frac{\partial\psi(z)}{\partial z}|_{z=\pm(w/2+\epsilon)} = \frac{1}{m^*_{\text{barrier}}} \mp \kappa\psi(z)|_{z=\pm(w/2+\epsilon)} \tag{9.16}$$

It is possible, although somewhat messy, to apply the boundary condition of Equation 9.16 to Equation 9.10 (or even to use the exact potential outside the well and match to decaying Airy functions in the barriers). We will instead introduce an approximate method that gives very good solutions, has a nice physical meaning, and will prove useful in estimating the magnitude of the matrix element corrections when extracting the DOS from our I-V curves.

---

[†] $m^*$ in GaAs is 0.067; $m^*$ in $\mathrm{Al}_{.3}\mathrm{Ga}_{.7}\mathrm{As}$ is 0.088. We need to include the possibility that the effective mass is different in the barriers and in the electrodes in our results.

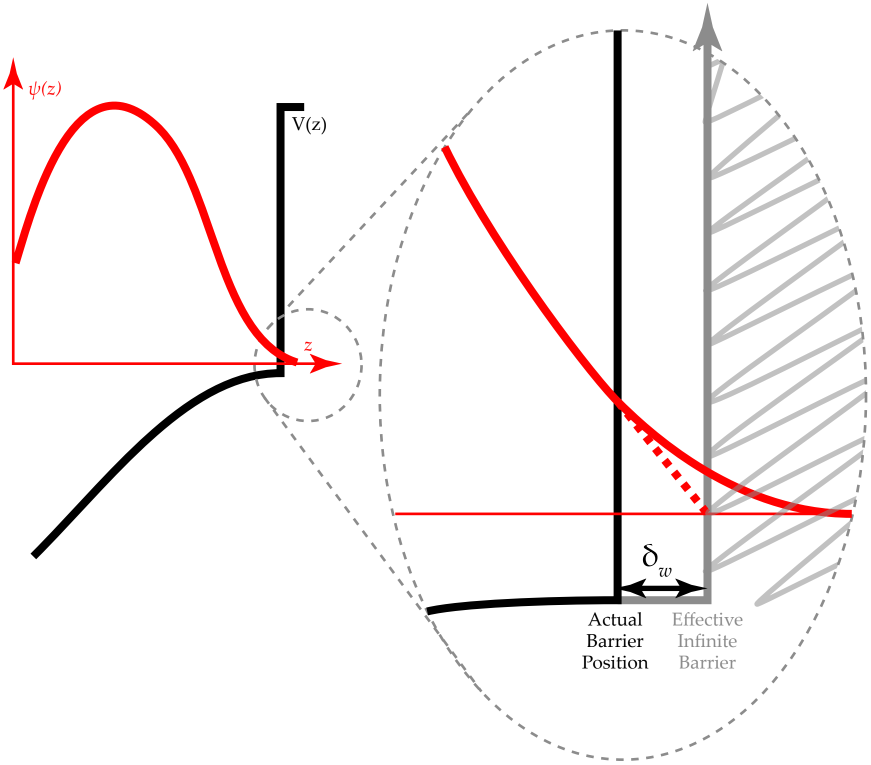


Figure 9-6: Real confining potential and wavefunction, showing the replacement of the potential barrier with an infinite barrier a distance $\delta_w$ inside of the real barrier.

Notice that the bound state energy and eigenfunction inside of the well are identical to those inside of a wider well, with finite wavefunction amplitude at the real barrier. If the amplitude of the wavefunction just inside the barrier is sufficiently small, we can use a linear approximation to find the location of its zero, where the barrier would be in a fictitious wider well (Figure 9-6). This zero will be at a location $w/2 + \delta_w$, with

$$\delta_w = \psi(z) \left( \left. \frac{\partial \psi(z)}{\partial z} \right|_{z=\pm(w/2-\epsilon)} \right)^{-1} = \mp \frac{m_{\text{barrier}}}{m_{\text{well}}} \kappa^{-1} \tag{9.17}$$

In other words, we can account approximately for the penetration of our wavefunction into a finite barrier by replacing our quantum well with an infinite well $2\frac{m_{\text{barrier}}}{m_{\text{well}}}\kappa^{-1} + O((k/\kappa)^{-2})$ wider than the actual one. (In fact, for a well with no electric field applied, the wavefunction inside the well is sinusoidal, and the first correction is $O((k/\kappa)^{-3})$.)

Notice that we have made no assumptions about the form of the wavefunction inside of the well; this is true for all of the bound states of the well as long as the amplitude at the barrier is small.

For the lowest bound state, $\frac{\partial\psi(z)}{\partial z}$ will be of the order $\pi/w$, so the wavefunction amplitude at the interface will be of the order $\frac{\pi}{\kappa w}$ times its peak value. For $x = .325$ $\mathrm{Al_xGa_{1-x}As}$, $\kappa^{-1} \approx 1.5\ nm$, and with a 17 nm quantum well, this gives an amplitude at the barrier of around 3% of its peak value; we expect this approximation to work very well in our samples. A 17.5 nm finite well is, in this approximation, equivalent to the 21.15 nm infinite well we have considered in this section. Exact results for a 17.5 nm finite well are plotted in each of the earlier plots in this section for comparison.

## 9.4 Self-Consistent Solutions

In all of this, we have neglected the effect of the electrons in the well on the potential profile in the well. However, the details of the exact process being simulated become very important when including self-consistency. In a DC bias sweep, the total charge in the well is determined by the bound state energy relative to the electrodes, while the electric field is nearly completely screened by the 2DEG before it reaches the tunneling electrode. On the other hand, when considering the instantaneous response to an applied pulse, no charge can enter or exit the well, and the additional applied field is not screened at all (although the detailed potential distribution inside of the well can change dramatically; in extreme cases, the 2DEG can slosh back and forth between the two edges of the well quite suddenly).

Thus, aside from a few notes, we will leave detailed discussion of self-consistency to Chapter 11 and Chapter 12.

In general, the applied electric field will tend to pull the charge distribution towards one end of the well through the well polarizability, while self-consistency will tend to flatten the applied field at the far end. We can think of self-consistency as adding a small attractive perturbation in the unoccupied end of the well (with an appropriate choice of boundary conditions for the electric field). Because the perturbation from self-consistency is largest exactly where the wavefunction is smallest, we do not in general expect it to dramatically change the above results. Nonetheless, where appropriate, we will refer to self-consistent simulations as described in Appendix C.

## 9.5 Summary

We have established that although our well is largely two-dimensional from the point of view of the intra-well physics we are interested in, its three-dimensional nature becomes evident with applied electric field. We have evaluated exactly, for non-interacting electrons, the bound state energy and charge center-of-mass for the lowest few bound states in an infinite quantum well as a function of applied electric field, and developed a very simple technique for dealing with finite barrier height in real quantum wells.

Finally, it is worth noting that the largest energies from the Fermi surface that we typically examine in this work are about 15 meV, well short of the energy of the second bound state. We do not usually expect to see it directly in any of our spectra, so the only contribution from the three-dimensional nature of the quantum well in our experiment will be the somewhat confounding effects it has on putting precise scales on the density, energy, and DOS scales of our spectra.

# Chapter 10

# Calibrating the Density of States

In Chapter 3, we presented a quick argument that the conductivity $\frac{dI}{dV}$ should be proportional to the tunneling density of states. However, a more detailed examination of our structure shows that this is wrong in several non-trivial ways. We will examine corrections to this simple idea, and develop ways to correct for variations in the tunneling matrix element.

## 10.1 Correcting for Matrix Elements

For the sake of simplicity, we consider only zero temperature.

Recall from Fermi's golden rule that the tunneling rate through our barrier out of any given state in the well is given by

$$r = \frac{2\pi}{\hbar} \left| \langle \phi_i | M | \phi_f \rangle \right|^2 \rho_f \left( E_f \right) \tag{10.1}$$

where $r$ is the transition rate, $M$ is the matrix element connecting the initial states to the final states, $\langle \phi_i |$ is the initial state, $| \phi_f \rangle$ is the final state, and $\rho_f$ is the density of final states.

At finite bias, our quantum well actually has a continuum of initial states, and we can integrate across them using the initial density of states $\rho_i$ to get the total transition rate $R$:

$$R = \frac{2\pi}{\hbar} \int_0^{e\mathrm{V}} \left|\langle \phi_i | M | \phi_f \rangle\right|^2 \rho_f(\varepsilon)\, \rho_i(\varepsilon)\, d\varepsilon \tag{10.2}$$

where V is the applied voltage between the 2DEG and the electrode and we measure energies with respect to the Fermi energy of the tunneling electrode.

In our simplified discussion, we assumed that the "uninteresting" part of this expression, $\left|\langle \phi_i | M | \phi_f \rangle\right|^2 \rho_f(\varepsilon)$ was constant, and thus found

$$\frac{d\mathrm{I}}{d\mathrm{V}} = e\frac{dR}{d\mathrm{V}} \propto \rho_i(e\mathrm{V}) \tag{10.3}$$

However, in our real sample, our pulses are comparable in energy to the 3D electrode's Fermi energy, and the 3D density of states varies appreciably. Furthermore, we will find that the tunneling matrix element $M$ changes significantly both with the momentum of the carriers in the 3D system, and as the quantum well polarizes itself in response to applied electric fields. In this case, we can no longer simply use the derivative to remove the integral when computing $\frac{d\mathrm{I}}{d\mathrm{V}}$, but must apply Leibniz's rule, giving

$$\frac{d\mathrm{I}}{d\mathrm{V}} = \frac{2\pi e}{\hbar} \left( \left|\langle \phi_i | M | \phi_f \rangle\right|^2 \rho_f(e\mathrm{V})\, \rho_i(e\mathrm{V}) + \int_0^{e\mathrm{V}} \frac{\partial}{\partial \mathrm{V}} \left[ \left|\langle \phi_i | M | \phi_f \rangle\right|^2 \rho_f(\varepsilon)\, \rho_i(\varepsilon) \right] d\varepsilon \right) \tag{10.4}$$

The first term in this is just proportional to the density of initial states, times our usual "uninteresting" matrix element and density of final states, and gives us the change in current because of electrons that are now able to tunnel because of the change in voltage. The second term, however, is an integral over the change in the "uninteresting" matrix element, and both densities of states. It gives the contribution in current due to electrons that were already able to tunnel, but now tunnel at a different rate because of the applied voltage. This second term is especially worrying; it mixes the actual DOS with a phase-shifted copy of itself, which could lead to substantial pulling of peaks and offsets in the spectra.

As discussed in Chapter 9, the 2DEG itself has no direct sensitivity to the applied electric field, so $\frac{\partial}{\partial V}\rho_i\left(\varepsilon\right) = 0$, but all of the remaining elements in this term can and do vary with the applied voltage. In trying to deal with this, we will initially assume our 2DEG is an ideal, noninteracting 2DEG with no applied magnetic field. This will allow us to simplify the problem we face.

## 10.2 Momentum Conservation and the 3D Density of States

One of the chief advantages of TDCS is that the structure used in it has translational symmetry in the $\hat{x}\hat{y}$ plane, rendering it insensitive to the in-plane conductivity of the 2DEG. This translational symmetry also implies that the transverse momentum $\hbar k_\perp$ is a conserved constant of motion in tunneling, which enormously restricts the states into and out of which electrons can tunnel[93]. Assuming that the effective mass is the same in the well and the electrode, conservation of energy then gives us, with $\delta_V$ the energy offset between the bottom of the 2D band and the bottom of the 3D band,

$$\begin{aligned} E_i &= E_f && (10.5) \\ \frac{\hbar^2 k_\perp^2}{2m^*} + \delta_V &= \frac{\hbar^2}{2m^*}\left(k_\perp^2 + k_z^2\right) && (10.6) \\ \frac{2m^*\delta_V}{\hbar^2} &= k_z^2 && (10.7) \end{aligned}$$

Thus, the band offsets between the 2D and 3D system uniquely determine the wavevectors $k_z$ in the 3D that can tunnel. Notice that all of the particles able to tunnel out of the 2DEG for a given density and applied voltage tunnel at the same $k_z$, regardless of $k_\perp$.

This suggests a geometric representation of the conserved quantities in tunneling. We begin by plotting the Fermi surface of the 3D electrode in the $k_x, k_y, k_z$ basis. On top of this, we can plot the 2D Fermi surface, assigning to it a fictitious $k_z = \frac{2m^*\delta_V}{\hbar^2}$. With this choice, states that occupy the same point in this diagram can tunnel between each other while conserving both energy and momentum. States that are in the same $k_x, k_y$ but different $k_z$ would have to violate energy conservation to tunnel, while those in different $k_x, k_y$, but the same $k_z$ would have to violate momentum conservation. Thus, the question of "how many states can tunnel (into/out of) the 2DEG without violating conservation laws" can be mapped to the question "what is the area of overlap between the (unoccupied/occupied) states of the 2D disc with the (occupied/unoccupied) states in the 3D sphere".

When the 2DEG and 3D are at equilibrium, the edge of the occupied Fermi disc will be the intersection between the disc and the sphere, and only the states along this edge can tunnel; this is shown in Figure 10-1.

If we then disequilibrate the 2D and 3D (Figure 10-2) by an energy $\Delta$, we will translate the Fermi disc up or down in $k_z$ by $2m^*\Delta/\hbar^2$, and the band of occupied states outside the Fermi sphere or unoccupied states inside the Fermi sphere will be able to tunnel. Notice that regardless of the relative densities of the 2D and 3D or the sign or magnitude of any applied pulse, all of the electrons able to tunnel will tunnel from the same value of $k_z$ in the 3D. This will greatly simplify our tunneling problem. We now have enough information to determine $\rho_f$ in Equation 10.2; $k_x$ and $k_y$ are conserved, so we are left with the 1D density of states in $k_z$ in the 3D; $\rho_f \propto 1/k_z$, with $k_z$ as above. Note that this means $\rho_f$ does not depend on $\varepsilon$, so we can take it out of the integral in Equation 10.4.

It is worth noting that for any dispersion relation other than $\mathrm{E} = \hbar^2 k^2/(2m^*)$ with $m^*$ the same in the 2D and the 3D (such as when a magnetic field is applied and the energy spectrum breaks up into discrete Landau levels), there will no longer be a single unique $k_z$ for the entire electron gas. This will "warp" the Fermi disc (Figure 10-3a). Further, because our quasiparticles have a finite lifetime, the Fermi disc will also acquire some width, as shown in Figure 10-3b. Of course, in a fully interacting theory, there is an incoherent background as well as the quasiparticle peaks, and our DOS is better thought of as a cloud of states which are eigenstates of neither energy nor momentum (Figure 10-3c). We cannot

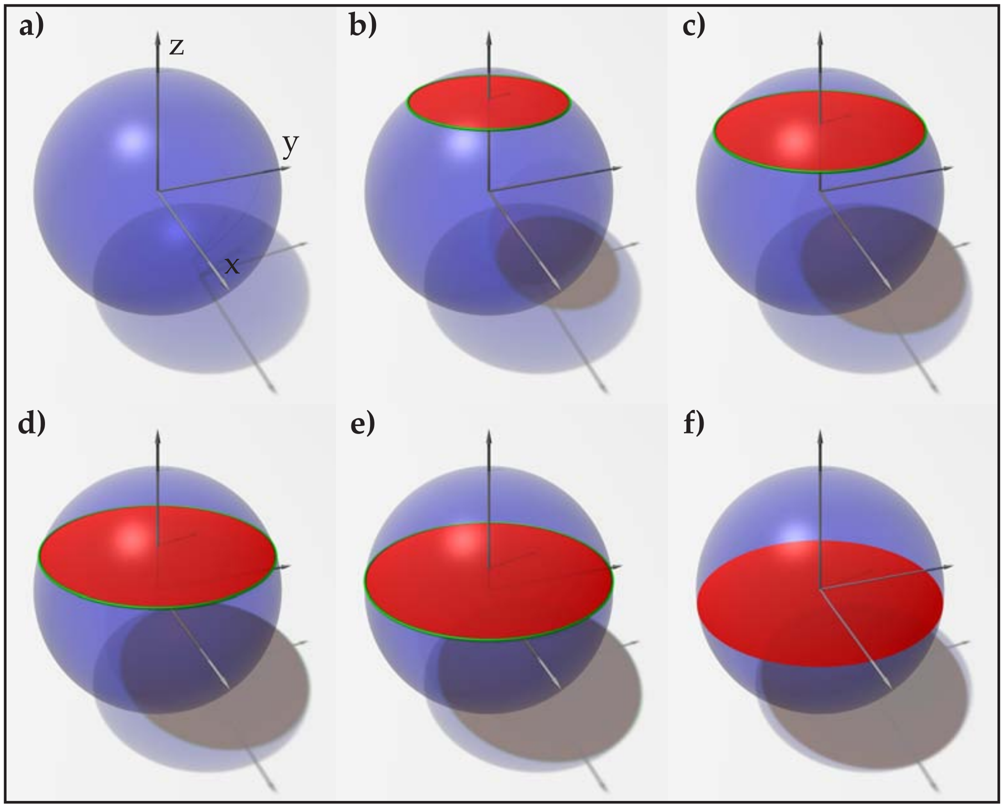


Figure 10-1: **The 3D Fermi sphere and 2D Fermi disc at equilibrium: a)** The 3D Fermi sphere, shown with the axes labels. We take the plane of the 2D to be the $x - y$ plane. **b-e)** The Fermi sphere, with an equilibrium Fermi disc superimposed at a variety of increasing densities. For clarity, only filled states are shown in the 2D. States able to tunnel are highlighted in green. **f)** Once the Fermi wavevector in the 2D becomes larger than the Fermi wavevector in the 3D, equilibrium tunneling is no longer possible unless momentum or energy conservation is broken.

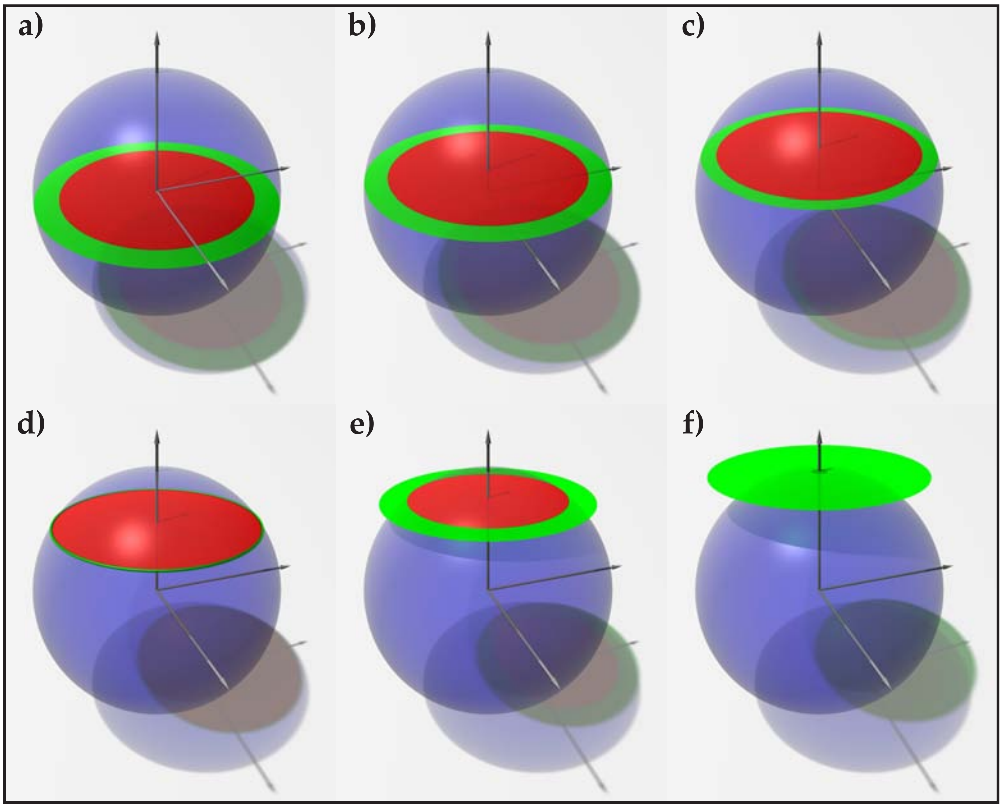


Figure 10-2: **The 3D Fermi sphere and 2D Fermi disc out of equilibrium:** The band of states able to tunnel for a variety of injecting (**a-c**) and ejecting (**e-f**) pulses. Note that if the injecting pulse in **a** were any larger, the 2D would be pulled down in $k_z$ through the origin of the sphere, and it would no longer be possible to simultaneously conserve energy and momentum. This is the region we substitute a black bar for in our matrix-element corrected plots. In addition, notice that in **f**, all of the states in the 2DEG are able to tunnel, and further increases in the pulse height will not change the number of states able to tunnel. This situation corresponds to the band edge in our DOS plots.

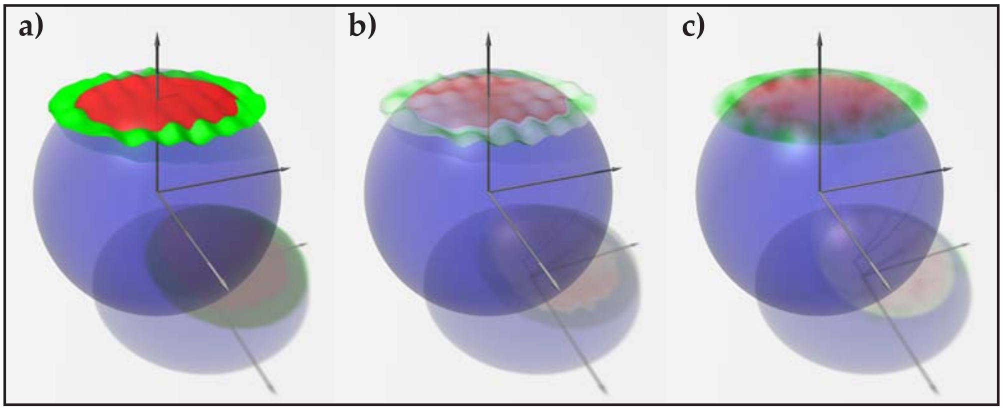


Figure 10-3: **The 3D Fermi sphere and 2D Fermi disc with interacting electrons:** **a)** The 2D Fermi disc will not be flat in our diagrams if the 2D dispersion relation is not the same as the 3D relation. **b)** Finite lifetimes create some spread in energy in the 2DEG, and a corresponding spread in the fictitious $k_z$ we assign to the 2DEG. **c)** The full quasiparticle spectral functions are not described by a single peak and lifetime, but by a cloud of probability, sharply peaked at the quasiparticle energies.

directly image this momentum dependence in our experiment. For the purposes of our matrix-element correction, which is approximate anyway, we will assume the dispersion of the electrons is similar to that for the zero field non-interacting electron gas. This approximation will become progressively worse as we raise the magnetic field.

Finally, we assume the 3D density of states is featureless here. In high magnetic fields, the 3D will actually form "Landau tubes", resulting in substantial variation of the DOS with $k_\perp$. However, the donors in the 3D provide strong sources of scattering, and we do not expect these Landau levels in the 3D to become well resolved until $l_b \sim N_d^{-1/3}$, at a field of around 16 T in the current sample.

## 10.3 Transfer Matrix Formalism

Next, we need to deal with the tunneling matrix element. Fermi's golden rule is a perturbational result, with $M$ a small correction to the Hamiltonian $\mathcal{H}$. In our

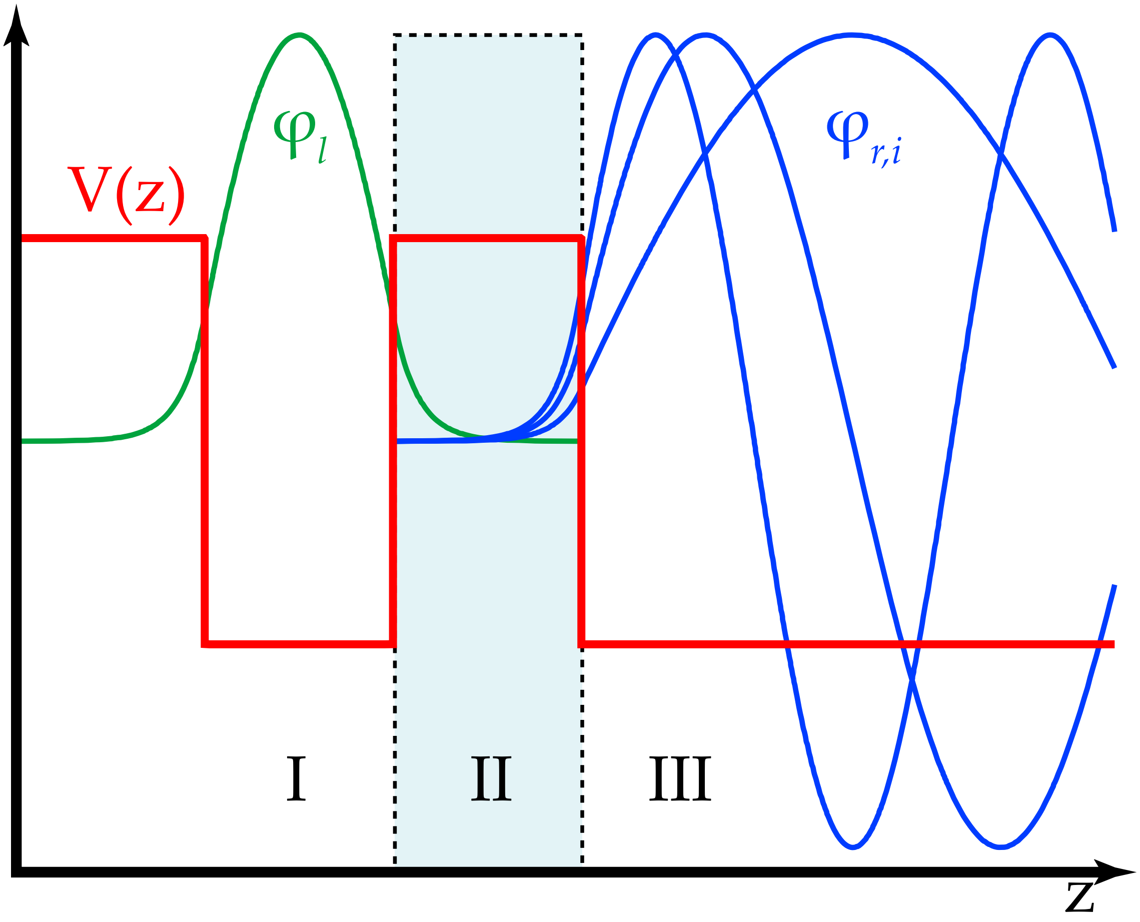


Figure 10-4: Regions and eigenstates for the transfer matrix formalism

case, although we can easily find the eigenstates of our two systems in isolation, it is less than obvious how to construct a small perturbation that joins them together through a tunneling barrier. One could imagine separating them by an infinite barrier and lowering that barrier as a perturbation, but the resulting correction to the Hamiltonian would be divergent, not small. Fortunately, Bardeen considered this problem in the context of metal-insulator-superconductor tunneling. He provides a simple formalism calculating tunneling matrix elements [94]. Because the result is unfamiliar to many readers, and we must deal with a position dependent $m^*$, we provide a brief derivation here.

Assume we have a Hamiltonian $\mathcal{H} = \vec{\nabla} \cdot \frac{1}{m^*} \vec{\nabla} + V(x)$ that describes the combined well and electrode system, and have already found a state vector $|\phi_{l,0}\rangle$

which is an eigenstate of the quantum well with energy $E_0$, meaning that inside of the quantum well and barrier (regions I and II in Figure 10-4),

$$\mathcal{H}\phi_{l,0} = E_0\phi_{l,0} \qquad \text{Inside regions I and II} \tag{10.8}$$

and that deep inside region III $\phi_{l,0}$ drops to zero. We further assume we have an orthonormal set of states $|\phi_{r,j}\rangle$ with energy $E_j$, which are similarly good eigenstates of the electrode and barrier in regions II and III, but not in region I. We will examine the time evolution of these states, and construct an approximate matrix element $M_j$ which, if applied as a perturbation to a system whose eigenstates are exactly $|\phi_{l,0}\rangle$ and $|\phi_{r,j}\rangle$ with energies $E_0$ and $E_j$, will reproduce the time evolution of the actual states.

We begin by constructing a time-dependent wavefunction

$$\Psi(t) = a(t)e^{-i\omega_0 t}|\phi_{l,0}\rangle + \sum_i b_i(t)e^{-i\omega_i t}|\phi_{r,i}\rangle \tag{10.9}$$

where $a(0) = 1$ and $b_i(0) = 0$.

Substituting into the time-dependent Schrödinger equation,

$$\begin{aligned} &a(t)\mathcal{H}|\phi_{l,0}\rangle + \sum_i b_i(t)\mathcal{H}|\phi_{r,i}\rangle = \\ &i\hbar\left\{\left(a'(t) - iw_0 a(t)\right)|\phi_{l,0}\rangle + \sum_i \left(b_i'(t) - iw_i b_i(t)\right)|\phi_{r,i}\rangle\right\} \end{aligned} \tag{10.10}$$

Projecting this result onto $\langle\phi_{r,j}|$, using $\langle\phi_{r,i}|\phi_{r,j}\rangle = \delta_{ij}$, and evaluating at $t = 0$, we find

$$b_j'(t) = \frac{a}{i\hbar}\langle\phi_{r,j}|\mathcal{H} - E_0|\phi_{l,0}\rangle - a'(t)\langle\phi_{r,j}|\phi_{l,0}\rangle \tag{10.11}$$

The term proportional to $a'$ is there because $|\phi_{r,j}\rangle$ and $|\phi_{l,0}\rangle$ are not quite orthogonal, and so as the occupation of $|\phi_{l,0}\rangle$, changes, there is a small change in the expected occupation of $|\phi_{r,j}\rangle$. For an adequately thick barrier, this term is negligible. What remains can be compared with the standard result from time dependent perturbation theory, giving a perturbational matrix element $M$

$$M_j = \langle \phi_{r,j} | \mathcal{H} - E_0 | \phi_{l,0} \rangle \tag{10.12}$$

This result alone is somewhat disturbing; it depends on an integral over all space, but each of our wavefunctions is only well defined in a small region. If we are to believe it has any physical significance, it should only depend on the value of the wavefunctions in regions in which they are well defined. Writing out $\mathcal{H}$ explicitly gives a hint as to how to accomplish this; it is possible to imagine using the divergence term in it in the divergence theorem, to reduce our volume integral to a surface integral.

$$M_j = \iiint \phi_{r,j}{}^* \left( \vec{\nabla} \cdot \frac{1}{2m^*} \vec{\nabla} + V(\vec{x}) - E_0 \right) \phi_{l,0} dV \tag{10.13}$$

Note that because of Equation 10.8, this integral vanishes except inside of region III, allowing us to take it only over region III and an arbitrary part of region II. We can attempt to remove the dependence on the potential by noting that, using our definition of $|\phi_{r,j}\rangle$, that $(\mathcal{H} - E_j)\phi_{r,j} = ((\mathcal{H} - E_j)\phi_{r,j})^* = 0$ inside of regions III and II, and subtracting it times $\phi_{l,0}$ from the integrand, leaving

$$\begin{aligned} M_j &= \iiint_{\text{II,III}} \phi_{r,j}{}^* \left( \vec{\nabla} \cdot \frac{\hbar^2}{2m^*} \vec{\nabla} + V(\vec{x}) - E_0 \right) \phi_{l,0} - \\ &\qquad \phi_{l,0} \left( \vec{\nabla} \cdot \frac{\hbar^2}{2m^*} \vec{\nabla} + V(\vec{x}) - E_j \right) \phi_{r,j} \, dV && (10.14) \\ &= \iiint_{\text{II,III}} \phi_{r,j}{}^* \vec{\nabla} \cdot \frac{\hbar^2}{2m^*} \vec{\nabla} \phi_{l,0} - \phi_{l,0} \vec{\nabla} \cdot \frac{\hbar^2}{2m^*} \vec{\nabla} \phi_{r,j} \, dV && (10.15) \\ &= \iiint_{\text{II,III}} \vec{\nabla} \cdot \left( \phi_{r,j}{}^* \frac{\hbar^2}{2m^*} \vec{\nabla} \phi_{l,0} - \phi_{l,0} \frac{\hbar^2}{2m^*} \vec{\nabla} \phi_{r,j} \right) dV && (10.16) \\ &= \iint_{\text{S}} \frac{\hbar^2}{2m^*} \left( \phi_{r,j}{}^* \vec{\nabla} \phi_{l,0} - \phi_{l,0} \vec{\nabla} \phi_{r,j} \right) \cdot d\text{S} && (10.17) \end{aligned}$$

where S is any surface passing through region II, and we have used $E_0 \approx E_j$ in the energy range of interest for Fermi's golden rule.

This result is pleasing in that it resembles the familiar current operator, and it only depends on the values of the wavefunction inside of the tunneling barrier. For the case of a 1D potential, we can separate out the $x - y$ dependence of the wavefunctions by taking $\phi_{(lr),i}(\vec{x}) = \phi_{(lr),i}(z)e^{i(k_{lr,x}x+k_{lr,y}y)}(2\pi)^{-2}$, and perform the integral immediately, giving [95]

$$M_i = \frac{\hbar^2}{2m^*}\left(\phi_{l,0}\frac{\partial \phi_{r,i}{}^*}{\partial z} - \phi_{r,i}{}^*\frac{\partial \phi_{l,0}}{\partial z}\right)\bigg|_{z=z0} \delta(k_{r,x} - k_{l,x})\delta(k_{r,y} - k_{l,y}) \tag{10.18}$$

reconfirming our momentum selection rules from earlier in the chapter.

