Single Electron Charging Effects in Quantum Dot Nanostructures

by

Arvind Kumar

S.M., Electrical Engineering and Computer Science
Massachusetts Institute of Technology, 1988
S.B., Electrical Engineering and Computer Science
Massachusetts Institute of Technology, 1988

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1994

Author .....
Department of Electrical Engineering and Computer Science
April 29, 1994

Certified by .....
Dimitri A. Antoniadis
Professor of Electrical Engineering
Thesis Supervisor

Certified by .....
Terry P. Orlando
Professor of Electrical Engineering
Thesis Supervisor

Accepted by .....
Frederic R. Morgenthaler
Chairman, Departmental Committee on Graduate Students

Single Electron Charging Effects in Quantum Dot Nanostructures

by

Arvind Kumar

Submitted to the Department of Electrical Engineering and Computer Science
on April 29, 1994, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

This thesis focuses on the study of GaAs quantum dot devices, in which an electron gas is electrostatically confined to a small conducting island. The device dimensions are sufficiently small that striking effects due to the capacitive charging of the island by a *single* electron can be observed in the transistor characteristics, leading to a *periodic* dependence of the current on gate voltage. In particular, we study a quantum dot structure in which a novel gate geometry allows the island to be contacted by three electron reservoirs. When the dot charge is well-confined, periodic conductance oscillations due to Coulomb charging are observed in-phase with each other at two of the leads in response to a small excitation voltage at the third. As the tunnel barriers are made softer by changing the gate voltage, a strikingly different phenomenon is observed: conductance peaks at the two output leads evolve from perfect *correlation* to perfect *anti-correlation* with each other. Two simple models of transport in the weakly blockaded regime are presented as possible explanations.

Thesis Supervisor: Dimitri A. Antoniadis
Title: Professor of Electrical Engineering

Thesis Supervisor: Terry P. Orlando
Title: Professor of Electrical Engineering

Acknowledgments

First, I would like to thank my co-advisors Dimitri Antoniadis and Terry Orlando for the support they have shown during this project. I would also like to thank them for giving me the independence which allowed me to pursue several projects and collaborations during my graduate career. Finally, I would like to thank them for the financial and other support that enabled me to travel and meet with colleagues all around the world.

Although it has become a cliché by now, this thesis would have really been impossible without the help of many suberbly talented individuals.

I thank Michael Rooks of the National Nanofabrication Facility for his perfectionist lithography. I thank Michael Melloch of Purdue who supplied our group with very high quality heterostructures.

From the day that I met him, Gee Rittenhouse has been a true friend in every sense of the word. He selflessly took time out of his own work to teach me low-temperature measurements and helped me on so many occasions that I can only say thanks for everything.

I thank Nathan Belk whose generous help with the electronics was crucial to this thesis. Thanks are also due to Ethan Foxman, Paul Belk, and David Abusch for the many lunchtime discussions.

On the device side, I have had the pleasure to collaborate with Cris Eugster, whose fabrication wizardry was essential to this thesis. When I've needed a helping hand, Dave Carter has on numerous occasions gone far beyond the call of duty, and I wish him luck continuing this project. Martin Burkhardt has also been a continuing source of support and help, particularly with data acquisition. I am grateful to Jarvis Jacobs for his help with computer problems.

I have also had the pleasure to collaborate with some outstanding theorists. I especially thank Jari Kinaret for his persistent interest in our results and for inviting me to visit Nordita. Since I met Mike Stopa in Japan, he has on numerous occasions given me the benefit of his clear thinking, in addition to his optimism. I am also grateful to Konstatin Matveev for useful discussions.

Phil Bagwell taught me basic transport theory and instilled in me a great respect for the Landauer formula. We collaborated on a number of papers, which are reprinted in Appendix B.

I thank Prof. Hank Smith for his help in the NSL and for improving my speaking style by always demanding clearer explanations of me. Thanks are also due to Prof. John Graybeal, who allowed me the use of his measurement facilities during periods when I had none. I am grateful to Prof. Jesus del Alamo for help with fabrication issues.

I thank Philippe Debray for inviting me to spend three months at Saclay, where the magnetic field data in this thesis were taken. During my stay at Saclay, Jerome Blanchet helped me on countless occasions with equipment, rides, and French. I thank Julien Nagle of CSF Thomson for supplying us with wafers when we were in a bind. I also profitted greatly from Christian Glattli's deep understanding and insight and am sure that it will be a pleasure to work with him as a postdoc.

In parallel with my graduate career at MIT, I have had the pleasure to work with Frank Stern on the numerical modeling of quantum structures. To start, I thank him for introducing me to the field which later became the focus of my graduate work. Frank's treatment of me and of the people around him is truly exemplary. I especially thank him for the numerous occasions that he invited me to dinner during my many visits to Yorktown, and for hosting my visit to the Max Planck Institute. The papers on which he, Steven Laux, and I collaborated are reprinted in Appendix D.

I thank Steven Laux for teaching me all that I know about numerical methods.

I thank Dieter Kern and his group for allowing me access to their state-of-the-art electron beam lithography system during the early stages of this work.

I thank Cristina Gordy, Megan Jamison, Donna Martinez, and Kevin Meyer-Golden for the extremely efficient way in which they handled all my special administrative problems.

I thank Bill Chu, Yao-Ching Ku, and Ray Ghanbari for introducing me to the NSL, and Mark Mondol, Bob Sisson, and Jimmy Carter for the many evaporations I required. I owe a special thanks to Rich Perilli for bending over backwards to do evaporations when I desperately needed them.

I thank the Semiconductor Research Corporation for fellowship support and the IBM Corporation for funding for my visits to IBM Yorktown through a joint study agreement.

Finally I thank my family and friends for their encouragement.

Contents

List of Figures

Chapter 1

Introduction

1.1 Motivation and Background

Recent advances in nanolithography have made possible the fabrication of semiconductor structures in which electrons can be confined to a small conducting island, generally called a "quantum dot." At low temperatures, striking phenomena can be observed due to the capacitive charging of the island by just a *single* electron. This effect can be exploited to realize a "single electron transistor," so called because it can be cycled from an insulating to a conducting back to an insulating state on changing the effective channel charge by just one electron. Quantum dots also sometimes called "artificial atoms" since the effective channel consists of just a few tens or hundreds of electrons which are bound to discrete energy levels by a three-dimensional potential. This thesis focuses on the experimental study of quantum dot structures in gallium arsenide.

Because the blockage of current can be controlled by a gate electrode, such devices have potential applications as amplifiers, detectors, and switches. Thus, there is an immediate motivation from the basic research viewpoint to investigate whether novel effects in such devices might be exploited to set the stage for a possible revolution in electronics [1]. In

contrast to conventional field-effect transistor action, charging effects due to single electrons improve, rather than degrade, with shrinking dimensions. Finally, the natural discreteness of quantum-mechanical phenomena in parameters such as charge, energy, and spin could be exploited to encode bits. Upbeat prospectuses are given by Bate [2] and Likharev [3] whereas a more pessimistic outlook is given by Landauer [4].

The ability to fabricate structures in which electrons can be confined to reduce their effective dimensionality has also created a fertile playground for solid state physicists. Advances in microtechnology fabrication over the last decade have fueled discoveries in the field of "mesoscopic physics," loosely signifying the regime where both quantum mechanics and statistical mechanics apply. Much of the early work [5] concentrated on the study of transport in the *quantum diffusive regime*, where the device dimensions are smaller than the inelastic mean free path but longer than the elastic mean free path. As it became possible to fabricate smaller structures, interest turned to the *quantum ballistic regime*, where the device dimensions are shorter than both the elastic and inelastic mean free paths. An elegant demonstration of ballistic transport came in the discovery that the conductance of a narrow constriction – in which electrons are confined to a width comparable to the Fermi wavelength λ_F – increases in steps of $2e^2/h$ as each new one-dimensional channel is opened [6, 7].

These experiments, along with previous work in our group [8], probed the quantum-mechanical *wave* nature of the electron, observable when the confinement is comparable to the Fermi wavelength (about 50 nm in a typical two-dimensional inversion layer). The study of Coulomb blockade effects, which exploit the *granular* nature of the electron, was launched by a curious experiment on double-gate silicon MOSFETs by Scott-Thomas et al. [9]. In a dirty semiconductor with many impurities, it would be expected that the conductance would exhibit highly irregular (but reproducible) oscillations as the density is varied, due to impurity potential fluctuations. Thus it came as a complete surprise when the measured conductance was a *periodic* function of gate voltage, as shown in Fig. 1-1. Because the

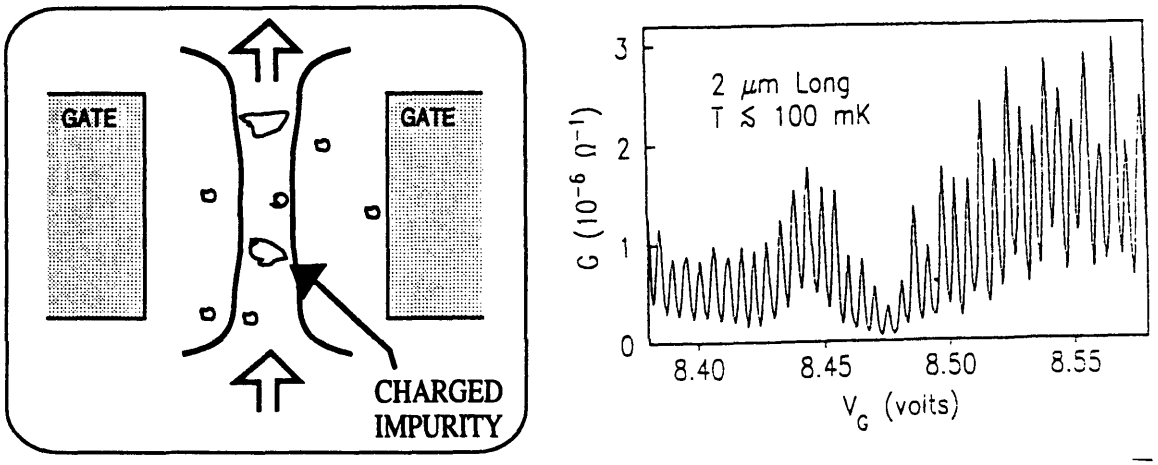
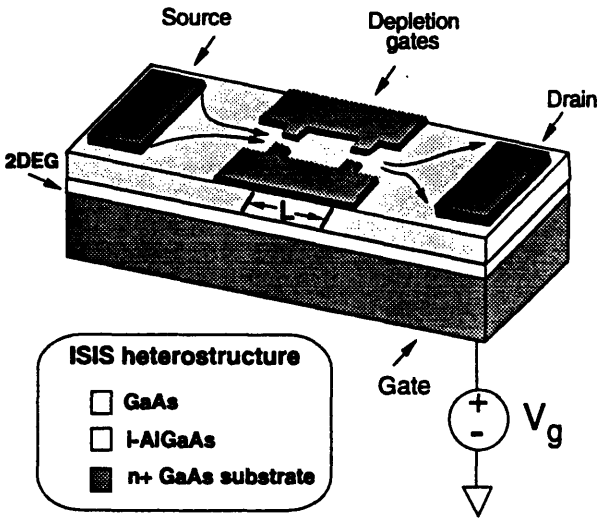


Figure 1-1: (a) Schematic of the electron channel in the structure of Scott-Thomas et al. [9]. Two dominant charged impurities were believed responsible for pinching off an isolated segment of the channel, giving rise to periodic conductance oscillations in (b).

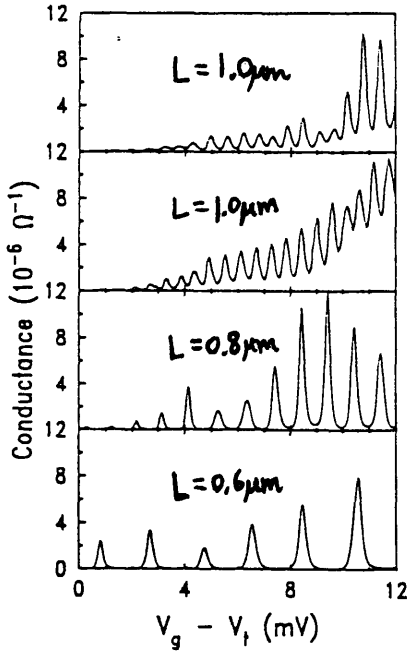
period of the oscillation varied randomly from device to device, it was hypothesized that two dominant charged impurities, positioned randomly in the channel, were responsible for the effect. The experimental failure of the peaks to undergo Zeeman splitting in a high magnetic field ruled out a simple explanation of resonant tunneling of *non-interacting* electrons.

Soon after the data of Scott-Thomas et al. [9] was published, van Houten and Beenakker [10] proposed the *Coulomb blockade* of *interacting* electrons as a possible explanation. Basically, they hypothesized that the two impurities delimited a segment of the channel, forming a natural quantum dot. The Coulomb blockade is based on the condition that the dot can contain only an integer number of electrons and hence its charge changes discretely in multiples of the electron charge. Each conductance peak represents the addition of a single electron to the dot.

The two-impurity hypothesis was verified by Meirav et al. [11], who used the structure shown in Fig. 1-2(a) to produce in a controlled fashion a narrow electron channel interrupted by two potential barriers. The structure consists of two Schottky gates used to squeeze



(a)



(b)

Figure 1-2: (a) Schematic of the device structure of Meirav et al. [11]. A negative voltage on the two Schottky gates on the top depletes the electron gas at the lower GaAs/AlGaAs interface below, forming a narrow electron channel interrupted by two potential barriers at the constrictions. The n+ GaAs substrates serves as a back gate to modulate the channel electron density. (b) Periodic oscillations in the conductance as a function of back gate voltage for four devices of the type shown in (a). The oscillation period correlates inversely with the spacing between constrictions L.

the electron gas laterally, while the n^+ substrate serves as a back gate to independently modulate the channel electron density. As shown in Fig. 1-2(b), the voltage separation between successive conductance peaks was found to depend on the separation between the two potential barriers.

In the next section, we discuss the structure studied in this thesis, but before doing so let us recap why this particular effect shows promise for future applications. The word "transistor" understates the functionality of this device: it can be swept through a *periodic series* of conducting and insulating states, with each peak in the conductance elegantly counting the addition of a *single* electron to the dot.

1.2 Coupled Quantum Dot Structure and Thesis Outline

This thesis was originally motivated by the prospect of studying *two* quantum dots in sufficiently close proximity that they can interact with each other. Coupling between neighboring quantum dots can take two forms. If the coupling between the dots is simply capacitive, changes in the electrostatic potential of one dot affect the potential of the other dot. If the potential barrier separating the dots is made sufficiently thin, however, the two dots can also exchange particles with each other. For future applications, a double-dot system would be the first step toward a quantum-based architecture in which two elements could interact with each other without interconnects. From the viewpoint of fundamental physics, a coupled dot system is interesting because of the interplay between two different charging energy scales in the problem.

