AN ABSTRACT OF THE THESIS OF

Allen David Gunther for the degree of Doctor of Philosophy in Physics presented on March 10, 2000.

Title: Transport in Silicon Metal Oxide Semiconductor Quantum Dots.

Redacted for Privacy

Abstract approved:

Stephen M. Goodnick

Herein, a program of research is detailed related to transport through the Si metal oxide semiconductor (MOS) quantum dots fabricated in a process flow compatible with modern ULSI (ultra large scale integrated circuit). Silicon quantum dots were fabricated by placing split gates within a MOSFET structure. Quantum dots of several sizes and geometries were fabricated by this process for the purpose of investigating the effects of size and shape on quantized transport through the dots.

The transport properties of the different quantum dot sizes and shapes were investigated at low temperatures, and compared to normal MOSFETs fabricated by the same technology. Equilibrium measurements with the device biased in the regime from the onset of weak inversion to just past the onset of strong inversion revealed strongly oscillatory behavior in the tunneling conductance. The conductance peaks appear to map an energy level spectrum in the dot as the inversion and depletion gates are separately swept. Symmetric devices, biased both symmetrically and asymmetrically, show two groups of "branches" which evolve with different slopes in the $V_{\mathrm{Inv}} - V_{\mathrm{Depl}}$ plane. An asymmetric device studied shows three groups of branches. In addition, a fine structure is observed in the conductance peak behavior of two devices.

This apparent energy level structure is compared to the body of literature on the so-called artificial atoms, as well as self-consistent three dimensional quantum mechanical solutions for the energy levels in the same dot structure, which qualitatively agree with the overall slope of the observed data. However, the calculations reveal only the multiple sets of slopes when asymmetrically biased. These multiple slopes are postulated to arise due to the splitting of the degenerate states of the symmetric structure as the bias makes the structure increasingly asymmetric. Finally, a simplified model is presented which shows how slight asymmetry in the dot confining potential can give rise to both a fine structure and multiple slopes in the branches, and several alternative mechanisms are presented to explain the origin of the fine structure observed.

Transport in Silicon Metal Oxide Semiconductor Quantum Dots

by

Allen David Gunther

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed March 10, 2000
Commencement June 2000

Doctor of Philosophy thesis of Allen David Gunther presented on March 10, 2000

APPROVED:

Redacted for Privacy

Major Professor, representing Physics

Redacted for Privacy

Chair of the Department of Physics

Redacted for Privacy

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Allen David Gunther, Author

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF TABLES

DEDICATION

To Joni

TRANSPORT IN SILICON METAL OXIDE SEMICONDUCTOR QUANTUM DOTS

1. INTRODUCTION

Semiconductor nanostructures have received much attention in recent years, driven strongly by the reduction of electronic devices towards diminishingly small physical dimensions. Smaller device dimensions are required to lower the cost and increase the functionality and speed of ultra large scale integrated (ULSI) semiconductor microchips. Silicon metal oxide semiconductor field effect transistors (MOSFETs) and complementary MOSFETs (CMOS) have long been the dominant technology for logic devices, and thus have driven advances in the industry. In order to reduce the size of these devices, it is necessary to reduce the gate length, which leads to high fields in the drain region of the device unless the operating voltages are correspondingly scaled. These high fields lead to problems such as hot carrier effects, including carrier multiplication due to impact ionization, carrier injection into the oxide, excessive substrate currents, and velocity saturation. These effects lead to a shift in the threshold voltage (device turn on voltage) causing logic errors, and multiplication-induced feedback effects leading to higher currents than designed, which can lead to snapback breakdown. In order to correct for these degrading effects, thinner gate oxides, shallower Ohmic contact junction depths, and complex drain-engineered structures are required. These modifications in turn, lead to problems from increased tunneling through the gate oxide to increased complexity in fabrication. It is even conceivable that channel lengths may decrease to the point where the discrete nature of dopant impurities will cause non-uniformities in the doping properties, which could cause additional device-to-device fluctuations in

threshold voltages. However, from the perspective of the electronics industry, the issue of greatest importance may turn out to be whether electronic transport continues to behave in a macroscopic manner. The study of these issues, meanwhile, has lead to a wealth of phenomena physically interesting in their own right.

Historically, transport in semiconductor devices has been described using the Boltzmann transport equation (BTE) and its simplifications such as the drift-diffusion equations. However, as devices are scaled down, one or many of the assumptions inherent to the BTE may break down as characteristic device dimensions such as the gate length approach certain length scales. A brief discussion of the assumptions involved in solving the BTE in the context of charge transport are in order [1, 2, 3]. The interested reader is referred to the books by Conwell [4] and Ziman [5], for example. The important assumptions are:

1. The external potential varies slowly enough that the effective mass approximation is still good, so a semiclassical description of transport is possible

2. The system behaves as a dilute gas, with density large enough that statistical principles may be applied but small enough to minimize inter-particle interactions.

3. Scattering processes are highly localized in phase space $(\mathbf{r}, \mathbf{k}, t)$, considered to be instantaneous, weak, and non-interacting. In other words, the time between collisions is much greater than the scattering time, τ_{sc}, and the de Broglie wavelength, λ_d, is much less than the distance between scattering centers.

4. Collisions are random, un-correlated events in which no phase memory is conserved between events, and there is no dependence on the previous history of the particle.

The first and second assumptions are somewhat coupled. The third and fourth assumptions allow the decoupling of multiple particle distribution functions. In addition, they allow a treatment the particles using classical trajectories between scattering events.

It is now appropriate to discuss the important length and time scales and their association with transport behavior in semiconductor nanostructures. In a two-dimensional electron gas structure (2DEG), such as that found at the $Si - SiO_2$ interface of an MOSFET structure, the Fermi velocity is closely related to the sheet density of the 2DEG. The Fermi wave vector is given by $k_F = (2\pi n_{2D})^{1/2}$, and the Fermi velocity is $v_F = \hbar k_F/m^*$, where n_{2D} is the 2DEG density, and m^* is the effective mass. Then the elastic mean free path is defined as the distance an electron may travel at the Fermi velocity without encountering a collision, $l_e = v_F\tau_{sc}$, which depends on the scattering time, τ_{sc}. The scattering time is also related to the diffusion constant, defined in terms of the Fermi velocity and the dimensionality, d, of the system as $D = v_F^2\tau_{sc}/d$. The distance which a particle may travel without encountering an inelastic or phase breaking scattering event is defined as the inelastic mean free path, $l_{IN} = v_F\tau_{IN}$, where the inelastic time, τ_{IN} has been introduced. Inelastic scattering occurs due to electron-phonon or electron-electron processes, which are assumed to destroy the phase coherence of an electron. Lastly, define the phase coherence length as $l_\phi = (D\tau_\phi)^{1/2}$, where τ_ϕ is the phase coherence time. l_ϕ may be different from l_{IN} due to.

While these length scales are usually quite small at room temperature, cooling the system to a sufficiently low temperature suppresses inelastic scattering mechanisms that are strongly temperature dependent, and these characteristic length scales increase and may actually surpass the characteristic length L of the device. When this crossover occurs, interesting transport phenomena occur. For example, When L becomes on the order of or less than l_{IN}, carriers may move through the

Table 1.1: A summary of the mesoscopic length scales typical measured for experimental operating conditions of devices studied in this work.

$n_s = 1.2 \times 10^{12}/cm^2$	$D = 27.51\, cm^2/s$
$k_F = 1.94 \times 10^6\, cm^{-1}$	$\tau_\phi = 0.57 \times 10^{-11}\, s$
$v_F = 1.4 \times 10^7\, cm/s$	$l_{\mathrm{IN}} = 7.98 \times 10^{-5}\, cm$
$l_e = 3.93 \times 10^{-6}\, cm$	$l_\phi = 1.25 \times 10^{-5}\, cm$

device without encountering an inelastic scattering event. Further reduction in L to the phase coherence length l_ϕ allows for quantum phase interference of particles to be observed in transport measurements. At this point, the fourth assumption of the BTE has been violated! As the L is reduced below the Fermi wavelength, $\lambda_F = 2\pi/k_{\mathrm{F}}$, and the phase coherence length l_ϕ, momentum and energy become quantized in that direction. Then it is no longer possible to treat the motion between collisions as having classical trajectories because the particle is not localized within the device. Values for these length and time scales calculated for typical device operation from the experimental data within this work are summarized in Table 1.1 using the expressions for the phenomenological parameters defined above and the magnetoconductance data to be presented in Chapter 4.

In addition to these effects, as the device size is decreased its associated capacitance decreases, and the charging energy becomes large enough that the discrete nature of charge itself can be physically observed. Such effects include steps in conductance as the excitation bias is swept and peaks in conductance as a gate bias is swept, and are referred to as Coulomb blockade phenomena.

Through advances in fabrication techniques, it has become possible to study electrical systems in which the discrete nature of charge and the discrete energy spectrum of allowed states are physically observable in transport through the system. Using advanced epitaxial growth techniques such as molecular beam epitaxy (MBE) and metallo-organic chemical vapor deposition (MOCVD), careful control of compound semiconductor heterojunctions on the scale of a few monolayers can be achieved. Through careful engineering of these heterojunctions and/or by applying an electric field normal to the interface, confinement of carrier motion along this direction results in a quantum well. While the carrier motion in the plane of the hetero-interface is free electron-like, the motion normal to the interface is quantum mechanically quantized into electron sub-bands. Since motion parallel to the interface is essentially two-dimensional, the electrons are sometimes referred to as a two-dimensional electron gas (2DEG).

The addition of some sort of lateral patterning in the plane of the interface can be used to remove the second degree of freedom, quantizing motion in one lateral dimension while leaving the particle free in the third. Such a structure can be achieved by placing depletion gates on the semiconductor surface, etching away most of the conducting layers leaving only a narrow strip of the channel, or even by etching away some of the channel and then oxidizing part of the remainder. Such a structure is termed a quasi-one dimensional (1D) system or a quantum wire.

A quantum dot is created by confining the carriers laterally in the other spatial dimension as well. Now the motion of the carriers is fully quantized, and the energy levels within the dot are fully discrete. Such a system behaves as a zero-dimensional system, and is termed a quantum dot, artificial molecule, or 0D system. Devices have been devised to operate using the quantum mechanical behavior manifested in these 1D or 0D structures based on the optical properties [6], the wavelike properties including interference [7], charge transfer devices such as pumps and turnstiles [8],

and charge storage devices such as memory cells [9]. More recently, exciting progress has been made in fabrication technology which has allowed the study of the interaction of few electron systems starting from one elecron [10]. These structures have been dubbed "artificial atoms" in the literature, behaving in some ways like atomic systems.

While quantization normal to the oxide-semiconductor inteface was studied in MOSFETs 15 years earlier than in GaAs heterostructures, in recent years the system of choice for the study of these mesoscopic effects has been GaAs/AlGaAs heterojunction structures, due in part to the higher mobility attainable because of the material properties and advanced growth techniques.

In the present investigation, the Si/SiO_2 or metal oxide semiconductor (MOS) system has been used to study the effects of lateral confinement in 0D systems. Here, the 2DEG is created by applying an electric field normal to the oxide-semiconductor interface. At this stage, the device behaves exactly as a MOSFET which is already well known. Additional confinement in the plane of the interface is achieved by depositing small confining depletion gates on the thin oxide close to the $Si\text{-}SiO_2$ interface, forming a quantum dot. While the mobility of the Si system is much lower than the GaAs system, SiO_2 has the advantage of being a much better insulator than semi-insulating GaAs. That translates into the ability of forming a much thinner insulating layer between the depletion gates and the 2DEG, which allows for closer gate spacing and a much sharper confining potential profile for the dot, as well as much higher densities in the 2DEG. In addition, interesting effects should be observable due to the many-valley properties of Si. This work is then dually motivated by the technological importance of the $Si\text{-}SiO_2$ system for the semiconductor industry and the increased complexity expected due to the more complicated band structure of Si, the higher density of carriers possible, and the higher degree of control possible due to the thin oxide.

The organization of the rest of this thesis goes as follows: In Chapter 2, the relevant background theory of transport in nanostructures will be developed, with particular emphasis on the work relating closely to this project. In parallel to this development will be a review of the literature regarding these subjects, which, while not exhaustive, should give the reader a good point of reference. In Chapter 3, the fabrication of the devices which make up the systems of study will be outlined, and a description of the experimental setup and method will be given.

The main experimental results in Chapter 4 focus on the equilibrium transport phenomena in silicon MOSFET quantum dots. A sequence of conductance peaks is observed whose positions are controllable by changing the applied bias on the gates defining the dot and the inversion layer, and are strongly dependant on the physical geometry of these gates. The analysis of these results will be presented in Chapter 5. The peak positions are studied under the assumption of pure Coulomb blockade and pure electronic confinement, and compared to theoretical calculations. It is found that the peak behavior is more consistent with pure electronic confinement dominated transport, and that the peak positions map the spectrum of electronic states n the dot as a function of external bias. Finally, conclusions and proposed future work are given in Chapter 6.

2. THEORY AND LITERATURE REVIEW

2.1 Introduction

The basic building block for most all semiconductor nanostructure devices is a two dimensional electron gas (2DEG) formed at the semiconductor surface. This 2DEG may be created by either applying an electric field normal to the semiconductor-insulator interfacial plane, forming an inversion layer at that interface, or by growing heterostructures of different bandgap semiconductors which confine carriers in a narrow region of space forming a quantum well. In this research project, the former method was used to create vertical confinement within a Si MOS (metal-oxide-semiconductor) structure. In the remainder of this chapter, the previous work in nanostructure transport relevant to the experimental results presented in this work is reviewed. This will serve to formulate the basic background required to explain the results to be presented in Chapter 5.

2.2 Silicon Band Structure

Before discussing the electronic structure of an inversion layer formed at the Si-SiO_2 interface, it is necessary to discuss the band structure of the bulk Si system. Bulk Si has an indirect band gap, with the conduction band minima near the **X** point. Near these conduction band minima, the constant energy surfaces can be described well by ellipsoids of revolution, as illustrated in Fig. 2.1. The symmetry of the **X** point in Si gives six equivalent constant energy ellipsoids, or valleys, which are degenerate with each other in the bulk.

The transport properties within a given valley are different depending on the orientation of the principal axis of the valley with respect to the surface or interface. For a Si (100) interfacial plane, there are three types of valleys, which are denoted

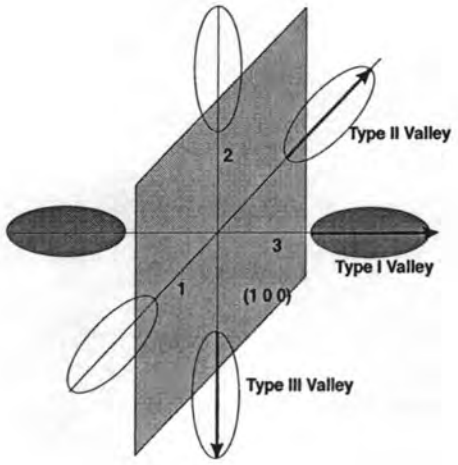

Figure 2.1: A schematic of the constant energy ellipsoids, where the shaded plane indicates the interfacial plane and reference valleys are labeled and orientation indicated by the bold arrows.

reference valleys of type I, II, and III, which are shown in Fig. 2.1. Due to the symmetry of the valleys, the type I valley have a different mass compared to the type II and III valleys with respect to transport parallel to to surface. Transport quantities are described with respect to the principal axes within the reference valleys must then be related back to the sample coordinate system.

2.3 Inversion Layers

Under conditions usually present in n-type metal oxide semiconductor (NMOS) structures, band bending occurs with zero bias applied to the gate electrode (e.g., by positively charged impurities in the oxide or work function mismatch from the gate electrode). Application of a bias voltage to a metal electrode deposited on the surface of a semiconductor creates an electric field normal to the interfacial plane of the semiconductor, resulting in bending of the energy bands i.e., the so-called 'field effect'. For a negative bias with respect to the substrate, an accumulation layer of holes forms at the interface. If a positive bias is applied, the majority carriers at

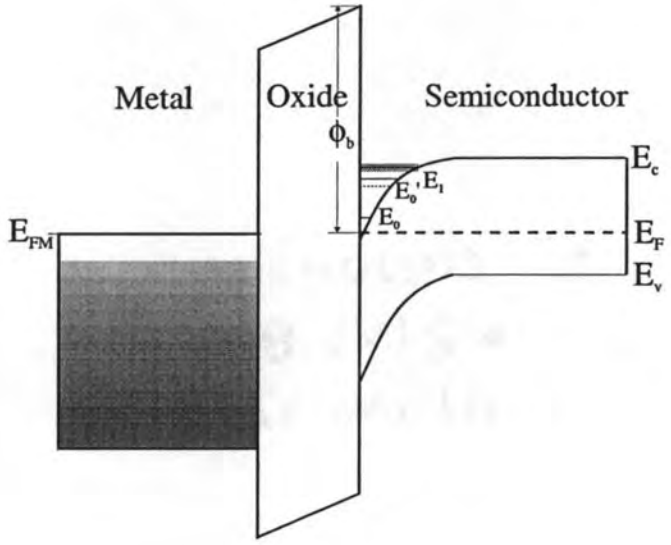

Figure 2.2: A schematic of the band diagram for a Si inversion layer. The energy levels are labeled as E_0, E_1, E_2 and ϕ_b is the barrier height for electrons entering into the oxide.

the interface are first depleted with increasing bias. If the applied field is strong enough, this bending may move the conduction band edge near and then below the Fermi level in the bulk creating a potential well, confining motion of the minority carriers normal to the semiconductor interface, and forming quantized electronic subbands labeled E_0, E_o', and E_1 *etc* as shown in Fig. 2.3. Under these conditions, the carriers behave as a 2DEG, and the thin layer is referred to as an inversion layer. An inversion layer in an acceptor-doped (or p-type) substrate is referred to as an n-channel inversion layer, as shown in Fig. 2.3, where the carriers are electrons; a donor doped (or n-type) substrate yields a p-channel inversion layer with holes forming the channel.

2.3.1 Subband Structure

In solving for the electronic subband structure for the inversion layer, the most common approach is the effective mass approximation (EMA), in which the periodic potential of the lattice is separated from the slowly varying potential due to the band bending at the interface. With this assumption for the Hamiltonian, the solution for the electronic wave function is a product of the bulk Bloch function an envelope function given by the effective mass Hamiltonian.

In the presence of an external potential, the form of the Schrödinger equation in the effective mass approximation may be reduced for symmetric surfaces to the familiar form [11]

$$\left[-\frac{\hbar^2}{2}\left(\frac{1}{m_x}\frac{\partial^2}{\partial x^2}+\frac{1}{m_y}\frac{\partial^2}{\partial y^2}+\frac{1}{m_z}\frac{\partial^2}{\partial z^2}\right)-V(\mathbf{R})\right]\psi(\mathbf{R})=E\psi(\mathbf{R}) \tag{2.1}$$

Here, z is the direction normal to the interface, $\mathbf{r}=(\mathbf{x},\mathbf{y})$ is in the parallel direction, and $\mathbf{R}=(\mathbf{r},z)$. Assuming that the Hamiltonian is separable, near the band edge (small $\mathbf{k}$) the wave function assumes a product form [1]

$$\Psi(\mathbf{R})=u_{n,\mathbf{K}}(\mathbf{R})\psi(\mathbf{r},z), \tag{2.2}$$

where $u_{n,\mathbf{K}}(\mathbf{R})$ is a Bloch function with band index n and wave vector $\mathbf{K}$ on the conduction band minimum. If the potential varies only in the direction normal to the interface, the envelope wave function may be decoupled so that the motion in the plane of interface is unrestricted and electrons behave as free electrons, $\zeta(\mathbf{r})=A\exp[i(\mathbf{k}\cdot\mathbf{r})]$ where $\mathbf{k}$ is the component of the electron wave vector parallel to the interface. The envelope wave function $\xi_i(z)$ for the motion normal to the interface satisfies the one-dimensional Schrödinger equation

$$\left(-\frac{\hbar^2}{2m_z}\frac{\partial^2}{\partial z^2}-V(z)\right)\xi_i(z)=E_i\xi_i(z), \tag{2.3}$$

where the potential $V(z)$ includes the effective potential due to the conduction band offset beteween Si and SiO_2, as well as the electrostatic potential due to Poisson's

equation

$$\nabla^2 V(z) = -\frac{4\pi\rho(z)}{\epsilon}, \tag{2.4}$$

where the charge density $\rho(z)$ includes not only that due to the depletion region, but also the many body contribution due to the electons in the inversion layer itself [12]. The composite envelope function in Eq. 2.3 is then written

$$\psi(\mathbf{r}, z) \equiv \zeta(\mathbf{r})\xi_i(z) = \mathrm{e}^{i(\mathbf{k}\cdot\mathbf{r})}\xi_i(z) \tag{2.5}$$

where i represents the i^{th} eigenstate or subband for motion parallel to the surface.

The energy dispersion measured from the conduction band minimum is given by the equation

$$E = E_{n,\mathbf{k}} + E_i + \frac{\hbar^2}{2}\left(\frac{k_x^2}{m_x} + \frac{k_y^2}{m_y}\right) \tag{2.6}$$

where $E_{n,\mathbf{k}}$ is the energy at the conduction band edge, where n is the band index, E_i is the sub-band energy due to the vertical confinement, and the remaining portion is the kinetic energy of the electrons in the plane of the interface. Within the Hartree approximation, the one-electron envelope equation 2.3 is solved self-consistently with the Poisson equation 2.4 assuming an average charge density associated with the envelope function itself for electrons in the inversion layer.

Due to the different effective mass projections along the 100 axis onto a (100) surface, the six-fold degeneracy within the bulk system is split into two sets of equivalent degenerate valleys. The two valleys with the larger effective mass (longitudinal) projecting in the direction normal to the interface (type I) have the lower energy level structure; in the terminology of Stern [12], these are the Δ_2 ladder of subbands. The four valleys that have principal axes oriented tangential to the interface (type II, III) have the transverse effective mass in the direction of the field, and thus have higher subband energies. These valleys are termed the Δ_4 ladder of subbands. The effective mass values in Si are tabulated in Table 2.1. In the

Table 2.1: Effective mass properties for symmetrically orientated Si inversion layers

Effective Mass Properties of Silicon Inversion Layers[1]						
surface orientation	m_x[2]	m_y	m_z	m_c[3]	m_{de}[4]	g_v
{1 0 0}	m_t	m_t	m_l	$\frac{m_t}{2}$	m_t	2
	m_t	m_l	m_t	$\frac{m_t m_l}{m_t+m_l}$	$\sqrt{m_t m_l}$	4
{1 1 0}	m_t	$\frac{(m_l+m_t)}{2}$	$\frac{2m_l m_t}{m_t+m_l}$	$\frac{m_l m_t+m_t^2}{m_l+3m_t}$	$\sqrt{\frac{m_l m_t+m_t^2}{2}}$	4
	m_t	m_l	m_t	$\frac{m_t m_l}{m_t+m_l}$	$\sqrt{m_t m_l}$	2
{1 1 1}	m_t	$\frac{m_t+2m_l}{3}$	$\frac{3m_l m_t}{m_t+2m_l}$	$\frac{m_t^2+2m_l m_t}{4m_t+2m_l}$	$\sqrt{\frac{m_t^2+3m_l m_t}{3}}$	6

extreme quantum limit (EQL) corresponding to low temperatures and low carrier density, all of the electrons are in the lowest subband. In this case, several approximate forms have been found to describe the energies and wave functions. Two such approximate solutions are the triangular approximation and the variational wave function approximation. For more details, see the excellent review article by Stern, *et al.* [12].

---

[1] After Stern and Howard (1967) [11]

[2] Effective masses for given direction in silicon. All effective masses are in units of the free electron mass.

[3] the conductivity effective mass is defined as $m_c = \frac{m_x m_y}{m_x+m_y}$

[4] the density of states mass per valley is defined by $m_{de} = \sqrt{m_x m_y}$

2.3.2 Density of States

The density of states for a pure two-dimensional system is constant. Hence, for a multi subband 2DEG the density of states is step-like, given by [12]

$$D(E) = \sum_i D_i(E) = \sum_i \frac{g_s g_v m_{\mathrm{de}}^{\mathrm{v}}}{2\pi\hbar^2}\theta(E - E_i) \tag{2.7}$$

where the index v is for the valley, g_v is the valley degeneracy, g_s is the spin degeneracy, and the density of states effective mass is defined as $m_{\mathrm{de}}^{v} = \sqrt{m_{v,1}m_{v,2}}$. Using Fermi-Dirac statistics for the equilibrium distribution function,

$$f(E) = \frac{1}{\exp[\frac{E_f - E}{k_B T}] + 1}. \tag{2.8}$$

The sheet density in the 2DEG is given by $n = \sum_i n_{2\mathrm{D}}^i$ where $n_{2\mathrm{D}}^i$ is the 2D density in the i^{th} subband

$$n_{2\mathrm{D}}^i = \int_0^\infty D_i(E) f(E)\, dE. \tag{2.9}$$

Since D_i is constant, Eq. 2.9 becomes a Fermi-Dirac integral of zero order, which can be performed analytically, yielding

$$n_{2\mathrm{D}}^i = \frac{g_s g_v m_{\mathrm{de}}^{\mathrm{v}} k_{\mathrm{B}} T}{2\pi\hbar^2} \ln\left[1 + \exp\left(\frac{E_{\mathrm{F}} - E_i}{k_{\mathrm{B}} T}\right)\right] \tag{2.10}$$

2.3.3 Magnetic Field

If a magnetic field is introduced into the system normal to the interfacial plane, an additional potential is added to the Hamiltonian due to the Lorentz force. Classically, these carriers are confined to circular orbits in the interfacial plane with angular frequency $\omega_c = eB/m_c$, provided there are no other forces present in the lateral direction. Quantum mechanics requires that additional energy quantization occurs. For small fields, the Hamiltonian is modified by using the Peierl's substitution for the momentum operator, such that $\mathbf{p} \rightarrow \mathbf{p} + e\mathbf{A}$ where $\mathbf{A}$ is the magnetic

vector potential. If one chooses the Landau gauge, $\mathbf{A} = (0, Bx, 0)$ then the single band effective mass Hamiltonian is once again separable, with

$$\Psi(\mathbf{r}) = \xi(z)\chi(x, y). \tag{2.11}$$

Now the energy levels are quantized with an additional energy spectrum due to the magnetic field such that the energy levels (termed Landau levels) are

$$E = E_i + E_n; \quad E_n = \hbar\omega_c \left(n + \frac{1}{2}\right), \; n = 0, 1, 2, \ldots \tag{2.12}$$

The effect of this additional quantization is to introduce a singular density of states, with each level being highly degenerate. Within each Landau level lie all the 2DEG states in the range $\pm\hbar\omega_c/2$, and the Fermi energy is pinned to the highest occupied Landau level. This gives the number of electrons within a given Landau level as

$$D = g_{\mathrm{s}} g_{\mathrm{v}} \frac{eB}{2\pi\hbar}. \tag{2.13}$$

This singular density of states is indeed observable in the oscillatory magnetoconductance observed as the magnetic field is swept, called Shubnikov de-Haas (SdH) oscillations. Peaks in the conductance occur when the Landau level passes through the Fermi circle of the 2DEG, which is pinned to the highest occupied Landau level. These oscillations are periodic in $1/B$ with the spacing being

$$\Delta\frac{1}{B} = \frac{g_{\mathrm{s}} g_{\mathrm{v}}}{2\pi\hbar n_{\mathrm{2D}}}. \tag{2.14}$$

From this equation, it is easy to determine the 2DEG density, n_{2D}. The magnetoconductance oscillations can be described in weak fields in the EQL by the form derived by Ando using linear response theory. This form is discussed in more detail in a review article [13]. For weak magnetic fields ($\omega_c\tau \leq 1$), the longitudinal magnetoconductance is written

$$\sigma_{xx} = \sigma_0 \frac{1}{1+\omega_c^2\tau^2}\left[1 - \frac{2\omega_c^2\tau^2}{1+\omega_c^2\tau^2}\frac{2\pi^2 k_B T_e}{\hbar\omega_c} \times \mathrm{csch}\left(\frac{2\pi k_B T_e}{\hbar\omega_c}\right)\cos\left(\frac{2\pi^2 E_f}{\hbar\omega_c}\right)\exp\left(-\frac{\pi}{\omega_c\tau}\right) + \ldots\right] \tag{2.15}$$

where the zero-field conductance is defined as $\sigma_0 = n_{2D}e^2\tau/m^*$ and τ is the scattering time for electrons at the Fermi energy. The first order term contains the non-oscillatory component which depends on the Fermi energy and the inverse magnetic field, corresponding to the SdH oscillations described above.

Another interesting property of the 2DEG system is that of the quantum Hall effect (QHE) first reported in 1980 by von Klitzing, *et al.* [14]; at precisely the same field as the longitudinal magnetoconductance vanishes, there occur plateaus in the Hall resistance (transverse magnetoresistance) which have the magnitude

$$R_H = \frac{V_H}{I} = \frac{h}{e^2}\frac{1}{n}, \qquad n = 1, 2, 3, \ldots \tag{2.16}$$

where n is precisely the number of spin degenerate Landau levels occupied. Figure 2.3 shows the SdH oscillations and the Hall plateaus for one of the MOS samples studied in the present work. As can be seen there, plateaus occur at the same magnetic field as peaks in the transverse magnetoresistance; in the measurement, no attempt was made to source a constant current, so the Hall resistance could not be calculated.

The 2DEG conductivity mobility can be determined from the zero field resistance by the relation $\rho_{2D}|_{B=0} = 1/(\mu n_{2D}e)$, where $\rho_{2D}|_{B=0}$ is the zero field longitudinal resistivity. The 2DEG mobility and density can be determined for samples without perfect Hall bar geometry by the van der Pauw method [15].

2.4 Quantum Wires

If carriers within the 2DEG discussed in Section 2.3 are confined in one additional dimension, a so-called quantum wire is produced. Now carriers are free to move in only one spatial direction, and they behave in a quasi-one-dimensional manner.

Quantum wires are traditionally fabricated in one of two manners: either a 2DEG channel is patterned with an etch mask and the channel is etched to leave

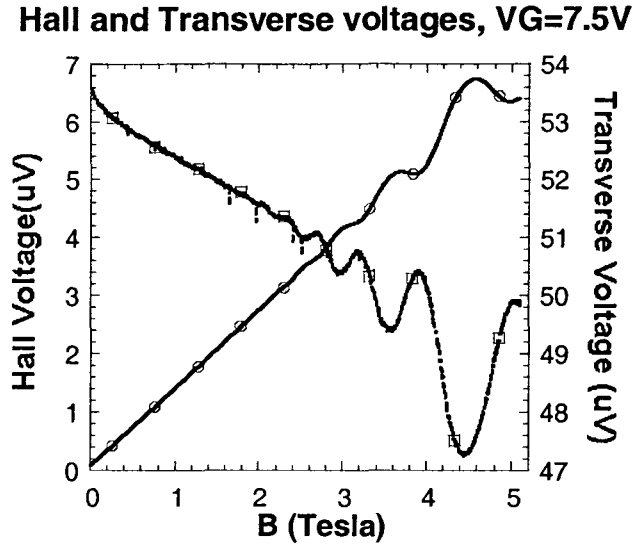

Figure 2.3: A plot of the Hall voltage (open circles) and longitudinal voltage (open squares) for a voltage biased MOSFET. The data is from device M1 measurement H6.

only a narrow conducting pathway, or metal gates are patterned on the surface of a structure under which a 2DEG exists such that when the gates are negatively biased, the channel is electrostatically constricted to allow only a narrow conducting path. Both methods result in essentially the same soft sort of confining potential for different reasons. In the former case, surface roughness induced from the wet or dry etching technique lead to surface states and roughness scattering which produces at best a parabolic potential profile, limiting the strength of the confinement energy. In addition, if the physical channel is made too narrow, the depletion at the air-wire interface can result in complete pinch off of the conducting path. In the latter method, also called the split-gate technique, the smearing of the electrostatic confinement results in a parabolic confining potential and the same end result of a relatively weak energy level separation. In GaAs/AlGaAs technology, quantum

wires are achieved confining an existing quantum well structure by one of the methods just mentioned. In Si, this is a little more complicated in that an additional oxide must be added to use the split gate technique.

There have been other approaches to attaining a quantum wire in Si inversion layers. Fowler, *et al.* [16] used a MOSFET structure in which lateral confinement was achieved by implanting the regions adjacent to the channel and reverse biasing these regions to cause lateral depletion. F. Scott-Thomas *et al.* [17] used a dual gate MOSFET structure in which a narrowly separated ($\approx$ 70 nm) split-gate structure is placed on top of a thermal oxide, and then a CVD oxide is grown on top. Finally, an upper gate is placed above the entire structure. Only the space between the narrow split gates forms an inversion layer that is the quantum wire.

Ishikuro, *et al.* [18] used a (100) oriented silicon-on-insulator (SOI) wafer substrate, selective oxidation and anisotropic etching combined with electron beam lithography to fabricate quantum wires. First, a mesa is created under an etch mask formed with Si_3N_4. The edges are faceted in the (111) direction by use of an anisotropic dry etchant. The (111) facet sidewalls of the Si mesa are selectively oxidized through the mesa mask. Electron beam lithography (EBL) is used to expose the length adjacent to the (111) facet and this is also etched anisotropically and the whole triangular shaped wire is oxidized to form a gate oxide. Poly-Si gates are formed over the top, leaving wires which are 100 nm long in the (110) direction and roughly 10nm wide to form triangular shaped quantum wires with a poly-Si gate.

Je *et al.* [19] used SOI, EBL and anisotropic etching to form a wire. In their structure, a wire is defined by patterning the Si layer with EBL and etching away all Si to the buried oxide. All of the oxide around the wire is etched away (including underneath) using buffered oxide etch (BOE). The Si wire is then further constricted by oxidizing the remaining Si, and a poly-Si gate is deposited.

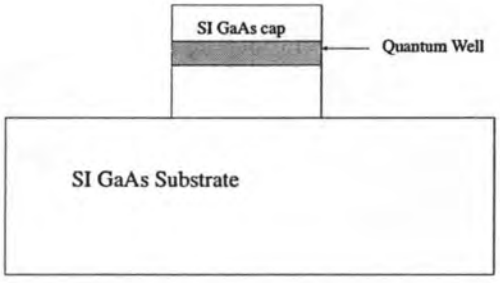

Figure 2.4: An example of an etched quantum wire in a quantum well

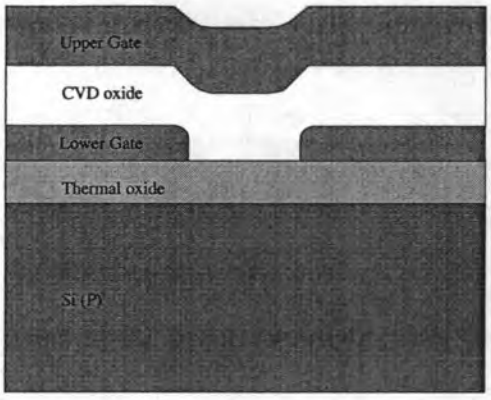

Figure 2.5: An example of a split-gate quantum wire after F. Scott-Thomas *et al.*,1989

Ohdomari and Shinada [20] used SOI wafers to form quantum wires. In their structure, the wafers were oxidized and then the oxide surface was modified by scanning a focused ion beam (FIB) across the surface. The regions of oxide thus modified exhibited enhanced, anisotropic etching and were removed in hydrazine water, leaving a wire. These approaches are illustrated by the cross sections in Fig. 2.4.

2.4.1 Quantum Wire Electronic States

For a quantum wire system in which the channel in the direction of free propagation is long enough, we assume that translational symmetry holds. Then, within the effective mass approximation (EMA) we assume that the wave function is separable into a product form of free carrier and confined states for a single band

$$\Psi(\mathbf{r}, z) = \phi_{n,m}(\mathbf{r}) e^{ik_x x}/L, \tag{2.17}$$

where L is the normalization, and the wave function in two dimensions satisfies a one-dimensional Schrödinger equation.

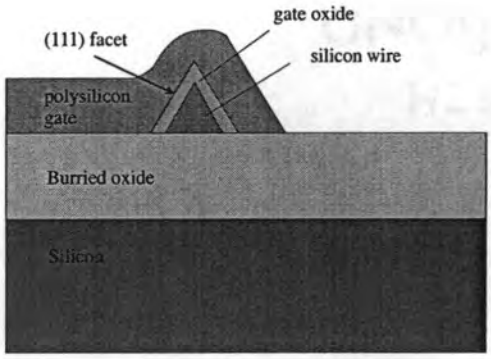

Figure 2.6: A quantum wire fabricated by anisotropic etch and selective oxidation, after Ishikuro *et al.*, 1996

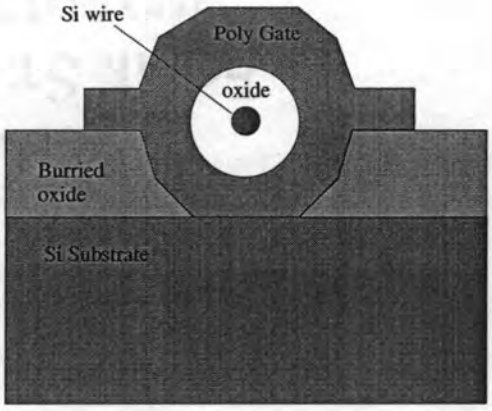

Figure 2.7: A quantum wire fabricated by the gate-around technique, after Je *et al.*, 1998

The total energy associated with the quantum wire overlap function is the sum

$$E_{k,i,j} = \frac{\hbar^2 k_x^2}{2m_x} + E_i + E_j \tag{2.18}$$

where E_i is defined in Sec. 2.3 and E_j is the eigenvalue for the j^{th} one dimensional subband; it will be convenient to introduce the shorthand notation $E_{i,j} = E_i + E_j$ below.

2.4.2 Density of States in One Dimension

The density of states (per unit length along the wire axis) for a quasi-one-dimensional wire is given by the equation

$$D(E) = \alpha_s g_v \frac{m_x}{2\pi\hbar^2} \left(\frac{\hbar^2}{2m_x}\right)^{\frac{1}{2}} \sum_{i,j} (E - E_{i,j})^{-\frac{1}{2}} \Theta(E - E_{i,j}). \tag{2.19}$$

From Eq. 2.19 it can be seen that there are divergences in the DOS as the energy goes through each discrete energy level associated with either the vertical confinement or the confinement normal to the wire axis in the plane of the 2DEG, which go as

$1/\sqrt{E - E_{i,j}}$. Experimentally, these divergences are seen in optical spectroscopy, but are not seen in transport. This is probably due to broadening of the levels by surface states and roughness as discussed previously. Self-consistent calculations of the quantum wires in Si [21] and in GaAS [22, 23], have shown that the typical subband spacing is on the order of 5-10 meV.

2.5 Quantum Dots

By constricting the semiconductor in the remaining free dimension, carriers may be confined in all three spatial dimensions, forming a quantum dot. The conventional methods of introducing this confinement are the split gate technique, hard wall etching, and in-plane gating. The first two methods are exactly analogous to the same techniques for wire formation, and need no further discussion. The third method is similar to the hard wall etching method, except that only a narrow trench is etched through the 2DEG region. Another approach is growth of self assembled quantum dots. In self assembled, or self organized growth, a thin (typically about one monolayer) epitaxial layer (epilayer) of crystal is grown by MBE on top of a substrate with slightly different lattice constant. As the layer increases, strain develops between the epilayer and the substrate, and at some critical thickness the strain relaxes by forming clusters of atoms which are of non-stoichiometric composition. Such growth is called Stranski- Krastanov (SK) growth, and the dots formed by this procedure are called self-organized or self-assembled quantum dots. The material systems of choice are typically $In_xGa_{1-x}As/In_yGa_{1-y}As$, where the value of y is intentionally larger than x to dilate the epilayer lattice constant, or $GaAs/GaAs_{1+x}$ where the fraction x is achieved by growing the epilayer at a lower temperature than is needed to grow stoichiometric GaAs.

A representative group of studies for the strained InAs, InGaAs, and Sb containing compound semiconductors on GaAs substrates include a study of the tem-

perature and monolayer coverage effects [24], a detailed study of the critical layer thickness [25], growth on vicinal GaAs (001)-(2×4) surfaces [26], ordering phenomena [27], and growth of InSb, GaSb, and AlSb dots on GaAs [28]. It was also shown that the strain field induced by the self organized dots could in fact create confinement in a 2DEG located underneath the dots both experimentally by Lipsanen *et al.* [29] and theoretically by Tulkki and Heinänäski [30].

Arsenic precipitants in low-temperature-grown GaAs of p-doped and n-doped substrates was studied as a function of anneal time [31], and their structure was found to be of hexagonal single crystal structure [32]. Another approach to getting the non-stoichiometric deposition leading to precipitants is the use of GaAs ion implantation and thermal annealing [33].

The problem with SK growth is that it is difficult to make contact to the semiconductor island regions, the size of the islands is some statistical distribution so it is not easy to be sure of the size of the dot grown, and the location of the islands on bare wafer surfaces is random. There are at least two known methods to get around this difficulty. One is to grow the strained layer on slightly off-axis wafers; the second is to grow strained layers on patterned (typically by EBL) substrates. Both approaches use the fact that these clusters tend to precipitate along sidewalls, present due to the crystal monolayer plane steps or by the patterning technique. By patterning trenches on a GaAs substrate, followed by subsequent deposition by MBE, chemical beam epitaxy (CBE), or metallo-organic vapor phase epitaxy (MOVPE), it was found that InAs [34, 35] and GaInP [36] would precipitate preferentially only on the sidewalls, the ridge tops, or mesa edges at the bottom of the trenches. Interestingly, it was found that using low-temperature-grown GaAs on unpatterned substrates, self-organized dots could be formed preferentially by patterning deposited InGaAs followed by SiN after the low-temperature growth [37]. The chains were found to grow with as little as a 23 nm edge-to-edge spacing. It

would be interesting to study the transport properties of dots that formed a one-dimensional chain, which could behave as a near ideal one dimensional crystal. However, these spacings are still not close enough. It may be possible to overcome this difficulty, however, by stacking in the vertical direction, as was demonstrated possible by Miller, *et al.* [38], who stacked InAs islands in GaAs layers on a patterned GaAs substrate. They were able to have spacings as close as 5.4 nm for 2 nm high by 20 nm wide islands.

Optical studies on these self-assembled quantum dots have shown that they are optically active. Photoluminescence (PL) on unpatterned wafers [39, 40, 41] have shown phenomena characteristic of discrete electronic states. The results of optical studies on the strain-induced dots were even more conclusive [29]. Cathodoluminescence studies on single dots have shown sharp luminescent features [42]. Optical results on chains and stacks of self assembled quantum dots have also shown that they are active [34, 38].

In the $Si - SiO_2$ system, one approach for fabricating quantum dots has evolved around variations of a dual-gate split-gate structure along the same lines as Fig. 2.5. Alsmeier *et al.* [43] started with a thin NiCr mesh embedded in a MOS structure which they probed with far infra red spectroscopy to show that they could tune the electron number by the field effect. Matsuoka *et al.* [44, 45] used an arrangement where an inversion gate was placed on a lower thermal oxide, and then overlaid this with a depletion gate arrangement. Their depletion gate architecture was a series of periodically spaced lines ranging from one to several, which created a series of quantum dots, on which they carried out transport measurements. Matsuoka *et al.* [46, 47, 48] followed this with a Π-shaped gate to form one quantum dot with the same dual gate structure.

Several groups have taken the approach of etching the top Si layer of SOI wafers between source and drain contacts forming a narrow constriction and overlaying it

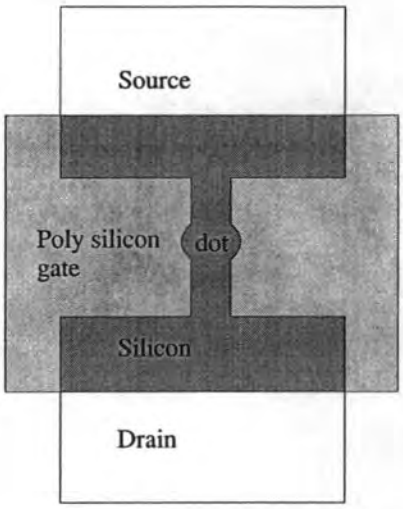

Figure 2.8: A schematic of the Leobandung device geometry, after Leobandung *et al.*, 1995

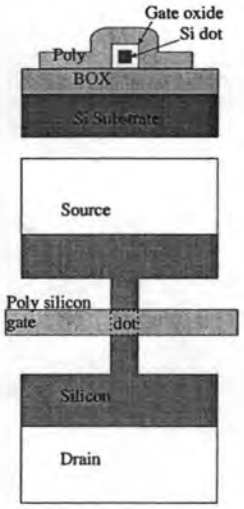

Figure 2.9: Example of the etch and oxidize confinement showing the cross section of a dot and device geometry. After Peters *et al.*, 1998

with a poly gate to form a dot. Chou *et al.* [49, 50, 51] used EBL and reactive ion etching (RIE) to form a wire channel with a node in the middle which served as the dot. This dot was overlaid with a poly-Si gate, as is illustrated in Fig. 2.8. Uchida *et al.* [52] and Peters *et al.* [53] etched the top Si layer of an SOI channel to a wire, oxidized the wire to reduce the channel to as small as 100 nm, and overlaid with a poly-Si gate going across the channel to yield dots as small as 300 nm by 100 nm. Fujiwara *et al.* [54] used pattern-dependent oxidation (PADOX), which takes advantage of enhanced oxidation due to geometry, to oxidize a narrow wire formed by etching on a SOI wafer, forming narrower constrictions on either side of the dot. Fig. 2.9 illustrates the basic idea of this technique.

Si nano-crystals are also quantum dots, which may be made by several approaches. Kanemitsu *et al.* [55] used laser breakdown of silane gas in a vacuum (P $< 10^{-6}$ Torr) to deposit the nano-crystals on Ge and fused quartz substrates. They were then oxidized at room temperature in a clean air box. Photoluminescence stud-

ies show that there is a peak in the spectrum which is independent of the size of the nano-crystals. Yamada, *et al.* [56] created nano-crystals on Si and fused quartz substrates by laser ablation of crystalline Si in a vacuum with constant He background pressure. The size of the nano-crystals was controlled using the helium background pressure to have a mean diameter of 13-17 nm. Prior to PL measurements, the samples were oxidized in dry oxygen at 800 C. PL data at room temperature yield strong peaks; one independent of dot size, and one dependent of dot size. Fukuda *et al.* [57] fabricated nano-crystal Si on SiO_2 substrates by depositing pure silane in a low pressure CVD reactor at differing temperatures from 550 to 600 C, yielding dot diameters from 10 to 20 nm. The dots were then oxidized 1 nm at room temperature or 800 C. An atomic force microscope tip (AFM) was used as a probe to measure the tunneling current between the tip and the substrate; a room temperature peak was observed, which could be modeled by resonant tunneling.

2.5.1 Electronic Structure of Quantum Dots

The electronic structure of quantum dots has been studied quite extensively in recent years. Theoretical approaches range from non-interacting to self-consistent including exchange and correlation effects, and have been calculated using exact diagonalization of the Hamiltonian, density functional and Hartree-Fock formulations. In this work, we will begin by presenting a simple non-interacting picture and build upon it. In most cases, the details will be left to the references listed.

2.5.1.1 Single Particle Electronic Structure

First, we shall consider the case of a single, non-interacting electron in a quantum dot. Begin again with the effective mass Hamiltonian, Eq. 2.1, assuming that it is separable and the wave functions may be written in a product form (suppressing

the Bloch wave functions for simplicity)

$$\psi(x,y,z) = \chi(x)\phi(y)\xi(z) \tag{2.20}$$

and that the z-component again satisfies Eq. 2.3. If we once again assume non-interacting electrons, we can simply choose an approximate potential model (for example, hard wall or harmonic oscillator) and write down the wave function as a product of the solutions. The hard wall and harmonic confinement cases have been shown to be exactly soluble for the case of magnetic field normal to the interfacial plane. In the general case including a magnetic field, one starts with the Hamiltonian

$$H = \frac{1}{2m}\left(\mathbf{p} + e\mathbf{A}\right)^2 + V(R), \tag{2.21}$$

where $\mathbf{p}$, $\mathbf{A}$, and $\mathbf{R}$ were defined earlier. Geerinckx *et al.* [58] showed that for isotropic effective mass in the plane and a cylindrically symmetric confinement potential

$$V(r) = \begin{cases} 0 & \text{for} \quad r < R_0, \\ \infty & \text{for} \quad r > R_0, \end{cases} \tag{2.22}$$

that the solution for the motion in the plane is given by

$$\Psi_{n,l}(r,\phi) = \frac{1}{\sqrt{C_{nl}}}\frac{\mathrm{e}^{il\phi}}{\sqrt{2\pi}}\left(\frac{m\omega_c}{\hbar}\right)^{\frac{1}{2}}\left(\frac{r^2}{2l_B^2}\right)\exp\left(\frac{r^2}{4l_B^2}\right)\left[{}_1F_1\left(-\alpha; 1+|l|; \frac{r^2}{2l_B^2}\right)\right] \tag{2.23}$$

where ${}_1F_1()$ is a confluent hypergeometric series given, for example, by Morse and Feschbach [59], the magnetic length is defined as $l_\mathrm{B} = \sqrt{\hbar/eB}$, single valuedness requires that $l = \pm 1, \pm 2, \ldots$, and the normalization is given by

$$C_{nl} = \int_0^{x_0} dx\, \mathrm{e}^{-x} x^{|l|}\left[{}_1F_1(-\alpha; 1+|l|; x)\right]^2. \tag{2.24}$$

The energy levels are given by

$$E_{nl} = \hbar\omega_c\left(\alpha_{nl} + \frac{l+|l|}{2} + \frac{1}{2}\right), \tag{2.25}$$

where α_{nl} are the roots to the equation

$$ {}_1F_1\left(-\alpha; 1+|l|; \frac{R_0^2}{2l_B^2}\right) = 0, \tag{2.26} $$

or, equivallently,

$$ \alpha_{n,l} = \frac{E}{\hbar\omega_c} - \frac{1}{2}(l+|l|+1). \tag{2.27} $$

This expression, however, is singular at zero magnetic field, which is a non-physical result The only orbital dependence comes from the azimuthal portion of the wave function. The energy levels for negative orbital quantum number are always lower than the positive values, and so non-degenerate. Classically, this means that there is a difference between clockwise and counterclockwise orbits. At zero temperature, the Fermi level is pinned to the $(N/2)^{\text{th}}$ level, and is discontinuous but approaches $\hbar\omega_c/2$ at high field.

As mentioned in the context of quantum wire systems, a more realistic choice for the potential is that of a harmonic oscillator. It has also been shown that the solution to this potential can be expressed by an analytic form in the presence of a normal magnetic field, and the solutions are the so-called Darwin-Fock solutions [60, 61], which have been investigated more thoroughly by Dingle [62] in the context of large systems of electrons, and later by Bockelmann [63]. Start with the Hamiltonian, Eq. 2.21 where $V(x,y) = m/2\omega_0^2(x^2+y^2)$ and make the following substitutions:

$$ \Omega = \sqrt{\omega_0^2 + \frac{1}{4}\omega_c^2}, \quad W = \frac{2\pi}{\hbar}\frac{E}{\Omega}, \quad \alpha = \sqrt{\frac{\hbar}{m\Omega}}, \tag{2.28} $$

and make a transformation $\vec{r} = \rho\alpha(\cos(\phi), \sin(\phi))$. We end up with the wave function

$$ \Psi_{n,l}(\mathbf{r}) = \left(\frac{n!}{\pi\alpha^2(n+|l|)!}\right)^{\frac{1}{2}} \rho^{|l|} L_n^{|l|}(\rho^2) \mathrm{e}^{-\frac{\rho^2}{2}} \mathrm{e}^{-il\phi} \quad (n = 0, 1, \ldots), \tag{2.29} $$

where the associated Leguerre polynomials $L_n^{|l|}(\rho^2)$ are defined in Gradshteyn and Ryzhik [64], and the energy levels are given as

$$E_{n,l} = (2n + |l| + 1)\hbar\Omega - \frac{1}{2}\hbar\omega_c l. \tag{2.30}$$

The level structure associated with the harmonic potential evolves linearly with confinement strength, and, once again, the degeneracy in the orbital quantum number l is lifted by its sign. As expected, the spacing of the ground states are uniform with a value $2n\hbar\Omega$.