## 10.4 Tunneling Matrix Elements

Just as in Section 9.3, the tunneling barrier is infinitely high compared to the bound state energies, the applied voltage, and the offset of the 2DEG band bottom from the 3D band bottom, so we can do quite well in approximating it as square. For the sake of concreteness, we put our barrier in the region $(-W, 0)$, and give it height $V_0$ above the bound state energy of our well . In this case, inside of the barrier,

$$\phi_l = A_l e^{-\kappa(z+W)} \phi_r = A_r e^{\kappa(z)} \tag{10.19}$$

where $\kappa = \sqrt{\frac{2m^* V_0}{\hbar^2}}$, and $A_{lr}$ is the amplitude of the wavefunction of the 2DEG and 3D where they enter the barrier. Substituting into Equation 10.18, this gives us our tunneling matrix element

$$M = A_l A_r^* \frac{\hbar^2 \kappa}{m^*_{\text{barrier}}} e^{-\kappa W} \tag{10.20}$$

Before we turn our attention to calculating $A_l$ and $A_r^*$, we need to revisit our neglect of the variation in $\kappa$. Although the fractional changes in $\kappa$ due to the applied electric are small, the product $\kappa W$ is quite large in our structure (about 10), and we are exponentially sensitive to variations in $\kappa W$ in the last term of Equation 10.20. This demands a more careful treatment of our tunneling barrier;

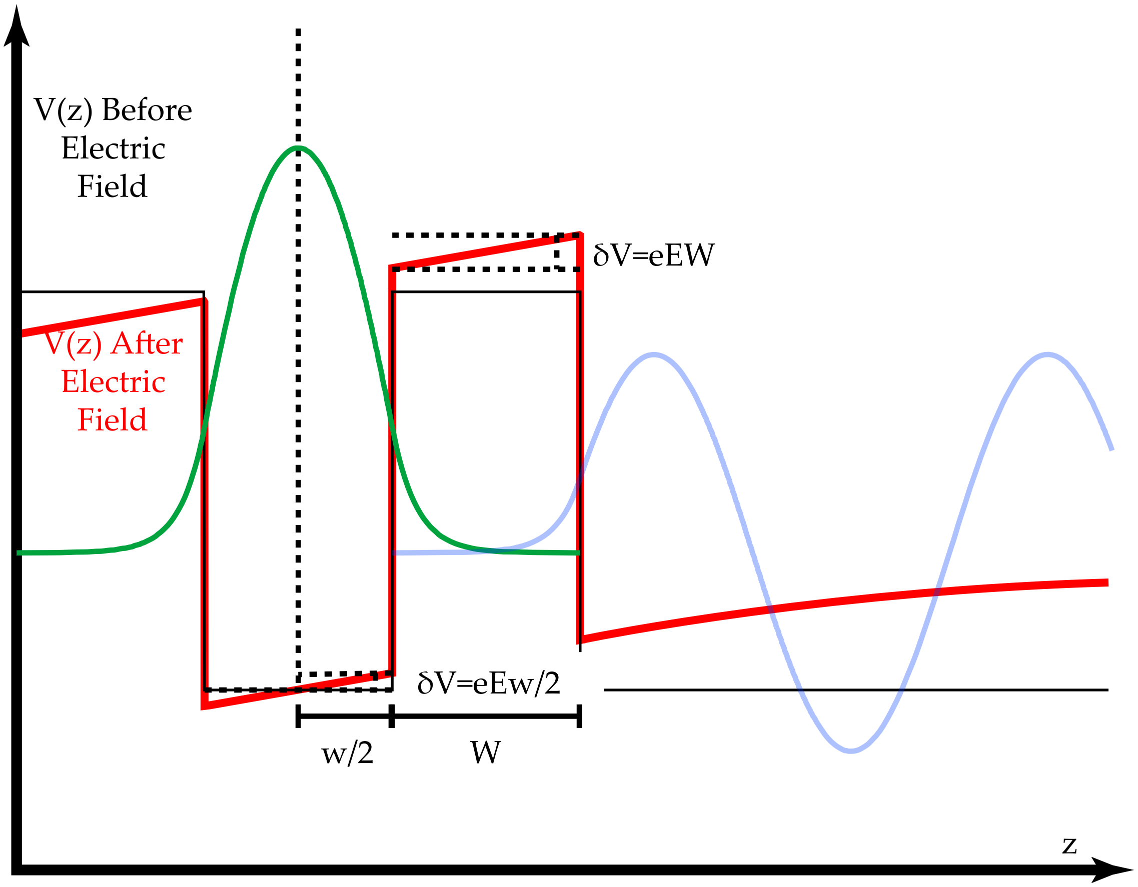


Figure 10-5: **The trapezoidal tunnel barrier** we use in approximating our matrix element

it is both slightly different in height from $V_0$, and it is trapezoidal due to the applied electric field (Figure 10-5).

With this small variation in the barrier shape, we can write our decay constant $\kappa$ as

$$\kappa = \sqrt{\frac{2m(V_0 + \delta_V + Eez)}{\hbar^2}} \approx \sqrt{\frac{2mV_0}{\hbar^2}} \left(1 + \frac{\delta_V}{2V_0} + \frac{Eez}{2V_0}\right) \qquad (10.21)$$

We will deal with these two corrections separately. Taking $\kappa \rightarrow \kappa + \kappa\delta_V/(2V_0)$, and assuming $\kappa W \delta_V/(2V_0) \ll 1$,

$$e^{-\kappa W} \rightarrow e^{-\kappa W} e^{-\frac{\kappa \delta_V W}{2V_0}} \approx e^{-\kappa W}(1 - \kappa W \frac{\delta_V}{2V_0}) \qquad (10.22)$$

For the trapezoidal barrier top, we use the WKB approximation inside of the barrier, and under the same assumptions,

$$\begin{aligned}
\psi(z) &= \frac{1}{\sqrt{\kappa(z)}} \exp\left(-\int \kappa(z) dz\right) \qquad (10.23)\\
&= \frac{1}{\sqrt{\kappa + \kappa \frac{\mathrm{E}ez}{2V_0}}} \exp\left(-\int \kappa + \kappa \frac{\mathrm{E}ez}{2V_0} dz\right) \qquad (10.24)\\
&= \frac{1 - \frac{\mathrm{E}ez}{4V_0}}{\sqrt{\kappa}} e^{-\kappa z} e^{-\kappa \frac{\mathrm{E}ez^2}{4V_0}} \qquad (10.25)\\
&= \frac{1 - \frac{\mathrm{E}ez}{4V_0}}{\sqrt{\kappa}} e^{-\kappa z} \left(1 - \kappa \frac{\mathrm{E}ez^2}{4V_0}\right) \qquad (10.26)\\
\frac{\psi(W)}{\psi(0)} &= \left(1 - \frac{\mathrm{E}eW}{4V_0}\right) e^{-\kappa W} \left(1 - \kappa W \frac{\mathrm{E}eW}{4V_0}\right) \qquad (10.27)\\
&= e^{-\kappa W} \left(1 - (\kappa W + 1) \frac{\mathrm{E}eW}{4V_0}\right) \qquad (10.28)
\end{aligned}$$

Notice that if $\kappa W \gg 1$ and $\kappa W \delta_V / (2V_0) \ll 1$, the transmission of the trapezoidal barrier is just equal to the transmission of a barrier whose height is the average height of the barrier.

Using this approximation and putting these together, we get a revised matrix element:

$$M = A_l A_r^* \frac{\hbar^2 \kappa}{m^*_{\text{barrier}}} \left(1 - (\kappa W) \frac{e\mathrm{E}w + e\mathrm{E}W}{4V_0}\right) e^{\kappa W} \qquad (10.29)$$

In the 3D, outside of the barrier, our wavefunctions are given by

$$\phi_r = \frac{1}{4\pi} \sin(k_z(z + z_0)) \qquad (10.30)$$

For an infinite barrier, $z_0$ is zero, and we simply have a standing wave near the wall. As the barrier is lowered, $z_0$ becomes very slightly positive, and the first zero of the wavefunction is moved slightly towards the barrier. Matching boundary conditions as in Equation 9.16, we find

$$\begin{aligned}
A_r &= \frac{1}{4\pi} \frac{k_z}{\kappa} \frac{m^*_{\text{barrier}}}{m^*_{\text{3d}}} \qquad (10.31)\\
z_0 &= \frac{1}{\kappa} \frac{m^*_{\text{barrier}}}{m^*_{\text{3d}}} \qquad (10.32)
\end{aligned}$$

Once again, just as for the quantum well, the finite but tall barrier is equivalent to an infinite barrier positioned a distance $z_0$ inside of the actual barrier.

Inside of the quantum well, we can use the first order perturbational result Equation 9.13 to approximately evaluate $A_l$; inside of the well

$$\phi_l \approx \sqrt{\frac{2}{w'}} \left[ \sin\left(\frac{\pi z}{w'}\right) - \mathrm{E}e \frac{32 m^*_{\mathrm{well}} w'^3}{27\pi^4 \hbar^2} \sin\left(\frac{2\pi z}{w'}\right) \right] \tag{10.33}$$

with $w' = w + \frac{2}{\kappa} \frac{m^*_{\mathrm{barrier}}}{m^*_{\mathrm{well}}}$ is the effective width of our quantum well. We have already approximated the derivative $\frac{\partial z}{\partial \phi_l}$ as constant over the distance $w' - w$;

$$\left.\frac{\partial \phi_l}{\partial z}\right|_{z=0} = \sqrt{\frac{2\pi^2}{w'^3}} \left(1 - \mathrm{E}e \frac{64 m^*_{\mathrm{well}} w'^3}{27\pi^4 \hbar^2}\right) \tag{10.34}$$

This gives an amplitude at the interface a distance $(w - w')/2$ away of

$$\begin{aligned} A_l &= \frac{1}{\kappa} \frac{m^*_{\mathrm{barrier}}}{m^*_{\mathrm{well}}} \frac{\partial \phi_l}{\partial z} && (10.35) \\ &= \frac{1}{\kappa} \frac{m^*_{\mathrm{barrier}}}{m^*_{\mathrm{well}}} \sqrt{\frac{2\pi^2}{w'^3}} \left(1 - \mathrm{E}e \frac{64 m^*_{\mathrm{well}} w'^3}{27\pi^4 \hbar^2}\right) && (10.36) \end{aligned}$$

Substituting $A_l$ and $A_r$ into Equation 10.20 gives

$$\begin{aligned} M &= \frac{\hbar^2 k_z}{\kappa m^*_{\mathrm{well}} m^*_{\mathrm{3D}} \sqrt{8 w'^3}} \left(1 - \mathrm{E}e \frac{64 m^*_{\mathrm{well}} w'^3}{27\pi^4 \hbar^2} - (\kappa W) \frac{e\mathrm{E}w + e\mathrm{E}W}{4V_0}\right) e^{-\kappa W} && (10.37) \\ &\propto k_z \left(1 - \mathrm{E}e \frac{64 m^*_{\mathrm{well}} w'^3}{27\pi^4 \hbar^2} - (\kappa W) \frac{e\mathrm{E}w + e\mathrm{E}W}{4V_0}\right) && (10.38) \end{aligned}$$

where in Equation 10.38 we have only included terms we expect to vary within a measurement. The term in this proportional to E, the applied electric field, deserves some comment. The DC applied field we use to gate the sample will typically be screened by the 2DEG, and very little of it will reach the tunneling barrier. Because of this, we expect it to have very little impact on the tunneling matrix element (we can, and will, include this screening in numeric models of the structure; see Appendix C). For now, we neglect the DC applied field, and replace E with $V_{\text{AC}}/d$, where d is the width of the entire structure, and $V_{\text{AC}}$ is the applied pulse height. Note that $M$ does not depend on $\varepsilon$, so it can also be pulled out of the integral in Equation 10.2.

We can now substitute our 3D density of states and tunneling matrix element into Fermi's golden rule (Equation 10.2), giving

$$\begin{aligned}
R &\propto \left[k_z\left(1-\frac{V_{\text{AC}}e}{d}\frac{64m^*_{\text{well}}w'^3}{27\pi^4\hbar^2}-(\kappa W)\frac{e\text{E}w+e\text{E}W}{4V_0}\right)\right]^2\times \\
&\quad \frac{1}{k_z}\int_0^{e\text{V}}\rho_i\left(\varepsilon\right)d\varepsilon && (10.39)\\
&\propto k_z\left(1-\frac{V_{\text{AC}}e}{d}\frac{64m^*_{\text{well}}w'^3}{27\pi^4\hbar^2}-(\kappa W)\frac{e\text{E}w+e\text{E}W}{4V_0}\right)^2\times \\
&\quad \int_0^{e\lambda V_{\text{AC}}}\rho_i\left(\varepsilon\right)d\varepsilon && (10.40)\\
&\propto \sqrt{V_{\text{AC}}-V_{\text{align}}}\left(1-\frac{V_{\text{AC}}e}{d}\frac{64m^*_{\text{well}}w'^3-(\kappa W)\frac{e\text{E}w+e\text{E}W}{4V_0}}{27\pi^4\hbar^2}\right)^2\times \\
&\quad \int_0^{e\lambda V_{\text{AC}}}\rho_i\left(\varepsilon\right)d\varepsilon && (10.41)
\end{aligned}$$

where $V_{\text{align}}$ is the density dependent AC voltage that precisely aligns the band bottoms of the 2DEG and 3D electrode. Note that because both of these terms depend on $V_{\text{AC}}$, if we attempt to differentiate our signal and then correct $\frac{dI}{dV}$ the correction will be quite complicated. We will have to both subtract off the second term in Equation 10.4, and divide the remaining term by the matrix element and 3D density of states. However, $\frac{R}{|M|^2\rho_f}$ is proportional to just $\int_0^{e\text{V}}\rho_i\left(\varepsilon\right)d\varepsilon$, and if we differentiate this, we are left with exactly what we want: the density of states in the quantum well.

This gives the following procedure for extracting the DOS from the raw tunnel current:

1. For each density, estimate $V_{\text{align}}$. Except for where it diverges, our correction is only weakly dependent on this number, so the estimate does not have to

be terribly good. Typically, we pick the value of $V_{\mathrm{AC}}$ at which $I \rightarrow 0$ as an estimate of $V_{\mathrm{align}}$. After finding this point for two different DC biases, we extrapolate linearly to other densities, although we can also use first-principles estimates or fits to the zero field data with similar results.

2. The remaining terms in the correction are all known from the sample geometry and the lever arm calibration to within a few percent(see Chapter 12). Substitute these in, and prepare a "corrected" current

$$\begin{aligned} I' &= \frac{I}{\sqrt{V_{\mathrm{AC}} - V_{\mathrm{align}}}\left(1 - \frac{V_{\mathrm{AC}} e}{d} \frac{64 m^*_{\mathrm{well}} w'^3}{27\pi^4\hbar^2} - (\kappa W)\frac{e\mathrm{E}w + e\mathrm{E}W}{4V_0}\right)^2} \qquad (10.42) \\ &\approx \frac{I}{\sqrt{V_{\mathrm{AC}} - V_{\mathrm{align}}}\left(1 - 5.4 V_{\mathrm{AC}}\right)^2} \qquad (10.43) \end{aligned}$$

   where we have evaluated the constant for a 175 nm quantum well with x = 0.325 barriers 13 nm thick and a total capacitor thickness of 100 nm, and $V_{\mathrm{AC}}$ is measured in volts.

3. The correction diverges as $V_{\mathrm{AC}} \rightarrow V_{\mathrm{align}}$; set $I'$ to zero when these numbers become close enough together that uncertainty in $V_{\mathrm{align}}$ renders the correction meaningless. Typically, the tunneling current is close enough to zero at this point that the only points we are removing from the spectra are pure noise anyway. This step results in the black band in the high energy, high density corner of our corrected spectra.

4. Differentiate $I'$ to get a number proportional to the single-particle density of states in the 2DEG.

The results of this correction at zero field are shown in Figure 10-6. Note that while it normalizes the contrast and makes the data much easier to understand, it does not add any new features, nor does it substantially move any features that are already present. Fundamentally, this correction should be thought of as a best-effort attempt at normalizing the spectra in a fashion that makes them easier to examine and fit, not as a quantitative or exact process.

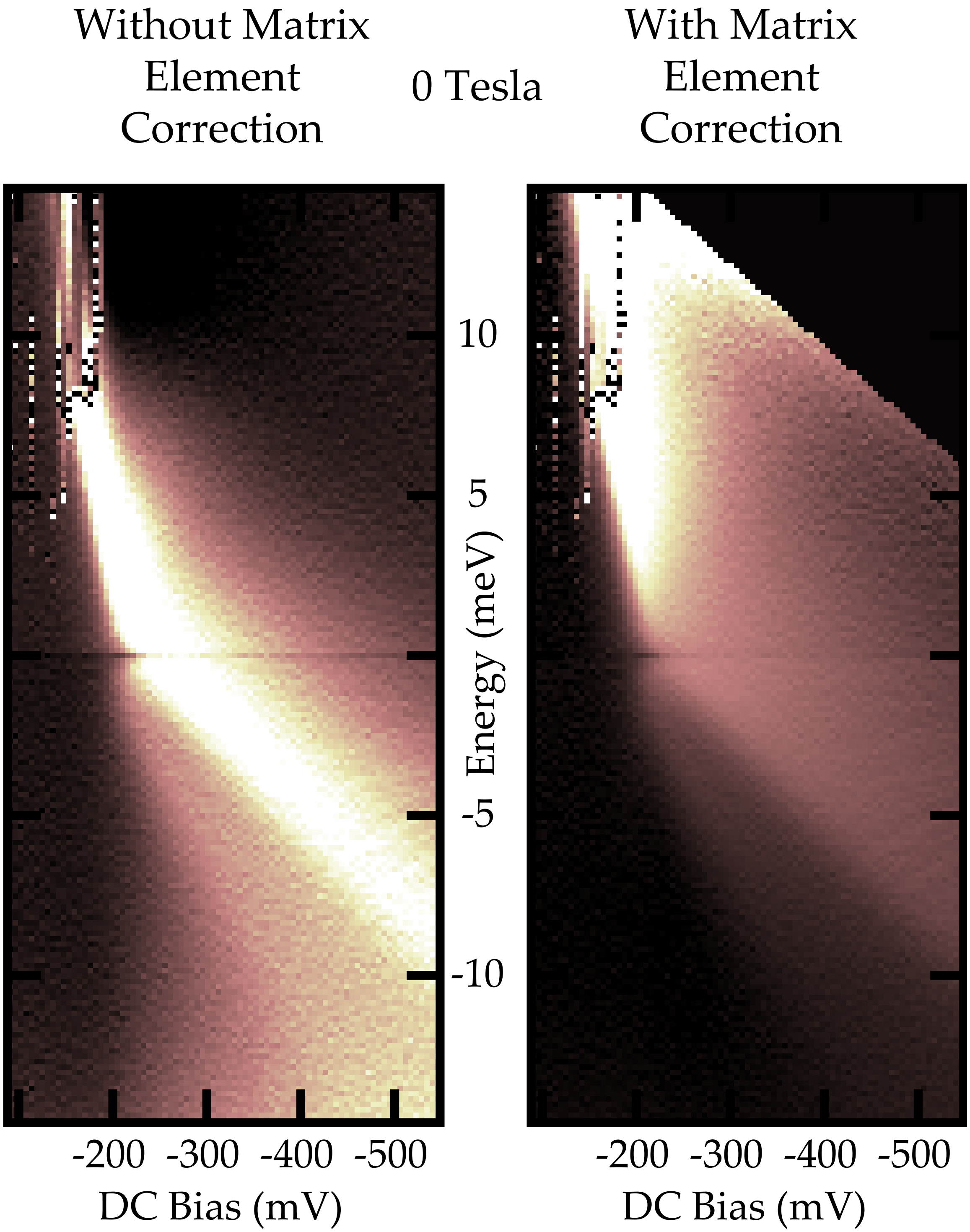


Figure 10-6: Example of matrix element correction: Zero field data

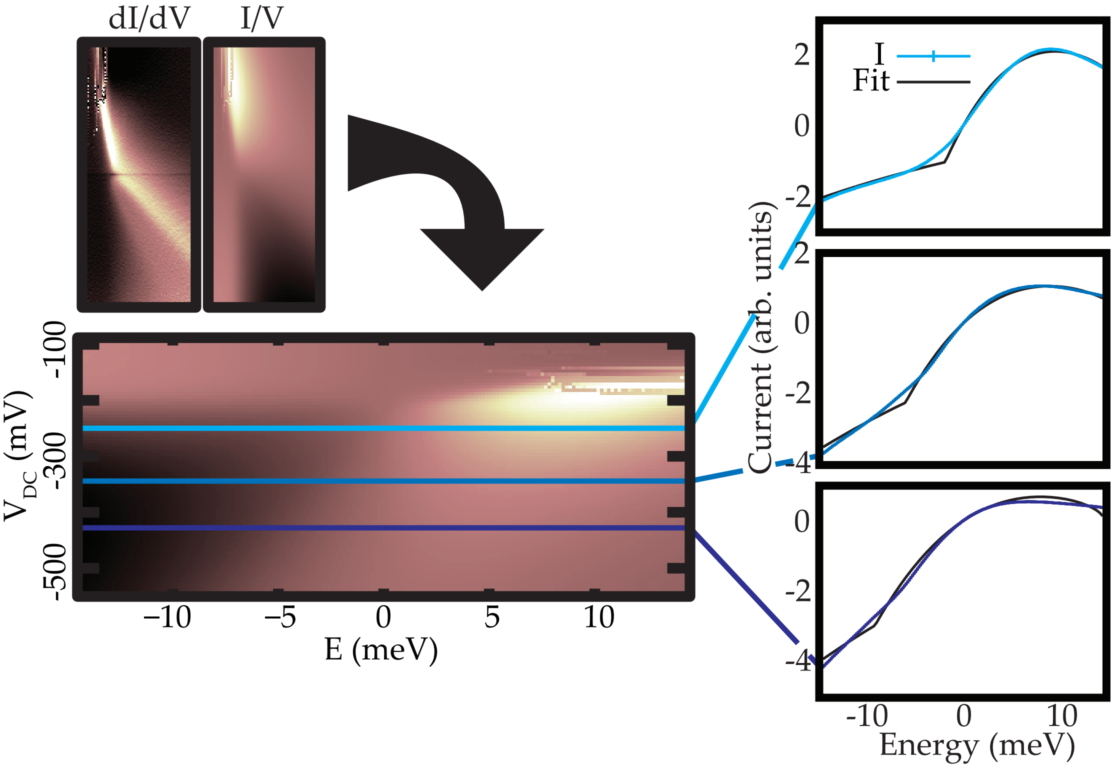


Figure 10-7: Fitting the matrix element at zero field

We can check this result by attempting to fit this current to our experimental data at zero field (Figure 10-7). There are three degrees of freedom in this fit; an overall current scale, $V_{align}$, and the 2D Fermi energy. As one might expect, we see substantial deviations at large pulse height, where higher order corrections become important, and at large density, where self-consistency corrections become significant. Nonetheless, we find good enough agreement to be confident applying the correction described above will largely remove matrix element effects from the data.

## 10.5 Extrinsic Errors

In addition to the systematic corrections from the 3D density of states and tunneling matrix elements, there are errors extrinsic to the sample originating from our pulse and amplifier quality.

During the measurement interval, within the first few hundred nanoseconds of the applied pulse, there are noticeable deviations of our pulse from perfectly flat. Although they are not large enough to substantially change the applied bias across the tunneling barrier, they couple through to our charge amplifier through the geometric capacitance of the sample, where they are indistinguishable from the signal from the displacement current. The largest deviations can be placed into two basic categories: those independent of the pulse height from our signal generator, and those proportional to the pulse height.

The pulse height independent deviations are mostly from direct coupling of the input to the pulse shaper to the output, and are thus an attenuated copy of the ringing from the HP signal generator. Those proportional to the pulse height come from reflections in the cables, and "droop" in the top of the pulse due to the bias tee. It is difficult to eliminate either of them completely, but we will see it is easy to remove them from the data in post-processing.

Linear fits of functions have properties that are very convenient in considering the consequences of these signals; if we have a set of sample points $(x_i, y_i)$, all of which have the same standard deviation, the least-squares fit to the slope $m$ is given by $m = \frac{\langle xy \rangle - \langle x \rangle \langle y \rangle}{\langle x^2 \rangle - \langle x \rangle^2}$, where $\langle q \rangle$ is the average value of $q$.

If we replace the points $y_i$ with a set of points $y'_i = y_i + \delta_i$ where $\delta_i$ is the signal from the deviation of our pulse from flatness, we can use the fact that expectation value is linear to write

$$\begin{aligned} m' &= \frac{\langle x(y+\delta) \rangle - \langle x \rangle \langle y+\delta \rangle}{\langle x^2 \rangle - \langle x \rangle^2} & (10.44) \\ &= \frac{\langle xy \rangle + \langle x\delta \rangle - \langle x \rangle \langle y \rangle - \langle x \rangle \langle \delta \rangle}{\langle x^2 \rangle - x^2} & (10.45) \\ &= m + m_\delta & (10.46) \end{aligned}$$

where $m_\delta$ is the least squares fit to the slope of the noise signal.

Using this, we can see that our measured current will be given by

$$I' = I + m_{\text{constant}} + V m_{\text{proportional}} \tag{10.47}$$

where $m_{\text{constant}}$ is the slope (with respect to time in the measurement interval) of the deviation signal that is constant with respect to the pulse height, and $m_{\text{proportional}}$ that which is proportional to pulse height.

In the absence of matrix element correction, the constant term does not appear in the $dI/dV$ spectrum, while the proportional term appears as an offset in the measured density of states. (Actually, the constant term tends to reverse sign when the sign of the pulse is changed, so it appears as a sharp spike at zero bias). However, the matrix element correction mixes them into the DOS spectrum in strange ways, and they are worth removing. This is easily done by emptying the well entirely by applying a large DC bias, fitting each polarity of the remaining I-V curve with a straight line, and subtracting this from the measured spectra before matrix element correction. We find this correction drifts very little; we use the same straight lines for months at a time, only repeating the fitting procedure when the wiring or pulse shaper are modified.

## 10.6 Summary

The spectra we take cover a range of energies and densities large enough that the 2D Fermi energy, the applied voltage across the tunneling electrode, the 3D Fermi energy, and the excitation energy to the next bound state are all of the same order. Because of this, there are substantial variations of the tunneling matrix element and 3D density of states over the range of the measurement. For a non-interacting electron gas in zero field, these are the same for all of the electrons able to tunnel at any given bias and density, and so they are easily corrected for before differentiation.

# Chapter 11

# Calibrating the 2D Density

The second axis that we need to calibrate to make TDCS a quantitative technique is the 2D density. Typically, densities in 2DEGs are calibrated using the locations of the integer and fractional quantum Hall plateaus, which are easily found within our spectra. Thus, roughly, even a raw spectrum with no labels on the x-axis can be thought of as having a fairly good density calibration. Using the Landau level crossings, we can locate integer filling fractions, and we can simply interpolate between these points.

However, the compressibility of the 2DEG is non-uniform. Inside of the chemical potential jumps between Landau levels, applied DC voltages change the density very little, but instead go towards changing the chemical potential. Inside of the Landau levels where the density of states is large, applied DC voltages cannot change the chemical potential without adding a large amount of charge, and instead change the density in the nearly metallic 2DEG. We are interested in knowing the density with sub-Landau level precision. This is important, for example, in determining the variation of the exchange gap with density within a Landau level.

Thus, if we are interested in measuring the 2DEG density accurately *in-between* the integer quantum Hall plateaus, we need to turn to more sophisticated techniques. The density, as a function of the applied DC voltage, is dependent on the 2DEG's thermodynamic density of states, and through that, on the applied magnetic field and the electron temperature. As such, we must perform separate density calibrations for each condition under which we measure.

## 11.1 Extracting Density from Capacitance

Much work has been done on extracting chemical potential and density data from equilibrium capacitance data[33, 48]. Here we present a vastly simplified discussion of this problem in which we only concern ourselves with measuring the density in the 2DEG, and closely examine the errors inherent in the technique.

The mechanisms that determine the total charge density in the quantum well are quite complicated. As we vary the electric field, the bound state energies in the quantum well shift and the 3D doped electrodes form depletion regions. In addition, the thermodynamic density of states of the electrons in the well can be complex, even showing counterintuitive and difficult to deal with features such as negative compressibility.

However, instead of dealing with all of these complexities we can think of the structure as a "black box," which changes density with applied voltage. We concern ourselves with the purely electrostatic problem of what capacitance we will measure, given the positions of the electrodes and the charge density and position of the 2DEG as a function of applied voltage. We will take advantage of the fact that any finite charge distribution is, from the outside, indistinguishable from a sheet charge layer at its center of charge,

$$\langle x \rangle = \frac{\int \rho(x) x dx}{\int \rho(x) dx} \tag{11.1}$$

This allows us to replace our somewhat messy structure with an idealized capacitor with a sheet of charge between two perfect metallic electrodes, although both electrodes and the charge sheet will move around a little bit with applied field (Figure 11-1). Note that we are not making any approximations (yet); we are simply abstracting the physics we do not want to deal with into the positions and

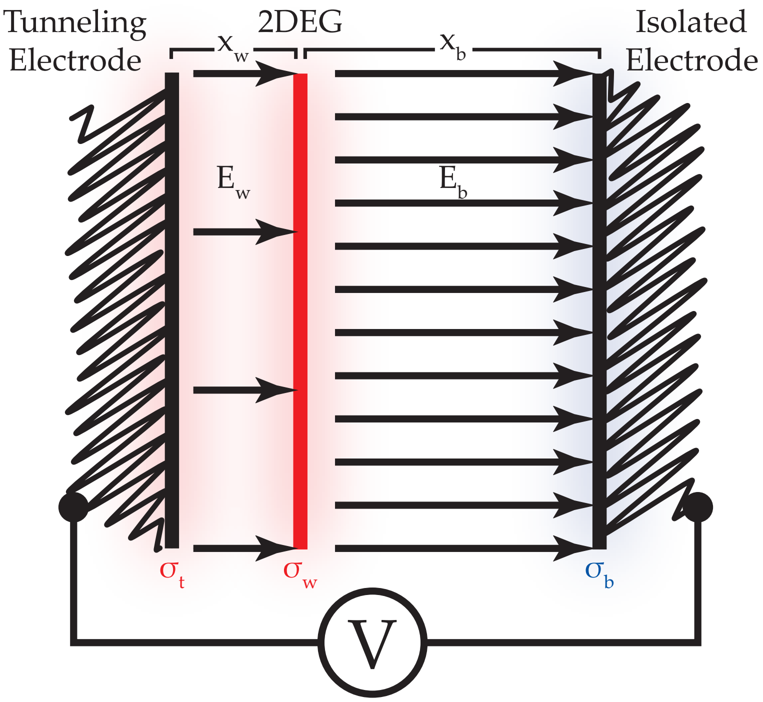


Figure 11-1: **Idealization of the device** as a sheet of charge inside a metallic capacitor

charge densities functional dependences.

Given the distances and electric fields as described in the diagram, we can write the total voltage across the structure as

$$V = E_w x_w + E_b x_b \tag{11.2}$$

Using the result $\sigma = \epsilon \Delta E$ for the change in electric field across a sheet of charge and the fact that the E field vanishes in the electrodes, we can rewrite the electric fields in terms of charge densities and solve for $\sigma_b$

$$\sigma_b = -\frac{\epsilon V + \sigma_w x_w}{x_b + x_w} \tag{11.3}$$

Recall that $C = \frac{\partial \sigma}{\partial V}$, where we use italicized $C$ to denote capacitance per unit area. The charge on the blocking electrode is the total charge on that plate of our tunnel capacitor, so

$$
\begin{aligned}
C &= \frac{\partial \sigma_b}{\partial V} && (11.4)\\
&= \frac{d}{dV}\left(\frac{\epsilon V + \sigma_w x_w}{x_b + x_w}\right) && (11.5)\\
&= \overbrace{\frac{\epsilon}{x_b + x_w}}^{\text{Geometric Capacitance}} + \overbrace{\frac{x_w}{x_b + x_w}}^{\text{Lever Arm}} \frac{\partial \sigma_w}{\partial V} + \\
&\quad \underbrace{\epsilon V \frac{d}{dV}\left(\frac{1}{x_b + x_w}\right) + \sigma_w \frac{d}{dV}\left(\frac{x_w}{x_b + x_w}\right)}_{\text{“Motional” Capacitance}} && (11.6)
\end{aligned}
$$

Equation 11.6, with its dependence on $\frac{\partial \sigma_w}{\partial V}$ forms the basis by which we extract the density from our capacitance measurements.

The first two terms are just the geometric capacitance of our structure, without the quantum well, and a charging term from charge entering and leaving the quantum well, which is the term of interest to us. The final two terms have been coined as the “motional capacitance” terms [33], and originate with the motion of charge already in the structure as the plates move around. (These are similar in nature to the well-known Kelvin probe signal, where two metallic objects with a potential between them are moved closer together and farther apart, resulting in a current proportional to the voltage between them.) We can quickly estimate the magnitude of these terms: for the first,

$$
\begin{aligned}
\epsilon V \frac{d}{dV}\left(\frac{1}{x_b + x_w}\right) &= \left(-\frac{\epsilon V}{x_b + x_w}\right)\left(\frac{\frac{\partial x_b + x_w}{\partial V}}{x_b + x_w}\right) && (11.7)\\
&= -C_{geom} V \frac{\frac{\partial w}{\partial V}}{w} && (11.8)
\end{aligned}
$$

where $w$ is the width of the structure. In the full-depletion approximation, we treat the 3D electrodes as if they are completely depleted up to a distance $t$ into the electrodes [96], at which point they become charge neutral (Figure 11-2).