Figure 1-3 shows conceptually two coupled quantum dots. The dots can be accessed by four contacts at the corners, allowing operation either (a) in parallel or (b) in series. As discussed in Sect. 2.1, the electrostatic potential of a dot varies periodically in a sawtooth fashion as the gate voltage is swept. When the dots operate in parallel (Fig. 1-3(a)), the periodic variation of the potential of one dot swept through a series of conductance peaks

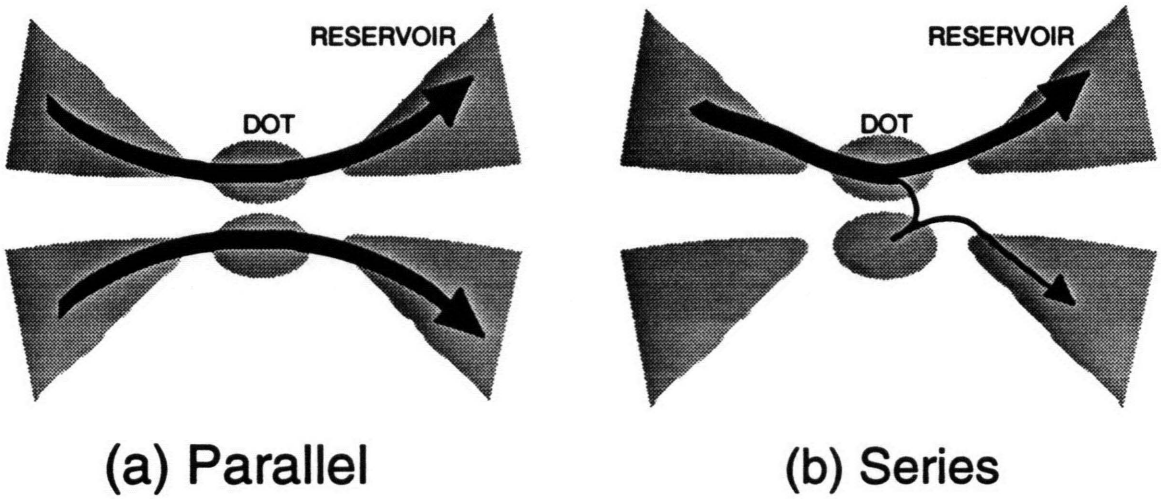


Figure 1-3: Conceptual picture of two coupled quantum dots: (a) shows current flow for study of parallel transport through dots; (b) shows current flow for study of series transport through dots.

should be detectable by a change in the current through the other dot. Such a capacitive effect has been studied by Field et al. [12], in which a narrow constriction was used as a "detector" rather than another dot. It would be particularly interesting to study the parallel configuration in the time domain, where it might be possible to detect the changing of the number of particles in one dot as a current flows through it.

The advantage of the proposed configuration is that it also offers the possibility to study two quantum dots operating in series, as shown in Fig. 1-3(b). Using this structure, the current through one of the dots can be compared directly with the series current through two dots. The theory of transport through two coupled dots in series is discussed in Sect. 2.1.

The introduction of an additional tunnel barrier to allow particle exchange between two dots opened up the possibility to perform an experiment not forseen at the outset. In the series configuration depicted in Fig. 1-3(b), an electron which has tunneled into the upper dot can exit the dot by tunneling into an empty state either in the upper right reservoir *or* in the lower dot. If the electrostatic confinement of the lower dot is removed, a single quantum dot contacted by three electron reservoirs is realized, as depicted in Fig. 1-4. The study of

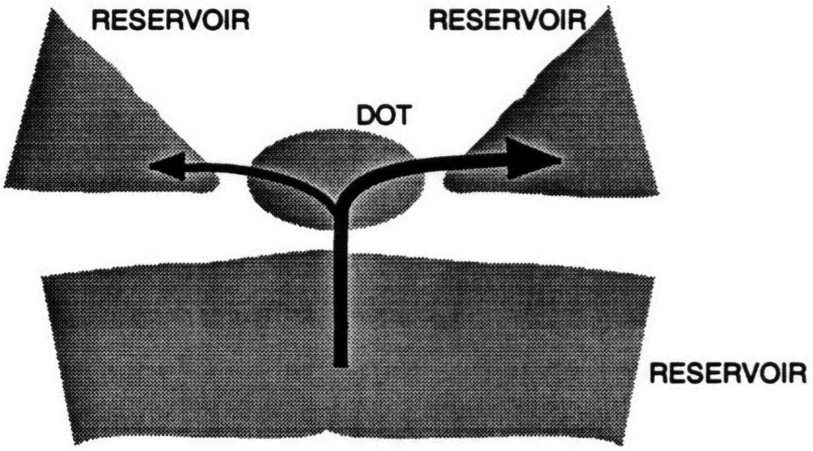


Figure 1-4: Electron distribution and current flow in a three-terminal quantum dot. The current into the dot divides itself between the left reservoir and the right reservoir.

a three-terminal quantum dot, not carried out previously, produced some fundamental new results reported in this thesis.

Chapter 2 reviews basic Coulomb blockade theory and recent experiments. Chapter 3 traces the path from designing devices to measuring them. Chapter 4 presents experimental results, focusing on the study of the three-terminal quantum dot. Chapter 5 investigates possible models for the experimental results. Chapter 6 is the conclusion and discusses directions for future work.

Chapter 2

Review of Theory and Experiments

2.1 Coulomb Blockade: Basic Concepts

The outline of this section is as follows. First, we understand the classical Coulomb blockade phenomenon by studying a metal grain, where energy quantization can be neglected. Then we generalize the treatment to the case of a semiconductor quantum dot in equilibrium. Finally, we summarize the approach taken by Beenakker [13] to calculate the current, which utilizes a non-equilibrium rate equation method.

Consider a metal grain with charge Q and at potential $\phi(Q)$ capacitively coupled to several conductors, as depicted in Fig. 2-1(a). The other conductors may be gate electrodes or the leads from which electrons can tunnel on to or off of the grain. The charge Q on the

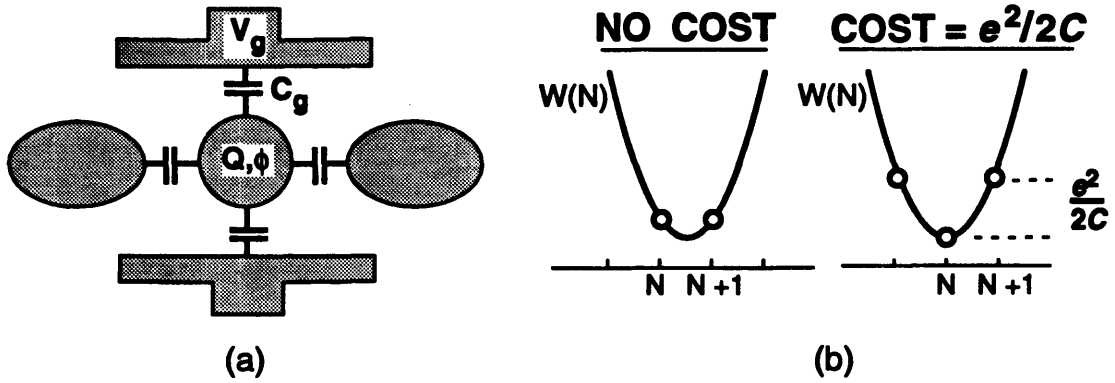


Figure 2-1: (a) Metal grain with charge Q at potential ϕ capacitively coupled to a number of other conductors, which may be leads or gates. (b) Electrostatic energy $W(N)$ showing conditions under which an electron can be added to the grain at no cost in energy and under which the energy cost is maximum.

grain is related to the voltages on the other conductors $\{V_i\}$ by the capacitances $\{C_i\}$:[1,2]

$$Q = \sum_i C_i(\phi - V_i). \tag{2.1}$$

Thus the electrostatic potential of the grain is

$$\phi(Q) = \frac{Q}{C} + \frac{1}{C}\sum_i C_i V_i, \tag{2.2}$$

where $C = \sum_i C_i$. The electrostatic energy required to bring the charge Q onto the grain is

$$W(Q) = \int_0^Q \phi(Q')dQ' = \frac{Q^2}{2C} + \left(\sum_i \frac{C_i}{C} V_i\right) Q = \frac{(Q - Q_0)^2}{2C} - \frac{Q_0^2}{2C}, \tag{2.3}$$

where $Q_0 = -\sum_i C_i V_i$. The grain contains an integer number N of electrons so that its

[1]The analysis assumes that the charge on the grain is fully imaged on the conductors and thus there is no charge at infinity.

[2]An alternate formulation uses capacities instead of capacitances. See Ref. [14].

charge is

$$Q = -Ne \tag{2.4}$$

and Eq. (2.3) becomes

$$W(N) = \frac{e^2}{2C}(N - N_0)^2 - \frac{e^2 N_0^2}{2C} \tag{2.5}$$

where $N_0 = -Q_0/e = \sum_i C_i V_i/e$. Figure 2-1(b) illustrates this parabolic dependence in two special cases. To minimize its energy, the grain must have N electrons when $N - \frac{1}{2} < N_0 < N + \frac{1}{2}$ and $N + 1$ electrons when $N + \frac{1}{2} < N_0 < N + \frac{3}{2}$. However, if N_0 is set to be $N + \frac{1}{2}$ exactly, then $W(N) = W(N + 1)$ so that the number of electrons can fluctuate freely between N and $N + 1$. As a result, current can flow. If the voltage V_g on one electrode is swept while the other voltages are fixed, this degeneracy occurs when

$$V_g = \frac{e}{C_g}\left(N + \frac{1}{2} - \sum_{i \neq g} \frac{C_i V_i}{e}\right) \tag{2.6}$$

yielding a periodicity e/C_g in V_g. When V_g is not at one of these special degeneracy points, the flow of current through the grain is impeded by the *Coulomb blockade*; when $V_g = eN/C_g$, the Coulomb charging energy takes on its maximum value of $e^2/2C$.

Quantum-mechanical effects become important when the Fermi wavelength of the electrons is comparable to the length scale of the confinement. This condition is much easier to satisfy in a semiconductor with low carrier density than in a metal. As a first approximation, the effect of energy quantization can be accounted for by adding the energies of the occupied single-particle levels to the electrostatic energy. Using Eq. (2.5), we obtain for

the ground-state energy $U(N)$

$$U(N) = W(N) + \sum_{p=1}^{N} E_p = \frac{(Ne)^2}{2C} - Ne\phi_{ext} + \sum_{p=1}^{N} E_p, \tag{2.7}$$

where $\phi_{ext} = \sum_i C_i V_i / C$.[3] Here the $\{E_p\}$ are the single-particle energy levels, relative to the bottom of the potential well, in the *bare* confining potential ϕ_{ext}. Quite generally one can calculate the energy cost of adding the $(N+1)^{st}$ electron to a dot with N electrons as

$$\mu(N+1) = U(N+1) - U(N). \tag{2.8}$$

For the simplified $U(N)$ in Eq. (2.7), we find

$$\mu(N+1) \;=\; (N+\frac{1}{2})\frac{e^2}{C} - e\phi_{ext} + E_{N+1}. \tag{2.9}$$

The quantity $\mu(N+1)$ is the electrochemical potential, which is composed of the chemical potential E_{N+1} plus the electrostatic potential energy $-e\phi(Q)$ with effective charge $Q = -(N+1/2)e$. Inclusion of the discrete energy levels leads to a spacing in gate voltage between successive current peaks of

$$\Delta V_g = \frac{e}{C_g} + \frac{C}{C_g}\frac{E_{N+1} - E_N}{e} = \frac{C}{C_g}\frac{\mu(N+1) - \mu(N)}{e}, \tag{2.10}$$

where the factor C_g/C is the "lever arm" by which the gate voltage moves the bottom of the potential well. For a set of energy levels with twofold spin degeneracy, ΔV_g will alternately have and not have a contribution due to the energy level spacing. For the experiment of Meirav et al. [11], $e/C_g \approx 800$ μV whereas the energy level spacing is estimated to be 100 μeV [15, 16], so that the classical model is still a good lowest-order approximation.

In order for current to flow, the dot must be coupled to a reservoir with which it can

[3] A term accounting for the potential due to charged donor ions can also be included in ϕ_{ext}.

exchange particles. Consider the series of diagrams in Fig. 2-2 as the gate voltage sweeps the device from a conductance minimum to a maximum to a minimum. The quasi-Fermi levels of the two reservoirs are displaced by a very small voltage V_{ds}. At the conductance minima (diagrams 1 and 3), the gate voltage sets ϕ_{ext} so that the quasi-Fermi level of the reservoirs lies midway between successive electrochemical potentials. In this case, there is a *Coulomb gap* of approximately e^2/C between the highest filled state and the lowest empty state: an energy of $e^2/2C$ is required to bring an electron at the quasi-Fermi level to the lowest empty state in the dot or to bring an electron at the highest filled state in the dot to the quasi-Fermi level. At the conductance maximum, the electrochemical potential lies between the quasi-Fermi levels of the reservoirs. When an electron from the left reservoir tunnels onto the dot, all the levels are shifted up by e^2/C, which prevents another electron from tunneling into the dot. The electron can then tunnel into an empty state in the right reservoir, the levels shift back down, the process repeats itself, and current flows. A key point is that, at a conductance maximum, the electrons flow through the dot *one at a time.*

The above description is sufficient to understand the periodicity of the oscillations but not their amplitude or width. The calculation of the conductance of a quantum dot was carried out independently by Beenakker [13] using a rate-equation approach and by Meir, Wingreen, and Lee [17] using an Anderson model. Here we will summarize the main result of Ref. [13], along with the underlying assumptions. A very clear, pedagogical introduction is given in Ref. [16].

The reservoirs are assumed to be composed of a continuum of states whose occupancy is given by the Fermi-Dirac distribution function

$$f(E) = \frac{1}{1+\exp[E/k_BT]}, \tag{2.11}$$

where for convenience we reference all energies relative to the Fermi energy. The dot is assumed to be composed of a number of single-particle levels $\{E_p\}$, each of which is

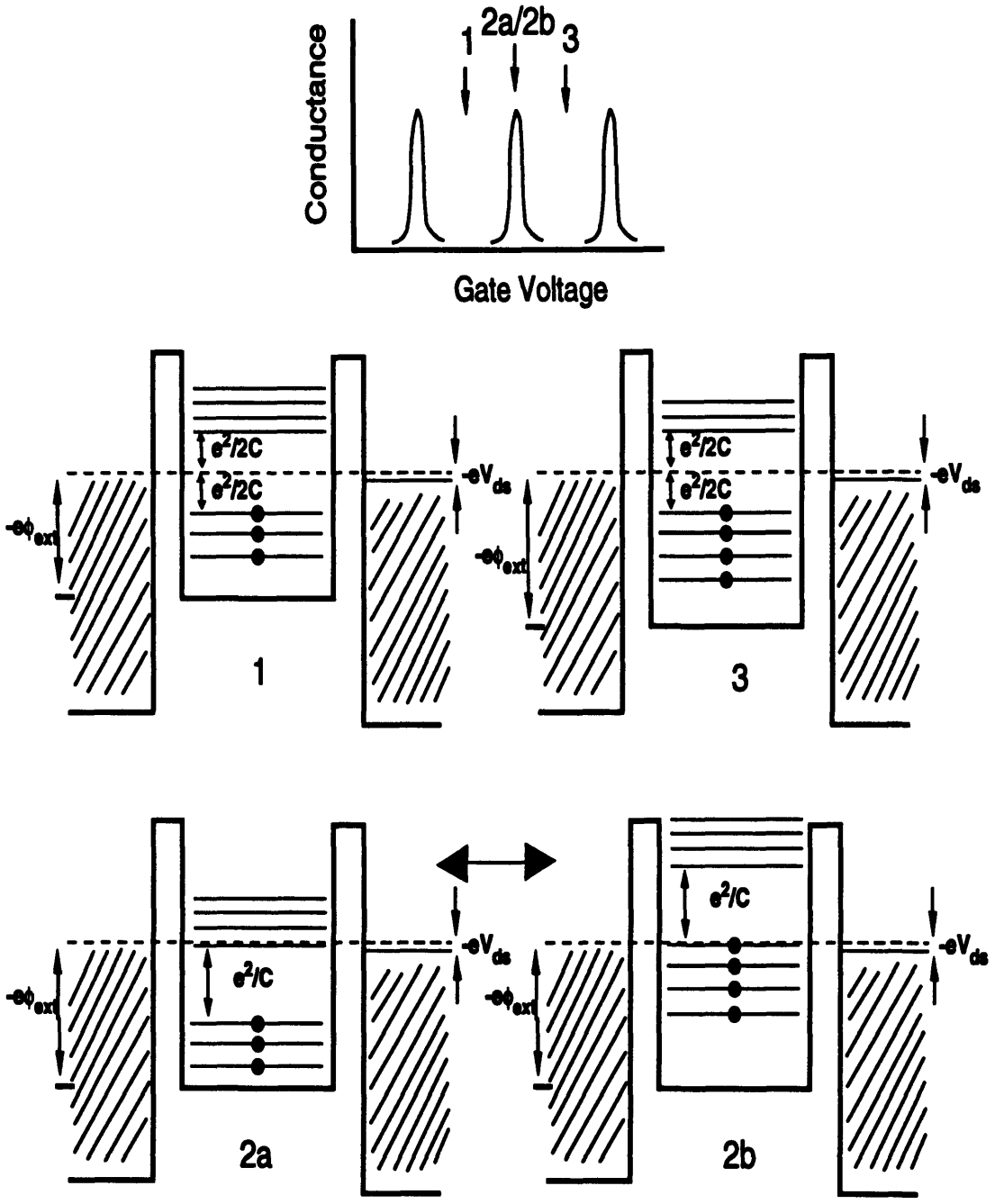


Figure 2-2: Quantum dot coupled to two reservoirs showing conditions under which current is blockaded (1 and 3) and under which the Coulomb blockade is lifted (2a and 2b).