2.5.1.2 Multi-Particle Electronic Structure

We now move to the case where there are many electrons in the quantum dot. Here it is necessary to include the effect of the electron-electron interaction, and the Fermion nature of electrons (the wave functions must be anti-symmetric with respect to the exchange of position). We begin with the general form of the multi-electron Schrödinger equation, given for example by Ashcroft and Mermin [1]:

$$\mathcal{H}\Psi(\mathbf{r_1}, s_1, \mathbf{r_2}, s_2, \ldots \mathbf{r_i}, s_i, \ldots \mathbf{r_N}, s_N,) = E\Psi(\mathbf{r_1}, s_1, \mathbf{r_2}, s_2, \ldots \mathbf{r_i}, s_i, \ldots \mathbf{r_N}, s_N,) \tag{2.31}$$

where s_i are the spin indices. Note that the wave equation is written to explicitly show the spin and position of each particle in the dot. The Hamiltonian is written as the sum of single particle Hamiltonians and the inter-particle interaction Hamiltonian:

$$\mathcal{H} = \sum_{i=1}^{N} H_0(\mathbf{r_i}) + \mathcal{H}_1(\mathbf{r_i} - \mathbf{r_j}). \tag{2.32}$$

The single particle Hamiltonians in the presence of a magnetic field are given by

$$H_0^i(\mathbf{r_i}) = \frac{1}{2m^*}\left(\mathbf{P_i} + \frac{e}{c}\mathbf{A}(\mathbf{r_i})\right)^2 + V_0^i(\mathbf{r_i}) + \frac{g_m^*\mu_B}{\hbar}\mathbf{S_i}\cdot[\mathbf{B}(\mathbf{r_i}) + \nabla\mathbf{V_0}(\mathbf{r_i})\mathbf{L_i}], \tag{2.33}$$

where here the spin-orbit coupling is included with g_m^* being the effective Lande g-factor, and the interaction part is given by the following,

$$\mathcal{H}_1(\mathbf{r_i} - \mathbf{r_j}) = \frac{1}{2\kappa_s \epsilon_0} \sum_{i \neq j} \frac{e^2}{|\mathbf{r_i} - \mathbf{r_j}|}. \tag{2.34}$$

There have been many attempts in the recent literature to solve the multi-particle quantum dot problem. The most direct method is to solve Eq. 2.31 exactly by diagonalizing the Schrödinger equation for a suitably chosen basis set of wave functions. Of course, it is necessary to eventually truncate the number of terms in the series expansion so at this point it becomes no longer exact. The total wave function is built up from a series expansion of products of the single particle basis states such that

$$\Psi(\mathbf{r_1}, s_1; \mathbf{r_2}, s_2; \ldots \mathbf{r_i}; s_i, \ldots \mathbf{r_N}; s_N,) = \prod_{i=1}^{N} \psi_i(\mathbf{r_i}, s_i) \tag{2.35}$$

$$\psi_i(\mathbf{r_i}, s_i) = \langle \mathbf{r_i}; s_i | \sum_{n,\vec{\sigma}} a_{n,\vec{\sigma}} | \varphi(\mathbf{r_i}; \chi_{\vec{\sigma}}) \rangle \tag{2.36}$$

$$a_{n,\vec{\sigma}} = \langle \varphi(\mathbf{r_i}; \chi_{\vec{\sigma}}) | \Psi(\mathbf{r_1}, s_1, \mathbf{r_2}, s_2, \ldots \mathbf{r_i}, s_i, \ldots \mathbf{r_N}, s_N,) \rangle \tag{2.37}$$

where the spin portion of the wave function is $|\chi_{\vec{\sigma}}\rangle$ and the spin index $\vec{\sigma}$ can be either $|\uparrow\rangle$ or $|\downarrow\rangle$. Exact solutions to Eq. 2.31 have been reported by several groups [65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75]. The natural choice of basis states is the set of solutions to the single particle Hamiltonian, where the bare potential is chosen to be harmonic. With the inclusion of the Coulomb interaction, the radial quantum number mixes, and only the total angular momentum $\mathcal{L} = \sum_i l_i$ is conserved, while the total spin is constrained to minimize the Zeeman term [75]. However, Bryant [65], Chakraborty [66, 69], Merkt *et al.* [68], and Pfannkuche [70] used the center-of-mass and relative coordinates, showing that they gave the same results as the single-particle basis states. In the case where the confinement is assumed to be harmonic and the Zeeman term is neglected, the Hamiltonian is

separable as

$$\mathcal{H} = H_{CM} + \mathcal{H}_{\rm rel}, \tag{2.38}$$

where the various terms are defined as usual by

$$H_{CM} = \frac{1}{2M}\left(\mathbf{P} + Q\mathbf{A}(\mathbf{R})\right)^2 + \frac{1}{2}M\omega_0^2\mathbf{R}^2, \tag{2.39}$$

$$\mathcal{H}_{\rm rel} = \sum_{i(i\neq j)}^{N-1}\left[\frac{1}{2\mu}\left(\mathbf{p_i} + \frac{e}{c}\mathbf{A}(\mathbf{r_i})\right)^2 + \frac{1}{2}\mu\omega_0^2\mathbf{r_i}^2 + \frac{e^2}{\kappa_s\epsilon_0}\frac{1}{|\mathbf{r_i}|}\right], \tag{2.40}$$

$$\mathbf{R} = \frac{1}{N}\sum_{i=1}^{N}\mathbf{r_i}, \quad Q = Ne, \quad \mathbf{P} = \sum_{i=1}^{N}\mathbf{p_i}, \quad M = \sum_{i=1}^{N}m^*, \tag{2.41}$$

$$\mathbf{p_i} = \mathbf{p_i} - \mathbf{p_j}, \quad \mathbf{r_i} = \mathbf{r_i} - \mathbf{r_j}, \quad \mu = \frac{m^*}{2}. \tag{2.42}$$

The wave functions for the center of mass and relative coordinate systems are identical (in the absence of Coulomb interaction) to the case of the single-particle Hamiltonian if we make the following substitutions:

$$\rho \rightarrow \rho_{\rm CM} = \frac{\hbar\mathbf{R}}{M\Omega} \tag{2.43}$$

$$\rho \rightarrow \rho_{\rm rel} = \frac{\hbar\mathbf{r_i}}{\mu\Omega}. \tag{2.44}$$

When the Coulomb interaction is taken into account, the wave functions become mixed in radial quantum number n for both the center of mass (CM) and single electron basis states. The CM system has the advantage that the Coulomb interaction affects only the motion of the relative wave function [75]. It was shown by Bryant [65] that this was very important in the case of zero magnetic field. Ezaki *et al.* [76] later diagonalized the N-particle Hamiltonian for up to 12 electrons in the dot, and allowed deformation of the cylindrical symmetry (still parabolic confinement) by changing the form of the external potential to be of the form

$$V(x,y) = \frac{1}{2}m^*(\omega_x^2 x^2 + \omega_y^2 y^2)\left\{1 + \alpha\frac{2}{7}\cos(3\phi)\right\}, \tag{2.45}$$

while Pfannkuche and Gerhardts allowed deviations from parabolic confinement [70].

For greater than ~ 10 electrons in the dot, the exact diagonalization of Eq. 2.31 becomes computationally taxing, and other approximation methods become necessary. The simplest approximation replaces the electron-electron interaction with a Coulomb interaction between one electron and an average charge distribution from the remaining electrons in the dot, the so-called Hartree approximation,

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V_0(\mathbf{r}) + \left(e^2\sum_j \int d\mathbf{r}'\,|\psi_j(\mathbf{r}')|^2\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right) - \mathcal{E}_i\right]\psi_i(\mathbf{r}) = 0, \tag{2.46}$$

and the set of i equations are the Hartree equations. Two shortcomings of the Hartree equations are that it does not properly treat the Fermion property of antisymmetry with exchange of particle position in the product wave function (Pauli principle),

$$\begin{aligned}&\Psi(\mathbf{r_1}, s_1; \ldots; \mathbf{r_i}, s_i; \mathbf{r_j}, s_j; \ldots; \mathbf{r_N}, s_N)\\ &\quad = -\Psi(\mathbf{r_1}, s_1; \ldots; \mathbf{r_j}, s_j; \mathbf{r_i}, s_i; \ldots; \mathbf{r_N}, s_N)\end{aligned} \tag{2.47}$$

and that it includes a self interaction in the Coulomb interaction term. Antisymmetrization of the wave equation can be ensured by taking the Slater determinant of the product wave function.

The over counting of the Coulomb interaction by inclusion of the self-interaction term is corrected for by the inclusion of an additional factor leading to the Hartree-Fock equations:

$$\begin{aligned}&\left[-\frac{\hbar^2}{2m}\nabla^2 + V_0(\mathbf{r}) + \left(e^2\sum_j \int d\mathbf{r}'\,|\psi_j(\mathbf{r}')|^2\frac{1}{|\mathbf{r}-\mathbf{r}'|}\right) - \mathcal{E}_i\right]\psi_i(\mathbf{r})\\ &\qquad - \sum_j \int d\mathbf{r}'\,\frac{e^2}{|\mathbf{r}-\mathbf{r}'|}\psi_j^*(\mathbf{r}')\psi_i(\mathbf{r}')\psi_j(\mathbf{r})\delta_{s_i s_j} = 0\end{aligned} \tag{2.48}$$

The last term in the Hartree-Fock equations, called the exchange term, makes the set of equations integro-differential equations, which are much more computationally complex.

The Hartree approach to solving the multi-particle quantum dot has been reported by Kumar *et al.* [77], Stopa *et al.* [78], and Gudmundsson and Gerhardts [79]. In these references, no attempt is taken to account for the electron spin, but the full three dimensional problem is studied by the first two authors.

The unrestricted Hartree-Fock equations including spin were solved for a two-dimensional parabolic potential by Palacios *et al.* [80], and for a three dimensional anisotropic harmonic potential of the form $1/2(\omega_x^2(x^2+y^2)+\omega_z^2 z^2)$ by Fujito *et al.* [81]. Ezaki [82] later solved the Hartree-Fock equations for a self-consistent three-dimensional potential.

Another approach that has been used to determine the energy levels within a quantum dot is the density functional method [83, 84, 85, 86]. Stopa [83] solved for the external confinement potential self consistently by solving the z-component of the Schrödinger-Poisson equation at every point in the $x-y$ plane and then making an expansion in terms Darwin-Fock states or Bessel functions to reduce the problem in the plane of the dot. Macucci *et al.* [84, 85, 86] solved the Schrödinger equation including Coulomb interaction, exchange, and correlation for a quasi-parabolic potential.

2.5.1.3 Experimental Observation of Energy Level Structure in a Quantum Dot

Experimentally, the level structure of quantum dots has been observed optically and by transport measurements. I will delay discussion on the latter until the discussion of transport in quantum dots. Intense interest has been directed toward the potential application to quantum dot lasers, so that there have been numerous investigations of the optical properties. Studies in PL and cathodoluminescence have been discussed in connection with self assembled quantum dots. Another optical probe of the level structure within a quantum dot is far infrared absorption (FIR) which has been investigated on arrays of quantum dots [87, 88, 89, 90, 91]. The

arrays are fabricated using holographic double exposure lithography and reactive ion etching of high mobility samples such as GaAs/AlGaAs and InSb, and include a gate to couple to the chemical potential of the dots. In this manner, the number of electrons can be tuned and the size of the dots is very uniform (on the order of 100 nm). In the experiments, FIR radiation is incident on the sample which is placed in a superconducting magnet equipped cryostat. The detector, which measures optical conductivity, is typically coupled out of the cryostat by waveguide. The data reveals peaks in the optical conductivity which can be modeled by the classical conductivity model

$$\sigma_{\pm}(\omega) = \frac{en_0\mu}{1 + \left[(\omega_0^2/\omega) - \omega \pm \omega_c\right]^2 \tau^2}, \tag{2.49}$$

where τ is a phenomenological relaxation time and the mobility $\mu = e\tau/m^*$. As the magnetic field is increased, branches $\omega_{\pm}$ are observed at the allowed dipole transitions with dispersion

$$\omega_{\pm} = \sqrt{\omega_0^2 + \left(\frac{\omega_c}{2}\right)^2} \pm \frac{\omega_c}{2}. \tag{2.50}$$

Typically, only the lowest one or two branches are observable. Demel *et al.* [88] observed an anti-crossing of the ω_{1+} and ω_{2-} levels. Meurer *et al.* [91] studied the integrated FIR absorption strength as a function of gate voltage to correlate the FIR spectrum to electron number. The expression for relative transmission is

$$\frac{\Delta T}{T} = -2\Re\left\{\frac{\sigma(\omega)}{[(1+\sqrt{\epsilon}+r_v/r_g)\epsilon_0 c]}\right\}, \tag{2.51}$$

where r_v is the optical vacuum impedance and r_g is the impedance of the gate contacts. They found that the integrated conductance followed sharp steps when allowing for a superposition of dipole frequencies. It is curious that so few resonances appear in the FIR spectrum; far fewer features are observable than predicted from the theory presented for the parabolic quantum dots in a magnetic field, in the previous section. However, the results agree well with theory; Maksym and

Chakraborty [66] showed that the FIR energy couples only to the sum of the radial coordinates, and therefore couples only to the CM system, which is not affected by the electron-electron interaction. This result was confirmed theoretically by Peters [92] and by Brey *et al.* [93]. Therefore, FIR measurements only show the single electron energy states. Two possible models were given for the anti-crossing observed by Demel *et al.* [88]. Chakraborty *et al.* showed that for strictly parabolic confinement, a coupling between $\mathcal{L}_{(\mathrm{CM})} = 2$ and $\mathcal{L}_{(\mathrm{rel})} = 5$ would cause a splitting and also between $\mathcal{L}_{(\mathrm{CM})} = 2$ and $\mathcal{L}_{(\mathrm{CM})} = 5$ would cause splitting. Pfannkuche and Gerhardts [70] showed that this splitting could be a result of a non-parabollic confining potential. It occurs when the ground state changes from $\mathcal{L}_{(\mathrm{CM})} = 0$ to $\mathcal{L}_{(\mathrm{CM})} = 1$.

In principle, it should also be possible to measure thermodynamic quantities to get at the additional structure of the energy levels in quantum dots. The specific heat capacity of electrons has been calculated as a function of magnetic field by Maksym and Chakraborty [66], given by $C_v = \left(\frac{\partial\langle U\rangle}{\partial T}\right)\Big|_V$, who found that the specific heat capacity including electron interactions was far different from the case where the electron-electron interactions were neglected. Oscillatory structure was observed where competing ground states of total $\mathcal{L}$ were observed. The specific heat capacity of two-dimensional electron gases has been measured by Gornik *et al.* [94] using a heat pulse method. In principle, it should be possible to measure C_v for an ensemble of dots.

Maksym and Chakraborty [95] showed another possible probe for the electronic structure in a magnetic field is the magnetization. Their calculation of

$$\mathcal{M} = \left(\frac{\partial\langle F\rangle}{\partial B}\right)\Big|_{T,V}, \tag{2.52}$$

whose matrix elements are evaluated from the operator expression

$$\mathcal{M} = -\frac{e}{2m^*}\sum_{i=1}^{N}\mathbf{r_i}\times(\mathbf{p_i} + e\mathbf{A}), \tag{2.53}$$

for parabolic confinement found that there were discontinuities in the magnetization when the ground state switched to different values of total angular momentum $\mathcal{L}$ and the magnetization was sensitive to the electron-electron interaction. Störmer *et al.* [96, 97] have measured the magnetization for a two-dimensional electron gas; in principle it should be observable in a quantum dot (or array of quantum dots).

Some of the most important experimental verification which has been done on the level structure of quantum dots are the transport measurements. Discussion of these in terms of discrete level structure will be deferred to Section 2.6.3.

2.6 Transport in Nanostructures

The subject of transport in nanostructures is introduced in this section, dealing with the effects encountered in this research project. It should be emphasized that this is intended only to serve as a review of a very narrow portion of a very rich field of study. A review of a more broad scope is the book by Ferry and Goodnick [3] with a more complete list of references. Beginning with single barriers as the most simple unit and progressing through quantum dots, a review of some of the more significant results is pointed out below.

2.6.1 Quantum Point Contacts

One of the most basic nanostructures is a quantum point contact (QPC) within a two dimensional electron gas (2DEG). The structure may be considered to be an ideal 1D conductor into which electrons may be injected from a pair of phase randomizing reservoirs with chemical potential μ_l and μ_r, separated by a barrier which constricts the channel, only allowing a few electrons at a time to pass. Phase randomized means that there is no phase relation between carriers injected into the conducting channel. A typical example is that of a pair of split gates which separate two sides of a modulation doped field effect transistor (MODFET), the

Ohmic contacts serving as the reservoirs. The current allowed to pass through the QPC can be written as the sum of the left going and right going portions,

$$I = \frac{2e}{2\pi}\left[\int_0^\infty dk\, v(k) f_1(k) T(E) - \int_0^\infty dk'\, v(k') f_2(k') T(E')\right], \tag{2.54}$$

where the prefactor is the 1D density of states in k-space, $v(k)$ the 1D velocity, and $T(E)$ is the transmission coefficient for the barrier. At low temperatures, carriers are injected with a maximum energy equal to the chemical potentials in the leads. In the electron quantum limit, and in the linear transport regime, the transmission coefficient $T(E)$ may be considered to be energy independent since linear transport assumes that $E_l \approx E_r$. With these approximations, the resulting expression is the Landauer formula [98, 99],

$$I = \left(\frac{2e}{h}\right)\frac{T}{1-T}(\mu_l - \mu_r) \quad \text{or} \quad G = \frac{I}{V} = \left(\frac{2e^2}{h}\right)\frac{T}{R}. \tag{2.55}$$

Note that the pre-factor for G is significant; it represents the fundamental quantum of conductance. This value is experimentally observable, having the value

$$G_Q = \frac{2e^2}{h} = \frac{1}{12.907} k\Omega^{-1} = \frac{1}{R_Q}. \tag{2.56}$$

Equation 2.56 corresponds to making a four-terminal measurement where the potential drop across the sample is measured without including the potential drop across the leads. If making two terminal measurements, the value of conductance differs from Eq. 2.55 by a factor of $1/(1-T)$ due to the contact potential drop, giving

$$G = \left(\frac{2e^2}{h}\right) T, \tag{2.57}$$

which is lower than Equation 2.55. The generalization to multiple channels (i.e. conduction through more than one 1-D subband) is quite straightforward [100]. Experimentally, quantized conductance steps have been observed in split-gate structures fabricated in high mobility material ($\approx 1 - 2 \times 10^6\, cm^2/Vs$), where the sheet charge density is on the order of $10^{11} - 10^{12}/\, cm^2$. In a typical experiment, a fixed

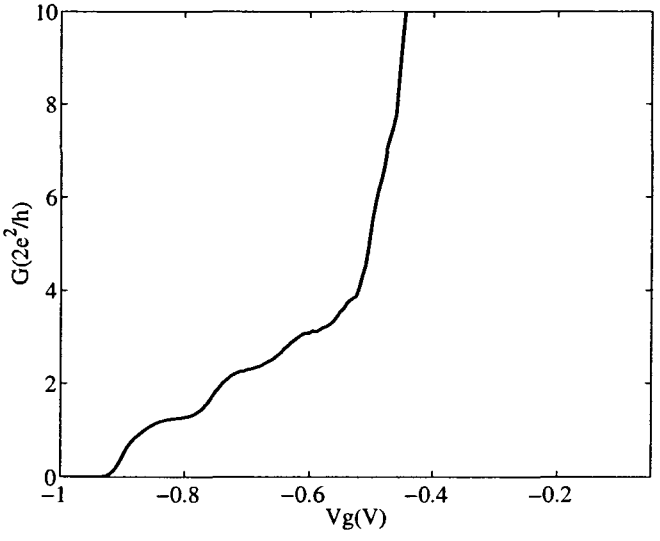

Figure 2.10: Conductance steps in an GaAs/AlGaAs QPC device at 100 mK, After A. Krishnaswamy, 1998

source drain bias is applied and the depletion gates are swept from more open to more closed while the conductance is measured. Examples of such experimental results were first reported in papers by van Wees *et al.* [101, 102, 103] and Wharam *et al.* [104, 105], who found steps in the conductance in integer multiples of G_0 with depletion gate voltage. Figure 2.10 shows an example of these conductance steps, measured in GaAs/AlGaAs QPCs measured at the University of Oregon.

One of the earliest attempts to solve for the transmission coefficient and energy levels at which conductance steps occur for a realistic potential (albeit simple one) was by Kemble [106] using a one-dimensional Schrödinger equation and the WKB approximation. Later, Connor [107], Miller [108], Fertig *et al.* [109], and Büttikker [110] extended the parabolic case to include a saddle point potential in a magnetic field. Fertig [109], solved this case by three methods (exact, WKB approximation, and method of complex variables), and showed that the exact solution

to a potential saddlepoint of the shape

$$V(x,y) = V_0 - U_x x^2 + U_y y^2, \tag{2.58}$$

yields a transmission coefficient of the form ,

$$T(E) = \frac{1}{1+\exp(-\pi\epsilon)}. \tag{2.59}$$

The expression for ϵ is

$$\epsilon \equiv (E_G - V_0)/E_1, \tag{2.60}$$

$$E_G = E - (n+\frac{1}{2})E_2, \quad n = 1,2,\ldots \tag{2.61}$$

and the E_1 and E_2 are functions of the confinement U_1 and U_2 and the cyclotron frequency ω_c. In the absence of a magnetic field, this expression has a nice interpretation in terms of open channels of transport [110, 111]. Classically, all electrons with energy up to the Fermi level when

$$E_F \geq V_0 + \frac{1}{2}\hbar\omega_y\left(n+\frac{1}{2}\right) \tag{2.62}$$

are unbound, and thus open channels to transmission; i.e. $T(E) = 1$; then $V(x,y)$ can be viewed as the n^{th} subband edge of a conducting channel. Due to the quadratic nature of the potential, there is no mixing between subband indexes; the transmission coefficient for electrons with energy above this classically accessible subband edge reduces to [110, 107, 108]

$$T_{nm}(E) = \frac{1}{1+\exp(-\epsilon\pi)}\delta_{nm}. \tag{2.63}$$

This quantum mechanical tunneling probability leads to rounding of the conducting steps as classical transmission is not possible above the subband index n, as seen in the experimental data where $G = G_0 T$ and $T = \sum_n T_{nm}$.

The shape of the conductance plateaus is dependent on the ratio of ω_y/ω_x. For $\omega_y \gg \omega_x$, they are nearly flat, while for $\omega_y \approx \omega_x$ there is little structure. When

there is a magnetic field, the same holds true with the inclusion of the cyclotron frequency; for $\omega_c + \omega_y \gg \omega_x$, the plateaus are nearly flat, while for $\omega_c + \omega_y \approx \omega_x$, there is little structure [110].

This leads to some discussion of the importance of the shape of the potential. For smooth merging of the quantum point contact potential to the channel potential (adiabatic potential), one only needs to calculate the conductance due to the barrier for accuracy up to only small corrections. Using a hard wall potential model, Glazman *et al.* [111] showed that conductance steps appear (meaning reflection coefficient $R \ll T$) if the potential radius of curvature is sufficiently smooth; i.e. $\pi^2\sqrt{2R/d} > 1$ where R is the radius of curvature and d is the diameter of the QPC constriction.

For finite magnetic field, we can predict how the conductance plateaus should behave by examining how many subbands should be classically accessible for transport. This is accomplished by re-examining the expression for the dimensionless parameter ϵ: the number of subbands should be the number such that the numerator of $\epsilon \geq 0$ at the Fermi level of the system

$$E_F - V_0 - \left(n + \frac{1}{2}\right) E_2 \geq 0 \rightarrow n = Int\left\{\frac{E_F - V_0}{E_2} - \frac{1}{2}\right\}. \tag{2.64}$$

Since E_2 is monotonically increases with ω_c, this expression says that as the magnetic field increases, there should be fewer available channels for transport, and the number of plateaus should decrease. This effect is termed the magnetic depopulation of subbands [112].

When the temperature of the system is finite, this also introduces broadening into the shape of the conductance plateaus, as expected by the rounding of the distribution function. Such thermal broadening was indeed observed by van Wees *et al.* [113].

2.6.2 Transport in Quantum Wires

While the topic of transport in quasi-one-dimensional systems is indeed interesting, it is recognized that many of the effects seen in these systems are similar or identical in nature to the phenomena in quasi-zero-dimensional systems. Coulomb blockade will be discussed in some detail in the following section (section 2.6.3). Other interesting effects are due to the discrete density of states, or mode matching which is similar to electromagnetic waves propagating in wave-guides. For some discussion and many references on these matters, the reader is referred to chapters 2 and 3 of the book by Ferry and Goodnick [3].

2.6.3 Transport in Quantum Dots

The topic of transport in quantum dots is a large field. The discussion in this section will be only a narrow subsection, intended to develope the background for the results to be presented in Chapters 4 and 5.

2.6.3.1 Coulomb Blockade

One of the most remarkable observables to come out of nano-structure systems is due to the discrete nature of charge, which is manifested in the phenomenon of Coulomb blockade (CB).

To understand how this effect may be seen in transport measurements, consider the relationships between charge, capacitance, and energy stored in a charged capacitor. The electrostatic potential energy stored on two charged conductors with associated capacitance C is simply

$$E = \frac{Q^2}{2C}, \tag{2.65}$$

where the linear relationship between charge and external electrostatic potential is given by $Q = CV$. To extend this to a system of multiple conductors, an $N \times N$ matrix $\mathbb{C}$ with elements C_{ij} describing the capacitance between each pair of

conductors is introduced, and the total charge may be written by the superposition of the linear relation

$$Q_i = \sum_{j=1}^{N} C_{ij} V_j. \tag{2.66}$$

Extending Eq. 2.65 for the energy to multiple conductor systems, we have

$$E = \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} (\mathbb{C}^{-1})_{ij} Q_i Q_j \quad \text{or equivalently,} \tag{2.67}$$

$$E = \sum_{i=1}^{N} \sum_{j=1}^{N} (C)_{ij} V_i V_j. \tag{2.68}$$

When $\mathbb{C}^{-1}$ is very small, the characteristic charging energy, E, may be on the order of the thermal voltage $k_B T$ for charge corresponding to a single electron. We are not interested in the change in background charge, or polarization charge, on the capacitor, which may be varied continuously by means of the external potential on the i^{th} conductor, but instead the change in number of electrons occupying the dot. In transport measurements, CB is observable only when the total current through the sample is on the order of single electrons at a time. Typically, CB is observed in what is referred to in the language of mesoscopic physics as the single electron tunneling (SET) regime. Currents are on the order of ~nA-pA range, and are governed by the transmission rates through barriers in the system. Within this regime, CB is observable only when the energy associated with charging the capacitor system Eq. 2.65 or Eq. 2.68 is greater than the energy of thermal lattice vibrations, $E > k_B T_l$. If these conditions are met, CB can be understood by considering the schematic for an energy diagram shown in Fig. 2.11. In order for an electron to tunnel into the dot, there must be an available energy state in the dot, and electrons in the channel available at this same energy. The addition of a single electron to the dot system creates an energy barrier for the next electron to enter the dot, and another electron cannot tunnel into the dot until either that electron has an energy to overcome this barrier, or the first electron tunnels out of the dot.

The energy level separation between the N^{th} level and the $(N+1)^{th}$ level is given by the simple expression (with N being the quantum number for number of electrons in the dot)

$$E_{N+1} - E_N = \Delta E_n + \frac{e^2}{2C}, \tag{2.69}$$

where the contribution due to Coulomb charging is $E_C = e^2/2C$ and $\Delta E_n = E_{n+1} - E_n$, and E_n is the bare energy for an electron in an energy eigenstate with principle quantum number n). This results in tunneling events which are single correlated tunneling of electrons, or single electron tunneling (SET). The region in which tunneling is allowed may be changed by either changing the position of the energy levels in the dot by means of an external gate bias, or by changing the chemical potential in one side of the channel with respect to the other (which in effect results in higher electron energy). Sweeping an external gate results in oscillatory conductance as the energy levels in the dot sweep through alignment with the chemical potential in the channel. If the spacing of the peaks is periodic with period E_C, the Coulomb charging dominates the energy level spacing and we have Coulomb blockade transport. However, it is also possible to have oscillatory conductance due to uncorrelated resonant tunneling if the term ΔE_n dominates; this type of transport is not termed Coulomb blockade. For pure Coulomb blockade transport, changing the the chemical potential on one side of the dot with respect to the other results in a step in the conductance for each multiple of $e^2/2C$ difference between the right and left chemical potential in the channel, known as the Coulomb staircase. For dots where the term ΔE_n dominates, it is possible to see regions of negative differential conductance, as in the case of resonant tunneling diodes.

To investigate this idea mathematically, consider the following simple model for a quantum dot. Two ideal reservoirs at chemical potential μ_1 and μ_2 are separated from each other by a perfectly conducting channel. In the middle of the channel, an isolated island is created by placing a tunnel barrier on either side of

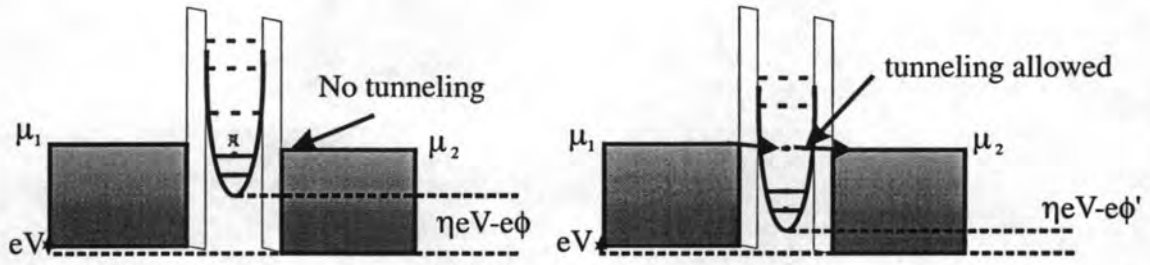

Figure 2.11: Left: When the chemical potentials in the channel do not align with an allowed state in the dot, no tunneling is allowed. Right: Alignment between the chemical potential in the channel and an allowed energy level in the dot leads to allowed tunneling.

the central region. This central conducting region becomes the dot. Phenomenologically, these barriers can be described in terms of a tunneling resistance and a capacitance, $R_t^{l,r}$ and $C_{l,r}$, respectively. The relative chemical potentials between the two reservoirs may be changed by applying a bias across the channel, and the chemical potential in the dot may be coupled externally by capacitively coupling a voltage source (with a Schottky gate) to the dot. An equivalent circuit model is schematically illustrated in Fig. 2.12. To understand the Coulomb blockade easily, consider a case where the dot is large, so that the spacing of energy levels is small compared to the charging energy; $\Delta E_n \ll E_C$. In terms of our phenomenological parameters, the energy can be described by considering the charge stored on the system of capacitors and the work done by the voltage source transfering the electrons in and out of the dot. For our model system, the total energy stored may be written [3]

$$E_s = \frac{1}{C_\Sigma}\left[C_g C_l (V_d - V_g)^2 + C_l C_r V_a^2 + C_g C_r V_g^2 + Q^2\right], \tag{2.70}$$

$$Q = e(n_l - n_r), \quad C_\Sigma = C_l + C_r + C_g, \tag{2.71}$$

where Q is the charge stored in the dot, and n_l and n_r are the number of electrons tunneling through the left QPC into the dot and from the dot through the right

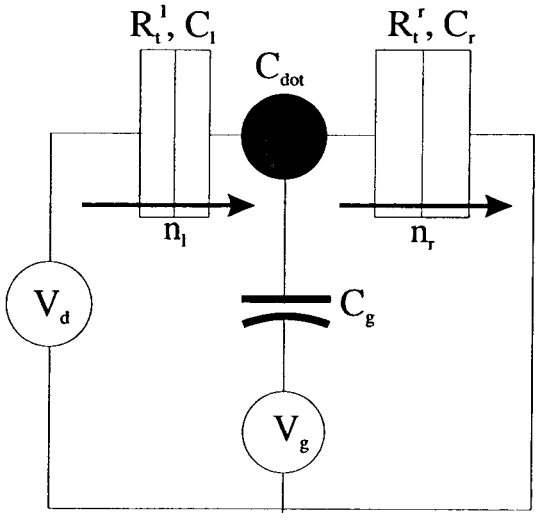

Figure 2.12: An equivalent circuit model for a quantum dot capacitively coupled to an external electrostatic potential by means of a gate.

QPC, respecitvely. The work done by the external voltage sources, W_s, to force the electrons through the tunnel barriers is

$$W_s \equiv \int dt\, V_d I(t); \quad I_{l,r}\Delta t = n_{l,r} e \equiv \Delta Q, \tag{2.72}$$

$$\begin{aligned} W_s(n_l) &= -n_l \left[\frac{C_r}{C_\Sigma} eV_d + \frac{C_g}{C_\Sigma} e(V_d - V_g)\right], \\ W_s(n_r) &= -n_l \left[\frac{C_l}{C_\Sigma} eV_d + \frac{C_g}{C_\Sigma} eV_g\right]. \end{aligned} \tag{2.73}$$

These equations can be combined to get the energy associated with a given charge state;

$$\begin{aligned} \Delta E_l^{\pm} &= \frac{e}{C_\Sigma}\left(-\frac{e}{2} \mp [en - Q_p + (C_g + C_r)V_d - C_g V_g]\right) \\ \Delta E_r^{\pm} &= \frac{e}{C_\Sigma}\left(-\frac{e}{2} \pm [en - Q_p - C_l V_d - C_g V_g]\right), \end{aligned} \tag{2.74}$$

where the $\pm$ indicies $+$ for an electron tunneling into the dot and $-$ for an electron tunneling out of the dot, and $n = n_r - n_l$ is the net number of electrons in the dot. Also, an additional term Q_p has been added to represent the polarization charge

coming from un intentional background impurities as well as other sources (oxide charge and work function differences, for example in Si inversion layers). From these equations, we can develop conditions in which tunneling is blocked, or equivalently, the number of electrons in the dot is stable. At low temperature, it is necessary that electron tunneling events leave the system in a state of lower total energy. If we redefine the gate voltage to include the stray background and polarization charges to be $V_g' = V_g + Q_p/C_g$, the conditions for tunneling are:

$$-\frac{e}{2} \mp [en + (C_g + C_r)V_d - C_g V_g'] > 0 \tag{2.75}$$

$$-\frac{e}{2} \pm [enC_l V_d - C_g V_g'] > 0. \tag{2.76}$$

In addition to the conditions for stability, the Heisenberg uncertainty principle, $\Delta E \Delta t > \hbar$, gives another condition for observation of CB. The time to charge the capacitance of the dot is $\tau = \Delta t = R_t C_\Sigma$, and the energy required for tunneling is $\Delta E \approx e^2/2C_\Sigma$. Combining these two, we have that $R_t \gg h/e^2 \gg 25k\Omega$. The case of a drain bias sweep experiment, where the relative chemical potential of the reservoirs are changed while the dot is left essentially constant, is perhaps easiest to understand. If the quantum point contact pairs defining the entrance and exit barriers are biased such that the exit barrier is much higher than the input barrier ($R_t^r \gg R_t^l$), CB manifests itself in steps in the current through the dot. The step height will be controlled by the portion of the drain bias which drops across the output barrier divided by the tunnel resistance

$$\Delta I \approx \frac{\Delta V_r}{R_t^r}. \tag{2.77}$$

For small V_{DS}, sweeping a gate capacitively coupled to the dot leads to periodic oscillations in the conductance through the dot (if the bare level spacing ΔE_n is small), according to Eqs. 2.75 and 2.76 equating to zero (and hence the Coulomb blockade is lifted).

This corresponds to the situation shown in Fig. 2.11 in which the energy of the next filled level in the dot is aligned with the Fermi energy in the left and right.

The theory for Coulomb blockade in quantum dots and wires has been well studied and agrees with experiment rather well. Both Beenakker [114] and Averin *et al.* [115] extended the theory presented by Kulik and Shekhter [116] for Coulomb blockade in metal tunnel junctions to include the effect of finite energy level spacings, such as in a semiconductor system. Following the development of Beenakker [114], the charging energy $U(N)$ in the dot is given by extension of Eq. 2.68

$$U(N) = \frac{(Ne - Q_{ext})^2}{2C_\Sigma} - \frac{Q_{ext}^2}{2C_\Sigma} \tag{2.78}$$

$$Q_{ext} = C\phi_{ext} \tag{2.79}$$

$$\phi_{ext} \approx \frac{C_g}{C_\Sigma} V_g \quad C_\Sigma = C_{\rm dot} + C_l + C_r + \sum C_{\rm gates}, \tag{2.80}$$

where $c_{\rm dot}$ is the self-capacitance of the dot. Using linear response theory, they calculate the conductance with the assumptions that the tunneling rates are small and temperature sufficiently low, *i.e*,

$$k_{\rm B}T, \min\{\Delta E_n, E_C\} \gg \hbar\Gamma^{l,r} \quad \text{and} \quad k_{\rm B}T \ll \min\{\Delta E_n, E_C\}, \tag{2.81}$$

where the tunneling rates, $\Gamma^{l,r}$, are defined by the sum of the transmission transmission coefficients over all energy levels from the left into the dot or from the dot to the right, respectively i.e.

$$\Gamma^{l,r} = \frac{2\pi}{\hbar} \sum_i |T_i^{l,r}|^2 \delta(\epsilon - \epsilon_i). \tag{2.82}$$

Due to the discrete nature of the allowed tunneling states in the dot, the expression for the conductance in the case of tunneling, (Eq. 2.54) is not given by an integral over continuous states, but rather as a sum over allowed energy states p and electron occupation number, N, as

$$\begin{aligned} G = \frac{e^2}{k_{\rm B}T} \sum_{p=1}^{\infty} \sum_{N=0}^{\infty} \frac{\Gamma_p^l \Gamma_p^r}{\Gamma_p^l + \Gamma_p^r} \\ \times P_{\rm eq}(N)[1 - F_{\rm eq}(E_p|N)] f(E_p + U(N) - U(N-1) - E_{\rm F}) \end{aligned} \tag{2.83}$$

where the probability that there are N electrons in the dot is $P_{\mathrm{eq}}(N)$ and the conditional probability that level p is occupied if N electrons are in the dot is $F_{\mathrm{eq}}(E_p|N)$ as defined by thermodynamic relations of the free energy and energy related to the charge, and $f(E_p+U(N)-U(N-1)-E_F)$ is the Fermi Dirac distribution function.

This theory predicts that there are periodic oscillations in the conductance, whose period is given by the condition that an allowed energy level in the dot is aligned between the chemical potential in the leads. Such oscillations are observed experimentally when a gate bias coupled to the dot is swept; the period is given as

$$\Delta V_g = \frac{C_\Sigma}{C_g}\left(\frac{E_{n+1}-E_n}{e}\right)+\frac{e}{C_g}. \tag{2.84}$$

It should be emphasized that there are really two transport mechanisms described by Eq. 2.84, distinguished by which term dominates the spacing. If the first term in Eq. 2.84 dominates, the conductance oscillations should be referred to as resonant tunneling, as the discrete levels of the dot dominate. If the second term dominates, we have true Coulomb blockade oscillations. In the case of CB, true SET behavior occurs, meaning that only one electron tunnels through the dot at a time due to a charging energy barrier. These tunneling events are thus correlated. However, in resonant tunneling, tunneling is restricted by discrete allowed states; the barrier due to Coulomb charging is small, and thus the tunneling is not correlated.

An example of ideal Coulomb blockade oscillations due to a gate bias sweep is shown in Fig. 2.13 (assuming the first term in Eq. 2.84 is small). The Beenakker model [114] described by Eq. 2.83 leads to the following expression for the line shape close to one of the conductance peaks

$$\frac{G}{G_0} \approx \frac{1}{\cosh^2\left(\frac{\delta_{E_c}}{2.5k_BT}\right)}, \tag{2.85}$$

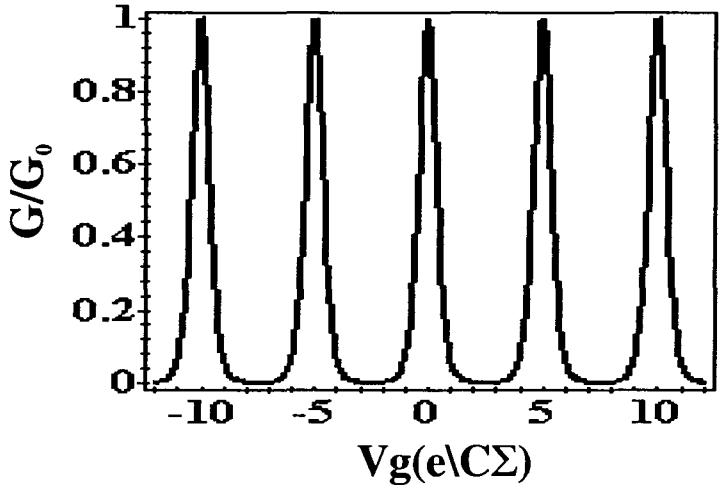

Figure 2.13: Coulomb Blockade oscillations according to the theory of Beenakker for $k_{\mathrm{B}}T = 0.1 \times e^2/2C_\Sigma$ and $C_g = 0.2 \times C_\Sigma$.

where the δE_C is defined as

$$\delta_{E_c} = \frac{eC_g}{C_\Sigma}(V_g^{\mathrm{res}} - V_g), \tag{2.86}$$

with V_g^{res} being the gate voltage of the resonance.

By pulling the lead chemical potential up or down with respect to an energy level in the dot by means of the gate bias, the region of drain bias in which Coulomb blockade occurs in the source-drain bias sweep can be shifted positive or negative. By combining gate bias sweeps with stepped drain bias, using the conditions for allowable tunneling, a stability plot can be generated in the plane of gate and drain bias, as in Fig. 2.14 and 2.15. Regions where tunneling is blocked by stable number of electrons in the dot are diamond shaped (the light regions in Fig. 2.14 and 2.15); these regions are bordered by regions which are unstable, meaning that they are energetically favorable to tunneling. Coulomb blockade oscillations are routinely observed in high mobility semiconductor material, and were first reported in the context of a semiconductor quantum dot by Meirav *et al.* [117] in 1990.

More recently, Coulomb blockade-like features have been reported in Si devices, although it is not clear that true CB is observed in any of the experiments. Examples include work by Matsuoka *et al.* [44, 45, 46, 47, 48] in a dual-gate split-gate structure, Chou *et al.* [49, 50, 51] with their etch-defined dot on SOI operating up to room temperature, Khoury *et al.* [118, 119, 120], Gunther *et al.* [121] (the work reported in this dissertation), and Fujiwara *et al.* [54].

2.6.3.2 Excited Discrete states

The variation in the gate period of the Coulomb blockade oscillations in quantum dots or resonant tunneling peaks in resonant tunneling diodes (RTD's) due to quantization of the energy levels in quantum dots predicted by Eq. 2.84 has been observed experimentally [122, 123, 124, 125, 126, 127]. Reed [122] found peaks in the resonant tunneling through small area vertical quantum well structures, which were attributed to the singular density of states associated with discrete energy levels. In lateral structures exhibiting Coulomb blockade behavior, discrete excited states show up in the stability plots in the $V_d - V_g$ plane as regions of suppressed current in single electron tunneling regions (SET- regions) along the drain bias (V_d) axis, as was seen in the work by Johnson *et al.* [123], Foxman *et al.* [124] and Weis *et al.* [125]. A schematic of these stability plots are given in Fig. 2.14 and 2.15. These variations in linear transport were in fact predicted by Averin *et al.* [115, 128]

By studying the linear and non-linear magnetoconductance spectrum of a quantum dot, Stewart *et al.* [127] found correlations between the m^{th} excited state of an N-electron dot and the ground state of an $(N + m)$ electron dot. These correlations suggested that these dots behave like collections of single-particle like systems, although there were also some characteristics of many bodied interactions. Ashoori *et al.* [129, 130, 131] developed another probe for studying the energy spectra of quantum dots by fabricating a quantum well with one barrier thin enough to allow

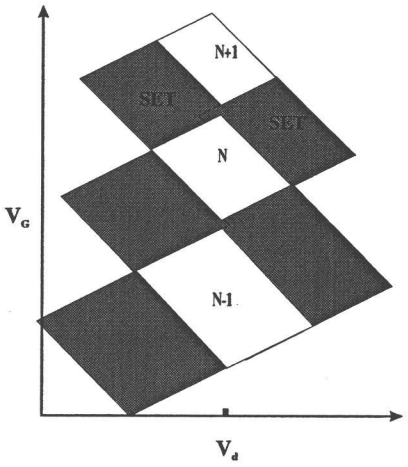

Figure 2.14: An example of a stability plot. In a gray-scale (dark for high conductance) plot of differential conductance, the Coulomb blockade regions of zero conductance correspond to the white areas.

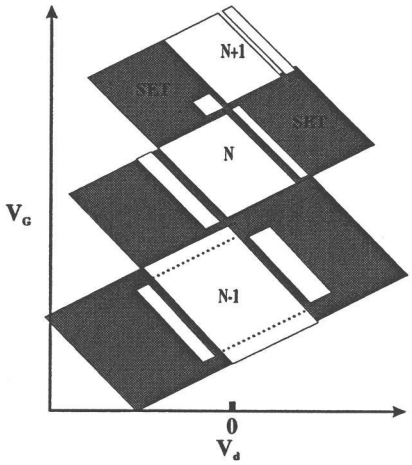

Figure 2.15: A stability plot where discrete states are important. The central regions are CB regions, and dotted lines indicate regions of suppressed but not zero conductance.

tunneling, and placing a gate structure which caused lateral confinement forming a quantum dot. By studying the peaks in the capacitance as a function of gate voltage, the addition energy was determined. The dependence of these energy levels in a magnetic field was shown to behave as the single particle Darwin-Fock states (Eq. 2.29, 2.30) discussed in Section 2.5.1 with the exception of the observation of a spin triplet to singlet transition and some features which point towards the fractional quantum hall effect (which is a many body effect).

2.6.3.3 Few Electron Dots and Artificial Atoms

When the number of electrons in a quantum dot approaches zero, the conductance characteristics in the dot reveal a rich spectrum of peaks remarkably similar to the discrete levels of atoms [132, 133]. Perhaps the nicest example of such devices are those of Tarucha *et al.* [10, 134, 135], who fabricated a cylindrically symmetric quantum dot structure based on a vertical resonant tunneling diode structure,

which was etched into a cylindrical pillar and a Schottky side gate was evaporated around the base and side wall to controllably deplete the dot. The measured addition energy (the energy required to add two successive single electrons to the dot) has a filling structure much like the atomic shell structures in atoms. Filled shells give rise to large addition energies (relative to other filling numbers) for the second, fourth, sixth, 12th, and 18th electrons, in agreement to exact and Hartree-Fock calculations [86, 76, 82] assuming a harmonic confinement potential model. In low magnetic field there appears to be a spin singlet (anti-parallel$|\downarrow\uparrow\rangle$) to spin triplet (parallel $|\downarrow\downarrow\rangle|\uparrow\uparrow\rangle$)transition with increasing field. Evidence for such a transition has in fact been observed in other experiments [136, 130, 137], as was predicted by Wagner *et al.* [73]. In addition, Tarucha *et al.* observed transitions between ground states and excited states which allow for reduction in Coulomb energy and exchange energy, crossings between excited states, and crossings between parallel and anti parallel spins [135], which at higher magnetic field switches to a spin-degenerate pairing behavior, in agreement with Hund's rule.

2.6.3.4 Spin Blockade

Another interesting feature seen in small systems with correlated multiple electrons is that of spin blockade of linear and non-linear transport. In several experiments in which excited states were accessible in the transport (typically in the non-equilibrium regime), regions of negative differential conductance were observed [123, 138, 139, 125]. A possible explanation for this was presented by Weinmann *et al.* [140, 141] in terms of spin blockade. There, transition rates were calculated taking spin into account by including Clebsch-Gordan coefficients:

$$\begin{aligned}\Gamma_{ji}^{l,r} = \frac{T^{l,r}}{2}\left|\langle S_i, M_i, \frac{1}{2}, \pm\frac{1}{2}|S_j, M_j\rangle_{\mathrm{CG}}\right|^2 \\ \times \left(f_{l,r}(E)\delta_{n_j,n_{i+1}} + [1 - f_{l,r}(-E)]\delta_{n_j,n_{i-1}}\right).\end{aligned} \tag{2.87}$$

What comes out of Eq. 2.87 are selection rules which lead to slow depopulation of states. This slow depopulation leads to reduced conductance (or negative differential conductance) because the current is due essentially to electrons oscillating between accessible states. In the Coulomb blockade regime, the accessible states are electron numbers N and $N-1$. In one type of spin blockade, the oscillations are transitions between correlated many electron states with electron number and maximal spin which must end in lower total spin states; $|N, S = N/2\rangle \rightarrow |N-1, S' = (N-1)/2\rangle$. But states with $S < N/2$ may go either to higher or lower total spin states, so depopulation is slower from states in the spin polarized state, leading to a negative differential conductance.

Another anomalous behavior is that of increase in the Coulomb blockade peak amplitude with increasing temperature, as seen in experiments by Nichols *et al.* [142]. The possible spin blockade mechanism responsible for suppressing this peak at very low temperatures is due to transitions between ground states requiring $\Delta S > 1/2$, forbidden for ground states at zero temperature; i.e. $\{E_0(N), S\} \leftrightarrow \{E_0(N-1), S'\}$; $|S - S'| > 1/2$. At higher temperatures or under non-equilibrium conditions, excited states may become available to allow these forbidden transitions.

2.6.3.5 Coulomb Blockade of Activated Conduction

Matveev and Glazman [143] showed that it is possible to have Coulomb blockade in transport characterized by activated conductance over barriers rather than tunneling as discussed earlier as the dominant mechanism of transport. They start with the expression for the limiting conductance over a barrier,

$$G_b = \frac{2e^2}{h} \int_{-\infty}^{\infty} \left[-\frac{\partial n_F(\epsilon)}{\partial \epsilon} \right] T(\epsilon) d\,\epsilon, \tag{2.88}$$

where the tunneling coefficient $T(\epsilon)$ was that describing a saddle point contact

$$T(\epsilon) = \frac{1}{1 + \exp\left(\frac{U-\epsilon}{T_0}\right)}, \quad T_0 = \frac{\hbar\omega}{2\pi}, \tag{2.89}$$

which below is extended to include finite temperature. In the region of finite temperature $U \gg T > T_0$, it was shown that the conductance shape has the form

$$G(N) = G_b(U)\frac{\exp\left[\frac{\delta_{E_C}}{(2\lambda-1)k_BT}\right]}{\sinh\left(\frac{\delta_{E_c}}{k_BT}\right)} \tag{2.90}$$

$$G_b = \frac{2e^2}{h}\frac{\pi T_0}{k_BT\sin(\pi T_0/k_BT)}e^{-U/k_BT}. \tag{2.91}$$

In Eq. 2.91, the parameter λ describes the asymmetry of the tunneling process,

$$\lambda = U - (E_N - E_{N-1}). \tag{2.92}$$

The expression δ_{E_c} was defined earlier in Eq. 2.86 This expression gives peaks in conductance similar to those predicted by the expression of Beenakker [114], Eq. 2.85, but are shifted from the resonance and are asymmetric in shape.