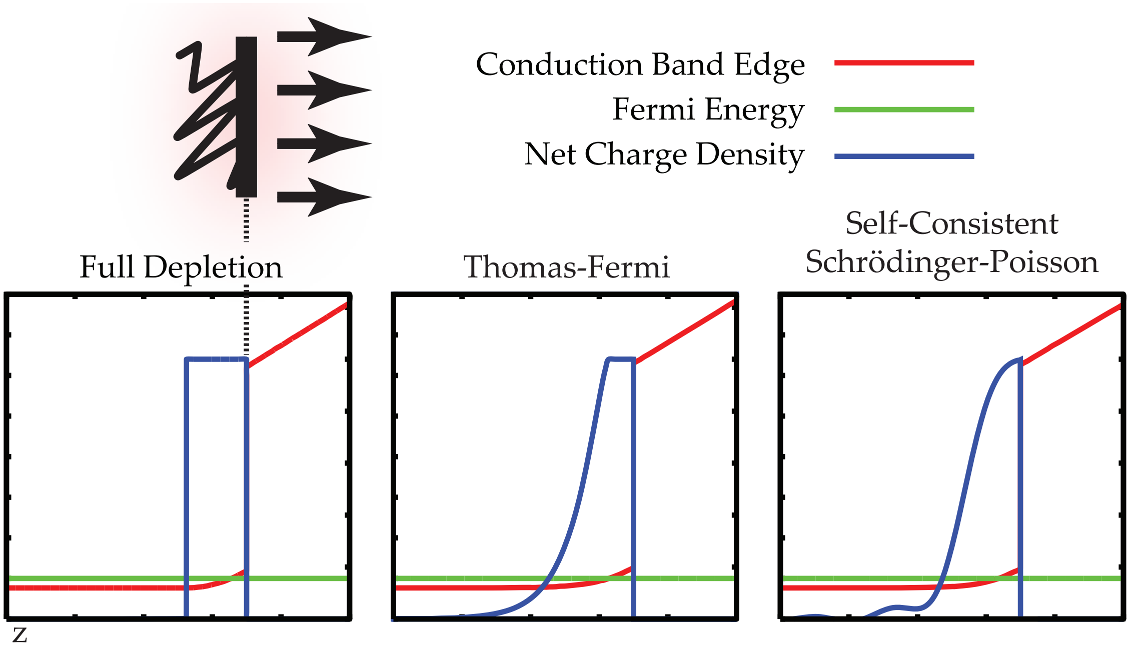


Figure 11-2: **Depletion of 3D electrodes:** Charge density and conduction band edge in a partially depleted electrode, showing the full-depletion approximation, the Thomas-Fermi approximation, and a fully self-consistent solution to the coupled Schrödinger and Poisson equations (Appendix C). Note the similar depletion depths in all three approximations, and that for a depleted electrode the Thomas Fermi approximation is quite good.

Although this is only a good approximation if the depletion width is much wider than the Fermi wavelength, it is sufficient to estimate the order of magnitude of these motional terms. In this limit, $t = \sigma/N_d$, where $N_d$ is the donor density. Thus, $\frac{\partial w}{\partial V} = 2\frac{\partial \sigma}{\partial V}/N_d = 2C_{geom}/N_d$. Using this, we find the contribution from this term to be

$$\delta_C = -C\left(\frac{2CV}{N_d w}\right) = C\frac{t}{w} \tag{11.9}$$

Thus, the fractional change in capacitance is just the fractional change in the width of the capacitor due to the applied voltage, which for a 3D density of about $10^{18}$ cm$^{-3}$ and a 2D density of about $10^{11}$ cm$^{-2}$, is about a 1 nm depletion width on a 100 nm structure, or about 2% of the capacitance.

Using our results for the polarizability of the quantum well, we can show the second term to be of similar order.

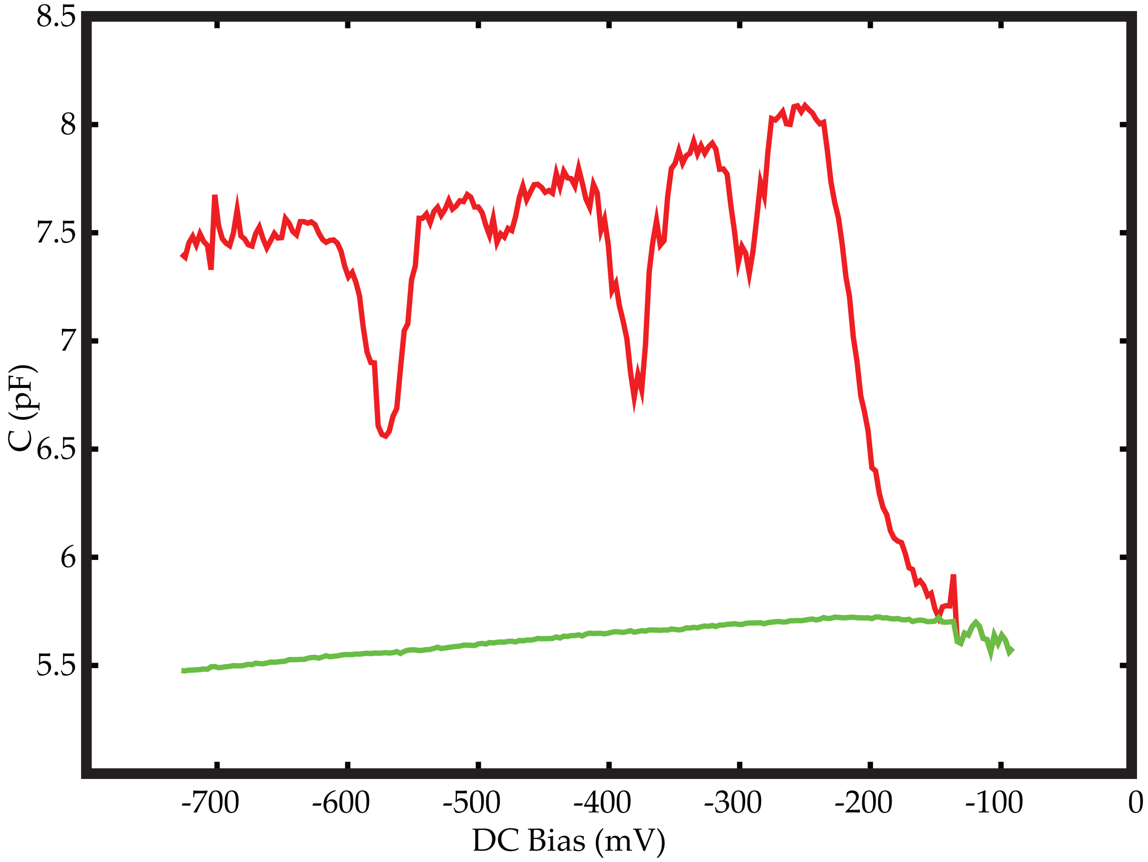


Figure 11-3: **High and low frequency capacitances measured at 3 Tesla,** showing the decrease in the low frequency capacitance in regions where the density of states is low, so the well does not charge effectively. The overall slope to the high-frequency capacitance is from depletion of the electrically isolated electrode as the well is charged.

These terms are neglected in all prior work on extracting carrier density from equilibrium capacitance measurements[48, 33]. We do so for the moment, but will revisit this approximation later.

In a typical measurement, the capacitance is measured at each DC voltage both at low frequencies, where the well has time to equilibrate fully , and at very high frequencies, where no charge has time to enter or leave the well (Figure 11-3).

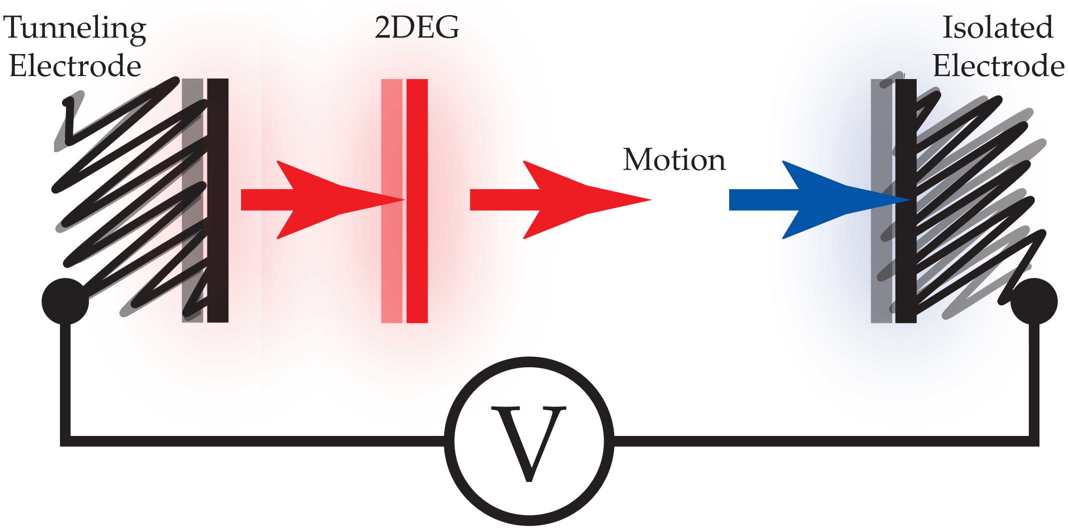


Figure 11-4: **Motion of electrodes as a voltage is applied:** As the voltage across the tunnel capacitor is changed, the 2DEG, the tunnel electrode, and the blocking barrier tend to move together, in the same direction.

Neglecting motional terms,

$$C_{low} = \frac{\epsilon}{x_b + x_w} + \frac{x_w}{x_b + x_w}\frac{\partial \sigma_w}{\partial V} \tag{11.10}$$

$$C_{high} = \frac{\epsilon}{x_b + x_w} \tag{11.11}$$

$$\tag{11.12}$$

Note that for each applied voltage, we have two measurements, but 3 unknowns; $x_b$, $x_w$, and $\frac{\partial \sigma_w}{\partial V}$. While it would be in principle possible to construct a sophisticated model for these to attempt to extract them from the observed data, there are several simplifying assumptions that typically cause errors smaller than the assumptions introduced by neglecting the motional terms.

Ashoori notes that typically, as we apply a voltage across the structure, the quantum well and the 3D spacer both move in the same direction at about the same rate (for our structure, the 3D depletes at a rate of about 3 nm/V, while the quantum well polarizes at a rate of about 8 nm/V Figure 9-4). Based on this, we treat $x_w$ as fixed as a function of voltage, while $x_b$ varies. Using this, $C_{high}$

becomes a measurement of $x_b + x_w$, and we can write

$$\frac{\partial \sigma_w}{\partial \mathrm{V}} = \left( \frac{C_{low}}{C_{high}} - 1 \right) \frac{\epsilon}{x_w} \tag{11.13}$$

By comparison, Dolgopolov [48] assumes the geometric lever arm $\dfrac{x_w}{x_b + x_w}$ is fixed, giving the simple result

$$\frac{\partial \sigma_w}{\partial \mathrm{V}} = \frac{x_b + x_w}{x_w} \left( C_{low} - C_{high} \right) \tag{11.14}$$

This approximation is somewhat hard to justify, but has the advantage that the lever arm is a simple quantity to discuss, and that it provides the simple result that the change in charge density is proportional to the area between the low and high frequency capacitance curves.

Each of these leaves two parameters to be calibrated. One is just the density at a single bias point to establish where "zero density" is. The other is, in the case of Equation 11.13, $\frac{\epsilon}{x_w}$, and in the case of Equation 11.14, $\frac{x_b + x_w}{x_w}$. These are just simple constants of proportionality for extracting the density. These are easily calibrated by forcing the density at the chemical potential dips in the spectrum to have densities that differ by the known Landau level degeneracy $D = eB/h \approx \mathrm{B} \cdot 2.54 \times 10^{10} \mathrm{cm}^{-2}$, either over the entire range of the capacitance spectrum, or using a piecewise linear calibration inside each Landau level.

Unfortunately, on using either of these calibration approaches, we find that the density at large depleting voltages is negative (Figure 11-5). A first thought on explaining this discrepancy might be to examine the motional capacitance that we neglected. We estimated its magnitude at a few percent of the total capacitance, so although it would be a stretch to explain a 10% error in the total charge of $\nu = 1$, it is conceivable for our rough estimate to be off by a factor of 5.

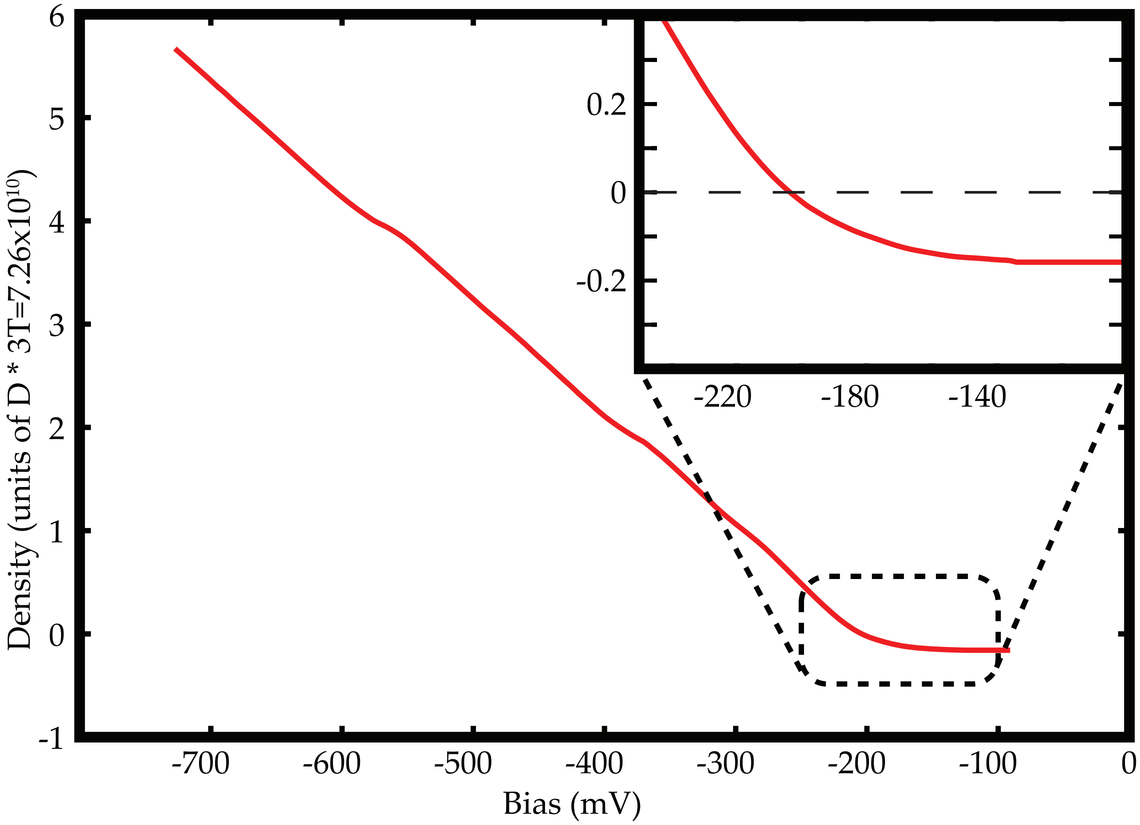


Figure 11-5: **Extracted density at 3 Tesla:** Change density in the quantum well extracted as a function of applied gate voltage using Equation 11.13 above. There is no noticeable difference if Equation 11.14 is used instead. Note that at large depleting voltages (inset) the extracted density is negative.

## 11.2 Corrections due to Motional Capacitance

To examine this possibility, we simulate the low and high frequency capacitances of our structure using our self-consistent simulation code (Appendix C). Because we have access to the charge density as a function of position in these simulations, they provide us with a test-bed for different algorithms for extracting density from capacitance measurements, including algorithms that have "oracular" knowledge of the depletion of the electrodes we have no access to in the real experiment.

To extract the low frequency capacitance, we simulate the structure at equilibrium at a variety of DC bias voltages, and compute $d\sigma/d\mathrm{V}$ on the electrically isolated electrode. To extract the high frequency capacitance, for each bias in the DC simulation, we allow the chemical potential in the quantum well to float while holding the charge density fixed, and simulate the structure again with a slightly different applied bias. Recomputing $d\sigma/d\mathrm{V}$ then gives us the AC capacitance of the structure. Alternatively, we can compute the center of charge-locations of the two electrodes and the quantum well, and evaluate Equation 11.10 and Equation 11.11 to find the capacitance neglecting the motional terms. The difference between these results shows the relative importance of the motional capacitance at different points in the density sweep (Figure 11-6).

Notice the motional capacitance is quite small where the DOS in the well is small, between Landau levels. In these regions, the tunneling and electrically isolated electrodes will tend to shift together so the separation between the plates of the capacitor only changes slowly. On the other hand, where the well screens the applied electric field effectively, the motional capacitance gives a large positive contribution. The net effect of this will be to tend to overestimate slightly the amount of charge in each Landau level when neglecting the motional terms. At large depleting voltages there is also a sizable contribution from the motional terms; in this region, the location of the tunnel electrode has been pulled up to and pinned against the tunnel barrier. The isolated electrode continues to move, so the overall width of the capacitor changes rapidly. Here, the tunnel capacitor

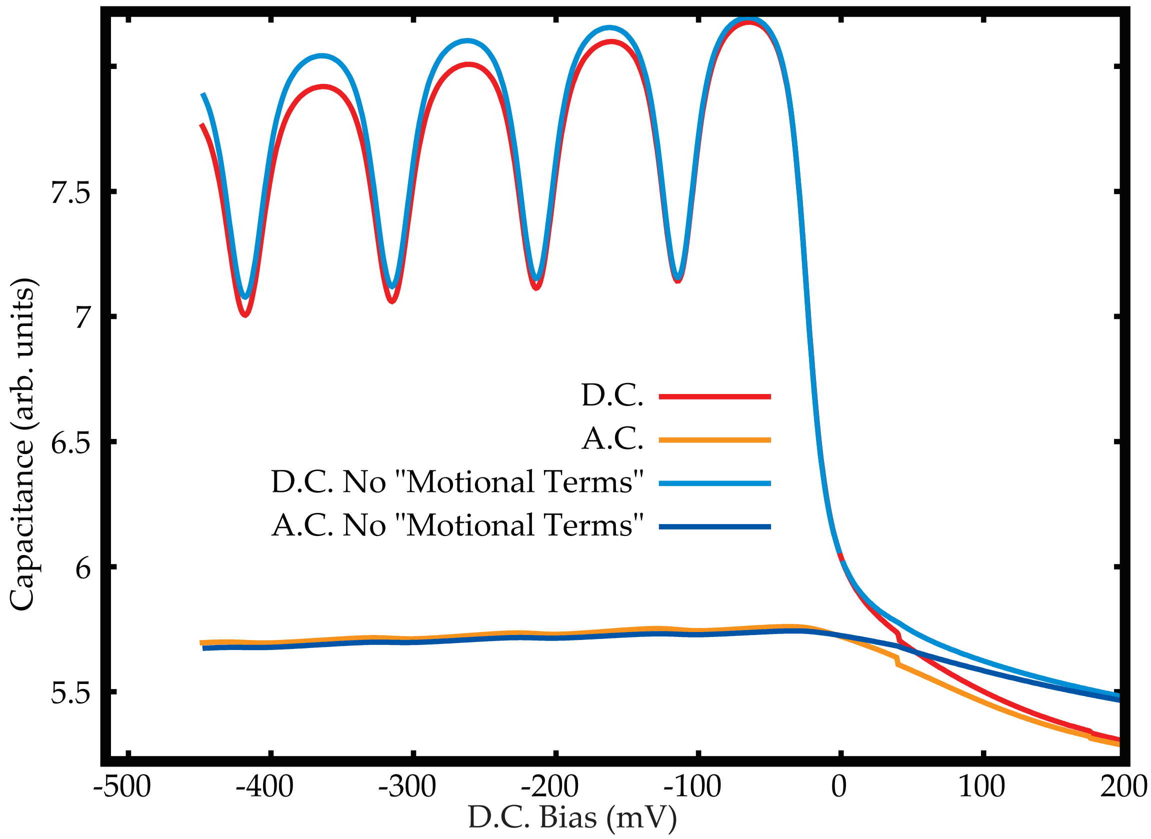


Figure 11-6: **Simulated capacitance curves at 3 Tesla:** Note the locations where the motional terms give a large contribution the capacitance. This simulation uses "spin-less" electrons (or equivalently has exchange splitting equal to $\hbar\omega_c$).

acts like a heavily biased abrupt varactor [96], and neglecting the motional terms gives large errors. Fortunately, the error is largely a uniform shift of the AC and DC capacitances.

As one might expect from our analysis of the magnitude of the motional capacitance, it is most important where the sample is under large bias (notice that the gap between the curves with and without the motional capacitance grows at larger bias voltages). Because of this, we will assign anomalously small total densities to high numbered Landau levels, and approximately correct densities to low level numbered Landau levels. When we then calibrate the lever arm using the entire spectrum, this will result in the total density in the lower Landau levels exceeding the Landau degeneracy, and give us the kind of error we see in our extracted densities.

We can check this idea by extracting densities from our simulated capacitance data, and comparing the resulting densities to the known density in the quantum well from the simulation. In order to simplify calibration of the lever arm and density offset, we simply pick them to set the density error to zero at two points. Note that we could substantially improve our calibration by using a different value for these parameters in each Landau level, but this choice will suffice for the purposes of comparison. Figure 11-7 shows the error in the extracted densities.

The black trace is extracted using Equation 11.6, and full knowledge of the positions of both electrodes and the quantum well provided by a hypothetical oracle. The variation in this trace from zero reflects the finite precision of the simulation. In the blue trace, we once again use oracular knowledge of the positions of the electrodes, but neglect the motional terms. Finally, in the red and green traces we use Equation 11.13 and Equation 11.14 respectively, with no oracular knowledge. Interestingly, the errors introduced by Ashoori's assumptions partially cancel the errors from neglecting the motional terms, and it gives noticeably better results than even the "oracular" case.

The most striking feature of this plot is the scale of the errors; they are at most $5 \times 10^9/\text{cm}^2$, or an error of about 0.06 of a Landau level, even using this simple-

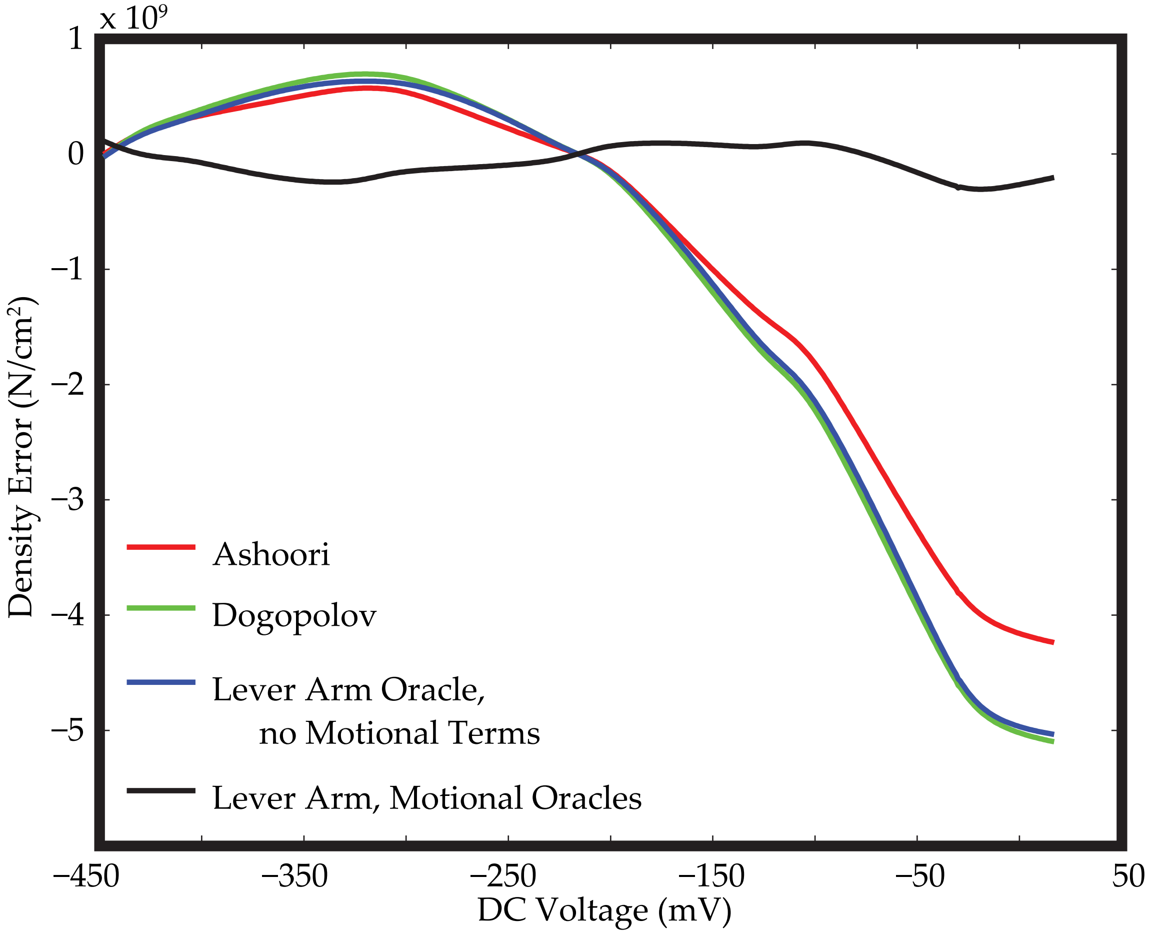


Figure 11-7: **Simulated extracted densities at 3 Tesla:** Densities extracted from simulated capacitance data in Figure 11-6. See text for a description of techniques.

minded calibration of the density offset and lever arm. This is some three times smaller than our observed error. We are forced to conclude that while motional corrections are noticeable, they are not large enough to explain our "negative" filling fractions.

## 11.3 Calibration Errors

Thus, we conclude that given a moderately good calibration of the lever arm and density offset, either of the approaches described should give much better results than we find in our data. It then behooves us to consider the assumptions that went into this calibration.

First, we assumed that the minima our capacitances traces correspond to the points where the chemical potential lies midway between two Landau levels. One could imagine this assumption not being met because, for example, the motional capacitance systematically pulls the minima in the capacitance slightly to one side of the DOS minima. This would be surprising given that we know the contribution of the motional capacitance is small in the dips in the capacitance. We nonetheless examine the possibility by once again taking advantage of our ability to extract parameters from our simulation that we cannot physically measure. In this case we plot the capacitance versus the chemical potential in the quantum well $\mu$ (Figure 11-8), and find the capacitance minima do correspond to the minima in the DOS.

Secondly, we assumed that the density changes by the Landau degeneracy, $\mathrm{D} = \mathrm{B} \cdot 2.54 \times 10^{10}/(\mathrm{cm}^2\ \mathrm{Tesla})$, in between each one of these dips. Assuming Lorentzian line-shapes for the Landau levels with reasonable widths for our Landau levels, we can easily integrate our DOS to find the total density as a function of chemical potential (Figure 11-9).

On so doing, we notice that at high densities, the densities at the chemical potential minima do indeed correspond to densities that are an integer multiple of D. However, at low densities, the actual density in the well can be significantly

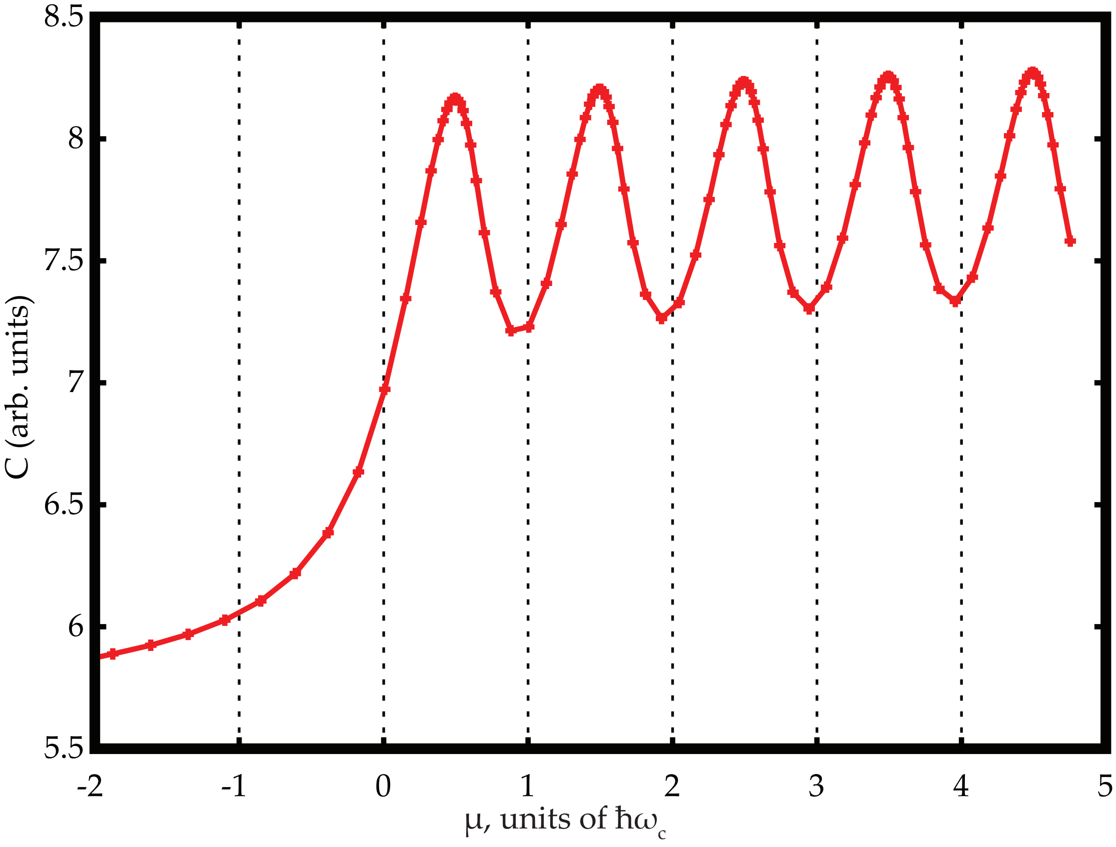


Figure 11-8: **Simulated capacitance as a function of chemical potential** $\mu$ in the quantum well. Note that to within the precision of the simulation, the capacitance dips fall precisely on the chemical potential minima.

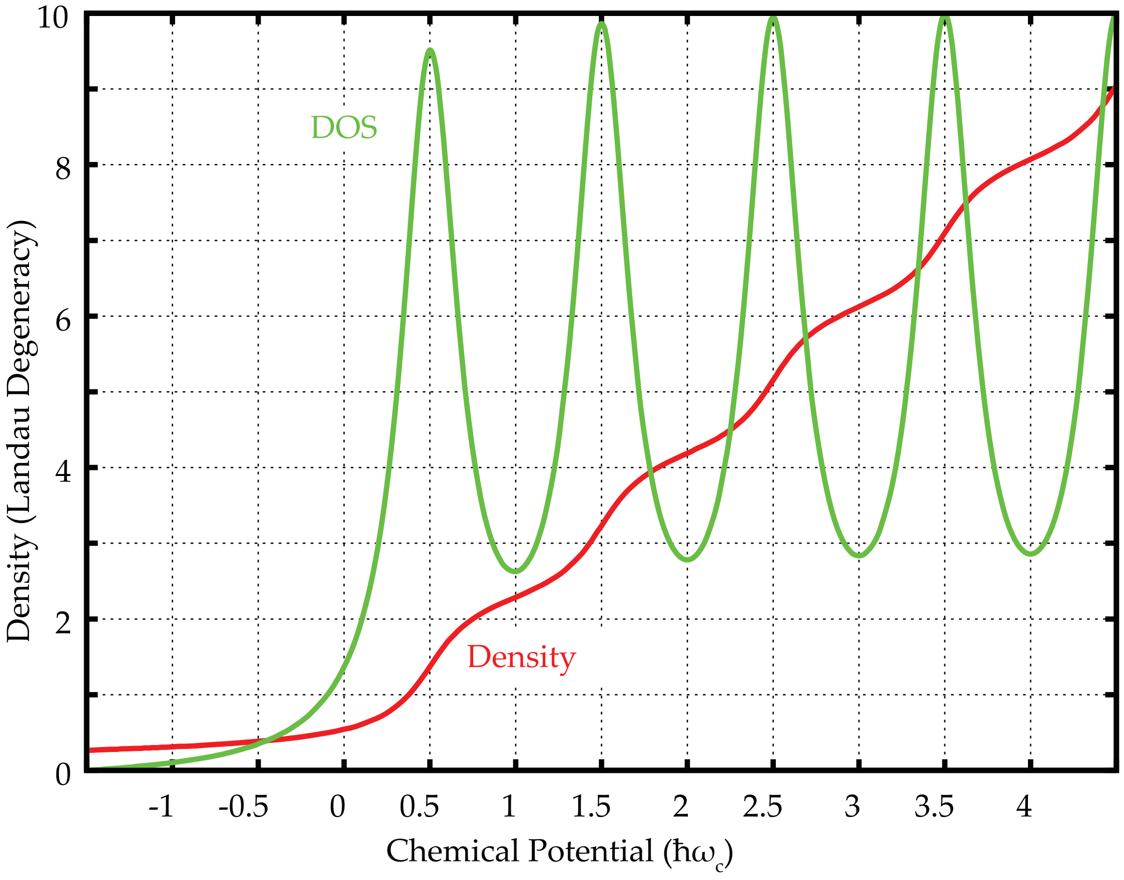


Figure 11-9: **Total density and DOS as a function of chemical potential:** Density is measured in units of the Landau level degeneracy, and the levels have HWHM of around 0.5 $\hbar\omega_c$.

larger; in this case, at filling fraction $\nu = 2$, the density is almost 10% more than 2D. The cause is clear; the long tails of the higher Landau levels are substantially raising the density. At large filling fractions, the missing density from the high energy tails of the lower Landau levels is equal to the extra density contributed from the low energy tails of higher Landau levels. As we reduce the filling fraction, the amount of "missing" density decreases, as there are fewer low-lying Landau levels whose tails are being lost, while the amount of extra density stays the same, resulting in higher densities than expected. Ultimately, when $\mu = 0$, only the extra density from the tails of the Landau levels remains, and we are left with finite density at the last (imaginary) Landau level minimum at $\mu = 0$.

This means that the *actual density change* in going from $\nu = n$ to $\nu = n + 1$ is, in fact, larger than the Landau level degeneracy. Because we fix our density offset by assuming the density at these dips is given by $\nu$D, this means should expect to extract a negative density for a heavily depleted well; the last Landau level has a substantial amount of "extra" charge. This effect has the correct sign to explain our offset, and the correct magnitude. We expect this effect to become small at large magnetic fields where the cyclotron energy separates the Landau levels well, but we have very little high field data yet.

## 11.4 Applying the Calibration to the Spectra

It is now a straightforward matter to distort our spectra so that the $x$ axis is linear in density instead of voltage (Figure 11-10). In general, the compressibility of the 2DEG is high, so the density is a linear function of gate voltage, and the two spectra are quite similar. However, the density calibration makes the transitions between Landau levels substantially steeper, as well as the region where the well is initially filled. This calibrated data is more useful than the uncalibrated for quantitative comparisons with theory, especially when considering the shape of the exchange gap as a function of filling.