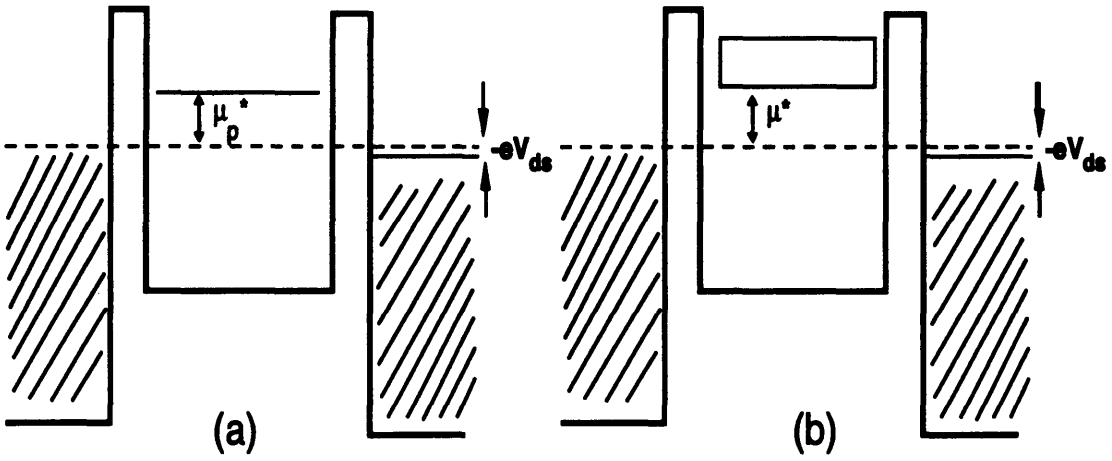


Figure 2-3: Limiting cases of (a) a single level and (b) a continuum of levels.

characterized by tunnel rates Γ_p^l and Γ_p^r (with $k_BT \gg h(\Gamma_p^l + \Gamma_p^r)$) to the left and right reservoirs, respectively. The state of the quantum dot is characterized by a set of occupation numbers $\{n_p\}$, where $n_p = 1$ if level p is occupied and $n_p = 0$ if it is unoccupied. The probability that the quantum dot is in a particular state is found by solving a kinetic equation.

The main result of Ref. [13] is that the linear conductance (i.e., $eV_{ds}/k_BT \to 0$) is given by

$$G = \frac{e^2}{k_BT} \sum_{p=1}^{\infty} \sum_{N=0}^{\infty} \frac{\Gamma_p^l \Gamma_p^r}{\Gamma_p^l + \Gamma_p^r} P_{eq}(N)[1 - F_{eq}(E_p|N)] f(\mu_p^*), \tag{2.12}$$

where $\mu_p^* = E_p + W(N+1) - W(N)$. Here $P_{eq}(N)$ is the equilibrium probability that the dot contains N electrons, evaluated in the grand canonical ensemble; $F_{eq}(E_p|N)$ is the conditional probability that level p is occupied given that there are N electrons in the dot. Physically, the product of the distribution functions represents the probability that an electron in a filled state of the reservoir has the requisite energy to tunnel into an empty state of the dot. If $k_BT \ll e^2/2C$, $P_{eq}(N)$ will be heavily weighted about one particular

value of N.

Two limiting cases, depicted in Fig. 2-3, are of interest. If $k_BT \ll \Delta E \ll e^2/C$ (ΔE is the energy level spacing), only a single level p of the dot is accessible. Then Eq. (2.12) reduces to

$$G = -e^2 \frac{\Gamma_p^l \Gamma_p^r}{\Gamma_p^l + \Gamma_p^r} f'(\mu_p^*) = \frac{e^2}{4k_BT} \frac{\Gamma_p^l \Gamma_p^r}{\Gamma_p^l + \Gamma_p^r} \cosh^{-2}\left(\frac{\mu_p^*}{2k_BT}\right), \tag{2.13}$$

where $f'(E)$ is the derivative of the Fermi-Dirac function. The classical limit, originally studied by Kulik and Shekhter [19], is obtained when $\Delta E \ll k_BT \ll e^2/C$. In this case, the summation over p reduces to an integral, yielding

$$G = G_{max} \frac{\mu^*/k_BT}{\sinh(\mu^*/k_BT)} \approx G_{max} \cosh^{-2}\left(\frac{\mu^*}{2.5k_BT}\right), \tag{2.14}$$

$$G_{max} = \frac{e^2}{2\Delta E} \frac{\Gamma^l \Gamma^r}{\Gamma^l + \Gamma^r}, \tag{2.15}$$

where Γ^l, Γ^r are evaluated at μ^*.

The $1/k_BT$ dependence obtained for a single level is cancelled by the number $k_BT/\Delta E$ of levels that participate in the conduction. However, it is interesting to note that the peak conductance G_{max} is only one-half the classical series conductance obtained in the limit $k_BT \gg e^2/C$ when the particle granularity is unimportant. The physical reason for this factor of 2 is that the electrons can pass through the dot only one at a time, even though the Coulomb blockade is lifted.

As the temperature is increased from $k_BT \ll \Delta E$ to $k_BT \gg \Delta E$, the amplitudes evolve from having a T^{-1} dependence to being temperature independent. The lineshape maintains the same approximate functional form, but the FWHM is roughly 1.25 larger in the classical limit. The transition from the resonant tunneling to the classical regime is nicely illustrated

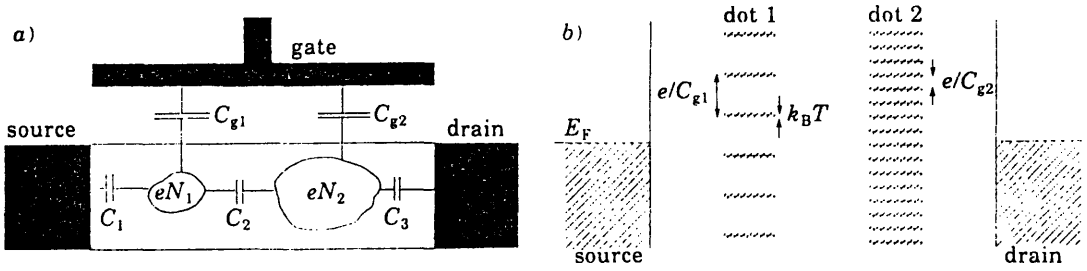


Figure 2-4: (a) Schematic of a two-dot system. The two dots are capacitatively coupled to each other, gate, and leads. (b) Energy ladders for a two-dot system in the limit $C_1 = C_2 = C_3 \ll C_{g1}, C_{g2}$. The incommensurate energy scales cause the conductance to evolve from periodic resonances at higher temperatures to quasi-random oscillations at lower temperatures. From Ref. [18].

in Ref. [16].

Some theoretical work on Coulomb blockade transport through two coupled dots connected in series has been done recently [18]. Figure 2-4 depicts two dots in series, coupled by capacitances C_{g1} and C_{g2} to the gate electrode. The main result is that, if $C_{g1} \neq C_{g2}$, then the conductance will be marked by periodic oscillations at higher temperatures which give way to quasi-random oscillations at lower temperatures. The reasoning behind this is that, at higher temperatures, the conductance of only the dot with greater charging energy will be modulated by Coulomb blockade; states of the other dot will appear as a continuum. As the temperature is lowered further, transport through both dots will be modulated by Coulomb blockade, but generally it will not be possible to simultaneously align states in the two dots with each other. A conductance maximum occurs when a state in one dot is aligned with the quasi-Fermi level, but electron hopping to or from the corresponding state in the other dot will be thermally activated. This thermal activation at the conductance

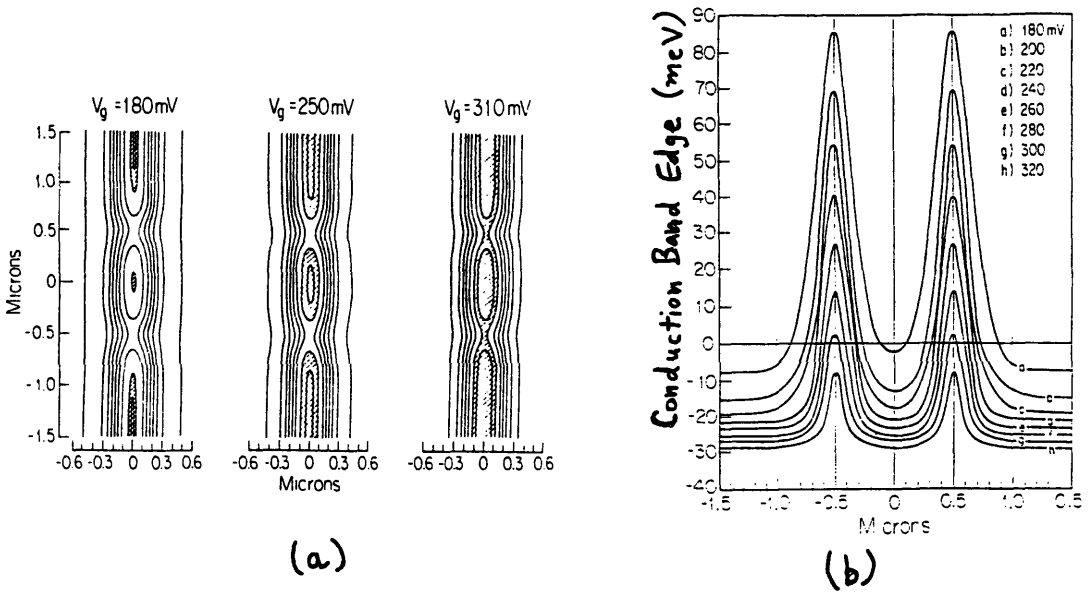


Figure 2-5: (a) Equipotential contours calculated for the structure of Meirav et al. [11]. (b) Conduction band edge versus position along a longitudinal cut through the center of the dot.

maxima is in contrast to the single dot case, where the conductance peaks are classically temperature-independent.

2.2 Numerical Simulations

The previous section has given a textbook description of Coulomb charging effects in idealized systems. Numerical simulations are often a useful tool to better understand Coulomb and quantum-mechanical effects in actual experimental structures. Through a numerical simulation, one can obtain approximate potential profiles and electron densities for arbitrary layer structures, device geometries, and bias conditions. This section summarizes the results of several papers written in collaboration with Frank Stern and Steve Laux of IBM Yorktown; Refs. [20, 21, 22] are reprinted in Appendix D.

Following the initial observation of periodic conductance oscillations, we carried out numerical simulations to help determine the applicability of the Coulomb blockade pic-

ture [22, 23, 24] which was controversial at the time [25, 26]. Figure 2-5(a) shows calculated equipotential contours in the plane of the 2DEG (at the GaAs/AlGaAs interface) for the structure of Meirav et al. [11]. The rounding of the contour lines can be understood in terms of the attenuation of high wavevector Fourier components of the potential, as discussed in Ref. [20]. The areas occupied by electrons, such that the Fermi level is above the conduction band edge, are shaded. A well-defined dot between the tunnel barriers is formed which evolves into a continuous channel with increasing back gate voltage. This is also seen in Fig. 2-5(b), which shows the conduction band edge along a longitudinal cut through the center of the dot for various back gate voltages. To make connection with experiment, we calculated the number of electrons in the dot as a function of back gate voltage V_g, as plotted in Fig. 2-6. For experimental structures having periodicities of 0.6, 1.0, and 1.8 mV (see Fig. 1-2(b)), we found the bottom gate voltage required to add a single electron to the island to be 0.65, 0.9, and 1.2 mV, respectively.[4] Even with uncertainties about doping levels and other parameters, *the one electron per oscillation picture is the only plausible conclusion.*

The simulations are done by numerically solving the three-dimensional Poisson equation

$$\nabla \cdot \epsilon \nabla \phi(x, y, z) = \rho(\phi) \tag{2.16}$$

iteratively, where $\phi(x, y, z)$ is the electrostatic potential, ϵ is the dielectric permittivity, and the charge density $\rho(\phi)$ includes contributions from electrons, holes, and ionized dopants. The above simulations were done in the Thomas-Fermi approximation, where the electron charge is evaluated semiclassically. The Thomas-Fermi approximation is sufficient for calculating quantities such as capacitance and potential profiles, but ignores the quantum mechanics which governs the electron distribution in the dot.

[4] These numbers are extracted from the nearly linear regime indicated in Fig. 2-6. In this regime, the Fermi level just crosses the tops of the tunnel barriers, giving reasonable values of the transmission coefficients when the lateral confinement is taken into account.

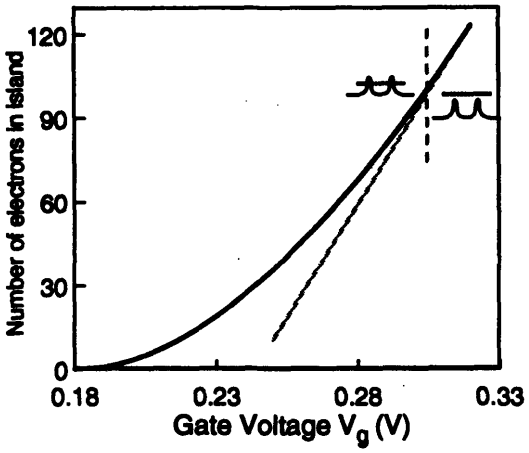


Figure 2-6: Calculated number of electrons in dot as a function of gate voltage.

Quantum mechanics is included by using the Schrödinger equation to evaluate the electron charge density. The self-consistent problem in Eq. (2.16) is rendered much more difficult because changes in the potential at one location affect the quantum charge everywhere, resulting in a *nonlocal* dependence of charge on potential.[5] In the Hartree approximation, the Poisson equation (2.16) is solved self-consistently with the one-electron Schrödinger equation, yielding the single-particle eigenstates. Self-consistent calculations in the Hartree approximation were first carried out for a three-dimensional GaAs quantum dot [27] by Kumar, Laux, and Stern [21]. In three dimensions the calculations are very time consuming and thus limited to dots containing only a few (< 12) electrons. However, Stopa [28] has recently introduced a technique in which the effective dimensionality can be reduced from three to two, making possible a self-consistent solution for a dot containing 70 electrons. Reference [28] also generalizes Eq. (2.12), the master equation for the conductance, to the case of self-consistently calculated energy levels.

[5]Reference [14] discusses the implications of this nonlocal dependence on the definition of capacitance.