3. FABRICATION AND EXPERIMENTAL SETUP

3.1 Device Overview

The quantum dots studied in this work are based on an n-channel Si MOSFET structure, with the modification of having thin metal depletion gates sandwiched between two oxide layers underneath the metal inversion gate used to induce the inversion layer. This structure is shown in the cross-sectional diagram of Fig. 3.1. The lower of these oxides is a 5 nm thick thermal oxide, while the upper is a 95 nm thick oxide deposited by the Plasma Enhanced Chemical Vapor Deposition (PECVD) technique. The inversion gate partially overlaps the source and drain ohmic contacts to ensure the conducting channel extends over the entire active region of the device. In addition, Ohmic contacts are provided in the Hall configuration surrounding the device for additional characterization of the inversion layer, with legs of the inversion gate extended to partially overlap them. To minimize background leakage current, the channel region as defined by the inversion gate is surrounded by the channel stop region formed by implantation.

There were four different geometries of depletion gates fabricated, of which three produced useful transport measurements. The generic structure is shown in Fig. 3.2. The size of the dot refers to the inner length of the box defined by the depletion gates, which produce the electronic confinement potential. The square geometries consisted of 500 nm and 200 nm dots. Also fabricated were 100 nm $\times$ 200 nm asymmetric dots. A summary of the device structures measured is given in Table 3.1.

A sketch of the 200 nm symmetric configuration is given in Fig. 3.2. The inset shows an SEM micrograph of the depletion gate structure. The lower gate forms quantum point contacts (QPC's) with the left and right gates shown in Fig. 3.2,

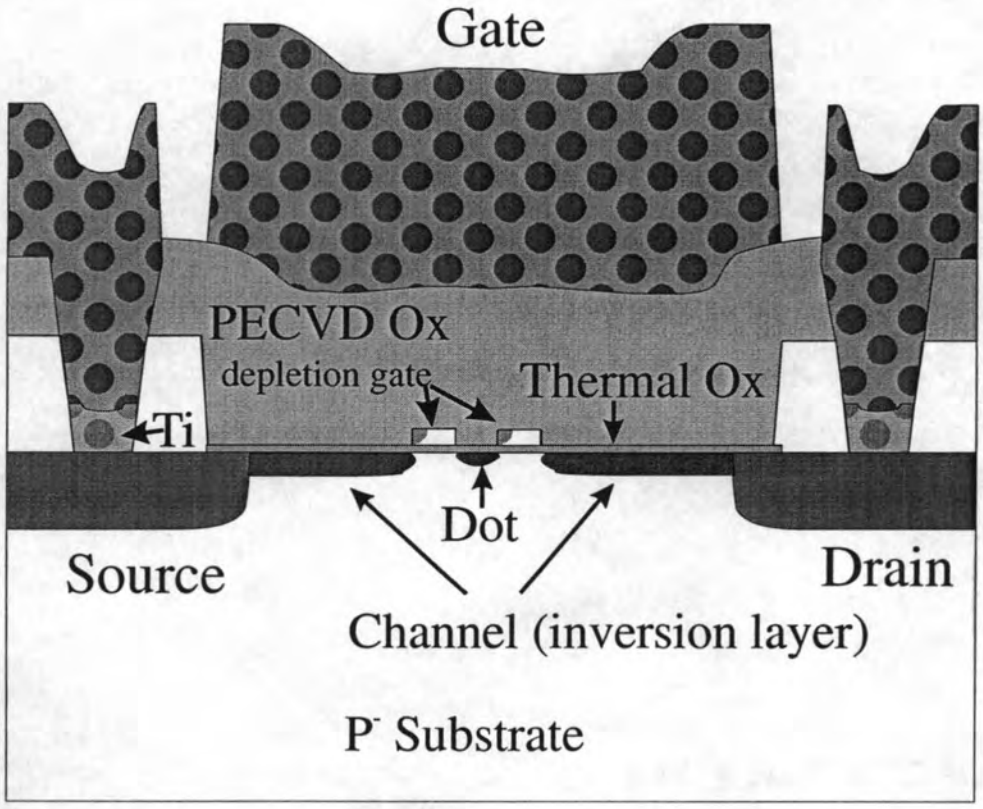

Figure 3.1: A schematic of the vertical layer structure for the quantum dots. The vertical scale has been exaggerated to allow all features to be represented.

Table 3.1: A summary of the device geometries used for the experimental measurements

	Symmetric		Asymmetric
lateral dimension	500 nm × 500 nm	200 nm × 200 nm	100 nm × 200 nm
device	D2, D10	D9, D12	D11

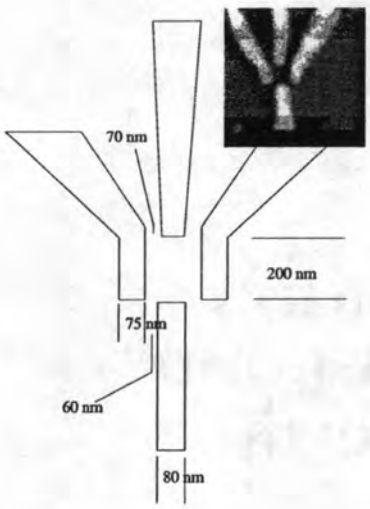

Figure 3.2: A sketch of the symmetric 200 nm depletion gate geometry. The dimensions of the as drawn metal lines (before liftoff) are given. Inset: An SEM micrograph of the 200 nm gate geometry.

controling transport into and out of the dot. The middle upper gate controls the geometry of the dot itself, and is referred to as the 'plunger'. These depletion gates will often be referred to in the remainder of this work by abbreviations SGL, SGM, SGR, SGB, for side gate left, side gate middle, *etc* while voltages applied to these gates will often be denonted as V_{SGL}, V_{SGM}, V_{SGR}, or V_{SGB}.

3.2 Fabrication

In this section, the fabrication of devices will be described in some detail. When introducing a new fabrication technique, a full description will be included. In later cases, these methods will just be referred to by name. As there were several revisions to the process as development progressed, an explanation of these differences and the reasons for changes will be included. The included figures illustrate the layer structure and layout of the devices at each stage of the process.

Also, differences from typical industry NMOS practices will be pointed out. An excellent reference for silicon processing in general is the series by Wolff and Tauber [144].

3.2.1 p^+ REGION

This step in the initial process flow is used to make degenerately doped Ohmic contact regions to the substrate. It introduces many of the process steps used during the device fabrication.

Piranha clean wafers: the Piranha clean procedure is as follows: Immerse wafers in hot piranha (7:3 sulfuric acid:hydrogen peroxide) for 5 min; rinse in running deionized water (DI) for 10 min; etch wafers in dilute hydrofluoric acid:deionized water (1:100) for 10 min; spin dry in rinser/dryer.

400 nm wet thermal oxide growth: Immediately remove wafers from rinser/dryer and insert into mouth of wet oxide tube which is preheated to 700 C . Wait 5 min, and insert to center of 3-zone furnace. Ramp temperature of furnace tube to 1050 C . Oxidize in O_2-H_2O ambient for 50 min. The ratio of flows for the gases should be approximately 2 parts H_2 to one part O_2, which produces a water vapor with a slight excess of oxygen due to unreacted gases. Remove the wafers at a rate of six inches per minute to avoid warping the wafers.

Pattern wafers with implant mask: Spin on Hexamethyl Di-silazane (HMDS) at 5000 rpm for 30 s. Spin on positive photoresist OCG825 for 30 s at 5000 rpm; Soft bake 15 min at 80 C; Expose wafers in the Canon contact aligner using light integra setting 6.0; spray-develop wafers for approximately 30 s; rinse in running DI and blow dry with dry N_2 gas. Post bake the wafers for 20 min at 130 C to harden resist.

Etch wafers in buffered oxide etch (BOE): Immerse wafers in BOE, allowing sufficient etch time for slight over-etching of SiO_2; rinse for 5 min in running DI. Rinse and spin-dry in the rinser/dryer. Strip the photoresist: Immerse the wafers

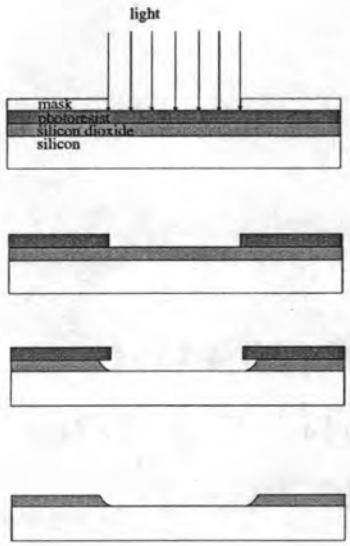

Figure 3.3: A schematic of the wet etch technique

in hot Microstrip 2000 photoresist stripper followed by hot methanol, acetone, and running DI (*AMD clean*). The etch process is shown schematically in Fig. 3.3.

Ion implantation: Ion implantation was done externally at Ion Implantation Services. The implantation has the following specifications: Implant BF_2 at 40keV acceleration to an area dose of $2 \times 10^{15}/\,\mathrm{cm}^2$ with 7 degrees of tilt. Strip the oxide in BOE and rinse the wafers in running DI and spin dry.

A process revision replaces the p^+-ohmic contact with a channel stop implant. The dosage is reduced to $1 \times 10^{15}/\,\mathrm{cm}^2$ and only the active channel region is masked off by the implant mask. The reasoning behind this change is to reduce the background leakage around the channel. This step is also more in line with a standard MOS process.

3.2.2 Field Oxide/n^+ Ohmic Contacts

Piranha clean wafers; Grow 650 nm wet thermal oxide at 1050 C; pattern the wafers in the Canon Aligner for the n^+ ohmic contact regions.

Post bake the wafers at 130 C for 30 min to harden the photoresist mask. Then etch the wafers in BOE to the Si surface; strip photoresist; then send wafers out for implantation with the following specifications: Implant at 40 keV As at a dose of $2 \times 10^{15}/\,\mathrm{cm}^2$ at a tilt of 7 degrees.

In the second revision to the process, several changes were made to this step: after stripping the photoresist, perform a hot Piranha clean; re-grow a thin dry thermal oxide (about 10 nm) at 1000 C. The purpose for this thin oxide is to improve on the randomization of the incoming ion beam during implantation to minimize channeling. The wafers are sent out for implantation, where a dual implant is performed: first implant $3 \times 10^{15}/\,\mathrm{cm}^2$ As at 50 keV followed by P at 50 keV to a concentration of $1 \times 10^{15}/\,\mathrm{cm}^2$. This implantation scheme should leave a higher concentration of less mobile As impurities close to the surface with further thermal steps, ensuring degenerate contacts at low temperatures. In addition, the more mobile P ions penetrate further into the substrate, and diffuse laterally during thermal processes later to form a doping profile which reduces the fields at the drain region in high-bias operation, which leads to hot carrier effects and merging of the source and drain depletion regions for high drain bias for short-channel MOSFETs.

This process step differs greatly from most modern MOS practice. There, the field oxide is grown using a local oxidation process (LOCOS) in which the active region of the device is masked off using Si_3N_4 as an oxygen diffusion barrier. This step serves to make a smooth profile, allowing for better step coverage on later deposition steps. Formation of the source and drain Ohmic contacts is delayed until after the gate is formed, utilizing a self aligned process. In addition, the implant step is usually followed by a drive-in step to diffuse the dopant impurities to the desired junction depth.

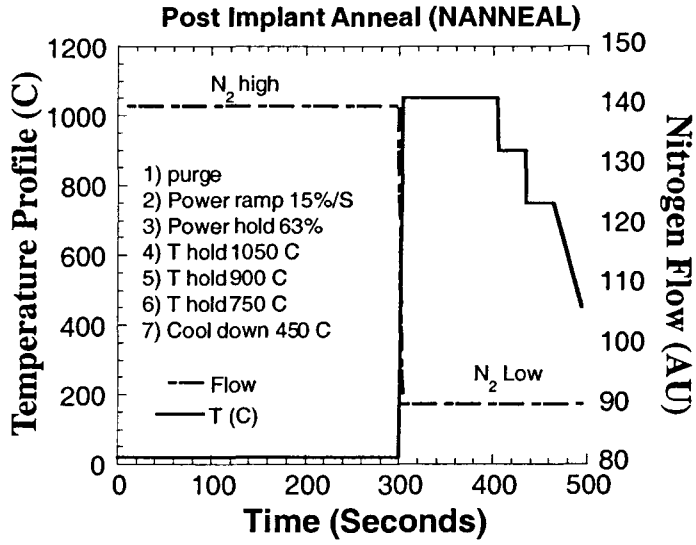

Figure 3.4: A flow diagram showing the process variables in the post-implant anneal step.

3.2.3 Post Implant Anneal

Piranha clean the wafers. Next, in a Tamarack rapid thermal processing reactor, under N_2 ambient, a ràpid thermal anneal is performed at 950 C for 60 s. The susceptor for this reactor is a clean Si wafer. This step is done to anneal out damage incurred during the ion implantation and activate the dopant species. Rapid thermal processing is chosen to minimize the thermal load on the samples, which causes unwanted redistribution of dopant impurities. It is also chosen to minimize propagation of defects which may otherwise occur at lower temperatures which would be a unavoidable with conventional annealing. Fig. 3.4 gives the relevant times, temperatures, and gas flows for the post-implant anneal.

Following the post implant anneal, the upper 250 nm of oxide is stripped in BOE to remove the dopant-rich oxide layer.

3.2.4 Define Active Region and Gate Oxide Growth

Pattern the wafers with the active region mask layer using the Canon aligner and OCG 850 PPR. Post bake wafers at 130 C for 20 min and then etch the field oxide in BOE to expose the Si surface. Finally, strip the photoresist. Perform a *piranha clean*, and insert the wafers into the mouth of the dry oxide tube heated to 950 C. After a three minute delay, push the wafers to the center zone at a rate of 6 in per minute. Insert a delay time to allow for thermal equilibrium to be achieved, and then grow a 5 to 10 nm oxide in pure O_2 ambient. Ramp the furnace to 1000 C, and then anneal 15 min in pure N_2 ambient. Remove the wafers at 6 in per min.

As previously mentioned, this step differs in modern MOS processes by the use of a local oxidation (LOCOS) process in which the active region would be masked off during the field oxide growth by a silicon nitride (Si_3N_4) diffusion barrier, which would be removed following the oxidation, a sacrificial oxide would be grown and stripped, and then the active region would be prepared.

3.2.5 Electron Beam Guide Pattern

This first metalization step lays down a pattern that aids in the alignment process for electron beam lithography and demonstrates another basic process which will be repeated several times, that of *lift-off*. Illustrated in the schematic Fig. 3.5, this process step goes as follows: pattern the wafers using the electron beam guide mask in the Canon aligner with OCG 825 PPR. After development and rinsing, dry wafers very thoroughly and place in the vacuum jar of the electron beam evaporator (either the CHA or TorrVac) and evacuate to as high vacuum as can be reached within the time allotment (typically, about 2×10^{-6} Torr). Sequentially evaporate 40 nm Cr followed by 8 nm Pt. Lift off all unwanted metal by sequentially immersing the wafer in hot Microstrip 2000, followed by hot methanol, acetone, methanol,

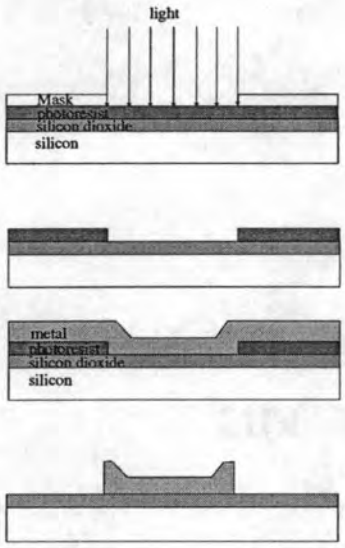

Figure 3.5: A schematic of the liftoff technique.

and then rinse in running DI. It may be necessary to use a jet of Microstrip or use ultrasonic agitation. In this way, the pattern may be transferred to the metal without the need for etching the metal.

The metals for this process were chosen for the following reasons: Cr adheres nicely to, does not interact heavily with, nor diffuse through Si or SiO_2. Pt has a high Z number, thus gives off a high yield of secondary electrons and is thus easily seen under the scanning electron microscope (SEM) but diffuses rapidly through Si and SiO_2, beads up on the surface of Si, and forms a highly conducting silicide. Thus, Cr acts as a wetting and adhesion layer for the Pt while acting like a diffusion barrier, and the Pt serves the purpose of alignment marks for the SEM. Other possible choices studied included Co [145, 146], Ti [147], Nb [147, 148], V, Ta, and Zr [149].

This process step has no analog in the modern MOS process. In addition, it should be noted that in Si processing, the liftoff technique is rarely used. Instead, uniform layers of metal are deposited and then the wafers are patterned such that

the photoresist masks off areas in which metal should not be etched. Then the unmasked areas are etched off, typically using a plasma etching process.

3.2.6 Titanium Contact Metal/ First Via layer

The first major revision to the process was to add a $TiSi_x$ contact metal to enhance the contacts at low temperature. Ti is chosen for several reasons. It reacts strongly with SiO_2, dissolving thin layers of the oxide and moving the oxygen away from the Si surface to form TiO_x, which is also conducting. TiO_x is highly conducting and stable, and the reaction occurs even at room temperature [150]. Overall, the effect is to remove the native oxide which is always present and forms a conducting silicide in its place. In addition, it forms a diffusion barrier against Al spiking into the Si. In the industry, a Ti/W barrier layer is often used due to its better endurance. Alternatively, a Ti/TiN stack may be used to replace the Ti.

The order in which this step occurs in the process went through several changes, eventually occurring after the guide pattern and before the electron beam lithography in the final process sequence for efficiency. The process follows: Pattern the wafers with the first via mask in the Canon aligner using OCG 825 PPR. Do a post bake for 15 min at 130 C. Etch in BOE long enough to slightly over-etch through the oxide layer, rinse and thoroughly dry. Immediately place the patterned wafers into the electron beam evaporator and evacuate. Sequentially evaporate 30 nm Ti followed by 30 nm Al, remove from the vacuum jar, and lift off.

3.2.7 Define Depletion Gates/Electron Beam Lithography

Electron beam lithography (EBL) is the key research tool, which allowed fabrication of the nano-scale electronic devices studied here. It has advantages and disadvantages over conventional photolithography. The key advantage is the ability to achieve dimensions not yet available by photolithography. The major disadvan-

tage is that it takes a great deal more time per exposure compared to photolithography because EBL is a direct write technology, whereas photolithography utilizes a blanket exposure masked off by some sort of optical mask. In either case, the same principle of operation is used. A resist thin film is placed on the surface of the sample to be patterned, and then selectively exposed to energetic particles. The sample is then placed in some developer chemical, which reacts to the exposed regions. In the case of EBL, the particles are electrons that are rastered across the surface by controlling the scan coils of a scanning electron microscope which has been modified for this purpose. A computer program is used to translate a pattern file, typically a computer aided design (CAD) drawing, into an analog signal. In the case of this research project, the program and lithography package used was ELPHY plus, manufactured by Raith GmbH. The electron beam lithography process steps will be discussed in detail below; there are several good references on the subject of electron beam lithography [151, 152, 153, 154].

Other approaches to achieving small pattern size include patterning self-assembled monolayers with electron beam lithography capable of producing sub-10 nm lines [155], and patterning by scanned probe oxidation (using a biased atomic force microscope (AFM) tip in an oxygen ambient) capable of 10 nm lines [156, 157].

3.2.7.1 Sample Preparation

For EBL, the following process steps were used. Clean the wafer using an *AMD clean* with ultrasonic agitation. Spin on poly methyl methacrylate (PMMA) solution (2% PMMA in chlorobenzene) for 30 s at 4,000 rpm and bake at 170 C for at least 2 hours. Place the sample in the main chamber of the scanning electron microscope and evacuate to a base pressure of $\approx 10^{-6}$ Torr. Once the base pressure has been reached, the filament is brought to saturation at an acceleration of 40 keV. For lithography purposes, a high acceleration is desired to minimize secondary

exposure from secondary electrons reflected from the substrate. Higher acceleration causes the back scattering of secondary electrons to occur deeper in the sample. After the filament is saturated, the column must be carefully aligned electronically, and image shift between probe current modes (LM and STD) must be corrected for. Stigmation of the lenses must also be corrected for to achieve the smallest symmetric spot sizes.

3.2.7.2 Coarse Alignment

After the lenses have been corrected for image shift and stigmation, the sample coordinates must be set up. The electron beam lithography set up is equipped with a laser-interferometer to make this procedure convenient and accurate. The particular model is capable of controlling the position of the sample relative to some internal three dimensional axis to a precision of several nanometers through DC servo motors and piezoelectric transducers. The location is monitored by means of a laser interferometer. In the process of coarse alignment, the known dimensions and location of the optical alignment marks are determined, and the software performs a coordinate transformation. This process allows for moving from one pattern exposure to the next by simply typing in the desired sample coordinates.

3.2.7.3 Fine Alignment and Focus

After the coarse alignment is completed, the sample is moved to the location of the electron beam alignment metal pattern at the corners of the active region. The patterns are scanned into the computer at the magnification of writing, and their centers are picked manually by a mouse click. Now the best focus must be attained by the following procedure: in LM mode, get the best focus possible up to 70kX magnification. Increase the magnification to 250X, and use the spot scan mode for long enough to burn a hole in the PMMA; then focus on this spot.

3.2.7.4 Exposure

After alignment and focus are completed, the pattern is exposed. First the current must be accurately measured to ensure that the correct dosage of charge is delivered. This is done by moving to the Faraday cup which is mounted on the stage, and measuring the current with a Keithley electrometer. Then move back to the working position and expose the pattern.

Due to the extremely small nature of the patterns used for these devices, and the fact that the surface is a dielectric which charges and deflects the trajectory of the electron beam, a strategy was necessary to achieve fine features. This strategy was manifested in the drawing of the pattern file. Issues that arose include:

- Since the current supplied by the filament drifts over the time of the exposure, the most critical features are patterned first to ensure the closest possible dose.
- To counter the jitter which occurs when the beam settles into a new location, features are always written to scan from far away from critical features toward the feature.
- To counter exposure by secondary electrons from pattern features in close proximity to other features, the relative dose is reduced. Typically, the dosage in these regions is reduced to a factor of .90 to .95 of full scale.
- Small, isolated features need higher dose to ensure full exposure, so these are typically scaled to 1.1 to 1.2 of full scale.
- The smallest lines are written as a series of single pass lines. The uniformity of dose across the lines in ensured by increasing the dose on the outer line relative to inner line of a series.
- three layers of exposure are performed, with each having its own exposure parameters. The finest, most sensitive layer is exposed first, followed by a

slightly less fine portion of the pattern further from the center. These are exposed in the STD current mode, which has a higher focus resolution. Large area patterns are written after all fine patterns have been exposed using the high current, lower resolution LM mode to reduce the exposure time. Some overlap is required to ensure that no lines are severed between layers after re-positioning and fine aligning on the second pass.

Before an actual pattern is written, it is necessary to optimize exposure parameters for the exact conditions a given sample will have. These parameters will change with the age of the resist and filament, as well as the topography of the sample. Therefor, optimization must occur on identical samples with nearly identical resist and filament conditions and requires a full metal deposition and liftoff. For the patterns developed for this research project, the parameters were generally very close to the parameters given in Table 3.2.

Table 3.2: Exposure parameters for Electron Beam Lithography

A summary of typical EBL exposure parameters yielding successful results								
Mode	LM			STD				
Parameter	S [5]	AS [6]	Mag.	S	AS	SPLD [7]	SPLS [8]	Mag.
Layer 1	NA	NA	NA	272	2 (.0024)	2.5	2.3685	3.77kX
Layer 2	NA	NA	NA	312	2 (.0024)	NA	NA	3.77kX
Layer 3	285	3 (.0036)	1kX	NA	NA	NA	NA	NA

3.2.7.5 Develop PMMA

After the sample has been exposed and the SEM has been cooled down, the sample is removed and then developed. Immerse the sample in PMMA developer (CEL:MIBK:MIK) for 20 s, followed by 20 s in isopropyl alcohol (IPA) and then a thorough rinsing in running DI. The sample is next inspected under an optical microscope to ensure that the patterns were exposed properly. If the patterns do

---

[5]Sensitivity (nC/cm^2)

[6]Area line spacing (Pixels) (μm)

[7]single pixel line dose (nC/cm)

[8]single pixel line spacing (nm)

not come out well, the whole EBL process must be repeated. If the patterns are sufficiently well defined, deposition of the depletion gate metal follows.

3.2.7.6 Depletion Gate Metalization

Thoroughly dry (both dry N_2 and 100 C oven bake for about 5 min) the sample and place in the vacuum jar of the E-beam evaporator and evacuate to $\approx 10^{-6}$ Torr. Evaporate 30 nm Cr at a rate of 0.1 nm/s. Remove the sample from the vacuum jar and liftoff in hot Microstrip, followed by a thorough cleaning including ultrasonic agitation in hot methanol, acetone, and then running DI. Careful inspection is important at this stage to ensure that the metalization and liftoff are complete and the pattern transfer is good. Note that the smallest patterns are much below the resolution limits of optical microscopy at this point, but the larger features are clearly observable.

As mentioned in Section 3.1, three different depletion gate geometries were fabricated by electron beam lithography:

- symmetric 500 nm × 500 nm (D2, D10)
- symmetric 200 nm × 200 nm (D9, D12)
- asymmetric 100 nm × 200 nm (D11)

SEM Micrographs of the first three geometries are shown Fig. 3.6, which gives a view of the devices before the second oxide is laid down. After the oxide is deposited, it is no longer possible to obtain an SEM image, due to excessive charging from the oxide. The quoted dimensions of the dot are the distance between the inside edges of the metal lines. In the Fig. 3.6, the scale is not the same between micrographs.

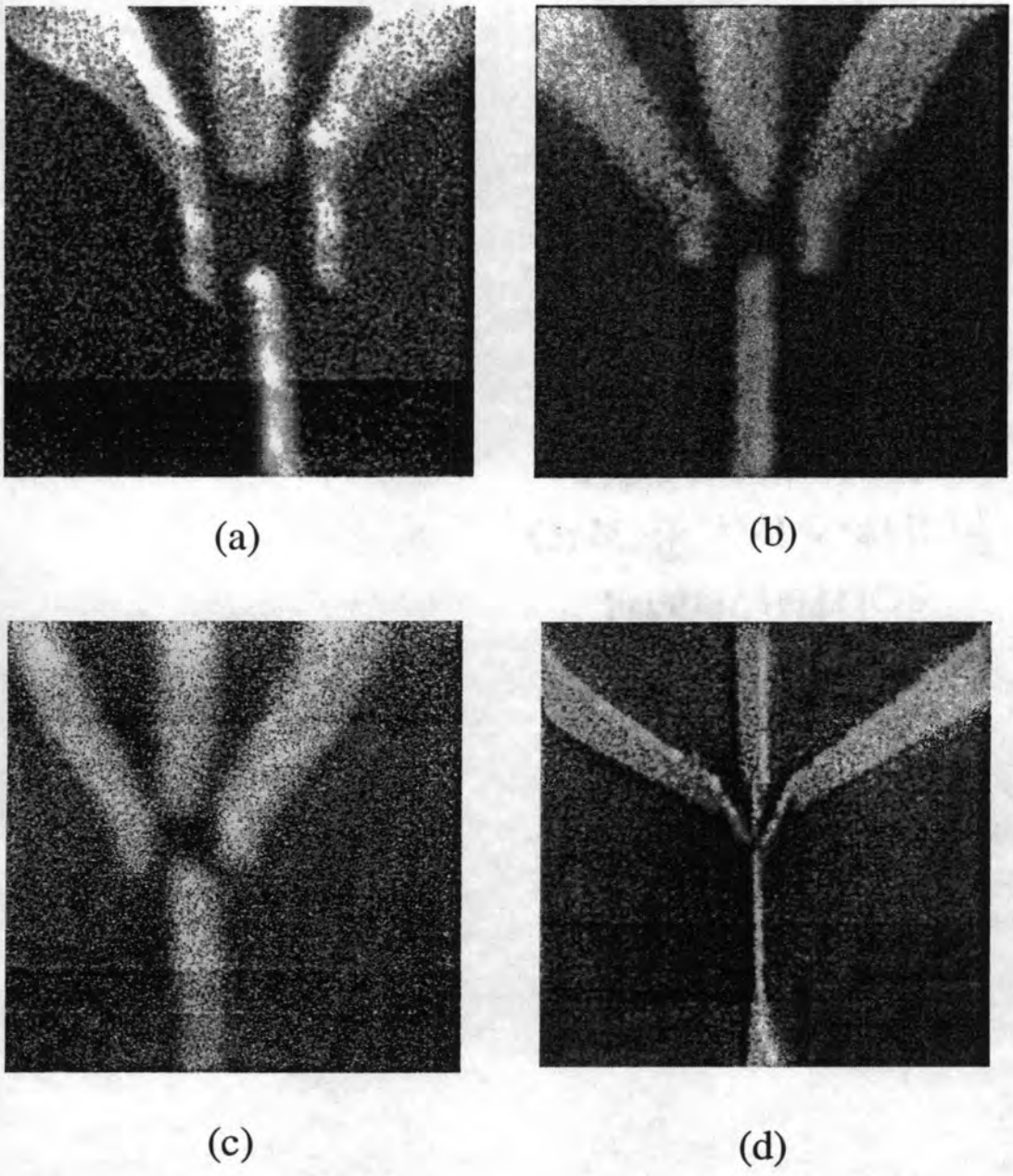

Figure 3.6: SEM micrographs showing: a) a symmetric 500 nm dot (D10), b) a asymmetric 100 nm × 200 nm dot (D11), c), a symmetric 200 nm dot (D12), and d) lower magnification of the 200 nm dot.

3.2.8 Second Gate Oxide

Immediately after the liftoff of the depletion gate metal, the upper gate oxide is deposited by remote plasma enhanced chemical vapor deposition (RPECVD). As metals are already present in the structure, thermal loads must be minimized to eliminate cross contamination by metal migration. This precaution is also necessary to minimize the redistribution of dopant species by diffusion. RPECVD allows a reduced temperature of 350 C in this process. Besides the electron beam lithography, this step required the most time to develop.

The sample is first thoroughly cleaned and dried, and placed in the loading chamber of the PlasmaQuest CVD system, which is a downstream reactor design. Here it is allowed to further dry in a flowing N_2 gas ambient for 5 min. A cantilever exchange arm loads the sample, and the reactor chamber is pumped down to as low a pressure as is achievable for the system. It is critical that generous time is allowed to reach the process temperature, as the oxide performance is highly sensitive to temperature. A He purge is performed, and then a stable flow of the process gases He and N_2O is achieved. Now the RF power is introduced, and the plasma is tuned to a maximal value in the minimum time. Finally, silane (SiH_4) is introduced, the plasma re-tuned, and the deposition is timed based on a previous calibration of deposition rate. Fig. 3.7 follows the pressure, gas flow rates, and RF power as a function of time for the process. Note that the times for the steps where the plasma are tuned (steps 4 and 5) are kept to the minimum time required for the tuning of the plasma to be achieved.

Initially, the thickness of the deposited oxide was 50 nm, but this was increased in the third revision of the process to 90 nm to allow for more planarization of the active region, reducing edge-induced electric breakdown.

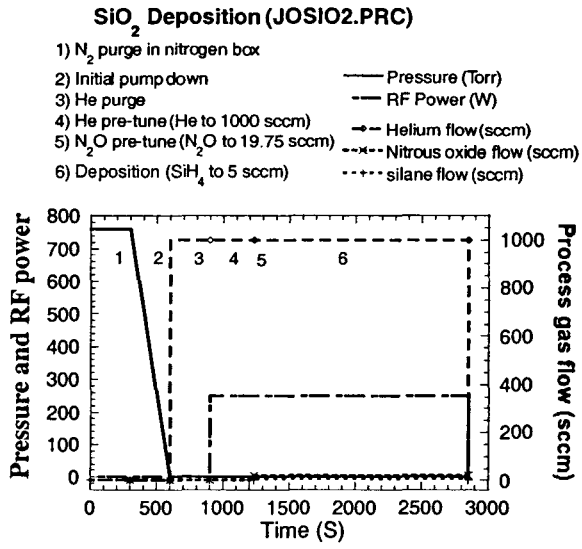

Figure 3.7: A flow diagram showing the process variables for the RPECVD second gate oxide step.

This process does not have any analog in the modern MOS process, although PECVD oxides are commonly used as passivation layers or inter-metal dielectrics in multi-metal processes.

3.2.9 Second Via Level

The second via level goes through the RPECVD oxide to make contact to all underlying electrical contact regions. As the etch rate of the RPECVD oxide is much higher than for thermal oxide, it must be determined to minimize the the undercutting by lateral etching. After it has been determined, the process is standard: pattern the wafers using OCG 825 PPR in the Karl Suss contact aligner with a 30 s exposure through the second via level mask. Develop for approximately 30 s and rinse thoroughly in running DI, then blow dry with dry N_2 gas. Post bake for 12 min at 130 C. Etch in BOE for sufficient time to completely etch through the deposited oxide, and rinse in running DI. Strip the photoresist, clean and blow

dry. Note that in a modern MOS process in the sub-micron range, vias are typically etched using a reactive ion etching process (RIE), resulting in very steep sidewalls. The vias are then plugged with CVD-deposited W to ensure good step coverage.

3.2.10 Final Metalization

This step defines the inversion gate, metal interconnects, and makes contact to previously defined contacts as well as forming interconnects and bonding pads. The basic steps are:

Pattern the sample in the Karl Suss aligner using OCG 825 PPR and the final metal mask, develop, dry, and insert into the vacuum jar of the electron beam evaporator. Deposit 350 nm of Al, remove from the evaporator, and liftoff.

In the fourth revision, the Al metal was replaced by Si doped Al (Al:Si 3%) to prevent spiking into the Si Ohmic contacts. In the fifth revision, a major change in the processing occurred due to the break down of the TorrVac evaporator. The CHA requires the use of a planetary multiple wafer mounting fixture which rotates the sample through all angles to ensure step coverage. Unfortunately, this makes lift off difficult, as the solvent can not attack the underlying photoresist. To alleviate this, a new photoresist was selected because it was possible to spin a much thicker resist film. This also allowed for deposition of a much thicker Al layer, which helped in the bonding process further down in the fabrication sequence. The revised process steps are: spin on HMDS for 30 s at 4000 rpm and immediately spin AZ4210 PPR at 4000 rpm for 30 s. Bake the sample on an 85 C hot plate for 1.5 min, expose in the Karl Suss aligner for 30 s, and develop for 20 s in AZ 400K developer diluted 3:1 H_2O:AZ 400K. Evaporate 450 nm Al:si in the electron beam evaporator and lift off.

In the sixth revision, a stack of Cr and Si doped Al was used in the thicknesses of 30 nm and 400 nm, respectively. Also, the hot plate bake was replaced by an oven bake at 85 C.

In a standard MOSFET process, the gate would be formed from a uniformly deposited poly-Si, which is defined by etching away everything but the gate. The source and drain would be implanted at the same time the gate is doped using a self-aligned gate process. In a standard self aligned process, the gate is used as an implant mask for the channel, aligning the gate to the source and drain Ohmic contacts through windows etched through the poly-Si and other existing layers.

3.2.11 Contact Sinter

The final step involving all the devices as a whole is the Al sinter step, which is done to improve the metal conductance and to hydrogen-passivate the dangling Si bonds at the Si-SiO_2 interface.

The steps are: Perform an *AMD clean*, and very thoroughly dry sample first by blowing with dry nitrogen gas, and then baking at 110 C in an oven. Load the sample into the center chamber of the annealing furnace heated to 425 C under forming gas ambient. Sinter the sample for 12 min and remove from the annealing furnace tube.

These process steps are shown pictographically in the figures on the following three pages, in which both the significant topographical and layer structure changes which occur with each process step. To give some impression of the scale, the width of the gate is 20 μm. In the close-up, (Fig. 3.13) the depletion gates are just barely visible through the oxide and the inversion gate.

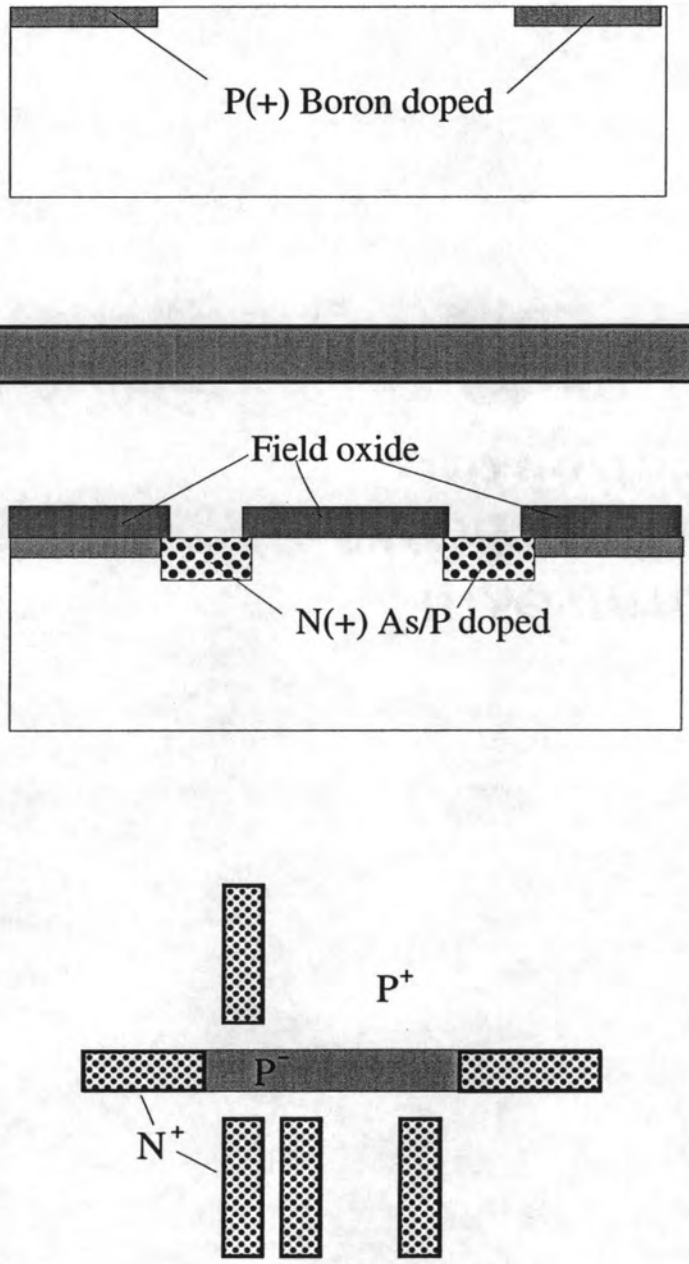

Figure 3.8: Process steps 1 (top two) and 2 (bottom two)

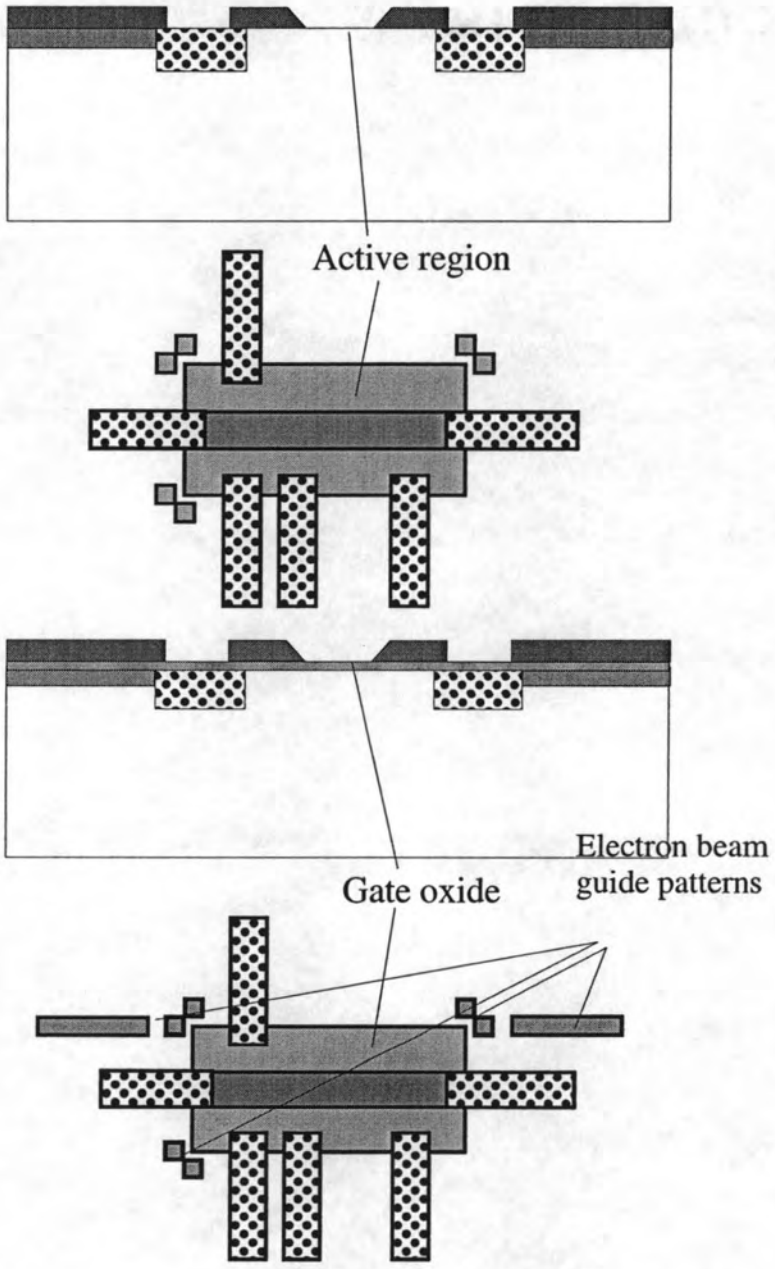

Figure 3.9: Process steps 4 (top two) and 5 (bottom two)

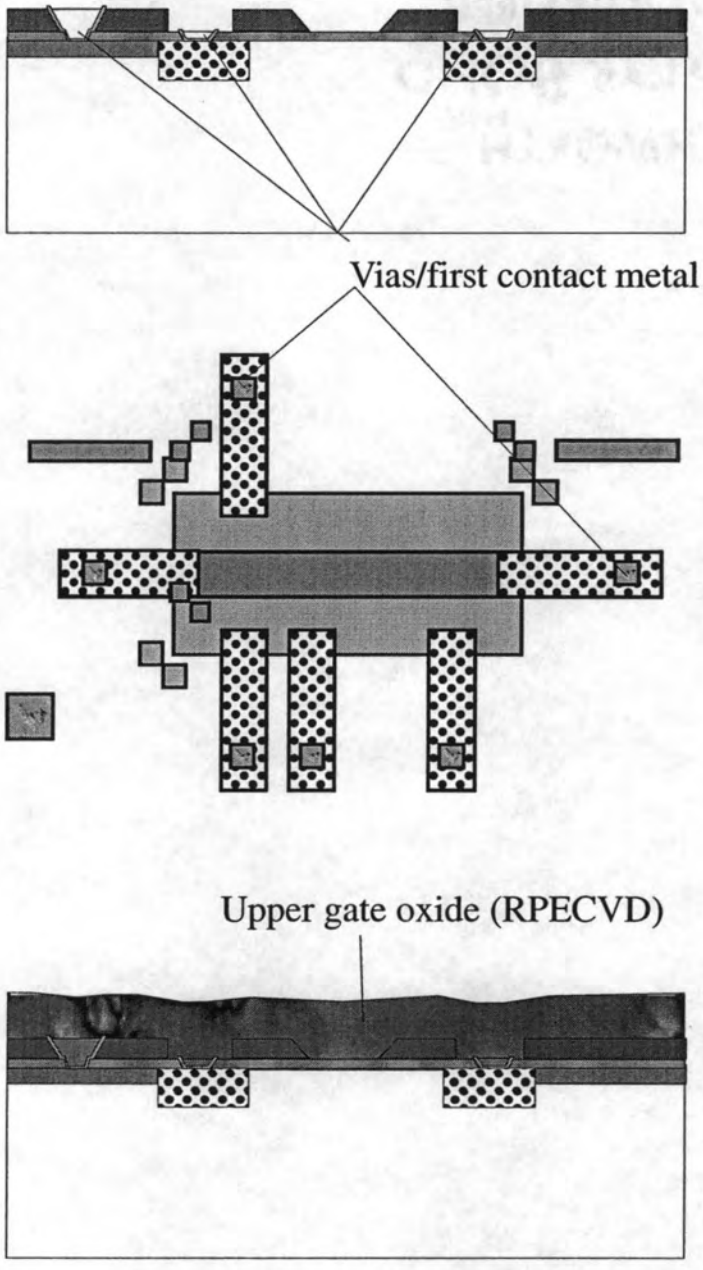

Figure 3.10: Process steps 6 (top two) and 8 (bottom)

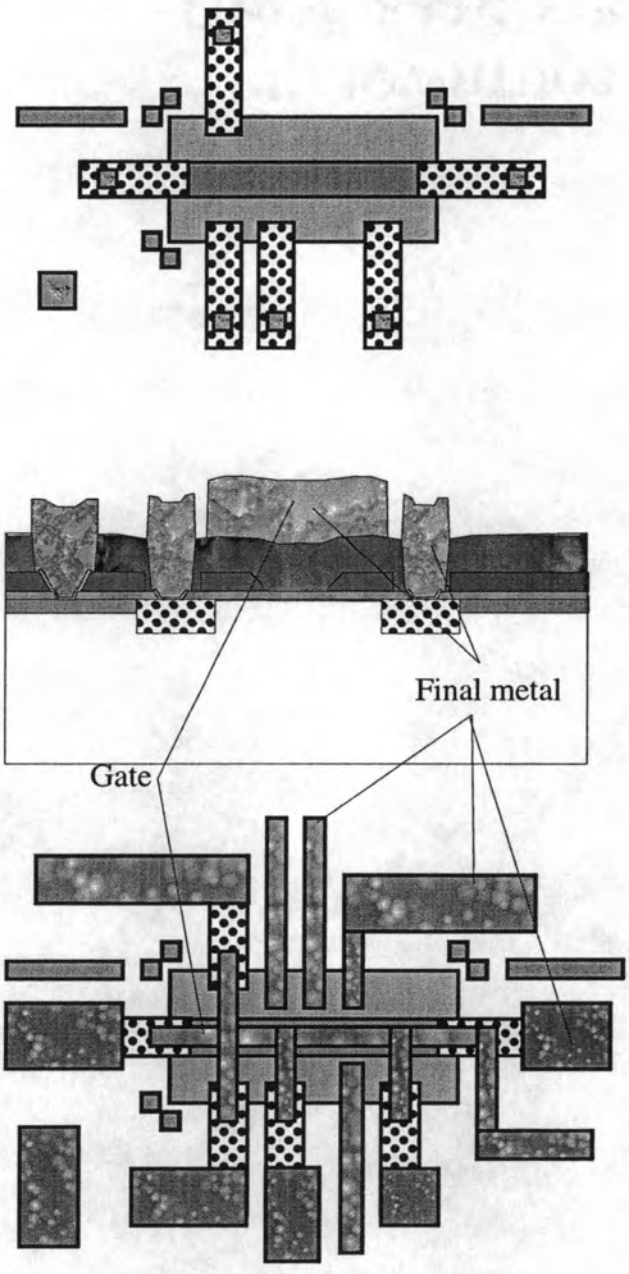

Figure 3.11: Process steps 10 (top) and 11 (bottom two)

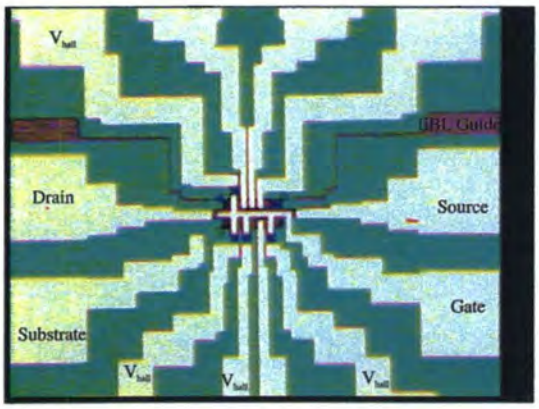

Figure 3.12: An optical micrograph showing all of the features of a complete device. At this scale (200X magnification), it is not possible to distinguish the depletion gates.

All that remains in the device fabrication is to break the chips into individual devices, mount them in header packages that are suitable for cryostatic measurements, and bond the electrical leads on the package to the leads on the device. Optical micrographs of the devices to this point are shown in Figs. 3.12 and 3.13.

3.2.12 Cleaving Chip

The chip is inspected under the optical microscope to determine which individual devices will be tested. After selecting the desired devices, they are extracted from the chip by cleaving the chip. This is done by applying pressure along the cleavage planes ($\langle 100 \rangle$ directions) with a diamond scribe.

3.2.13 Mounting Devices

An *AMD clean* is performed on the selected devices in ultrasonic baths, they are dried, and placed into 16 pin DIP packages using old, thick photoresist as an adhesive. They are then baked at 170 C for 20 min, and checked to make sure that they are firmly mounted.

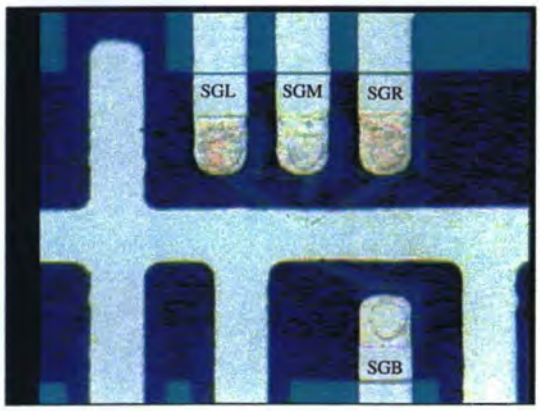

Figure 3.13: A close-up showing the active region of the device. At this scale (2000 magnification), the depletion gates are barely visible as shadows in the inversion gate and the gate oxide.

3.2.14 Ball-Bonding

Last, the devices are bonded to the lead frame packages in an thermosonic wire bonder using Au wire. The packages are heated to a temperature of 180 C, and ultrasonic power is supplied. At this point, electrostatic precautions are very important to the electrical survival of the devices; both a grounding strap and electrostatic-protective gloves are used whenever it is possible to make electrical paths between the handler and the device.

3.3 Cryogenic Setup

Low temperatures are required to observe the phenomena of interest in these nanostructure systems. At higher temperatures, inelastic scattering of electrons with phonons reduces the phase coherence length to the point where quantum interference effects are no longer visible. Coulomb blockade effects also require that $e^2/2C \gg k_B T$. Kinetic energy may be imparted to electrons in the collisions making confinement effects negligible. Cooling to low temperatures makes the phonon modes available in a thermal equilibrium distribution small enough to reduce the

effects of the inelastic scattering and thus allows the transport effects to be seen. Several references are available for review of cryogenic systems [158, 159, 160]. In the following, a description of the cryogenic systems used for the experiments reported in Chapter 4 is given. Three different cryogenic systems were used in this research briefly discussed here.

3.3.1 Introduction to Cryogenics

Cryogenic systems utilize liquefied inert gases to cool specimens to low temperature. N_2 gas liquefies at 77 K at atmospheric pressure. By pumping on the vapor above the liquid, it is possible to lower the temperature, but violent bubbling and freezing of the liquid in simple systems makes it an inpractical working fluid for temperatures less than 77 K.

The isotope ^{4}He liquefies at 4.2 K, and by pumping on the vapor above the L^{4}He it is possible to reduce the temperature of evaporation to approximately 1.2 K. For temperatures between 4.2 K and 77 K, an arrangement can be used which allows heated ^{4}He vapor to pass by the sample and maintain it at intermediate temperatures. It is possible to reach 300 mK by pumping on the isotope L^{3}He, and below 300 mK, dilution refrigeration is required.