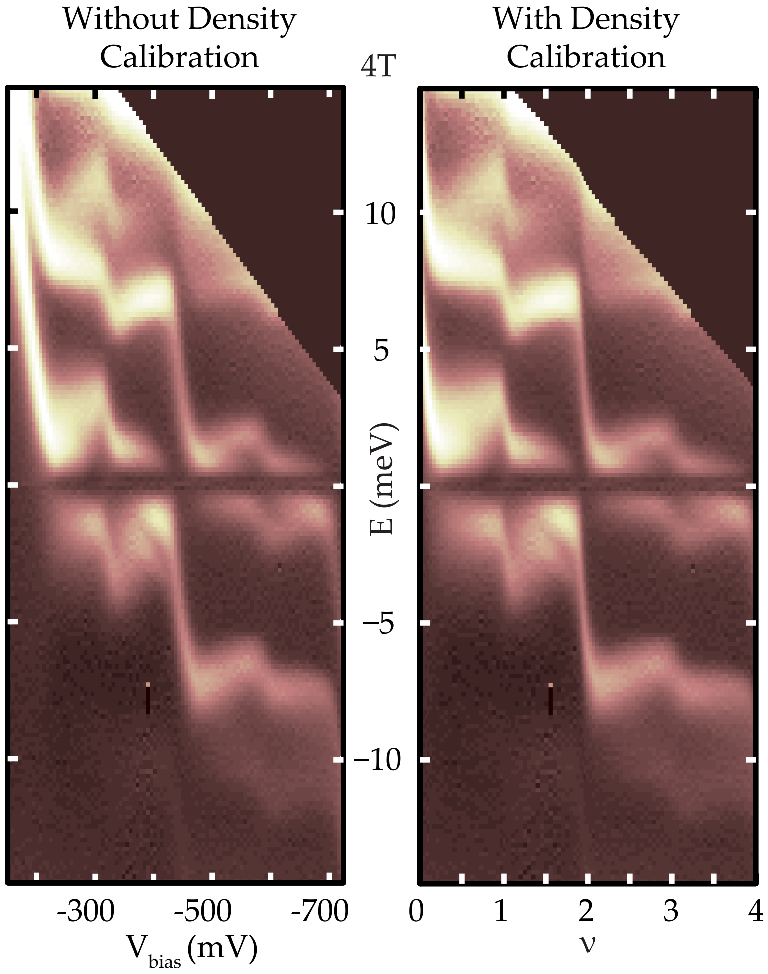


Figure 11-10: **A comparison of TDCS spectra with and without the density calibration:** The transitions between Landau levels and filling of the well are made significantly sharper by the density calibration.

## 11.5 Implications

The periodic structure imposed by Landau levels is the current accepted "gold standard" for determining the density in 2D systems. Landau level broadening tends to cause this technique to *underestimate* the 2D density at low filling fractions. This may be of some importance in determining the density of very low density samples where only a few SdH oscillations are visible. However, most measurements are performed at fixed density by varying magnetic field, and averaging across many quantum Hall plateaus should greatly reduce the impact of this effect, as will the fact that the lowest Landau levels are only visible at high magnetic fields, where the densities should be similar to those expected in the limit of infinitely sharp Landau levels.

Within TDCS, however, it introduces an interesting ambiguity. We could choose to label our plots with absolute densities, using measured line-widths to estimate the actual density at each minimum of the DOS. However, the physics we are concerned with is largely a function of which Landau level we are populating, not the exact density. As such, we choose to label the density axis of our plots in filling fraction $\nu$, with a scale that is linear in density. This captures the density dependence of our data, marking the most important points nicely. This does cause our data to extend to slightly negative $\nu$ at some points, reflecting the fact that the density in the lowest Landau level exceeds D slightly.

## 11.6 Stability

Typically, we accumulate DOS spectra under any given set of conditions for a few weeks or even months. As such, a substantial amount of time can pass between the acquisition of the spectrum and the acquisition of the capacitance traces for density calibration. During this time, any drift in the sample, perhaps through changes in donor ionization, drift in the DC bias supplies, or even changes in the sample leakage or gate leakage of our cryogenic amplifier can change the density of the 2DEG for any given bias voltage applied at room temperature. If such drifting has occurred, it is visible as an offset between the compressibility dips as seen in the capacitance traces and the chemical potential jumps in the pulsing traces. In general, with the feedback on the drain current on, our setup appears to be remarkably stable as long as the sample space temperature is constant. Thermally cycling the sample space, even just up to 1K and back, seems to cause a noticeable drift. It is unclear if this drift is related to sample leakage, changes in donor ionization, or changes in the cryogenic amplifier. However, our refrigerator is highly reliable and easily able to maintain a constant temperature over several months, so this is not a serious concern.

## 11.7 Summary

We have established a calibration technique for transforming the DC bias axis of our spectra to density. It has systematic errors of order 5%, largely from the neglected motional capacitance terms. Close examination of the density as a function of bias has revealed an unexpected increase in the total change in charge density between filling fractions at low Landau level indices that originates with the long Lorentzian tails of high-lying levels.

# Chapter 12

# Calibrating the Energy

The voltage difference between the 2DEG and the nearby electrode is, at the moment the pulse is applied, given by the height of the applied voltage pulse times a geometric lever arm given by the ratio of the distance from the quantum well to the tunnel electrode to the total thickness of the structure (Figure 3-6).

The real world is never quite so simple. Because the Fermi wavevector in the electrodes is finite, there is no sharp edge from which to measure the thicknesses. Instead, there is a finite 1D charge distribution. As we learned in Chapter 9, the 2D electron gas has a bound state energy and center-of-charge that both shift noticeably with applied voltage. Even putting aside these variations of the lever arm, we need to find a feature in the spectrum whose energy spacing we know with enough precision to use it to calibrate our energy scale; a "standard candle."

In this chapter, we will briefly discuss our choice for this standard candle. We will then quantitatively estimate the magnitude of the variation of the lever arm within the range of our experiment. Finally, we will present some numeric modeling of the geometric lever arm that will provide a more accurate estimate of the lever arm variability.

## 12.1 A Standard Candle

In astronomy, the use of standard candles provides a simple way of determining the distance to an object. If one has, a priori, a good estimate of the brightness of an object, one can use its observed brightness at the Earth to estimate the distance to it. The stellar bodies for which good a-priori estimates are available are known as standard candles. However, this leaves the difficult problem of identifying objects whose brightnesses are accurately known.

Similarly, in our spectra we see a complex interplay between screening, disorder, interactions, and simple single-particle physics. As of this writing, there are no theories that accurately reproduce all the features we see in our spectra; there are very few features whose spacing we know from first principles.

Historically, the cyclotron energy has been used to calibrate measurements of chemical potential and energy spacings in thermally activated conductivity and capacitance measurements [33, 48, 42]. However, as we have learned, the exchange interaction modifies the cyclotron gap between Landau levels separated by changes in the orbital quantum number (for example, between $\nu = 2$ and $\nu = 3$) substantially. Because of these many body effects, we cannot use these cyclotron gaps to calibrate our energy scale.

However, in TDCS we have a region of spectrum available that is inaccessible in traditional capacitance and conductivity measurements: the area where the quantum well is empty. In this region, we have effectively turned off electron-electron interactions; there are no carriers with which the injected charge can interact. In this region, we do expect the cyclotron gap to be equal to $\hbar\omega_c$. This part of the spectrum forms our standard candle.

We use relatively low magnetic fields to calibrate our lever arm. This allows us to see many Landau levels in the empty well, and provides a good baseline for fitting. We find the Landau levels to be evenly spaced over the range from 0 to 55 meV that we can measure, to within experimental precision (Figure 12-1). While there is some question that the experimentally available cyclotron mass

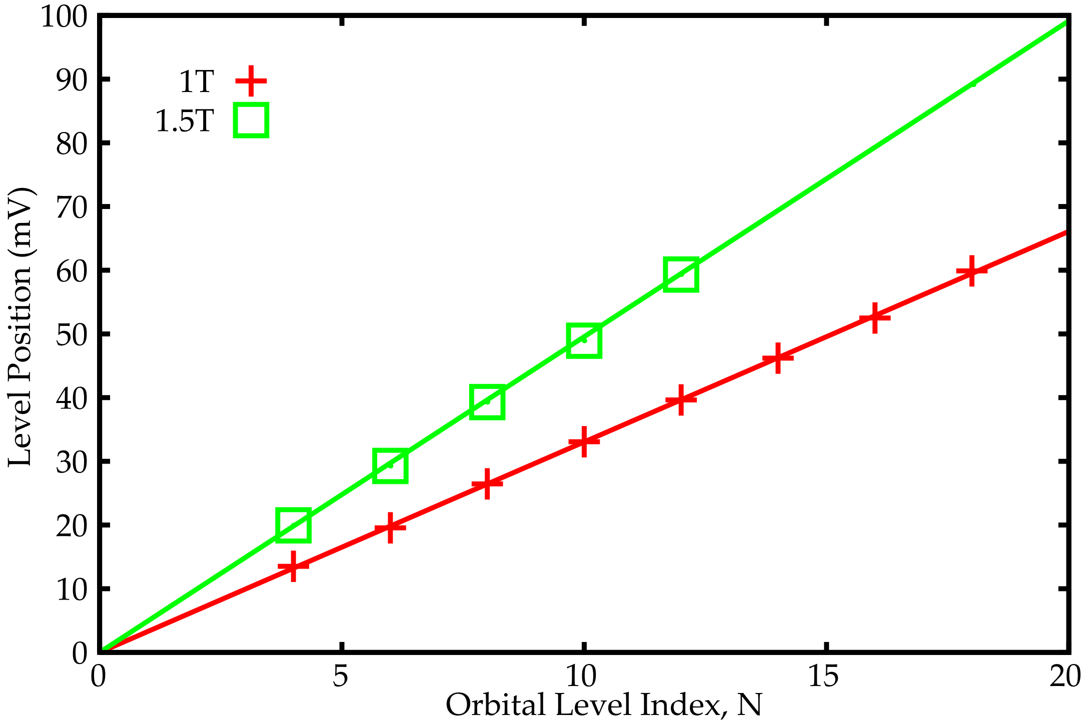


Figure 12-1: Fitting the cyclotron energy in an empty well

$m^* = 0.067$ is applicable when injecting carriers into a completely empty band[61], this at worst renormalizes our energies by a few percent, on the same order as our expected systematic errors. For our current sample, we extract a lever arm $\lambda = 0.267$.

## 12.2 Variations in the Lever Arm; The Quantum Well

Everything in our structure squishes around as we apply AC and DC voltages across it. We will find that motion of the 3D electrodes dominates lever arm variations, although the polarization of the quantum well also gives noticeable corrections.

We only care about shifts in the bound state energy in response to DC voltages to the extent to which they change the response to AC voltages. Unlike attempts to extract the chemical potential from equilibrium capacitance measurements [33], we do not need to keep track of the change in the bound state energy as we vary the DC bias. This simplifies our problem enormously, as the DC response is complicated by the screening of the applied voltage by the 2DEG itself.

We have already calculated the bound state energy shift of the quantum well with applied electric fields (Chapter 9). Recall that by measuring energies from the center of the quantum well, the linear shift vanished, and we were left with a quadratic term. The largest AC field we apply is $\sim$55 mV/100 nm, giving us a bound state energy shift of 100 $\mu$eV. Since our lever arm is about one quarter, this is equivalent to changing the pulse height by .4 meV, or an energy error of less than 1%. Note that this bound state energy shift includes the polarization of the quantum well. This AC shift does not depend on the density of states in the quantum well; it occurs before any charging happens. Finally, because it originates with a quadratic correction, this error vanishes as we approach the Fermi energy; at energies of 5 meV, it will give a contribution of less than a part in a thousand. Because our tunneling matrix elements typically vanish well before we reach a full 15 meV above the Fermi energy, we have substantially overestimated the practical importance of this error.

We apply DC fields substantially larger than the (55 mV)/(100 nm) AC fields we typically use; we bias until the sample starts to leak appreciably at around 300 mV applied voltage. Neglecting screening by the 2DEG, this DC field moves us to a region on the bound state energy curve in which we are much more sensitive to applied electric field (Figure 12-2). This is not surprising; the quantum well is symmetric, so we know that the bound state energy shift as a function of electric field must have a local extrema at zero electric field. We see that at 300 mV, the bound state energy shift from a 0 mV pulse to a 55 mV pulse is around 1.2 meV. This bound state energy shift can be linearized in the AC pulse height, so we can interpret it as an effective 8% increase in the lever arm caused by the applied DC bias. A simple intuitive explanation of this increase is that the applied DC bias tends to pull the 2DEG closer to the isolated electrode, increasing $x_w$ in calculating the geometric lever arm. It will turn out that we have enormously overestimated

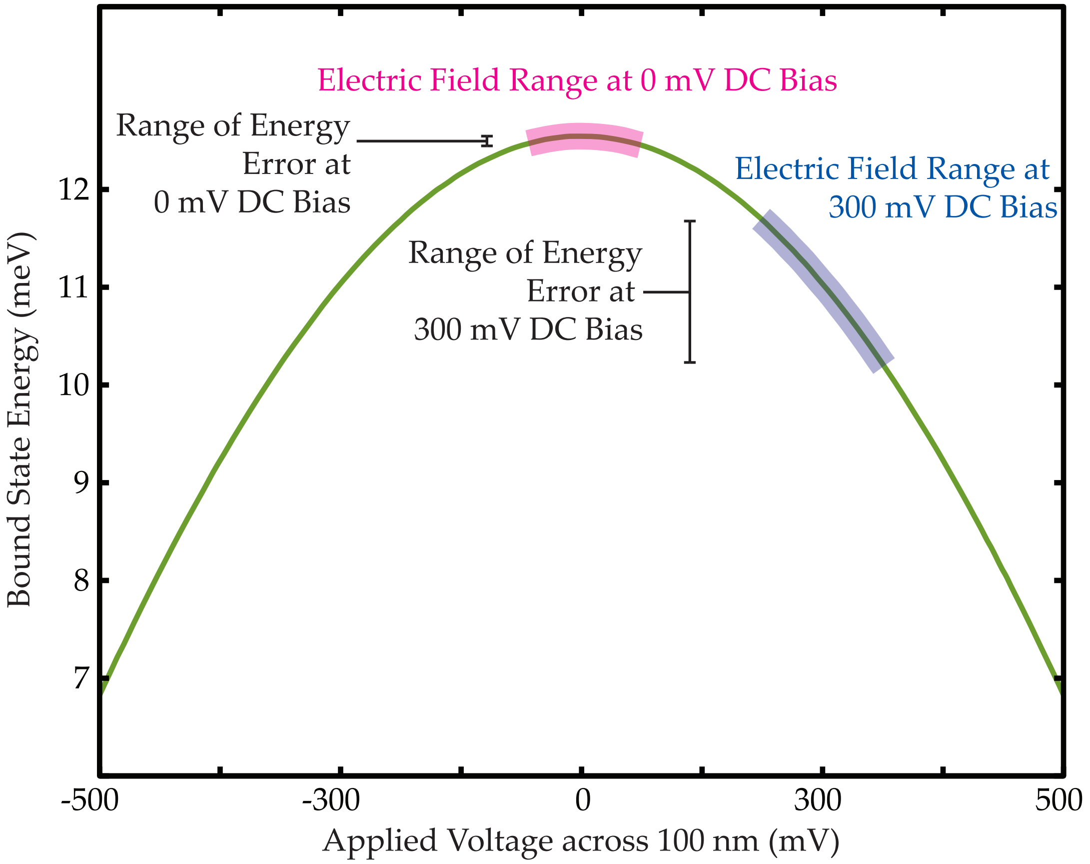


Figure 12-2: **Combining the AC and DC voltages in the quantum well polarization:** The highlighted areas show the range of electric fields induced by pulses from -55 mV to 55 mV at DC biases of 0 mV (pink) and 300 mV (blue). As the DC bias is increased, we see larger energy errors due to the bound state energy shift. Results shown are the exact shift for a 21.15 nm infinite well, which are appropriate for a 17.5 nm finite well (Chapter 9).

this number; we have neglected the screening of the DC field by carriers in the well, which tends to make the quantum well look rectangular, and as such flatten the parabola in Figure 12-2. This screening becomes more complete as the DC bias and density in the quantum well are increased. This will be treated more exactly numerically at the end of this chapter. It will also turn out that, as discussed in Chapter 11, the quantum well and the tunneling electrode tend to shift together and in the same direction. Because of this, the depletion of the 3D electrode will partially cancel the lever arm shift due to the polarization of the quantum well.

## 12.3 Variations in the Lever Arm; DC Motion of the Electrodes

If we examine the high frequency capacitance of our tunnel capacitor, we typically find that it is peaked at around zero bias, and drops steadily off as we both enhance and deplete the 2DEG (Figure 12-3). This high frequency capacitance is, neglecting motional terms, a measure of the total thickness of the structure from the isolated electrode to the tunneling electrode. The quantum well itself is unimportant in the high frequency capacitance.

At large enhancing (negative in Figure 12-3) voltages, the high density of states in the quantum well screens the tunnel electrode from the applied DC electric field, so that electrode is stationary as we vary the DC voltage. The isolated electrode, on the other hand, feels the entire applied electric field, and moves as we vary the voltage. In this domain, changes in the high frequency capacitance reflect motion of the isolated electrode.

Similarly, at large depleting voltages (near zero volts in Figure 12-3), the 2DEG is completely empty, and so it no longer screens the electric field. Because of this, the tunnel electrode depletes as we vary the applied voltage. This electric field also enhances (increases the negative charge on) the isolated electrode. This enhancement, however, does not actually change the effective position of the

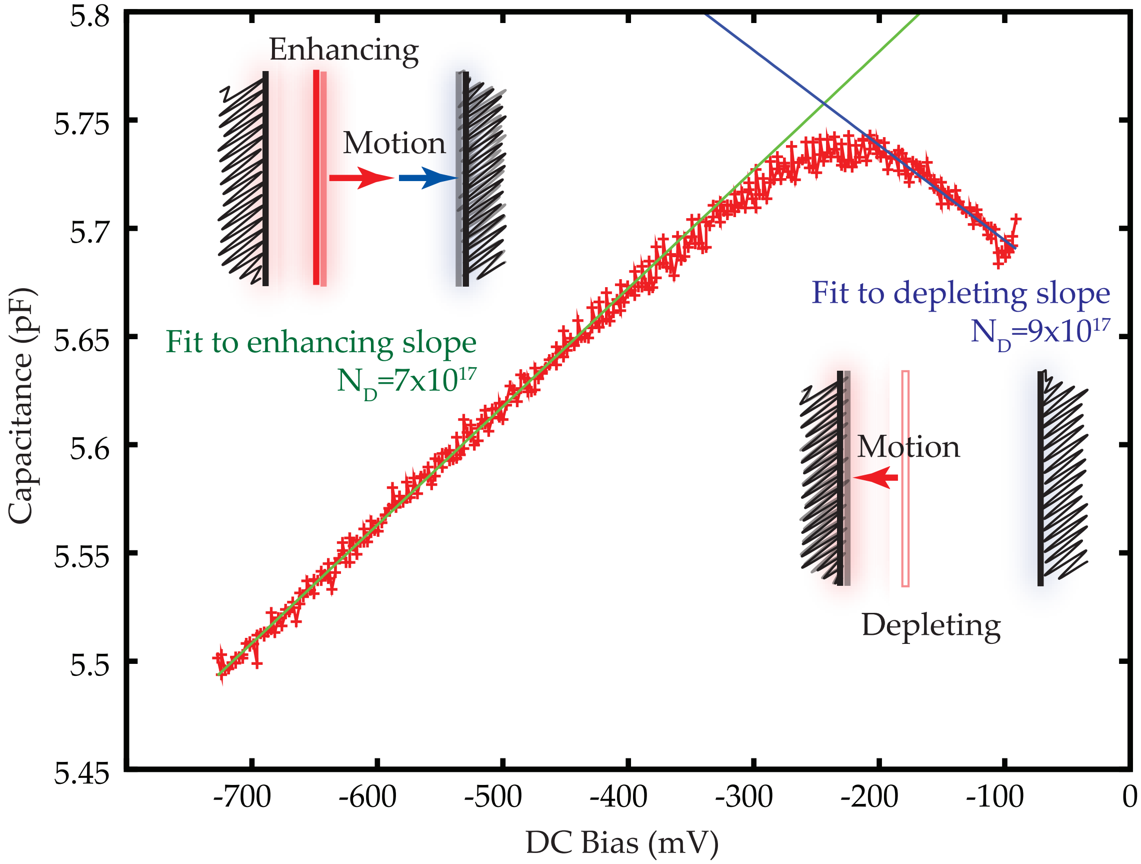


Figure 12-3: **High frequency capacitor of one of our samples** at 100 mK, 0 T. At large negative voltages, the 2DEG has finite density and screens the electric field from reaching the tunneling electrode, so changes in capacitance reflect the donor density in the isolated electrode. Similarly, near zero voltage, the isolated electrode begins to enhance near the blocking barrier and is unable to move, so changes in capacitance reflect the donor density in the tunneling electrode.

isolated electrode; it is pinned approximately 1/4 Fermi wavelength away from the blocking barrier. Thus, for depleting voltages, changes in the high frequency capacitance reflect motion of the tunnel electrode.

We can use the slope, $d\mathrm{C}/d\mathrm{V}_{\mathrm{DC}}$ of the high frequency capacitance curve to extract experimentally the donor density in the electrodes. Recall that in the full depletion approximation the depletion thickness $t$ is given by $t = \sigma/N_d$, where $N_d$ is the donor density and $\sigma$ is the induced sheet charge. As the donor density increases, it takes a larger sheet charge to move the edge of the electrode the same distance; this makes the electrode "stiffer", and $d\mathrm{C}/d\mathrm{V}_{\mathrm{DC}}$ smaller.

More precisely, working in capacitances per unit area ($\tilde{\mathrm{C}}$), and assuming we are considering small changes in capacitances away from a value $\tilde{\mathrm{C}}_0 = \epsilon/d$ caused by a change in the capacitor thickness $d \to d + \delta_d$,

$$\begin{aligned}
\tilde{\mathrm{C}}(\mathrm{V}) &= \frac{\epsilon}{d+\delta_d} && (12.1)\\
&= \frac{\epsilon}{d}\frac{1}{1+\delta_d/d} && (12.2)\\
&= \tilde{\mathrm{C}}_0\left(1-\frac{\delta_d}{d}\right) && (12.3)\\
&= \tilde{\mathrm{C}}_0\left(1-\frac{\sigma}{N_d d}\right) && (12.4)\\
&= \tilde{\mathrm{C}}_0\left(1-\frac{\tilde{\mathrm{C}}_0\mathrm{V}}{N_d d}\right) && (12.5)\\
\frac{\partial\tilde{\mathrm{C}}}{\partial\mathrm{V}} &= \frac{-\tilde{\mathrm{C}}_0^2}{N_d d} && (12.6)\\
N_d &= \frac{\tilde{\mathrm{C}}_0}{\frac{\partial\tilde{\mathrm{C}}/\tilde{\mathrm{C}}_0}{\partial\mathrm{V}}d} && (12.7)\\
&= \frac{\epsilon}{\frac{\partial\tilde{\mathrm{C}}/\tilde{\mathrm{C}}_0}{\partial\mathrm{V}}d^2} && (12.8)
\end{aligned}$$

giving us our desired expression for the doping density. For a 100 nm thick structure in GaAs, this works out to $N_d = 7.2 \times 10^{16}\ \mathrm{cm}^{-3}/\frac{\partial\tilde{\mathrm{C}}/\tilde{\mathrm{C}}_0}{\partial\mathrm{V}}$, with V measured in volts. For the data in Figure 12-3 taken on our sample, this gives densities of $7 \times 10^{17}\ \mathrm{cm}^{-3}$ for our isolated electrode and $9 \times 10^{17}\ \mathrm{cm}^{-3}$, very close

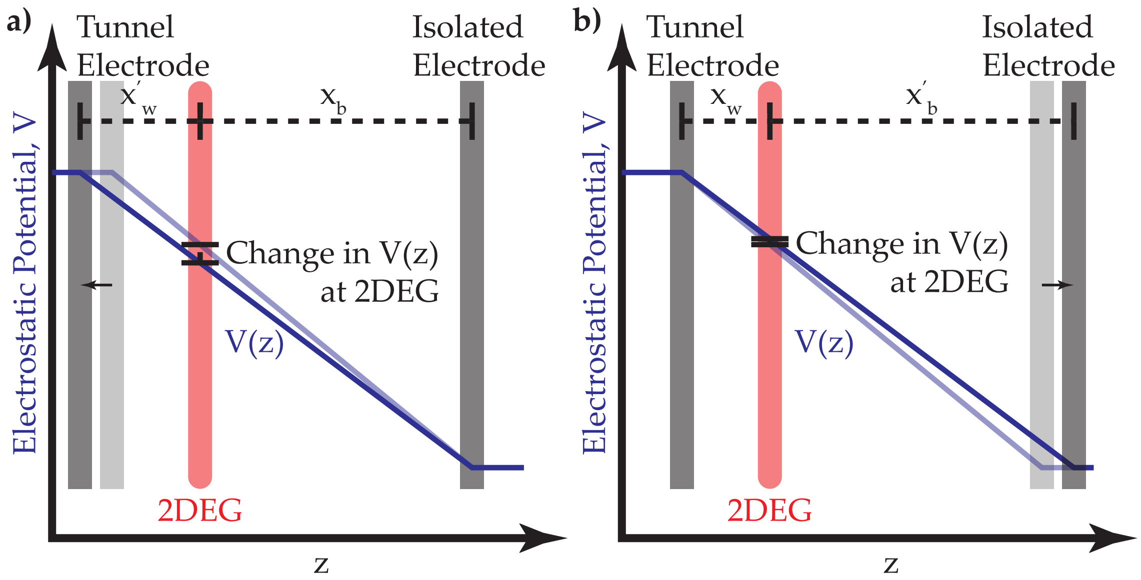


Figure 12-4: **Variation of the lever arm as the electrodes move:** Note that the bound state energy changes by much less in response to motion of the isolated electrode than it does in response to the motion of the tunnel electrode.

to the designed density of $1 \times 10^{18}$ cm$^{-3}$.

As these electrodes enhance and deplete in response to the DC bias, the lever arm changes. Because of the asymmetry in the structure, we are much more sensitive to the motion of the tunnel electrode than we are to the motion of the isolated electrode (Figure 12-4). Over a 300 mV range of DC bias, an electrode with a donor density of $1 \times 10^{18}$ cm$^{-3}$ will move by about 2 nm. If we assume the 2DEG screens the applied bias perfectly, in the enhancing direction only the blocking electrode moves. The 2 nm of motion results in a decrease of the lever arm by about 2% in response to the DC bias. In the depleting direction, the 2DEG will no longer screen the DC bias, and the same 300 mV would increase the separation between the 2DEG and the tunnel electrode by 2 nm, resulting in a massive 10% increase in the lever arm. Fortunately, once the 2DEG is empty, we stop applying DC bias, so in practice we never measure in this regime.

Thus, over our measurement range, we expect an approximately 2% variation of the geometric lever arm from the depletion of the isolated electrode. Note that while the polarization of the quantum well tended to increase our geometric lever arm by pulling the 2DEG closer to the isolated electrode, the isolated electrode tends to move away at the same time. This partial cancellation will tend to reduce the change in the geometric lever arm to less than 6% over the range of our measurement.

## 12.4 Variations in the Lever Arm; AC Motion of the Electrodes

In response to AC pulses, the 2DEG is unable to screen the applied electric field. Because of this, both the isolated and tunneling electrode tend to shift together and by approximately the same amount when the pulses are applied. (Note that this assumes that the tunnel electrode is not has not moved so far that it is pinned against the tunneling barrier. In practice, our simulations show the large spacer between the barrier and the doped region ensures this is not the case.)

As discussed above, the motion of the tunnel electrode is around 4 times more important than the motion of the isolated electrode. Focusing on this, in response to our 55 mV AC pulse, we expect around 4 Å of motion, resulting in a 2% change in the bound state energy of the quantum well. For enhancing pulses, this will tend to decrease the lever arm as we pull the electrode closer to 2DEG, while for depleting pulses, we push the electrode further from the 2DEG, increasing the lever arm. Note that this motion actually tracks well with both the direction and order of magnitude of the motion of the quantum well. Because the degree of polarization of the quantum well varies with DC bias both because of self-consistency and because it originates with a quadratic correction, the exact degree of cancellation will vary, but it gives us some confidence that this effect will reduce our current energy error estimate of around 6%. A more exact understanding of this reduction can be found by numeric simulation of the structure.

## 12.5 Numeric Simulations

Because our structure is essentially one-dimensional, it is easy to simulate its behavior numerically (Appendix C). In Chapter 11 we took advantage of this to provide a model sample in which it was possible to know exactly what was happening in one of our structures in response to applied DC biases and small AC excitations. We can easily extend this to simulate the response to large AC pulses. In extracting the lever arm from the structure, it is important to include the effects of donor migration during the epitaxial growth process[97, 98]. We approximate this migration by, instead of suddenly dropping the donor density from $1 \times 10^{18}$ to zero in the tunnel electrode, tapering it off in several steps approximating an exponential decay with a decay constant of 8 nm. This is appropriate for a 655° Celsius growth temperature.

We first simulate the structure at each DC bias at equilibrium, building up a table of 2D density in the quantum well as a function of bias. With the aid of this table, we are now ready to simulate the structure far out of equilibrium. For each DC bias, we change the bias across the structure by the pulse height being simulated while modifying the chemical potential in the quantum well to keep the 2D density constant at the equilibrium value we computed earlier. This allows us to model accurately the situation in the moment after the pulse is applied but before any charge has had time to tunnel.

Because the 2D density of states in the well is the same before and after the pulse, the difference between the lowest bound state energy and chemical potential in the well is fixed; we can compute the geometric lever arm using either. In practice, we will use the bound state energy because this quantity also enters into attempts to extract the well chemical potential from AC capacitance traces[33].

In Figure 12-5a, we show the bound state energy as a function of AC pulse height and DC bias over a range comparable to our experiment. Figure 12-5d shows a simulated dI/dV spectrum to allow the readers to orient themselves on

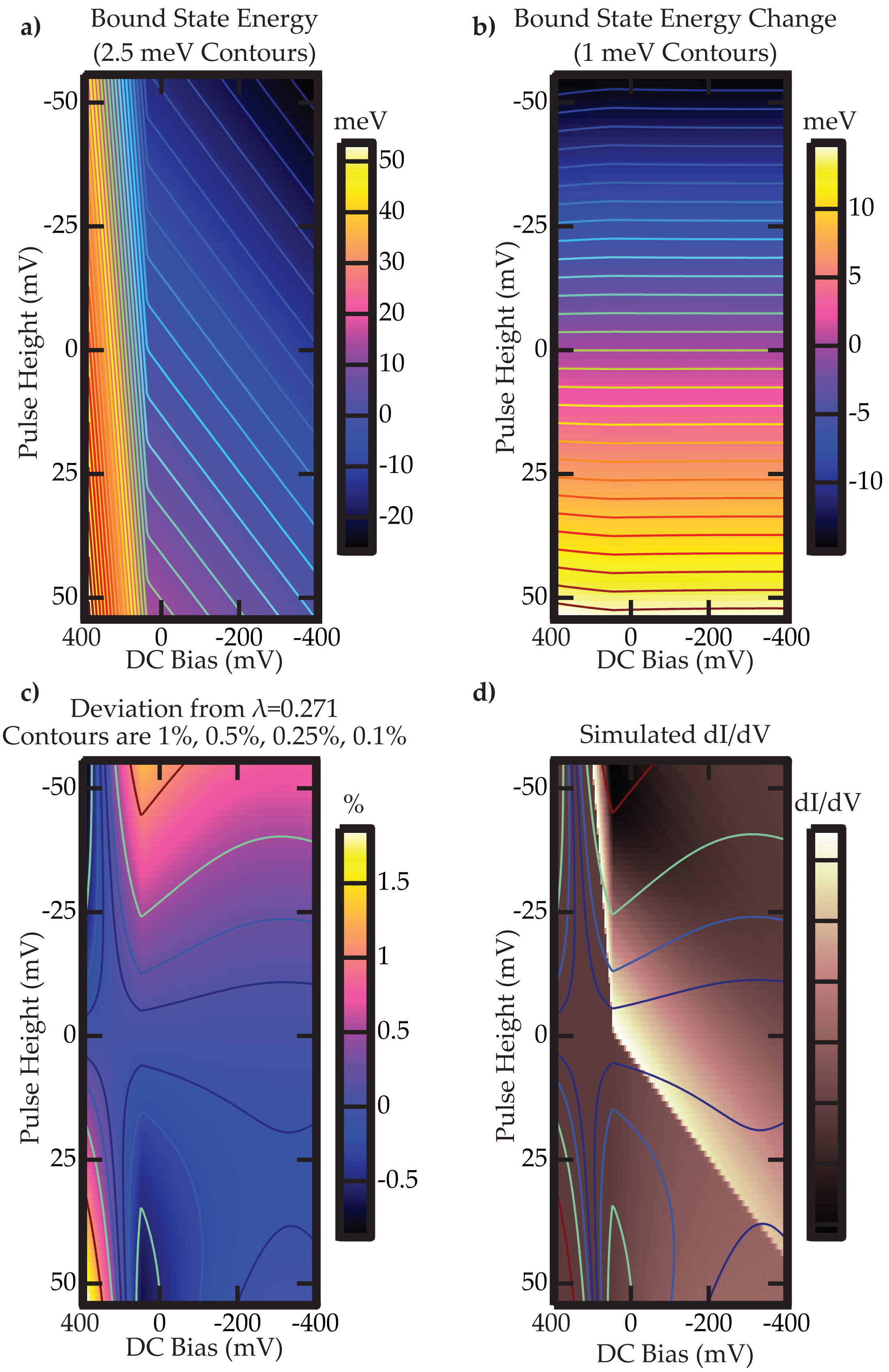

a)
Bound State Energy
(2.5 meV Contours)
meV
50
40
30
20
10
0
-10
-20
Pulse Height (mV)
-50
-25
0
25
50
400
0
-200
-400
DC Bias (mV)
b)
Bound State Energy Change
(1 meV Contours)
meV
10
5
0
-5
-10
Pulse Height (mV)
-50
-25
0
25
50
400
0
-200
-400
DC Bias (mV)
c)
Deviation from $\lambda$=0.271
Contours are 1%, 0.5%, 0.25%, 0.1%
%
1.5
1
0.5
0
-0.5
Pulse Height (mV)
-50
-25
0
25
50
400
0
-200
-400
DC Bias (mV)
d)
Simulated dI/dV
dI/dV
Pulse Height (mV)
-50
-25
0
25
50
400
0
-200
-400
DC Bias (mV)

Figure 12-5: **Calculated Lever Arm Errors: 0 Tesla** (previous page) **a)** Calculated bound state energy as a function of DC bias and pulse height. Note the bound state energy changes quite rapidly as a function of DC bias when the well is empty, and the 2DEG is unable to screen the applied electric field. TDCS is not sensitive to this overall shift in the bound state energy, however; only to the change in the bound state energy in response to applied pulses. 2.5 meV contours are superimposed for clarity. **b)** Change in the bound state energy in response to the applied pulse; this is obtained by subtracting the bound state energy after the pulse is applied from the bound state energy before the pulse is applied. 1 meV contours are superimposed; the uniformity of the lever arm can be seen in the even spacing of the contours, however, some narrowing of the contour interval, caused by an increase in the lever arm, can be seen at the far left edge where the well is fully depleted. **c)** Calculated deviations in the bound state energy from an assumed constant $\lambda = 0.272$ lever arm. Note the maximum deviation is 1.5%, while in the region of interest where there is charge in the well deviations from ideality are much smaller. **d)** Calculated dI/dV on the same scale as **a-c)**, to provide some basis for comparison.