The Hartree approximation is a single-particle picture which omits many-body effects [29]. The consequences of many-body interactions on the energy spectrum are discussed for model potentials in Refs. [30, 31, 32, 33, 34]. One error made in the single-particle Hartree approximation that an electron sees the effect of its own charge on the potential which determines its wave function. Such an error worsens as the number of particles is reduced. In the local density functional approximation this error can be partially compensated by reducing the effective potential energy in the Schrödinger equation at places of higher electron density [35]. However, little change was found in the electron addition spectrum of a realistic device geometry when this correction was included in the local density functional approximation [23].

2.3 Review of Experiments

Since the initial work of Scott-Thomas et al. [9] and Meirav et al. [11], there has been a tremendous interest in the field of few-electron semiconductor nanostructures. Naturally most experiments have focused on gated structures in which a quantum dot is defined in a controlled fashion; however, periodic oscillations have also been observed in disordered GaAs channels [36, 37], in a Si quantum point contact [38], in nearly insulating indium wires [39], and in a GaAs gate finger possibly damaged during processing [40]. This section is an attempt to summarize some of the many recent developments. Some reviews of early experiments are given in Refs. [41, 42].

The evolution of the single-particle eigenstates of a quantum dot with a magnetic field has been a subject of great interest. Basically, the classical circular orbits interact with the walls of the confining potential, producing the well-known "edge states," as explained nicely in an article by Lent [43]. It seems plausible that one could map out the single particle energy spectrum of the dot as a function of magnetic field simply by subtracting off the e/C_g term from the spacing between successive peaks in Eq. (2.10). Furthermore, variation of the

amplitudes of the peaks with magnetic field gives information about the coupling between the lead and the particular eigenstate of the dot through which conduction occurs. This was the broad motivation for the original work by McEuen et al. [44], using the quantum dot structures of Meirav et al. [11]. However, this simple picture was later found to be wrong [45]. The main fault is that the simple decomposition of the energy (2.7) into electrostatic and single-particle level terms is really not valid; instead the energy of the interacting electron system must be calculated self-consistently with the quantum-mechanical density of the electrons, as discussed in Sect. 2.2.

A device structure is typically composed of several gates, each of which has its own capacitance "lever arm" to the dot. The period of the conductance oscillations depends on the capacitance of the particular gate being swept. By contrast, the charging energy $e^2/2C$ required to bring an electron onto the dot at a conductance minimum depends on $C = \sum_i C_i$, where the sum is over all conductors. Since the finite valley conductance is thermally activated, this charging energy can be estimated by fitting the conductance to a thermally broadened lineshape. A puzzle in early experiments was that the sum of the gate capacitances obtained from the various periodicities was appreciably smaller than the value of C extracted from lineshape analysis [25, 46].

In later work by Foxman et al. [15], it was hypothesized that the missing capacitances are the capacitances between the dot and the source/drain leads. To estimate the lead capacitances, they performed a *nonlinear* measurement. Application of a finite (but not too large) drain-source bias V_{ds} allows one to probe excited states of the dot *at fixed number.* This leads to additional conductance resonances which can be mapped out in the $V_g - V_{ds}$ plane. If one traces these resonances in the $V_g - V_{ds}$ plane, they will move with slope $dV_{ds}/dV_g = C_g/(C - C_{lead})$, allowing a direct measurement of C_{lead}. This relation follows directly from Eq. (2.9) by equating μ with the shifted Fermi level $-eV_{ds}$ of the biased lead. The physical reason for this formula is that changing the lead voltage has *two* effects – namely, to shift the bottom of the potential well *and* to alter the resonance condition.

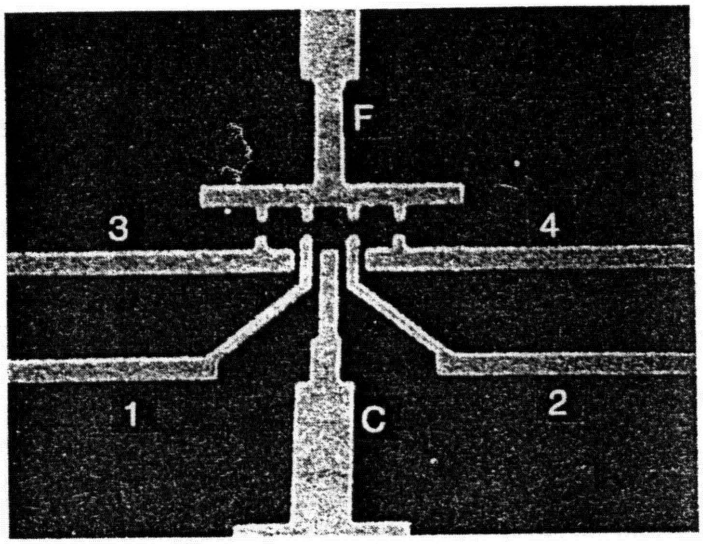


Figure 2-7: Scanning electron micrograph of the gate geometry of the structure used by Kouwenhoven et al. [48]. Gates 1 and 2 are used to form independently-controlled QPC's. Gates 3 and 4 are not used.

The quantum dot structure of Meirav et al. [11] was fabricated on an inverted GaAs heterostructure [47], which is not widely available. Kouwenhoven et al. [48] were able to reproduce the basic results of Meirav et al. [11] using the geometry shown in Fig. 2-7 deposited on a conventional modulation-doped GaAs/AlGaAs heterostructure, whose principle is described in Sect. 3.1. This geometry has become a prototype for the many further transport experiments carried out in GaAs quantum dots.

Two particularly noteworthy results became possible because of the ability to control the left and right quantum point contacts (QPCs) *independently* using gates 1 and 2. The conductance through each QPC alone is the number of propagating one-dimensional channels times the quantum of conductance $\sigma_Q \equiv 2e^2/h$. If the leak rate through either QPC is too large, the number of electrons in the dot will not be well-defined and there will be no Coulomb blockade. Roughly, this will occur when the conductance of the QPC is $\sim \sigma_Q$ since the leak rate will be $\sim \sigma_Q/C$ leading to an energy uncertainty $\sim e^2/C$. Kouwenhoven et

al. [48] found that the Coulomb blockade oscillations disappear just when the conductance through one of the QPCs exceeds σ_Q. However, it should be noted that this may not be an absolute upper limit, as Pasquier et al. [49] have found that small oscillations persist even when both QPCs have conductances exceeding σ_Q.

By modulating the two QPCs with two phase-shifted rf signals, Kouwenhoven et al. [50] have implemented a "turnstile" device, in which only an integer number of electrons can tunnel through the dot during each rf cycle. The current is therefore an integer multiple of the electron charge times the rf frequency, with the integer controlled by the drain-source voltage. Such a device might have applications as an accurate current standard.

With finite dot-lead coupling, there are corrections to the single electron tunneling picture leading eventually to the Ohmic regime where no oscillations are observable. The first correction is a process called co-tunneling, in which current flows through the dot by simultaneous tunneling of an electron from source to dot and from dot to drain [51, 52]. This problem has been studied in depth in experiments by Pasquier et al. [49], who found that a co-tunneling model accounts well for the finite conductance obtained at the valleys. Basically, co-tunneling with an algebraic dependence on temperature dominates over simple thermal activation as the dot-lead coupling increases with gate voltage. This is to be contrasted with the work of Ref. [15], where the washout of the Coulomb blockade is viewed as a geometrical effect in which the divergence of a capacitance quenches the charging energy.

In the structure of Field et al. [12], a quantum dot is coupled to a narrow constriction, as shown in Fig. 2-8. As the dot is swept through the Coulomb blockade oscillations, the potential of the dot changes in a sawtooth fashion. In the tunneling regime, the resistance of the constriction is very sensitive to changes in the potential of the quantum dot, with which it is capacitively coupled. If a calibration procedure is applied to the resistance of the constriction, the electrostatic potential of the dot can be obtained.

In a high perpendicular magnetic field, transport through the dot occurs via the edge

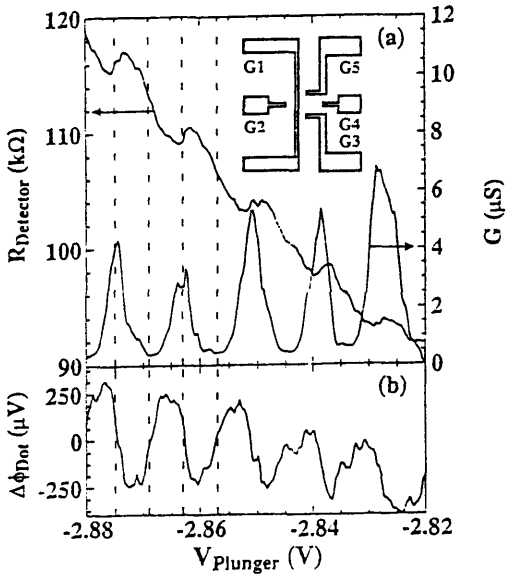


Figure 2-8: (a) Coulomb blockade oscillations and resistance through the constriction "detector" for the structure of Field et al. [12], pictured in the inset. (b) Dot potential obtained from calibration procedure applied to constriction resistance.

channel of the lowest Landau level and is essentially one-dimensional. Because of the absence of backscattering, the chance for *coherent* resonant tunneling, in which the electron maintains its phase memory, is greatly improved. Johnson et al. [53] have studied this regime, where the transmission probability through both barriers can approach unity even though each barrier, taken by itself, is highly reflective. Alphenaar et al. [54] have studied the interplay of Coulomb blockade with the Aharanov-Bohm effect, which also results in periodic conductance oscillations. Finally, Weis et al. [55] have studied a quantum dot in a *parallel* magnetic field, so that spin-splitting effects are enhanced relative to orbital-splitting effects.

All of the experiments discussed so far have focused on probing the properties of quantum dots by passing a small current through the dot, known as transport spectroscopy. Semiconductor quantum dots have also been studied by capacitance spectroscopy and by far infrared spectroscopy. In capacitance spectroscopy [27, 56], the dot is coupled to only

one reservoir, and a peak is observed in the capacitance signal every time the Fermi level of the reservoir aligns with the energy required to add a new electron to the dot. Single electron effects have also been observed dramatically using far infrared spectroscopy. In the experiments of Meurer et al. [57], it was found that the absorption increased in a stepwise fashion as a single electron was added to each dot in a large array.

Finally, it should be mentioned that the study of single electron effects in metals has also been a very active field. Indeed the basic Coulomb blockade explanation can be traced back to Gorter [58] in 1954. Coulomb charging effects have been studied in a metallic island isolated by tunnel junctions [59] and in a metal droplet coupled to an STM tip [60]. Reference [61] discusses some applications of practical interest in metrology.

Chapter 3

Device Design, Fabrication, and Measurement

3.1 GaAs/AlGaAs Heterostructures

Formation of a two-dimensional electron gas (2DEG), whose physical properties are reviewed in Ref. [62], is the first step in achieving confinement to create a quantum dot. To date the highest electron mobilities have been attained using modulation-doped GaAs/AlGaAs heterostructures grown by molecular beam epitaxy, in which the 2DEG is physically separated from the ionized dopants by a spacer layer. Cross sections of two modulation-doped GaAs/AlGaAs heterostructures used in this work are shown in Fig. 3-1. The calculated conduction band edge for wafer MBE18 is plotted as a function of position in Fig. 3-2. Note that surface states pin the Fermi level at the surface near the middle of the bandgap. The conduction band edge discontinuity at the GaAs/AlGaAs interface confines electrons in a steep triangular potential of which only the lowest subband is populated. The density of carriers in the 2DEG is proportional to the donor doping in the AlGaAs.

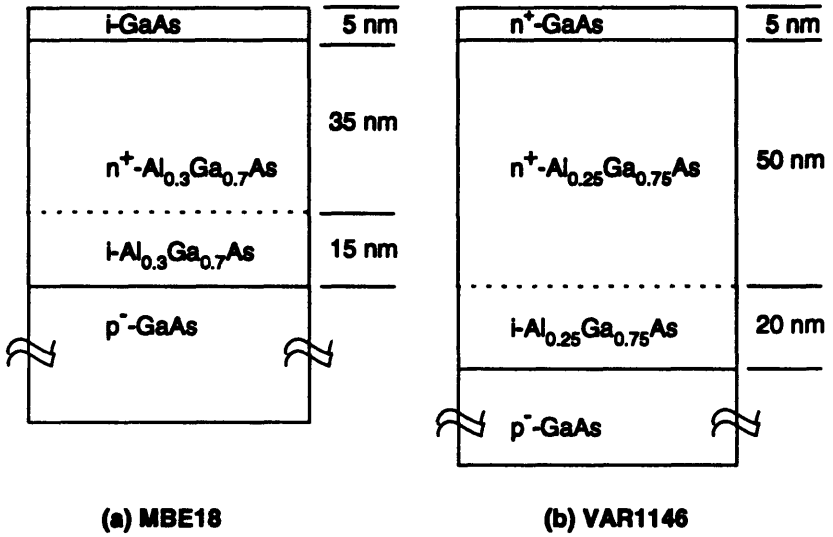


Figure 3-1: Two modulation-doped GaAs/AlGaAs heterostructures used in this work. The structures are grown by molecular beam epitaxy over a semi-insulating GaAs substrate. The n^+ regions are typically Si-doped 10^{18} cm^{-2}. (a) Shallow structure with mobility 2.5×10^5 cm^{2}V^{-1}s^{-1} and electron density 3×10^{11} cm^{-2}. (b) Deeper heterostructure with mobility 5.8×10^5 cm^{2}V^{-1}s^{-1} and electron density 4.6×10^{11} cm^{-2}.

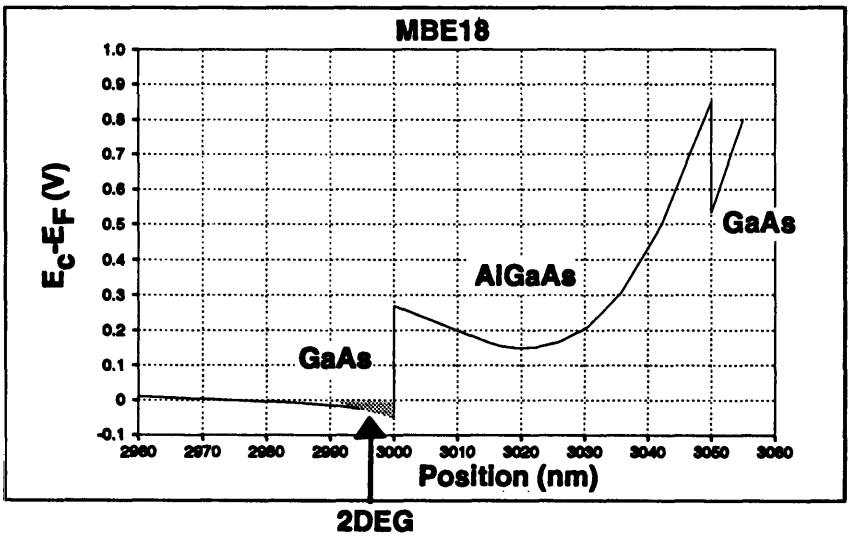


Figure 3-2: Conduction band edge as a function of position for heterostructure MBE18, found by solution of Poisson's equation. Surface states pin the Fermi level at the surface near the middle of the bandgap. The conduction band edge discontinuity at the GaAs/AlGaAs interface confines electrons in a steep triangular potential of which only the lowest subband is populated, resulting in formation of a 2DEG.

Heterostructure MBE18, grown by Prof. Michael Melloch of Purdue University, has a low temperature mobility of 2.5×10^5 cm^{2}V^{-1}s^{-1} and electron density 3×10^{11} cm^{-2}. The relatively low mobility, corresponding to a mean free path of 1.6 μm, is a consequence of the thin spacer layer. Heterostruture VAR1146, grown by Dr. Julien Nagle of CSF Thomson, has a low temperature mobility of 5.8×10^5 cm^{2}V^{-1}s^{-1} and electron density 4.6×10^{11} cm^{-2}. The tradeoff of the higher mobility is a larger separation between the gates and the 2DEG, resulting in less sharply confining potentials [20, 63].