3.3.2 L^{4}He Transport Dewar Cryostat

The simplest cryogenic arrangement used was to insert a dipstick into a 100 l liquid He transport dewar. This system allowed for low noise and simple grounding of the cryostat to the instrumentation by use of a grounding strap. The insert for this arrangement consisted of a hollow tube through which a bundle of insulated copper wires led to an 18-pin DIP socket. Coaxial contacts provided electrical connection to the instrumentation.

3.3.3 Janis Research Cryostat

For early measurements, a Janis Research cryostat was used. The Janis cryostat uses a liquid nitrogen jacket as the outer insulation and the sample chamber is isolated from the He bath by an inner vacuum jacket. Liquid He may be introduced to the sample chamber by means of a needle valve, and the temperature may be reduced below 4.2 K by pumping continuously with a small mechanical pump or increased by heating the vapor which passes through the sample space with a resistor and monitored and controlled with a carbon resistor bridge and a Lakeshore cryogenics controller. Magnetoconductance measurements are possible with a 9 T superconducting magnet. The chronic problem with this cryogenic setup was a bad grounding arrangement for the leads in the time between connection to the insert and connection for measurement. This problem almost always lead to destruction of devices before it was possible to get any reasonable transport data. One notable exception was the device M1, which was a simple MOSFET used for calibration of the 2DEG properties.

3.3.4 American Magnetics Cryostat

The American Magnetics cryostat was equipped with a 9 T superconducting magnet with a lambda plate (allowing for magnetic fields up to 11 T) and a superinsulated outer shell. Two insert configurations were used: The dipstick of the transport dewar cryostat arrangement, and a L^3He insert made by RMC systems (a now defunct corporation). For the purposes of this research project, the L^3He insert was used as a sealed cryostat. A rough sketch of the American Magnetics cryostat is given in Fig. 3.14.

The RMC systems cryostat consisted of a vacuum can housing a ^4He-pot connected to the He bath by a continuous fill impedance and to an external mechanical pump through a hermetically sealed tube which ran the length of a sliding seal to

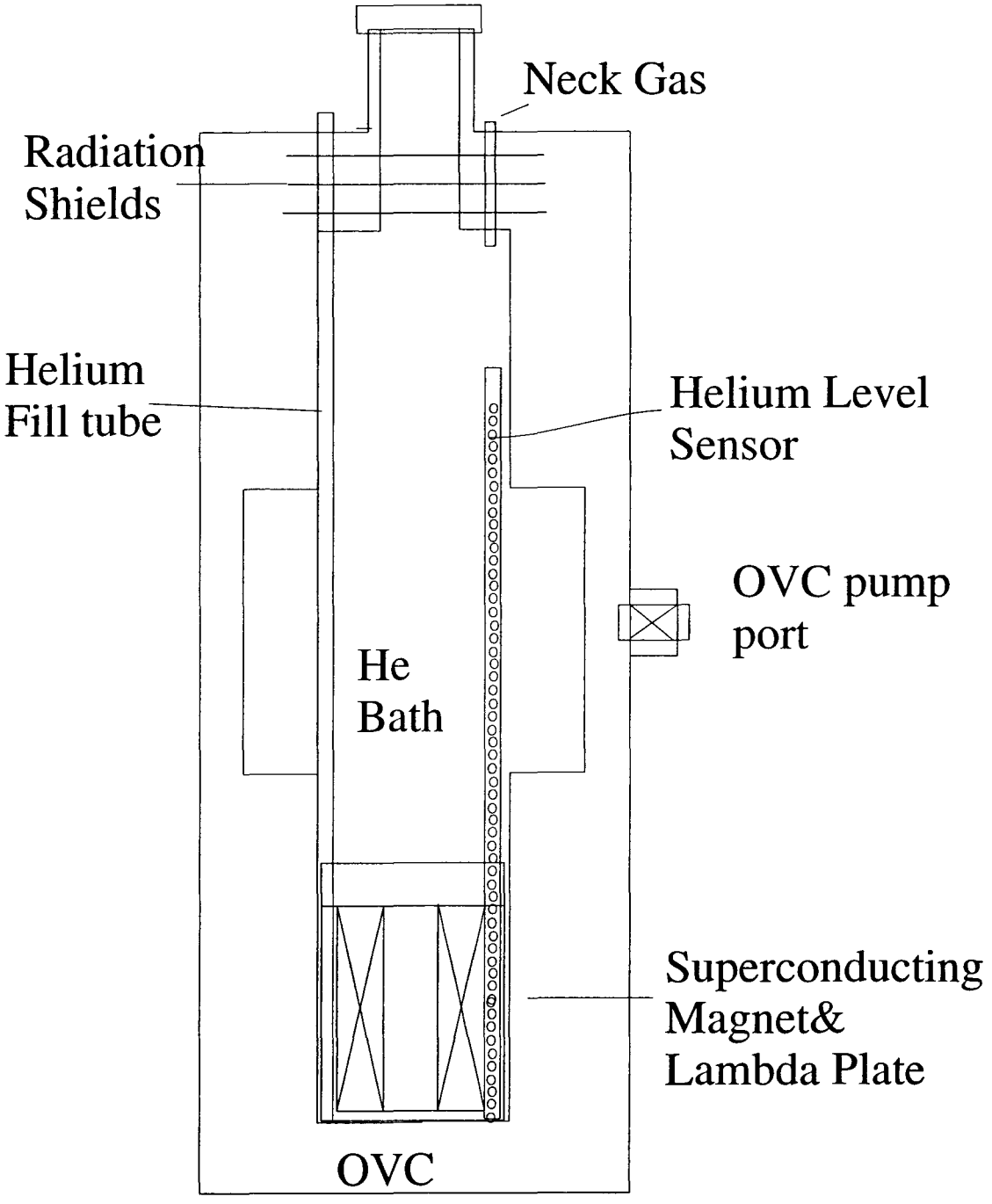

Figure 3.14: A sketch of the American Magnetics cryostat.

the pump port. The signal and control wiring bundles also ran the length of the sliding seal tube. These bundles consisted of insulated 30 gauge twisted pair copper wire. A ^{3}He-pot was connected to the ^{4}He-pot, and this was thermally connected to the sample stage and heater assembly. The sample stage was made of Cu, and housed both a calibrated carbon resistor and a calibrated Si diode for temperature measurement. The signal and control wire bundle were anchored at the vacuum can, at the ^{4}He-pot and ^{3}He-pot. This length of wiring consisted of twisted pairs of constantine wire.

3.4 Experimental Measurement

In this section, the electrical measurement setups will be briefly described. Simplified schematic diagrams will illustrate the basic configuration of the external electrical connections for the current carrying nodes.

3.4.1 Instrumentation

A wide variety of electrical measurement instruments was used in acquiring data during experiments. There were two classes of measurement: near-equilibrium and non-equilibrium; nearly all measurements were of the first kind. All DC measurements used a Hewlett Packard 4156b semiconductor parameter analyzer. AC measurements used lock in amplifiers as both reference source and measurement instruments. Lock in amplifiers used included the Stanford Research 830, and the EG&G 7265 and 5210. Additional voltage sources were provided by the Keithley 1000 voltage source and a Keithley 236 source measure unit, while a Hewlett Packard 2000 digital multimeter was used to measure magnetic fields.

3.4.2 Magnetoconductance Measurements

Three types of magnetoconductance measurements were performed: AC near-equilibrium, DC near-equilibrium, and AC non-equilibrium. Most of these measurements were done in order to characterize the 2DEG density and mobility.

3.4.2.1 AC Near-Equilibrium

Near-equilibrium AC magnetoconductance measurements were performed by a current-biased AC lock-in technique. Early experiments were done in the Janis Research cryostat with a Stanford Research SR 830 lock-in amplifier (LIA) providing the excitation voltage. The gate was held at a constant bias by using a Keithley 236 source measure unit (SMU). In order to divide out line noise, which was damaging to the gate, a 5:1 voltage divider was used. A schematic of the current biased lock-in circuit is given in Fig. 3.15. The magnetic field was determined by measuring the voltage dropped across a shunt resistor. All voltages were acquired from a DAQ card. LabVIEW software controlled the measurements over a GPIB interface.

Later experiments (in the American Magnetics cryostat) used a EG&G lock-in amplifier, with the voltage across the shunt being measured with a Keithley 2000 digital multimeter.

3.4.2.2 AC Non-Equilibrium

An experiment was performed to purposely force the 2DEG system out of equilibrium. This was done using a voltage-biased drain bias excitation, in which a DC voltage is placed in series with a small AC signal. The DC signal was varied between 200 μV and 2 mV, ensuring that the bias was higher than $4k_BT$, the expected energy of the thermal fluctuations of the lattice. A schematic of this setup is also shown in Fig. 3.15.

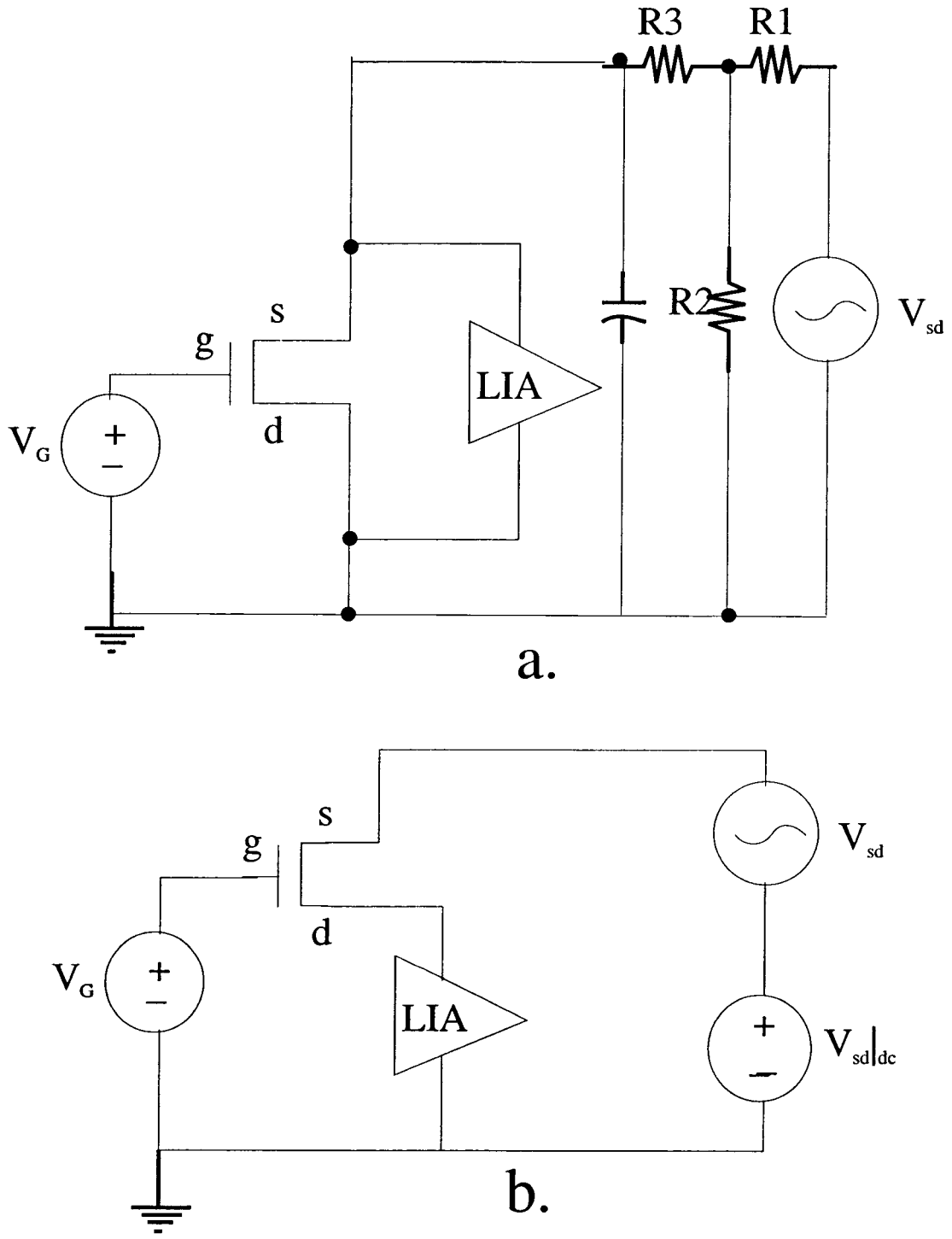

Figure 3.15: Schematics of the circuits used in the AC (a) near- and (b) non-equilibrium measurements.

3.4.2.3 DC Equilibrium Magnetoconductance

DC magnetoconductance measurements were performed using a Hewlett-Packard 4156b semiconductor parameter analyzer. The measurement uses a voltage-biased circuit, shown in Fig. 3.16.

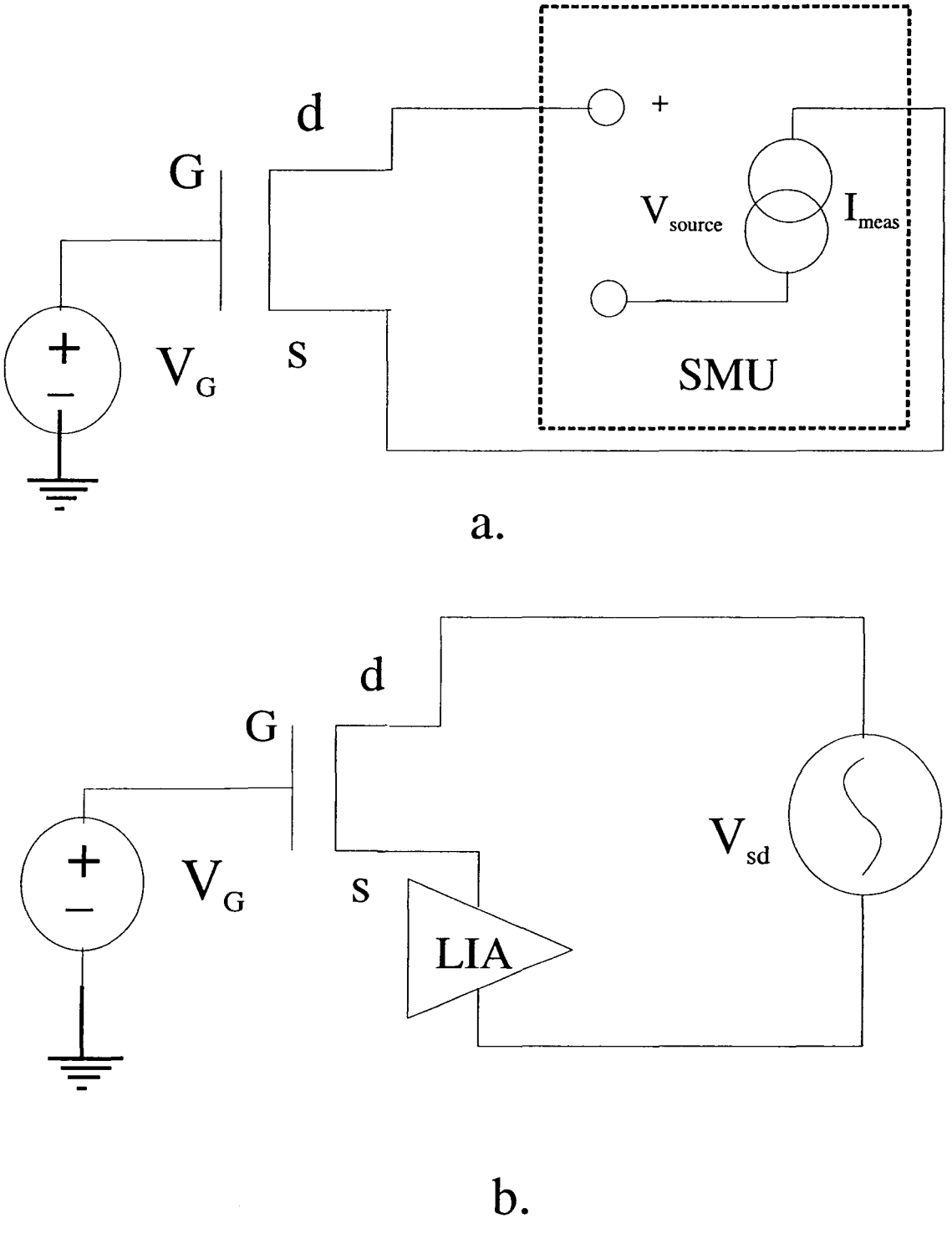

Figure 3.16: A schematic for measurement circuits used in a) voltage-biased DC near-equilibrium transport measurements, and b) voltage-biased AC near-equilibrium transport measurements.

4. EXPERIMENTAL RESULTS

4.1 Characterization of MOSFET Structure Without Dot Depletion Gates

Prior to transport measurements on the quantum dot structures themselves, the properties of the 2DEG were determined in the absence of depletion gates to establish a baseline reference point using devices M1 and D11. Device M1 is a MOSFET without additional depletion gates to define the quantum dot; its lower gate oxide is $10\,nm$. Device D11 is full quantum dot structure with the symmetric 200×200 nm geometry. The lower oxide thickness is $\approx 5\,nm$. Oscillatory magnetoconductance measurements were performed to determine the sheet charge density as a function of inversion gate bias and estimate the mobility. Also, the current as a function of gate bias was measured to establish the threshold voltage and transconductance characteristics of the transistor structure itself.

4.1.1 MAGNETOCONDUCTANCE

Figure 4.1 shows the magnetoresistance plotted for various inversion gate biases as a function of the magnetic field normal to the $Si - Si0_2$ interfacial plane taken with current-biased AC lock-in measurements. A schematic of the Hall bar configuration used in these measurements is shown in Fig. 4.2. At low magnetic fields, the longitudinal magnetoresistance is measured through the full MOSFET structure using the source-drain contacts to source current and measuring the voltage drop across the sample using the contacts labeled V_{xx} and V_{xy}, as was done in the case of sample M1 with a four-terminal method. It may also be measured in the full quantum dot structure by measuring two terminal between V_{yy} and V_{xy}, as was done in device D11. An example of Shubnikov-de Haas (SdH) oscillations in the longitudinal magnetoresistance at 4.2 K plotted against $1/B$ is shown in Fig. 4.3.

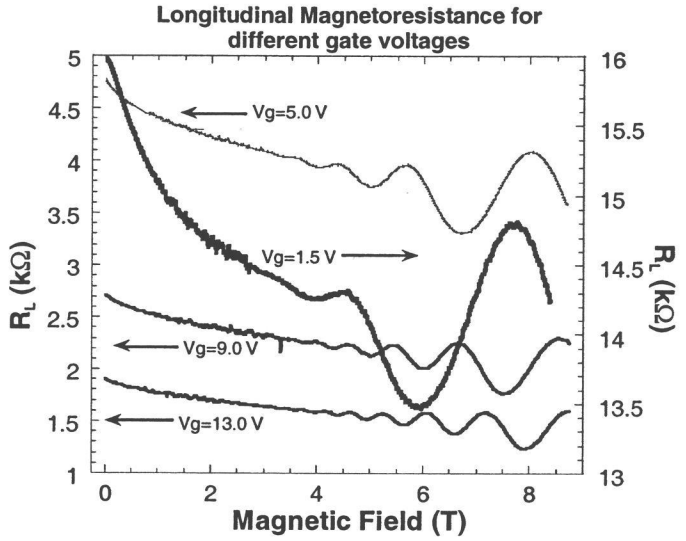

Figure 4.1: Magnetoresistance plotted as a function of normal magnetic field for various values of the inversion gate bias

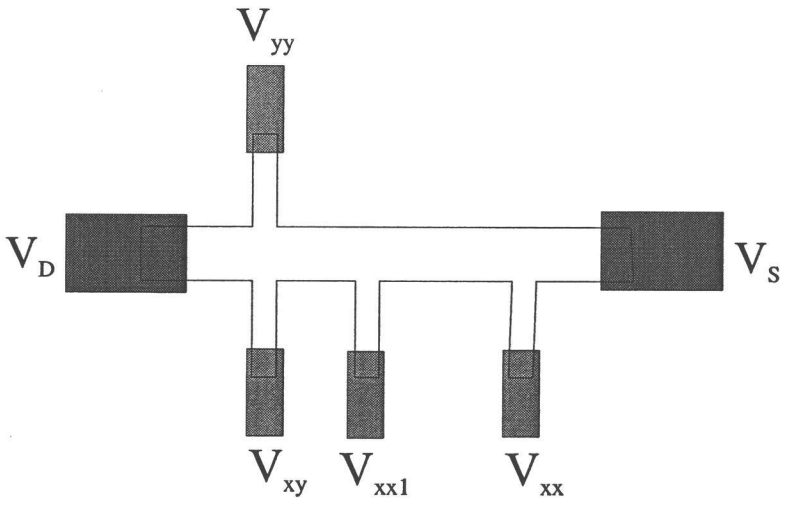

Figure 4.2: A schematic of the Hall bar configuration used in the transport measurements. The channel is the region subtended by the inversion gate, which allows for additional fingers connecting the voltage probes for four terminal and Hall measurements. Ohmic contacts are the gray regions.

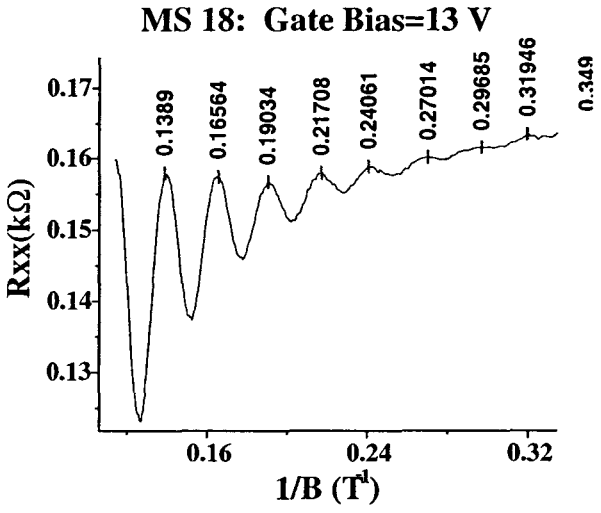

Figure 4.3: Longitudinal magnetoconductance oscillations in sample D11 for gate bias of 13V. The sheet density for this bias is approximately $n_{2D} \approx 3.5 \times 10^{12}/\text{cm}^2$

Using Eq. 2.14, the 2DEG sheet density for the average periodicity shown in Fig. 4.3 gives a density of approximately $3.5 \times 10^{12}/cm^2$, where the degeneracy factors (g_s, g_v) have been taken to be 2, meaning that the electrons are in the electron quantum limit.

4.1.2 Non-Equilibrium Magnetotransport

In order to study non-equilibrium effects in the transport properties of the material, a DC bias voltage was supplied in series with the AC voltage , and the AC response was measured. The intent was to observe carrier heating effects on the mobility and possibly see carriers populating the Δ_4 ladder of subbands. The results of this experiment are shown in Fig. 4.4, where the only observable effect is greater zero field longitudinal resistance for an inversion gate bias of 2 V and drain bias of 2 mV, which is due to lower mobility.

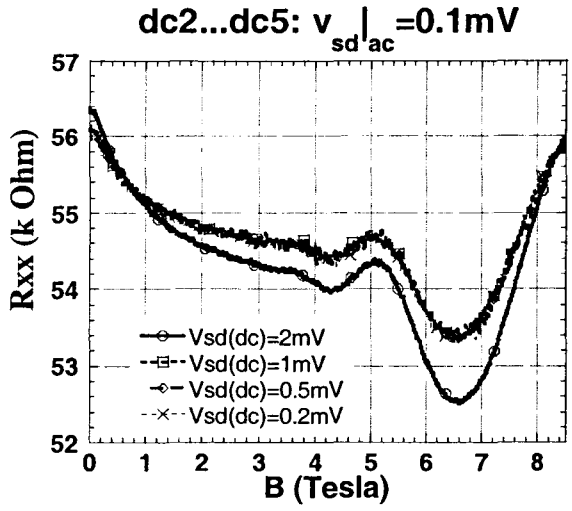

Figure 4.4: Semi-Equilibrium magnetoconductance measurement showing a slight increase in the zero-field resistance and oscillation amplitude at gate bias of $2\,V$.

4.1.3 Gate Response Characterization

In order to understand the transport through the quantum dot structure, it is necessary to understand the transport properties of the 2DEG channel which connects the dot to the reservoirs. Of most interest is the response to the inversion gate bias of MOSFETs. For a thorough discussion of the fundamental operating characteristics, the reader is referred to the text by S. M. Sze [161]. In an ideal MOSFET, there are two basic regions of operation: the linear region and the saturation region. For small drain bias and gate bias above the threshold voltage V_{T}, the current is characterized by a linear increase from zero bias, following the relation

$$I_D \cong \left(\frac{W}{L}\right) \mu_{\mathrm{eff}} C_{\mathrm{ox}} (V_{GS} - V_{\mathrm{T}}) V_{DS} \quad \text{for } V_{DS} \ll V_{GS} - V_{\mathrm{T}} \tag{4.1}$$

In this region of operation, the differential conductance, g_D, and transconductance, g_m, are given by the equations

$$g_D = \left.\frac{\partial I_D}{\partial V_{DS}}\right|_{V_{GS}=const.} = \frac{W}{L}\mu_{\text{eff}}C_{\text{ox}}(V_{GS} - V_{\text{T}}), \tag{4.2}$$

$$g_m = \left.\frac{\partial I_D}{\partial V_{GS}}\right|_{V_{DS}=const.} = \frac{W}{L}\mu_{\text{FE}}C_{\text{ox}}V_{DS}, \tag{4.3}$$

where $W$and L are the channel width and length, respectively. The differential conductance and transconductance give a method for solving for the so-called effective mobility (μ_{eff}) and field effect mobility (μ_{FE}), which differ in that the effective mobility is generally greater. Below V_{t}, the current has very weak dependence on the drain bias (ideally none) and a nearly exponential dependence on the gate bias in weak inversion.

As the drain bias is increased, the depletion region associated with the p-n junction at the drain increases. At pinch off, this depletion layer pinches off the inversion layer, and increasing the drain bias results in approximately no increase in the current. The saturation voltage is approximately given by $V_{DS\text{Sat}} \approx V_{GS} - V_{\text{T}}$. In the saturation region, the current is ideally given by

$$I_{D\text{Sat}} \cong \left(\frac{mW}{L}\right)\mu_{\text{eff}}C_{\text{ox}}(V_{GS} - V_{\text{T}})^2, \tag{4.4}$$

where m is a dimensionless constant whose value is a function of the doping concentration. The transconductance in saturation is given by

$$g_m = \frac{2mW}{L}\mu_{\text{FE}}C_{\text{ox}}(V_{GS} - V_{\text{T}}). \tag{4.5}$$

As in the inversion gate characteristics of a MOSFET without depletion gates were investigated, with measurements both by the AC the DC methods. Figure 4.5 shows the gate voltage dependence of the source-drain current and conductance for device M1. From this measurement the threshold voltage is determined by interpolating the linear portion of the current to zero through the point of maximum

Table 4.1: The threshold voltage for device M1 for temperatures of 300, 77, and 4.2 K

Temperature (K)	300	77	4.2
Threshold (V)	-0.25	0.5	1.5

g_m. The zero intercept gives the gate voltage at the onset of strong inversion, defined in terms of the surface potential as the point where $\psi_{\text{surf}} = 2\psi_{\text{B}}$, where ψ_{B} is the bulk potential defined as the difference of the bulk Fermi energy from the intrinsic Fermi energy. Also seen in the figure is the onset of weak inversion as indicated by the rapid increase in the voltage drop across the channel, around $1.5\,V$. The area of weak inversion, which is roughly the range of gate bias where the surface potential is $\psi_{\text{B}} \leq \psi_{\text{surf}} \leq 2\psi_{\text{B}}$ is shown in Fig. 4.6 shows a nearly exponential current increase with gate bias. The threshold voltages for various temperatures are given in Table 4.1.

A striking point to notice about Fig. 4.5 is that the conductance in this range of measurement is far below the first quantum of conductance $2e^2/h \approx 77.5\,\mu\text{S}$. In fact, due to the geometry and low mobility of the sample, it was not possible within the working range of the gate bias to achieve such a base conductance. This was the case with all devices fabricated during the project.

4.1.4 Drain Response Characteristic

The response of the MOSFET to finite drain bias for constant inversion gate bias gives important information about sub-threshold current, Ohmic contacts, and

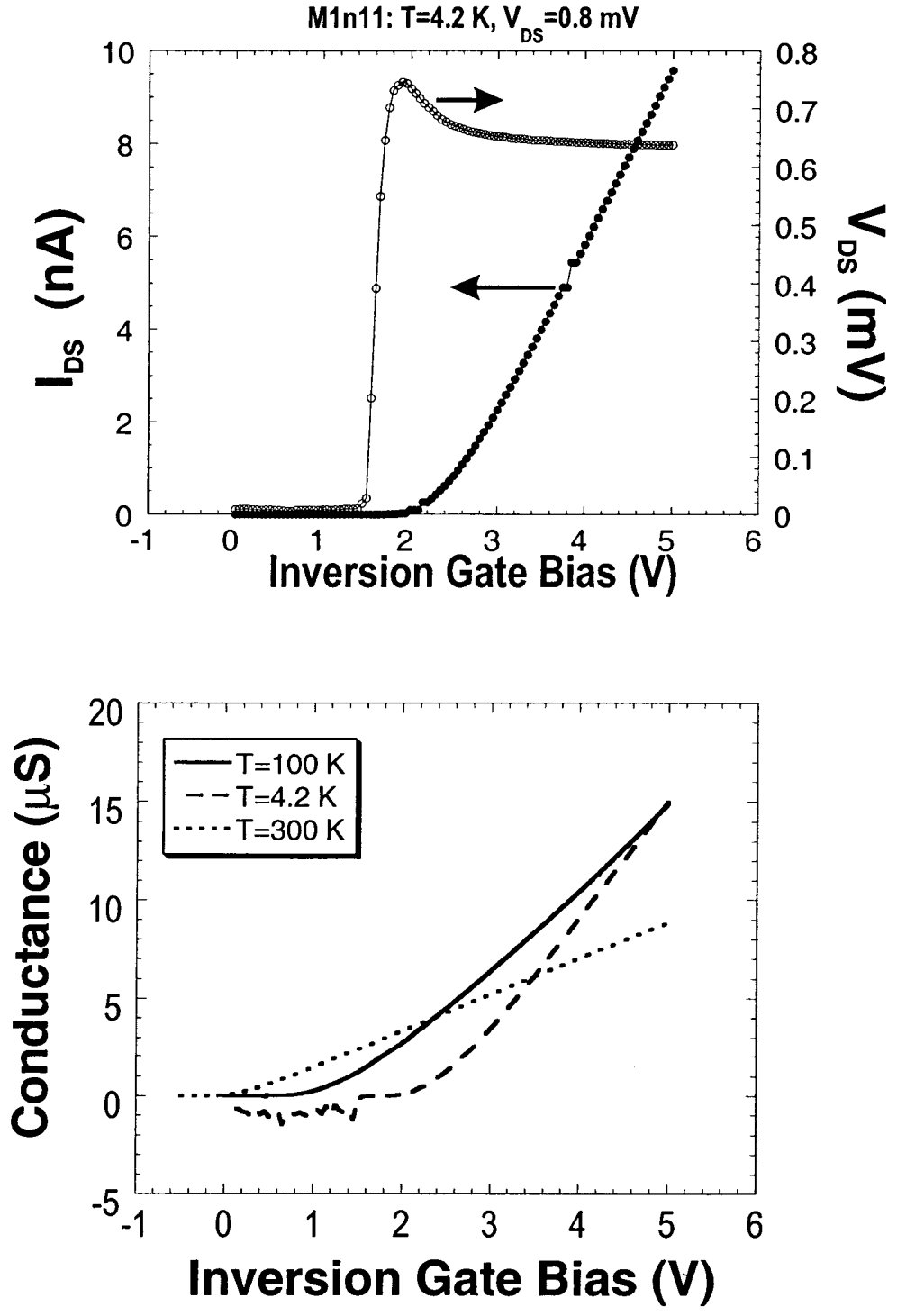

Figure 4.5: Top: four-terminal gate characteristic of device M1 at 4.2 K. Bottom: Conductance for device M1 measured by using an AC 4 terminal technique for drain excitation $\approx 2 \times k_B T$.

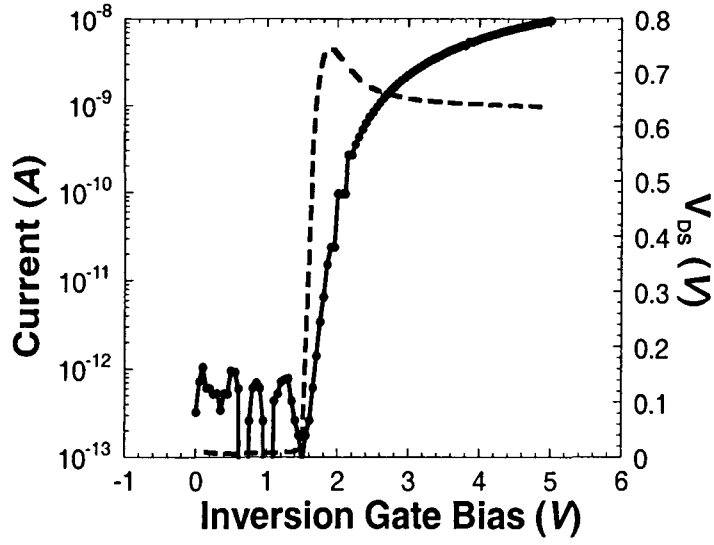

Figure 4.6: The current in the weak inversion regime as a function of gate bias.

device saturation characteristics. Measurements of the drain response characteristic were performed using a DC technique. An example of a good device performance at room temperature is given in Fig. 4.7. This device is operating below saturation for all sweeps except for the second sweep, where the gate bias is at 1.5 V.

4.1.5 Contact Resistance

One of the test structures placed on the die was a TLM structure, used in measuring the resistance of the Ohmic contacts [162] by the transfer length method (TLM). In this method of measuring the specific contact resistance, the resistance is measured between Ohmic contacts at different spacings and plotted as a function of the contact separation. The intercept with zero distance is then twice the contact resistance. From this measurement, it was determined that the contact resistance did not significantly limit the conductance of the device.

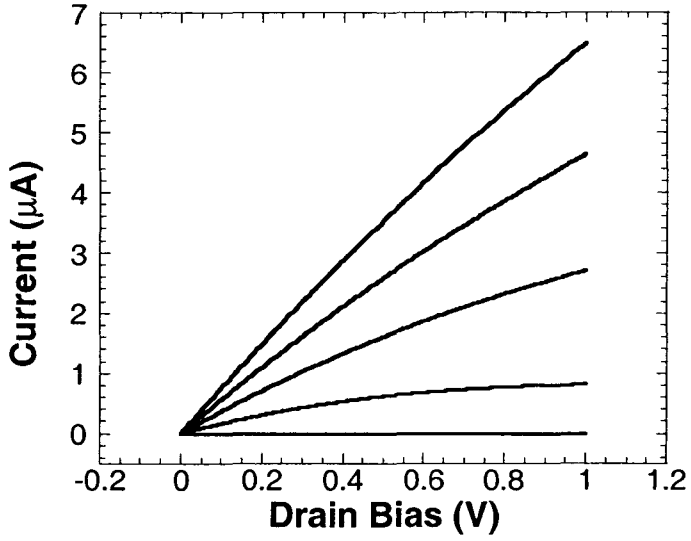

Figure 4.7: Drain characteristic for device M1 at 300 K for gate voltages of 0,1.5, 3, 4.5, and 6 V from bottom to top. The device is operating below saturation.

4.2 D10: 500 nm Symmetric Dot

Device D2 was the first fully functional dot measured with features reportable in this work. The depletion gate geometry is 500 nm×500 nm, which is larger than the other devices reported on later in this work.

4.2.1 Inversion Gate Sweep

The first type of experiment done after confirming that there is functionality in the device gates is to sweep the inversion gate bias for fixed drain and depletion gate biases. In effect, this changes the chemical potential in the leads. As shown in Fig. 4.5, in an ordinary MOSFET, the $I_D - V_{GS}$ curve in the linear region exhibit a smooth turn on followed by a linear increase of drain current past threshold. As shown in Fig. 4.8, peaks in the source-drain current are seen as the inversion gate is swept in device D10 from from below the onset of weak inversion to slightly above the onset of strong inversion. For a range of depletion gate biases, these peaks are

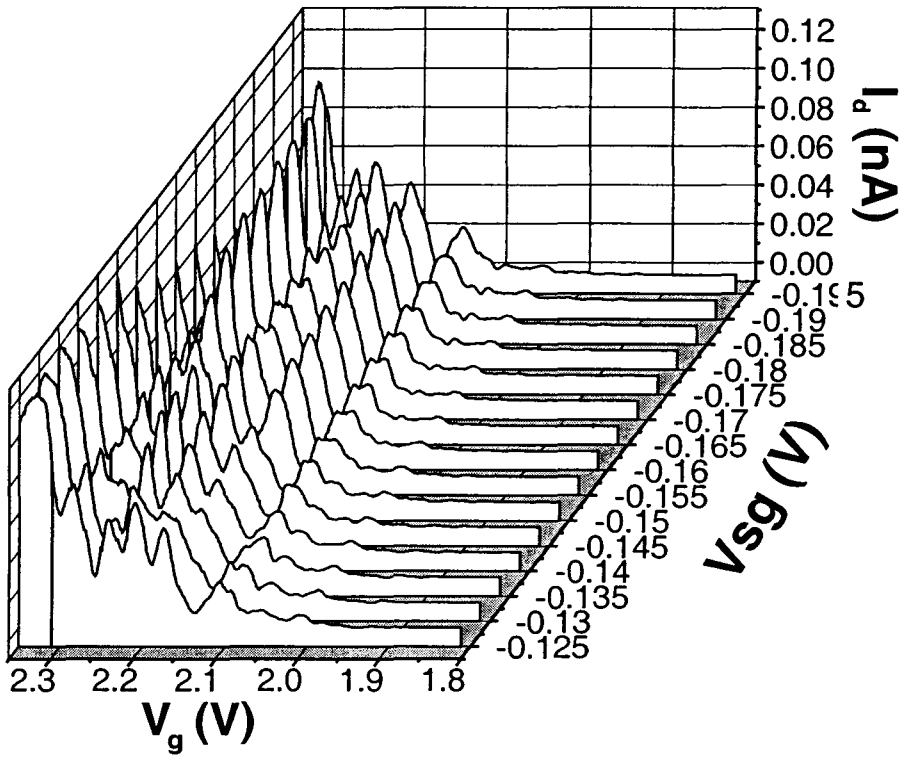

Figure 4.8: A plot of the drain current through device D10 versus inversion gate voltage for a sequence of depletion gate biases. Some splitting and merging of the peaks is observed as the depletion gate bias is changed.

pronounced and evolve in a systematic way as the depletion gate bias is stepped as a parameter. All of the depletion gates are tied to the same bias voltage and stepped from -100 mV to -200 mV. For depletion gate biases less negative than -0.1 V, the behavior of the peaks becomes less ordered. Visible in Fig. 4.8 are some apparent merging and splitting of peaks, which is reminiscent of crossings or anti-crossings in electronic energy levels observed in other reported dot structures, as discussed in Section 2.5.1.

4.2.2 Non-zero Drain Bias

In order to study carrier heating effects on the peaks observed in Fig. 4.8, different source-drain biases were applied to the sample. The range of bias ($\approx$ 2mV) were larger than k_BT at 4.2 K, hence heating is expected. In these measurements, the drain bias is increased in successive increments while the inversion gate bias is swept, as shown in Fig. 4.9 as a gray-scale composite plot. The main effect is that as the amplitude of the current increases with increasing bias, the peak to valley ratio for the most part decreases; eventually, the peaks are broadened and not distinguishable from the background conductance. Another feature is that the minimum current valleys vary in position as the drain bias is changed. It is seen that there is not a large change in the position of the current peaks as the drain bias changes. While there appears to be a large variation as the drain bias goes from 0.1 mV to 0.5 mV, this is likely to be due to the shift from the oxide charging discussed more in Section 4.3. After this initial range, the change is largely linear.

As the drain bias is further increased, the device operates less and less in the linear transport regime. A sequence of such drain sweeps has been plotted in a gray scale semi-log plot of the current magnitude in Fig. 4.10. The main feature is that the position of minimum current has some small variation in position about zero drain bias.

4.3 D9: Symmetric 200 nm Dot

Two symmetric 200 nm devices were fabricated which showed some interesting conductance features. These devices are labeled D9 and D12. D9 is discussed in the present section (4.3) while D12 is discussed separately in Section 4.4.

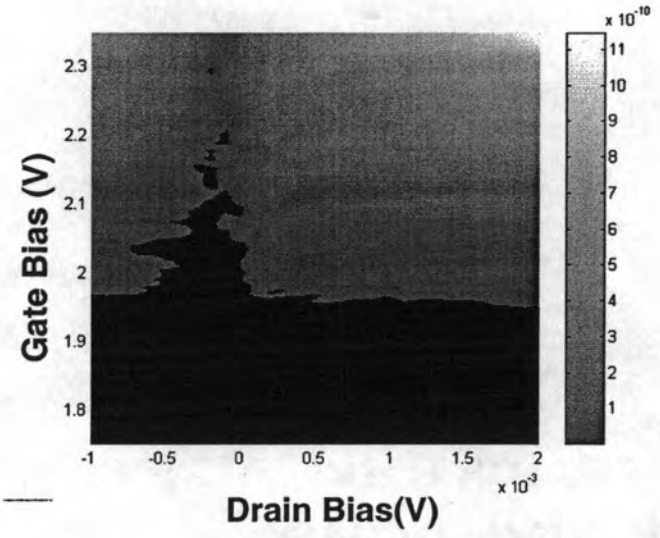

Figure 4.9: A composite plot in the drain bias-inversion gate bias plane for device D10 in which the current is represented by the gray scale color map (dark is lower current).

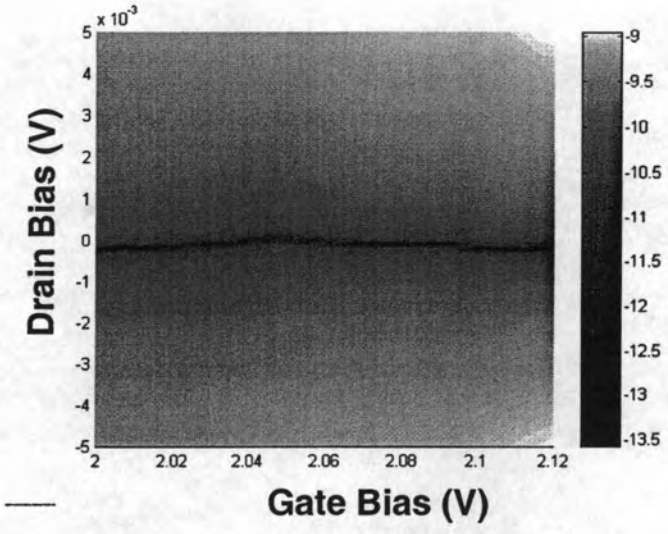

Figure 4.10: Magnitude of I_D through device D10 plotted on a semi-log scale in the $V_{GS} - V_{DS}$ plane. The gray scale represents the magnitude of the current.

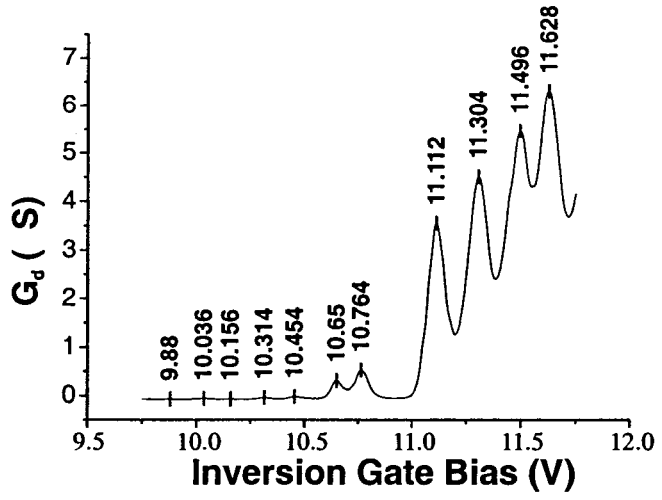

Figure 4.11: An example of an inversion gate sweep for fixed depletion and drain bias. The peaks at have been marked and labeled. The values for gate biases are: $V_{SGR} = -30\,\mathrm{mV}$, $V_{SGL} = 25$ mV, $V_{SGB} = V_{SGM} = 0.0$, $V_{DS} = 100\,\mu\mathrm{V}$.

4.3.1 Inversion Gate Sweep

The symmetric 200 nm dot D9 exhibited oscillatory features as a function of the inversion gate bias. For low source drain bias ($100\,\mu$V), aperiodic peaks in the drain conductance occurred. An example of these peaks is shown in Fig. 4.11, where the peak positions have been picked by a numerical peak selection routine and labeled. Because the background conductance is a strong function of the inversion gate bias, (see, for example Fig. 4.5), peaks at lower gate bias have amplitude below the visible threshold in the plot. These peaks have a shape very similar to the form for Coulomb blockade given by Beenakker.

Multiple sweeps of the inversion gate bias for different bottom depletion gate bias (V_{SGB}) results in the same sequence of peaks, with the position of the peaks changing in inversion gate voltage. These measurements are performed with the sequence:

- The drain bias is set. This value does not change until the end of the measurement.
- The depletion gates are set and held for each individual sweep.
- The initial value of the gate bias is set, and held for a hold time set by the amount of time for which further increases do not change the initial value of a given measurement due to capacitive charging effects associated with the gates.
- The sweep is performed, with a delay between each data point in the sweep run.
- The new depletion gate values are set and the inversion gate bias is reset to the initial value in a given sweep.

To see the effects of changing the bottom depletion gate bias on a larger scale, a composite of each of the individual sweeps has been generated and plotted in a two dimensional contour plot. The color scale represents the conductance amplitude. This composite is shown in Fig. 4.12. The order of the individual sweeps was mixed up for the data in the image, and this is manifested in the shifting of the positions of the ridges going laterally across the plot; each start of a new series of sweeps had its own offset, which resulted in a shift in the position of the ridges associated with that sweep.

4.3.2 Bottom Depletion Gate Sweep

Since the position of the peaks change as a function of the depletion gate bias, it should be possible to see oscillations in the conductance if the inversion gate bias is maintained at constant value and the depletion gates are swept independently. This is indeed the case, as shown in Fig. 4.13, where oscillations are seen on a linear scale. It is observed that

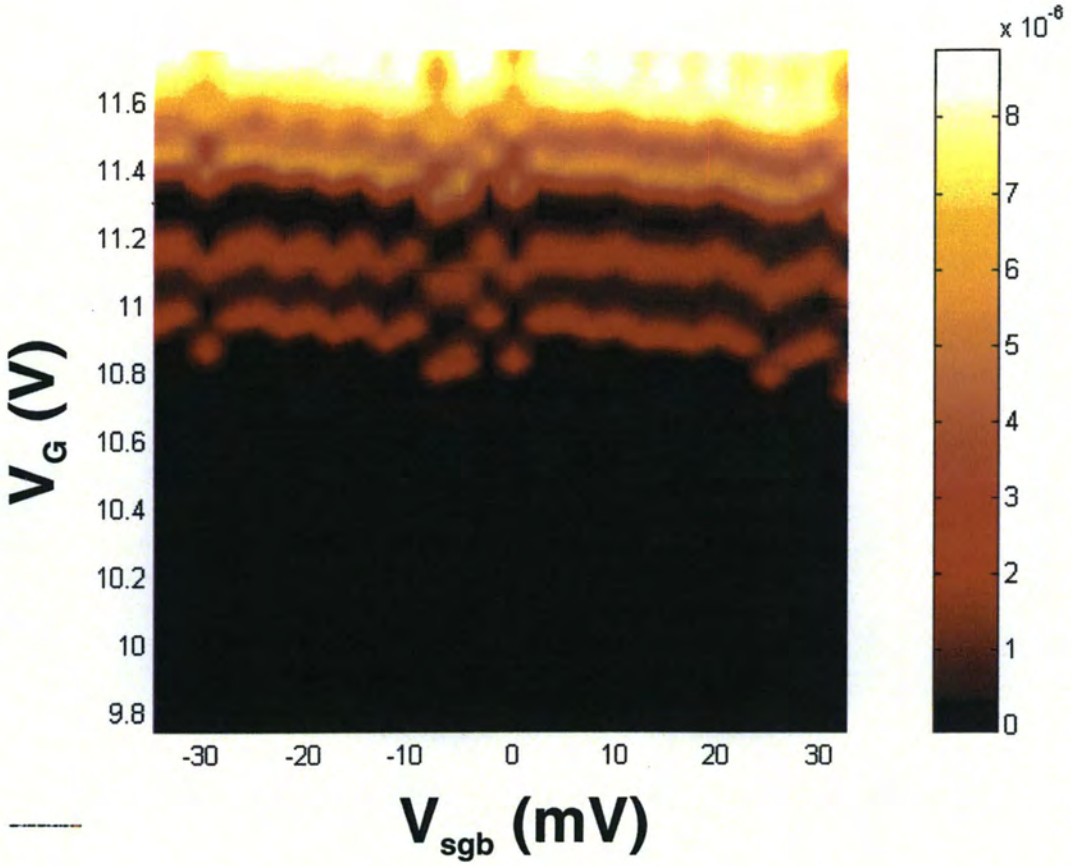

Figure 4.12: A composite plot of inversion gate sweeps of device D9 for different bottom depletion gate biases. The color scale represents the conductance amplitude. Apparent shifting in the conductance peak positions is an artifact of the order in which the individual sweeps were taken and the offset due to a presumed oxide charging mechanism.

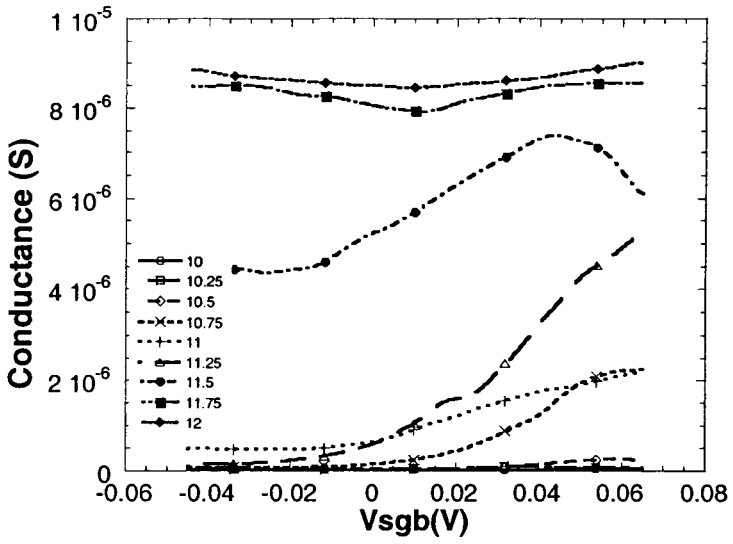

Figure 4.13: A linear plot showing the dependence of the current through device D9 on the bottom depletion gate for various inversion gate biases.

- the spacing in inversion gate bias is not fine enough to detect any significant trends

- The spacing between peaks at different inversion gate biases is a strong function of inversion gate bias, as is shown in Fig. 4.12.