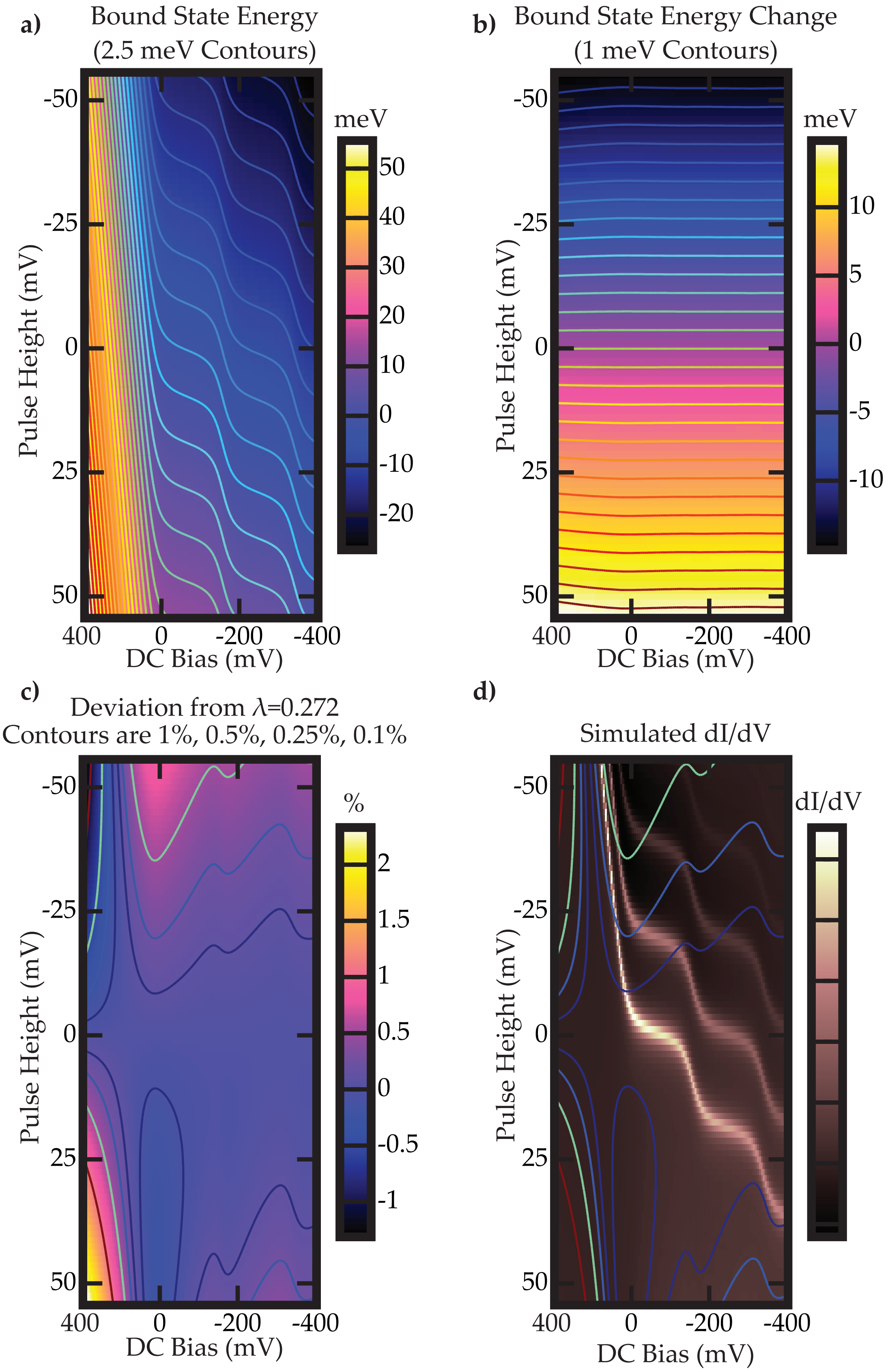

a)
Bound State Energy
(2.5 meV Contours)
meV
50
40
30
20
10
0
-10
-20
Pulse Height (mV)
-50
-25
0
25
50
400
0
-200
-400
DC Bias (mV)
b)
Bound State Energy Change
(1 meV Contours)
meV
10
5
0
-5
-10
Pulse Height (mV)
-50
-25
0
25
50
400
0
-200
-400
DC Bias (mV)
c)
Deviation from $\lambda$=0.272
Contours are 1%, 0.5%, 0.25%, 0.1%
%
2
1.5
1
0.5
0
-0.5
-1
Pulse Height (mV)
-50
-25
0
25
50
400
0
-200
-400
DC Bias (mV)
d)
Simulated dI/dV
dI/dV
Pulse Height (mV)
-50
-25
0
25
50
400
0
-200
-400
DC Bias (mV)

Figure 12-6: **Calculated Lever Arm Errors: 3 Tesla** (previous page) **a)** Calculated bound state energy as a function of DC bias and pulse height. Note the bound state energy changes quite rapidly as a function of DC bias when the well is empty or in the chemical potential jump in between Landau levels, preventing the 2DEG from screening the applied electric field. TDCS is not sensitive to this overall shift in the bound state energy, however; only to the change in the bound state energy in response to applied pulses. 2.5 meV contours are superimposed for clarity. **b)** Change in the bound state energy in response to the applied pulse; this is obtained by subtracting the bound state energy after the pulse is applied from the bound state energy before the pulse is applied. 1 meV contours are superimposed; the uniformity of the lever arm can be seen in the even spacing of the contours, however, some narrowing of the contour interval, caused by an increase in the lever arm, can be seen at the far left edge where the well is fully depleted. Some small wiggles are also present near the chemical potential jumps. **c)** Calculated deviations in the bound state energy from an assumed constant $\lambda = 0.272$ lever arm. Note the maximum deviation is 1.5%, while in the region of interest where there is charge in the well deviations from ideality are much smaller. There is much more fine structure for the 3 Tesla case than there is in the zero field case; this fine structure exemplifies the difficulties in trying to account more accurately for the variations of the lever arm. **d)** Calculated dI/dV, on the same scale as **a**-**c)**, to provide some basis for comparison.

this plot. At the far left of the spectrum, the 2DEG is fully depleted and is no longer able to screen the applied DC bias, resulting in a rapid change in the bound state energy. This reflects a changing “DC” lever arm; where the quantum well is able to charge, very little electric field penetrates it, so it takes large applied voltages to create small changes in chemical potential. The enormous fluctuations in this “DC” lever arm make it difficult to accurately extract the chemical potential from equilibrium capacitance measurements; for this reason we do not do so.

However, if we are only interested in the distance from features in the spectrum to the chemical potential, we need only worry about the difference between the chemical potential (or bound state energy) before and after the pulse is applied. Figure 12-5b shows this quantity. The evenness of the contour intervals reflects the accuracy of our fixed lever arm approximation. Only a very slight narrowing of the interval at the left edge of the figure where the well is depleted, reflecting the increase of the lever arm discussed above, is visible.

In our actual experiment, we fit the Landau level spacing in the region where the well is fully depleted to extract the geometric lever arm. We approximate that procedure by fitting the plane plotted in Figure 12-5c over the DC bias range from 50 mV to 100 mV to the form $E_{\text{bound state}} = \lambda V_{\text{Pulse}}$. We extract $\lambda = 0.272$; this a-priori calculation from the as-designed structure is in excellent agreement with the experimental value $\lambda = 0.267$. In our experiment, we approximate this lever arm as constant across the entire range of measurement. Doing the same, we extend this plane and plot the difference between the fit and the actual bound state energy in Figure 12-5c. Here, we plot the error in the bound state energy, $E_{\text{bound state}} - \lambda V_{\text{Pulse}}$ as a percentage of the full scale range of bound state energies, taken to be 15 mV. Contours are plotted at 0.1%, 0.25%, 0.5%, and 1%. With the exception of a small region in the bottom left corner corresponding to trying to eject electrons from an already empty well, the errors are all much less than 1%, and in the region within a few millivolts of the Fermi surface where we perform the bulk of our measurements, the errors are less than 0.1%.

One possible cause for concern is that the further application of a magnetic field creates additional regions where the 2DEG is unable to screen. In these regions we expect the quantum well polarization and the tunnel electrode depletion to both change quickly; this results in a rapid change in the bound state energy and the lever arm, just as in the fully depleted region at zero field. However, the average density of states in the quantum well remains fixed, so while this should cause some local perturbation in the lever arm, we do not expect this to result in global distortions. Furthermore, because the bound state energy changes very quickly in these regions, it only takes very small changes in the DC bias to move through them, minimizing their effect. This is evident in the small and narrow dips in the simulated and real capacitance traces in Figures 11-3 and 11-6.

We can demonstrate this by repeating our simulation, substituting a Landau level density of states for the flat density of states used in Figure 12-5; this is shown in Figure 12-6. We neglect all many body effects (including exchange splitting and negative compressibility) in this simulation. Note that although we see more fine structure in the lever arm, the errors are similar in magnitude and direction to those at zero field. This fine structure demonstrates the difficulty in attempting any higher order correction of the lever arm; any such correction is dependent on the physics of the quantum well.

It is possible to include some short-time capacitance experimental data to give an idea of how far the electrodes have depleted as a function of both DC bias and pulse height. However, the lever arm we observe comes from the partial cancellation of the well and electrode polarization, so using such data to generate higher order corrections would require using a detailed model of the entire structure similar to our simulations. We are hesitant to include such detailed corrections as part of our measurement.

## 12.6 Summary

The lever arm that enters into the TDCS energy calibration depends only on the distances between electronic states in our structure. By comparison, attempts to extract chemical potential jumps from capacitance data depends on a DC lever arm that incorporates screening by the 2DEG, while RTD's are sensitive to voltage drops across two separate barriers. Because the energy calibration in TDCS depends only on geometric factor, it is comparatively simple and insensitive to the detailed physics of the structure being studied. This gives well calibrated and undistorted energy scales in TDCS spectra compared to other electronic techniques.

We estimate energy errors of a few percent using qualitative arguments; detailed simulations showed that inclusion of the screening of the applied electric field by the 2DEG reduce these lever arm variations to around 150 $\mu$eV, or less than 1% of full scale. Errors in the region of the most interest near the Fermi surface are even smaller.

Reducing these errors further by accounting for the variability of the lever arm seems unlikely; even though we have some information about the depletion of the electrodes from the short time capacitance, the polarization of the quantum well is an important term in the lever arm. We have no direct measure of this polarization.

One interesting question is how to design a sample to reduce the lever arm variation. Such an attempt would have to simultaneously narrow the quantum well, reduce its polarization, and modify the donor density in the 3D electrodes to improve the cancellation between the polarization of the quantum well and the depletion of the electrodes. While it would be tedious, the simulation software developed in Appendix C would provide an invaluable tool in such attempts.

# Appendix A

# Signal Generation and Transmission

Careful attention needs to be paid to both the generation of the pulses used to excite the sample and their transmission into the sample space in order to meet the strict requirements on the pulse shape and amplitude control needed to acquire accurate and noise-free spectra. Many of the traditional tricks used in RF design only work over a narrow range of frequencies, and do not carry over well to time domain measurements. Here, we discuss what has been done and what options are available for improving the quality of the pulse arriving at the sample.

## A.1 The Pulse Shaper

Commercial signal generators are available with extremely short rise-times: times much faster than 100 picoseconds are readily available. However, this rise time is typically the 10% - 90% value; we find the pulses generated continue to ring on the order of a few percent, typically for several microseconds, after generation.[*] For TDCS, we only need moderately good rise times; the 50 Ω cable driving our 10 pF sample limits us to a 2 ns rise time (to 99%). However, we do need our pulses to be flat to within our energy resolution over the entire pulse width after the first few nanoseconds, and flat to around 100 times better than that

---

[*] Pulse generators designed for time domain reflectometry are available with much better characteristics, but typically do not provide amplitude control.

during the measurement interval to allow accurate slope determinations. We also need unusually precise amplitude control to allow accurate balancing and energy determination.

We achieve these using a device we call a "pulse shaper." This is a high gain amplifier designed to saturate sharply at an adjustable voltage. We use a commercial pulse generator to generate a signal that drives it into saturation, where the few percent ringing causes a negligible change in its output voltage. The current pulse shaper features some refinements over that discussed in [99] to maintain a better 50 Ω fixed output impedance.

The saturation voltage is adjustable by a DC control input. Currently our amplitude resolution is limited by the quantization of the voltage source used to generate this input, while the pulses generated are flat to within better than we can directly measure after 1 nanosecond has passed by. Subtracting two pulses from each other, we see some high frequency components on the order of one part in ten-thousand; this results in direct capacitive feed through of ringing from our signal generator to the output of the pulse shaper. It could be further reduced by chaining two pulse shapers, or by careful filtering of the signal generator's output. However, such filtering would have to be done with care; asymmetries in the threshold voltages of the amplifiers in the pulse shaper can turn slow input rise times into offsets in time of the generated pulses.

## A.2 There and Back Again; Reflections

Of course, the quality of the pulse generated is irrelevant if we are unable to transport it to the sample without distortion. Every component between the pulse shaper and the sample introduces an additional source of reflection and attenuation, and deserves careful consideration. On the excitation half of the bridge, the AC component is of the order 50 mV, but all but approximately a part in one thousand of that will be balanced off by the bridge, so distortions of the order of a part in one-hundred thousand will be clearly visible in the experiment;

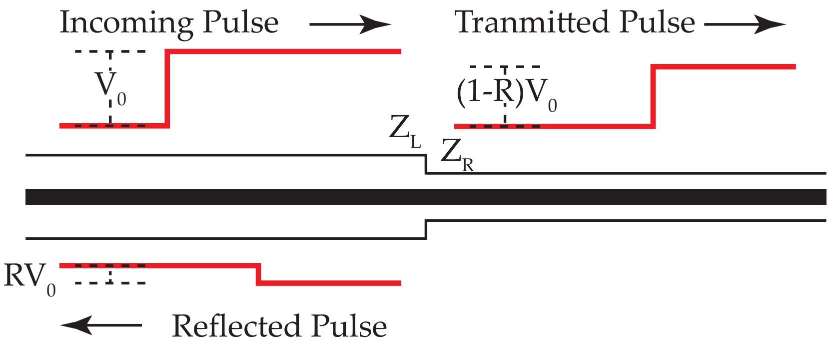


Figure A-1: **Single impedance mismatch:** When a step is sent through a single impedance mismatch, it is attenuated but not distorted.

even the smallest reflections are critical.

The simplest way to distort a pulse is with an impedance mismatch. Recall that if there is a sudden discontinuity in the characteristic impedance of a transmission line, there is a reflection, with the voltage reflection and transmission coefficients for a right-going wave given by[100]

$$R = \frac{Z_R - Z_L}{Z_R + Z_L} \approx \frac{\delta Z}{2Z} \tag{A.1}$$

$$T = 1 + R \approx 1 + \frac{\delta Z}{2Z} \tag{A.2}$$

Assume, for the moment, that the source and load in the experiment are correctly terminated. In that case, a single impedance mismatch will generate a reflection that will propagate backwards towards the signal generator where it will be absorbed. The net effect, then, is to slightly reduce the amplitude of the pulse arriving at the sample (Figure A-1), which will be automatically absorbed into our lever arm and density calibrations. Thus, a single reflection is not a problem.

If we introduce a second impedance mismatch, the situation becomes more complicated. Consider, for example, inserting a short section of cable of length $l$ with an impedance $Z_1$ in the center of a section of cable of impedance $Z_2$. This can be used as a simplistic model of a damaged or dented cable, a connector, or even a poorly designed strip-line on a circuit board.

Typically, our reflection coefficients will be small enough that it will be adequate to consider only the first few reflections. For the case of our short mismatched segment of line, noting that the reflection coefficient for entering the mismatched section of line has the opposite sign and the same magnitude $R$ as the reflection coefficient on exiting it, and taking for concreteness $Z_1 > Z_2$, we can consider terms of up to order $R^2$ and easily reason through the shape of the reflected and transmitted pulses (Figure A-2).

Imagine the step incident on the mismatch arrives at the front surface at $t = 0$. A small, negative step of magnitude $R$ is reflected back towards the signal generator, while a slightly attenuated pulse of magnitude $1 - R$ travels onward towards the back surface of the mismatch. At time $t = l$ (using units where $c = 1$), the attenuated pulse reaches the back surface. Here, a small positive step of magnitude $(1 - R)R \approx R$ is reflected back towards the signal source, while a step of total magnitude $(1 - R)(1 + R) = 1 - R^2$ is transmitted on in the cable.

At a time $t = 2l$, the positive reflected step reaches the front surface of the mismatch again. The portion transmitted through the front surface, back towards the signal source, has an amplitude $R(1 + R) \approx R$, and cancels the original reflected step. The net reflection from the mismatch is, to order $R$, a pulse of length $2l$ and amplitude $R$.

On the other hand, the portion reflected by the front surface, back towards the output, has an amplitude of $R^2 + \mathrm{O}(R^3)$. At a time $t = 3l$, this reaches the back surface of the mismatch again. The transmitted portion has an amplitude of $R^2 + \mathrm{O}(R^3)$. This is enough to restore, to order $R^3$, the pulse to its original amplitude.

Thus, the net effect of a short, small mismatch of length $l$ is to take a little "mouse bite" of length $2l$ out of the leading edge of a step, and to reflect a pulse of length $2l$ back towards the signal generator.

Our experiment has several connectors and a bias tee on the path from the signal generator to the sample. Roughly speaking, each of these take a mouse

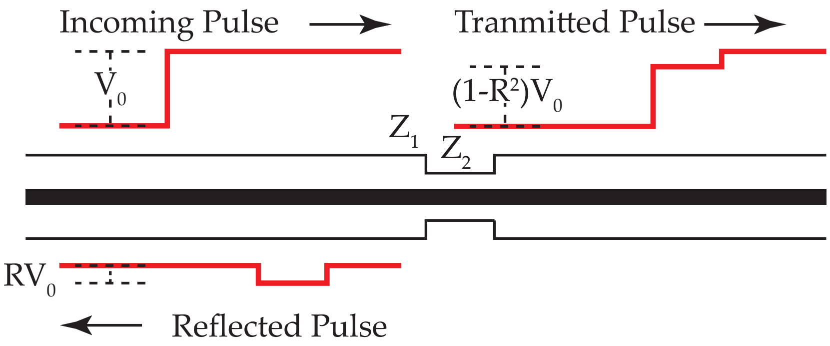


Figure A-2: **A double impedance mismatch:** The reflections off the front and back surface of the mismatch nearly cancel, so a double mismatch takes a short nibble out of the front of a transmitted pulse, and reflects a short pulse back towards the signal generator.

bite out of the front of the pulse, but because they are all small ( typically $\sim$ 100 ps long), the direct effect on the pulse is negligible.

However, our transmission line is not terminated in the cryostat; the power consumption of doing so would be unacceptable. Our sample is approximately a 10 pF capacitor to ground; although it can store some power, it cannot absorb any. Thus, at long times, it gives a reflection coefficient of +1; at short times, it looks like a short to ground, and gives a reflection coefficient of -1. When a step of amplitude V impinges on the sample, a reflected "glitch" that swings from -V to +V over the RC time of a few nanoseconds is reflected back towards the signal generator.

If our cables were perfect and the signal generator were an exact 50 Ω source, the signal generator would absorb this reflection and it would not matter. However, each connector and bias tee on the way back towards the signal generator acts as a short mismatched section. When this reflected glitch hits them, they reflect short, attenuated copies of the glitch back towards the sample. If they occur a distance $d$ down the sample, these glitches arrive at the sample at a time $2d$. The first of these mismatches is the SMA connector at the top of the cryostat, approximately 1 m from the sample. Thus, after we apply a pulse, we see a few nanosecond rise time, then approximately 10 ns of very clean pulse,

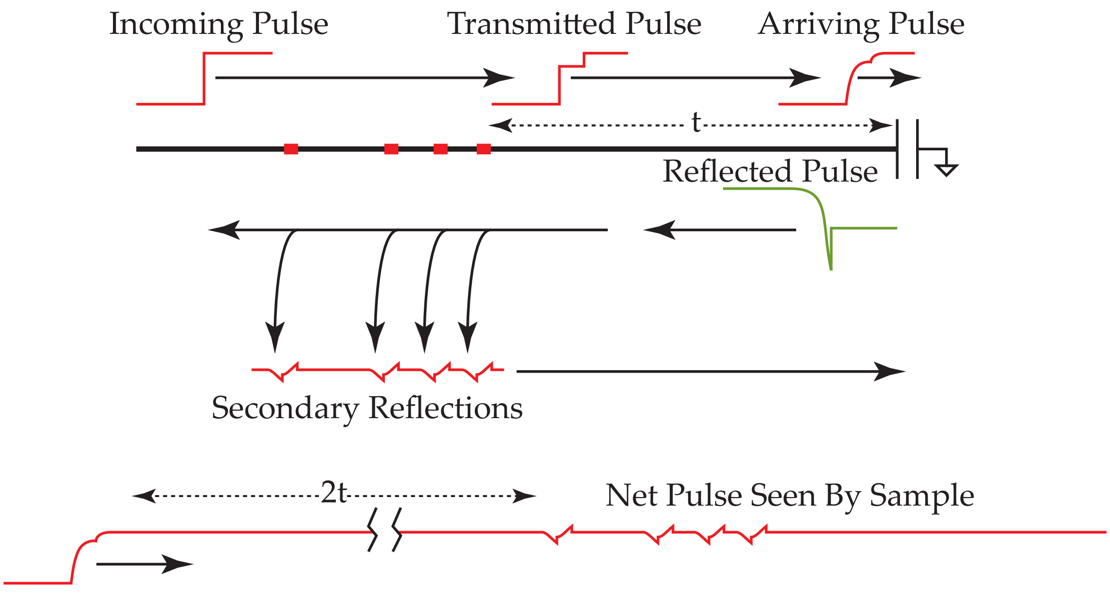


Figure A-3: **Net pulse:** On the way to the sample, multiple short mismatches created by connectors and the bias tee create a slight stoop at the front of the pulse. The RC time of the line charging the sample smoothes this stoop into an approximately exponential rise. However, the reflection caused by the charging of the sample has twice the peak-to-peak amplitude of the arriving pulse. This reflection travels back towards the signal generator, and additional power is reflected back towards the sample at each of the impedance mismatches. Because of the cable length, these arrive long after the initial pulse transient has died away, and provide a sizable contribution to our detected signal.

then we see a series of glitches corresponding to this connector, a bias tee, and the first attenuator on our cables, spanning the period from roughly 10 ns to 20 ns (Figure A-3).

We keep the magnitude of these attenuated glitches small by careful impedance matching, using high quality carefully tightened SMK connectors, and using semi-rigid coaxial cables manufactured with tight tolerances. With these efforts, only the first reflection off each connector is visible; pulses that reflect off the sample, off an imperfection, back off the sample again, and off an imperfection again are not visible when the apparatus is adjusted well.

The contribution from later connectors is usually negligible; there is a 20 dB attenuator immediately before the bias tee. A reflection off an imperfection on the far side of this attenuator passes through it twice: once on the way from the sample to the imperfection, and a second time on the way from the imperfection back to the sample. This 40 dB attenuation is typically enough to render the already-small reflection imperceptible. Careful placement of attenuators improves the quality of the pulse seen by the sample enormously.

For very fast samples, we can work in the relatively flat period between the initial transient and the arrival of the first reflections. However, there is some global distortion in this region from the RC rise time of the sample and the frequency-dependent attenuation of our cryogenic coaxial cables. For slower samples, we typically work after the reflected glitches have mostly died away and rely on the differentiation step to remove any systematic bias introduced by them.

## A.3 Detection

The problem in getting the signal back out of the cryostat again is enormously easier; the lion's share of the pulse has been cancelled by the balancing, so reflections on the order of a part in a thousand that would be a concern on the signal generation side are unimportant on the detection side. Our cryogenic amplifier has an approximately 50 Ω output impedance, but this does not have to be maintained with very much care as each of the later stages of amplification has a fairly good 50 Ω input and generates very little reflection.

With carefully selected low-pass filters to match the output bandwidth to our sampling rate, we see approximately 1 $\mathrm{nV}/\sqrt{\mathrm{Hz}}$ noise rated to the input of our cryogenic amplifier. This noise is dominantly from the thermal noise of our first room temperature amplifier. It is necessary to use constant-phase low pass filters to avoid distorting the received signal. It is also important to remember that most low-pass filters use LC circuits to provide the filtering; these do not absorb high frequency power, but reflect it back towards the source. Such a filter cannot be put directly on the output of the cryogenic amplifier; its poor impedance match will result in high frequency ringing; rather, it must be placed between two room temperature gain stages or directly on the input of the averager.

## A.4 Amplifier Design

> ... and that if any one ſhould doubt, whether the electrical matter paſſes through the ſubſtance of bodies, or only over and along their ſurfaces, a ſhock from an electricfied large glaſs jar, taken through his own body, will probably convince him.
>
> Watson, Wm. and Franklin, Benjamin. *An Account of* Mr. Benjamin Franklin's *Tratiſe, lately publiſhed, intituled,* Experiments and Obſervations on Electricity, made at Philadelphia in America; *by* Wm. Matſon, F.R.S.. *Philisophical Transactions* (47):202-211, June 1751.

Charge amplifiers sufficient to detect somewhat smaller charges than were used in Mr. Watson and Franklin's experiments are needed to detect the image charge of the displacement currents we use in our measurements. We use an amplifier based closely off a design with a long history in AC capacitance measurements [101, 33, 29]. The amplifier used in the past is a current-biased single or multiple stage common-source HEMT based amplifier. It has a power dissipation of a few $\mu$W, a gain of order 5, an input capacitance of around 300 fF, and a bandwidth of about a megahertz. We will sacrifice input capacitance and gain to extend the bandwidth to a few gigahertz.

The basic idea behind this tradeoff is the observation made in Gary Steele's thesis[29]; increasing the gate width increases the transconductance of the amplifier. In the bulk of our capacitance measurements, we are measuring very small capacitances, so the increase in the transconductance of the transistor is balanced by the decreased signal at the gate because of the shunting capacitance of the amplifier. In this case, the advantages of a wider gate amplifier are negligible.

However, in our case, the sample capacitance is large, so we can afford a very wide gate. We do not have the ability to fabricate our own HEMTs; however, we can effectively create a wide-gate amplifier by placing many HEMTs in parallel.

Given that we expect the transconductance to increase linearly with the number of HEMTs we use, it would seem to make sense to use enough to get the input capacitance of our amplifier to be about equal to our sample capacitance; in our case, this would mean using about 30 in parallel. Besides being impractical and unreliable, such a choice would give our cryogenic amplifier an unreasonably low output resistance.

This suggests a simple choice for the number of HEMTs; we can put enough in parallel to set the output resistance to 50 Ω , giving us an enormous bandwidth improvement over the 1 kΩ output impedance more typically used.

In order to pick the number of HEMTs to use, we need to choose an operating point, in terms of a drain current and voltage for our amplifier. Once we have selected that point, we can reach it by tuning the gate voltage of the HEMT and the current source used to bias the transistor. The main constraint on the operating point of our amplifier is power; if we want to consume 3 $\mu$W of power, it limits us to the set of points on the I-V curve of the transistor where $\mathrm{I} \times \mathrm{V} = 3\ \mu\mathrm{W}$; a single line. We can then examine that line and, for any given number of amplifiers, pick the point where the gain of the amplifier peaks. Finally, we look at the output impedance $\frac{\partial \mathrm{I}_{\mathrm{drain}}}{\partial \mathrm{V}_{\mathrm{drain}}}$. for any number of HEMTs, and pick the number of HEMTs that makes the output impedance closest to 50 Ω ; this is 3 or 4 for FHX35Xs depending on the batch of transistors. With our typical settings, the amplifier has a voltage gain of about one when driving a 50 Ω load.[†] Although the resulting amplifier is simplistic, we find it compares favorably in bandwidth and power consumption, if not gain, to those reported elsewhere[102].

Once we have cooled our amplifier, we measure the output impedance of our amplifier as a function of gate bias, and fine tune the gate bias to set the output impedance to exactly 50 Ω .

Our amplifier slowly drifts over time. The origin of this drift is not completely clear, but it seems to originate with DC leakage off the input of the amplifier slowly changing the bias point. In order to compensate for this drift, we continuously monitor the drain voltage on the HEMT, and feed back the DC gate

---

[†] This is, however, an enormous power gain.

bias to keep it fixed at the value that gave us a 50 Ω output impedance we found at the start of the cool down.

## A.5 Summary

A mixture of careful cabling, judicious use of attenuators, and a carefully designed pulse shaper allow us to deliver steps of unparalleled fidelity and amplitude control to a cryogenic load. Reflections off cables and direct capacitive feed-through dominate distortions in the voltage step delivered to the sample. Improvements are possible at all stages.

# Appendix B

# Experimental Setup

An 8131A two-channel pulse generator (**a**) is set to generate pulses with a high repetition rate; typically 10 to 100 kHz. Each repetition is used to generate one complete pulse sequence, so averages are typically accumulated at the rate of 10,000–100,000 per second. The trigger output ($T_1$) is optionally coupled through a hand-wound balun (**b**) to isolate the ground, and passed to the trigger in on the repetitive signal averager (**c**). Channel 1 is set to generate the gate signal for the measurement pulse ($P_1$), typically a few $\mu$s after the trigger to allow time for any transients from the trigger pulse to die down. It is sent as a complementary pair to the pulse shaper (**d**). The pulse shaper uses a complementary pair of fixed amplitude gate pulses combined with DC control inputs to generate a very flat, continuously variable amplitude pulse. The 8131A only provides 1% control over pulse width; this is inadequate for the discharge pulse. A DG535 (**e**), providing better than 1 ns resolution with a 1 ns rise time, is used to generate the gate signal for the discharge pulse. It generates high frequency noise while timing (in between being triggered and generating the falling edge of the last pulse it generates); this is acceptable during the discharge pulse, but generates excessive interference if the DG535 is running during the measurement pulse. To circumvent this, Channel 2 on the 8131A is set to generate a trigger pulse ($T_2$) at the end of the measurement pulse. This DG535 begins the discharge pulse gate signal ($P_2$) immediately after being triggered.

Four DC input voltages ($V_P^{1-4}$) are used by the pulse shaper to control the heights of the measurement pulse on the excitation, the discharge pulse on the excitation, the measurement pulse on the standard, and the discharge pulse on the standard. These are generated by 20 bit unipolar D/A converters based around AD760s. At present, the quantization error in these outputs dominates our pulse amplitude control errors. The pulse shaper has unipolar outputs; in order to generate negated pulses, we reverse the duty cycle on the pulses. This introduces a DC offset that is removed later by AC coupling. The 8131A is able to complement its outputs, but the DG535 cannot. Four 50 Ω relays (**g**) are used to optionally swap the outputs of the DG535, effectively complementing them. The control coil is electrically isolated from the signal path; this does not connect the computer's ground to the experiment's.

In order to preserve the flatness of the pulses generated by the pulse shaper, the cabling from this point out is done with care, using semi-rigid coaxial cables. Immediately before entering the cryostat, 3 to 30 dB of broadband attenuators (**h**) are used to reduce the impact of any reflections. RF leakage from the large amplitude (2.5 V) gate pulses can interfere with the low level detected signals. A separate Faraday cage surrounds the signal-generation portion of the apparatus to minimize this leakage (**i**). A stripline RC bias tee (**j**) is used to add the sample bias $V_{DC}$. This bias is generated by an independent set of D/A converters (**k**) to reduce any stray coupling between the signal generation apparatus and the sample excitation.

A cryogenic amplifier (Section A.4, **l**) provides somewhat less than unity voltage gain to match the high-impedance sample to the 50 Ω cabling leaving the cryostat. It is current biased through $R_{D2}$, which is integrated with the AC coupling on the first stage of room temperature amplification. $R_{D1}$ is used to allow easy adjustment of the effective bias voltage and low pass filter the output of the low-drift +10 VDC fixed-output voltage source (**m**). A voltage meter (**n**) is used to monitor the DC drain voltage of the transistor, and thus the drain current, through a large isolation resistor $R_m$. This signal is used to continuously and slowly tune the gate voltage of the cryogenic amplifier $V_{\text{Gate}}$ to maintain a fixed output impedance despite small drifts in the cryogenic transistors.

Two to three series Mini-Circuits Gali amplifiers (**o**) are used to provide 44-66 dB of gain from 100 Hz to 1 GHz. This bandwidth is limited on the low end by AC coupling capacitors and the high end by room temperature amplifier selection. Our overall noise is limited by the input noise of these amplifiers; typically approximately 1 $\mathrm{nV}/\sqrt{\mathrm{Hz}}$ referenced to the gate of the cryogenic transistor. The apparatus as a whole is grounded through these amplifiers.