3.2 Electron Confinement and Device Design

The most versatile way to achieve one-dimensional confinement on a GaAs/AlGaAs heterostructure is by the *split-gate technique* illustrated in Fig. 3-3. As a negative voltage is applied, the electrons directly underneath the gates are the first to be depleted. The voltage required to fully deplete the carriers directly underneath the gate is called the 2-d threshold, V_{2d}. Electrons in the channel are then depleted by the fringing fields of the gates, but at a slower rate than in the 2-d regime. The voltage at which the channel is fully pinched off is called the 1-d threshold, V_{1d}. Davies [64] has derived a useful expression for the ratio of these two threshold voltages: [1]

$$\frac{V_{1d}}{V_{2d}} = (1 - \frac{2}{\pi} \arctan \frac{w}{2d})^{-1}, \tag{3.1}$$

where w is the slit width between the gates and d is the depth of the 2DEG below the surface. The w/d ratio should be chosen so that a reasonable 1-d regime is obtained without inducing appreciable gate leakage. For typical doping levels, $V_{2d} \sim -0.5$ V, so that a width-to-depth ratio $w/d \sim 3$ yields a reasonable value of $V_{1d} \sim -1.3$ V. A comprehensive study of the conditions required for observation of the 1-d regime is given in Ref. [65].

[1] Equation (3.1) assumes a uniform dielectric constant, which is a good approximation for GaAs and AlGaAs.

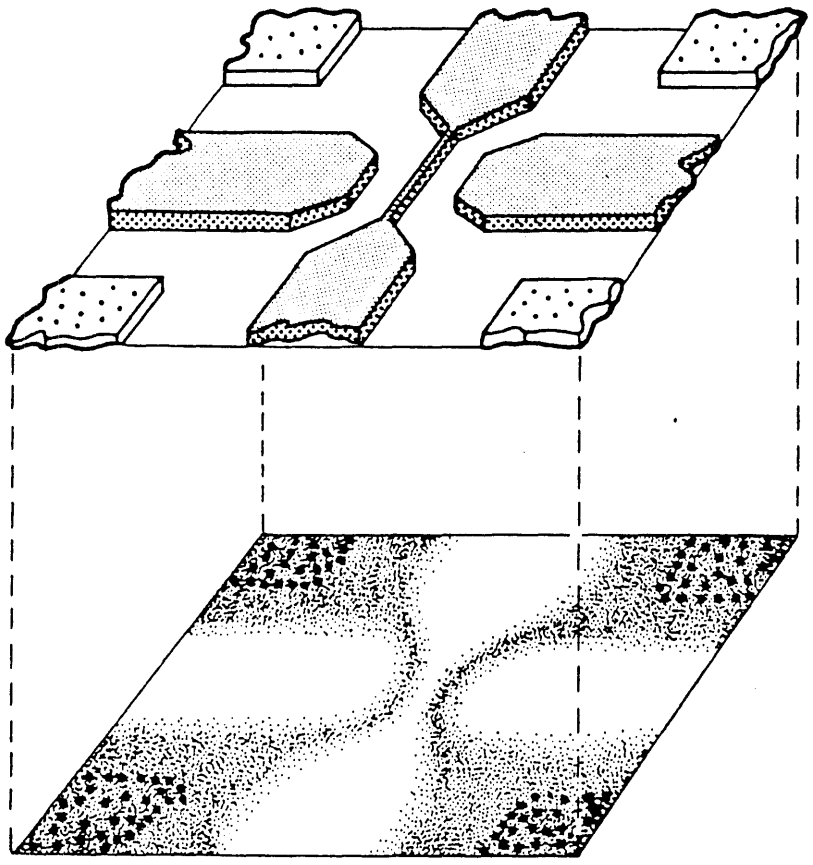


Figure 3-3: Conceptual picture of a 1-d channel formed using the split-gate scheme. A negative voltage applied to the gates first depletes the electrons directly underneath. Electrons in the channel are then depleted by the fringing capacitance. Figure courtesy of Cris Eugster.

Figure 3-4 shows scanning electron micrographs of the gate geometry for three possible schemes to realize the coupled quantum dot structure described in Sect. 1.2. The gates are patterned on a GaAs/AlGaAs heterostructure using electron beam lithography, as described in the next section. For all three schemes, the lithographic width of the quantum point contacts (QPCs), which couple the dot to the leads, was chosen to be about 3 times the depth of the 2DEG, in accordance with Eq. (3.1) above. Geometry I, the simplest design, consists of three depletion gates. The thin middle gate is used to create a tunnel barrier. This technique has been used by Eugster et al. [63, 66] to study the coupling between two electron waveguides. The linewidth of the middle gate is crucial, as the current through a thick barrier will be dominated by thermionic emission rather than by tunneling.

Geometry II utilizes a split-gate technique to implement the tunnel barrier. In this scheme, some degree of independent control is available between the left and right QPCs. Geometry III allows the maximum degree of independent control at the expense of complexity. The conductances of the left and right QPCs can be set independently and the dot size can be changed using the finger gate located inbetween.

For simplicity, the first sample contained only Geometry I and Geometry II. These three-gate and four-gate structures were compatible with an existing mask set and electron-beam lithography process. The first sample also included two structures to reproduce the basic Coulomb blockade effect in a single dot. One of these structures has Geometry I except that the middle gate was intentionally made very wide; the other structure is similar to that of Kouwenhoven et al. [48] pictured in Fig. 2-7. Results obtained on devices with Geometry I and Geometry II motivated us to process a later sample with Geometry III using a revised optical mask set; unfortunately the lithographic yield on that sample turned out to be too poor to find a working device.

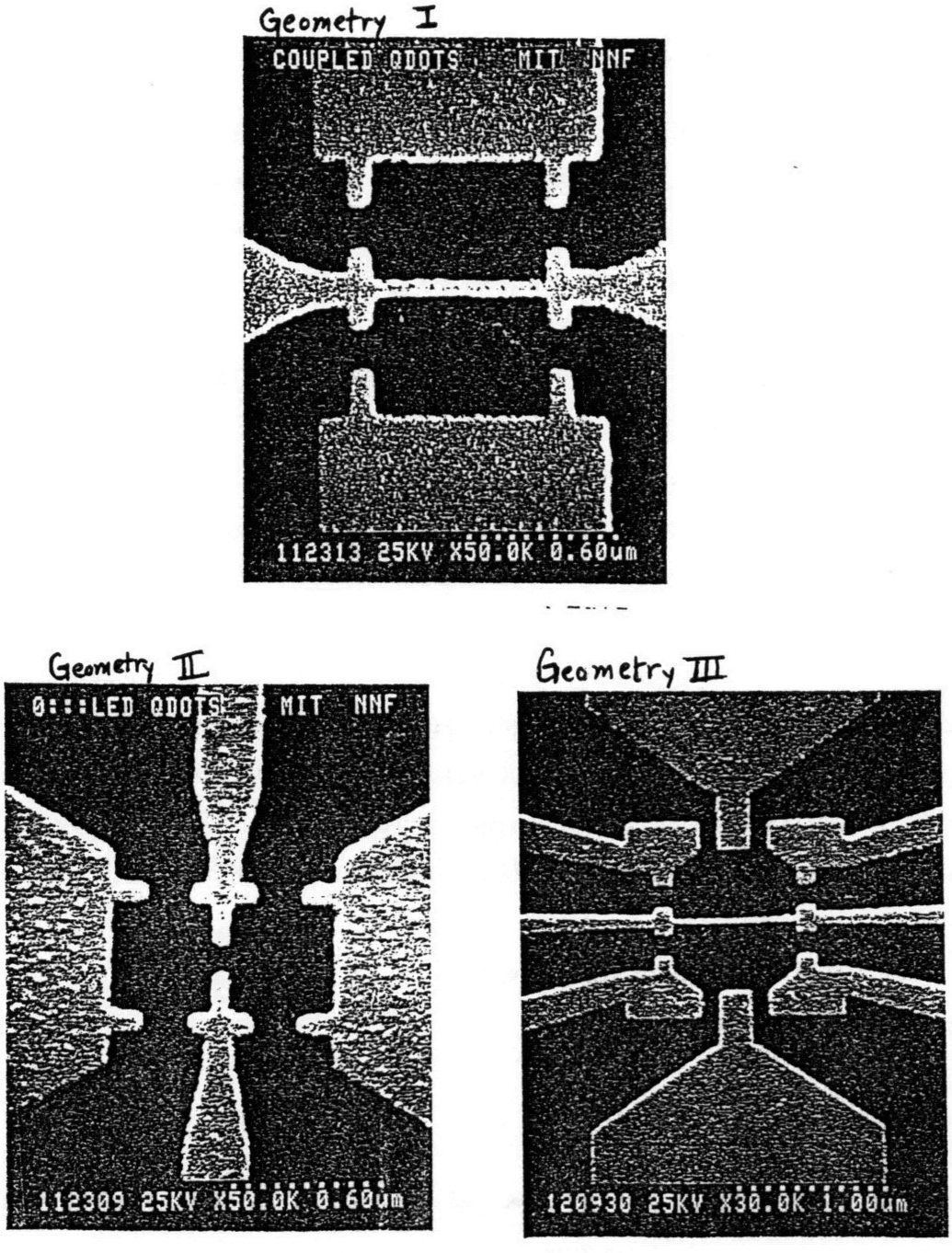


Figure 3-4: Scanning electron micrographs of three geometries used to realize a coupled dot structure. Geometry I and Geometry II are the simplest designs, whereas Geometry III offers the maximum degree of independent control.

3.3 Device Processing

Sample dies from two optical mask set designs are shown in Figs. 3-5 and 3-6. The mask set shown in Fig. 3-5 is the same as in Ref. [63]. It was used for the first sample because of its immediate compatibility with the electron beam lithography process at NNF-Cornell, described in Ref. [67]. The revised mask set shown in Fig. 3-6 was designed to accommodate the independent control devices discussed in the last section.[2]

Both mask designs were designed to facilitate the bonding of a number of devices to a 44-pin chip carrier. In the original mask design, one ohmic contact is shared among all devices to allow the accommodation of 7 devices with 7 connections per device. In this case, the ohmic contacts of all devices must be left floating during the measurement, which was found to increase significantly the noise level. The shared ohmic contact was eliminated in the revised design containing 4 devices with 11 connections per device.

The structure in the upper left corner is a van der Pauw structure for measuring the carrier density and mobility [68]. The structure in the lower left corner is a transmission line model which allows the contact resistance and sheet resistance to be extracted by finding the resistance of a fixed mesa width for various gaps between contacts. The structures in the upper and lower right corners are auxiliary devices with connections for 4 ohmic contacts and 4 gates.

Device fabrication consists of five main steps:

(1) mesa isolation of devices,
(2) ohmic contact formation,
(3) 1st level gate metallization,
(4) electron beam lithography of gates, and
(5) 2nd level gate metallization.

Steps (1),(2),(3), and (5) involve features of several microns and can be done using optical lithography at 400 nm. The gates required for electron confinement are of a sub 100 nm

[2]The design of the revised mask set and processing of the second sample were done in collaboration with David Carter.

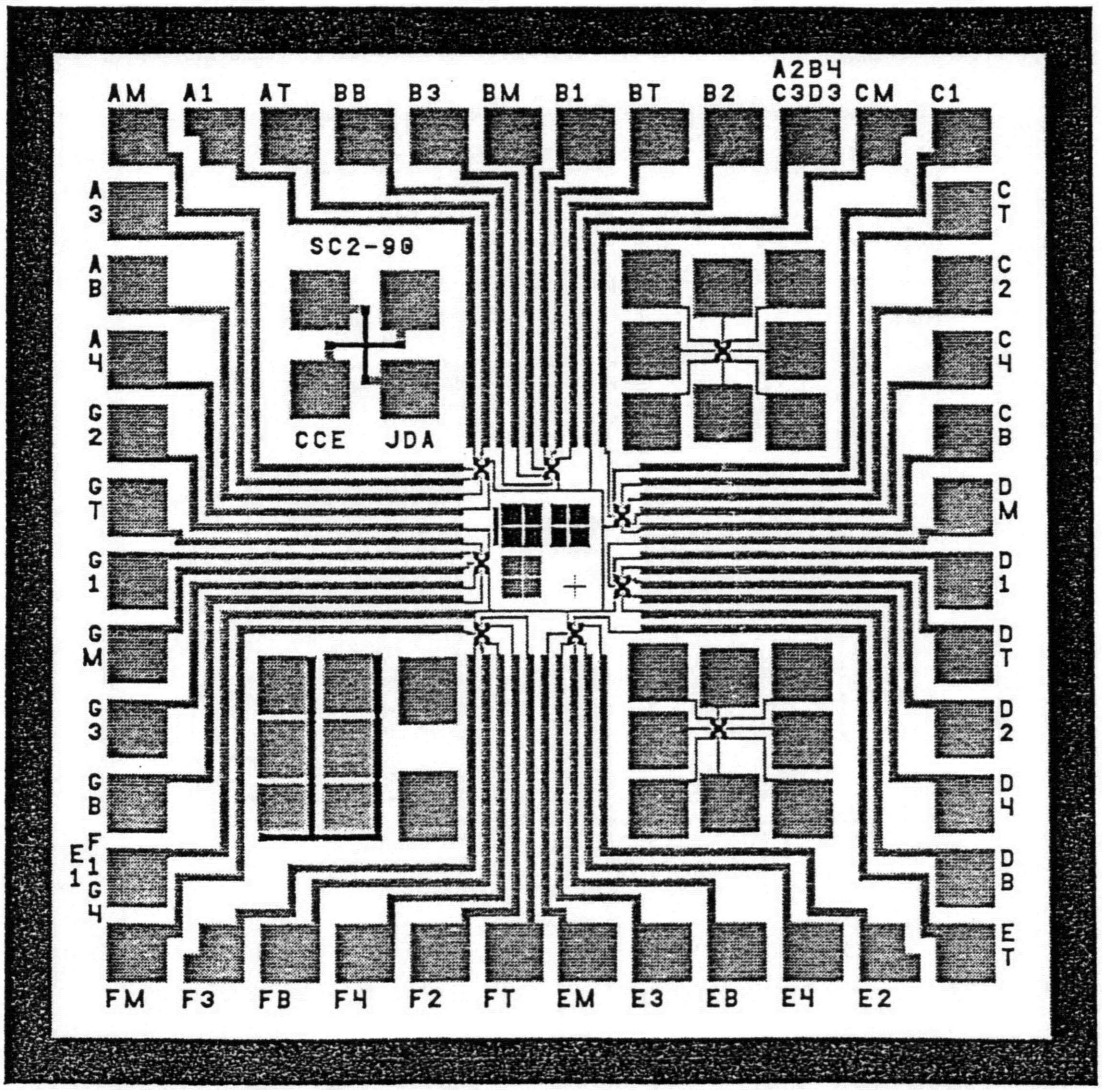


Figure 3-5: One die from the original mask set design. From Ref. [63]. The mask set is designed to facilitate the bonding of the 7 interior devices to a 44-pin chip carrier. Each of the 7 interior devices has connections for 4 ohmic contacts and 3 gates.