4.3.3 Middle (Plunger) Depletion Gate

Although the plunger gate was not connected in a configuration which allowed for a sweep-type of measurement (due to a limitation in number of source-measure units on the HP4156b), the inversion gate sweep was done for different plunger gate voltages. The results of this measurement are very similar to those for the bottom depletion gate in Section 4.3.1, Fig. 4.12, except for a smaller range of inversion gate bias. A notable difference is that the offset due to the oxide charging seems to be reduced (but clearly not non-existent), likely due a greater amount of relaxation

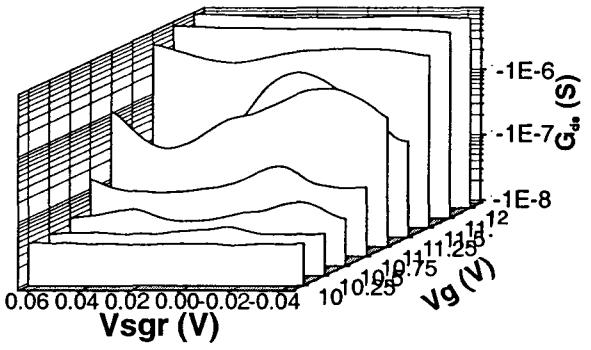

Figure 4.14: A quasi-3D plot showing the dependence of the conductance on the right depletion gate of device D9 for different inversion gate biases.

time between sweeps because the plunger gate bias had to be manually configured on the Keithley 1000 source unit between successive sweeps.

4.3.4 Right Depletion Gate

The characteristic of the right depletion gate has also been measured for different inversion gate biases, and the results shown in Fig. 4.14, plotted as a semi-log quasi-3D plot. It is observed that there is a rapid increase in current within the range of bias.

4.3.5 Effect of Drain Bias

Large drain excitation bias sweeps were taken for different inversion gate biases, as shown in Figs. 4.15 and 4.16, where the current as a function of drain bias is plotted for several values of the inversion gate bias. In these sweeps, the input and output barriers to the contacts have been purposely biased asymmetrically to look

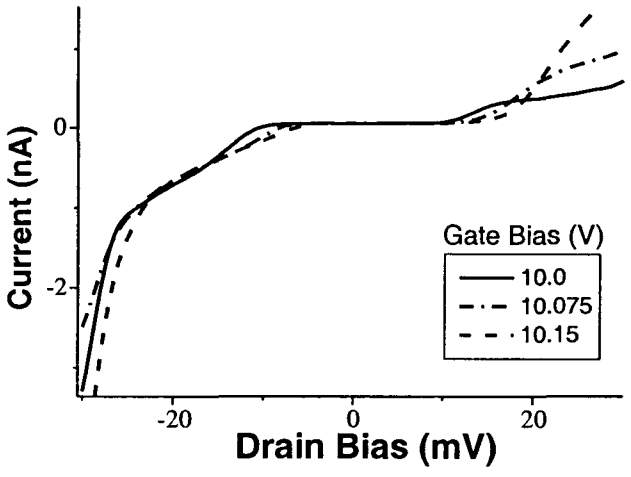

Figure 4.15: Drain bias sweeps of device D9 for different values of the inversion gate bias. The depletion gate biases are -30 mV, 35 mV, 35 mV, and -15 mV for the left, middle, right, and bottom respectively.

for the Coulomb staircase discussed in Section 2.6.3. Although this phenomenon was not seen, there are obvious kinks in the characteristic for gate bias in the range of 10.0-10.15 V plotted in Fig. 4.15; for slightly larger inversion gate bias these kinks give way to a smooth increase in current which is much greater (Fig. 4.16). Both Figs. 4.15 and 4.16 are for inversion gate bias significantly lower than that where a sharp turn on in current is observed in equilibrium transport observed in Fig. 4.11.

4.4 D12: Symmetric 200 nm Dot

Device D12 is a symmetric dot with the same pattern file (and thus intended) gate structure as D9. However, it will be seen that the features are somewhat different.

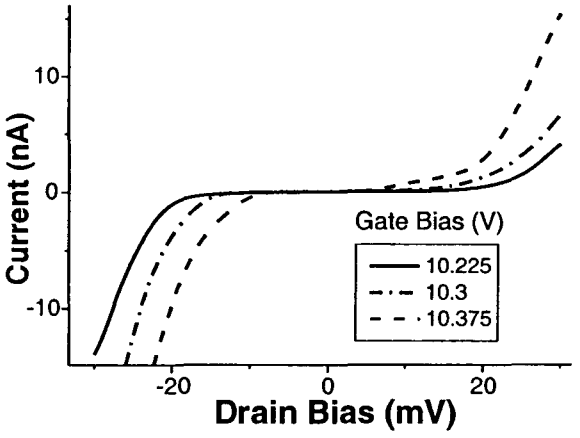

Figure 4.16: Drain bias sweeps of device D9 for various values of the inversion gate bias. The depletion gate biases are -30 mV, 35 mV, 35 mV, and -15 mV for the left, middle, right, and bottom respectively.

4.4.1 Inversion Gate Sweep

The first type of experiment on this device is a sweep of the inversion gate bias with all of the depletion gates held at zero bias. The results of a sweep with the drain bias held at $100\,\mu V$ are shown in Fig. 4.17. In this sweep, there were some major peaks, with some minor peaks, shoulders and inflection points superimposed. These points have again been picked by a numerical peak finding routine and the bias values labeled. A major feature of the data is the presence of a negative current. It is believed that this is due to an unresolved ground loop in the measurement system which serves as a lower resistance path in the subthreshold region of the MOSFET operation.

4.4.2 All Depletion Gates Tied

The $I-V$ characteristics were also measured for the case where the depletion gates are all tied to the same bias, which is swept for a fixed source-drain and

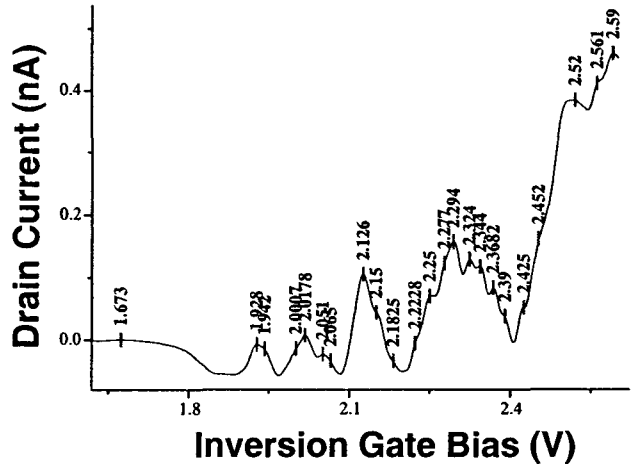

Figure 4.17: A single inversion gate sweep of device D12 where all of the peaks and inflection points have been identified and labeled.

inversion gate bias. An example of a single sweep is shown in Fig. 4.18, where the source-drain bias is held at $100\,\mu$V and the inversion gate bias is 1.615 V. In the figure, peaks in the source-drain current are observed. If we now generate a composite image of all such sweeps taken for a range of inversion gate biases, the position of these peaks is observed to evolve as a function of both inversion and depletion gate bias. The source-drain current is plotted in the depletion gate-inversion gate plane in Fig. 4.19, where the amplitude is represented by a color map. Superimposed on the color mapped plot are the positions of peak conductance (circles) and shoulders (open squares) discussed earlier in regard to Fig. 4.17. The main overall feature of this plot is that the positions of the peaks evolve in an almost linear fashion with depletion gate bias with slope α. If one examines Fig 4.19 closely, there are several other subtle features:

- There appear to exist two different ladders of lines which evolve with different slopes, α_1 and α_2, which produce crossings.

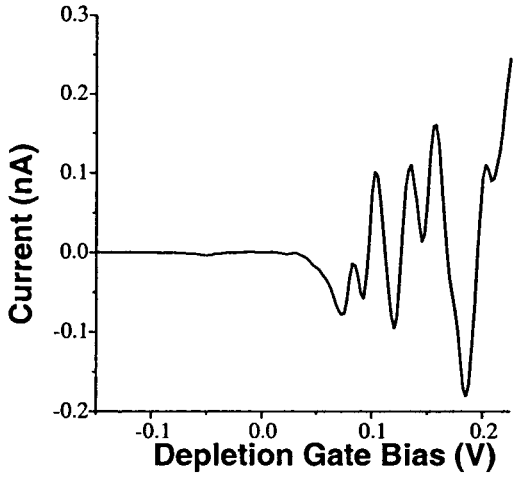

Figure 4.18: A single sweep of the depletion gate bias in device D12 for fixed source-drain and inversion gate biases of 1.615 V and 0.1 mV, respectively.

- There are several regions away from these crossings where one line jumps to another line in the same ladder.
- Some lines go from series of peaks to series of shoulders, and some which begin as series of shoulders disappear only to reappear at lower gate bias.

One would expect that in changing the depletion gate bias, the confining energy would change with it. However, it is equally likely that in changing the inversion gate bias, the shape of the potential would also be changed by such effects as carrier screening. From the data, it appears that the changes due to the inversion gate bias are equally balanced by the changes due to the depletion gate bias, keeping the change linear.

4.4.3 Plunger Depletion Gate

Rather than varying all of the depletion gates simultaneously, measurements were performed in which only the plunger gate (shown in Fig. 3.2) was swept, while

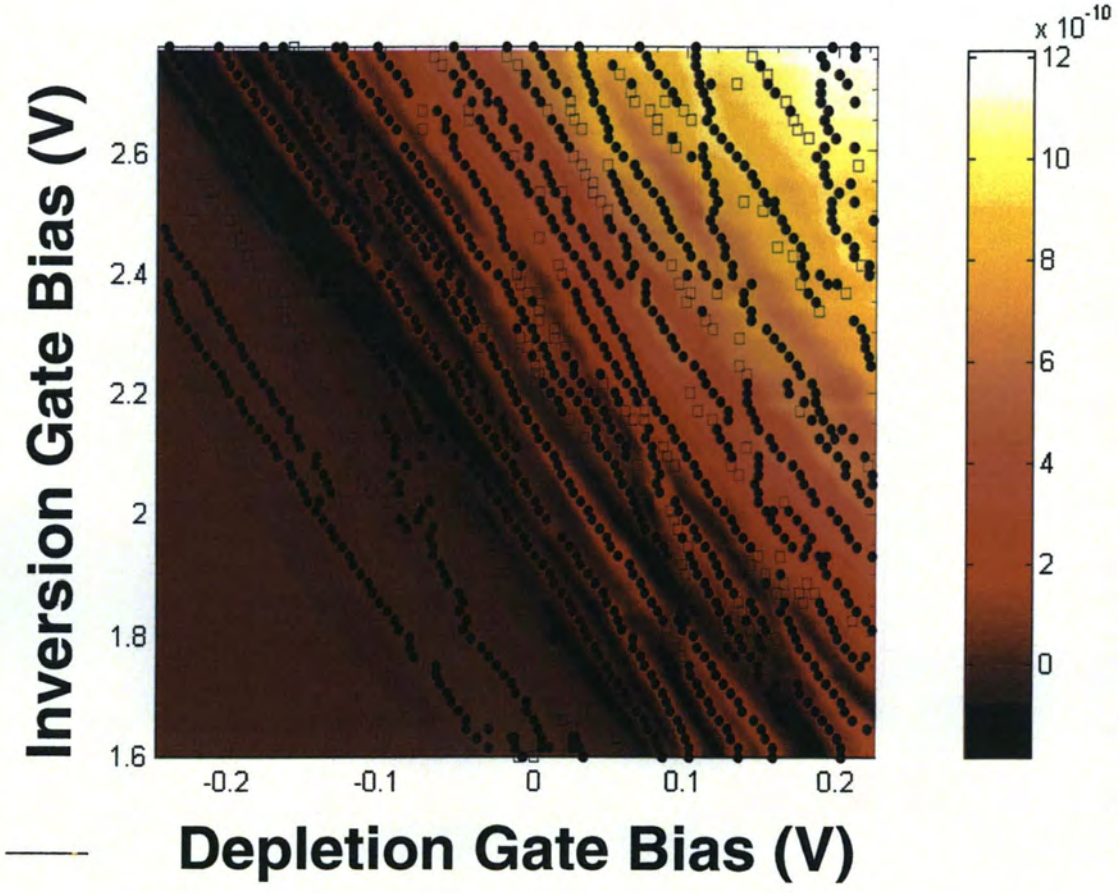

Figure 4.19: A composite plot of individual sweeps of the depletion gate bias in device D12 for different inversion gate bias with the position of the peaks overlaid. The filled circles are the position of the peaks, and the open boxes are the position of clear shoulders.

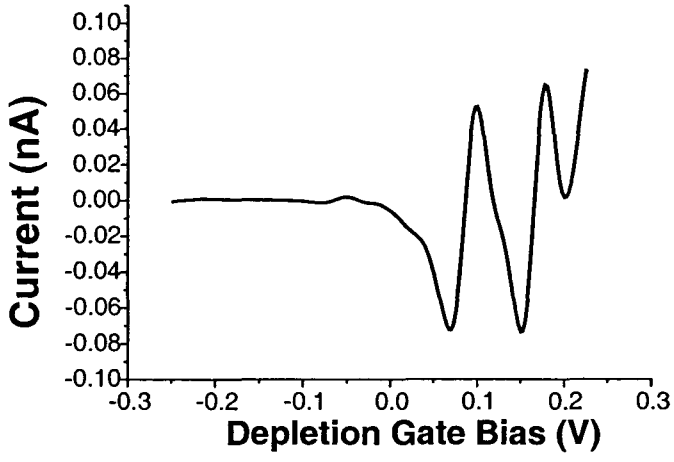

Figure 4.20: A single sweep of the plunger gate bias in device D12 (with all other depletion gate biases held at ground) for fixed source-drain and inversion gate biases of 0.1 mV and 1.735 V, respectively.

the other depletion gates were held at ground. The plunger controls the input and output channels directly, and should in principle have less effect on the overall shape and size of the dot itself compared to the biasing all gates simultaneously. It also allows us to look at asymmetry effects in the dot shape. An example of an individual sweep is shown in Fig. 4.20.

Figure 4.21 is a composite of the sequence of sweeps for the plunger gate. The color map has the same meaning as before, and the peak positions are again superimposed over the color map plot. Here, it is even more evident that there is an additional ladder of peaks that crosses the first ladder. Also, it is not clear that the first peaks which are visible in this plot correspond to starting with zero electrons in the dot.

Viewing this sequence of sweeps in a quasi-three dimensional plot allows us to visualize the peak structure in a different way. Figure 4.22 shows a waterfall plot

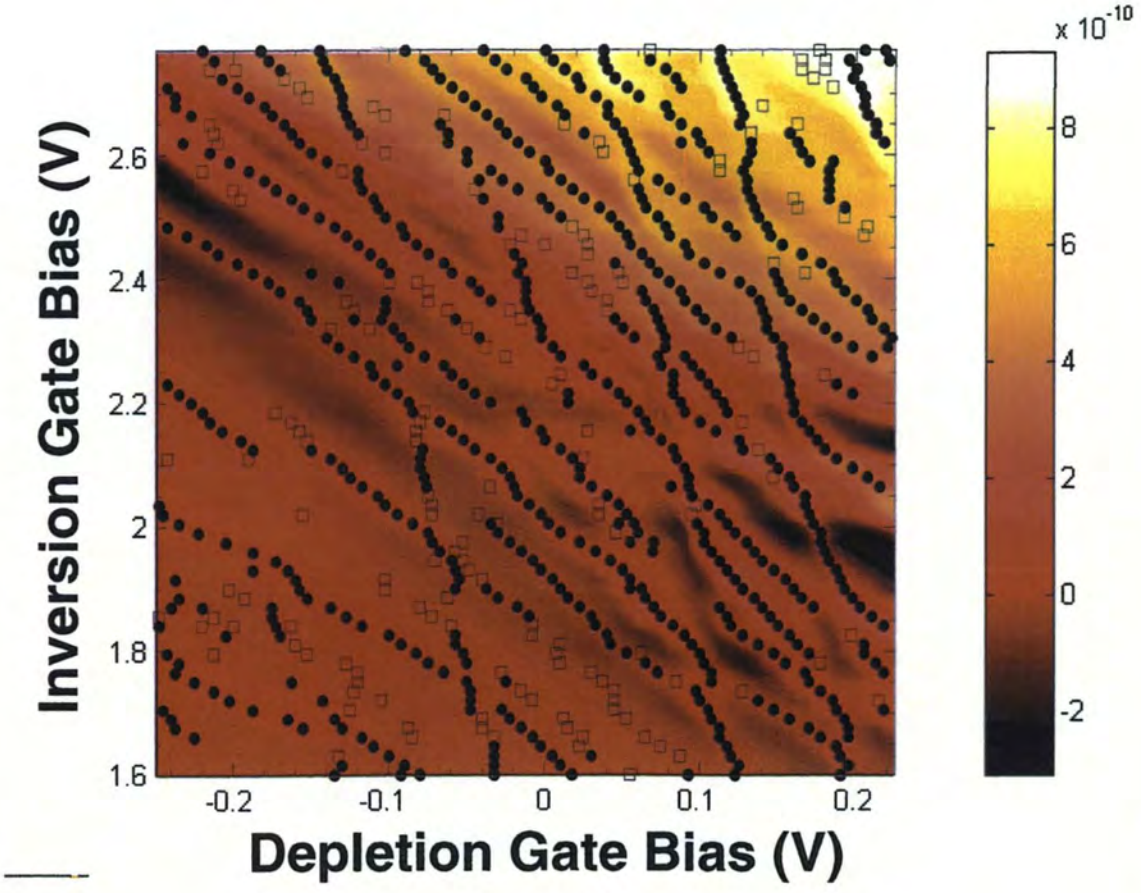

Figure 4.21: A composite plot of individual sweeps of the depletion gate bias of device D12 with the position of the peaks (filled circles) and shoulders (open boxes) overlaid.

that has been rotated in such a way that the crossing points are easily visualized, showing that the peaks that have the higher slope α, also have larger amplitude.

4.4.4 Non-Zero Drain Bias

To study non-equilibrium effects, the inversion gate is swept for small finite drain bias. The results are shown in Fig. 4.23 and Fig. 4.24 for various positive and negative drain biases, respectively. These figures show that the peak current magnitude increases monotonically as the excitation bias is increased, as expected.

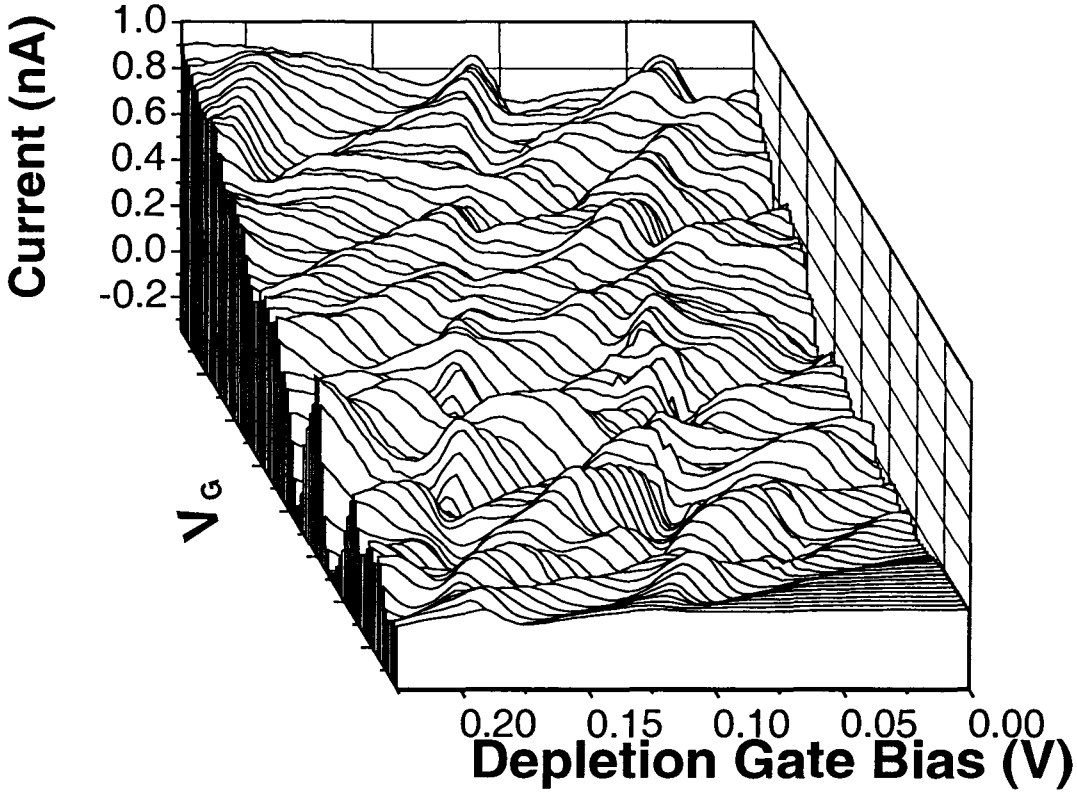

Figure 4.22: A waterfall plot of the plunger gate sweeps oriented with the view along the higher ladder of peaks.

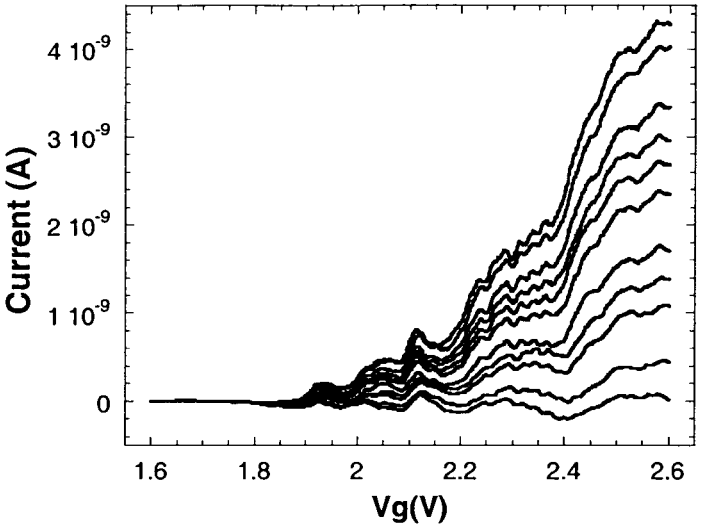

Figure 4.23: A set of increasingly positive drain excitation bias in device D12, showing oscillations with a fine structure which are persistent in the range of measured drain excitation of 100 μV to 1 mV DC.

There is an oscillatory structure which is persistent above the 1 mV excitation bias. However, the peak to valley ratio decreases as the excitation bias is increased. In addition, there is a fine structure, which remains approximately the same, whose peak to valley ratio does not decrease with the increasing drive bias.

The existence of current with zero source-drain excitation bias suggests that there is a ground loop in the system, to which the negative currents observed could also be attributed, as previously mentioned. A composite plot of all of the individual sweeps is shown in a semi-log plot of the current magnitude in Fig. 4.25. In the color map, the darker color is smaller current. It can be seen that changing the inversion gate bias moves the position of the gap back and forth in drain bias.

Larger bias sweeps of the source drain excitation bias were performed as shown in Figs. 5.24 and 5.27. As mentioned in Section 2.6.3, observation of the Coulomb staircase, which should occur due to single electron charging, gives information about the charging energy in the dot and its coupling to the drain bias. If the

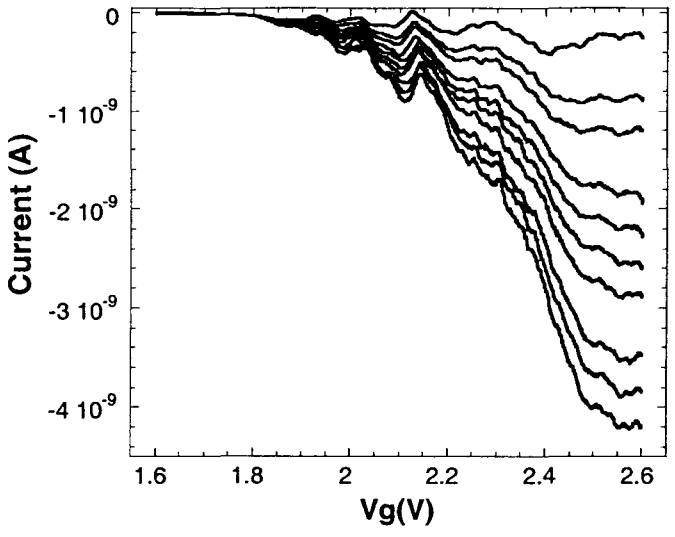

Figure 4.24: A set of increasingly negative drain excitation bias, showing oscillations with a fine structure in device D12, which are persistent in the range of measured drain excitation of 0.0 mV to -1 mV DC.

Coulomb charging energy were the dominant energy scale and the dot barriers were asymmetrically biased and the overall conductance of the channel were $\gg 2e^2/h$, the Coulomb staircase would be evidenced.

4.5 D11: Asymmetric 100 nm by 200 nm Dot

Device D11 was fabricated with an asymmetric geometry (as shown in Fig. 3.6 frame b) with the idea of specifically probing the effects of an asymmetric confinement potential and therefore help in understanding the transport spectroscopy of the symmetric devices as well.

4.5.1 Inversion Gate Sweep

The measured inversion gate sweeps for constant drain bias are shown in Fig. 4.26 for several depletion gate biases for sample D11. The inset is a smaller interval with range similar to that of depletion gate sweep measurements. The major

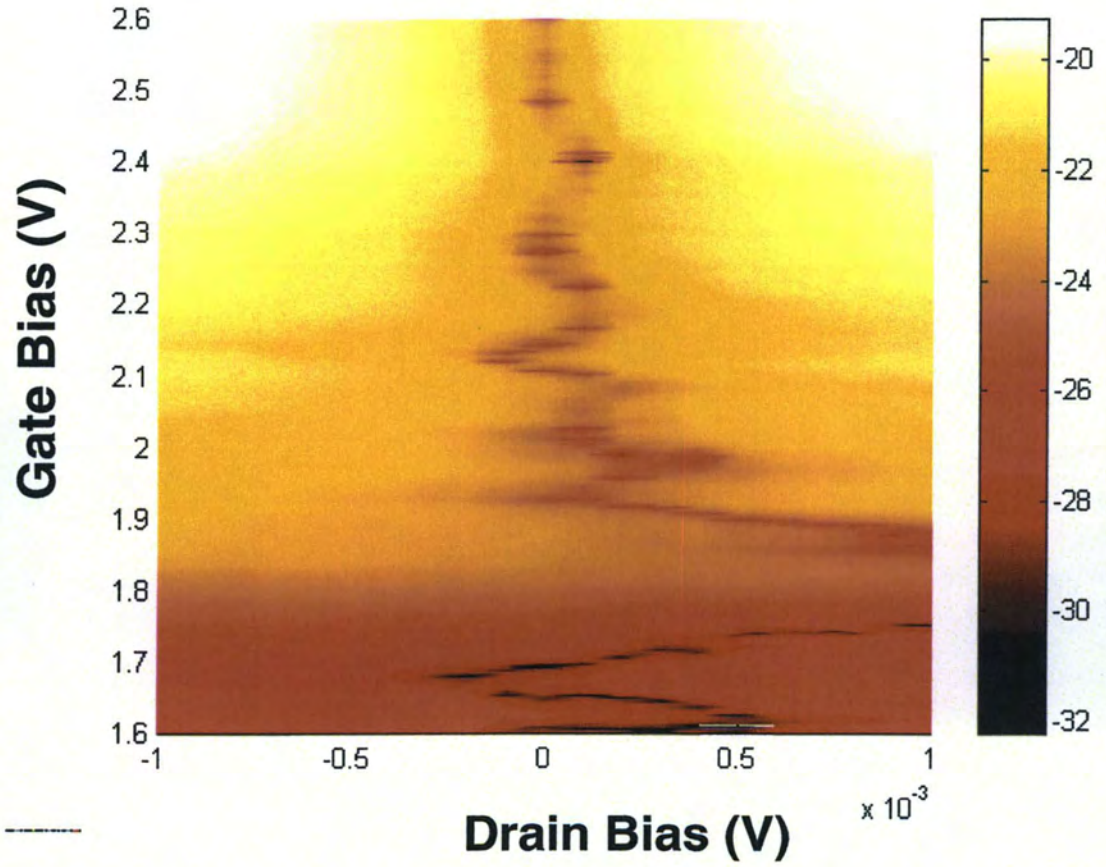

Figure 4.25: A semi-log plot of I_d sweeps of device D12 for finite drain bias. The gap in current is seen to move back and forth in drain bias as the inversion gate bias is changed.

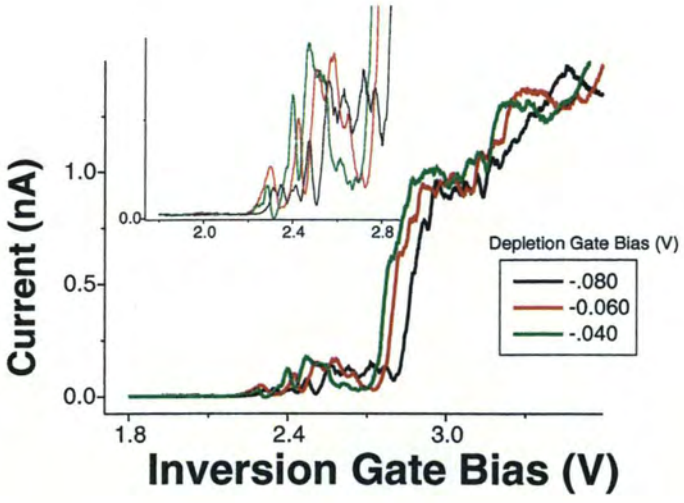

Figure 4.26: Inversion gate sweeps of device D11 for several increments of the depletion gate bias, with all depletion gates tied to the same voltage source.

features are similar to the other devices measured. In this device, there was a very dramatic turn on and some fine structure similar to device D12.

4.5.2 All Depletion Gates Tied

Most of the measurements were performed with all of the depletion gates tied together. These measurements gave similar results to devices D10 and D12. Figure 4.27 shows a sequence of sweeps of all depletion gates tied together for several successively larger inversion gate biases. Note how the relative importance of each peak changes with changing inversion gate bias as the sweeps progress. Also note how the current amplitude is much lower than for the other devices for similar biasing conditions.

A composite plot of all depletion gate sweeps is shown in Fig. 4.28, with the peak positions superimposed over the color mapped plot. An interesting feature of this plot is that there are apparently three different types of lines. Members of

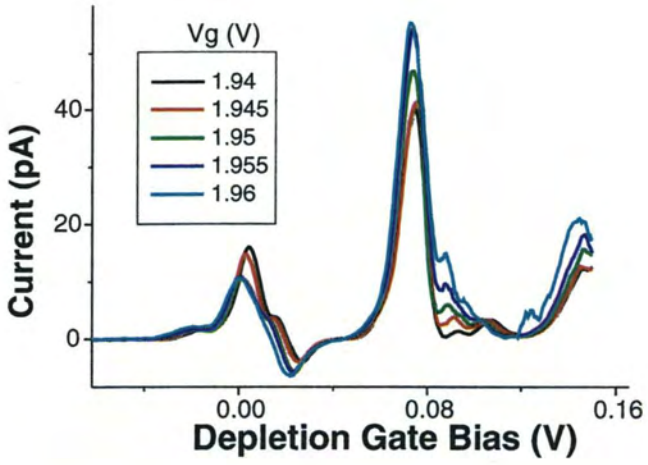

Figure 4.27: A sequence of several depletion gate sweeps of device D11 for successively larger inversion gate biases.

each group have the same slope in the $V_{\text{Inv}} - V_{\text{Depl}}$-plane, which is different from the slopes in the other groups. This is discussed in more detail in Chapter 5.

4.5.3 Non-zero Drain Bias

Figure 4.29 illustrates the effect of non-zero drain bias on the $I-V$ characteristics, where the top inversion gate bias is swept while the drain bias is incremented from 100 μV to 3.4mV. As the inversion gate bias is swept for increasing drain bias, the position of the current peaks evolve linearly with positive slope. The positive slope is a departure from D12. For drain bias around 3.2 meV, most of the features have broadened to a point that they are indistinguishable from the background current increase.

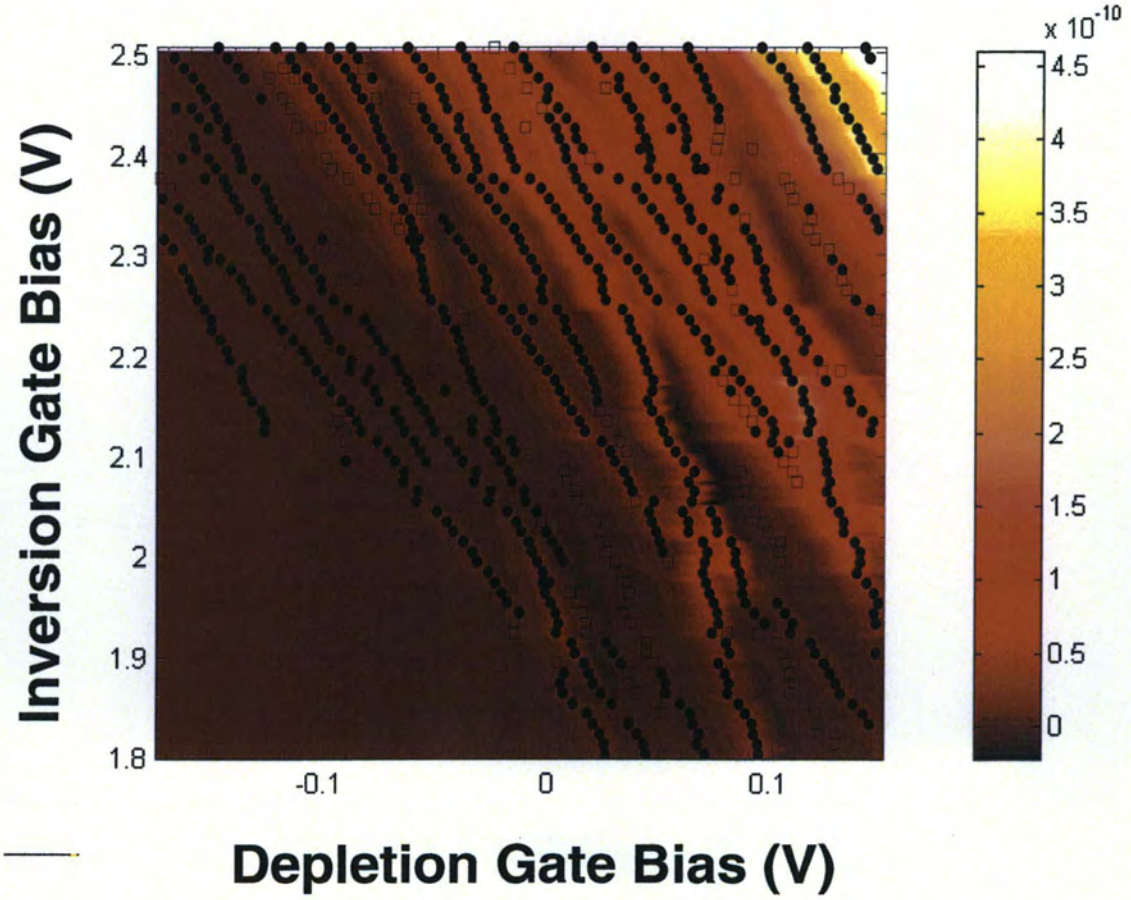

Figure 4.28: A composite plot of all depletion gate sweeps of device D11 for increasing inversion gate bias. The peak (filled circle) and shoulder (open square) positions have been overlaid.

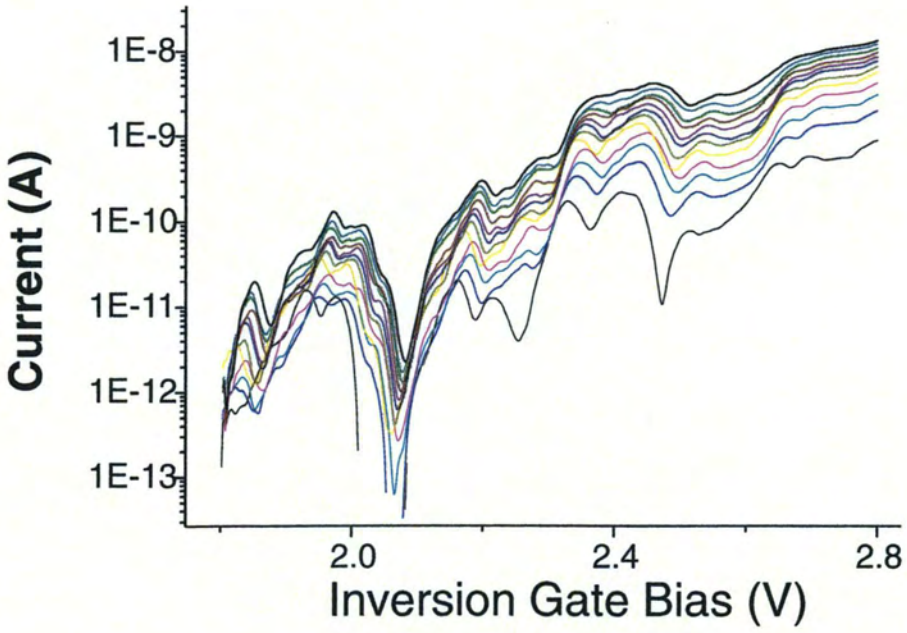

Figure 4.29: A sequence of inversion gate sweeps of device D11 for increasing drain bias from 0.1 mV to 3.4 mV in 0.3 mV steps.

5. ANALYSIS

5.1 Introduction

In this chapter, an analysis of the transport data for the devices D10, D9, D12, and D11 will be presented. The calibration of the inversion charge density and reservoir Fermi energy as a function of the inversion gate bias is first discussed in Section 5.2. This will be followed by a discussion of the dot capacitance and size in Section 5.3. In Section 5.4, the effort to calculate the quantum dot energy levels within the tunneling regime is discussed, followed in Section 5.5 by a comparison of the experimental addition energy spectrum. Section 5.6 follows with a discussion of how the peak shape characterization can be related to the various dot capacitances and energy level fine structure. In Section 5.7 the behavior of dot D11 in a magnetic field will be discussed, followed in Section 5.8 by a discussion of the results for a finite source-drain bias. Perhaps the most important results in this work are presented in Section 5.9, the evolution of the peak conductance positions in the inversion gate bias-depletion gate bias plane. Included are simulations which attempt to explain these results in terms of the calculated energy level structure presented in Section 5.4.

5.2 Conversion of Inversion Gate Bias to Reservoir Fermi Level

Here the conversion from the raw magnetoresistance data to inversion layer electron density and the calibration of the Fermi Level in the channel as a function of inversion gate bias will be discussed.

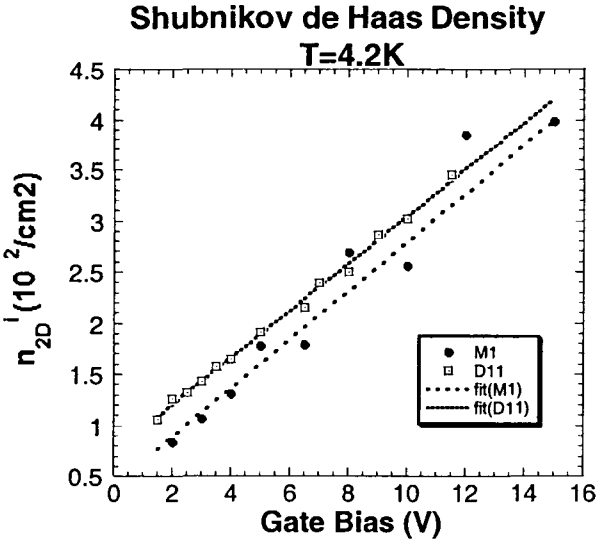

Figure 5.1: Inversion charge density as determined by SdH measurements at 4.2K for devices M1 and D11.

5.2.1 Magnetotransport and Density Determination

As discussed in Section 5.7, the density of a 2DEG can be determined from the periodicity of Shubnikov de Haas oscillations with inverse magnetic field, as given by Eq. 2.14, or solving in term of the sheet charge density,

$$n_{2D}^{i} = \frac{g_s g_v}{2\pi\hbar\Delta\left(\frac{1}{B}\right)}. \tag{5.1}$$

As discussed in Section 2.3.1, the valley degeneracy g_v for the lowest silicon (1 0 0) surface (in the absence of valley splitting) is 2, and the spin degeneracy of the lowest lying set of Δ_2 subbands g_s is 2. The 2DEG density as a function of inversion gate bias is plotted in Fig. 5.1 for devices M1 and D11, extracted using Eq. 5.1 and the positions of the SdH peaks extracted numerically. As expected, they can be well fit by a straight line. The extracted relation between the inversion charge density n_{2D} and the gate voltage for the data shown in Fig. 5.1 is

$$n_{2D} = 2.31(3.20 + V_g) \times 10^{11}(\#/\mathrm{cm}^2). \tag{5.2}$$

In an MOS capacitor, the relation between the 2DEG density and the capacitance and threshold voltage is given by

$$n_{2D} = \frac{C_{\text{ox}}}{q}(V_g - V_{\text{Th}}) \tag{5.3}$$

where the capacitance per unit area is given by $C_{\text{ox}} = \epsilon_{\text{ox}}\epsilon_0/x_{\text{ox}}$. In an ideal MOSFET, the threshold voltage depends on the temperature, dopant density, and substrate bias. However, imperfections in the oxide (such as oxide charges), interface states, and the electron work function for the metal may shift this threshold voltage. Comparing the results of the SdH measurements to the expected value in a device with a 95 nm oxide, we would expect a slope of $2.27 \times 10^{11}/\text{cm}^2$, which is in good agreement to our measured slope of $2.31 \times 10^{11}/\text{cm}^2$.

If the position of the SdH oscillation peaks are plotted as a function of the inversion gate bias, a fan diagram results in which the dependence of the Landau levels on the gate bias can be seen. This diagram helps identify spin and valley splitting, and the population of higher 2D subbands. An example of a fan diagram for the data of D11 is given in Fig. 5.2. One can associate sequences of peaks (shown by the straight lines) with the evolution of a given Landau level with magnetic field. In Fig. 5.2, it is speculated that peak labeled 1' may be attributed to spin or valley splitting. The first eight Landau levels from the lower ladder of subbands (Δ_2 ladder) fill to the right. No evidence for upper subband occupancy is observed, which supports the assumption that electrons may be treated in the electron quantum limit (EQL) in calculations.

5.2.2 Fermi Level Calculation

It was shown in Section 2.3.2 that for a pure 2DEG system, the 2DEG density and the Fermi energy are related by Eq. 2.10. For low temperatures, this may be

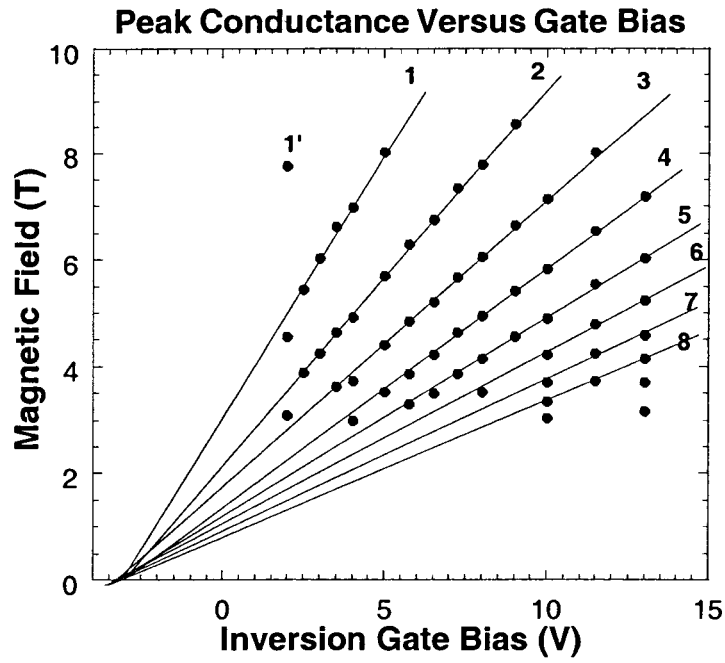

Figure 5.2: A fan diagram for the SdH data from D11. the Landau levels corresponding to the filling of the Δ_2 ladder of sub-bands are denoted by un-primed labels; 1' denotes filling of the lowest Δ_4 Landau level.

accurately approximated by the expression

$$n_{2D}^{i} = \frac{g_s g_v m_{de}^{v}(E_F - E_i)}{2\pi\hbar^2}, \tag{5.4}$$

which, when solved for the Fermi energy relative to the energy at the bottom of the subband of index i, gives the linear relation

$$E_F - E_i = \frac{2\pi\hbar^2 n_{2D}^{i}}{g_s g_v m_{de}^{v}}. \tag{5.5}$$

As was determined in the previous section, the devices measured in this work behaved as a single subband 2DEG for all experimental conditions encountered. Using Eq. 5.5 and the experimentally determined inversion gate 2DEG density dependence, the Fermi energy in the reservoirs can be calculated as a function of the inversion gate bias. Using the experimental data for the 2DEG density presented in

the previous section (Eq. 5.2) and Eq. 5.5, we have

$$E_F - E_i = 5.81(3.2 + V_{Inv})\,\text{meV}, \tag{5.6}$$

for the Fermi level referenced to the bottom of the lowest Δ_2-subband ($E_i = E_0$).

5.3 Estimation of Dot Sizes and Capacitance

In this section, the quantum dot size will be estimated first based on Coulomb blockade analysis (that is to say that it is assumed that transport behaves as if purely Coulomb blockade limited), and then under the assumption that the confinement energy dominates the transport.

5.3.1 An Estimate of Dot Size Based on Coulomb Blockade Analysis

If CB is the limiting factor in transport (*i.e.* the dominant term in Eq. 2.84) then it is possible to make an estimate of the dot size based on the effective capacitance of the dot. While there are fluctuations in the periodicity in each of the gate voltage sweeps which suggest the probable role of electron-electron interactions and the underlying discrete energy level structure, a median value for the capacitances will be assumed a valid measure of the system. The total capacitance of the system is given by equation

$$C_\Sigma = C_{\text{Self}} + C_{\text{gates}} + C_{\text{Bias}} \tag{5.7}$$

where C_{gates} is the algebraic sum over all the capacitively coupled gates in the system, C_{Bias} will be discussed below, and C_{Self} is the self-capacitance of the dot itself, which is ultimately what we seek. For device D12, the algebraic sum of gate capacitances may be computed as follows:

- The inversion gate capacitance is taken from the mean voltage splitting between consecutive central peaks and Eq. 2.84, to be $C_{\text{Inv}} = 1.08$ aF.

- From Table 5.3, the average value of the sum of depletion gate capacitances is $C_{\text{Depl}} \approx 6.65$ aF. These values can be obtained by either fitting the capacitance using Eq. 2.85 or the peak splitting using Eq. 2.84.

- The total capacitance of the system can be taken as the mean value as obtained from the peak fits from Tables 5.3 and 5.4 to be $C_\Sigma \approx 67.62$ aF. Note that not all devices could be fit in this manner; an alternative approach is to assume that the Coulomb gap voltage (*i.e.*, the voltage range around zero bias where current is supressed) gives the value of C_Σ using Eq. 2.84. This method was used in Table 5.1, where C_{Bias} is neglected. Using this method for D12 gives $C_\Sigma \approx 32$ aF.

- The capacitance of the biased lead will be neglected in this section; discussion of this quantity will be deferred to Section 5.8

Subtracting from the total capacitance these individual parts gives the dot capacitance

$$C_{\text{self}} = C_\Sigma - [C_{\text{Depl}} + C_{\text{Inv}}] \approx 24.3\,\text{aF}. \tag{5.8}$$

To estimate the size, a model for the quantum dot self capacitance must be found. Several possible model geometries and their associated capacitances are [163]

$$\text{Parallel plates:} \quad C = \frac{\epsilon A}{d} \text{ assuming } A \gg d \tag{5.9}$$

$$\text{disc above conducting plane:} \quad C = 8\epsilon R \tag{5.10}$$

$$\text{Conducting sphere:} \quad C = 4\pi\epsilon a \tag{5.11}$$

conducting sphere above grounded plane:

$$C = 2\pi\epsilon\sqrt{d^2 - 4a^2}\sum_{j=0}^{\infty}\left(\coth\left[(j+1/2)\text{arcosh}\left(\frac{d}{2a}\right)\right] - 1\right) \tag{5.12}$$

where a is the sphere's radius and d is the distance from the sphere's center to the grounded plane. The most popular model among experimentalists is that of the

conducting disc, Eq. 5.10, since the 2DEG is narrow and the electrostatic confinement gives a rounded lateral potential. For Eq. 5.10 to hold, one must assume that R is much smaller than the distance above the plane. Here, since the distance from the Si-SiO_2 interface is $\approx$ 95 nm, this is a questionable assumption. The solution incidentally is given as a problem in Jackson's book on Electrodynamics [164]. Taking the relative dielectric constant for Si as 11.7, Eq. 5.10 gives the radius of the dot as $R_{\text{dot}} \approx 29.3$ nm, or a diameter of 58.6 nm. These results are significantly larger than those reported by Khoury *et al.* [165] for a similar structure.

However, it is equally reasonable to consider the case of a sphere above a plane (Eq. 5.12), since in a MOS structure there is a finite width to the inversion layer on the order of 10 or 20 nm. As the expression 5.12 does not lend itself easily to inversion for a, instead we will use Eq. 5.11 as an approximation. Table 5.1 gives a summary of the capacitances for devices D9, D10, D11, and D12, as well as the calculated radii for all but D11. It should be noted that the value for C_Σ for device D12 is the second value stated above, which gives a result much more consistent with the other devices measured. In all cases except D10, the derived radius is much less than the lithographic dimensions of the dots shown in Fig. 3.6, which is expected based on the depletion potential around the gates which extends inwards from the metal edge itself.

5.3.2 An Estimate of Dot Size Based on Discrete Energy Levels

For a device the size given in Table 5.1, one should certainly expect that the discrete nature of the dot states would be a dominating factor in the tunneling transport. In this section, it is assumed that the period of tunneling current peaks is due solely to the spacing of confinement energy states in the dot rather than CB. Theoretical calculations based on the same dot structure [166] (to be presented in Section 5.4) show that the lower lying energy eigenvalues are grouped with nearly

Table 5.1: Capacitance and radius values for dots, assuming pure Coulomb blockade transport.

Device	C_Σ (aF)	C_{Depl} (aF)	C_{Inv} (aF)	C_{Self} (aF)	R_{Dot} (nm)
D9	24.6	7.36	0.917	16.4	12.6
D10	267	1.49	4.76	261	199
D11	16.9	2.20	0.903	13.9	?
D12	32.0	6.65	1.08	24.3	29.3

uniform spacing, as predicted for dots with a harmonic lateral confinement potential (see Section 2.5.1). These calculations appear to be supported by the data to be presented in Section 5.9.

In the following, it is assumed that the spacing of conductance peaks is due solely to the alignment of the quantum dot energy eigenvalues with the Fermi energy in the leads. For simplicity assume that the wave function is that of a one dimensional simple harmonic oscillator, which leads to well known results for the lateral extent of the wave function given by (see, for example the text by Cohen-Tannoudji *et al.* [167]) the root mean square deviation of the spatial extent (X) of the wave function

$$\begin{aligned} \Delta X = \sqrt{(\Delta X)^2} &= \left[\langle\psi_n(x)|X^2|\psi_n(x)\rangle\right]^{1/2} \\ &= \frac{1}{\beta\sqrt{2}}, \quad \beta = \sqrt{\frac{m^*\omega}{\hbar}}, \end{aligned} \tag{5.13}$$

where $\hbar\omega = \Delta E$, the confinement energy for electrons in the dot and β has dimensions of inverse length. For the ground state $\Delta X = 2R_{\text{Dot}}$. The values for

Table 5.2: A summary of the estimated dot size and capacitance based on the assumption that the confinement energy dominates the transport

Device	$\delta\bar{E}$ (meV)	R_{Dot} (nm)	C_{Self} (aF)
D9	0.282	17.6	22.9
D10	0.066	115.1	149.8
D11	0.106 (0.358)	28.6 (15.6)	28.7
D12	0.141	24.8	32.3

each quantum dot measured are given in table 5.2, where the mean energy splitting between tunneling peaks are taken from Section 5.9. Again, the radius is less than the lithographic dimensions of the dot.