The signal averager is grounded to the (otherwise floating) computer through a USB connection. Current-generation USB isolators do not support the transfer rate needed by this experiment. This is unfortunate, but unavoidable. Noise contributions from the computer are minimized by removal of the video card. Digital dithering ([27]) is accomplished using the input offset circuitry internal to the signal averager.

Figure B-1: **Block Diagram of Experiment:**

**a** Two channel pulse generator, used to control the pulse sequence and generate the gate for the measurement pulse

**b** Balun, only used sometimes, to break the ground loop between the 8131A and the signal averager

**c** Digital signal averager

**d** Pulse shaper; generates variable height pulses from fixed-height gate pulses (complementary pairs) and DC control voltages

**e** Delay generator, used to generate discharge pulses with precisely controlled widths

**f** DC Voltage source, generates control voltages for the pulse shaper

**g** 4-Relay controllable swap to allow complementing of the delay generator output

**h** 3 – 30 dB attenuation to limit the impact of reflections off the sample

**i** Copper mesh Faraday cage to prevent RF leakage from the gate signals from interfering with the experiment

**j** Bias tee, approximately 100 Hz low frequency cutoff.

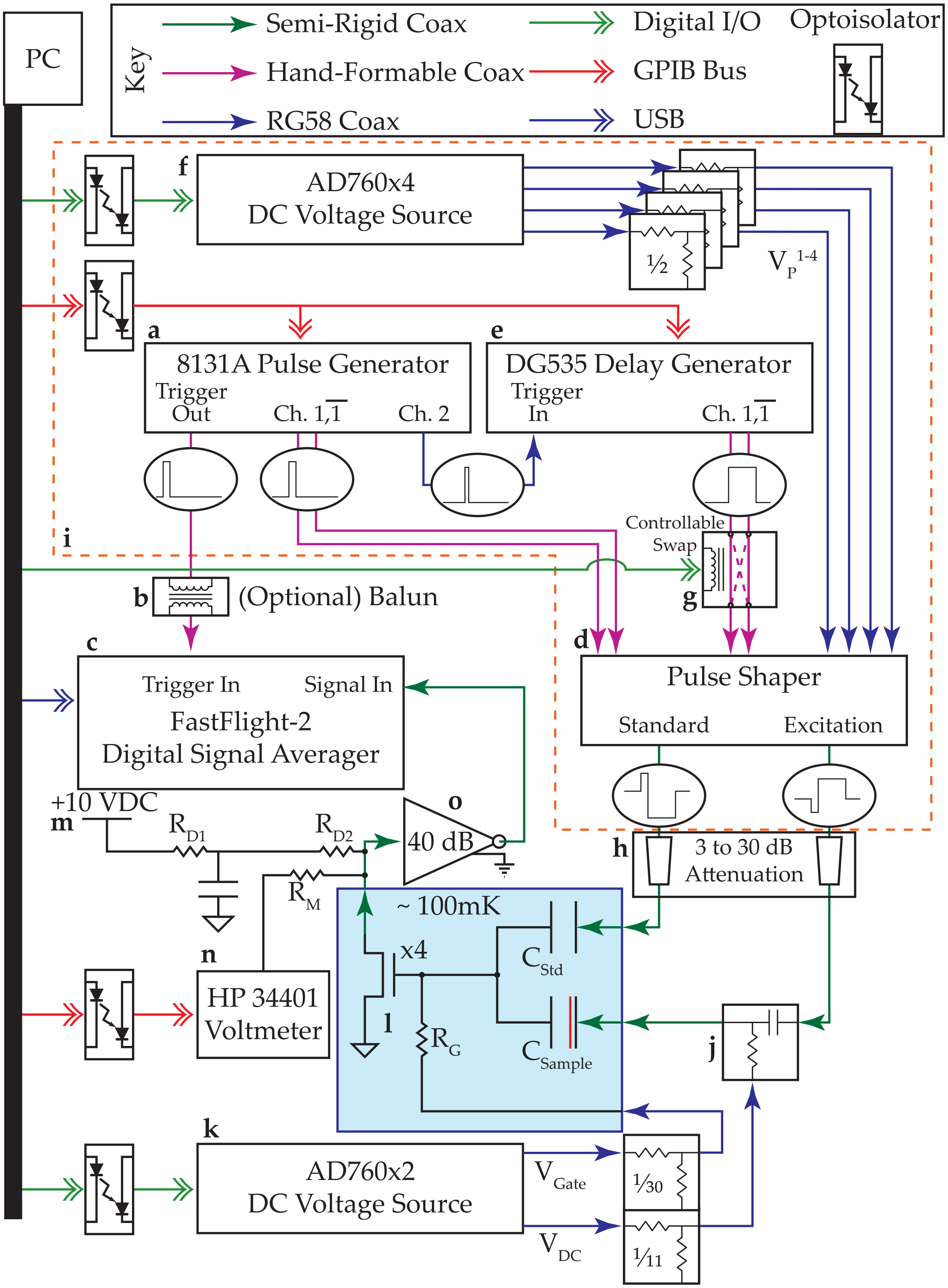
**k** DC voltage source, isolated from **f** to prevent RF coupling

**l** Cryogenic amplifier, approximately unity gain, 50 $\Omega$ output; see Section A.4

**m** 10 VDC fixed DC voltage source

**n** DC voltage meter, monitors drain current for feedback to $V_{\text{Gate}}$

**o** Room-temperature amplifiers


PC
Key
Semi-Rigid Coax
Hand-Formable Coax
RG58 Coax
Digital I/O
GPIB Bus
USB
Optoisolator
f
AD760x4
DC Voltage Source
½
$V_P^{1-4}$
a
8131A Pulse Generator
Trigger Out
Ch. 1,$\overline{1}$
Ch. 2
e
DG535 Delay Generator
Trigger In
Ch. 1,$\overline{1}$
i
Controllable Swap
g
b
(Optional) Balun
c
Trigger In
Signal In
FastFlight-2
Digital Signal Averager
d
Pulse Shaper
Standard
Excitation
+10 VDC
m
$R_{D1}$
$R_{D2}$
o
40 dB
$R_M$
h
3 to 30 dB Attenuation
~ 100mK
x4
$C_{Std}$
l
$R_G$
$C_{Sample}$
n
HP 34401 Voltmeter
j
k
AD760x2
DC Voltage Source
$V_{Gate}$
$V_{DC}$
1⁄30
1⁄11

# Appendix C

# Computer Simulations

Numeric simulations of our structure have provided essential insight into TDCS; not only have they helped us understand the limitations of our density calibration (Chapter 11) and the error bounds on our energy calibration (Chapter 12), but they provide a basis for rational sample design, in which we can start from the desired properties of our structure in terms of tunneling rate, energy range, energy accuracy, density range, and sub-band separation, and develop a structure to suit. The ability to extract quantities that we cannot physically measure, such as the chemical potential in the quantum well or the center of charge in one of the electrodes, makes these simulations useful in testing ideas for new ways to analyze the data TDCS provides. We can ask questions like “if we were able to extract or guess the electrode positions, how much more accurate would our density calculations be if we still ignored motional capacitance?”

The simulations need to be used carefully; they do not attempt to include all of the physics of our structure, nor do they directly include the uncertainties of sample growth and inhomogeneity. They also need to be used in parallel with a well developed physical intuition as to what is actually happening in the structures; without that intuition, it is difficult to detect when something has gone wrong with the simulation code, causing it to pursue unphysical or even incorrect solutions, and it is also difficult to choose how to modify a structure to alleviate a problem. This last concern is responsible for the bulk of Chapters 9- 12; although we could simply simulate the structure in its totality to extract the desired results, we would not trust the simulations (or learn very much) without deriving the results, at least in order of magnitude, by hand.

The simulations discussed in this chapter are straightforward extensions of those possible with G. L. Snider's *1dpoisson* [103, 104]. Our own simulation code is able to simulate much larger structures, and provides some unusual features well suited to the non-equilibrium situations that arise in simulating our structures, as described below. We will discuss the basis of the simulations in order to understand the limitations and capabilities of this kind of approach and explain the differences between our code and prior work, and then discuss approaches for applying these simulations to our experiment.

## C.1 Poisson's Equation in One Dimension

We are trying to solve two coupled problems; one is, given a charge distribution $\rho(\vec{x})$ and dielectric constant $\epsilon(\vec{x})$, what potential field $\mathrm{V}(\vec{x})$ satisfies the macroscopic version of Poisson's equation,

$$\nabla \cdot \left( \epsilon\left(\vec{x}\right) \vec{\nabla} \mathrm{V} \right) = \rho\left(\vec{x}\right) \tag{C.1}$$

In one dimension, we can replace $\nabla$ by $\partial / \partial z$. We solve this problem by first discretizing our fields on a series of lattice points $z_i$ with such that $\mathrm{V}_i = \mathrm{V}(z_i)$, and approximating our derivatives using finite differences

$$\frac{\partial f(z_{i+0.5})}{\partial z} \approx \frac{f(z_{i+1}) - f(z_i)}{z_{i+1} - z_i} \tag{C.2}$$

where $z_{i+0.5}$ is a reminder that we have evaluated the derivative half-way between lattice points, not at a lattice point.

We can now rewrite our differential equation as a vector equation

$$\mathcal{A}\vec{\mathrm{V}} = \vec{\rho} \tag{C.3}$$

In one dimension, using the discretization above, the matrix $\mathcal{A}$ is band tridiagonal, as the value of $\rho$ at any point only depends on the value of V at that point and its two neighbors. This makes the 1D version of this problem easy compared to the problem in higher numbers of dimensions; Crout reduction[41, Chapter 2] can generate V for any $\rho$ in both linear time and memory.[†] This is not surprising; one can imagine picking a random electric field at one end of the problem, integrating Poisson's equation to the opposite end, and then, using the superposition principle, adding a uniform electric field to the entire solution to match the boundary conditions. Crout reduction uses some clever bookkeeping to do essentially this for a somewhat more general circumstance. In higher dimensions one needs to move to complicated relaxation techniques and/or much higher order computations[105, 29].

## C.2 Nonlinear Materials

In semiconductors, the charge distribution $\rho$ is not fixed, but is actually a function of the potential V. Let us assume we have a model for our semiconductor in the form of a function $\rho(V)$; that is, the charge density at any point in the semiconductor is only a function of the potential at that point. We can iteratively attempt to find $\mathrm{V}(z)$ by linearizing the charge density about our last guess at the potential. Defining $\mathcal{R}$ as the matrix whose diagonal elements $\mathcal{R}_{ii}$ are given by $\frac{\partial\rho(V^i)}{\partial V^i}$ for the $i$th iteration $\mathrm{V}^i(z)$,

$$\mathcal{A}\vec{\mathrm{V}}^{i+1} = \rho\left(\vec{\mathrm{V}}^i\right)\mathcal{R}(\vec{\mathrm{V}}^{i+1} - \vec{\mathrm{V}}^i) \tag{C.4}$$

$$\left(\mathcal{A} - \mathcal{R}\right)\vec{\mathrm{V}}^{i+1} = \rho\left(\vec{\mathrm{V}}^i\right) - \mathcal{R}\vec{\mathrm{V}}^i \tag{C.5}$$

For non-pathological charge functions, this will converge in only a few iterations; on a modern computer, a structure that consumes all available memory to store will typically converge within a few seconds.

---

[†] We cannot generate $\mathcal{A}^{-1}$ in linear time or memory; it is not sparse.

One oft-used $\rho(V)$ is that given by the Thomas-Fermi approximation[32]. Essentially, the charge density at any point is the same as it would be in an infinite semiconductor subject to the same chemical potential $\mu$. This approximation does not include any consideration of confinement energy, but is quite good if the potential variation is smooth on the scale of the Fermi wavelength. In this approximation, in a single band 3D semiconductor with a conduction band edge $\mathrm{E}_c$, an effective mass $m^*$, and with $\mathrm{E} = eV$

$$\rho = \Theta(\mathrm{E} - \mathrm{E}_c) e \frac{\left(2m^*(\mathrm{E} - \mathrm{E}_c)\right)^{3/2}}{3\pi^2} \tag{C.6}$$

where $\Theta(x)$ is the Heaviside step function.

Using this approximation, with a position dependent value of $\mathrm{E}_c$ and $m^*$, we can easily simulate 1D semiconductor heterostructures. However, the heterostructures of interest for us have quantum wells in them, where we know the confinement energy is of utmost importance. In the Thomas-Fermi approximation, we not only overestimate the charge density substantially due to the neglect of the bound state energy relative to the band edge, but we will tend to move all of the charge density to the edges of the well where the electric field is strongest (Figure C-1). In this region, we will need to solve the coupled Schrödinger-Poisson equations. For all of our simulation work, the only accounting we make of many-body effects in to use a self-consistent electrostatic potential when modeling our structures; no exchange or correlation effects are included. Much like the Poisson equation, when the Schrödinger equation is rewritten as finite-difference equations, the Schrödinger equation is band tridiagonal[104]. In fact, with an appropriate choice of unitary transformation, it can be made symmetric tridiagonal. There exist a large number of efficient eigenvector and eigenvalue solvers for matrices of this form[41, 106].

We will partition our problem into the region inside of the quantum well, where we know solving Schrödinger's equation is crucial, and the region outside where the Thomas-Fermi approximation may be appropriate. We will begin by

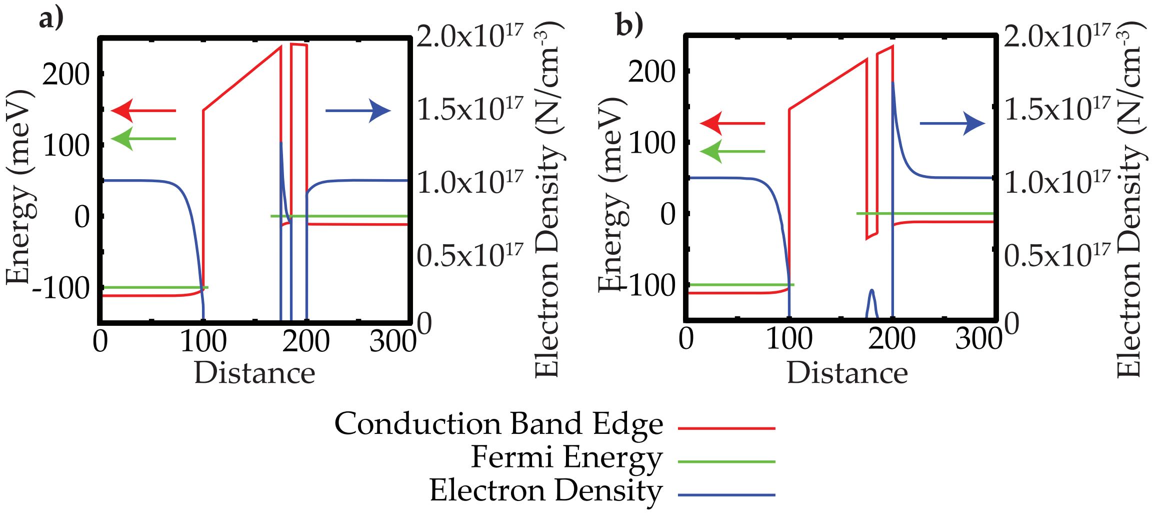


Figure C-1: **Different approximations to $\rho(z)$ inside the quantum well**: **a)** The quantum well simulated in the Thomas-Fermi approximation. **b)** The quantum well simulated using the coupled Schrödinger and Poisson equations, with the leads simulated using the Thomas-Fermi approximation.

finding a trial solution using the Thomas-Fermi approximation everywhere. We will then continue iterating using the Schrödinger equation in the well. After each iteration where we have generated a trial potential $\vec{\mathrm{V}}^i$, we will solve for the set of $N$, where $N$ is the number of $z_i$ inside of the quantum well, eigenvalues $\lambda_j$ and eigenvectors $\Psi_j(z_i)$ that solve Schrödinger's equation inside of the quantum well. We can then evaluate the charge density in this region as

$$\rho(z_i) = e \sum_{j=0}^{N} \left|\Psi_j(z_i)\right|^2 \mathrm{N}_{2\mathrm{d}}(\mathrm{E}_{\mathrm{fermi}} - \lambda_j) \tag{C.7}$$

$$\mathrm{N}_{2\mathrm{d}}(\mathrm{E}) = \frac{m^*\mathrm{E}}{\pi\hbar^2}\Theta(\mathrm{E}) \tag{C.8}$$

We will once again linearize the nonlinear part of the charge density. We will neglect the off diagonal coupling of the potential at one point in the structure to another, and using first order perturbation theory,

$$\frac{\partial\rho}{\partial V_i} = e \sum_{j=0}^{N} \left|\Psi_j(z_i)\right|^2 \mathrm{N}_{2\mathrm{d}}(\mathrm{E}_{\mathrm{fermi}} - \lambda_j - \delta_E) \tag{C.9}$$

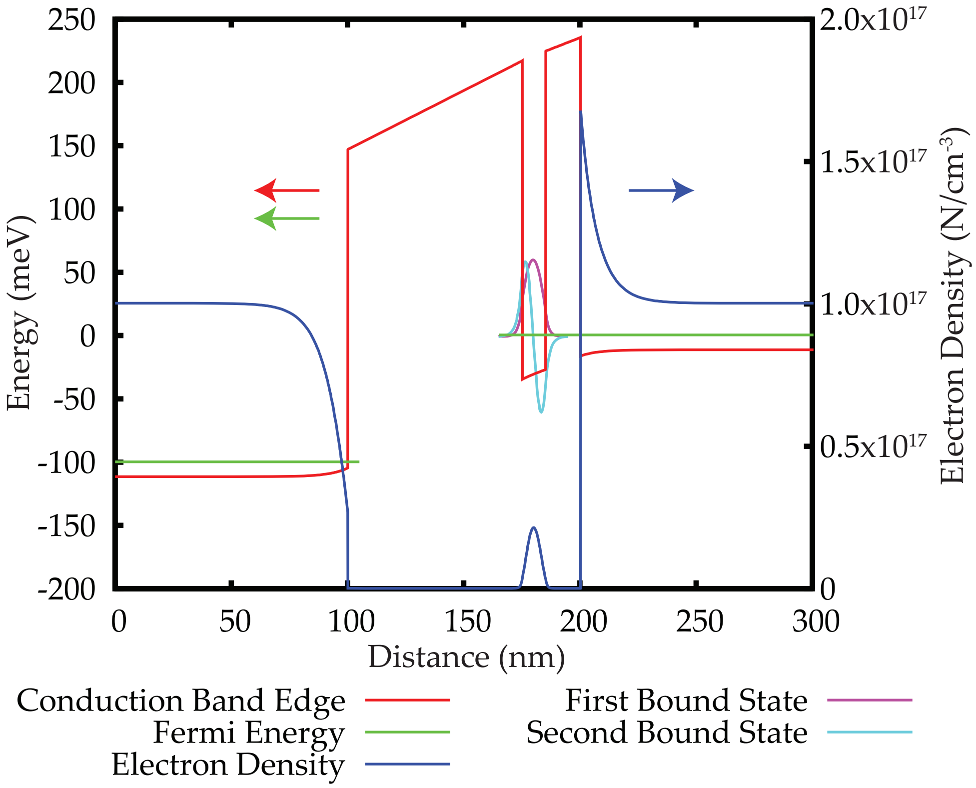


Figure C-2: **The first two bound states of a simple quantum well**: Only the lowest bound state is occupied.

$$\delta_E = e \left| \Psi_j(z_i) \right|^2 \tag{C.10}$$

Once again, convergence on iteration this process is fast for smooth densities of states like those in Equation C.8.

As a handy side benefit, this procedure also yields the bound state eigenfunctions and eigenvalues (Figure C-2). These are useful both as a sanity check and also in calculating tunneling rates through the structure when it is not at equilibrium.

## C.3 The Electrodes

On examining Figure C-2 with care, it becomes apparent that the Thomas-Fermi approximation is causing some difficulties in the leads; at around 150 nm into the structure, we see the edge of an electrode that has been enhanced somewhat by the electric field penetrating through the quantum well into it. In reality, we know that the charge density will go to zero[†] at the edge of the tall AlGaAs tunnel barrier; all of the wave functions in the electrode are forced to have a node here. As shown in Figure C-3, this zero in the wavefunction forces the effective edge of the electrode to be several nanometers away from the actual position of the barrier. The Thomas-Fermi approximation does not capture this. For the purposes of calculating our geometric lever arms and examining their variation with applied voltage, we need the region very close to the tunneling and blocking barrier to be accurately simulated.

Thus, we would like to solve the coupled Schrödinger-Poisson equations in the electrodes, as well as the quantum well, in order to work around this problem. The boundary condition at the edge of our simulation poses some difficulty, however; the correct boundary condition, that the electrode "continues on forever", allows the wavefunction to take on any value at the last element in the simulation. Because of this, the matrix representing the Schrödinger equation will be singular; this reflects the fact that the real boundary condition allows an infinite number of bound states.

On approach to this problem is to simulate the electrode as a very wide, but finite quantum well. We will replace our continuous spectrum in the electrode with a discrete one, but if we make the "electrode well" wide enough, this spectrum will be densely packed and will closely resemble the continuous spectrum. Similarly, our charge density will approach the true charge density. As we lengthen the quantum well, we will slowly add longer and longer wavelength eigenvectors to the simulation. In the case of the Thomas-Fermi approximation, we had a solution that was accurate at large length scales and approximate at short. In the case of using a finite well for an electrode, we will find solutions that are accurate at short length scales, but which deviate at long wavelengths.

It is important not to put too much trust in the wavefunctions and detailed

---

[†] Actually, because the barrier is finite, the density will be small but finite here. See Chapter 9 for some analytic approaches to dealing with finite barriers.

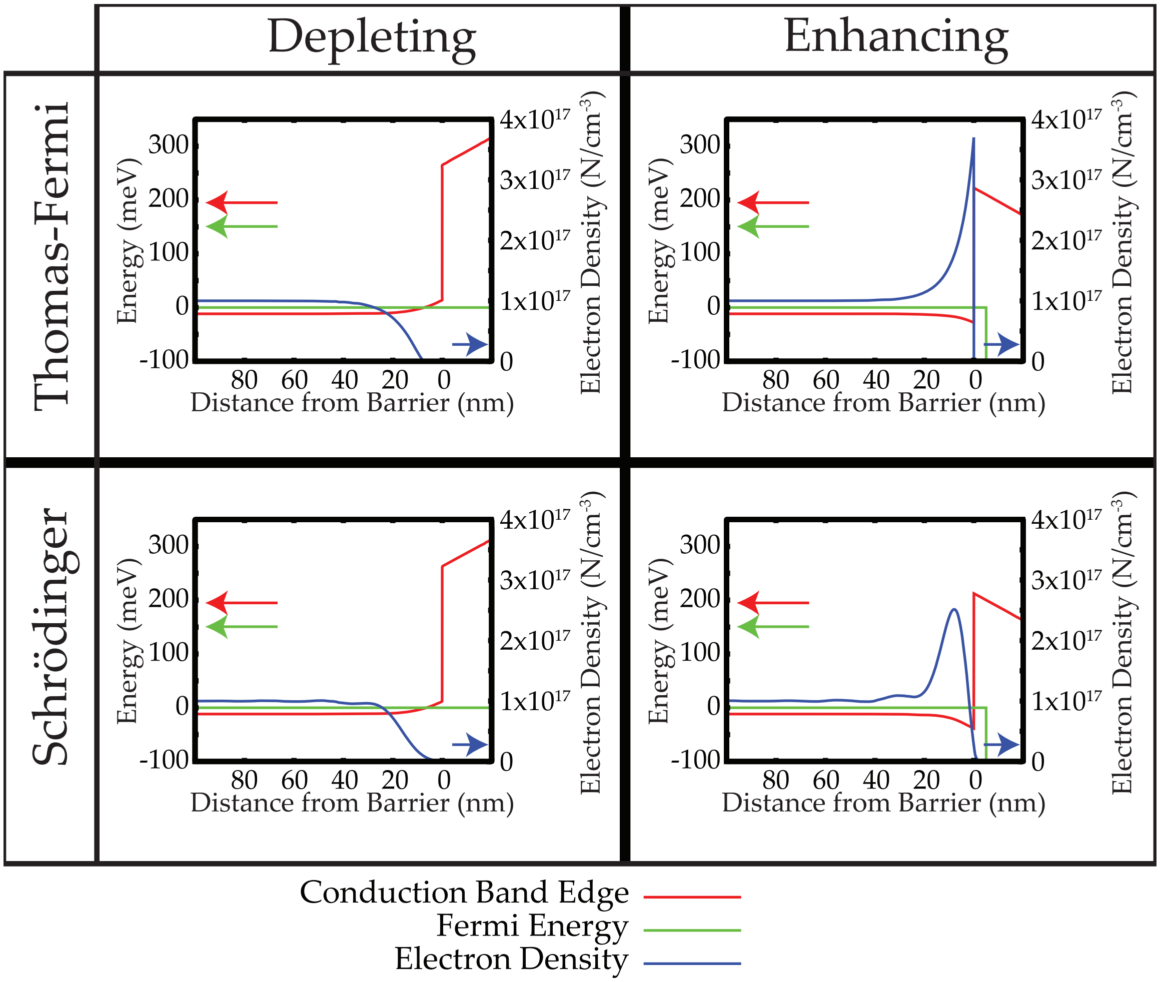


Figure C-3: **Charge densities near AlGaAs barriers** shown calculated near a barrier where a depleting or enhancing electric field has been applied, using either the Thomas-Fermi approximation or the full coupled Schrödinger-Poisson equations. The doping in these simulations extends all the way up to the barriers. A low donor density ($1 \times 10^{17} cm^{-3}$) emphasizes the importance of kinetic energy in the screening.

screening that come out of using wide quantum wells for the electrodes; the real quantum well has densely packed donors that act as scattering centers, while in our simulations we have a uniform charge density background. However, if our simulations can capture the zero in the charge density at the edge of the electrodes and the way the electrons nearest to the barriers move in response to small applied voltages, it will vastly improve the accuracy of our calculated lever arms. We typically use quantum wells 500-1000 nm thick in our calculations. Given that our top electrode is actually only 1000 nm thick, and the scattering length in our electrodes is on the order of 10 nm, these numbers seem more than adequate. Halving or doubling the widths of the electrodes does not change our results significantly.

Simulating these enormous quantum wells poses some numeric difficulties. In order to model accurately the region near the barrier where the potential is changing rapidly, we need to use closely spaced lattice points; typical values range from 0.1 Å to 1 Å. If we extend this across our entire electrode, we will have around $10^5$ lattice points in the electrode region, and thus $10^5$ eigenvalues and eigenvectors. Each eigenvector has $10^5$ elements, for a totally memory consumption of $10^{10}$ floating point numbers, requiring around 80 GB of storage. Calculating this full set of eigenvalues and eigenvectors would cause our modeling code to crawly to a halt. It is perhaps in light of this consideration that *1dpoisson* limits itself to a few thousand lattice points in its Schrödinger regions.

One approach would be to use a carefully chosen non-uniform lattice, with tightly packed points near the barriers and more widely spaced points in the bulk of the electrode. This seems overly complicated, would require careful tuning, and would inevitably carry with it some loss of precision. A more fruitful approach is to realize that, while there are hundreds of thousands of eigenvectors, we only care about the ones corresponding to occupied eigenstates, whose energies are below the Fermi energy. Typically, there will be a few hundred of these, reducing our CPU time and memory consumption by a factor of one thousand.

We use bisection of the secular equation [106, Chapter 8] to bracket all of the

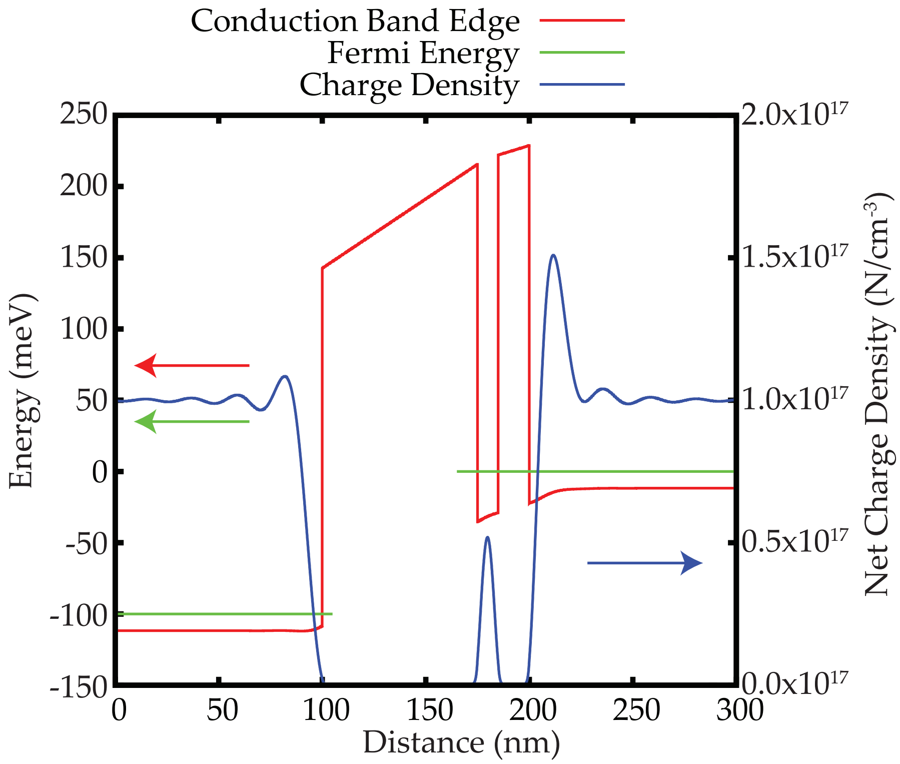


Figure C-4: **A fully self-consistent quantum well** modeled using the coupled Schrödinger-Poisson equations. The 3D electrodes extend 500 nm off the edges of the plot. A low donor density ($1 \times 10^{17} cm^{-3}$) emphasizes the importance of kinetic energy in the screening. There are no spacers in this structure; the donors extend right up to the edges in the AlGaAs barriers

eigenvalues below the Fermi energy, then use Brent's method[41, Section 9.3] to find the eigenvalue within each bracketed region. Once we have the eigenvalues, we can use inverse iteration[41, Section 11.7] to find the eigenvectors, once again using Crout reduction to solve the matrix equations. The net effect of this work is that, for an $n$ element lattice, we can find each eigenvalue and eigenvector in $\mathrm{O}(n)$ time. This allows us to rapidly model large structures without using the Thomas-Fermi approximation. For an example of such a simulation, see Figure C-4. This capability is one of the key features that separates our simulation code from *1dpoisson*.

## C.4 Non-equilibrium Structures

The simulation techniques we have discussed are only appropriate for equilibrium structures; we have not included any of the physics or techniques necessary to deal with current flows or non-thermal carrier distributions. Even when our structure is simply under a constant DC bias, it is not at equilibrium. Fortunately, because the blocking barrier is very thick, we can neglect current flow through the structure, and just treat it as two isolated, equilibrium systems that interact through the Poisson equation. In order to do this, we allow different Fermi energies in different parts of the structure, and offset our density functions by the Fermi energy at each point in the structure. Because the resistance of the tunnel barrier is negligible compared to that of the blocking barrier, we treat the quantum well as being at equilibrium with the tunnel electrode when the device is only under DC bias; see Figure C-5.

In order to accurately model the lever arm as a function of pulse height, we need to model the structure in the moments immediately after a pulse has been applied, but before any charge has tunneled. In order to simplify our modeling, we assume that enough time has elapsed after the application of the pulse for charge to flow in through the electrodes. This is true if the time after the pulse is much greater than $RC_{\text{sample}}$, where R is the sum of the impedances of the signal source, the ohmic contacts, and any bulk material the current must flow through, and $C_{\text{sample}}$ is the geometric capacitance of the sample. As long as the ohmic contacts are of high quality, R is dominated by the 50 Ω impedance of our signal source, and this time scale is a few hundred picoseconds. As we begin our actual measurements after several nanoseconds have elapsed, this is a good approximation. We also assume that the electrons in the quantum well remain in the ground state of the quantum well during the application of the pulse. Because our rise time is much longer than the period of Rabi oscillations between the first and second excited state of the quantum well, we expect the application of our pulse to adiabatically move the electrons across the well, not to excite transitions

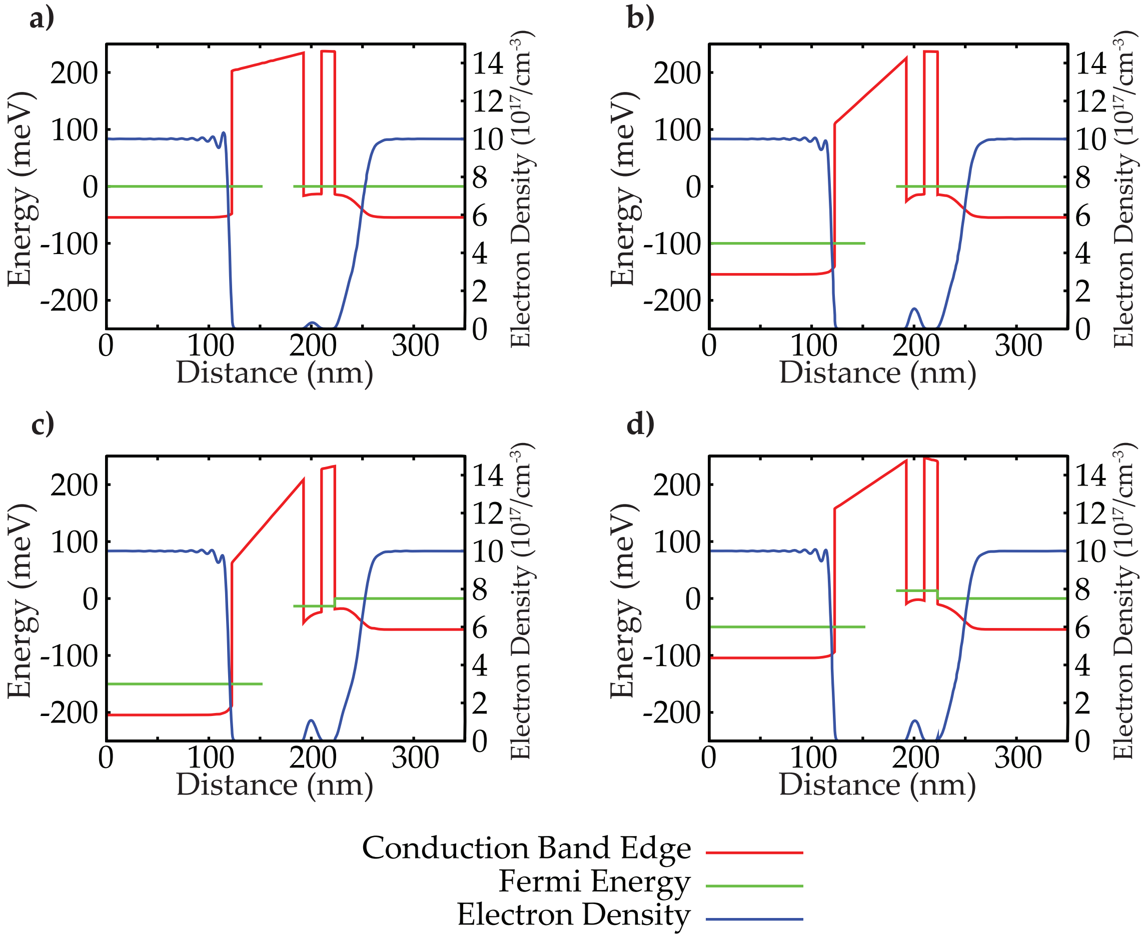


Figure C-5: **Simulations of non-equilibrium structures** modeled using a position dependent Fermi energy. **a)**: A simulation of PF061998-1 at zero bias. **b)**: The structure of **a**, simulated with 100 mV DC bias applied, increasing the charge density in the quantum well. **c)**: The structure of **a**, with 100 mV of DC bias applied, simulated immediately after the application of a 50 mV injecting pulse. **d)**: As **c**, but with an ejecting instead of injecting pulse.

to excited states in the quantum well.