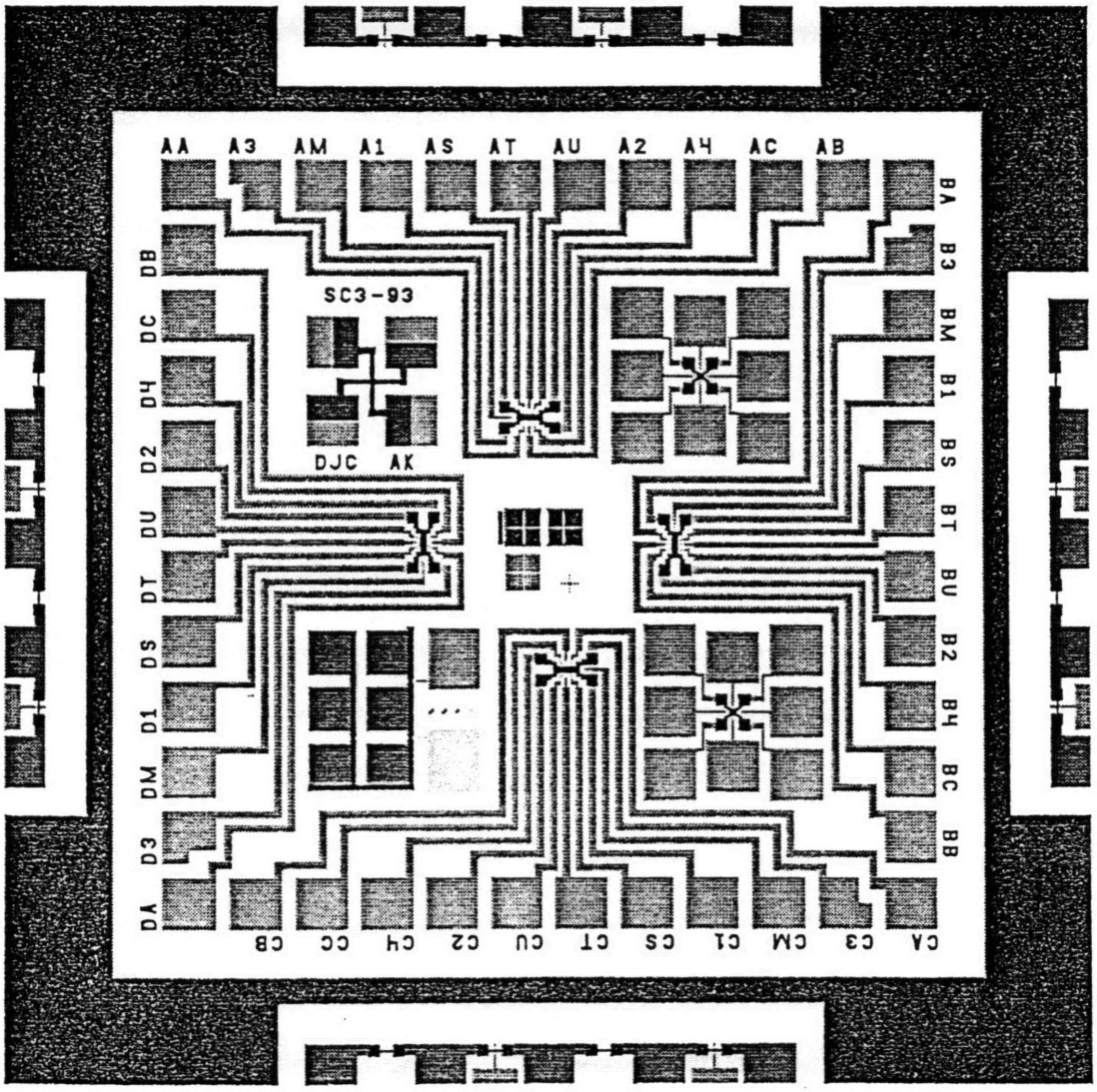


Figure 3-6: One die from the revised mask set design. The mask set is designed to facilitate the bonding of the 4 interior devices to a 44-pin chip carrier. Each of the 4 interior devices has connections for 4 ohmic contacts and 7 gates.

nature and must be done using either electron beam or x-ray lithography. The processing recipe was adapted from Ref. [63] with minor modification and is reprinted in Appendix A; the highlights are summarized below.

In Step (1), the 2DEG between devices is removed by immersing the sample in either an H_2SO_4 or H_2O_2/NH_4OH etchant. In Step (2), Ni/Au/Ge ohmic contacts are evaporated and sintered in a rapid thermal annealer at 425-430°C; the chemistry of ohmic contacts, not very well understood, is discussed in Ref. [69]. The contact resistance should be checked at room temperature using the transmission line model structure. If the contacts show any rectifying character or if the contact resistance exceeds 1 ohm-mm, the current-voltage characteristic should be checked at 77 K; test structures between adjacent dies of the revised mask set allow this to be done with little sacrifice to the sample.

In Step (3), Ti/Au or Cr/Au metal lines and bonding pads are evaporated to form the 1st level gates. The metal lines extend over the device mesas to make contact to the fine gates to be patterned in Step (4). Continuity between the fine gates with 15 nm metal and the optically patterned lines depends sensitively on the liftoff profile in Step (3). To check for continuity, the pinchoff characteristics of the devices should be measured at room temperature after Step (4); weak modulation is indicative of a continuity problem, in which case the 2nd level gate metallization in Step (5) should be performed.

In Step (4) the fine gates are patterned using a JEOL JBX 5DII electron beam lithography system at the National Nanofabrication Facility [67]. A bilayer pmma scheme, in which a less sensitive resist is spun on top of a more sensitive resist, is used to enhance the undercut profile. The critical middle gate is exposed by a single pass electron beam line at an optimally-calibrated dose. Exposure is followed by development in MIBK:IPA (1:3) and thermal evaporation of 15 nm Au/Pd (to minimize the grain size). The scanning electron micrograph in Fig. 3-7 shows that a linewidth of 30-40 nm is attainable.

Before each of the processing steps, a rigorous solvent clean consisting of a trichloroethane boil followed by ultrasound immersion in acetone and methanol is performed. Particular

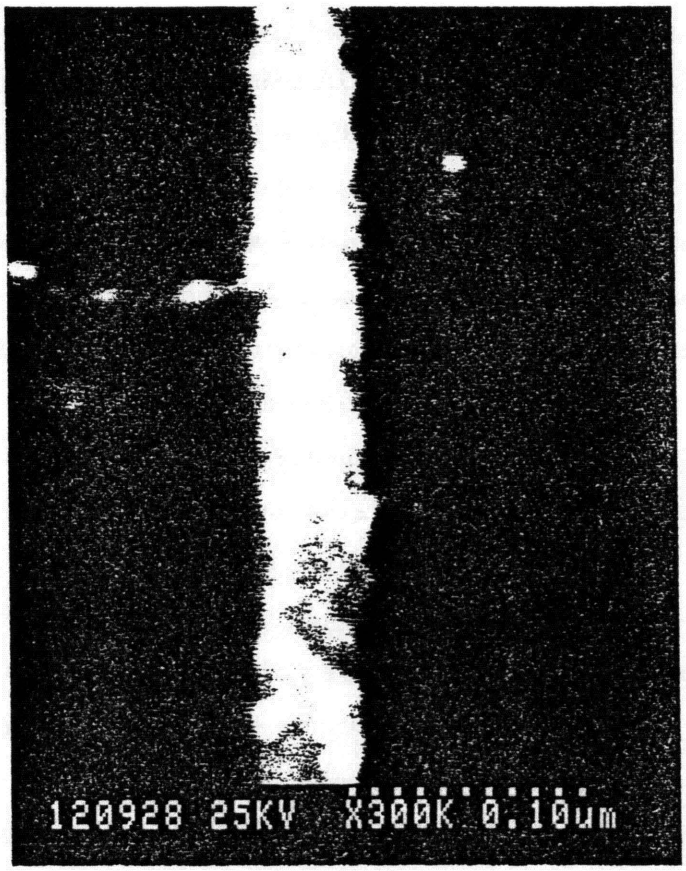


Figure 3-7: High resolution scanning electron micrograph of middle gate, showing linewidth of 40 nm.

care must be taken to insure that the surfaces are clean before the evaporation of any of the contact metals; the present process uses uv-ozone to remove organics and residual resist followed by immersion in Semico^{TM} to strip oxides. For all metallization steps after annealing, thermal evaporation is preferable to electron beam evaporation to reduce the possibility of mobility degradation [67]. To facilitate liftoff, Steps (2), (3), and (5) utilize a chlorobenzene soak procedure to create an overhang profile.

Before ending this section, it should be noted that x-ray lithography is a possible alternative to electron beam lithography. The x-ray process offers the advantage that much more metal can be evaporated using a thicker pmma layer because the x-rays do not scatter laterally during exposure. In principle, Steps (3) and (4) above could be combined and Step (5) would be eliminated in the x-ray process. However, the x-ray process is complicated by the initial requirement to fabricate and replicate an x-ray mask. Although some attempts were made to use x-ray lithography to pattern the fine features, a minimum linewidth comparable to that in the electron beam process was not achieved during these trials.

3.4 Cryogenic System

The Coulomb blockade oscillations wash out as k_BT approaches $e^2/2C$. To date, single-electron effects in semiconductor nanostructures have been observed only in the milli-Kelvin regime, as limited by the device dimensions which set the capacitance scale. To study Coulomb charging effects, our group acquired an Oxford Instruments Heliox insert with base temperature of 300 mK. A qualitative description of its operation is given below; the operating procedure is given in Appendix A.

A schematic of the Heliox is shown in Fig. 3-8, with the inner vacuum can (IVC) enlarged to show detail. The insert achieves its base temperature by reducing the vapor pressure of the ^{3}He using a sorb pump, which adsorbs gas when cooled below a certain temperature (around 12 K). After evacuating the IVC and pre-cooling in liquid nitrogen, the insert is

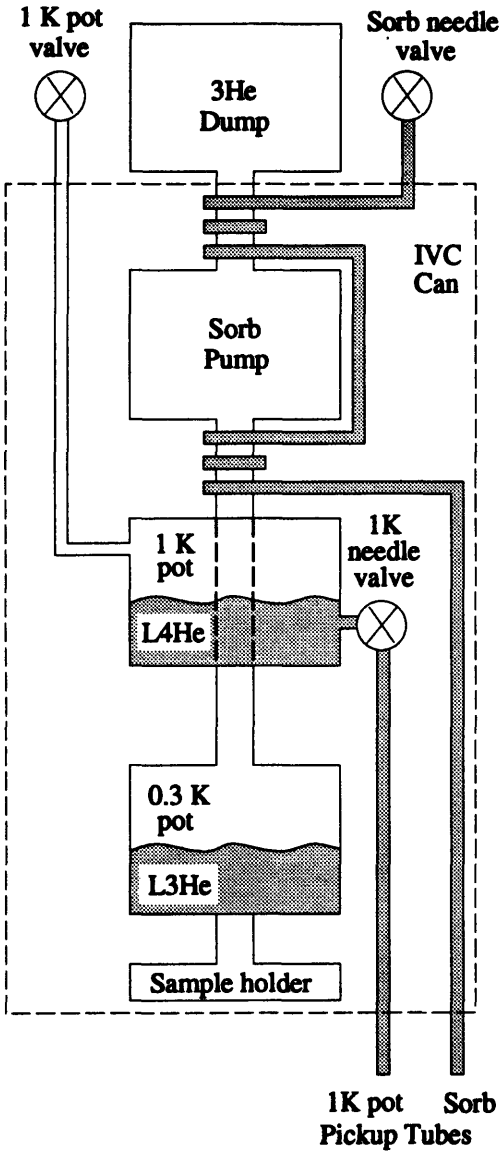


Figure 3-8: Schematic of Oxford Instruments Heliox 300 mK insert with the interior of the inner vacuum can (IVC) enlarged to show detail. The insert is lowered into a storage vessel of liquid ^{4}He. Schematic courtesy of Gee Rittenhouse.

lowered into a storage vessel of liquid ^{4}He. Operation begins by admitting some liquid from the ^{4}He bath into the 1 K pot through the 1 K pot pickup tube. The temperature at the 1 K pot is brought to 1.2 K by reducing the vapor pressure of the ^{4}He with an external rotary pump. The sorb pump is then heated; the ^{3}He (normal boiling point 3.19 K) which is released passes through the 1 K pot and condenses in the 0.3 K pot. The 1 K pot is kept cold during the heating of the sorb by flowing ^{4}He through the sorb pickup tube. When all the ^{3}He is condensed, the sorb heater is turned off and the temperature of the 0.3 K pot drops rapidly to 300 mK as the sorb pumps on the ^{3}He. The hold time, limited by the ^{3}He charge, is about 30 hours for our unit. No noticeable effect on device characteristics was seen from thermally cycling to 5 K, or to 77 K when the ^{4}He vessel is refilled.

The insert is wired with 48 manganin leads which are thermally sunk to the 0.3 K pot. The sample is affixed with silver paint onto a 44-pin chip carrier and bonded with an Al wedge bonder. The back of the chip carrier is thermally sunk to the copper sample holder by thermally conducting grease.

Some later measurements were taken at Saclay in a Tres Basse Temperatures dilution refrigerator with superconducting magnet. A dilution refrigerator is able to achieve temperatures < 50 mK by creating a very dilute phase of ^{3}He dissolved in ^{4}He [70]. Practical problems limited the base temperature of the Saclay fridge to 70 mK during these measurements.

3.5 Measurement Circuit

Figure 3-9 shows a schematic of the gate geometry and resulting electron distribution when a device with Geometry I in Fig. 3-4 is biased to realize a three-terminal quantum dot. The bottom gate is grounded, leaving a semi-infinite 2DEG. The voltage V_M is fixed so that the middle gate acts as a tunnel barrier, as evidenced by an exponential tail in the pinchoff characteristics. A small ac excitation voltage V_{ds} is applied at contact 3. The

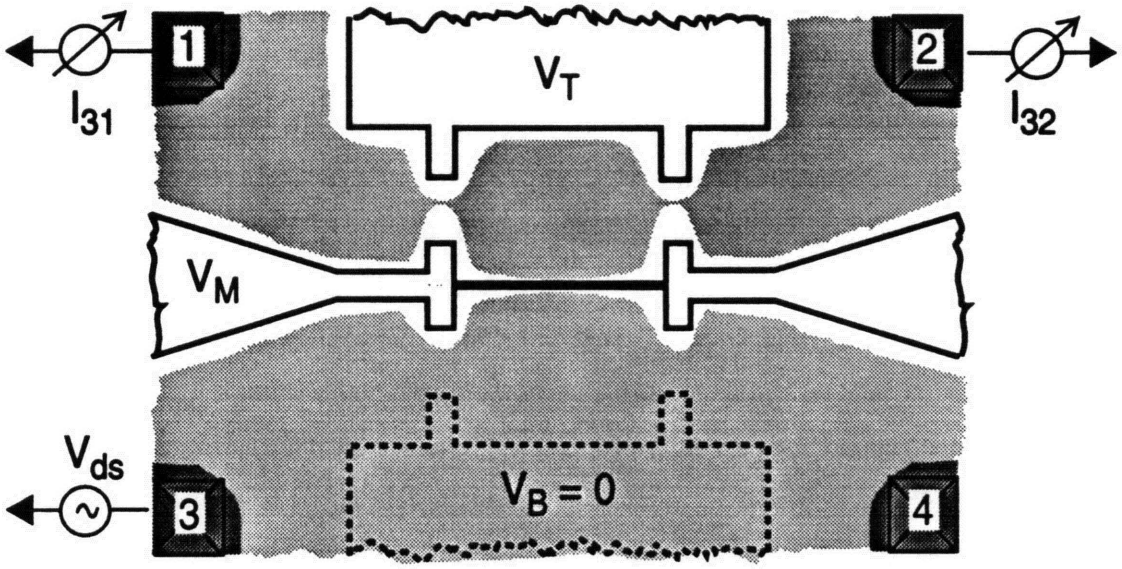


Figure 3-9: Schematic of the gate geometry of three-terminal quantum dot. Application of negative gate voltages V_T and V_M depletes the 2DEG underneath, leaving behind a dot coupled to three leads. Four ohmic contacts allow access to the 2DEG. In our measurements a voltage V_{ds} is applied at contact 3 and currents I_{31} and I_{32} are measured simultaneously as V_T is swept.

currents at the output leads $I_{31} = G_{31}V_{ds}$ and $I_{32} = G_{32}V_{ds}$ are measured simultaneously using two low-noise current amplifiers as the top gate voltage V_T is swept. Each current amplifier outputs a voltage which is measured using a lock-in amplifier.