By equating the charging energy to the confinement energy, ($E_C = E_Q$), the critical dot radius for which CB dominated energy gives way to confinement dominated energy:

$$R_{\mathrm{Dot}} = \frac{\hbar^2 \pi \epsilon_0 \epsilon_r}{q^2 m^*} = 1.413\,\mathrm{nm} \tag{5.14}$$

For all values of R_{Dot} greater than this value, theoretically the charging energy dominates.

5.3.3 Discussion

At this point, it is reasonable to question whether the estimates obtained in the previous two sections are in line with what can be expected from a theoretical perspective. As discussed in Chap. 1, one should expect to see the effects of quantization when the value of the phase coherence length, l_ϕ is on the order of the size

of the dot. Based on the experimental data characterizing the 2DEG (namely, the mobility and the sheet charge density extracted from the magnetoconductance and correlated to the field effect and effective mobility from gate sweeps), this length has been estimated to be $l_\phi \approx 1.25 - 1.9 \times 10^{-5}$ cm, or $\approx 125 - 190$ nm. Certainly, it should be expected that devices D9, D11, and D12 have smaller minimum dimensions, as was estimated in Section 5.3 and again in Section 5.4 to be on the order of 40 nm for the lithographic dimension of 200 nm. However, there is some question as to whether a device of the size of D10 or D2 (500 nm lithographic) should exhibit energy quantization effects. If we allow for similar lateral depletion (75 nm) in from each depletion gate edge, the electrostatic extent of the dot is still 350 nm, which is greater than two times l_ϕ. One should be skeptical that energy quantization should be observable in such a system. This suggests that possibly the energy quantization in the quantum point contacts defining the dot entrance and exit lead to the conductance oscillations. Another possibility is that Coulomb blockade dominates the conductance. For a dot with a radius of 175 nm (as defined by the lateral electrostatic depletion) the associated capacitance (using the disk model) is 146 aF with a corresponding charging energy $E_C = 548\,\mu$eV. Yet another possible explanation stems from theory developed to study energy relaxation in semiconductor quantum dot lasers [63]. In this work, it was found that the discrete allowed phonon states drastically retarded the phonon scattering rate; this leads to longer scattering times, which could enhance l_ϕ in semiconductor quantum dots. This effect has been studied more recently for the specific case of Si quantum dots by Dür *et al.* [168, 169], leading to similar findings. No attempt has been made to calculate the l_ϕ enhancement from this effect. It has also been shown that the role of multiple phonon [170], Auger [171], and spectral broadening [172] processes can reduce this so-called phonon bottleneck to phonon scattering.

Another scattering effect detrimental to the observation of quantization of energy levels which must be considered is level broadening. With $\tau_e = \mu m^*/e$ and using the Heisenberg uncertainty principle, it can be argued that level broadening occurs due to the scattering, which is estimated as $\Delta E = \hbar/\tau_e$. From the mobility of $3000 - 4000\,\mathrm{cm}^2/\mathrm{Vs}$ typical of the experimental conditions, the calculated energy level broadening is $125 - 200\,\mu\mathrm{eV}$. Level broadening of this magnitude should smear out all of the fine structure observed in device D12, as well as severely broaden other peak structures observed. This fact suggests once again the possibility of mobility enhancement in the dot. This mobility enhancement could make observation of these small quantization energies possible.

From the analysis in this section, combined with the theoretical calculations which will be presented in Section 5.4, the dot sizes calculated under the assumtion of pure electron confinement energy are more consistent with the as-drawn dot dimensions. There is some difficulty with the asymmetric device in coming up with a value; two values are reported in Table 5.2 for this reason. Also, it is believed that the calculated radius for D10 is underestimated by the confinement energy and overestimated by the capacitance method. It is a reasonable assumption that this is the dominant energy scale in the transport behavior observed in the quantum dots reported on in this work.

5.4 Self Consistent Calculations of Dot Structure

To calculate the energy level structure in these and other quantum dots fabricated in this work, Dragica Vasileska's group at Arizona State University recently developed a full 3-D Schrödinger-Poisson solver [166]. The calculations reported herein were performed by Srdjan Milicic at ASU. The calculations start by computing a self-consistent background potential by solving the Poisson equation taking into account detailed balance of dopant impurities and using Dirichlet boundary con-

ditions for the electrostatic potential applied to the depletion and inversion gates. For the solution of the 3D Poisson equation, either the ILU or the Bi-CGSTAB methods were used. Partial ionization of the dopant impurities are taken into account, but do not account for the exchange-correlation corrections to the ground state energy of the system.

This background potential is used as a starting point to self consistently solve the Schrödinger-Poisson equations for the bound and free electrons within the dot. For the solution of the 3D Schrödinger equations in the dot region, Lanczos/Arnoldi factorization is used to construct an orthogonal basis for a Krylov subspace that provides a way to implement the projection numerically. Implicit restarting is used to overcome the intractable storage and computational requirements in the original Lanczos/Arnoldi method. As such, this method allows calculation of a desired number of the lowest eigenvalues and the corresponding eigenvectors that describe the occupied states in the dot at low temperature. Stürm sequencing, combined with bisection and inverse iteration, is used to find the eigenvalues and the corresponding eigenfunctions of a tri-diagonal symmetric matrix that arises from the finite-difference discretization of the 1D Schrödinger equation for the 2DEG in the leads.

5.4.1 Energy Eigenvalues

in this section, the calculated energy eigenvalues will be presented for a symmetric dot with both symmetric and asymmetric biasing conditions. These calculated energy eigenvalues correspond to the dot geometries of devices D9 and D12. The energy eigenvalues for an asymmetric dot with symmetric biasing conditions will also be presented, corresponding to the dot geometry of device D11. The case of asymmetric dot geometry is not presented because there is not experimental data for the asymmetric dot to compare calculations to.

5.4.1.1 Symmetric Dot, Symmetric Depletion Gate Bias

The first 30 energy eigenvalues have been calculated by the methods described above for various values of inversion gate bias and with all depletion gates tied to the same bias for a quantum dot with the same geometry as that of devices D9 and D12. Fig. 5.3 is a plot of the energy eigenvalues for fixed inversion gate bias treating the depletion gate bias as a parameter, while Figure 5.4 plots the energy eigenvalues for fixed depletion gate bias treating the inversion gate bias as a parameter. It is observed that the effect of increasing the inversion gate bias is to pull down the energy of the lowest ground state eigenvalue (with the relative spacing of all eigenvalues becoming more dense), while the effect of biasing the depletion gates increasingly negative is to push up the lowest ground state energy eigenvalue. It is also observed in both plots that the lower eigenvalues evolve in groups which are nearly degenerate and have a nearly uniform spacing, while as the quantum number increases, the gap between groups of nearly degenerate eigenvalues decreases until the grouping structure is difficult to distinguish. In the first four groups, the membership increases by one per group so that the first 10 eigenvalues are grouped by 1, 2, 3, and 4 members. This grouping structure is in agreement with the degeneracy of an isotropic two dimensional harmonic oscillator potential ($g = n + 1$). The lowest lying states appear to be well described by harmonic oscillator-like solutions, whereas for higher energies, the potential is less symmetric, leading to a splitting of the degeneracies.

The dashed horizontal line in Fig. 5.3 is the position of the Fermi energy in the channel. States above this line are empty, states below it are filled, and states which cross this line are active tunneling channels at thermal equilibrium. The dependence of the ground state energy eigenvalues on the change in inversion gate bias (as a function of principal quantum number) is plotted in Fig. 5.5, while the effect of changing the depletion gate bias (as a function of principal quantum

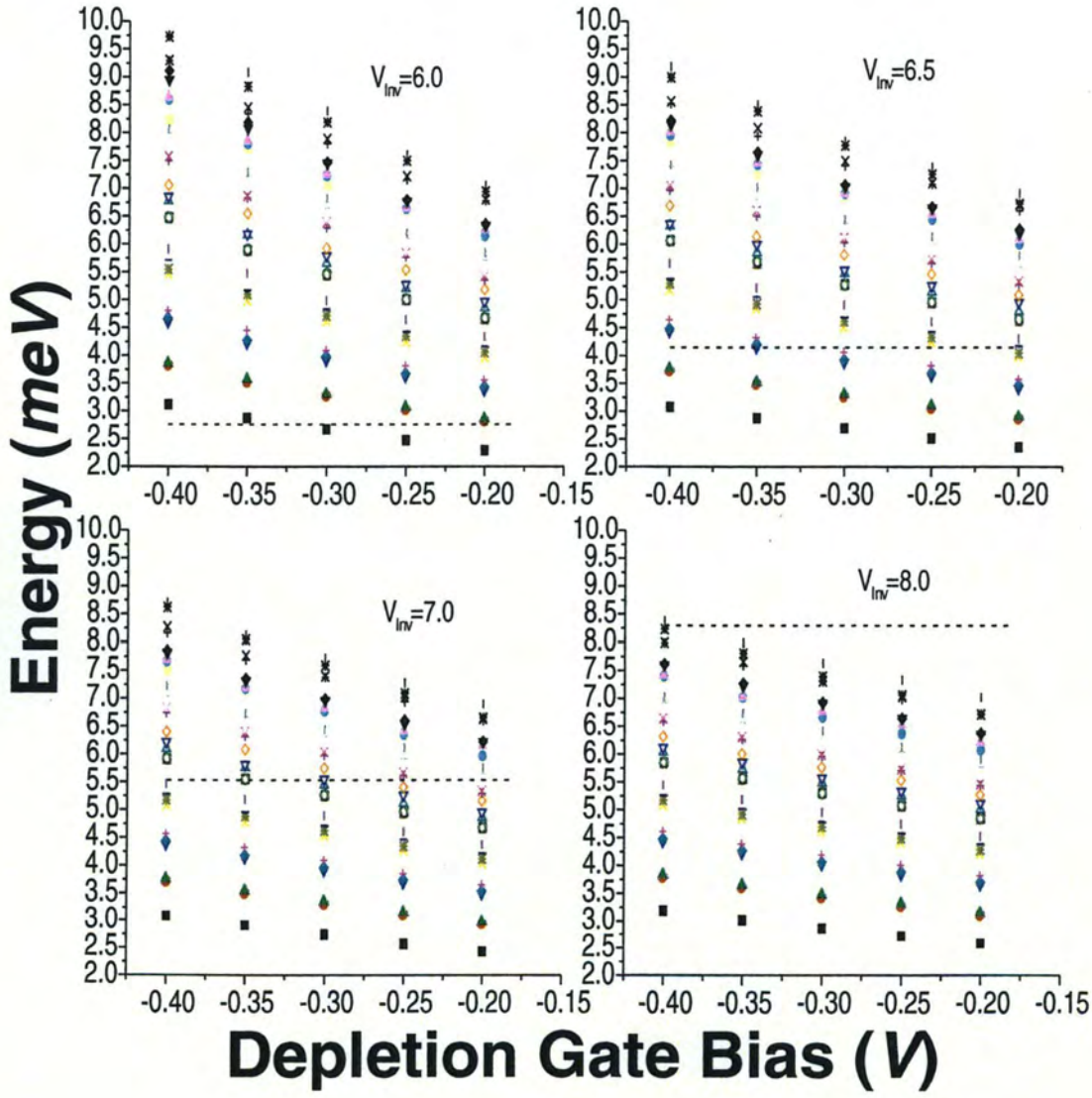

Figure 5.3: A plot of the first 30 energy eigenvalues for various inversion gate biases as a function of the depletion gate bias. The Fermi level in the channel is shown as a horizontal dashed line, with states below the Fermi energy being filled in the dot and those above it being empty.

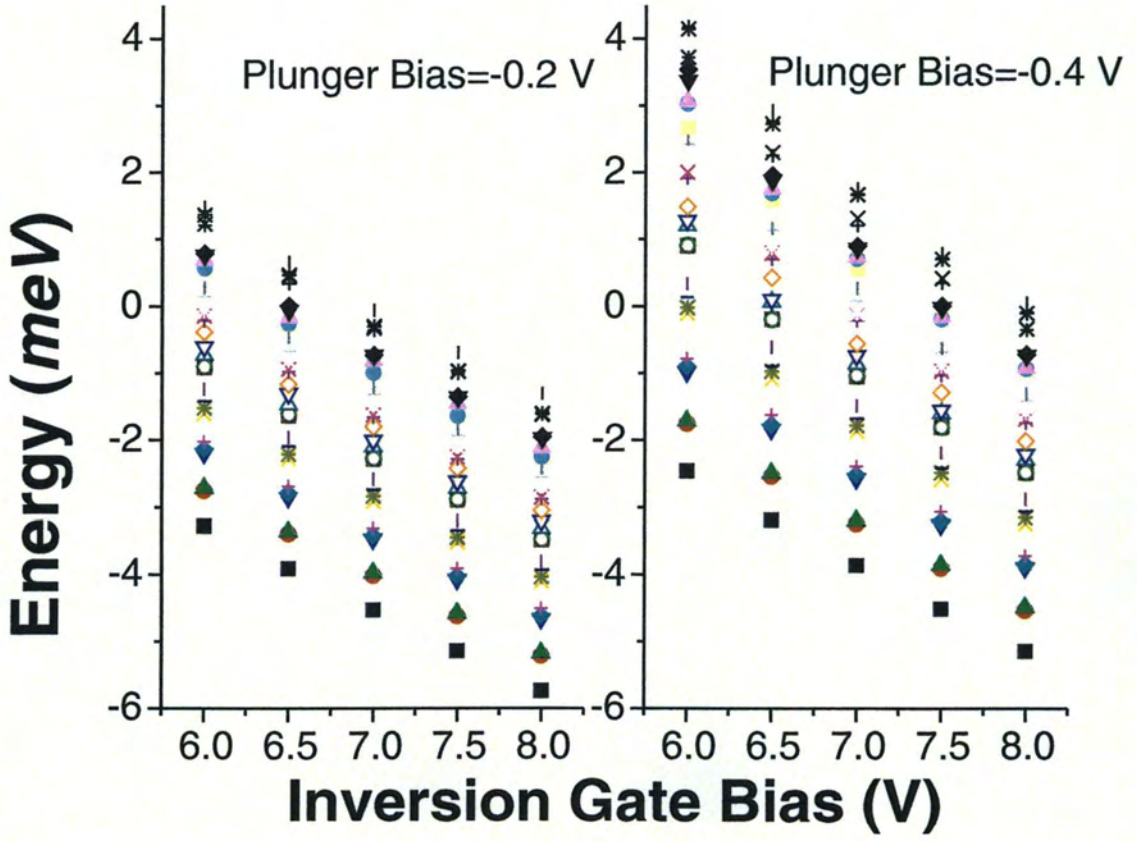

Figure 5.4: A plot of the first 30 energy eigenvalues for various depletion gate biases as a function of the inversion gate bias.

number) is plotted in Fig. 5.6. It is apparent from Fig. 5.3 that for higher energy level indices that there is deviation from a linear dependence for high n. The energy eigenvalues evolve with a nearly linear dependency of slope α_{Inv} (or slope α_{Depl}) defined by the relations

$$\alpha_{\text{Inv}} = \left.\frac{\partial E_n}{\partial V_{\text{Inv}}}\right|_{V_{\text{Depl}}=const.} \qquad \alpha_{\text{Depl}} = \left.\frac{\partial E_n}{\partial V_{\text{Depl}}}\right|_{V_{\text{Inv}}=const.} . \tag{5.15}$$

The same grouping as in Figs. 5.3 and 5.4 also clearly appear in Figs. 5.5 and 5.6. It appears that a more positive inversion gate bias leads to a smaller dependence on n in α_{Depl}, while a more negative depletion gate bias leads to a reduced dependence of α_{Inv} on n. Another feature to point out is that while the dependence on the inversion gate or depletion gate is increased for increasing n, the higher levels are occupied only for higher inversion gate bias, but the dependence is reduced as inversion gate bias is increased. An overall interpretation is that the higher lying states are more weakly bound as a result of this effect.

5.4.1.2 Symmetric Dot, Asymmetric Biasing

Figure 5.7 shows plots of the first 30 energy eigenvalues for a dot with symmetric gate bias, but for the biasing condition where the plunger gate is biased asymmetrically with respect to the remaining three depletion gates, which are held tied to ground. Again, the position of the Fermi energy in the channel (with respect to the lowest subband in the conduction band) is indicated by a dashed line. The primary difference from the case of symmetric biasing of a symmetric dot is that the plunger depletion gate coupling is weak compared to the case where all depletion gates are biased at the same value. The slope of the energy eigenvalues for change of plunger gate bias is plotted in Fig. 5.8 for several values of inversion gate bias. It is clearly observed that these slopes are smaller than those for symmetric depletion gate biasing in Fig. 5.5. Again, the same nearly degenerate grouping is observed in the plots in Figs. 5.7 and 5.8 as in the symmetrically biased counterparts.

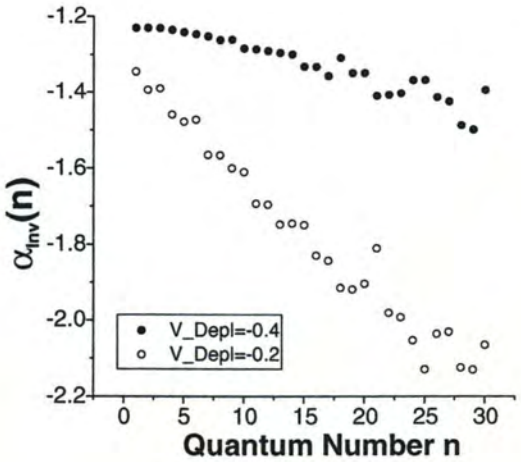

Figure 5.5: A plot of the slope α_{Inv} as a function of quantum number n for depletion gate bias of -0.2 V and -0.4 V.

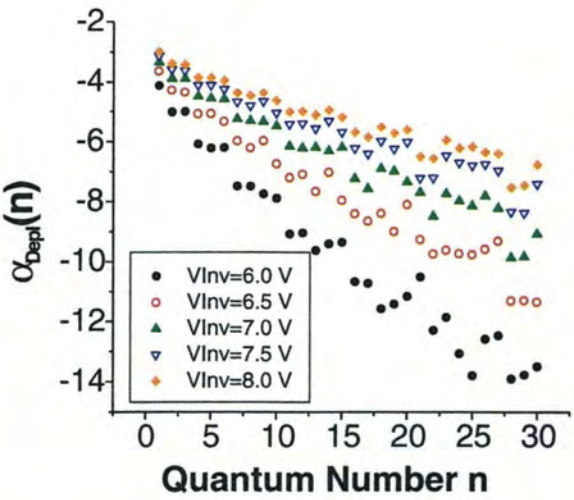

Figure 5.6: A plot of the slope α_{Depl} as a function of quantum number n for various inversion gate biases in a symmetrically biased symmetric quantum dot.

5.4.1.3 Asymmetric Dot

The first 28 to 30 energy eigenvalues for an asymmetric quantum dot with the same geometry as that of D11 are plotted in Fig. 5.9. Two obvious differences that can be observed are that the bottom of the energy band structure is much higher for the asymmetric dot geometry, and that (especially at lower inversion gate biases) there is a clear non-linearity in the dependence of the energy levels on the depletion gate bias. This non-linearity is significant enough that no attempt has been made to extract an α_{Inv} or α_{Depl}. The position of the Fermi level in the channel is marked for each inversion gate bias by the horizontal dashed line. Again energy levels below the dashed line are filled states, while energy levels above the Fermi level are empty. In general, as the inversion gate bias is increased, the dependence on the depletion gate bias is reduced, and the successive energy levels in the ladder are more closely

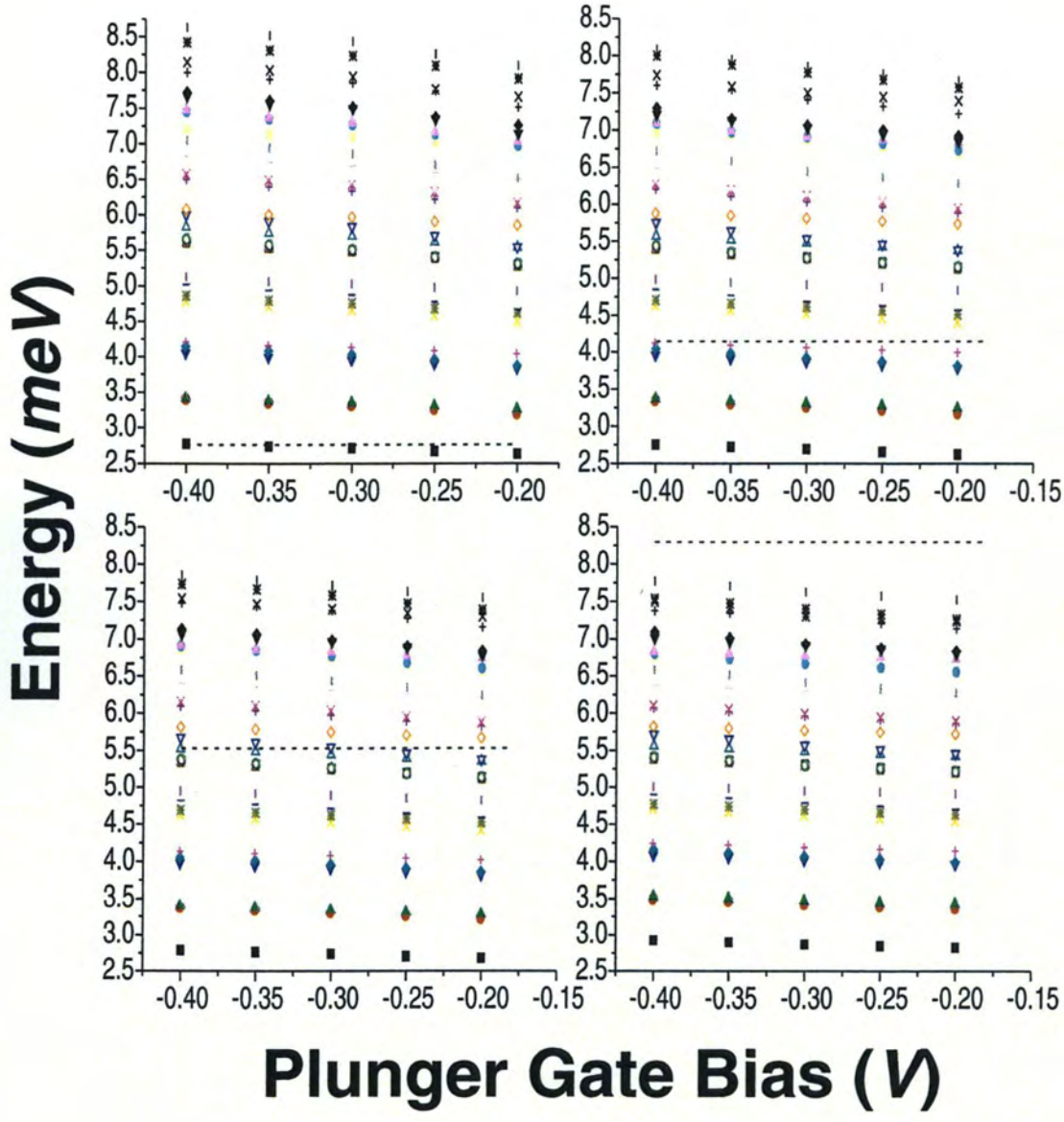

Figure 5.7: Plots of the first 30 energy eigenvalues for an asymmetrically biased symmetric quantum dot for various values of plunger gate bias. The inversion gate bias goes from right-to-left and top-to-bottom panes as 6.0 V, 6.5 V, 7.0 V, and 8.0 V.

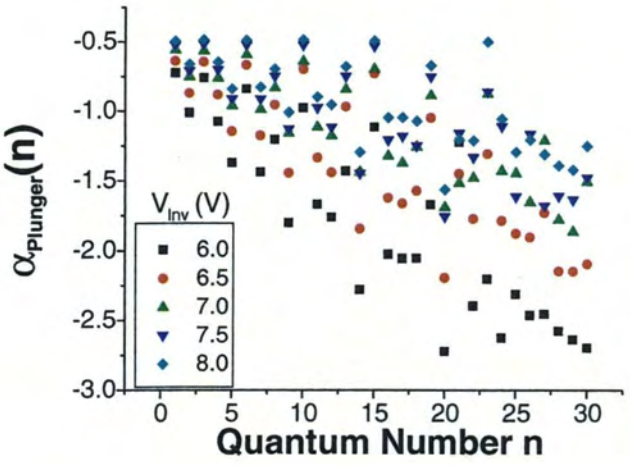

Figure 5.8: A plot of the slope α_{Plunger} as a function of quantum number n for various inversion gate biases for the asymmetrically biased symmetric quantum dot.

spaced. Another important feature is that the degeneracy structure, which was so apparent in the symmetric case, disappears, particularly as the inversion gate bias is increased.

5.4.2 Discussion

An important issue analyzed in the last section (Section 5.3) was the dot size. From the calculations of the background base potential profile, it appears that there is a lateral depletion from the depletion gates on the order of 50 nm in the symmetric devices. The calculations of electron wave function shows that there is considerable variation with quantum number, but that for the ground state the size is roughly 60 nm. This is in good agreement with the size calculated in Table 5.2.

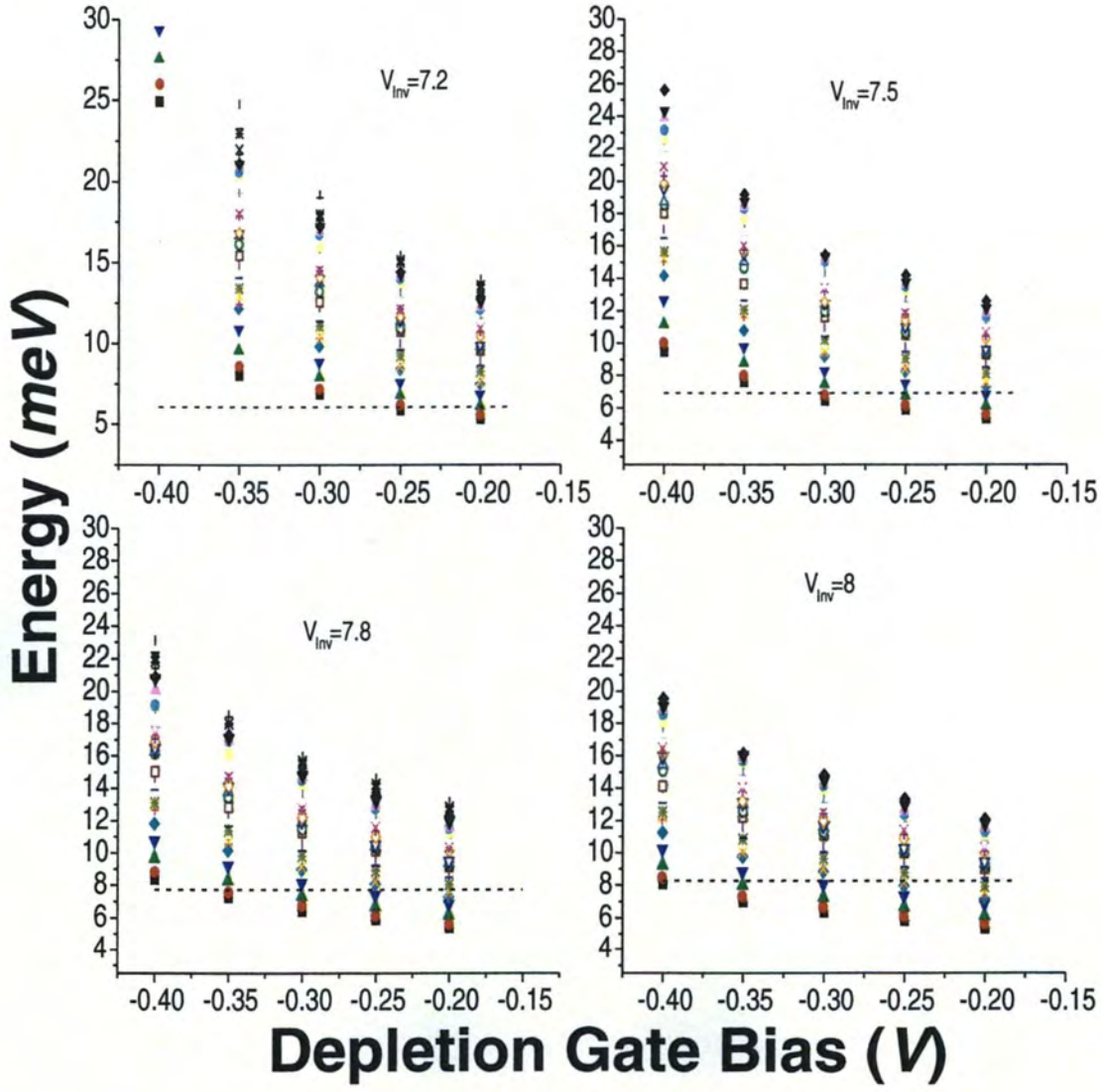

Figure 5.9: The energy level structure in a dot with asymmetric depletion gate geometry for various depletion gate biases. Each pane is for a different fixed inversion gate bias (which has been labeled); the position of the Fermi level in the channel is given by the horizontal dashed line.

5.5 Addition Energy

As discussed in Section 2.5.1, multi-electron quantum dots show an energy level structure which has features analogous to atomic spectra when the dots are small enough and are occupied by a small number of electrons. In Section 2.6.3, in the context of few electron dots and artificial atoms, it was discussed how a shell structure appeared to be observed in the tunneling conductance peak periodicity. Such structure is often discussed using the concept of an *addition energy*, ΔE, which is simply defined in terms of the incremental change in energy between two consecutive tunneling peaks, or

$$\Delta E \equiv E(n) - E(n-1), \tag{5.16}$$

where the argument n can be considered the principal quantum number, electron number N occupying the dot, *etc.* In other words, the addition energy is the energy needed to fill the next available state in the ladder of allowed energy levels given that the state preceding it is already filled. From this discussion, it is seen that it would also be appropriate to use the term *filling energy* interchangeably with the term *addition energy.* It is worth emphasizing that what is measured is the Fermi level in the channel at which the peak tunnel current is observed, so that in Eq. 5.16, $\Delta E = E_{\mathrm{F}}(n) - E_{\mathrm{F}}(n-1)$.

In this section, the addition energy spectrum will be presented from the experimental data for the quantum dots D9, D10, D11, and D12. As was discussed in Section 2.6.3, in tunneling transport through a quantum dot with fully discrete states, a peak in the conductance arises when the chemical potential in the channel (the leads of the quantum dot) aligns with an allowed state in the quantum dot, as was shown schematically in Fig. 2.11. In the following analysis, data from the inversion gate sweeps are used to generate an addition spectrum by using Eq. 5.6, the empirically derived relation between inversion gate voltage and 2D Fermi energy

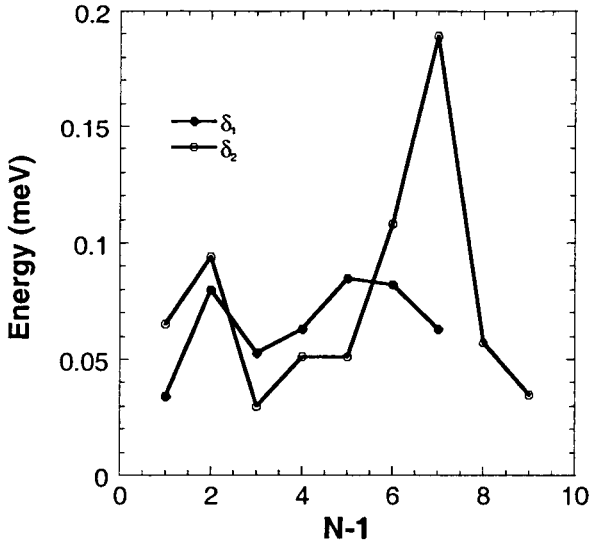

Figure 5.10: Filling energies for device D10, assuming that the two sets of peaks are independent of each other. These two sequences are labeled δ_1 and δ_2.

from the SdH data in Section 5.2, to correlate the inversion gate peak positions to the Fermi energy in the channel.

In this section these devices will be treated as if there are two independent addition spectra corresponding to two sets of "branches". As is discussed later in Section 5.9, several of the devices exhibit complex structure when the peak positions are plotted in the inversion gate-depletion gate bias plane. The reader is referred to that section for further discussion on this matter.

The addition spectrum for device D10 is shown in Fig. 5.10, as extracted from the data of Fig. 4.8 for a depletion gate bias of -100 mV, while the addition energy for device D9 is plotted in Fig. 5.11, as extracted from the data plotted in Fig. 4.11 for a similar sequence of depletion gate biases.

Figure 5.12 plots the addition energy for device D12 taken from the data presented in Fig. 4.17. For device D12, it is assumed that the fine structure is due to splitting of some degeneracy associated with the energy level, whose bare level

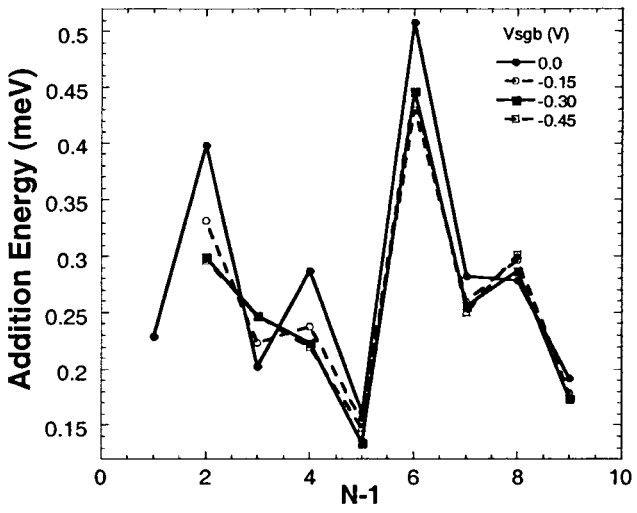

Figure 5.11: The addition energy spectra for several different bottom depletion gate voltages is plotted for device D9. There are local maxima in the spectra for filling number N=2, 4, and 6.

is located at the center of the major peak. Discussion of the meaning of this fine structure will be deferred to Section 5.6. Again, there appear to be two ladders of energy levels for device D12.

Figure 5.13 shows a plot of the addition spectrum for device D11 from the data presented in Fig. 4.26. For D11, there appear to be three ladders of energy levels.

5.5.1 Comparison of Mean Addition Energy

A simple comparison to make is that of the mean addition energy, which is essentially the same value as the mean spacing summarized in Table 5.2 in the first row. As one would intuitively expect, the addition energy is greater for the smaller devices (D9, D12, and D11) and much smaller for the larger device (D10).

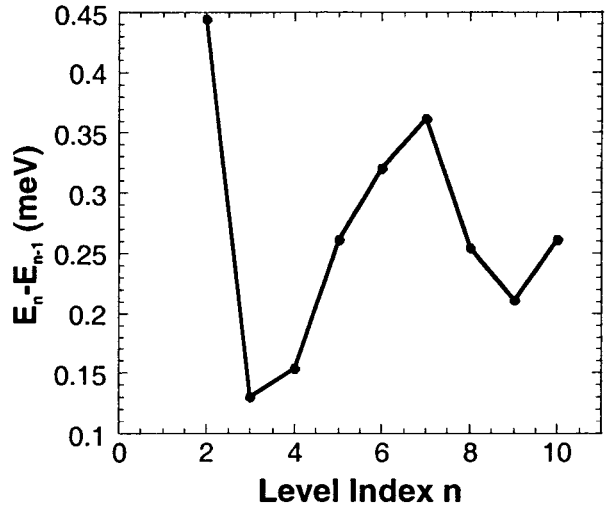

Figure 5.12: The addition energy spectrum for the ground state of the quantum dot D12, calculated assuming that the ground state energy is at the location of the central, dominant peaks. The depletion gate bias for these measurements is zero.

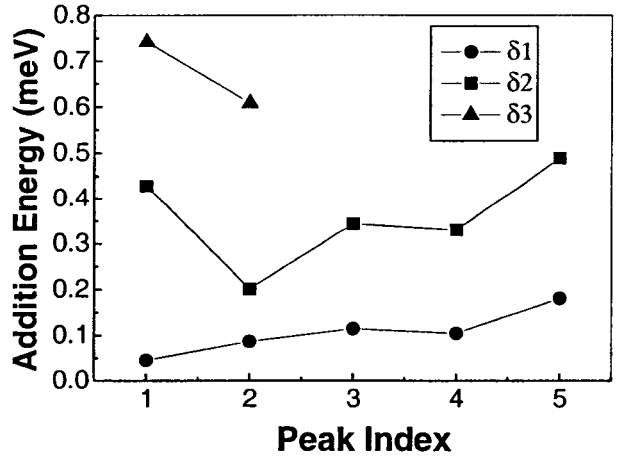

Figure 5.13: An estimate of the addition energies for each of the sets of branches 1, 2, and 3 plotted for device D11. The depletion gates have been biased at 0 V.

5.5.2 Comparison to Theory

A comparison to the theory of *artificial atoms* discussed in Section 2.6.3 is not necessarily straight-forward for the experimental data from this work for the following reason: in the experiments reported here, the Fermi energy in the channel does not remain fixed as the inversion gate is swept, while in the results discussed in the literature the Fermi energy should remain fixed. For the same reason, it is not necessarily straight-forward to compare the calculated energy levels with the addition energy structure reported in the literature; changing the inversion gate bias clearly leads to large variations in the bare level structure in the dot, as was evidenced by the plots for α_{Inv} and α_{Depl} in Figs. 5.5 and 5.6. What is observed in Figs. 5.10 through 5.13 is that only in Fig. 5.11 does there appear to be a shell structure similar to the literature; it is this apparent shell structure which will be considered first.

While addition energy in Fig. 4.11 is aperiodic, there are some features that have been observed in other systems [134], and have been predicted in the literature in conjunction with the theory of semiconductor *artificial atoms* [76, 82, 86]. In these previous works, it was found that there are special filling numbers for which the addition energy gives local maxima. These filling numbers have been attributed to filled electron shells, and correspond to: $N = \{2, 6, 12, \ldots\}$ with minor peaks at $N = \{4, 9\}$. Ezaki *et al.* [76, 82] showed through a theoretical study in which the exact diagonalization of the few particle Hamiltonian (as discussed in Section 2.5.1) that these minor peaks were in fact special to circular dots, and should not be seen in dots with square geometry (or at least should be suppressed).

While the effects of some shell structure for a circularly shaped dot potential is one possible interpretation of the data, another is that of a double-dot structure, which was studied theoretically by Matveev *et al.* [173, 174, 175] and experimentally by Waugh *et al.* [176] for intentionally formed double quantum dots. The theoretical

work found that if there were two dots coupled capacitively to each other and the reservoirs on either side by tunnel junctions, and the coupling between the two quantum dots and the channel were asymmetric, then the Coulomb blockade peaks would be grouped together in pairs, each pair separated by some small amount and with a nearly equal pair splitting $X_+ - X_-$ (defined below) which grows as the inter-dot coupling G_0 increases. The peaks were shown to occur when the charging energy on the two dot system satisfied the simple relation

$$E_{\mathcal{N}}(X^*) = E_{\mathcal{N}+1}(X^*), \quad \mathcal{N} = N_1 + N_2, \tag{5.17}$$

where, as usual $\{N_1, N_2\}$ are the number of electrons occupying dots $\{1, 2\}$. In their model, the peak sequences are related to the system capacitances by the following:

$$X_1^* = \frac{n + 1/2}{1 + \alpha/2} \quad X_2^* = \frac{n + 1/2}{1 - \alpha/2}, \; n = 0, 1, 2, \ldots, \tag{5.18}$$

where the asymmetry is defined by the capacitances:

$$\alpha = 2\left(\lambda + \frac{C_1 - C_2}{C1 + C2}\right) \tag{5.19}$$

$$\lambda = \frac{(C_2 + C_5 - C_1 - C_4)}{\sum_{i=1}^{5} C_i}. \tag{5.20}$$

While device D9 was not purposely fabricated to have this double dot structure, it is possible that there was an accidental dot formed by an ionized impurity or imperfections in the lithography. Such an accidental dot has been attributed to CB phenomena in other cases, such as work reported by F. Scott-Thomas *et al.* [17] in Si, or the work by the von Klitzing group. The accidental dot idea is illustrated schematically in Fig. 5.14, including the equivalent capacitive circuit used in Eq. 5.20. Figure 5.15 shows a plot of the peak positions for four successive (increasingly negative) bottom depletion gate biases are plotted versus energy (calibrated from the inversion gate bias) for various values of the bottom depletion gate bias. It appears that each sequence is the evolution of the previous

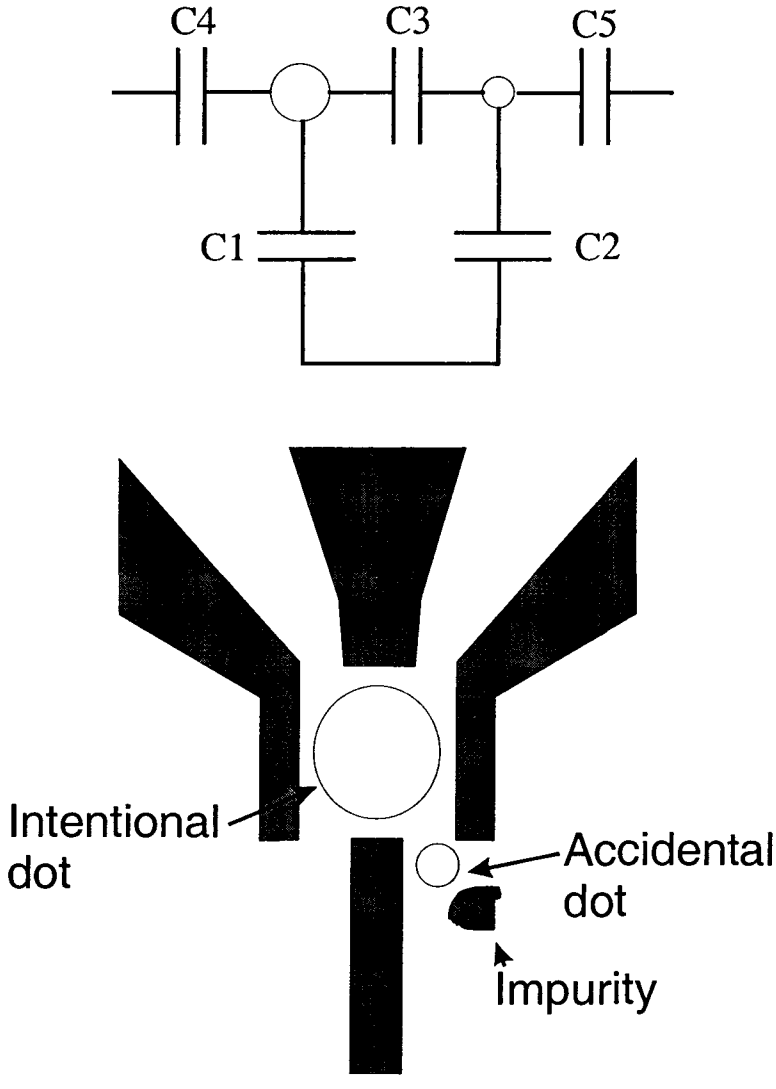

Figure 5.14: Top: The equivalent capacitance circuit for a double dot quantum dot structure. Bottom: a schematic of a double dot system accidentally formed by an impurity located near one of the point contacts defining the dot.

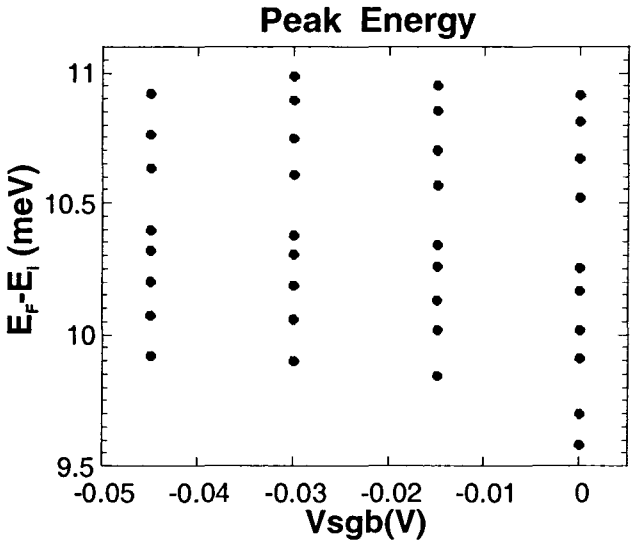

Figure 5.15: Positions of the peaks in inversion gate bias for different values of the bottom depletion gate voltage.

sequence, where the first peak in each sequence after the 0.0 V sequence is indistinguishable from the background, possibly suppressed due to oxide charging effects (recall that $I_d \propto (V_{\text{Inv}} - V_{\text{T}})$). Labeling the centroid of an apparent peak pair, $\bar{X} = (V_{\text{Inv}}(N) + V_{\text{Inv}}(N+1))/2$ and $\Delta X = \bar{X_N} - \bar{X_{N-1}}$, as the spacings between the centroids of successive pairs, the peak splittings ΔX in Fig. 5.15 have been plotted against the centroids of the pair positions $\bar{X}$ in Fig. 5.16. As predicted in the theory, one should expect that since the conductance is monotonically increasing with increasing gate bias, the peak splitting should increase monotonically as well. This is clearly not the case for this range of investigated inversion gate bias. In fact, from Fig. 5.15, it appears that the peaks occur in groups of 2, 4, and 4, meaning the shell structure model would be a much better explanation for the data.

The fact that devices D10, D11, and D12 did not exhibit this apparent shell structure does not mean that the role of electron-electron interactions is not important. Recently, Austing *et al.* [177] investigated ellipsoidally deformed vertical

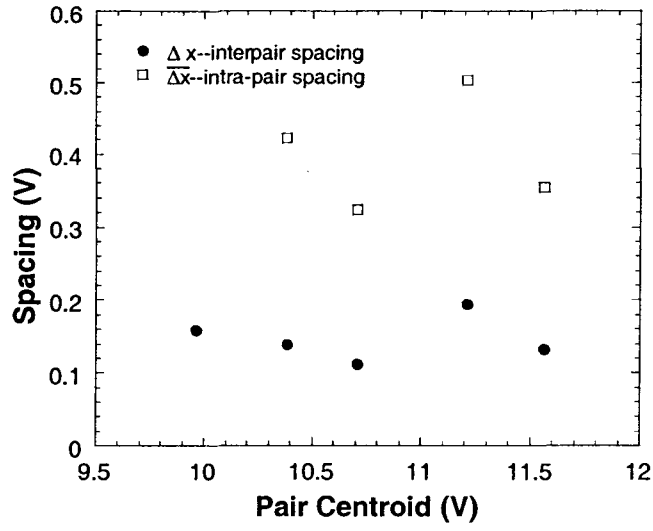

Figure 5.16: Inter-pair splitting (open squares) and intra-pair splitting (filled circles) plotted against the centroids of the pairs.

quantum dots of the same type as those reported by Tarucha *et al.* [134]. In their investigation, the symmetry was perturbed and the dot energy level structure was calculated within the spin density functional formalism, where it was found that by removing the rotational symmetry of the cylindrically symmetric dots by gradual deformation of the shape in any lateral direction, the shell structure was gradually removed. As is further discussed in Section 5.9, it is suspected that this deformation actually gives rise to several distinct spatial axes of energy quantization due to confinement in the plane.

5.6 Peak Shape Characterization

Analysis of the peak shapes can lead to useful information about the quantum dot. In particular, if pure Coulomb blockade transport is assumed, one can make fits to the peak shape using Eq. 2.85 to extract the total capacitance C_Σ and the capacitance of the gate being swept, C_G. These values can then be compared to the

capacitance obtained by the peak spacing using Eq. 2.84 assuming discrete energy level term makes no contribution to the peak period with gate bias.

In the quantum dots investigated in this study, the shape of the inversion gate sweep peaks is much more complicated than the depletion gate sweeps, and while they are important and will be discussed, they do not fit the Beenakker formula (Eq. 2.85). In contrast, for devices D11 and D12 the depletion gate sweep peaks do fit nicely to the Beenakker formula for low inversion gate bias. These results will be discussed first.

5.6.1 Depletion Gate Peak Shape

The peak lineshapes were fit using Eq. 2.85 using a least squares routine. Figure 5.17 plots the peak fits as a function of depletion gate bias, where all depletion gates are tied to the same bias (experimental data points are shown as circles). The fits assume the temperature is constant but allow the capacitances to vary with gate bias. Also, since the baseline current of the peaks varies from peak to peak, an additional parameter is included in the fit. The values of the capacitance obtained from the fit are then compared to values obtained from the peak spacing using Eq. 2.84 in Table 5.3. Again, it is important to point out that these fits assume that transport is purely Coulomb blockade, meaning in particular that the peak periodicity in Eq. 2.84 is simplified to be

$$\Delta V_{GS} = \frac{e}{C_G}. \tag{5.21}$$

In Fig. 5.17, only the larger peaks are fit, and smaller peaks or shoulders are indicated by arrows. Also, in Table 5.3, columns labeled $C|_{\text{fit}}$ refer to the results of the fit using Eq. 2.85 while those labeled $C|_{\text{exp}}$ are obtained from the peak periodicity using Eq. 5.21. From the plunger depletion gate sweep of device D12, the capacitance of the plunger gate has been fit using Eq. 2.85, and the capacitance values have been compared to Eq. 5.21 for the peak spacing. These results are summarized

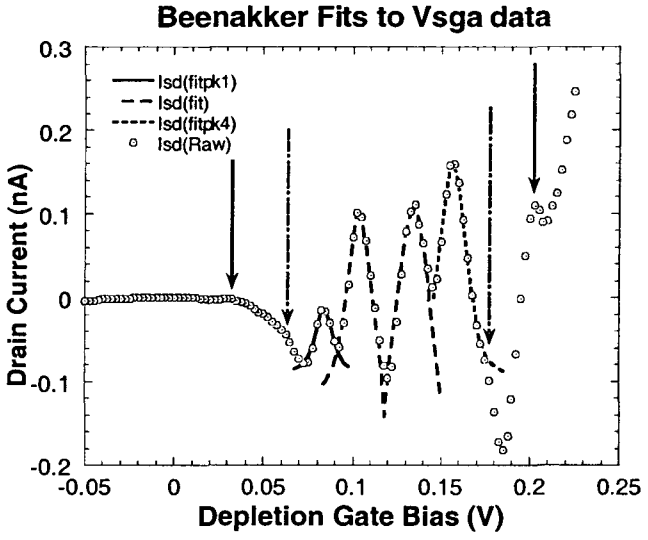

Figure 5.17: Current peaks from device D12 fitted to the Beenakker formula for CB oscillations (Eq. 2.85) for the case of all depletion gates tied. Arrows denote the position of shoulders or smaller peaks, which are used in the experimental calculation of the capacitance. Broken arrows denote shoulders.

Table 5.3: Fit and experimental values of capacitance for all depletion gates tied and swept for device D12.

peak #	$C_G\|_{\text{fit}}$ (aF)	$C_G\|_{\text{exp}}$ (aF)	$C_\Sigma\|_{\text{fit}}$ (aF)
1	3.355	5.41	23.48
2	7.802	8.21	76.57
3	4.255	5.18	76.57
4	5.408		60.2

Table 5.4: Fit and experimental values of capacitance for the plunger gate swept for device D12.

peak #	$C_G\|_{\text{fit}}\ (aF)$	$C_G\|_{\text{exp}}\ (aF)$	$C_\Sigma\|_{\text{fit}}\ (aF)$
1	2.865	3.15	74.03
2	3.865	3.77	48.1
3	1.646	5.93	28.59
4	2.24	4.035	70.27

in Table 5.4 Peaks from the depletion gate sweep for device D11 have also been fit in the same manner. In this case, however, the negative leakage current was not observed in the sweeps, allowing for the elimination of one of the fit parameters. The capacitance is estimated from the period of these peaks using Eq. 5.21 to give $C_{\text{SGA}} = 2.28$ and $2.14\ aF$ for the first and second peak pairs, respectively. Note that this is about 1/3 of the value for device D12, but is not completely surprising due to the differing geometry of the depletion gates. Also note that the peak spacing is a strong function of the inversion gate bias, with the associated capacitance increasing linearly with the gate bias.