Given these assumptions, both the electrodes and the quantum well will be in equilibrium states before we perform our measurement, but will be disequilibrated with each other. Much as with the DC bias, we can simulate the electrodes by setting their Fermi energies to reflect the voltage (AC + DC) applied. Setting the Fermi energy in the quantum well requires some care. We could simply use a geometric lever arm calculated from the geometry of the structure, but we are interested in using this model to extract accurate estimates of the lever arm and its dependence on the response of the structure to pulses, so this would be circular. Instead, we first simulate the structure with only the DC bias applied (Figure C-5b). We calculate the total charge in the quantum well with this DC bias. Then, we set the Fermi energies of the two electrodes appropriately for the sum of the AC pulse height and DC bias applied. We allow the Fermi energy in the quantum well to float, updating it in each iteration of the simulation to keep the charge density in the quantum well fixed. It is important for good convergence to remove the linearized dependence of the charge density on the electrostatic potential from the update equation Equation C.5 in the quantum well region; because we are allowing the chemical potential to float, $\frac{\partial \rho}{\partial V} = 0$.

Approaching the simulation of the structure immediately after the pulse is applied in this fashion allows us to realistically take into account the change in the shape of the bound state of the quantum well in response to the AC pulse. We can also easily extract the lever arm using the difference between the Fermi energy of the tunnel electrode and that of the quantum well. The ability to hold the charge density in a region fixed rather than the Fermi energy is another of the features that distinguishes our code from *1dpoisson*.

Detailed plots of lever arms extracted in this fashion can be found in Figure 12-5.

## C.5 Equilibrium Capacitance Measurements

A slightly specialized use of the ability to calculate the response of a structure after a pulse is applied but before any charge has tunneled is to calculate the expected response of the structure to equilibrium AC capacitance measurements. These calculations were key to understanding the errors inherent in extracting the quantum well density from capacitance measurements in Chapter 11.

We are now interested in the measured capacitance of our structure, $\frac{\partial \mathrm{Q}}{\partial \mathrm{V}}$, both at low frequencies where the quantum well and tunnel electrode have time to fully relax, and at high frequencies where no charge has time to tunnel out of the quantum well. Because our structure has to remain charge neutral, we know that at all times the sum of the charges on the isolated electrode, the tunnel electrode, and the quantum well will be zero. The charge that accumulates on the isolated electrode comes from one lead of our device, while that in the quantum well and on the tunnel electrode comes from the other. Thus, we could calculate Q above by considering either the charge on the isolated electrode or the sum of the charges on the tunnel electrode and quantum well; for convenience, we typically do the former.

In order to extract the low frequency capacitance, we simulate the structure at a variety of closely spaced DC biases $\mathrm{V}_i$, and find the charge density on the isolated electrode for each DC bias $\mathrm{Q}_i$. We then calculate the low frequency capacitance at $\mathrm{V}_i$ as

$$\mathrm{C}_{\mathrm{DC}}|_{\mathrm{V}_i} = \frac{\mathrm{Q}_{i+1} - \mathrm{Q}_i}{\mathrm{V}_{i+1} - \mathrm{V}_i} \tag{C.11}$$

In order to extract the high frequency capacitance, for each of the DC biases we simulate applying a very small pulse $\delta_{\mathrm{V}}$, typically 1 $\mu$V, as described above, and find the new total charge on the isolated electrode $\mathrm{Q}_i'$. We then simply compute the high frequency capacitance as

$$\mathrm{C}_{\mathrm{AC}}|_{\mathrm{V}_i} = \frac{\mathrm{Q}_i' - \mathrm{Q}_i}{\delta_{\mathrm{V}}} \tag{C.12}$$

Plots of typical simulated capacitances are shown in Figure 11-6

These simple simulations calculations allow us to find equilibrium capacitance values that reflect all of the physics of the structure, including the motional capacitance terms arising from the depletion of the electrodes, the polarization of the quantum well, and the depletion region near the AlGaAs barriers resulting from finite screening wavelengths.

## C.6 Non-uniform Densities of States

In order to provide contrast in our equilibrium capacitance calculations, we need to use a non-uniform density of states in our quantum well. This will also be useful in examining the importance of the screening of the quantum well in examining the variability of the lever arm.

To first order, we can simply substitute any function we desire for $\rho(V)$. For our density of states when a magnetic field is applied, we use a superposition of Lorentzians for $\frac{\partial\rho(V)}{\partial V}$, and thus a sum of arcsin functions for $\rho(V)$. Unfortunately, the rapid fluctuations this causes in $\frac{\partial\rho(V)}{\partial V}$ slows or entirely prevents convergence of our coupled simulation.

To overcome this, we use a fairly brute-force approach. We monitor convergence of our simulation, and if we find it is converging slowly, or not at all, we add a damping factor to our Newton's method update of the charge density. That is, after we have used $\vec{V}^i$ to find $\vec{V}^{i+1}$, we replace the calculated value of $\vec{V}^{i+1}$ by $\gamma\vec{V}^i + (1-\gamma)V^{\vec{i+1}}$, where $\gamma$ is a damping factor $0 \leq \gamma \leq 1$. We slowly increase $\gamma$ until the convergence improves, then reduce it again. With this addition, our simulations converge well. For our high field simulations, we typically use this Landau DOS in the quantum well, but a flat density of states in the electrodes. This greatly speeds the simulations (calculating arcsin(V) is slow, even aside from convergence issues), and can be physically motivated by the short scattering times in the 3D (Section 10.2).

It would be interesting to perform simulations that include a simple model of exchange like that used in Chapter 5. However, at present we are not able to deal with negative compressibility; it results in a function $\rho(V)$ that is double valued. It should be possible to, similar to Section 5.5, use a "fake capacitance" term to make $\rho(V)$ single valued, and add a term that updates the chemical potential in the simulation to remove this fake capacitance from the final results. This is left for future work.

## C.7 Calculating tunneling currents

In addition to calculate equilibrium properties of the quantum well, it is interesting to calculate the expected tunnel current through the barrier as a function of applied bias and pulse height. Because we have access to the wavefunctions in the quantum well and in the 3D electrodes, it is simple to apply the transfer matrix formalism discussed in Section 10.3. Our simulation code allows the regions in which we solve Schrödinger's equation to overlap inside of the tunnel barrier.[†] The option to solve Schrödinger's equation in multiple, overlapping regions is a third difference between our code and *1dpoisson*.

Once we have done this, we need to deal with the fact that our electrodes are finite, and so the wavefunctions we have there do not reflect a true continuum. In general, we will not find a wavefunction in the electrode with the same energy as the bound state we are considering in the quantum well. To overcome this, we linearly interpolate $\psi$ and $\psi'$, evaluated at the barrier, between the two bound states closest in energy in the electrode. This should work well as long as the eigenfunctions in the electrode vary smoothly with energy; if we had a super-lattice tunnel barrier creating wide fluctuations in the penetration with energy, this would not work well[‡]. We then compute the density of states in the 3D

---

[†] This means we have an over-complete basis inside of the tunnel barrier. This is a requirement for using the transfer matrix formalism, which uses the non-orthogonality of the wavefunctions for calculating transition rates

[‡] Typically, the super-lattice barriers we use have a lattice constant of 11 Å. At the energies we are interested in, this is much shorter than the wavelength of our electrons, and we will not see such oscillations

as the reciprocal of the energy spacing between these bound states. In order to avoid jumps in the tunnel current as we pass from one bound state in the 3D to another, we linearly interpolate the 3D density of states using the three bound states nearest in energy.

Examples of computed tunneling I-V spectra can be found in Figure 12-5 and Figure 8-5.

## C.8 Summary

We have developed simulation code, similar to Snider's *1dpoisson*, able to solve the coupled Schrödinger-Poisson equations in our structures. Our code is able to simulate non-equilibrium structures, providing facilities convenient for the calculations that arise in simulating the response of tunnel capacitors to sudden pulses. We include the ability to solve Schrödinger's equation over large regions, allowing us to include realistic screening in our device electrodes, and to calculate tunnel currents using the transfer matrix formalism.

# Appendix D

# Fabrication

This chapter is organized in two parts; first, we describe in general terms the steps in preparing one of our electron or hole doped samples. Then, we describe in detail the individual fabrication steps used in the first part.

We use both clear field and dark field versions of two masks in our fabrication of the mesa-isolated tunnel capacitors; a small area is diagrammed in Figure D-1, along with a photo of a small section of the mask. The labels in Figure D-1 will be used to refer to the different mask sections.

## D.1 Electron doped samples

**Remove Gallium** Remove the gallium on the back of the sample.

> This can be done by mechanical abrasion with a lapping machine.
>
> Alternatively, the sample can be mounted with the front surface against a glass slide using PMMA or mounting wax, and the gallium can be removed using a 20% hydrochloric acid solution. This will etch GaAs as well as gallium, so it is important that the entire front surface of the sample be covered.

**Backside Metallization** The wafer used in the electron doped samples is degenerately doped all the way through, so the contact to the tunneling electrode

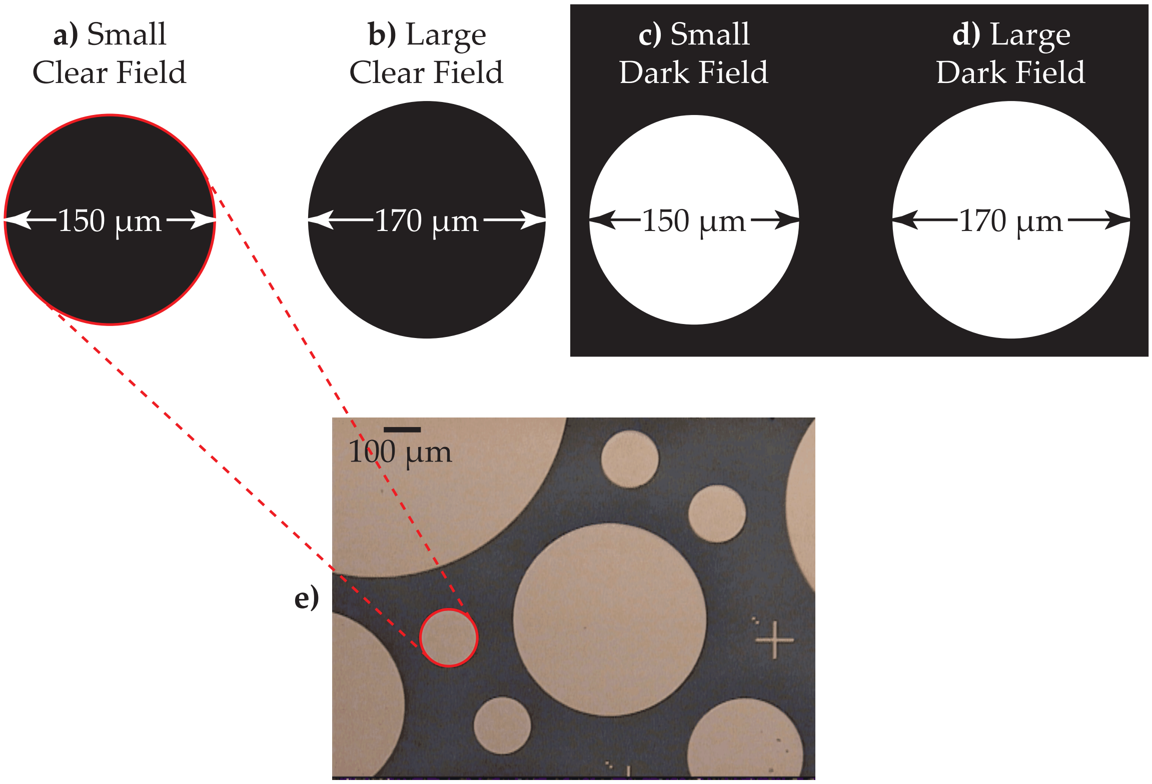


Figure D-1: **Masks** used in our fabrication processes. One mask has a clear field (**a-b**), while the other has a dark field (**c-d**); they are otherwise identical. Each has an array with a variety of different sizes of circles, as exemplified by a single 150 $\mu$m circle in **a**. The entire array is repeated with 20 $\mu$m larger diameter circles in a different section of the mask, as exemplified by the single 170$\mu$m circle in **b**. **e)** shows a photomicrograph of a small section of the actual mask, showing the pattern of different sized circles. We typically only use the smallest circles.

is made through the back side of the wafer. Mount, clean, and deposit metal on the back side of the wafer as described in subsection D.3.2, but do not yet anneal the sample. There will be a single anneal for all of the contacts at the end of the processing.

**Mesa Definition** Using the small, clear field mask and negative (Futurex) resist, define the mesa pattern as described in Section D.6.

**Mesa Metallization** Evaporate the top side contact, as described in subsection D.3.2. It may be useful to include an additional TiAu overcoat to ease bonding to the sample. Lift-off (Section D.8) the metal and anneal.

**Mesa Isolation Etch** Etch for 45 seconds (adjust, as appropriate, for the well depth), using 1:10 30% $H_2O_2$ :0.1N NaOH; see Section D.4.

**Profile** Using a profilimeter, check the etch depth to ensure it isolates the quantum wells from each other. Etch for longer, as needed.

**Back-side contact** Using H20E silver epoxy, glue the sample to a piece of GaAs that has been completely coated with Ti-Au. This will allow wire bonds to reach the back side contact. Be sure to stir the epoxy thoroughly before use.

## D.2 Hole doped samples

The wafer used for the carbon doped samples is semi-insulating, so both contacts to the tunnel capacitor need to be made through the front side of the sample. The current to the "back side" of the mesas will be carried through a thin ($\sim$ 1 $\mu$m) degenerately doped layer. Because this current flow is perpendicular to the applied magnetic field, care must be taken to reduce the distance the carriers must travel, in order to minimize the series resistance caused by this layer. We will use the entire sample away from the mesas to form our "back" contact, minimizing this distance.

We will take advantage of the undercut of the Futurex to provide a thin gap between the metallization at the bottom of the mesas and the mesa wall, and use the difference between the "large" and "small" masks to create a thin gap between the edge of the mesa and the top metallization; see Figure D-2. Extreme care is needed in alignment.

**Remove Gallium** Remove the gallium on the back of the sample.

This can be done by mechanical abrasion with a lapping machine.

Alternatively, the sample can be mounted with the front surface against a glass slide using PMMA or mounting wax, and the gallium can be removed using a 20% hydrochloric acid solution. This will etch GaAs as well as gallium, so it is important that the entire front surface of the sample be covered.

**Top Contact Definition** Using the small, clear field mask (Figure D-1a) and negative (Futurex) resist, define the top contact pattern as described in Section D.6; see Figure D-2b.

**Top Contact (Mesa) Metallization** Evaporate the top side contact, as described in subsection D.3.2, but do not anneal yet. It may be useful to include an additional TiAu overcoat to ease bonding to the sample. Lift-off (Section D.8) the metal. See Figure D-2c.

**Mesa Definition** Use the large, dark field mask (Figure D-1d) and negative photoresist to define the bottom contact pattern (Section D.6). Although good alignment to the top contact mask will give a high yield, the process is somewhat forgiving. If a small misalignment causes the resist undercut to touch the top contact metallization, the top contact metallization will act as a second mask, preventing most shorts.

The undercut provides the separation between the foot of the mesa and the bottom metallization (see Figure D-2d-f). In order to get a good separation between the foot of the mesa and the bottom contact, it is desirable to have a

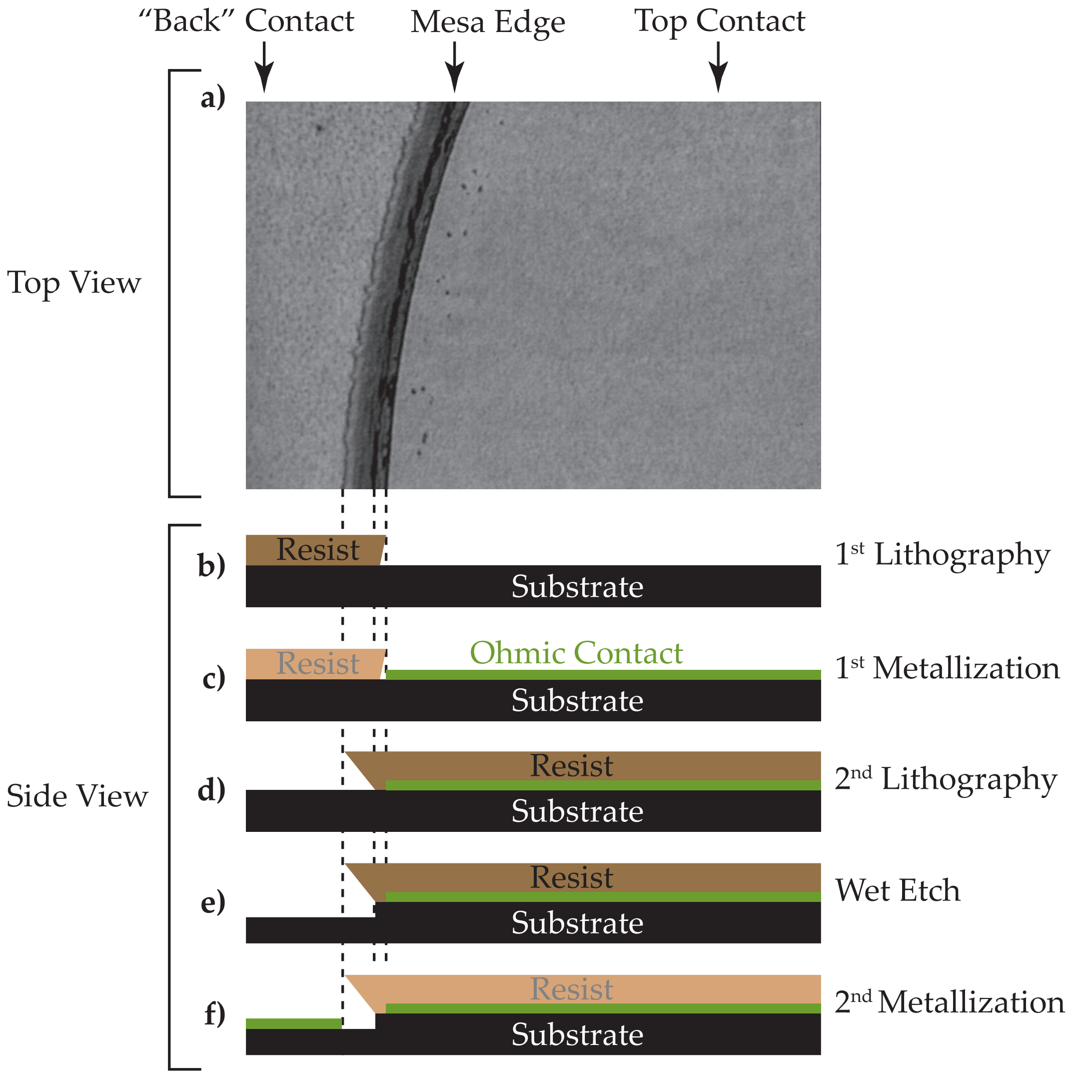


Figure D-2: **Hole doped sample fabrication**; showing **a)** a micrograph of the edge of a mesa in a completed sample. **b)** the sample after the first lithography step. **c)** the structure after the first ohmic contact metallization. Although the resist has been removed at this point as part of lift-off, it is still shown for illustrative purposes. **d)** the structure after the second lithography step. **e)** The structure after the mesa isolation etch. **f)** The structure after the final metallization; as with **c**, the resist is shown for illustrative purposes only.

large undercut in this step. Spinning the resist at 4k RPM instead of 5k will increase the undercut, as will reducing the exposure time to 45 seconds.

**Profile** In order to accurately measure the mesa etch depth, it is necessary to measure the thickness of the photoresist before starting the etch.

**Mesa Isolation Etch** Etch for 45 seconds (adjust, as appropriate, for the well depth), using 1:10 30% $H_2O_2$ :0.1N NaOH; see Section D.4.

**Profile** Using a profilimeter, check the etch depth to ensure it isolates the quantum wells from each other. Etch for longer, as needed.

**Bottom Contact Metallization** Evaporate the "bottom" contact, as described in subsection D.3.2, lift-off, and anneal.

## D.3 Ohmic Contacts to GaAs

Because the high density surface states in GaAs lie in the center of its band-gap, making low-resistance ohmic contacts to GaAs can be nontrivial. A wide variety of recipes exists in the literature[107, 108]. Here we summarize the three that we most commonly use in our devices.

### D.3.1 n-GaAs: Ni-Au-Ge Contacts

Nickel gold germanium contacts to GaAs are reliable, simple, and provide a low contact resistance. However, the gold in the metallization tends to form Gold-gallium alloyed spikes deep (several microns) into the substrate[109, 110]. If these spikes penetrate our tunnel capacitors, they will short them. Thus, use of NiAuGe contacts on the top of the mesas in our devices gives low yields and leaky capacitors. They are acceptable for use in the bottom contact, although in practice we normally use Pd-Ge (subsection D.3.2) contacts on both the top and bottom of the device for consistency.

**Mount** Place a small drop of 5% PMMA on a glass slide. Dip the bottom of the sample in the PMMA droplet and then place it on a different part of the slide. Bake the slide on a hot plate for approximately one minute, until the original drop of PMMA has dried. It may be useful to cut the glass slide before doing this so it fits conveniently into small beakers.

**Surface Clean** Clean the sample surface by dipping in 1:10 HCl:Water by volume, then rinsing in water. This will remove the surface oxide from the GaAs. The oxide will reform in a matter of minutes; it is important to quickly move the sample into the evaporator after this step. (A 5% Ammonia solution also works well for cleaning and oxide removal).

**Metallization**

| Material | Thickness |
|---|---|
| Ni | 5 nm |
| Ge | 35 nm |
| Au | 70 nm |
| Ti | 5 nm |
| Au | 200 nm |

If the contact is patterned, lift-off (Section D.8) should be performed at this point.

**Anneal** All of our ohmic contacts are annealed using a rapid thermal annealer. We anneal in a constant flow of forming gas (95% $N_2$ 5%$H_2$).

| Action | Temperature | Time |
|---|---|---|
| Ramp | 200° Celsius | 30 seconds |
| Hold | 200° Celsius | 30 seconds |
| Ramp | 415° Celsius | 22 seconds |
| Ramp | 435° Celsius | 12 seconds |
| Hold | 435° Celsius | 1 minute |
| Cool | Room Temperature | |

The Ge-Au layer forms a eutectic mixture, melting during the annealing process. The exact thicknesses of these layers can be modified more or less at will, as long as the gold is maintained as 88% of the total structure by weight (this corresponds to a 2:1 ratio of gold to germanium by thickness). Because this layer actually becomes liquid during annealing, gold-germanium contacts tend to spread and move slightly. This can be a problem if the contact is at the top of a mesa; the motion can cause it to short to the substrate. If used in this context (which is not recommended due to spiking), the anneal step should be performed before the mesa isolation so that, if the contact flows, it will mask the GaAs below it and make the mesa slightly larger instead of shorting.

The top Ti-Au layer is to ease wire bonding; the Ti layer prevents the top gold layer from interfering in the contact chemistry. It can be omitted if wire bonding is not necessary, and the gold can be made thicker or thinner at the user's discretion.

### D.3.2 n-GaAs: Pd-Ge-Au Contacts

Pd-Ge-Au[111] and Pd-Ge[112, 113] contacts are sintered rather than alloyed. This solid-phase reaction gives much smoother contacts than Au-Ge, and in the absence of gold, the Au-Ge spikes do not form. Because of this, the Pd-Ge contact is shallow, confined to within approximately 200 Å of the surface of the GaAs[114]. Use of Pd-Ge contacts virtually eliminates shorted devices in our structures.

**Mount** Place a small drop of 5% PMMA on a glass slide. Dip the bottom of the sample in the PMMA droplet and then place it on a different part of the slide. Bake the slide on a hot plate for approximately one minute, until the original drop of PMMA has dried. It may be useful to cut the glass slide before doing this so it fits conveniently into small beakers.

**Surface Clean** Clean the sample surface by dipping in 1:10 HCl:Water by volume, then rinsing in water. This will remove the surface oxide from the GaAs. The oxide will reform in a matter of minutes; it is important to quickly move the sample into the evaporator after this step. (A 5% Ammonia solution also works well for cleaning and oxide removal).

**Metallization**

| **Material** | **Thickness** |
|---|---|
| Pd | 30 nm |
| Ge | 40 nm |
| Au | 100 nm |

The thickness of the Pd-Ge layer can be varied as long as the stoichiometry is kept fix. The thickness of the Au layer can be varied as needed to make bonding easy; it does not participate in the contact geometry.

If the contact is patterned, lift-off (Section D.8) should be performed at this point.

**Anneal** All of our ohmic contacts are annealed using a rapid thermal annealer. We anneal in a constant flow of forming gas (95% $N_2$ 5%$H_2$).

| **Action** | **Temperature** | **Time** |
|---|---|---|
| Ramp | 260° Celsius | 60 seconds |
| Hold | 260° Celsius | 7 minutes |
| Ramp | 425° Celsius | 35 seconds |
| Hold | 425° Celsius | 30 seconds |
| Cool | Room Temperature | |

### D.3.3 p-GaAs: Au-Zn-Au Contacts

Zinc based ohmic contacts[115] to p type GaAs are reliable and safe. Beryllium gold contacts are also common, but suffer from some safety concerns. The first gold layer in this contact is essential as an adhesion layer; removing it will result in non-uniform contacts with high resistances. The second gold layer completes the contact, but also serves to “seal in” the zinc in the evaporator. Because zinc has a high vapor pressure at even moderate temperatures, this step is essential in keeping the base pressure of the system low, and in preventing zinc from being mixed into future depositions.

**Mount** Place a small drop of 5% PMMA on a glass slide. Dip the bottom of the sample in the PMMA droplet and then place it on a different part of the slide. Bake the slide on a hot plate for approximately one minute, until the original drop of PMMA has dried. It may be useful to cut the glass slide before doing this so it fits conveniently into small beakers.

**Surface Clean** Clean the sample surface by dipping in 1:10 HCl:Water by volume, then rinsing in water. This will remove the surface oxide from the GaAs. The oxide will reform in a matter of minutes; it is important to quickly move the sample into the evaporator after this step. (A 5% Ammonia solution also works well for cleaning).

**Metallization**

| **Material** | **Thickness** |
|---|---|
| Au | 30 nm |
| Zn | 100 nm |
| Au | 200 nm |

The thickness of the Zn and Au layers can be varied substantially without regard to stoichiometry without noticeably changing the quality of the contacts. A Ti-Au overcoat can make bonding to this contact easier.

If the contact is patterned, lift-off (Section D.8) should be performed at this point.

**Anneal** All of our ohmic contacts are annealed using a rapid thermal annealer. We anneal in a constant flow of forming gas (95% $N_2$ 5%$H_2$).

| Action | Temperature | Time |
|---|---|---|
| Ramp | 200° Celsius | 30 seconds |
| Hold | 200° Celsius | 30 seconds |
| Ramp | 415° Celsius | 22 seconds |
| Ramp | 435° Celsius | 12 seconds |
| Hold | 435° Celsius | 2 minutes |
| Cool | Room Temperature | |

## D.4 Wet Etching of GaAs

Just as with ohmic contacts, a wide variety of etchants are described in the literature[116].

GaAs is amphoteric, and can be etched by both acids and bases. In general, any mixture of an oxidizing agent and an acid or base will etch GaAs at a substantial rate. There is substantial variation of the etch rate with both Al mole fraction in $Al_xGa_{1-x}As$, and with crystallographic axis[107].

For mesa isolation etches, we typically want a very fast etch without much selectivity; highly selective etches complicate matters when trying to etch through the AlGaAs layers in our heterostructures. We commonly use two different etchants. A citric acid based has the advantage that its etch rate and selectivity can be tuned by varying the concentration of the hydrogen peroxide in it. A sodium hydroxide based etch seems more reliable on Carbon doped samples, and less sensitive to "dirt" on the sample surface (possibly it acts as a surfactant as well as an etch). However, it typically does not give a flat bottom, but rather leaves a valley that bows upwards in the center.

The etchant should be agitated with a stir bar for etches longer than a few seconds; this guarantees that the etch rate will not drop with time as the etchant near the substrate is consumed. It can be helpful to mount the sample on a large glass slide using mounting wax or PMMA to prevent it from being spun around the dish by the stir bar.

**Citric Acid and Hydrogen Peroxide[117, 118]** This is a selective etch; it removes $\sim$200 nm/minute GaAs, 20 nm/minute $Al_{0.3}Ga_{0.7}As$

10 parts 50% by weight citric acid solution

1 part 30% Hydrogen Peroxide

**Sodium Hydroxide and Hydrogen Peroxide[107, 119]**

This etch is relatively non-selective, and works at about 400 nm/minute.

1 part 1 Molar Sodium Hydroxide

1 part 30% Hydrogen Peroxide

10 parts Water

## D.5 Photolithography

We use both positive and negative resists in our lithography. In general, the undercut provided by the negative resists gives better performance in lift-off, while swelling features when used in subtractive processes. Positive resists, on the other hand, tend to be hard to use for lift-off, but give good feature control for etching.

## D.6 Negative Resist

We use Futurex's NR7-1000PY negative resist ("Futurex").

**Preparation** Preheat a hot plate to 100° Celsius , and turn on the mask aligner. Allow the aligner to warm up for at least 10 minutes before using.

**Surface Clean** If there are visible particles on the surface of the sample, ultrasound in a 2% $\mu$90 solution for 30 seconds, followed by a rinse in DI water.

If humidity is high, pre-bake the sample on a 100° Celsius hot plate to remove any moisture from the surface.

Place the sample on the spinner chuck, and start the spinner. From squirt bottles, spray the bottle with a steady stream of acetone for approximately 5 seconds, followed by 1 second of simultaneous acetone and isopropanol, followed by 5 seconds of isopropanol. Allow the sample to spin until dry.

**Spin** With the sample stationary, apply a small drop of Futurex NR7-1000PY with a clean dropper. There is no need to coat the entire surface of the sample.

Spin at 5000 RPM for 1 minute. (For thick metallizations, it may be useful to spin at a lower speed. For thin metallizations, a higher spin speed may be useful for fine features. Exposure and development times will have to be adjusted accordingly.)

**First Bake** Immediately remove the sample and place it directly on the 100° Celsius hot plate for 60 seconds. (For thick samples, or samples with a rough back, this may not adequately bake the resist. If the sample sticks to the mask in the next step, increase the time to 90 seconds.)

**Exposure** Remove the sample, align in the mask aligner, and expose for one minute at an intensity of 5.5 mW $\text{cm}^{-2}$[†]

**Second Bake** Bake the sample directly on the 100° Celsius hot plate for 60 seconds.

**Develop** Develop in Futurex RD6 for 15 seconds. If processing many samples, refresh the developer frequently. Rinse in two separate beakers of DI water for 15 seconds each.

**Examine** Examine the pattern in the Leitz microscope with the UV filter in place. If the sample looks underdeveloped (resist still in unexposed regions), develop longer. The undercut can be deepened by exposing the sample for less time or, with slightly less effect, developing for longer. The undercut can be decreased by exposing for longer, or, with slightly less effect, developing

---

[†] This number assumes the dose calibration on our aligner is accurate. There is no reason to assume this. Exposure times will need to be adjusted accordingly.

for a shorter period of time. Now is a good time to take a picture of the patterning to help resolve issues later in processing.

## D.7 Positive Resist

We use Shippley's Microposit 1813 positive resist. This resist is slower to work with than Futurex, so it is usually undesirable to use it unless Futurex's undercut is untenable.

**Preparation** Preheat an oven to 80° Celsius , and turn on the mask aligner. Allow the aligner to warm up for at least 10 minutes before using.

**Surface Clean** If there are visible particles on the surface of the sample, ultrasound in a 2% $\mu$90 solution for 30 seconds, followed by a rinse in DI water.

If humidity is high, pre-bake the sample on a 100° Celsius hot plate to remove any moisture from the surface.

Place the sample on the spinner chuck, and start the spinner. From squirt bottles, spray the bottle with a steady stream of acetone for approximately 5 seconds, followed by 1 second of simultaneous acetone and isopropanol, followed by 5 seconds of isopropanol. Allow the sample to spin until dry.

**Spin** With the sample stationary, apply a small drop of Shippley 1813 with a clean dropper. There is no need to coat the entire surface of the sample.

Spin at 4000 RPM for 1 minute. (For thick metallizations, it may be useful to spin at a lower speed. For thin metallizations, a higher spin speed may be useful for fine features. Exposure and development times will have to be adjusted accordingly.)