Figure 3-10 shows in detail the circuits used for the measurement.[3] The OP97 used in the current preamplifiers was chosen because of its low *current* noise which is important for high impedance measurements. The low bandwidth also helps reduce current noise but limits the maximum frequency at which lock-in measurements can be done. The use of active filters with differential input $(V_2 - V_1)$ to obtain V_G and V_{ds} isolates the grounds of the original voltage sources from the single-point ground of the measurement circuit. The OP27 was chosen for these circuits because of its low *voltage* noise. For the drive voltage V_{ds}, the oscillator output of one of the lock-ins is used as the differential input of the active

[3]The circuits in this section were suggested by Nathan Belk.

filter, which attenuates its input by 10000. For the dc gate voltages, a much lower cutoff frequency is used. The differential input is either a simple battery for a fixed gate or the Hewlett-Packard 3325 function generator for the gate being swept.

The current preamplifiers are isolated from the device by two 10 μF blocking capacitors. The blocking capacitors serve to shield the device from the input offset voltage of the operational amplifiers (typically 10-30 μV for the OP97). The input offset voltage will be dropped entirely across the capacitor as long as the resistance of the path to ground through the 10 ohm resistor of the drive voltage source is negligible compared to the leakage resistance of the capacitor. This condition is always satisfied for the high-quality capacitors used so that the two source reservoirs are virtual grounds of the circuit.

Several checks were done to verify that the offset voltages were properly compensated:

- When V_T is swept negative, the currents I_{31} and I_{32} are generally unequal due to intrinsic, unintentional asymmetries in the device. However, essentially the same traces for I_{31} and I_{32} were obtained if the current preamplifiers are interchanged.

- Any ac currents arising from dc offset voltages should be highly nonlinear in the drive voltage V_{ds}. However, the currents I_{31} and I_{32}, which are measured by an ac lock-in technique, were always found to be linear in V_{ds}.

All the circuits are assembled on a circuit board tied to a single ground point. The operation amplifiers are powered with simple batteries to avoid a ground loop. The entire assembly is placed in a metal box and connected to the cryostat wiring by two shielded twisted-pair cables. The voltage outputs of the two current preamplifiers are read by two PAR 5210 model lock-in amplifiers. Data is recorded by a computer linked to the lock-ins and HP 3325 by IEEE-488 buses.[4]

[4]The data acquisition system was set up by Martin Burkhardt based on software written by Gee Rittenhouse.

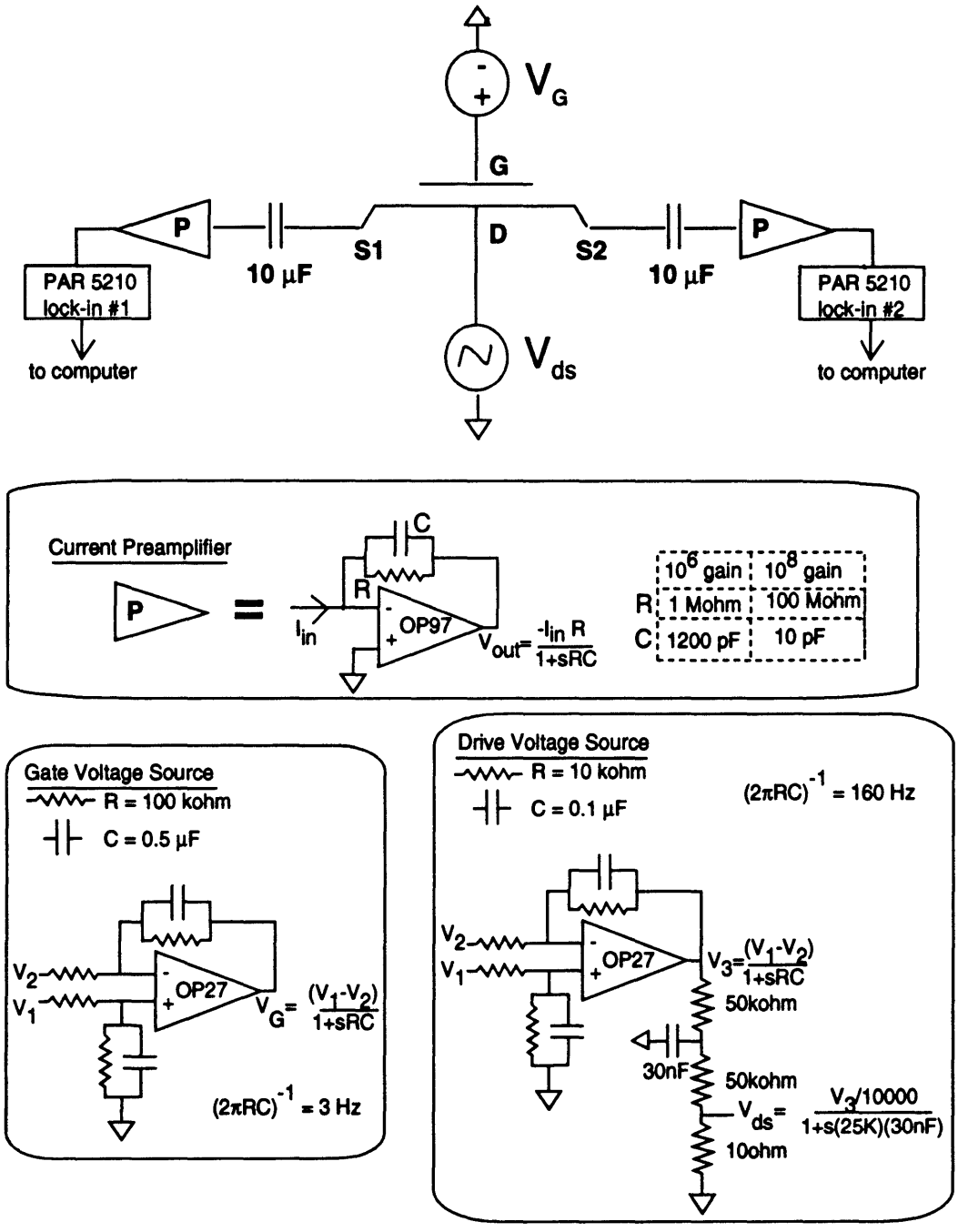


Figure 3-10: Circuits used to measure a three-terminal device.

Chapter 4

Experimental Results

4.1 Early Experiments

This section summarizes early experiments whose objective was to reproduce the basic Coulomb blockade oscillations in a single quantum dot. For this reason the first sample processed on MBE18 contained two types of test structures:

- A simple split-gate structure with Geometry I in Fig. 3-4 except that the middle gate was intentionally made impenetrably wide, creating effectively two independent single dots.
- A structure allowing independent control of the tunnel barriers, modeled after the structure of Kouwenhoven et al. [48] pictured in Fig. 2-7(a).

Measurement of the basic oscillations proved to be elusive until the filter circuitry described in Sect. 3.5 was implemented.

Figure 4-1 shows the Coulomb blockade oscillations first observed in the simple split-gate structure depicted in the inset. Periodic oscillations are observable as the top and bottom gates are tied together and swept. These results are consistent with the Meirav's finding

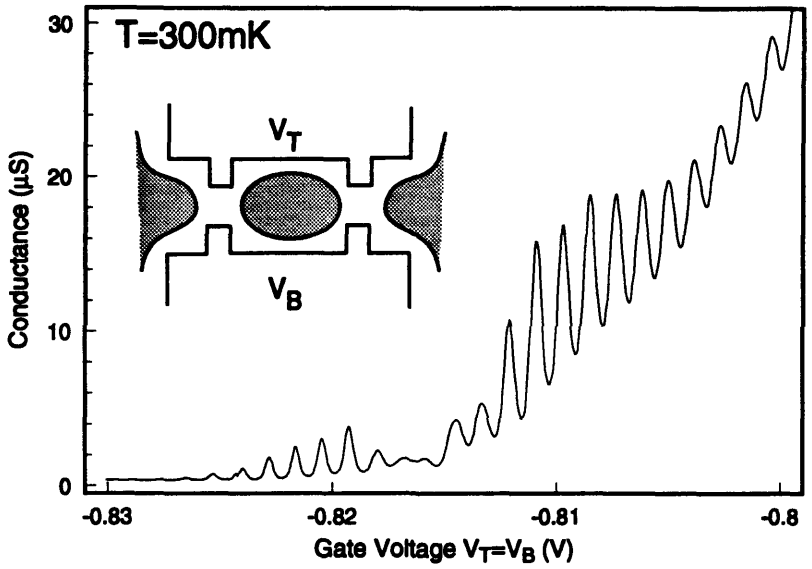


Figure 4-1: Periodic conductance oscillations observed in simple split-gate structure shown schematically in inset.

that clear oscillations could be observed by sweeping the upper depletion gates at a fixed back gate voltage [46]. In this type of structure, there is always some intrinsic asymmetry between the left and right tunnel barriers, due to impurities and lithographic imperfection. Coulomb blockade effects may not be observable in such a structure if the tunnel barriers have very dissimilar conductances. This is probably the reason why other devices with similar geometry on the same chip did not show Coulomb blockade oscillations.

By contrast, the structure depicted in the inset of Fig. 4-2 has a much higher probability of showing Coulomb blockade effects because the two tunnel barriers can be tuned independently. The finger gate (V_F) can then be swept with relatively little influence on the conductances through the QPCs.[1] The first step is to characterize each QPC independently. As shown in Fig. 4-3, the pinchoff characteristics are marked by a clear 2-d to 1-d transition. In the 1-d regime, both QPCs show well-resolved conductance steps quantized

[1]It is important to remember that the oscillations are observable as the voltage on any gate is varied.

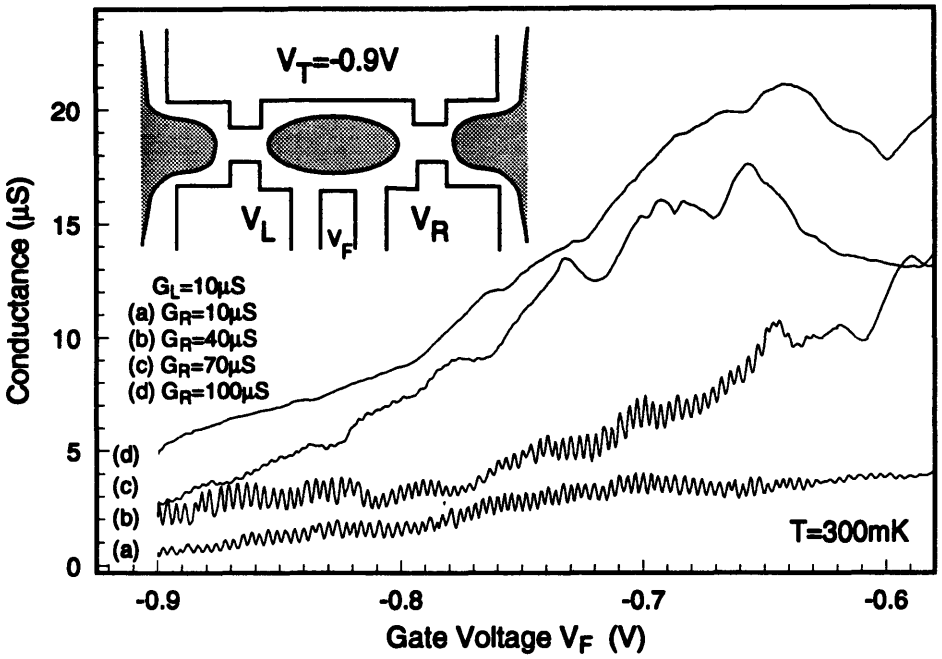


Figure 4-2: Conductance as a function of finger gate voltage V_F for four different values of G_R, the conductance through the right QPC. The voltage V_L is set so that the conductance through the left QPC is $G_L = 10\ \mu S$.

in multiples of

$$\sigma_Q = \frac{2e^2}{h} = 77.48\ \mu S = (12906\ \Omega)^{-1}. \tag{4.1}$$

Figure 4-2 shows the conductance obtained for various values of the right QPC conductance G_R when the left QPC is biased well into the lowest subband ($G_L = 10\ \mu S$). The Coulomb blockade oscillations are seen to wash out roughly when $G_R > 70\ \mu S$. This result is consistent with the findings of Kouwenhoven et al. [48]; however, Pasquier et al. [49] found that small Coulomb blockade oscillations persist up to $G_L + G_R = 3\sigma_Q$.

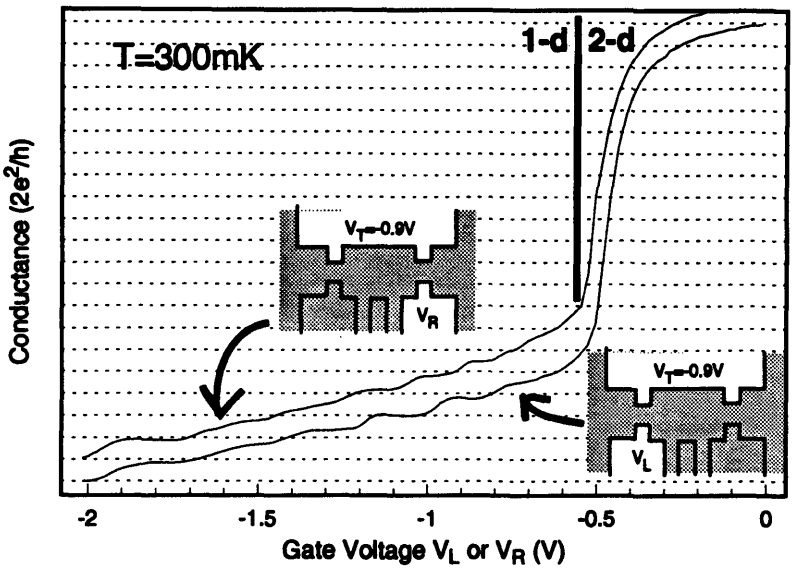


Figure 4-3: Quantized conductance steps for left (lower curve) and right (upper curve) QPCs. The curves are displaced by $2e^2/h$ for clarity. The top gate is fixed at -0.9 V while V_L or V_R is swept.

4.2 Three-Terminal Quantum Dot: MIT Experiments

This section and the next one deal with three devices having the simple split-gate geometry shown schematically in Fig. 4-4. The devices were all fabricated together on MBE18. Device A and Device C have $L = 0.6$ μm while Device B has $L = 1$ μm. All measurements in this section were taken at 300 mK in zero magnetic field, using the Oxford Instruments Heliox insert described in Sect. 3.4.