A peak from a device D11 depletion gate sweep has been fit with the Eq. 2.85, and an estimate of the algebraic sum of capacitances (C_Σ) found from the fit. The fit along with the original data are shown in Fig. 5.18. The value found for the capacitance is $C_\Sigma \approx 16.97\,aF$.

It is important to note that that the Beenakker formula is not the only one which fits a line of this shape; it is also well fit by a Lorentzian of the form of

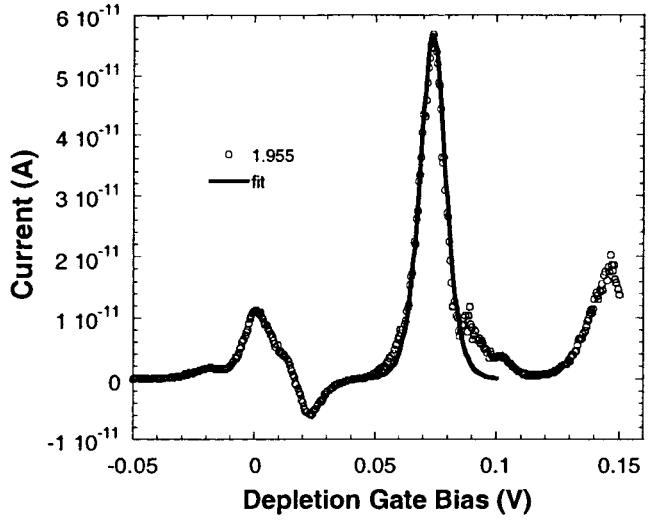

Figure 5.18: One peak in a depletion gate sweep of device D11 for an inversion gate bias of 1.955 V is fit using the Beenakker formula. For this fit, the depletion gate capacitance is chosen and total capacitance is the only fit parameter.

the Breit-Wigner formula for resonant tunneling, which was originally derived to describe tunneling in the nuclear theory of the decay of resonant states [178, 179]. The expression for the transmission coefficient in the Breit-Wigner formula is

$$T(E) \approx T_{\text{res}} \frac{\Gamma_n^2/4}{\Gamma_n^2/4 + (E - E_n)^2} \quad \text{where } (E - E_n)^2 = \left[\left(\frac{eC_g}{C_\Sigma}\right) - E_n\right]^2 \qquad (5.22)$$

where Γ_n is the full width at half maximum amplitude.

One should question why the effective capacitance, C_Σ should vary so greatly in Table 5.3 and 5.4, and that there is such large differences in the capacitance C_{sga}. This is evidence of the importance of other energy levels in the dot, as is the fact that there is such large variation from conductance peak to conductance peak. A body of evidence is beginning to come together which suggests that the charging energy is not the dominant energy in these dots, but rather, the electronic confinement energy of the dot itself.

5.6.2 Inversion Gate Peak Shape

Due to the nature the inversion gate sweeps, it is not expected that their shape should fit the Beenakker formula (Eq. 2.85), which assumes that the chemical potential in the leads (the channel on the source and drain in this case) are fixed as the gate is swept. However, in this case sweeping the inversion gate bias amounts to sweeping the chemical potential in the channel. In addition, it was shown in Section 5.4 that changing the inversion gate bias changes the position of the dot energy eigenvalues relative to the chemical potential in the channel. Thus, the shape of the conductance peaks can not be predicted by the Beenakker formula. This can in fact be verified for all devices by noting the asymmetry of peak shapes in Figs. 4.8, 4.11, 4.17, and 4.26 in Chapter 4; it is reasonable to expect a skewness in the shape of the peaks due to this coupling of the inversion gate bias to the channel chemical potential and the quantum dot energy levels with different strengths. The remainder of this section will be dedicated to discussing the complex nature of the conductance peaks in the inversion gate sweep of device D12 shown in Fig. 5.20 in the upper pane. These sweeps show what appears to be large central conductance peaks which each have their own fine peak structure superimposed. It will be discussed below how such an apparent structure can be observed for two different fundamental types of conductance behavior in a measurement at finite temperature (which induces thermal broadening of energy levels) or finite source-drain bias excitation (which could lead to transport through multiple channels at once), or both.

One possible peak behavior, which could be argued to model the experimental data, is a sequence of individual conductance peaks modulated by an envelope function, which are spaced closely enough that thermal broadening causes merging of the tails of peaks. Sequences of peaks modulated by envelope functions have been predicted in theory and observed experimentally. Stopa [180, 83] used Eq. 2.83

derived by Beenakker [114] for N-electron dots to calculate the conductance for a device structure with similar depletion gate geometry as those used in this work. He found that the conductance oscillations indeed appeared to follow an envelope function, with the relative transmission rates and level spacings changing with electron number. Such peaks have been experimentally observed in GaAs/AlGaAs dots by Staring *et al.* [181] and Heinzel *et al.* [182] in experiments where the quantum dot was strongly coupled to the leads (i.e. the transmission coefficient was on the order $0.1 < T < 1$) (Note: the condition of strongly coupled dots is usually presented as $G \approx G_0 = 2e^2/h$). They observed that the envelope function vanished with increasing magnetic field. This envelope function was also seen in dots weakly coupled to the leads by Sakamoto *et al.* [183]. They also observed an envelope function in the presence of weak magnetic field; in moderate magnetic fields they found that the number of CB oscillations per period decreased as the gate voltage became more negative. From this observation they estimated the confining energy of the dot. This possibility does not agree well with the data to be presented in Section 5.7.

On the other hand, a series of peaks with fine structure in each peak has also been observed in the inversion gate sweeps of devices D11 and D12, which will be shown to lead to two possible interpretations in terms of transport behavior. Such line shapes have been attributed to the transport properties of dots with a discrete confinement energy spectrum, and the excitation spectrum in particular. Individual peaks with a fine structure were observed by Johnson *et al.* [123] and Foxman *et al.* [124] in GaAs dots. In their experiments, the drain bias was held fixed and a gate bias (in these experiments a plunger gate) was swept. At low drain bias, there was a sequence of individual CB conductance peaks. For increasing drain bias, a fine structure evolved in each CB peak, with the amplitude of each fine peak decreasing with increasing gate voltage. This fine structure was interpreted in the following way: as the drain bias is increased, the window of excitation energy through which

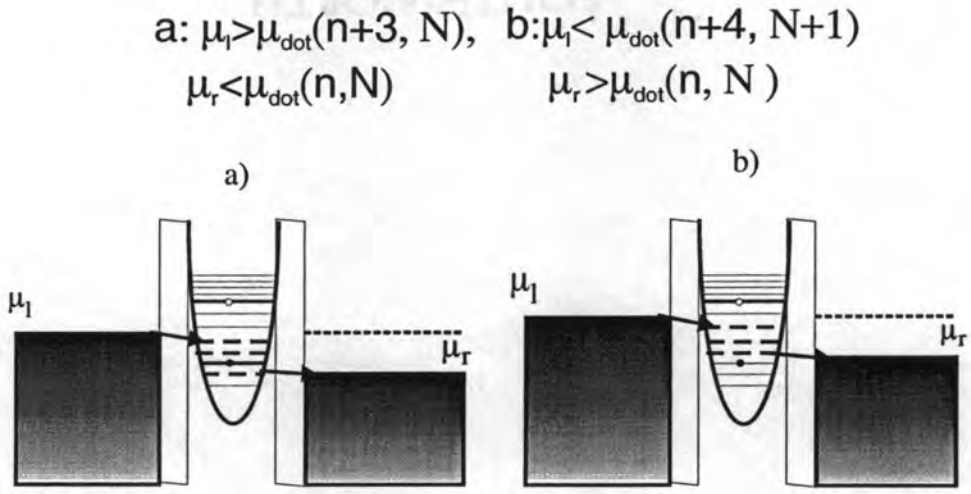

Figure 5.19: a schematic of the energy spectrum is shown when the Coulomb and discrete energy level spacing is appreciable. With the drain bias excitation greater than the energy level spacings, more than one tunneling channel is allowed. Active channels are depicted by thick lines, and inaccessible channels are shown as thin dotted lines.

tunneling may occur is increased. If the spacing between energetically accessible states is narrow enough, more than one tunneling channel may be accessible within this window. As the number of channels into and out of the dot changes, the conductance changes with it. For these experiments, the channels are interpreted as being the excited dot states. This idea is presented schematically in Fig. 5.19, where the active transport channels are represented by thick lines, while thin lines represent inactive channels which are inaccessible to electrons for tunneling. The ground state is represented with a circle, which is filled when the dot state is filled and open when the level is inaccessible for tunneling. In this picture, the Coulomb charging energy is much greater than the confinement energy spacing.

A possible alternative theoretical interpretation leading to observation of fine structure is for the discrete energy level spacing due to confinement in the dot $E(n+1) - E(n)$ to dominate the Coulomb charging energy $e^2/2C$, where n is some principal quantum number other than the dot electron occupation. In this model for

transport, the position of the central peaks is controlled by the confinement energy. The discrete levels have some degeneracy associated with, for example, spin and the valley symmetry within the Si band structure, as well as the degeneracy of the levels within the dot itself (which were observed in the calculated level structure plotted in Fig. 5.3). For the two types of degeneracy associated with the parent semiconductor Si, the degeneracy $g_s \times g_v$ would be 2×2 for the Δ_2 sub-band and 2×4 for the Δ_4 sub-band; however, it is not clear that splitting of this degeneracy would be observable. On the other hand, the calculations presented in Section 5.4 suggest that the splitting of degenerate dot energy levels should be appreciable. Due to the discrete nature of charge, the Coulomb charging energy leads to a degeneracy splitting which results in an observable fine structure. The degeneracy structure observed in these calculations is consistent with the degeneracy of a harmonic potential. Yet another mechanism for degeneracy splitting is through the spin interaction term in the single particle Hamiltonian, Eq. 2.33; discussion of this possibility will be deferred to Section 2.3.3.

Yet another cause for degeneracy breaking was recently simulated for the quantum dot structures reported on in this work by Milicic, *et al* [184]. In their work, the background potential calculated as described in Section 5.4 was modified after the initial self-consistent Poisson solver by adding a randomly distributed number of particles to simulate an actual ion implantion profile of the dot. It was observed that the neatly evolving group structure, which appeared to follow the splitting of the harmonic oscillator level degeneracies, was modified so that some members of the grouping of $g = n + 1$ split levels were additionally split by some small, apparently random energy. These results agree very well with the experimental fine structure observed in the inversion gate sweeps of devices D11 and D12.

Whatever the degeneracy breaking interaction, the shape of the current peaks is governed by a partial sum of the conductance through all of the tunneling channels

accessible in the energy window. This degeneracy breaking interpretation of the transport mechanism is shown schematically in Fig. 5.20.

We can directly address several of these possibilities by considering the experimental data. In particular, the explanation of the fine structure in terms of tunneling through additional channels in the form of multiple excited states is eliminated as a possibility. Examining the data presented in Section 4.4.4, Fig. 4.23 clearly shows that for increasing drain bias, which increases the excitation energy window, there are the same number of fine peaks observable up to the highest drain bias measured. This observation is in direct disagreement with the theory.

The data from Section 4.4.4 does not conclusively eliminate the first explanation presented, which says that this line shape is due to overlap of the tails of several individual peaks. The data is inconclusive because of the drain-bias and system temperature related energy compared to the fine structure splitting, which is on the order of $\delta E = 40\,\mu$eV. While the drain bias is smaller than the thermal equilibrium energy, both are greater than this splitting: $V_D|_{\min} \approx 100\,\mu$V, $k_B T \approx 362\,\mu\text{eV} > \delta E$. This means that multiple, independent channels could be accessible, and they could be thermally broadened such that the tails smear into each other. This leaves at least two possible interpretations for the line shape in the inversion gate sweeps.

It is strongly believed that the transport in these dots is dominated by the discrete levels caused by electonic confinement, by the arguments presented in Section 5.3, and supported by the calculations in section 5.4 and a simple model to be presented in the next section.

Before continuing to the analysis of the conductance peak positions in the inversion gate- depletion gate bias plane, one should assure him or herself that the fine structure is reproducible if the depletion gates are instead swept with the inversion gate held at a fixed bias. To check this, consider the plots in Fig. 5.21, which shows the results of an inversion gate sweep, and two sweeps generated by

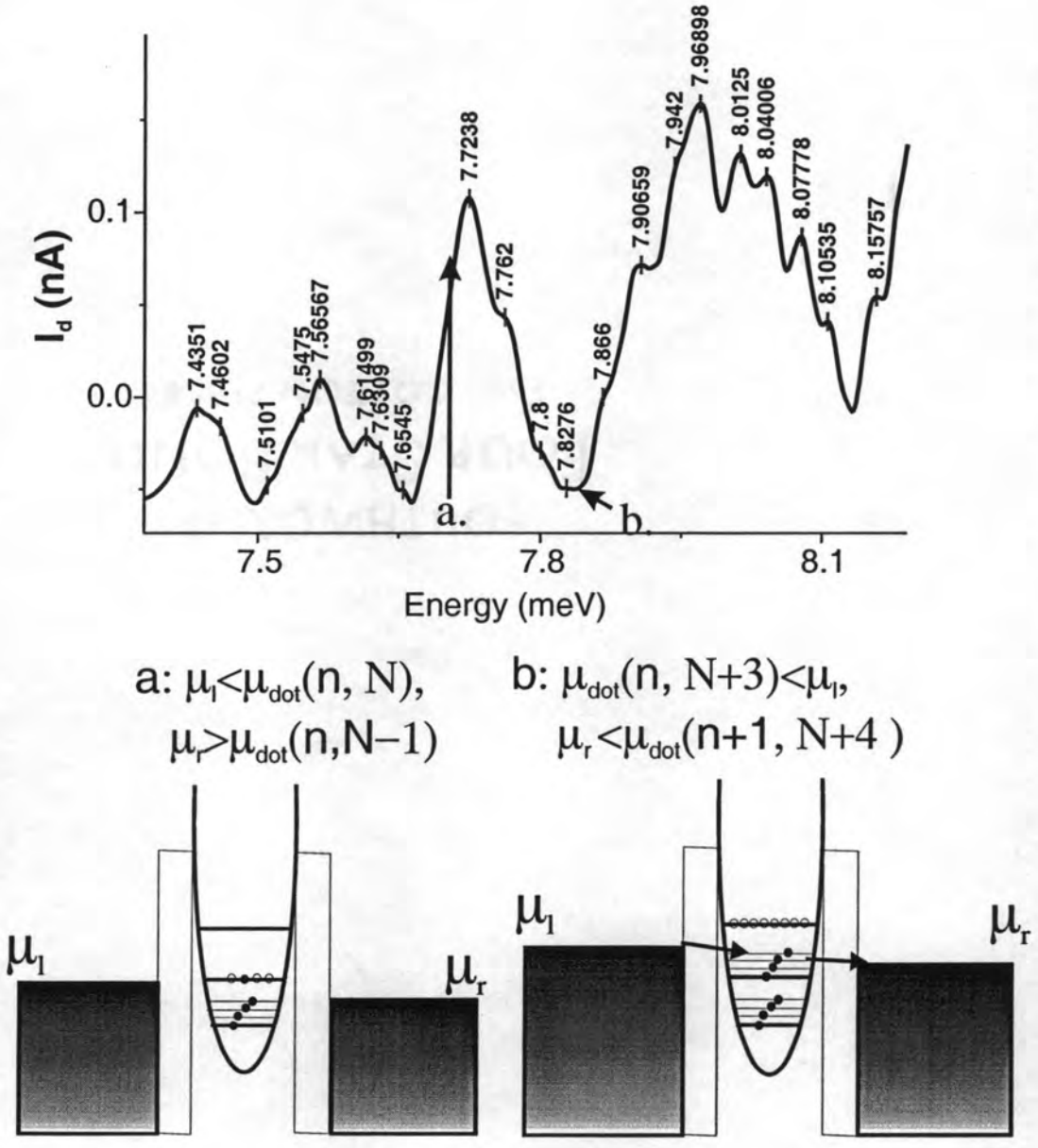

Figure 5.20: In the top pane, a portion of an inversion gate sweep is plotted for device D12. On the x-axis, the Fermi energy in the channel has been calculated in terms of the inversion gate bias . In the bottom panes, a schematic representing transport where the mechanism is transport through degenerate discrete levels. Point a (in the bottom left) represents the lead chemical potential just lining up with the unoccupied degenerate level. Point b (bottom right) shows this set of degenerate levels broken by some degeneracy splitting interaction.

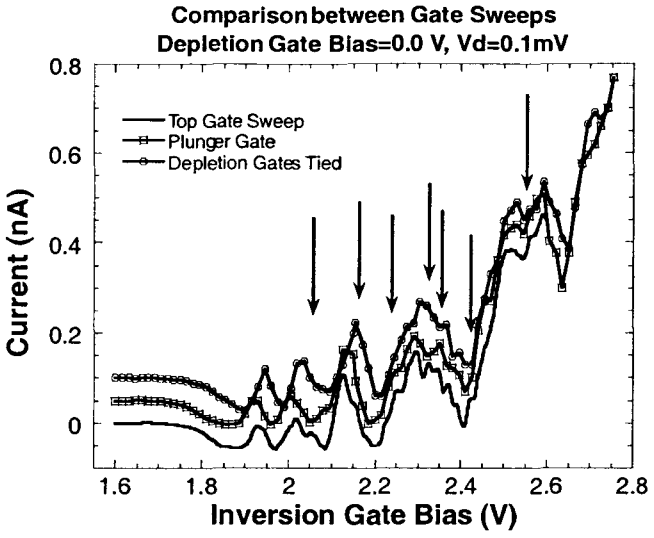

Figure 5.21: A comparison between the inversion gate sweep and depletion gate sweeps. The curves have been offset by 50 pA. Points where the fine structure are not seen in depletion gate sweeps are indicated by arrows.

taking the current values from sweeps of the plunger gate bias and of all depletion gates tied to the same bias at their zero crossing, for each value of the inversion gate bias. In Fig. 5.21, the two depletion gate sweeps have their current values offset by 5 pA from each other to better distinguish them. It is evident that while there is good agreement with the inversion gate sweep, some of the fine structure is missed by the generated sweep. At the same time, it gives clear evidence that the fine structure is present, and not due to any noise effect. However, it clearly shows that the data presented in the following section may in fact be only partially complete in terms of the conductance peak positions.

5.7 Magnetic Field Dependence of Peak Position

The evolution of the conductance peak positions in a magnetic field has recently proven to be a valuable probe of the shell structure of small, few-electron quantum dots, allowing for the determination of the orbital quantum states in the

conductance spectrum. Calculations have shown that the shell structure in few-electron quantum dots are expected to fill following Hund's rules for spin [185] even in the absence of magnetic fields.

Unfortunately, due to the somewhat fragile nature of the dots studied in this work, only one dot (D11) yielded any magnetoconductance data associated with the dot. Figure 5.22 plots a sequence of inversion gate bias sweeps (with symmetric depletion gate bias configuration, gates held to ground) for increasingly larger magnetic field. The change of the conductance peak position is very subtle. Before trying to analyze this behavior, it is useful to first consider the evolution of single, non-interacting quantum dot energy eigenvalues in a magnetic field, using the Darwin-Fock states described by Eq. 2.30. The first eight states have been plotted in Fig. 5.23 for a confinement energy of $250\,\mu$eV (roughly the same as device D12). It is not expected that there should be any crossings or large kinks in the evolution of any conductance peak until the fourth set of levels, and only at moderate field of $\sim 2.5\,T$. However, it should be noted that Eq. 2.30 is valid only for perfectly symmetric confining potential. Austing *et al* [186] did a systematic experimental study of elliptically deformed vertical quantum dots, and found that the dependence of conductance peak evolution is suppressed as the dot shape is made increasingly asymmetric. For a dot of roughly the same ratio of major to minor dot side length, their data showed no major bending in the conductance peak dependence until the 6th and 7th peaks. Because the cyclotron effective mass in Si is larger than that in GaAs, it is expected that in a similarly shaped Si dot the peak evolution should be correspondingly suppressed. The experimental findings were in good qualitative agreement with the theoretical results of Madhav and Chakraborty [187] for an analytic model of anisotropic dots in a magnetic field.

In fact, the observed evolution for the peak along the vertical line at $V_{\mathrm{Inv}} \approx 2.3\,\mathrm{V}$ in Fig. 5.22 shows a slight bending in qualitative agreement with the findings

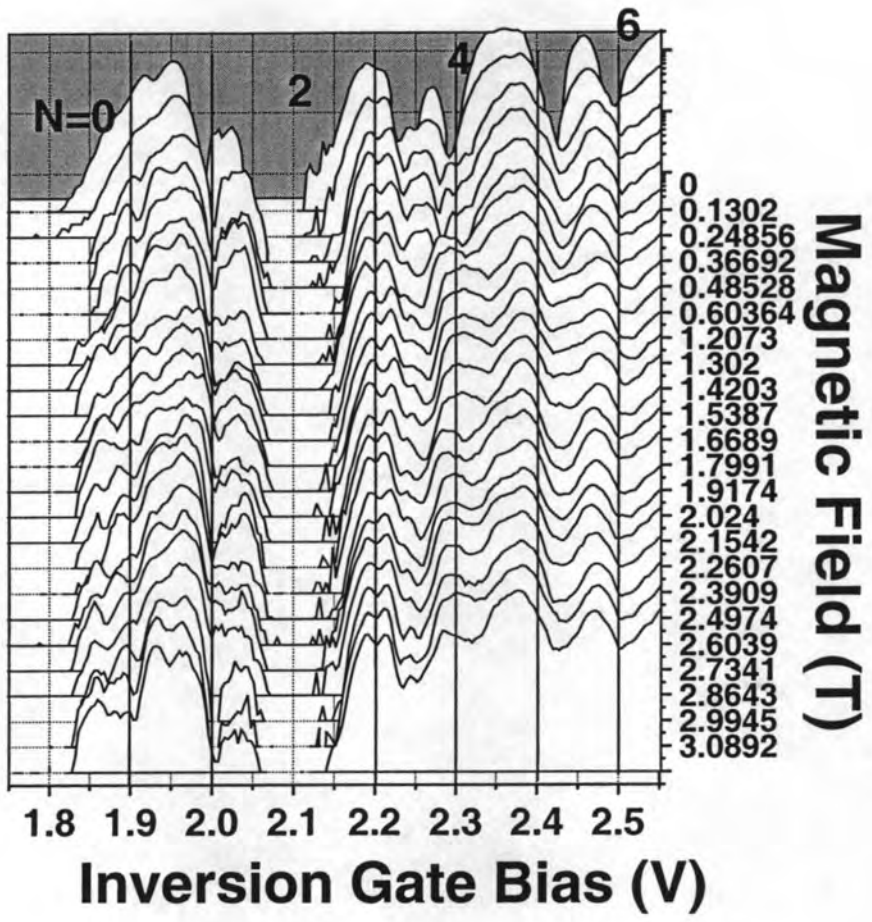

Figure 5.22: A semilog plot of a sequence of inversion gate sweeps for increasing magnetic field for the quantum dot D11, showing small changes in the conductance peak positions.

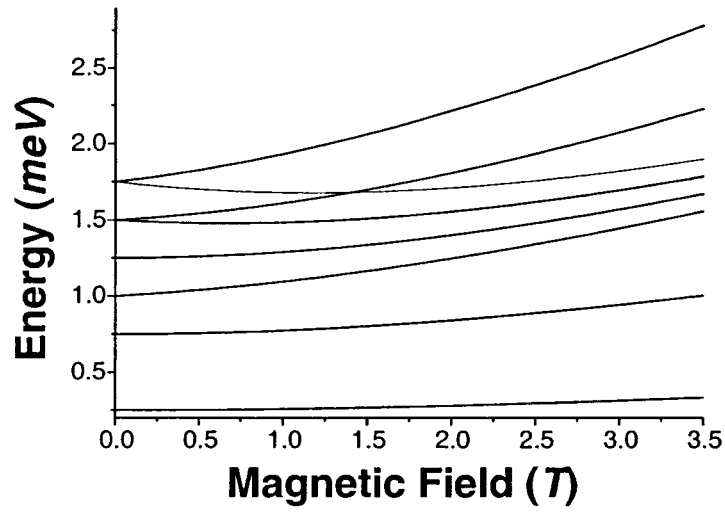

Figure 5.23: The magnetic field dependence of the single particle energy eigenvalues in a circularly symmetric harmonic confining potential.

of Austing *et al* [186]. The peak number of the first crossing observed in Fig. 5.23 is shifted only one position, and the relative shape qualitatively is similar to the experimental work for ellipsoidally deformed dots, appropriately suppressed due to the greater cyclotron mass.

From the results presented in this section, we can add an additional mechanism to the list of possible sources of splitting in the fine structure in the quantum dots D11 and D12. For sharply varying potential profiles, the term $\mathbf{S_i} \cdot [\mathbf{B}(\mathbf{r_i}) + \nabla \mathbf{V_0}(\mathbf{r_i})]$ in the single particle Hamiltonian (Eq. 2.33) could give rise to splitting due to the requirements of Hund's rules on spin filling of the electron shell. In particular, Hund's first rule states that electrons will be added to the shell in such an order that $\mathbf{S_z}$ is maximized, consistent with the Pauli exclusion principle. This effect could easily give rise to the fine structure observed in D11 and D12; unfortunately, with the absence of magnetoconductance data from device D12, it is impossible to confirm or exclude this possibility as the cause.

5.8 Effect of Finite Drain Bias

A brief discussion of the effects of finite drain bias on dot transport are given in this section. Mostly, the data for finite drain bias has served to rule out some of the interpretations of other data sets.

5.8.1 Drain Bias Sweeps

Sweeps of the drain bias for constant gate bias (both inversion gate and depletion gate) can give rise to steps in the drain current and oscillations in the differential conductance which are known and often-documented evidence for Coulomb Blockade and single-electron phenomena. However, observation of the steps (the so-called Coulomb staircase) is possible only when the barriers are asymmetric. For the asymmetrically biased device D9, there were no steps observed, but there were kinks in the $I - V$ curves (as shown in Fig. 4.15 and 4.16). The kinks observed in the lower inversion gate bias of Fig. 4.15 disappear as the inversion gate bias is increased. Note that the kinks, which could be associated with conductance through discrete states, disappear at higher inversion gate bias; this is consistent with the confinement energy decreasing as inversion gate bias increases.

Another sign of Coulomb blockade is the region of zero conductance about a range of source-drain bias near zero (the Coulomb gap), which is observed even when there is no Coulomb staircase. While there is a gap in the current at zero bias for device D9 (Fig. 4.15), there no gap for D11 and D12. The inset of Fig. 5.24 shows no gap in the current around zero V_{DS}; these facts together provide further evidence that the discrete states in the dot, rather than the Coulomb charging and Coulomb blockade.

These sorts of plots are presented as composites in the gate voltage-drain voltage plane, as in Fig. 5.25 and Fig. 5.26, which are plots of the current magnitude and the differential conductance versus gate bias, respectively. If transport is dominated

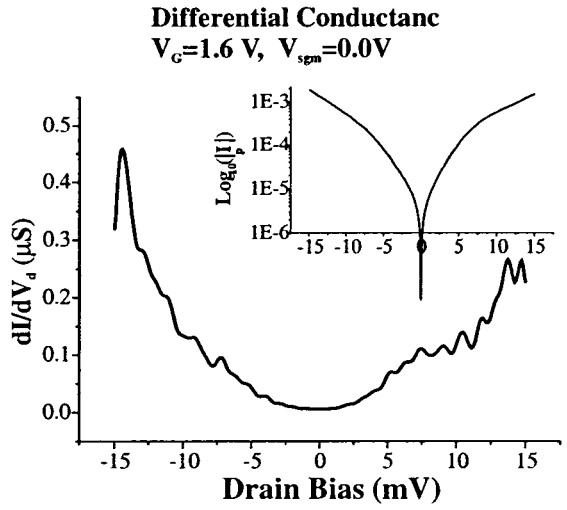

Figure 5.24: A plot of the differential conductance versus drain bias in device D12, showing peaks. The inset is a semi-log plot of the magnitude of the drain current for the same sweep showing a gap in the current about zero bias.

by pure Coulomb blockade, a diamond-like structure would be expected in Fig. 5.26, similar to that of Fig. 2.14 (although here we would have expected a skewed shape as sweeping the inversion gate should change both energy of electrons in the dot with respect to the Coulomb barrier and the energy of the electrons in the channel on either side of the barriers). However, no such structure is observed.

Plotting the differential conductance as a function of the drain bias can also yield information about the discrete energy level structure in the dot [123, 124]. Every time a discrete level in the dot aligns with the chemical potential in the biased lead (drain), a peak occurs in the differential conductance. The differential conductance for device D12 as a function of drain bias is shown in Fig. 5.24 for fixed gate biases. There are clear peaks in the differential conductance, re-enforcing other evidence presented that the discrete states in the dot in fact have a strong influence on the transport. The inset of Fig. 5.24 shows a semi-log plot of the current. The differential conductance versus drain bias for several different values of the plunger

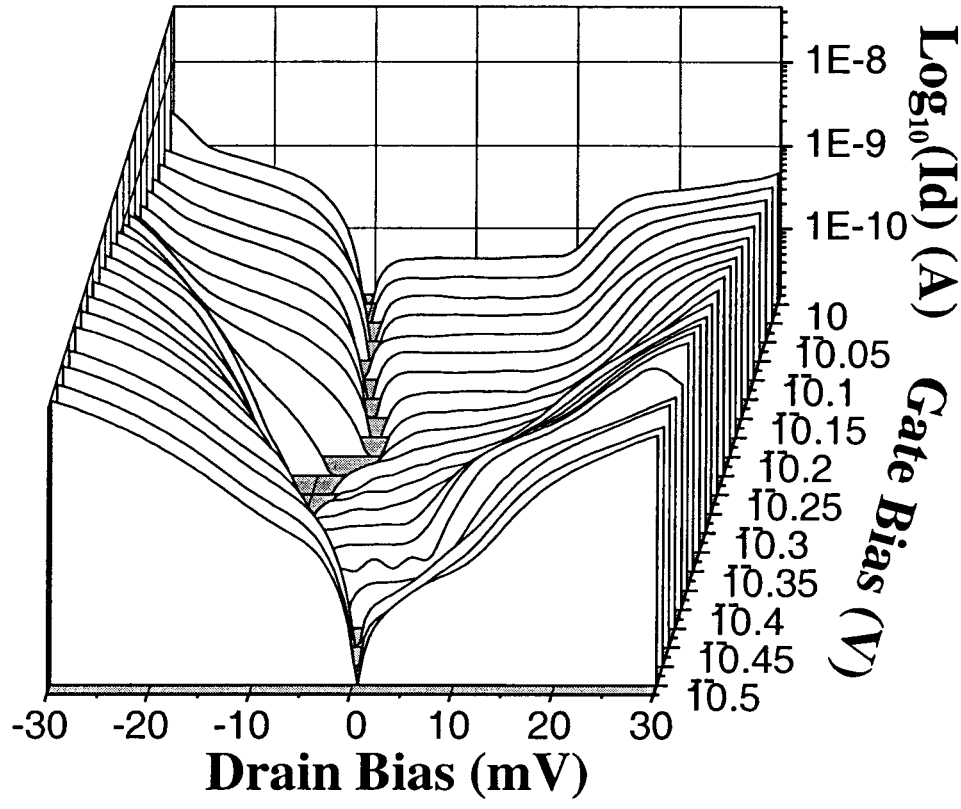

Figure 5.25: Drain current magnitude plotted versus drain bias for different values of the inversion gate bias on a semi-log scale. it can be seen that there are two flat regions of the current, and the gap about zero bias shifts in bias position as the gate bias increases.

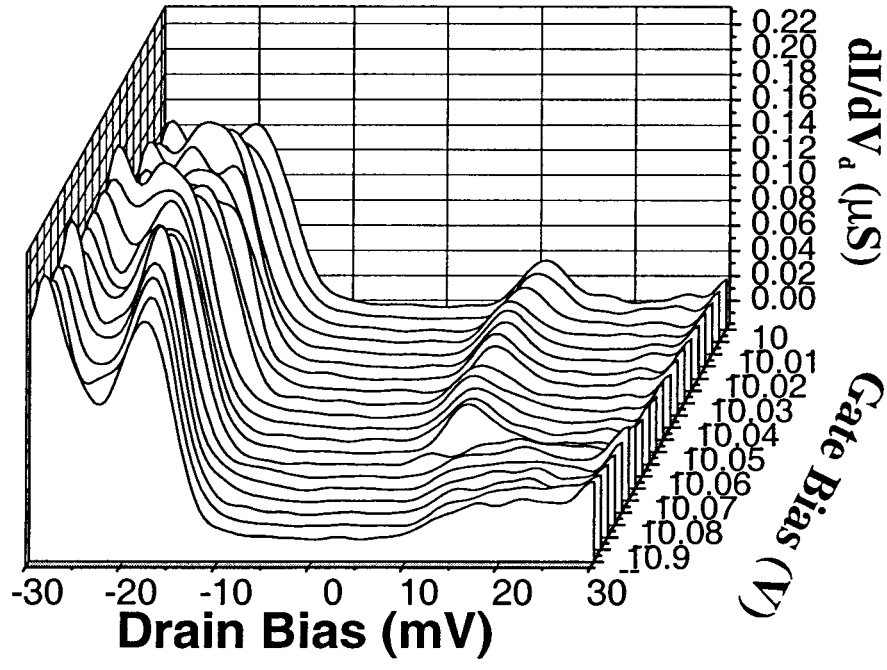

Figure 5.26: A plot of the differential conductance versus drain bias for a series of inversion gate biases. A peak separates the central minimum from a second valley, which could represent a second region of Coulomb blockade.

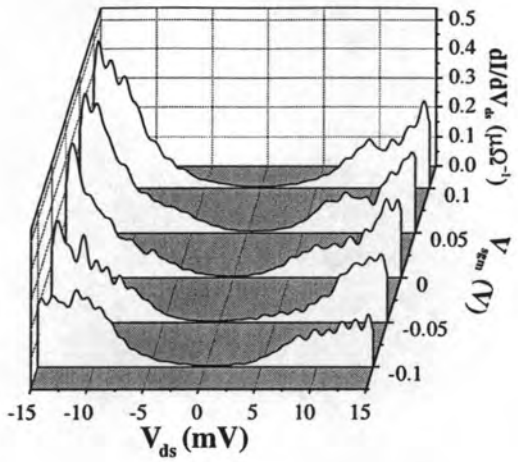

Figure 5.27: The differential conductance for several values of the plunger gate bias.

gate bias is shown in Fig. 5.27 for 1.6 V on the inversion gate. The density and range of these sweeps is too wide and short to support a composite plot showing the diamond structure in the Coulomb blockade regime. However, with the absence of other CB features, it is not expected that the modified diamond pattern similar to that in Fig. 2.15 would appear with more dense sweeps or greater range.

5.8.2 Stepped Finite Drain Bias

Stepping the drain bias and sweeping the gate bias yields information about how increasing the drain bias (and thus the tunneling energy of the injected electrons) changes the tunneling structure in the dot, and another measure of the confinement energy.

The meandering of the current minimum in the drain bias-gate bias plane shown in Fig. 5.25 is a feature common to all of the devices measured. It was also

observed in the data shown in Figs. 4.9, 4.10, and 4.25 for devices D10 and D12. This zigzagging of the minimum may have an analogy to the Coulomb blockade diamond structure.

The finite bias sweep may also be useful in determining the capacitance of the biased lead in quantum dots exhibiting Coulomb blockade oscillations; this parameter is often neglected in calculations of dot size, as it was in this work. The capacitance of the biased lead can be determined from the slope of the peak position in the drain bias-gate bias plane and the total capacitance by the relation [188]

$$\frac{dV_{\rm ds}}{dV_g} = \frac{C_g}{C_\Sigma - C_{\rm Bias}} \tag{5.23}$$

For the dots studied in this work, this slope is negative, so the value of capacitance obtained in this manner is unphysical, and thus was neglected.

As the bias is increased, features in the inversion gate and depletion gate sweeps become less distinct as the carrier energy becomes on the order of or greater than the characteristic features of the quantum system. The gross features of the dot conductance are clearly visible out to at least 3 mV as exhibited in device D11 (Fig. 4.29), and the fine structure is clearly visible in D12 to 1 mV as shown in Fig. 4.23. This does not suggest that the confinement energy in the dot is of that order; rather, it suggests that there is a finite drop between the drain and the first barrier, and from the second barrier to the source which is significant with respect to the energy drop across the dot.

5.9 Analysis of Inversion Gate-Depletion Gate Data

Analysis of the data presented previously in this chapter (Sections 5.4, 5.5, and 5.6) suggests that the conductance peak behavior is due to the presence of electronic confinement energy that dominates the Coulomb charging energy. The data presented in this section, in conjunction with the behavior of the energy level

structure found from the electrostatic calculations presented in Section 5.4 form perhaps the strongest argument in support of this assertion.

5.9.1 Experimental Analysis

By re-plotting the peak positions shown in the composite depletion gate-inversion gate bias figures for devices D12 and D11 (Fig. 4.19, 4.21, and 4.28) with the background color map removed and the inversion gate bias scaled using equation 5.6, and plotting the peak positions from the waterfall plot for device D10 (Fig. 4.8), we may investigate how the conductance peak positions evolve for changing depletion gate bias or channel electron energy. Plots of this dependence are shown for device D10 in Fig. 5.28, for device D12 in Fig. 5.29 and Fig. 5.30, and device D11 in Fig. 5.31. What is apparent in these plots are sequences of lines belonging in groups, where each member of each group has a distinguishable slope. Further, to some extent the members of each group appear to have a nearly uniform spacing along the energy axis. In the following, lines within a group whose members have the same slope will be referred to as "branches". In Figs. 5.28-5.31 these branches have been emphasized by drawing a straight line through each. Generally, the quantum dots with square geometry have two groups of branches, while it appears that the dot with asymmetric geometry has three groups of branches. Also, the mean periodicity for the branches in the larger dot geometry are smaller than the mean periodicity of branches in smaller dot geometries. The peak positions which form the members of each branch can be described empirically by the simple expression

$$E_i(n_i) = E_0^i + \alpha_i V_{\text{Depl}} + n_i \hbar\omega_i \tag{5.24}$$

where the sub- or superscript i refers to the groups of branches with nearly same slope and spacing, while the n_i refers to the branch index n in group i. likewise, the slope α_i refers to the slope of branches of members in group i. As described

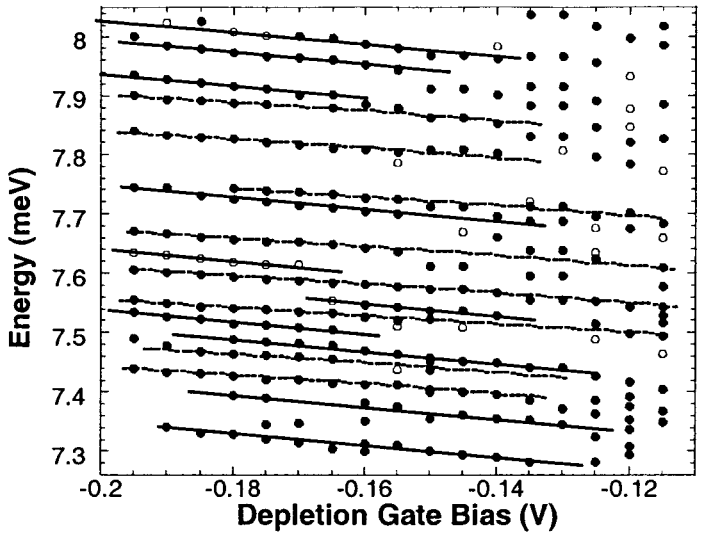

Figure 5.28: A plot of the peak position in the energy-depletion gate bias plane for device D10. Filled and open Circles denote peaks and shoulders, respectively. The straight lines serve as a guide to the eye.

above, for quantum dots with square geometry, $i = 1$ or 2, while for the asymmetric dot geometry $i = 1$, 2 or 3. Aside from these general differences, it is observed that the monatonic evolution of conductance peaks in device D10 becomes more disordered for depletion gate biases greater than about -100 meV. The values for the mean branch spacings $\delta_i = \hbar\omega_i$ and the slopes for branches n_i are summarized in Tables 5.5 and 5.6.

5.9.2 Discussion and Further Analysis

More discussion is needed in order to understand the complicated behavior of the conductance peaks in the inversion gate bias-depletion gate bias plane. One of the features of most interest is the existence of sets of branches with two different slopes for the symmetric quantum dot geometry and three for the asymmetric dot geometry. Three models are presented in the following discussion, with two of the possibilities being eliminated as improbable.

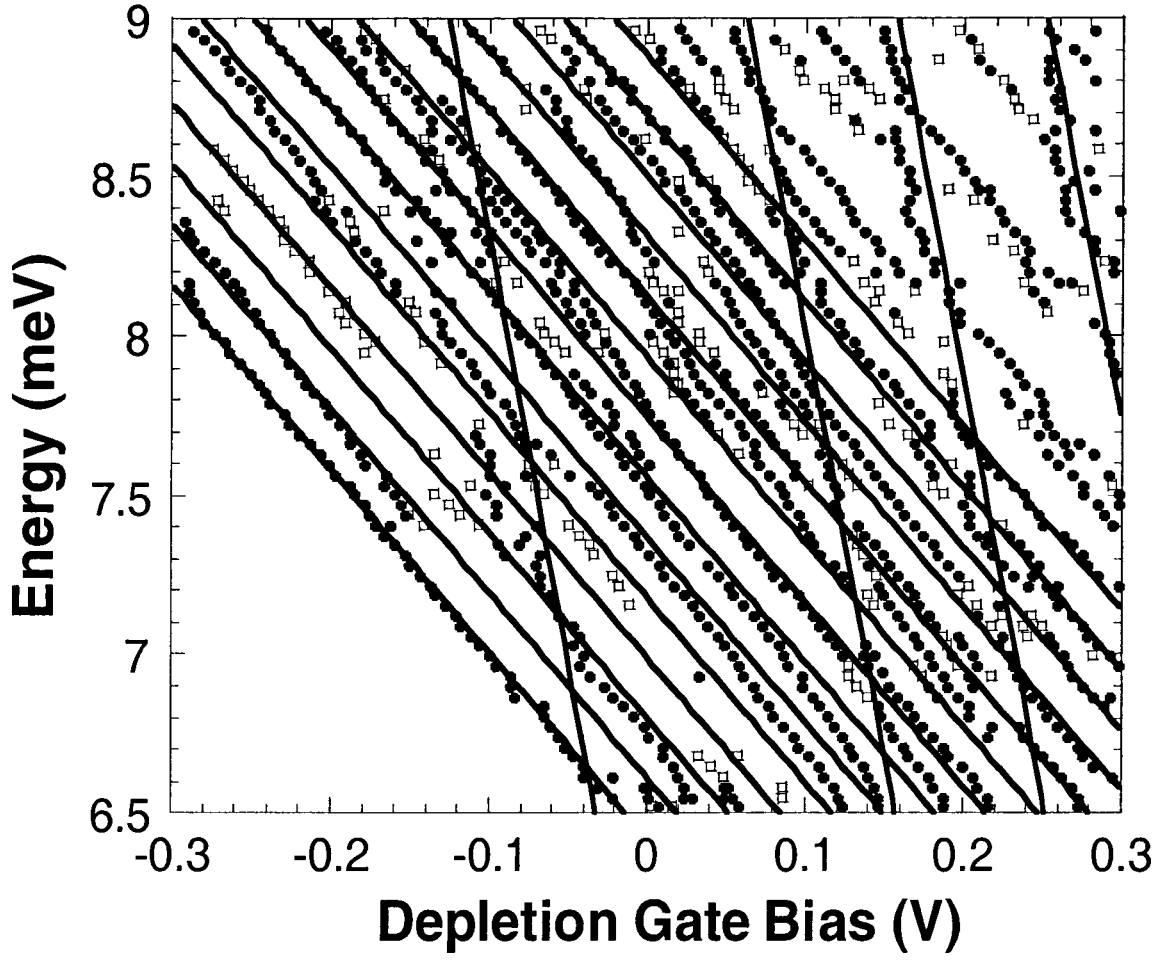

Figure 5.29: Straight lines overlaying the peak positions in Fig. 4.19. The separations along the energy axis are $\hbar\omega_1$ and $\hbar\omega_2$.

Table 5.5: A summary of the mean and maximum level spacings for the devices measured, in units of meV.

device	D9		D10		D11			D12	
branch	δ_1	δ_2	δ_1	δ_2	δ_1	δ_2	δ_3	δ_1	δ_2
mean	0.282	NA	0.0657	0.076	0.106	0.358	0.675	0.141	1.932
max	0.508	NA	0.085	0.189	0.181	0.489	0.742	0.445	1.932

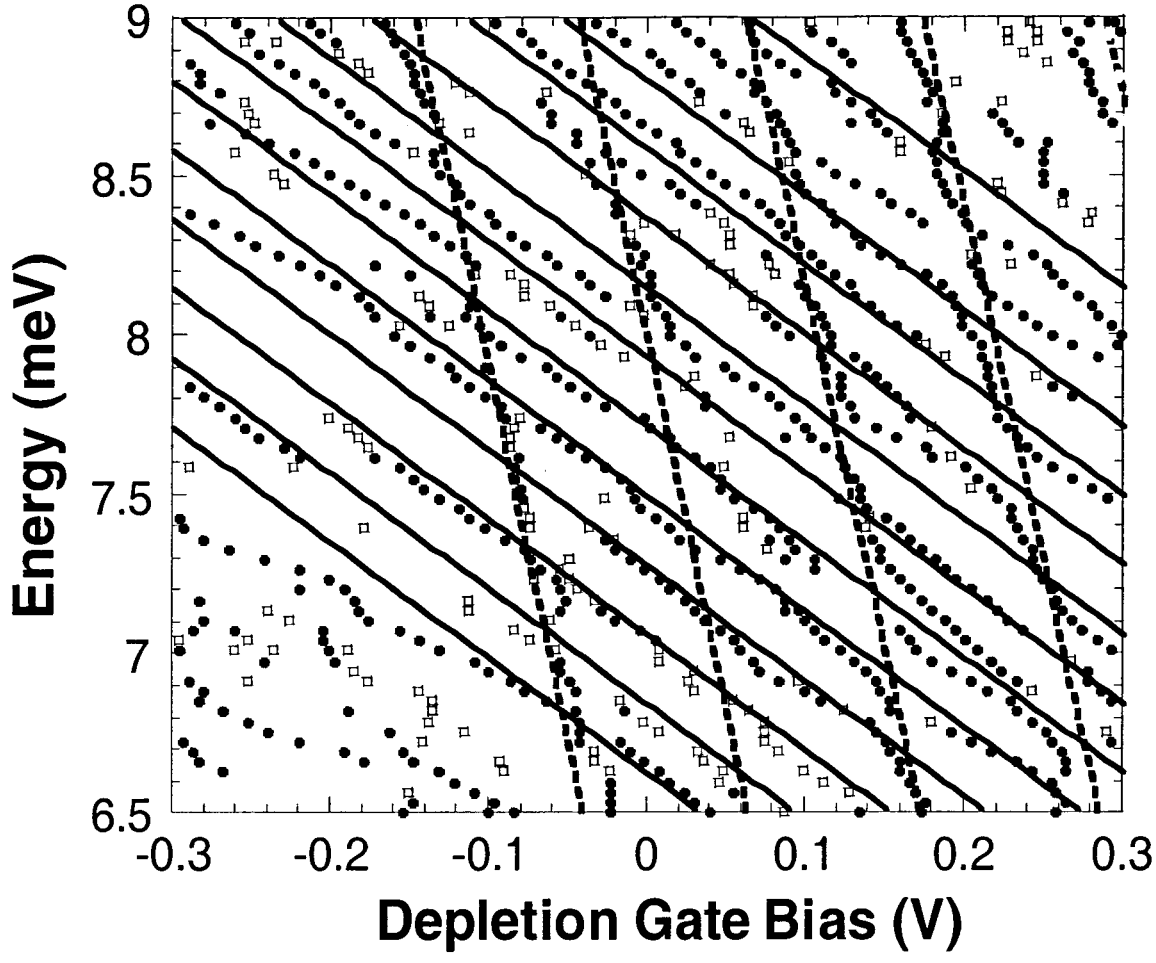

Figure 5.30: Straight lines overlaying the peak positions in Fig. 4.19 for the plunger gate bias sweep of device D12. The separations along the energy axis are $\hbar\omega_1$ and $\hbar\omega_2$. The slope of the second ladder of peaks is not constant with plunger gate bias, indicating that the confinement energy is dependent on the inversion gate bias.

Table 5.6: Summary of energy branch slopes in units of meV/V.

	D10	D11	D12
α_1	-0.64	-2.363	-5.164
α_2	-1.068	-5.152	-25.137
α_3	NA	-9.365	NA

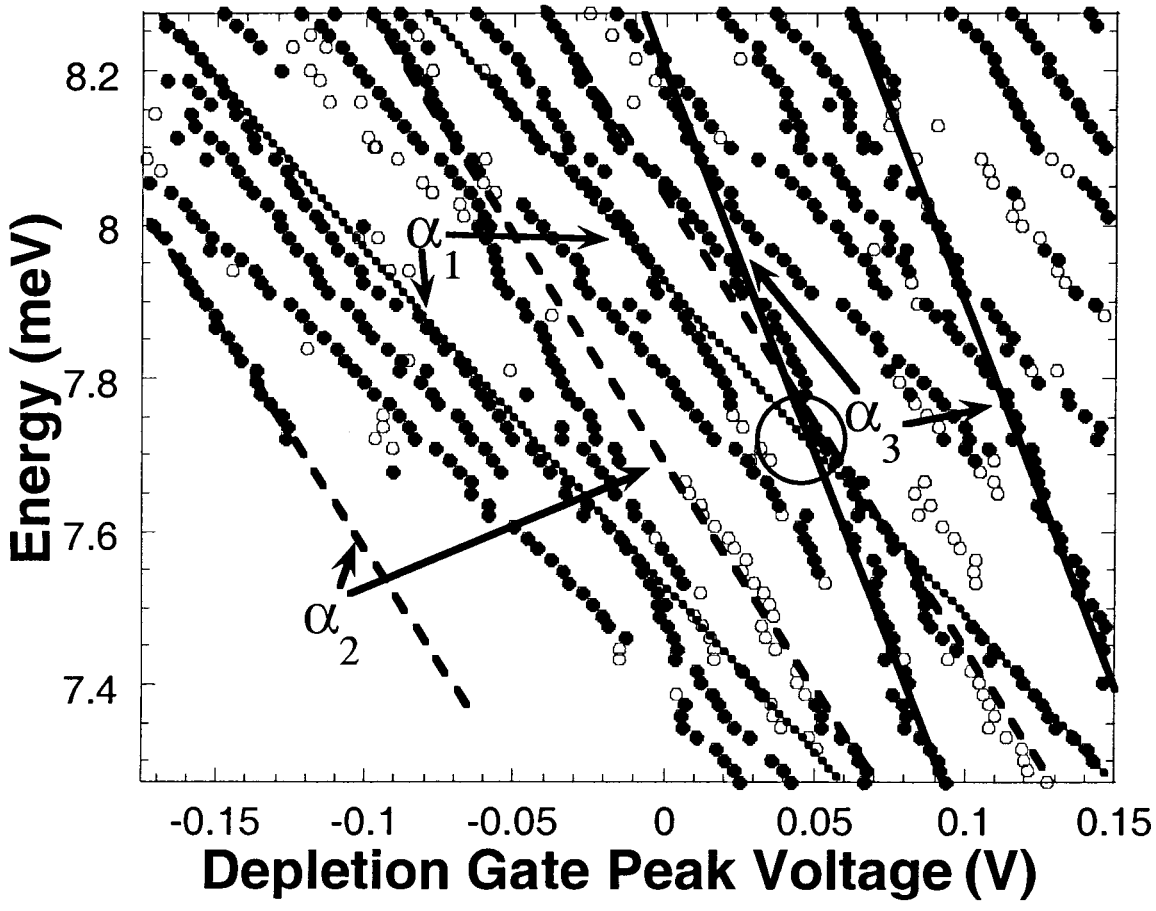

Figure 5.31: Peak and shoulder positions plotted in the depletion gate-channel Fermi energy plane for device D11. Lines with three different slopes have been identified and labeled by drawing lines through several branches, and a triple crossing point is emphasized by drawing a circle around it.