**Bake** Bake in an oven at 80° Celsius for one hour.

**Exposure** Remove the sample, align in the mask aligner, and expose for one minute at an intensity of 5.5 mW $\mathrm{cm}^{-2\dagger}$

**Develop** Develop in Shippley MF-319 for 45 seconds. If processing many samples, refresh the developer frequently. Rinse in two separate beakers of DI water for 15 seconds each.

**Examine** Examine the pattern in the Leitz microscope with the UV filter in place. If the sample looks underdeveloped (resist still in exposed regions), develop longer in 15 second increments. If the patterning is too poor to use, flood expose the entire sample by putting it in the aligner with no mask. It will make removing the resist with acetone much easier.

## D.8 Lift-Off

Pattern as described in Section D.5, ideally with a negative resist and good undercut. Evaporate the desired material on the surface of the sample.

Soak the sample in acetone for at least 5 minutes, or as long as several hours. Do not allow the sample to dry once it has been covered in acetone; it will redeposit dissolved resist on the surface, making it very difficult to do any further processing. If the metal does not readily lift off the surface of the sample, several options are available. Covering the beaker with a watch glass and putting it on a 60 degree hot plate will speed removal of the resist. Alternatively, using a syringe to strike the surface of the submerged sample with a jet of acetone will wash away the metal. Placing the sample in a low powered ultrasonic bath for several minutes can also help remove the metal. As a last result, a q-tip can be used to gently remove remaining metal.

Transfer the sample from the acetone into a beaker full of isopropanol without allowing it to dry. Gently agitate it in the beaker for several seconds to remove the acetone and then blow dry with dry nitrogen.

---

[†] This number assumes the dose calibration on our aligner is accurate. There is no reason to believe this. Exposure times will need to be adjusted accordingly.

# Bibliography

[1] Tsuneya Ando, Alan B. Fowler, and Frank Stern. Electronic properties of two-dimensional systems. *Rev. Mod. Phys.*, 54(2):437–672, Apr 1982. doi:10.1103/RevModPhys.54.437. 1

[2] A. B. Fowler, F. F. Fang, W. E. Howard, and P. J. Stiles. Magneto-oscillatory conductance in silicon surfaces. *Phys. Rev. Lett.*, 16(20):901–903, May 1966. doi:10.1103/PhysRevLett.16.901. 1, 2.2

[3] S. A. J. Wiegers, M. Specht, L. P. Lévy, M. Y. Simmons, D. A. Ritchie, A. Cavanna, B. Etienne, G. Martinez, and P. Wyder. Magnetization and energy gaps of a high-mobility 2D electron gas in the quantum limit. *Phys. Rev. Lett.*, 79(17):3238–3241, Oct 1997. doi:10.1103/PhysRevLett.79.3238. 1

[4] T. P. Smith, B. B. Goldberg, P. J. Stiles, and M. Heiblum. Direct measurement of the density of states of a two-dimensional electron gas. *Phys. Rev. B*, 32(4):2696–2699, Aug 1985. doi:10.1103/PhysRevB.32.2696. 1

[5] S. E. Barrett, G. Dabbagh, L. N. Pfeiffer, K. W. West, and R. Tycko. Optically pumped NMR evidence for finite-size skyrmions in GaAs quantum wells near Landau level filling $\nu = 1$. *Phys. Rev. Lett.*, 74(25):5112–5115, Jun 1995. doi:10.1103/PhysRevLett.74.5112. 1

[6] A. Pinczuk, J. P. Valladares, D. Heiman, A. C. Gossard, J. H. English, C. W. Tu, L. Pfeiffer, and K. West. Observation of roton density of states in two-dimensional Landau-level excitations. *Phys. Rev. Lett.*, 61(23):2701–2704, Dec 1988. doi:10.1103/PhysRevLett.61.2701. 1, 5.2

[7] J. R. Schrieffer. What is a quasi-particle? In *Electronic Density of States, NBS Special Publication 323*, pages 227–232. Nat. Bur. Stand., December 1971. 2, 2-1

[8] L. Hedin, B. I. Lundqvist, and S. Lundqvist. Beyond the one-electron approximation: Density of states for interacting electrons. In *Electronic Density of States, NBS Special Publication 323*, pages 233–249. Nat. Bur. Stand., December 1971. 2, 2

[9] Richard D. Mattuck. *A Guide to Feynman Diagrams in the Many-Body Problem*. Dover Publications, Inc., Second edition, 1976. 2, †, 7.2

[10] L. Hedin and S. Lundqvist. In H. Ehrenreich, F. Seitz, and D. Turnbull, editors, *Solid State Physics*, volume 23, New York, 1969. Academic. †

[11] C. Guillemot and F. Clérot. Coulomb lifetime and electronic distribution function in a drifting two-dimensional electron gas. *Phys. Rev. B*, 47(12):7227–7239, Mar 1993. doi:10.1103/PhysRevB.47.7227. 2-2, 8.1, 8-1

[12] L. Hedin, B.I. Lundqvist, and S. Lunqvist. New structure in the single-particle spectrum of an electron gas. *Solid State Communications*, 5:237–239, 1967. 2, †, 8.1, 8-1

[13] Joel I Gersten. Theory of surface plasmarons. *Surface Science*, 92:579–592, 1980. †

[14] L. D. Landau. The theory of a Fermi liquid. *Sov. Phys. JETP*, 3(6):920–925, Jan 1957. 2, 7.2

[15] L. D. Landau. Oscillations in a Fermi liquid. *Sov. Phys. JETP*, 5(1):101–108, Aug 1957. 2, 7.2

[16] Kai M. Siegbahn. Electron spectroscopy for atoms, molecules and condensed matter. In Tore Frängsmyr and Ekspång Gösta, editors, *Nobel Lectures, Physics 1981-1990*, Singapore, 1993. World Scientific Publishing

Co. Available from: `http://nobelprize.org/nobel_prizes/physics/laureates/1981/siegbahn-lecture.html`. 2.1

[17] Andrea Damascelli. Probing the electronic structure of complex systems by ARPES. *Physica Scripta*, 2004(T109):61–74, 2004. doi:10.1238/Physica.Topical.109a00061. 2.1

[18] S. Y. Zhou, G.-H. Gweon, J. Graf, A. V. Fedorov, C. D. Spataru, R. D. Diehl, Y. Kopelevich, D.-H. Lee, Steven G. Louie, and A. Lanzara. First direct observation of Dirac fermions in graphite. *Nat. Phys.*, (9):595–599, 2006. doi:10.1038/nphys393. 2-3

[19] P. D. Johnson and S. L. Hulbert. Inverse photoemission. *Review of Scientific Instruments*, 61(9):2277–2288, 1990. doi:10.1063/1.1141352. 2.1

[20] Friedrich Reinert and Stefan Hüfner. Photoemission spectroscopy&mdash;from early days to recent applications. *New Journal of Physics*, 7:97, 2005. doi:doi:10.1088/1367-2630/7/1/097. †

[21] Ivar Giaever. Energy gap in superconductors measured by electron tunneling. *Phys. Rev. Lett.*, 5(4):147–148, Aug 1960. doi:10.1103/PhysRevLett.5.147. 2-5, 2.1

[22] T. W. Hickmott. Capacitance measurements of magnetic localization and magnetic freezeout in $n^-$-type GaAs. *Phys. Rev. B*, 38(17):12404–12415, Dec 1988. doi:10.1103/PhysRevB.38.12404. ‡

[23] J. P. Eisenstein, T. J. Gramila, L. N. Pfeiffer, and K. W. West. Probing a two-dimensional Fermi surface by tunneling. *Phys. Rev. B*, 44(12):6511–6514, Sep 1991. doi:10.1103/PhysRevB.44.6511. 3

[24] L.P. Rokhinson, B. Su, and V.J. Goldman. Peak values of conductivity in integer and fractional quantum Hall effect. *arXiv.org:cond-mat*, page 9508081, 1995. arXiv:cond-mat/9508081. 3-2

[25] E. E. Mendez, L. Esaki, and W. I. Wang. Resonant magnetotunneling in GaAlAs-GaAs-GaAlAs heterostructures. *Phys. Rev. B*, 33(4):2893–2896, Feb 1986. doi:10.1103/PhysRevB.33.2893. 3

[26] P. C. Main, A. S. G. Thornton, R. J. A. Hill, S. T. Stoddart, T. Ihn, L. Eaves, K. A. Benedict, and M. Henini. Landau-level spectroscopy of a two-dimensional electron system by tunneling through a quantum dot. *Phys. Rev. Lett.*, 84(4):729–732, Jan 2000. doi:10.1103/PhysRevLett.84.729. 3

[27] H. B. Chan. *Tunneling Spectroscopy of the Two-Dimensional Electron Gas*. PhD thesis, Massachusetts Institute of Technology, 1993. Available from: `http://electron.mit.edu/theses/Ho_Bun_Chan.pdf`. 3.1, 3.7, 8.2, †, B

[28] H. B. Chan, P. I. Glicofridis, R. C. Ashoori, and M. R. Melloch. Universal linear density of states for tunneling into the two-dimensional electron gas in a magnetic field. *Phys. Rev. Lett.*, 79(15):2867–2870, Oct 1997. doi:10.1103/PhysRevLett.79.2867. 3.1, 3.7

[29] G. A. Steele. *Imaging Transport Resonances in the Quantum Hall Effect*. PhD thesis, Massachusetts Institute of Technology, 2006. Available from: `http://electron.mit.edu/theses/Gary_Steele.pdf`. 3.1, A.4, C.1

[30] H. L. Stormer. Electron mobilities in modulation-doped GaAs-(AlGa)As heterostructures. *Surface Science*, 132:519–526, 1983. 3.1

[31] B. E. Kane, L. N. Pfeiffer, and K. W. West. High mobility GaAs heterostructure field effect transistor for nanofabrication in which dopant-induced disorder is eliminated. *Appl. Phys. Lett.*, 67(9):1262–1264, 1995. doi:10.1063/1.114391. 3.1

[32] John H. Davies. *The Physics of Low-Dimensional Semiconductors*. Cambridge University Press, 1998. 3.1, C.2

[33] R. C. Ashoori. *The Density of States in the Two-Dimensional Electron Gas and Quantum Dots*. PhD thesis, Cornell University, 1991. arXiv:cond-mat/0607739. 3.3, 5.2, 11.1, 11.1, 11.1, 12.1, 12.2, 12.5, A.4

[34] V. G. Popov. Self-consistent equilibrium of a two-dimensional electron system with a reservoir in a quantizing magnetic field: Analytical approach. *Physical Review B (Condensed Matter and Materials Physics)*, 73(12):125310, 2006. doi:10.1103/PhysRevB.73.125310. 3.3

[35] R. C. Ashoori, J. A. Lebens, N. P. Bigelow, and R. H. Silsbee. Equilibrium tunneling from the 2-dimensional electron-gas in GaAs - evidence for a magnetic-field-induced energy-gap. *Physical Review Letters*, 64(6):681–684, 1990. doi:10.1103/PhysRevLett.64.681. 3.7

[36] J. P. Eisenstein, L. N. Pfeiffer, and K. W. West. Coulomb barrier to tunneling between parallel two-dimensional electron systems. *Phys. Rev. Lett.*, 69(26):3804–3807, Dec 1992. doi:10.1103/PhysRevLett.69.3804. 3.7

[37] Song He, P. M. Platzman, and B. I. Halperin. Tunneling into a two-dimensional electron system in a strong magnetic field. *Phys. Rev. Lett.*, 71(5):777–780, Aug 1993. doi:10.1103/PhysRevLett.71.777. 3.7

[38] F. G. Pikus and A. L. Efros. Coulomb gap in a two-dimensional electron gas with a close metallic electrode. *Phys. Rev. B*, 51(23):16871–16877, Jun 1995. doi:10.1103/PhysRevB.51.16871. 3.7

[39] Y. Hatsugai, P.-A. Bares, and X. G. Wen. Electron spectral function of an interacting two dimensional electron gas in a strong magnetic field. *Phys. Rev. Lett.*, 71(3):424–427, Jul 1993. doi:10.1103/PhysRevLett.71.424. 3.7

[40] A L Efros and B I Shklovskii. Coulomb gap and low temperature conductivity of disordered systems. *Journal of Physics C: Solid State Physics*, 8(4):L49–L51, 1975. doi:10.1088/0022-3719/8/4/003. 3.7

[41] William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing*. Cambridge University Press, Second edition, 1992. Available from: `http://www.nrbook.com/`. 3.9, C.1, C.2, C.3

[42] A. Usher, R. J. Nicholas, J. J. Harris, and C. T. Foxon. Observation of magnetic excitons and spin waves in activation studies of a two-dimensional electron gas. *Phys. Rev. B*, 41(2):1129–1134, Jan 1990. doi:10.1103/PhysRevB.41.1129. 5.2, 5.3, 12.1

[43] M. Hayne, A. Usher, J. J. Harris, and C. T. Foxon. Exchange enhancement of the Landau-level separation for two-dimensional electrons in GaAs/$\mathrm{Ga_{1-x}Al_xAs}$ heterojunctions. *Phys. Rev. B*, 46(15):9515–9519, Oct 1992. doi:10.1103/PhysRevB.46.9515. 5.2

[44] D. R. Leadley, R. J. Nicholas, J. J. Harris, and C. T. Foxon. Critical collapse of the exchange-enhanced spin splitting in two-dimensional systems. *Phys. Rev. B*, 58(19):13036–13046, Nov 1998. doi:10.1103/PhysRevB.58.13036. 5.2

[45] A. Schmeller, J. P. Eisenstein, L. N. Pfeiffer, and K. W. West. Evidence for skyrmions and single spin flips in the integer quantized Hall effect. *Phys. Rev. Lett.*, 75(23):4290–4293, Dec 1995. doi:10.1103/PhysRevLett.75.4290. 5.2

[46] D. K. Maude, M. Potemski, J. C. Portal, M. Henini, L. Eaves, G. Hill, and M. A. Pate. Spin excitations of a two-dimensional electron gas in the limit of vanishing Landé $g$ factor. *Phys. Rev. Lett.*, 77(22):4604–4607, Nov 1996. doi:10.1103/PhysRevLett.77.4604. 5.2

[47] A. Pinczuk, B. S. Dennis, D. Heiman, C. Kallin, L. Brey, C. Tejedor, S. Schmitt-Rink, L. N. Pfeiffer, and K. W. West. Spectroscopic measurement of large exchange enhancement of a spin-polarized 2D electron gas. *Phys. Rev. Lett.*, 68(24):3623–3626, Jun 1992. doi:10.1103/PhysRevLett.68.3623. 5-2, 5.2

[48] V. T. Dolgopolov, A. A. Shashkin, A. V. Aristov, D. Schmerek, W. Hansen, J. P. Kotthaus, and M. Holland. Direct measurements of the spin gap in the two-dimensional electron gas of AlGaAs-GaAs heterojunctions. *Phys. Rev. Lett.*, 79(4):729–732, Jul 1997. doi:10.1103/PhysRevLett.79.729. 5.2, 5.3, 11.1, 11.1, 11.1, 12.1

[49] J. P. Eisenstein, L. N. Pfeiffer, and K. W. West. Compressibility of the two-dimensional electron gas: Measurements of the zero-field exchange energy and fractional quantum hall gap. *Phys. Rev. B*, 50(3):1760–1778, Jul 1994. doi:10.1103/PhysRevB.50.1760. 5.2, 6.2

[50] I. V. Kukushkin and V. B. Timofeev. Radiative recombination of 2D electrons with nonequilibrium holes in silicon metal-insulator-semiconductor structures. *Sov. Phys. JETP*, 65(1):146–157, Jan 1987. 5.2

[51] I. V. Kukushkin and V. B. Timofeev. Magneto-optics of strongly correlated two-dimensional electrons in single heterojunctions. *Advances in Physics*, 45(3):147–242, 1996. doi:10.1080/00018739600101487. 5.2

[52] Sankar Das Sarma and M. A. Pinczuk, editors. *Perspectives in Quantum Hall Effects: Novel Quantum Liquids in Low-Dimensional Semiconductor Structures*. Wiley-Interscience, 1996. 5.2

[53] Walter Kohn. Cyclotron resonance and de Haas-van Alphen oscillations of an interacting electron gas. *Phys. Rev.*, 123(4):1242–1244, Aug 1961. doi:10.1103/PhysRev.123.1242. 5.2

[54] J. F. Janak. $g$ factor of the two-dimensional interacting electron gas. *Phys. Rev.*, 178(3):1416–1418, Feb 1969. doi:10.1103/PhysRev.178.1416. 5.2

[55] C. Kallin and B. I. Halperin. Excitations from a filled landau level in the two-dimensional electron gas. *Phys. Rev. B*, 30(10):5655–5668, Nov 1984. doi:10.1103/PhysRevB.30.5655. 5.2, 8.1

[56] M. A. Eriksson, A. Pinczuk, B. S. Dennis, S. H. Simon, L. N. Pfeiffer, and K. W. West. Collective excitations in the dilute 2D electron system. *Phys. Rev. Lett.*, 82(10):2163–2166, Mar 1999. doi:10.1103/PhysRevLett.82.2163. 5.2

[57] C. F. Hirjibehedin, A. Pinczuk, B. S. Dennis, L. N. Pfeiffer, and K. W. West. Evidence of electron correlations in plasmon dispersions of ultralow density two-dimensional electron systems. *Phys. Rev. B*, 65(16):161309, Apr 2002. doi:10.1103/PhysRevB.65.161309. 5.2

[58] V. S. Khrapai, A. A. Shashkin, E. L. Shangina, V. Pellegrini, F. Beltram, G. Biasiol, and L. Sorba. Spin gap in the two-dimensional electron system of GaAs/$\mathrm{Al_xGa_{1-x}As}$ single heterojunctions in weak magnetic fields. *Phys. Rev. B*, 72(3):035344, 2005. doi:10.1103/PhysRevB.72.035344. 5.2

[59] I. L. Aleiner and L. I. Glazman. Two-dimensional electron liquid in a weak magnetic field. *Phys. Rev. B*, 52(15):11296–11312, Oct 1995. doi:10.1103/PhysRevB.52.11296. 5.3, 5.3

[60] W Knap, V I Fal′ko, E Frayssinet, P Lorenzini, N Grandjean, D Maude, G Karczewski, B L Brandt, J Łusakowski, I Grzegory, M Leszczyǹski, P Prystawko, C Skierbiszewski, S Porowski, X Hu, G Simin, M Asif Khan, and M S Shur. Spin and interaction effects in Shubnikov–de Haas oscillations and the quantum Hall effect in GaN/AlGaN heterostructures. *Journal of Physics: Condensed Matter*, 16(20):3421–3432, 2004. doi:10.1088/0953-8984/16/20/013. 5.3, 6.4

[61] A. P. Smith, A. H. MacDonald, and G. Gumbs. Quasiparticle effective mass and enhanced g factor for a two-dimensional electron gas at intermediate magnetic fields. *Phys. Rev. B*, 45(15):8829–8832, Apr 1992. doi:10.1103/PhysRevB.45.8829. 5.3, 6.4, 12.1

[62] S. V. Iordanski and A. Kasbuba. Excitations in a quantum Hall ferromagnet with strong Coulomb interaction. *JETP Letters*, 75(7):348–353, 2002. doi:10.1134/1.1485265. 5.3, 6.4

[63] A. H. MacDonald, H. C. A. Oji, and K. L. Liu. Thermodynamic properties of an interacting two-dimensional electron gas in a strong magnetic field. *Phys. Rev. B*, 34(4):2681–2689, Aug 1986. doi:10.1103/PhysRevB.34.2681. 5.3, 5.5, 6, 6-1, 6.3, 6-4, 6.4

[64] I. L. Aleiner, H. U. Baranger, and L. I. Glazman. Tunneling into a two-dimensional electron liquid in a weak magnetic field. *Phys. Rev. Lett.*, 74(17):3435–3438, Apr 1995. doi:10.1103/PhysRevLett.74.3435. 5.3

[65] Tsuneya Ando and Yasutada Uemura. Theory of oscillatory g factor in an MOS inversion layer under strong magnetic fields. *J. Phys. Soc. Jpn.*, 37(4):1044–1052, 1974. doi:10.1143/JPSJ.37.1044. 5.5

[66] J. P. Eisenstein, L. N. Pfeiffer, and K. W. West. Negative compressibility of interacting two-dimensional electron and quasiparticle gases. *Phys. Rev. Lett.*, 68(5):674–677, Feb 1992. doi:10.1103/PhysRevLett.68.674. 5.5, 6.2, 8.5

[67] S. V. Kravchenko, J. M. Caulfield, J. Singleton, Hans Nielsen, and V. M. Pudalov. Electron-electron interactions in the two-dimensional electron gas in silicon. *Phys. Rev. B*, 47(19):12961–12963, May 1993. doi:10.1103/PhysRevB.47.12961. 5.5

[68] I. V. Kukushkin, J. H. Smet, D. S. Lyne Abergel, V. I. Fal'ko, W. Wegscheider, and K. von Klitzing. Detection of the electron spin resonance of two-dimensional electrons at large wave vectors. *Physical Review Letters*, 96(12):126807, 2006. doi:10.1103/PhysRevLett.96.126807. 6.4

[69] L. D. Landau. On the theory of the Fermi liquid. *Sov. Phys. JETP*, 35(8):70–74, Jan 1959. 7.2

[70] W. G. Baber. The contribution to the electrical resistance of metals from collisions between electrons. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, 158(894):383–396, Jan

1937. Available from: `http://links.jstor.org/sici?sici=0080-4630%2819370115%29158%3A894%3C383%3ATCTTER%3E2.0.CO%3B2-K`. 7.2, 7.4

[71] John W. Negele and Henri Orland. *Quantum Many-Particle Systems*. Perseus Boosk, 1998. 7.2

[72] Tsuneya Ando, Yukio Matsumoto, and Yasutada Uemura. Theory of Hall effect in a two-dimensional electron system. *Journal of the Physical Society of Japan*, 39(2):279–288, Aug 1975. doi:10.1143/JPSJ.39.279. 7.3

[73] A. V. Chaplik. Energy spectrum and electron scattering processes in inversion layers. *Sov. Phys. – JETP*, 33:997, 1971. 7.4

[74] A. V. Chaplik. *Zh. Eksp. Teor. Fiz.*, 60:1845, 1971. 7.4

[75] A. Yacoby, U. Sivan, C. P. Umbach, and J. M. Hong. Interference and dephasing by electron-electron interaction on length scales shorter than the elastic mean free path. *Phys. Rev. Lett.*, 66(14):1938–1941, Apr 1991. doi:10.1103/PhysRevLett.66.1938. 7.4

[76] S. Q. Murphy, J. P. Eisenstein, L. N. Pfeiffer, and K. W. West. Lifetime of two-dimensional electrons measured by tunneling spectroscopy. *Phys. Rev. B*, 52(20):14825–14828, Nov 1995. doi:10.1103/PhysRevB.52.14825. 7.4

[77] Gabriele F. Giuliani and John J. Quinn. Lifetime of a quasiparticle in a two-dimensional electron gas. *Phys. Rev. B*, 26(8):4421–4428, Oct 1982. doi:10.1103/PhysRevB.26.4421. 7.4

[78] Paul von Allmen. Plasmaron excitation and band renormalization in a two-dimensional electron gas. *Phys. Rev. B*, 46(20):13345–13350, Nov 1992. doi:10.1103/PhysRevB.46.13345. 8.1, 8-1

[79] R. Jalabert and S. Das Sarma. Quasiparticle properties of a coupled two-dimensional electron-phonon system. *Phys. Rev. B*, 40(14):9723–9737, Nov 1989. doi:10.1103/PhysRevB.40.9723. 8.1, 8-1

[80] J. R. Cuthill, R. C. Dobbyn, A. J. McAlister, and M. L. Williams. Search for plasmaron structure in the soft X-ray $L_{2,3}$ emission spectrum of Al. *Phys. Rev.*, 174(2):515–517, Oct 1968. doi:10.1103/PhysRev.174.515. 8.1

[81] Neville V. Smith and William E. Spicer. Photoemission studies of the alkali metals. I. Sodium and potassium. *Phys. Rev.*, 188(2):593–605, Dec 1969. doi:10.1103/PhysRev.188.593. 8.1

[82] Clas Blomberg and Birger Bergersen. Spurious structure from approximations to the Dyson equation. *Can. J. Phys.*, 50:2286–2293, 1972. 8.1

[83] B. Bergersen, F.W. Kus, and C. Bolmberg. Single particle Green's function in the electron-plasmon approximation. *Can. J. Phys.*, 51(102):102–110, 1973. 8.1

[84] R. Jalabert and S. Das Sarma. Many-polaron interaction effects in two dimensions. *Phys. Rev. B*, 39(8):5542–5545, Mar 1989. doi:10.1103/PhysRevB.39.5542. 8.1, 8.5

[85] C. Kallin and B. I. Halperin. Many-body effects on the cyclotron resonance in a two-dimensional electron gas. *Phys. Rev. B*, 31(6):3635–3647, Mar 1985. doi:10.1103/PhysRevB.31.3635. 8.1

[86] H. B. Chan, R. C. Ashoori, L. N. Pfeiffer, and K. W. West. Tunneling into ferromagnetic quantum Hall states: Observation of a spin bottleneck. *Phys. Rev. Lett.*, 83(16):3258–3261, Oct 1999. doi:10.1103/PhysRevLett.83.3258. 8.2

[87] Y.-W. Tan, J. Zhu, H. L. Stormer, L. N. Pfeiffer, K. W. Baldwin, and K. W. West. Measurements of the density-dependent many-body electron mass in two dimensional GaAs/AlGaAs heterostructures. *Physical Review Letters*, 94(1):016405, 2005. doi:10.1103/PhysRevLett.94.016405. 8.3

[88] Jainendra K. Jain and Michael R. Peterson. Reconstructing the electron in a fractionalized quantum fluid. *Physical Review Letters*, 94(18):186808, 2005. doi:10.1103/PhysRevLett.94.186808. 8.4

[89] S. Ilani, A. Yacoby, D. Mahalu, and Hadas Shtrikman. Unexpected behavior of the local compressibility near the $B = 0$ metal-insulator transition. *Phys. Rev. Lett.*, 84(14):3133–3136, Apr 2000. doi:10.1103/PhysRevLett.84.3133. 8.5

[90] K.S. Novoselov, A.K. Geim, S.V. Morozov, D. Jiang, M.I. Katsnelson, and I.V. Grigorieva. Two-dimensional gas of massless Dirac fermions in graphene. *Nature*, 438:197–200, 2005. doi:10.1038/nature04233. 8.6

[91] Yuanbo Zhang, Yan-Wen Tan, Horst L. Stormer, and Philip Kim. Experimental observation of the quantum Hall effect and Berry's phase in graphene. *Nature*, 438:201–204, 2005. doi:10.1038/nature04235. 8.6

[92] J. R. Albright. Integrals of products of Airy functions. *Journal of Physics A: Mathematical and General*, 10(4):485–490, 1977. doi:10.1088/0305-4470/10/4/011. 9.2

[93] J. A. Lebens, R. H. Silsbee, and S. L. Wright. Tunneling and transverse wave vector conservation in GaAs/AlGaAs heterostructures. *Appl. Phys. Lett.*, 51(11):840–842, 1987. doi:10.1063/1.98830. 10.2

[94] J. Bardeen. Tunnelling from a many-particle point of view. *Phys. Rev. Lett.*, 6(2):57–59, Jan 1961. doi:10.1103/PhysRevLett.6.57. 10.3

[95] J. Smoliner. Tunnelling spectroscopy of low-dimensional states. *Semicond. Sci. Technol.*, 11:1–16, 1996. doi:10.1088/0268-1242/11/1/005. 10.3

[96] Ben G. Streeman. *Solid State Electronic Devices*. Prentice Hall, 1995. 11.1, 11.2

[97] Loren Pfeiffer, E. F. Schubert, K. W. West, and C. W. Magee. Si dopant migration and the AlGaAs/GaAs inverted interface. *Applied Physics Letters*, 58(20):2258–2260, 1991. doi:10.1063/1.104915. 12.5

[98] Jr. A. P. Mills, L. N. Pfeiffer, K. W. West, and C. W. Magee. Mechanisms for Si dopant migration in molecular beam epitaxy $\mathrm{Al_xGa_{1-x}As}$. *Journal of Applied Physics*, 88(7):4056–4060, 2000. doi:10.1063/1.1308073. 12.5

[99] Aleem Siddiqui. *Use of Time Domain Capacitance Spectroscopy in the Study of Tunneling into Two-Dimensional GaAs/AlGaAs Heterostructures with an in-Plane Magnetic Field*. Master's thesis, Massachusetts Institute of Technology, 2003. A.1

[100] F. S. Crawford Jr. *Berkeley Physics Course 3. Waves*. McGraw-Hill Book Co., 1968. A.2

[101] S. Urazhdin, S. H. Tessmer, and R. C. Ashoori. A simple low-dissipation amplifier for cryogenic scanning tunneling microscopy. *Rev. Sci. Instrum.*, 73(2):310–312, Feb 2002. doi:10.1063/1.1433951. A.4

[102] B. H. Hu and C. H. Yang. A cryogenic voltage amplifier with 36 MHz bandwidth using discrete GaAs metal-semiconductor field-effect transistors. *Rev. Sci. Instrum.*, 76:124702, Dec 2005. doi:10.1063/1.2140224. A.4

[103] G. L. Snider and E. L. Hu. Electron states in mesa-etched one-dimensional quantum well wires. *J. Appl. Phys*, 68(6):2849–2853, Sep 1990. doi:10.1063/1.346443. C

[104] I-H. Tan, L. Snider, L. D. Chang, and E. L. Hu. A self-consistent solution of Schrödinger-Poisson equations using a nonuniform mesh. *J. Appl. Phys*, 68(8):4071–4076, Oct 1990. doi:10.1063/1.346245. C, C.2

[105] Clifford M. Krowne. Semiconductor heterostructure nonlinear Poisson equation. *J. Appl. Phys.*, 65(4):1602–1614, 1989. doi:10.1063/1.342952. C.1

[106] Gene H. Golub and Charles F. van Loan. *Matrix Computations*. The Johns Hopkins University Press, Third edition, 1996. C.2, C.3

[107] Ralph Williams. *Modern GaAs Processing Methods*. Artech House, 1990. D.3, D.4

[108] A. G. Baca, F. Ren, J.C. Zolper, R. D. Briggs, and S. J. Pearton. A survey of ohmic contacts to III-V compound semiconductors. *Thin Solid Films*, 308–309:699–606, Oct 1997. doi:10.1016/S0040-6090(97)00439-2. D.3

[109] Liliental-Weber Zuzanna, R. Gronsky, J. Washburn, N. Newman, W. E. Spicer, and E.R. Weber. Schottky and ohmic Au contacts on GaAs: Microscopic and electrical investigation. *J. Vac. Sci. Technol. B*, pages 912–918, Jul/Aug 1986. doi:10.1116/1.583536. D.3.1

[110] Masanori Murakami, K.D. Childs, John M. Baker, and A. Callegari. Microstructure studies of AuNiGe Ohmic contacts to n-type GaAs. *J. Vac. Sci. Technol. B*, 4(4):903–911, Jul/Aug 1986. doi:10.1116/1.583535. D.3.1

[111] C.L. Chen, L.J. Mahoney, M.C. Finn, R.C. Brooks, and A. Chu. Low resistance Pd/Ge/Au and Ge/Pd/Au ohmic contacts to n-type GaAs. *Applied Physics Letters*, 48:535–537, February 1986. doi:10.1063/1.96498. D.3.2

[112] E.D. Marshall, W.X. Chen, C.S. Wu, S. S. Lau, and T.F. Kuech. Non-alloyed ohmic contact to n-GaAs by solid phase epitaxy. *Appl. Phys. Lett.*, 47(3):298–300, August 1985. doi:10.1063/1.96198. D.3.2

[113] E. D. Marshall, B. Zhang, L.C. Wang, W.X. Jiao, and T. Sawada Chen. Nonalloyed ohmic contacts to n-GaAs by solid-phase epitaxy of Ge. *J. Appl. Phys.*, 62(3):942–947, Aug 1987. doi:10.1063/1.339705. D.3.2

[114] N. K. Patel, J. H. Burroughes, M. J. Tribble, E. H. Linfield, A. C. Churchill, D. A. Ritchie, and G. A. C. Jones. Independent contacting to electron layers in a double quantum well system using Pd-Ge shallow ohmic contacts. *Applied Physics Letters*, 65(7):851–853, 1994. doi:10.1063/1.112180. D.3.2

[115] D.E. Kren, A. A. Rezazadeh, and P. K. Rees. Low ohmic contacts to C-doped p-GaAs with Au/Zn/Au structure. *Electronics Letters*, 28(13):1248–1250, June 1992. Available from: `http://ieeexplore.ieee.`

`org/iel1/2220/3862/00144364.pdf?isnumber=3862&prod=STD&arnumber=144364&arnumber=144364&arSt=1248&ared=1250&arAuthor=Kren%2C+D.E.%3B+Rezazadeh%2C+A.A.%3B+Rees%2C+P.K.` D.3.3

[116] Kirt R. Williams. Etch rates for micromachining processing – part II. *Journal of Microelectromechanical Systems*, 12(6):761–778, 2003. doi:10.1109/JMEMS.2003.820936. D.4

[117] C. Juang, K. J. Kuhn, and R.B. Darling. Selective etching of GaAs and $\mathrm{Al_{0.30}Ga_{0.70}As}$ with citric acid/hydrogen peroxide solutions. *J. Vac. Sci. Technol. B*, 8(5):1122–1124, 1990. doi:10.1116/1.584928. D.4

[118] T. Kitano, S. Izumi, H. Minami, T. Ishikawa, K. Sato, T. Sonoda, and M. Otsubo. Selective wet etching for highly uniform $\mathrm{GaAs/Al_{0.15}Ga_{0.85}As}$ heterostructure field effect transistors. *J. Vac. Sci. Technol. B*, 15(1):167–170, 1997. doi:10.1116/1.589243. D.4

[119] Takeshi Kobayashi and Sugiyama Koichi. Etching of $\mathrm{Al_xGa_{1-x}As}$ in alkaline solution. *Japan. J. Appl. Phys.*, 12(4):619–620, 1973. doi:10.1143/JJAP.12.619. D.4