5.9.2.1 Population of Higher 2DEG Subbands

The first proposed model is that Pauli filling of energy eigenvalues forces electrons to reside in the Δ_4 subband of vertical quantum confined states. For electrons residing in the Δ_4 subband, the average distance from the semiconductor-insulator interface would be different from electrons populating the Δ_2 subband. Therefore, electrons would experience a different lateral confinement potential, leading to a different confinement energy, and quite possibly causing the depletion gates to have a different coupling strength to the evolution of the energy eigenvalues and thus conductance peak position. Since the lowest Δ_4 subband lies higher in energy than the lowest Δ_2-subband, it is expected that this occupation occurs at a higher inversion gate bias (and thus higher energy in the channel).

This model is somewhat supported by the fact that it does seem (in the case of D12) that the branches with larger slope (and larger average spacing $\hbar\omega$) begin to appear at higher Fermi energy in the channel. However, two or three stronger pieces of evidence point against this model. First, it is estimated that the average spacing between branches with lower slope is less than 150 μeV, while the lowest Δ_4 subband lies at least 10-15 meV above the lowest Δ_2 subband. With a degeneracy of at least two for the electron spin, this means that at least 133 electrons could occupy the dot before Pauli filling forces the occupation of the lowest Δ_4 subband. In addition, since there is an asymmetry in the effective masses projecting in the lateral directions, it is expected that there should be two different sets of energy eigenvalues resulting from the kinetic term of the Hamiltonian for the lateral portion of the wave functions. SdH data does not show population of the Δ_4 subband in the 2DEG.

Further, it was shown in Section 5.4 that this sort of behavior is not expected. This evidence suggests that Pauli filling forcing population of higher lying subbands is not responsible for the two slopes and spacings.

5.9.2.2 *Asymmetric Coupling of Multiple Axes of Confinement*

Another possible model to explain the existence of multiple slopes is due to asymmetric changes in the device geometry, as discussed in the last section. If the potential profile is perturbed away from square symmetry, such that the background potential becomes more rectangular, two separate axes of quantization will arise with average level spacings $\hbar\omega_1$ and $\hbar\omega_2$. Note that this condition in itself only gives rise to two different spacings. In addition, it is necessary that the coupling be different, $\alpha_{\text{Depl1}} \neq \alpha_{\text{Depl2}}$. Putting this into a parametric equation to describe the position of the energy levels, we have an expression like

$$\begin{aligned} E_{n_x,n_y}(V_{\text{inv}}, V_{\text{depl}}) = E_0 - \alpha_{\text{inv}}V_{\text{inv}} - \alpha_x V_{\text{depl}} + n_x\hbar\omega_x - \alpha_y V_{\text{depl}} + n_y\hbar\omega_y \\ - \{(\alpha_x + \alpha_y)V_{\text{depl}} + n_x\hbar\omega_x + n_y\hbar\omega_y\}, \end{aligned} \tag{5.25}$$

which it can be seen can even give rise to branches with three different slopes when the peak positions are solved for. Note that the labeling of x and y are arbitrary, but encourage one to visualize different ladders (sets) of branches arising from different lateral quantization axes. From Eq. 5.25 that electrons occupying branches with n_y zero, and electrons occupying branches with n_x zero can give two slopes. If an electron occupies both branches simultaneously, there could be three slopes observed (due to the $\alpha_x + \alpha_y$-term). Also included is a term which allows for a coupling to the inversion gate bias, as has been seen in the energy eigenvalue calculations from Section 5.4.

Simulations were performed by allowing for perturbation of both the coupling to the depletion gates (and thus the slopes α_i), and the confinement energy (and thus the branch spacing $\hbar\omega_i$). It can be shown that conductance peaks can arise from the above parameterization of the level spectrum (equation 5.25) which have two or three slopes for a given set of branches. In addition, it can be shown that this perturbation can give rise to a fine structure as well.

First, allow the confinement potential to be asymmetric in the x- and y- directions, deviating from square symmetry in the following way: $\omega_y = \omega_x(1+\delta)$. Applying this perturbation, the peak positions are found by calculating the intersection of the Fermi energy in the channel with the dot energy eigenvalues for a given branch as a function of the inversion gate and depletion gate bias. Plotting the positions of peak conductance found in this way gives the results shown in Fig. 5.32. This perturbation allows for branches with two different slopes for α_x, α_y and two spacings for integer values of δ, but gives rise to splitting appearing to be like the fine structure for $0 \leq \delta \leq 1$, shown in Fig. 5.32 for δ from zero to one.

If the coupling is allowed to be perturbed in the same way, where $\alpha_y = \alpha_x(1+\gamma)$, then we see that for small γ there appear to be only two slopes, whereas for larger γ an additional slope emerges. This is shown in Fig. 5.33 for γ from zero to one. The result of this simulation is that if we allow for a small δ and γ, we can recover something which is similar to the experimental results for D12, while if γ is larger, we get something which agrees with the experimental results for D11.

These simple model limits give an intuitive way to show how a square symmetric potential can be perturbed such that the experimentally observed results are obtained. They allow prediction of both a fine structure and multiple slopes arising from the coupling of the dot energy levels with the depletion and inversion gate biases. What these simulations predict is the qualitative behavior of two slopes. The coupling perturbation model agrees with the general behavior of D11 for large asymmetry ($\gamma = 1.0$) and of D12 for small asymmetry ($\gamma = 0.1$). It does not, however, agree with the slopes and branch spacings seen experimentally.

While it is important to note that these simulations give results which are not predicted in the calculated energy eigenvalues, it is equally important to understand that imperfections in the lithography and/or the presence of charged impurities in or near the vacinity of the quantum dot could easily cause perturbations with results

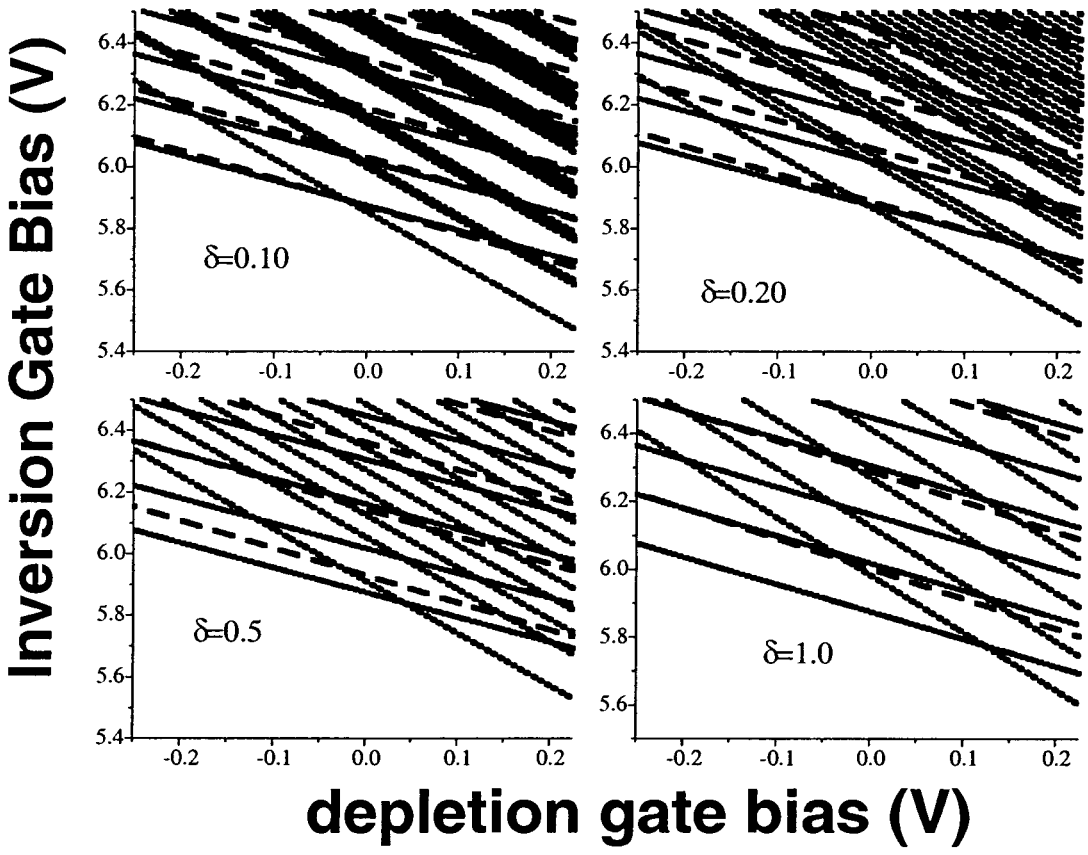

Figure 5.32: A plot of the model peak conductance positions where the perturbation of the confining potential is varied as a parameter from δ=0.1 to 1, for γ=0.1.

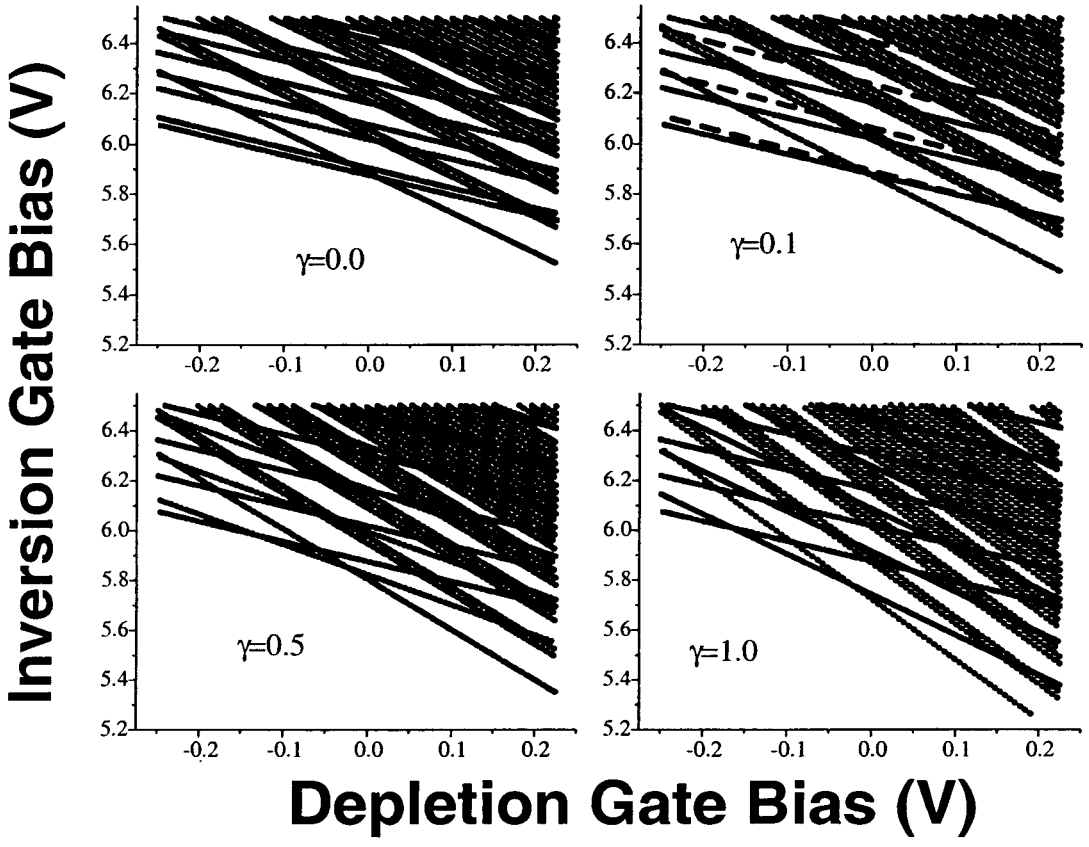

Figure 5.33: A plot of the model conductance peak positions where the perturbation of the coupling to the confining potential is varied from γ=0 to 1, with δ=0.1.

similar to the model here. In fact, preliminary studies at ASU including random impurity fluctuations indicate more complicated structure in the theoretical model. To date, the model presented here related to asymmetric deformation of the dot geometry seems most plausible model accounting for the experimental results.

5.9.3 Simulations Using Theoretical Calculations

It is a relatively straightforward procedure to invert the calculated energy level structure presented in Section 5.4 to obtain the predicted conductance peak positions in the inversion gate-depletion gate bias plane. Assuming that a conductance peak results from the chemical potential in the channel aligning with an allowed energy level within the quantum dot, one must simply solve for the crossings of the energy eigenvalues in the dot with the chemical potential in the channel for each inversion gate bias, using Eq. 5.6 to scale the bias with Fermi energy of the 2DEG. In practice, the solution for these values is obtained by linearly interpolating for the energy levels for successive depletion gate bias values, which was performed by S. Milicic at ASU. Figure 5.34 shows such a plot of the predicted conductance peak spectrum obtained from an inversion of the data in Fig. 5.3 for the symmetric biasing of a 200 nm dot. It appears that the degeneracy structure that was evident in the energy level spectrum is shifted so that the number of nearly degenerate members of each group changes, and that the splitting is nearly regular as the inversion gate bias is changed. In addition, some of the levels are predicted to be degenerate in the peak conductance spectrum, but split as the depletion gate bias is made increasingly positive. What one would expect to observe then at finite temperature is a dominant, broadened peak in which these degenerate or very nearly degenerate levels are concurrent, due to overlap of the tails of these peaks in the peak spectrum, and are thus indistinguishable from one another.

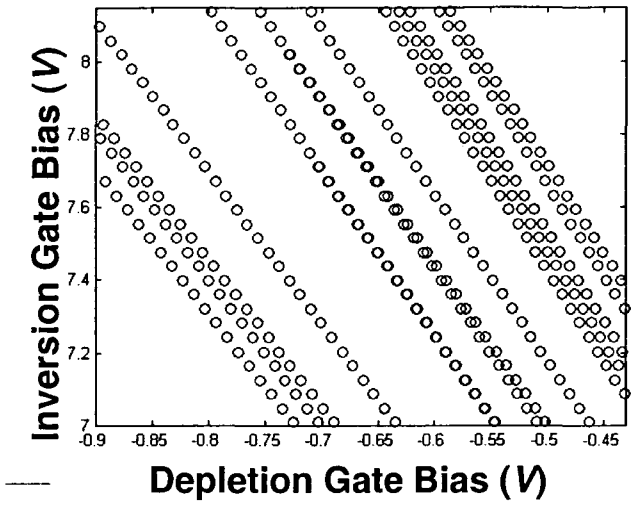

Figure 5.34: Simulated peak conductance positions for a symmetric 200 nm quantum dot with symmetric biasing conditions.

Figure 5.35 shows a plot of the predicted peak conductance positions based on the data plotted in Fig. 5.7 for the asymmetric (plunger) bias case. As expected, the asymmetry in the biasing condition leads to a crossing in the peak branch spectrum observed in Fig. 5.35 due to the breaking of the symmetry. It appears that two groups of branches are observed, with each group having nearly equal slopes and nearly equal branch to branch spacing. These two groups may be identified as states arising from quantization either parallel or perpendicular to the plunger direction. Since the plunger effectively changes the dimension of the dot in one direction primarily; states associated with this direction of confinement are more sensitive to plunger bias, leading to two sets of slopes.

Figure 5.36 shows a plot of the conductance peak positions for a symmetrically biased quantum dot with an asymmetric 100 $\times$ 200 nm geometry in the inversion gate bias-depletion gate bias plane.

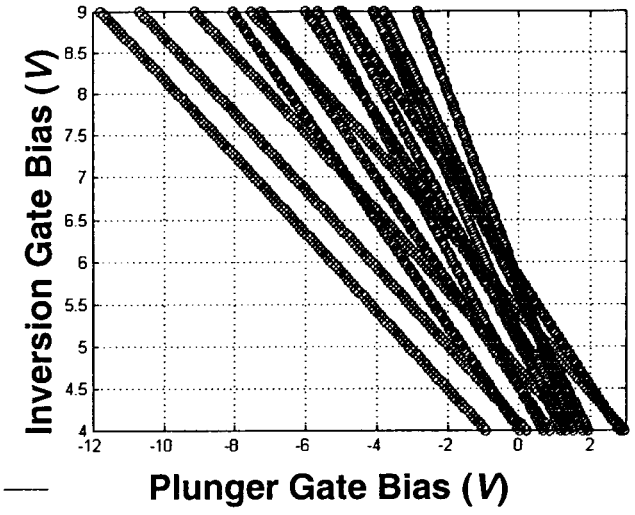

Figure 5.35: A plot of the predicted peak conductance positions based on the energy level calculations for an asymmetrically biased symmetric 200 nm quantum dot.

Unexpectedly, there is no real evidence of a crossing behavior in the theoretical curves, even though one would expect such behavior based on the arguments of the preceeding paragraph.

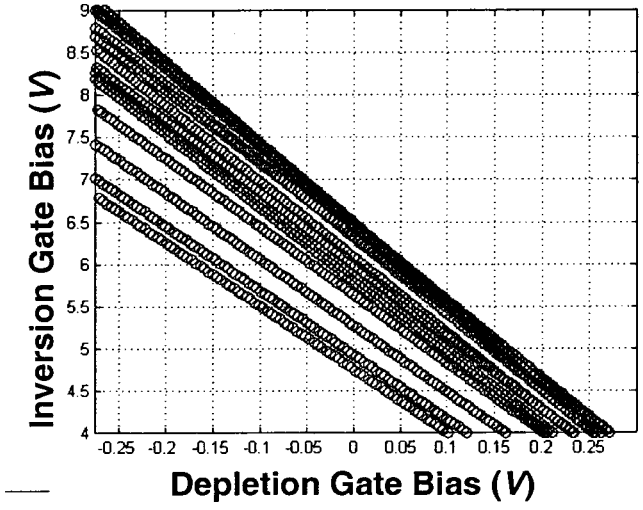

Figure 5.36: A plot of the predicted conductance peak positions for an asymmetric dot geometry biased symmetrically.

6. CONCLUSION AND FUTURE WORK

6.1 Conclusions

The simplest conclusion that can be drawn from this work is that it is possible to fabricate quantum dots using modern CMOS technology, and, with the exception of the minimum feature sizes required to define the depletion gate structures, all dimensions are achievable in present day CMOS processes. These devices are functional without using more expensive SOI substrates as have been used in other similar devices. However, it is equally valid to conclude that before more depth in understanding of the quantum effects observed in this project can be achieved, a more robust architecture must be implemented. Discussion of this topic will be deferred until the section on future work, Section 6.2.

The transport measurements were taken with biasing conditions such that the channel was biased in the regime from the onset of weak inversion to just around the onset of strong inversion, and the depletion gates and drain are biased so that the dot transport is in the tunneling regime. One of the main findings of this work is that the transport features are more consistent with the discrete confinement energies rather than Coulomb blockade effects due to single electron charging. The values for dot radii given by assuming that the peak splitting is due to electronic confinement are more consistent with the lithographic dimensions than the radii found assuming pure Coulomb blockade, although it is clear that a better model must be found for the dot self-capacitance. The large variations in the equivalent capacitance, C_Σ, the large variation in the peak amplitudes (in consectutive peaks), and the lack of a Coulomb gap in the current around zero source-drain bias also support this assertion. It is distinctly possible that the transport could be some combination of Coulomb blockade and discrete confinement energy-controlled tunneling. By sweeping the

inversion gates and depletion gates independently, what appears to be an energy level spectrum is mapped out, showing the effects of decreasing the bottom of the ladder of energy levels with the inversion gate, while increasing the confinement energy with the depletion gates.

Calculations of the quantum dot energy level using a self-consistent Schrödinger-Poisson solver confirmed that electrostatic confinement should be appreciable in dots of this dimension and shape, although of somewhat greater value than observed in the experimental data. These calculated energy levels behaved in a manner consistent with the observed experimental evolution of peaks in the $V_{\mathrm{Inv}} - V_{\mathrm{Depl}}$ plane. These calculations also showed that some lifting of degeneracy (where the degeneracy is the same as that for energy levels for a harmonic potential) in the dot are expected; this degeneracy breaking could explain the fine structure observed in the inversion gate sweeps, as well as multiple slopes in the data, although the experimentally observed fine structure is not as regular as the calculations predict. Still more recent calculations have shown that the presence of discrete dopant impurities can in fact modify the calculated degeneracy splitting to show results more consistent with the observed fine structure. These calculations do predict multiple slopes in the asymmetric bias case. They do not predict the multiple slopes of groups of branches observed in the actual physically drawn asymmetric device.

A simple model was presented, however, which showed that by allowing the spherically symmetric confinement to be split into two slightly different, independent confinement strengths oriented in two spatial directions allowed for qualitative reproduction of both the fine structure and multiple slopes observed in the experimental data. The model allows for asymmetry in both the confinement strength and the coupling of the bottom of the energy eigenvalue ladders, while independence between the two allows for x-like, y-like, and $x+y$-like ladders for three independent ladders of levels. This model is supported by the observation that apparent groups

of branches with three slopes occur for the asymmetric device, while there are only two apparent groups of branches for either of the two size symmetrical dots. This asymmetry is simulated in the numerical asymmetrical bias (plunger) simulation.

This good qualitative agreement between the presented empirical model, the simulations based on theoretical calculations, and the experimental data in the $V_{\mathrm{Inv}} - V_{\mathrm{Depl}}$ make the best evidence that the discrete energy level structure in the dot dominates Coulomb blockade in transport properties.

6.2 Future Work

Although a number of experiments have recently been reported on Si quantum dots, there remain large gaps in the knowledge of Si quantum dot systems. Particularly lacking are results specific to the multiple valley nature of the Si band structure. In the following, I propose future work, which should be performed in order to narrow these gaps, both experimental and theoretical.

6.2.1 Experimental

Before specific experiments are performed, the device structure which was used during this work must be modified to provide a much longer lifetime. In all devices measured, the devices failed far before a complete set of experiments could be performed on them. One of the weakest points in the device structure is the deposited gate oxide between the depletion gates and the inversion gate, which is prone to dielectric breakdown. A proposed improvement to this weak link is to replace the RPECVD process with a low temperature LPCVD oxide using TEOS, which results in a much higher density and thus more robust oxide quality. Since this is a rather slow process in terms of deposition rate, the thermal load could be reduced if this were a composite oxide, where the first part was LPCVD and the second part were RPECVD. In addition, there is a problem with charge in this

oxide layer. Preconditioning the surface prior to deposition with an oxide plasma treatment has been used to reduce this problem. The metal in the depletion gates could be protected by depositing a thin silicon layer by electron beam evaporation after the chrome is deposited. Additional improvement to the oxides can be obtained by using a local oxidation process to make the field oxide, or LOCOS process. This process allows for a more gradual profile for the metal layers to cover.

An additional limitation is the low overall channel conductance, which can be alleviated by creating a much more favorable channel geometry. By widening the channel, the overall conductance should be enhanced. Shortening the channel is more challenging, as it would require a careful study of MOS properties to determine the shortest length that could be used while still avoiding effects from the drain depletion region, high fields across the device, and allows de-phasing of injected carriers. In addition to this, better lithography tools would be required to allow for a large enough number of electrical contacts to connect to the active region.

Assuming these modifications lead to better performing and more robust devices, several experiments are very important in understanding the behavior of the devices:

- In order to understand the spin states of a few electron quantum dot, magnetic field studies of the conductance peaks must be obtained; specifically sweeps in both the inversion gate-magnetic field plane and the depletion gate-magnetic field plane.

- Plunger gate sweeps and sweeps of opposing depletion gate pairs for all of the dot geometries will allow a better understanding of the effect of the asymmetry in the dots.

- Large ranges of bias voltages must be taken to see where patterns break down in the evolution of the level spectroscopy.

- More complete sweeps of the differential conductance in the drain bias-plunger gate bias will yield more information about the evolution of levels in the dot. Larger drain bias and larger plunger gate biases are needed.

- The temperature dependence should be probed. Particularly, going to a lower temperature should allow for a much better resolution of the shapes of the minor peaks. The temperature dependence of the width and amplitudes of conductance peaks would be useful.

- Several additional sizes of dots should be fabricated to investigate the transition from a two dimensional continuum of states to fully discrete states within the dot, and also the onset of Coulomb blockade. Ideally this would be a continuous range of dots from 200 nm to the limit where conductance peaks are no longer observable, but practically this would include incremental increase to the regimes of two dimensional-one dimensional transition and to the limit where conductance peaks are no longer observable.

Beyond these experiments, optical experiments have not been performed, but in principle should be possible with a modification to the device structure (albeit a major one). By replacing the Al in the inversion gate with a conducting transparent oxide, such as ATO or ITO, and fabricating an ensemble of dots instead of a single dot, it should be possible to study at radiation absorption in these dots. In addition, other experiments could be performed to investigate whether selection rules were broken allowing optical transitions forbidden in bulk but allowed in the quantized systems.

By fabricating ensembles of dots, it should also be possible to measure the magnetization and the heat capacity, getting at the level structure, as proposed by Maksym and Chakraborty[66, 95]

6.2.2 Theoretical

Theoretical progress is being made on understanding the electronic structure and transport in Si quantum dots. However, more thorough transport calculations must be developed to support the findings of the research in this work.

Some calculations of the energy level structure for these quantum dots have been completed. However, the spin orbit interaction has not been taken into account, and some attempt should be made to calculate the splittings coming from the valley-orbit interaction, extending the theory for inversion layers and the bulk.

The second type of calculation should integrate the results of the first (the electronic structure) into the transport properties of the dot. These two calculations should support the findings of this work, and point out where future work is needed.

BIBLIOGRAPHY

1. See, for example N. W. Ashcroft, N. D. Mermin *Solid State Physics,* (Saunders College Publishers, New York, NY. 1976).

2. C. M. Wolfe, N. Holonyak, Jr., and G. E. Stillman, *Physical Properties of Semiconductors,* (Prentice-Hall, Englewood Cliffs, NJ USA, 1989).

3. D. K. Ferry and S. M. Goodnick, *Transport in Nanostructures,* (Cambridge University Press, New York, USA 1997).

4. E. M. Conwell, "High Field Transport in Semiconductors, in *Solid State Physics, supplement 9* (Academic Press, New York, USA 1967).

5. J. M. Ziman, *Electrons and Phonons,* (Oxford, 1960).

6. C. Weisbuch,and B. Vinter, *Quantum Semiconductor Structures: Fundamentals and Applications,* (Academic Press, San Diego, CA 1991).

7. T.F. Gaylord and K. F. Brennan, J. Appl. Phys. **65**, 814 (1989).

8. L. P. Kouwenhoven, N. C. van der Vaart, A. T. Johnson, W. Kool, C. J. P. M. Harmans, J. G. Williamson, A. A. M. Staring, and C. T. Foxon, Z. Phys. B. **85**, 367 (1991).

9. K. Yano, T. Ishi, T. Hashimoto t. Kobayashi, F. Murai, and K. Seki, IEEE Trans. on Elec. Dev. **41**, 1628 (1994).

10. S. Tarucha, D. G. Austing, T. Honda, R. J. van der Hage, and L. P. Kouwenhoven, Phys. Rev. Lett. **77**, 3613 (1996).

11. F. Stern and W. E. Howard, Phys. Rev. **163**, 816 (1967).

12. F. Stern, Phys. Rev. **B**5, 4891 (1972).

13. T. Ando, A. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982).

14. K. von Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).

15. see, for example D. C. Cook, *Electrical Characterization of GaAs Materials and Devices,* (J. Wiley & Sons, New York 1989).

16. A. B. Fowler, A. Hartstein, and R. A. Webb, Phys. Rev. Lett.**48**, 196 (1982).

17. F. Scott-Thomas, S. B. Field, M. A. Kastner, H. I. Smith, and D. A. Antoniadis, Phys. Rev. Lett. **62**, 583 (1989).

18. H. Ishikuro, T. Fujii, T. Saraya, G. Hashiguchi, T. Hiramoto, and T. Ikoma, Appl. Phys. Lett. **68**, 3585 (1996).

19. M. Je, S. Han, I. Kim, H. Kwon, H. Shin, and K. Lee, in press (NPMS'98)

20. I. Ohdomari and T. Shinada, in press (NPMS'98).

21. S.E. Laux and F. Stern, Appl. Phys. Lett. **49**, 91 (1986).

22. S.E. Laux, D. J. Frank, and F. Stern, Surf. Sci. **196**, 101 (1988).

23. K. Kojima, K. Mitsunaga, and K. Kyuma, Appl. Phys. Lett. **55**, 862 (1989).

24. G. S. Solomon, J. A. Trezza, and J. S. Harris, Jr., Appl. Phys. Lett. **66**, 991 (1995).

25. D. Leonard, K. Pond, and P. M. Petroff, Phys. Rev. B, **50**, 11687 (1994).

26. V. Bressler-Hill, A. Lorke, S. Varma, P. M. Petroff, K. Pond, and W. H. Weinberg, Phys. Rev. B **50**, 8479 (1994).

27. G. E. Cirlin, G. M. Guryanov,A. O. Golubok, S. Y. Tipissev, N. N. Ledentsov, P. S. Kop'ev, M. Grundmann, and D. Bimberg, Appl. Phys. Lett. **67**, 97 (1995).

28. B. R. Bennett, R. Magno, and B. V Shananbrook, Appl. Phys. Lett. **68**, 505 (1995).

29. H. Lipsanen, M. Sopanen, and J. Ahopelto, Phys. Rev. B **51**, 13868 (1995).

30. J. Tulkki and A. Heinänäski, Phys. Rev. B **52**, 8239 (1995).

31. C. L. Chang, K. Mahalingam, N. Otsuka, M. R. Melloch, and J. M. Woodall, J. of Electronic Materials **22**, 1413 (1993).

32. A. C. Warren, J. M. Woodall, J. L. Freeouf, D. Grischkowsky, D. T. McInturff, M. R. Melloch, and N. Otsuka, Appl. Phys. Lett **57**, 1331 (1990)

33. A. Claverie, F. Namavar, and Z. Liliental-Webber, Appl. Phys. Lett. **62**, 1271 (1993).

34. S. Jeppeson, M. S. Miller, D. Hessman, B. Kowalski, I. Maximov, and L. Samuelson, submitted to Appl. Phys. Lett.

35. D. S. L. Mui, D. Leonard, L. A. Coldren, and P. M. Petroff, Appl. Phys. Lett. **66**, 1620 (1995).

36. W. Seifert, N. Carlsson, L. -E. Wernersson, and L. Samuelson, ibid. **68**, 1684 (1996).

37. R. A. Keihl, M. Yamaguchi, O. Ueda, N. Horiguchi, and N. Yokoyama, Appl. Phys. Lett. **68**, 478 (1996).

38. M. Miller, S. Jeppesen, B. Kowalsky, M. Pistol, K. Georgsson, and L. Samuelson, submitted to Science.

39. D. Leonard, M . Krishnamurthy, C. M. Reaves, S. P. Denbaars, and P. M. Petroff, Appl. Phys. Lett. **63**, 3203 (1993).

40. J.-Y. Marzin, J.-M. Gérard, A. Izraël, D. A. Barrier, G. Bastard, Phys. Rev. Lett. **73**, 716 (1994).

41. P. Castrillo, D. Hessman, M. -E. Pistol, S. Anand, N. Carlsson, W. Seifert, and L. Samuelson, Appl. Phys. Lett. **67**, 1905 (1995).

42. M. Grundmann, J. Christen, N. N. Ledentsov, J. Böhrer, D. Bimberg, S. S. Ruvimov, P. Werner, U. Richter, U. Gösele, J. Heydenreich, V. M. Ustinov, A. Yu. Egorov, P. S. Kop'ev, and Zh. I. Alferov, ibid. **74**, 4043 (1995).

43. J. Alsmeier, E. Batke, and J. P. Kotthous, Surface Sci. **229**, 287 (1990).

44. H. Matsuoka, IEEE Electron Device Lett. **13**, 20 (1992).

45. H. Matsuoka, Appl. Phys. Lett. **64** 586 (1994).

46. H. Matsuoka, T. Ichiguchi, T. Yoshimura, and E. Takeda, J. Appl. Phys. **76**, 5561 (1994).

47. H. Matsuoka and S. Kimura, Jpn. J. Appl. Phys. **34**, 1326 (1995).

48. H. Matsuoka, Appl. Phys. Lett. **66**, 613 (1995).

49. E. Leobandung, L. Guo, U. Wand, and S. Chou, Appl. Phys. Lett. **67**, 938 (1995).

50. E. Leobandung, L. Guo, and S. Chou, ibid. **67**, 2338 (1995).

51. L. Zhuang, L. Guo, Y. Wang, and S. Chou, ibid. **72**, 1205 (1998).

52. K. Uchida, J. Koga, A. Ohata, and A. Toriumi, Proceedings of Silicon Nanoelectroncs Workshop '98, 59 (1998).

53. M. G. Peters, S. G. den Hartog, J. I. Dijkhuis, O. J. A. Buyk, and L. W. Molenkamp, J. Appl. Phys. **84**, 5052 (1998).

54. A. Fujiwara, Y. Takahashi, and K. Murase, Proceedings of the Fouth International Symposium on New Phenomena in Mesoscopic Structures, Kauai, HI USA, 230 (1998).

55. Y. Kanemitsu, T Ogawa, K. Shiraishi, and K. Takeda, Phys. Rev. B **48** , 4883 (1993).

56. Y. Yamada, T. Orh, I. Umezu, S. Takeyama and T. Yoshida, Jpn. J. Appl. Phys. Pt. 1 **35**, 1361 (1996).

57. M. Fukuda, N. Nakegawa, S. Miyazaki, and M. Hirose, Appl. Phys. Lett. **70**, 2291 (1997).

58. F. Geerinckx, F. M. Peeters, and J. T. Devreese, J. Appl. Phys **68**, 3435 (1990).

59. P. M. Morse and H. Feshbach, *Methods of Theoretical Physics,* (McGraw Hill, New York, 1953).

60. C. G. Darwin, Proc. Cambridge Philos. Soc. **27**, 86 (1930).

61. V. Fock, Z. Phys. **47**, 446 (1928).

62. R. B. Dingle, Proc. Roy. Soc. London **A211**, 500 (1952).

63. U. Bockelmann, Phys. Rev. B **50**, 17271 (1994).

64. J. S. Gradshteyn and J. M. Ryzhik, *Tables of Integrals, Series, and Products,* (Academic, New York, 1965).

65. G. W. Bryant, Phys. Rev. Lett. **59**, 1140 (1987).

66. P. A. Maksym and T. Chakraborty, Phys. Rev. Lett. **65**, 108 (1990).

67. Q. P. Li, K. Karraï, S. K. Yip, S. Das Sarma, and H. D. Drew, Phys. Rev. B **43**, 5151 (1991).

68. U. Merkt, J. Huser, and M. Wagner, ibid. **43**, 7320 (1991).

69. T. Chakraborty, V. Halonen, and P. Pietiläinen, ibid. **43**, 14289 (1991).

70. D. Pfannkuche,and R. R. Gerhardts, ibid. **44**, 13132 (1991).

71. J. M. Kinaret, Y. Meir, N. S. Wingreen, P. Lee, and X.-G. Wen, Phys. Rev. B **45**, 9489 (1992).

72. J. M. Kinaret, Y. Meir, N. S. Wingreen, P. Lee, and X.-G. Wen, ibid. **46**, 4681 (1992).

73. M. Wagner, U. Merkt, and A. V. Chaplik, Phys. Rev. B **45**, 1951 (1992).

74. P. A. Maksym and T. Chakraborty, ibid. **45**, 1947 (1992).

75. D. Pfannkuche, V. Gudmundsson, and P. A. Maksym, ibid. **47**, 2244 (1993).

76. T. Ezaki, N. Mori, and C. Hamaguchi, Phys. Rev. B **56**, 6428 (1997).

77. A. Kumar, S. E. Laux, and F. Stern, Phys. Rev. B **42**, 5166 (1990).

78. M. Stopa, Y. Aoyagi, and T. Sugano, Surf. Sci. **305**, 571 (1994).

79. V. Gudmundsson and R. R. Gerhardts, Phys. Rev. B **43**, 12098 (1991).

80. J. J. Palacios, L. M. Moreno, G. Chiappe, E. Louis, and C. Tejador, Phys. Rev. B **50**, 5760 (1994).

81. M. Fujito, A. Natori, and H. Yasunaga, Phys. Rev. B **53**, 9952 (1996).

82. T. Ezaki, "Study of Electronic States and Electrical Conduction of Semiconductor Quantum Devices," Ph. D. Thesis, Osaka University, (1997).

83. M. Stopa, Phys. Rev. B **54**, 13767 (1996).

84. M. Macucci, K. Hess, and G. J. Iafrate, Phys. Rev. B **48**, 17354 (1993)

85. M. Macucci, K. Hess, and G. J. Iafrate, J. Appl. Phys. **77**, 3367 (1995).

86. M. Macucci, K. Hess, and G. J. Iafrate, Phys. Rev. B **55**, R4869 (1997).

87. C. Sikorski and U. Merkt, Phys. Rev. Lett. **62**, 2164 (1989)

88. T. Demel, D. Heitmann, P. Grambow, and K. Ploog, Phys. Rev. Lett. **64**, 788 (1990)

89. C. Sikorski and U. Merkt, Surf. Sci. **229**, 282 (1990).

90. A. Lorke, J. P. Kotthous, and K. Ploog, Phys. Rev. Lett. **64**, 2559 (1990).

91. B. Meurer, D. Heitmann, and K. Ploog, Phys. Rev. Lett. **68**, 1371 (1992).

92. F. M. Peeters, Phys. Rev. B. **42**, 1486 (1990).

93. L. Brey, N. F. Johnson, and B. I. Halperin, Phys. Rev. B **40**, 10647 (1989).

94. E. Gornik, R. Lassnig, G. Strasser, H. L. Störmer, A. C. Gossard, and W. Wiegmann, Phys. Rev. Lett. **54**, 1820 (1985).

95. P. A. Maksym and T. Chakraborty, Phys. Rev. B **45**, 1947 (1992).

96. T. Haavasoja, H. L. Störmer, D. J. Bishop, V Narayanamurti,A. C. Gossard, and W. Wiegmann, Surf. Sci. **142**, 294 (1984).

97. J. P. Eisenstein, H. L. Störmer, V Narayanamurti, A. Y. Cho, A. C. Gossard, and C. W. Tu, Phys. Rev. Lett. **55**, 875 (1985).

98. R. Landauer, IBM J. Res. Dev. **1**, 223 (1957).

99. R. Landauer, Phil. Mag. **21**, 863 (1970).

100. M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Phys. Rev. B **31**, 6207 (1985).

101. B. J. van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marel, and C. T. Foxon, Phys. Rev. Lett. **60**, 848 (1988).

102. C. W. J. Beenakker and H. van Houten, "Quantum Transport in Semiconductor Nanostructures," in Solid State Physics **44**, ed. H. Ehrenreich and D. Turnbull (Academic Press, Boston, 1991) pp. 1-228.

103. H. van Houten, C. W. J. Beenakker and B. J. van Wees, "Quantum Point Contacts" in *Nanostructure Systems*, ed. M. Reed, Semiconductors and Semimetals **35** (Academic Press, Boston 1992), pp. 9-112.

104. D. A. Wharam, T. J. Thornton, R. Newbury, M. Pepper, H. Ahmed, J. E. F. Frost, D. C. Hasko, D. C. Peakock, D. A. Ritchie, and G. A. C. Jones, J. Phys. C **21**, L209 (1988).

105. D. A. Wharam, M. Pepper, R. Newbury, D. G. Hasko, H. Ahmed, J. E. F. Frost, D. A. Ritchie, D. C. Peacock, and G. A. C. Jones, "Ballistic Transport in Quasi-One-Dimensional Structures," in *Spectroscopy of Semiconductor Microstructures*, ed. G. Fasol, A. Fasolino, and P. Lugli, NATO ASI Series B: **206** (Plenum Press, New York, 1989) pp. 115-141.

106. E. C. Kemble, Phys. Rev. **48**, 549 (1935).

107. J. N. L. Connor, Mol. Phys. **15**, 37 (1968).

108. W. H. Miller, J. Chem. Phys. **48**, 1651 (1968).

109. H. A. Fertig and B. I. Halperin, Phys. Rev. Lett. **36**, 7969 (1987).

110. M. Büttiker, Phys. Rev. B **41**, 7906 (1990).

111. L. I. Glazman, G. B. Lesovik, D. E. Khemel'nitskii, and R. I. Shekter, JETP Lett. **48**, 239 (1988).

112. K.-F. Berggren, T. H. Thornton, D. J. Newson, and M. Pepper, Phys. Rev. Lett. **57**, 1769 (1986).

113. B. J. van Wees, L. P. Kouwenhoven, E. M. M. Willems, C. J. P. M. Harmans, J. E. Mooij, H. van Houten, C. W. J. Beenakker, J. G. Williamson, andC. T. Foxon, Phys. Rev. B **43**, 12431 (1991).

114. C. W. J. Beenakker, Phys. Rev. B **44**, 1646 (1991).

115. D. V. Averin, A. N. Korotkov, and K. K. Likharev, Phys. Rev. B **44**, 6199 (1991).

116. I. O. Kulik and R. I. Shekter, Sov. Phys. JETP **41**, 308 (1975).

117. U. Meirav, M. A. Kastner, and S. J. Wind, Phys. Rev. Lett. **65**, 771 (1990).

118. M. Khoury, A. Gunther, D.P. Pivin Jr., M.J. Rack, and D.K. Ferry, Nanotechnology, in press.

119. M. Khoury, A. Gunther, D.P. Pivin Jr., M.J. Rack, and D.K. Ferry, Japn. J. Appl. Phys. **38**, 469 (1999).

120. M. Khoury, A. Gunther, M.J. Rack, D.P. Pivin, Jr., and D.K. Ferry, Microelect. Eng. **47**, 189 (1999).

121. A. D. Gunther, S. M. Goodnick, M. Khoury, A. E. Krishnaswamy, M. J. Rack, T. J. Thornton, accepted for publication in Microelect. Eng.(2000).

122. M. A. Reed, J. N. Randall, R. J. Aggarwal, R. J. Matyi, T. M. Moore, and A. E. Wetsel, Phys. Rev. Lett. **60**, 535 (1988).

123. A. T. Johnson, L. P. Kouwenhoven, W. de Jong, N. C. van der Vaart, C. J. P. M. Harmans, and C. T. Foxon, Phys. Rev. Lett. **69**, 1592 (1992).

124. E. B. Foxman, P. L. McEuen, U. Meirav, N. S. Wingreen, Y. Meir, P. A. Belk, N. R. Belk, M. A. Kastner, and S. J. Wind, Phys. Rev. B **47**, 10020 (1993).

125. J. Weis, R. J. Haug, K. v. Klitzing, and K. Ploog, Phys. Rev. Lett. **71**, 4019 (1993).

126. D.C. Ralph, C. T. Black, and M. Tinkham, Phys. Rev. Lett. **78**, 4087 (1997).

127. D. R. Stewart, D. Sprinzak, C. M. Marcus, C. I. Duruöz, and J. S. Harris Jr., Science **278**, 1784 (1997).

128. D. V. Averin and A. N. Korotkov, J. Low Temp. Phys. **80**, 173 (1990).

129. R. C. Ashoori, H. L. Stormer, H. S. Weiner, L. N. Pfeiffer, K. W. Baldwin, and K. W. West, Phys. Rev. Lett. **68**, 3088 (1992).

130. R. C. Ashoori, H. L. Stormer, H. S. Weiner, L. N. Pfeiffer, K. W. Baldwin, and K. W. West, Phys. Rev. Lett. **71**, 613 (1993).

131. R. C. Ashoori, H. L. Stormer, H. S. Weiner, L. N. Pfeiffer, K. W. Baldwin, and K. W. West, Surf. Sci. **305**, 558 (1994).

132. M. A. Kastner, Physics Today **46**, 24 (1993).

133. R. C. Ashoori, Nature **379**, 413 (1996).

134. S. Tarucha, D. G. Austing, T. Honda, R. van der Hage, and L. P. Kouwenhoven, Japn. J. Appl. Phys. **36** pt. 1, 3917 (1997).

135. L. P. Kouwenhoven, T. H. Oosterkamp, M. W. S. Danoesastro, M. Eto, D. G. , T. Honda, and S. Tarucha, Science **278**, 1788 (1997).

136. B. Su, V. J. Goldman, J. E. Cunningham, Phys. Rev. B **46**, 7644 (1992).

137. T. Schmidt, Phys. Rev. B **51**, 5570 (1995).

138. N. C. van der Vaart, A. T. Johnson, L. P. Kouwenhoven, D. J. Maas, W. de Jong, M. P. de Ruyter van Stevenick, A. van der Enden, and C. J. P. M. Harmans, Physica **189B**, 99 (1993).

139. J. Weis, R. J. Haug, K. von Klitzing, and K. Ploog, Phys. Rev. B **46**, 12837 (1992).

140. D. Weinmann, W. Häusler, W. Pfaff, B. Kramer, and U. Weiss, Europhys. Lett. **26**, 467 (1994).

141. D. Weinmann, W. Häusler, and B. Kramer, Phys. Rev. Lett. **74**, 984 (1995).

142. T. J. Nicholls, J. E. F. Frost, M. Pepper, D. A. Ritchie, M. P. Grimshaw, and G. A. C. Jones, Phys. Rev. B **48**, 8866 (1993).

143. K. A. Matveev and L. I. Glazman, Phys. Rev. B **54**, 10339 (1996).

144. S. Wolf and R. Tauber, *Silicon Processing For the VLSI Era, Volume 1: Process Technology,* (Lattice Press , Sunset Beach, CA, 1986).

145. M. J. Rack, A. D. Gunther, M. Khoury, D. Vasileska, M. Siderov, and D. K. Ferry, Semiconductor Science and Technology **13**, A71 (1998).

146. A. E. Morgan, K. N. Ritz, E. K. Broadbent, and A. S. Bhansali, J. Appl. Phys. **67**, 6265 (1990).

147. S. Q. Wang and J. W. Mayer, J. Appl. Phys. **67**, 2932 (1990).

148. T. P. Chow, K. Hamzeh, and A. J. Steckl, J. Appl. Phys. **54**, 2716 (1983).

149. S. Q. Wang and J. W. Mayer, Thin Solid Films **202**, 105 (1991).

150. A. Wallart, H. S. Zeng, J. P. Nys, G. Dalmaai, and P. Friedel, J. Appl. Phys. **69**, 8168 (1991).

151. U. S. Tandon and W. S. Kohl, *Patterning Material Layers in Submicron Region,* (John Wiley & Sons, New York, 1993).

152. L. D. Jackel, R. E. Howard, P. M. Mankiewich, H. G. Craighead, and R. W. Epworth, Appl. Phys. Lett. **45**, 698 (1984).

153. K. A. Valiev, *The Physics of Submicron Lithography,* (Plenum Press, New York, 1992).

154. B. R. Brewer and J. P. Ballantyne, *Electron-beam technology in microelectronic fabrication,* (Academic Press, New York, 1980).

155. M. J. Lercel, H. G. Craighead, A. N. Parikh, K. Seshadri, and D. L. Allara, Appl. Phys. Lett. **68**, 1504 (1996).

156. E. S. Snow and P. M. Campbell, Science **270**, 1639 (1995).

157. J. A. Dagata, Science **270**, 1625 (1995).

158. A. Kent, *Experimental Low-Temperature Physics,* (American Institute of Physics, New York, 1993).

159. C. A. Bailey, *Advanced Cryogenics,* (Plenum Press, New York 1971).

160. D. S. Betts, *An Introduction to Milikelvin Technology,* (Cambridge University Press, New York 1989).

161. S.M. Sze, *Physics of Semiconductor Devices, 2nd Ed.*, (Wiley Interscience & Sons, New York, New York, 1981).

162. D. K. Schröder, *Semiconductor Material and Device Characterization,* (John Wiley & Sons, Inc. New York, NY 1990).

163. S. A. Roy, *Simulation Tools for the Analysis of Single Electronic Systems*, Ph. D. Theisis, University of Glasgow, (June 1994).

164. J. D. Jackson, *Classical Electrodynamics*, 2nd. Ed.(John Wiley & Sons, New York, 1965).

165. M. Khoury, M. J. Rack, A. Gunther, and D. K. Ferry, Appl. Phys. Lett. **74**, 1576 (1999).

166. D. Vasileska, S. Milicic, and F. Badrieh, private communication.

167. C. Cohen–Tannoudji, B. Diu, F. Laloë, *Quantum Mechanics* (John Wiley & Sons, New York 1977), pp. 533.

168. M. Dür, A. D. Gunther, D. Vasileska, and S. M. Goodnick, Proceedings of the Sixth International Workshop on Computational Electronics, Osaka, Japan, 1998.

169. M. Dür, A. D. Gunther, D. Vasileska, and S. M. Goodnick, Proceedings of the IEEE Silicon Nanoelectronics Workshop 1998, Honolulu, HI USA, 1998.

170. T. Inoshita and H. Sakaki, Solid State Electron. **37**, 1175 (1994).

171. Al. L. Efros, V. A. Kharchenko, and M. Rosen, Solid State Commun. **93**, 281 (1995).

172. I. Vurgaftman and J. Singh, Appl. Phys. Lett. **64**, 232 (1994).

173. K. A. Matveev, L. I. Glazman, and H. U. Baranger, Surface Science **361/362**, 623 (1995).

174. K. A. Matveev, L. I. Glazman, and H. U. Baranger, Phys. Rev. B **53**, 1034 (1996).

175. K. A. Matveev, L. I. Glazman, and H. U. Baranger, Phys. Rev. B **54**, 5637 (1996).

176. F. R. Waugh, M. J. Berry, D. J. Mar, R. M. Westervelt, K. L. Campman, and A. C. Gossard, Phys. Rev. Lett. **75**, 705 (1995).

177. D. G. Austing, S. Sasaki, S. Tarucha (unpublished).

178. G. Breit and E. Wigner, Phys. Rev. **49**, 519 (1936).

179. G. Garcoia Calderoon and R. E. Peierls, Nucl. Phys. A **265**, 443 (1976).

180. M. Stopa, Phys. Rev. B **48**, 18340 (1993).

181. A. A. Staring, B. W. Alphenaar, H. van Houten, L. W. Molenkamp, O. J. A. Buyk, M. A. A. Mabesoone, and C. T. Foxon, Phys. Rev. B **46**, 12869 (1992).

182. T. Heinzel, A. T. Johnson, D. A. Wharam, J. P. Kotthaus, G. Böhm, W. Klein, G. Tränkle, and G. Weinmann, Phys. Rev. B **52**, 16638 (1995).

183. T. Sakamoto, S. W. Hwang, Y. Nakamura, and K. Nakamura, Appl. Phys. Lett. **65**, 875 (1994).

184. S. N. Milicic, D. Vasileska, R. Akis, A. Gunther, and S. M. Goodnick, accepted for presentation at the International Conference on Modeling and Simulation of Microsystems, San Diego, CA USA, 2000.

185. M. Koskinen, M. Manninen, and S. M. Reimann, Phys. Rev. Lett. **79**, 1389 (1997).

186. D. G. Austing, S. Sasaki, S. Tarucha, S. M. Reimann, M. Koskinen, and M. Manninen, Phys. Rev. B **60**, 11514 (1999).

187. A. V. Madhav and T. Chakraborty, Phys. Rev. B **49**, 8163 (1994).

188. C. W. J. Beenakker and H. van Houten, in *Single Charge Tunneling*, eds. H. Grabert, J. M. Martinis and M. H. Devoret (Plenum, New York, 1991).