The first step is to characterize the middle gate. As shown in Fig. 4-5 for Device A, the pinchoff characteristic is marked by two threshold voltages, V_{p1} and V_{p2}. The first threshold at $V_M = V_{p1}$ occurs when the 2DEG underneath the wide portion of the middle gate is turned off. Because of the short-channel effect, a more negative threshold at $V_M = V_{p2}$ is required to turn off the 2DEG underneath the narrow portion of the middle gate. Physically, a more negative voltage is required because the electric field of a 1-d line of charge decays more slowly than that of a 2-d plane of charge. Close to pinchoff, the characteristic exhibits

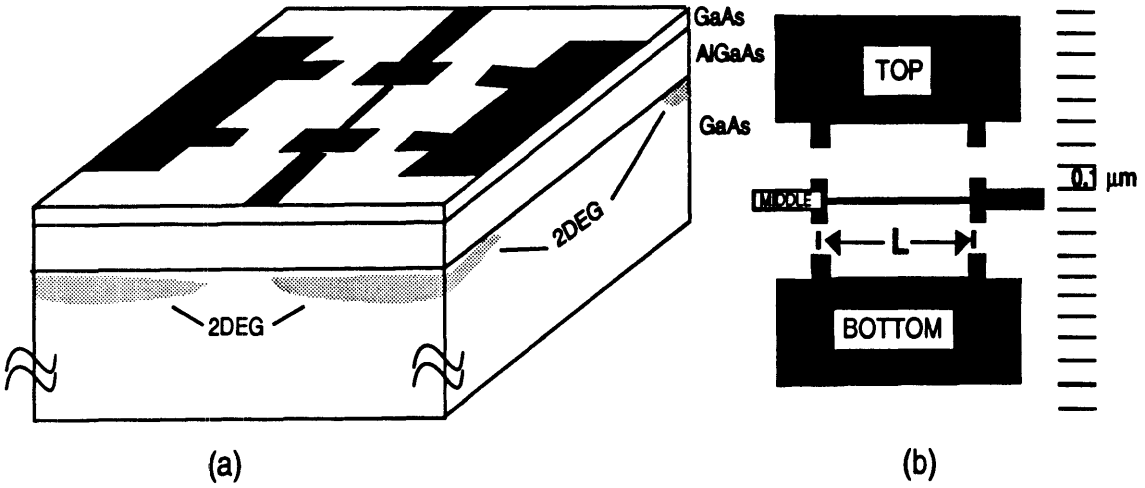


Figure 4-4: (a) Split-gate geometry over heterostructure MBE18 used to realize the devices studied. (b) Device A and Device C have $L = 0.6$ μm while Device B has $L = 1$ μm.

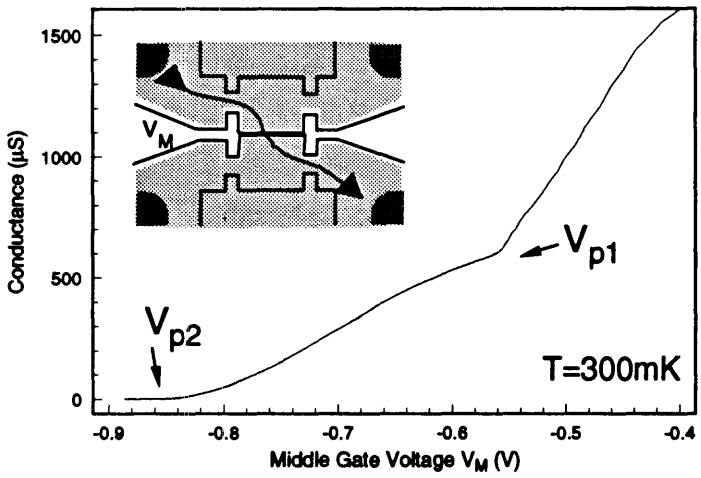


Figure 4-5: Pinchoff characteristic of middle gate, found by measuring the cross-current as shown with top and bottom gates grounded. The first threshold at $V_M = V_{p1}$ occurs when the 2DEG underneath the wide portion of the middle gate is turned off. Because of the short-channel effect, a more negative threshold at $V_M = V_{p2}$ is required to turn off the 2DEG underneath the narrow portion of the middle gate. Close to pinchoff, the characteristic exhibits an exponential tail.

an exponential tail. The middle gate should be biased in this regime if a tunnel barrier is to be formed.

The procedure for realizing a three-terminal quantum dot is as follows. First, the conductance through the tunnel barrier formed by the middle gate – henceforth called the injector barrier – is set to 20 μS. The conductance is measured by applying a small excitation voltage on one side of the injector barrier and measuring simultaneously the currents at the two ohmic contacts on the opposite side.[2] The gate (top or bottom) on the opposite side is then swept negative, as shown in Fig. 3-9. Due to intrinsic asymmetries between the left and right tunnel barriers, the currents become unequal.

Figure 4-6 shows the two conductances measured for Device A.[3] As originally anticipated, both conductances showed oscillations periodic in the gate voltage V_B. However, two striking features stood out:

- The peaks of one current do not occur at the same gate voltage as peaks of the other current. In fact, the peaks of one current sometimes coincide with the valleys of the other.
- One of the currents actually changes sign at some of the deep valleys.

Neither of these observations was consistent with conventional single-electron tunneling (SET) theory, according to which the oscillations must align with each other in gate voltage.

At this point there was strong motivation to see whether these effects could be repeated in other devices. We now discuss results of Device B, in which the effects seen in Device A were reconfirmed. Figure 4-7(a) shows the conductances G_{31} and G_{32}, measured concurrently, as the top gate voltage V_T is swept. [4] The middle gate voltage $V_M = -0.7$ V is kept fixed in the tunnel regime, as determined by an exponential tail in its pinchoff characteristic. Due to some intrinsic, unintentional asymmetry in our structure, the QPC 2 has a turn-on

[2] As in the discussion of Sect. 3.5, the conductance $G_{\ell m}$ refers to the current measured at lead m divided by the voltage $V_{d\ell}$ applied at lead ℓ.

[3] Weak oscillations were also visible in the top dot for the analogous biasing configuration.

[4] For an analogous biasing configuration, the bottom dot also showed the effect, although weakly.

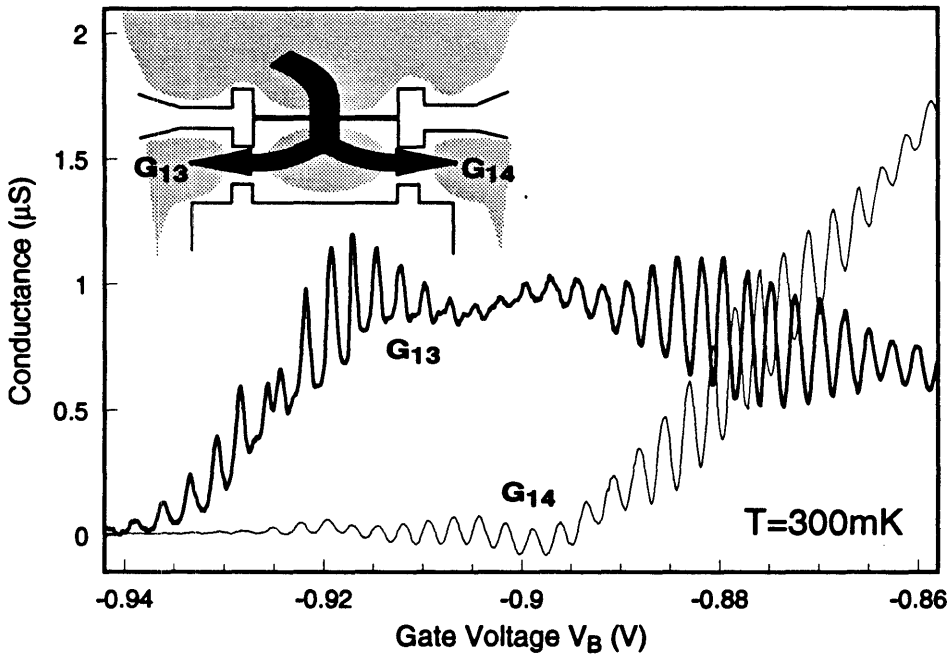


Figure 4-6: Conductances G_{13} and G_{14} for Device A as bottom gate voltage V_B is swept. The middle gate is fixed at voltage $V_M = -0.84$ V giving conductance $G_M = 20$ μS.

voltage about 40 mV higher than QPC 1.

Figure 4-7(b) shows in detail the conductance oscillations in G_{31} and G_{32} as QPC 2 just opens up. Also shown is the total conductance $G_{31} + G_{32}$ through the dot. The perfect alignment of the oscillations in G_{31} and G_{32} for $V_T \leq -0.555$ V can be understood simply from standard SET theory.[5] At a conductance maximum, an electron which has tunneled into the dot from the input lead has some probability of being transmitted through either one of the two output leads. As discussed in Chapter 5, the peaks in this regime are well fit by a thermally broadened lineshape, so that the finite off-resonance conductance is due primarily to thermally-excited carriers.

As the gate voltage is increased in Fig. 4-7(c), the peak-to-valley ratio in the total dot conductance $G_{31} + G_{32}$ drops markedly. However, instead of broadening accordingly, resonances in the individual conductances G_{31} and G_{32} evolve from being perfectly *correlated* to being perfectly *anti-correlated* with each other. Another striking feature of the anti-correlated regime is that the current at lead 2 actually changes sign at some of the deep valleys. The total conductance through the dot $G_{31} + G_{32}$ is, however, always positive.

The opening of QPC 2 is marked by a rapid rise in the peak amplitudes of G_{32}. During this regime, it is likely that the conductance of QPC 2 is the limiting conductance for G_{32}. As the top gate is swept more positive, the rapid rise in the envelope of G_{32} stops and it is likely that the limiting conductance for both G_{31} and G_{32} is that of the injector barrier. Here the peaks in G_{32} push ahead of those in G_{31} before giving way to a transition region without very clear resonances. Also notable is the beating pattern observed in G_{31} and G_{32}.

4.3 Three-Terminal Quantum Dot: Saclay Experiments

The measurements in this section were taken in collaboration with Philippe Debray at

[5]The oscillations are well-aligned even near the anomaly (probably caused by telegraph noise) at $V_T = -0.556$ V.

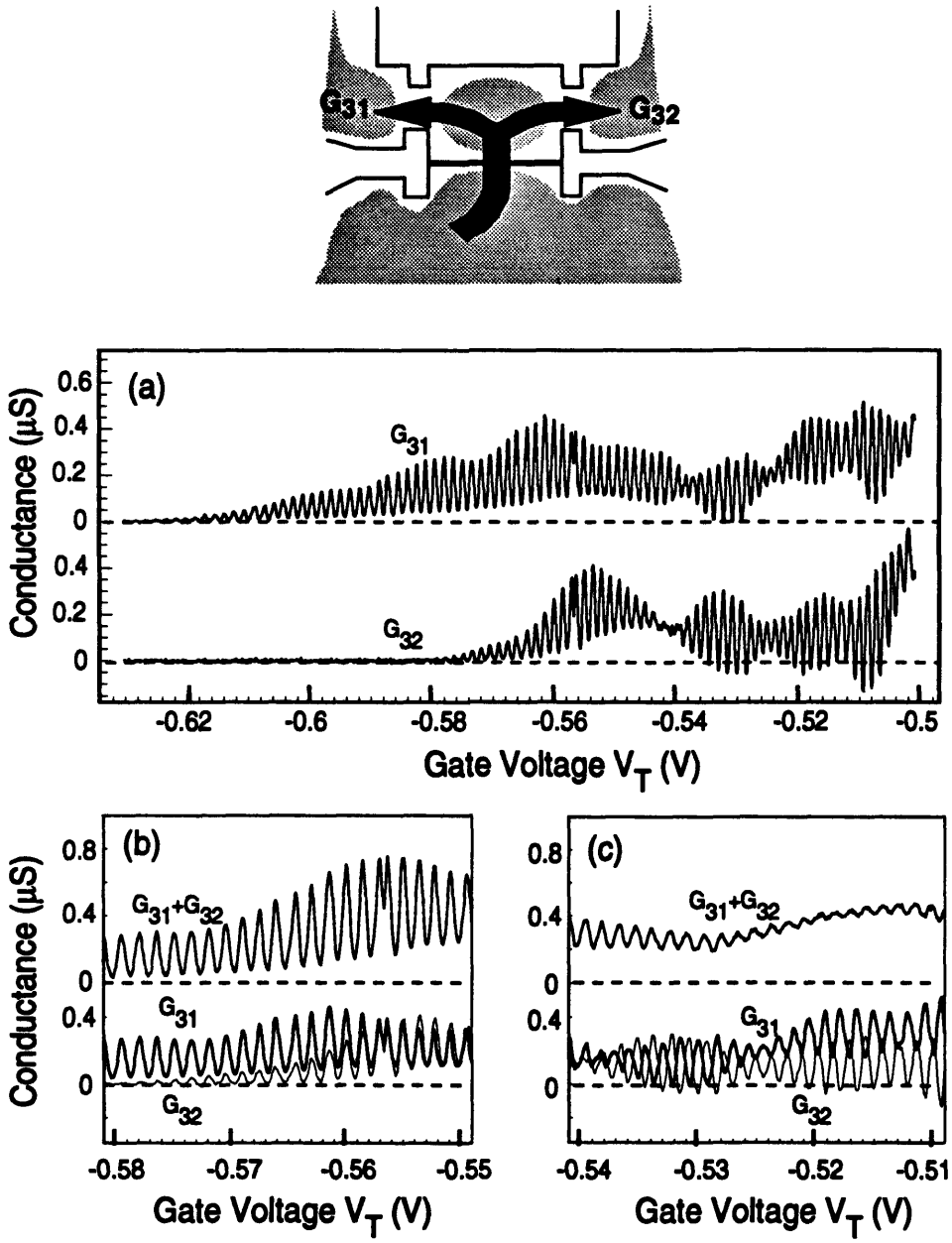


Figure 4-7: (a) Conductances G_{31} and G_{32} as the top gate voltage V_T is swept. The middle gate voltage $V_M = -0.7$ V is fixed in the tunnel regime. (b) Detailed plot of the first several resonances in (a), indicating that they are in-phase with each other. (c) Detailed plot of resonances at higher V_T, indicating that they have evolved from almost perfect *correlation* to almost perfect *anti-correlation* in gate voltage.

Centre d'Etudes de Saclay in France. The measurements were done in a dilution refrigerator at 70 mK. A superconducting magnet was also available.

Figure 4-8 shows the conductances G_{31} and G_{32} for Device C as the top gate voltage V_T is swept. In this case, clear regions of anti-correlated oscillations are seen (marked by an arrow) but they are separated by regions of varying phase between the oscillations. This intermixing was not seen in previous samples.

For comparison, the conductance G_{12} (scaled down by a factor 10) is also shown on the same graph. The conductance G_{12} was obtained in a different sweep by measuring the current at lead 2 in response to a voltage excitation at lead 1 while lead 3 and lead 4 were floating. Both conductances G_{12} and G_{31} pinch off at at the same value of V_T, confirming $V_T = -0.86$ V as the threshold voltage for QPC 1. The conductance G_{12} rises rapidly, reflecting the opening of QPC 1. The conductance G_{31} also rises rapidly at first, but becomes limited by the conductance of the injector barrier. After QPC 1 has opened up, it is reasonable to assume that both G_{31} and G_{32} are limited by the conductance of the injector barrier, since QPC 2 opens sooner than QPC 1. The background conductances of G_{31} and G_{32} follow each other, reflecting small changes in the transmission of the injector barrier as V_T is swept. Near the first regime of anti-correlated oscillations ($V_T \sim -0.83$ V), $G_{12} \sim 12$ μS. Since this is the minimum conductance of QPC 1 and QPC 2, it follows that both QPC 1 and QPC 2 must have conductances that are a substantial fraction of e^2/h. By contrast, the conductance of the injector barrier, which limits G_{31} and G_{32}, is less than 1 μS. The high conductances of QPC 1 and QPC 2 render the dot into the weakly blockaded regime, as evidenced also by the poor peak-to-valley ratio of G_{12}.

Figures 4-9 and 4-10 show the evolution of the conductances G_{31} and G_{32} as a small perpendicular magnetic field is applied. Taking the area of the dot to be very roughly $\pi(0.2\ \mu\text{m})(0.125\ \mu\text{m})=8\times10^{-10}$ cm^2, we estimate the magnetic field required to add a flux quantum $\Phi_0 = h/e = 4.14 \times 10^{-11}$T cm^2 to be 0.05 T. Little change is observed in the conductances until $B = 0.07$ T, where G_{31} is suppressed near the region of anti-correlated

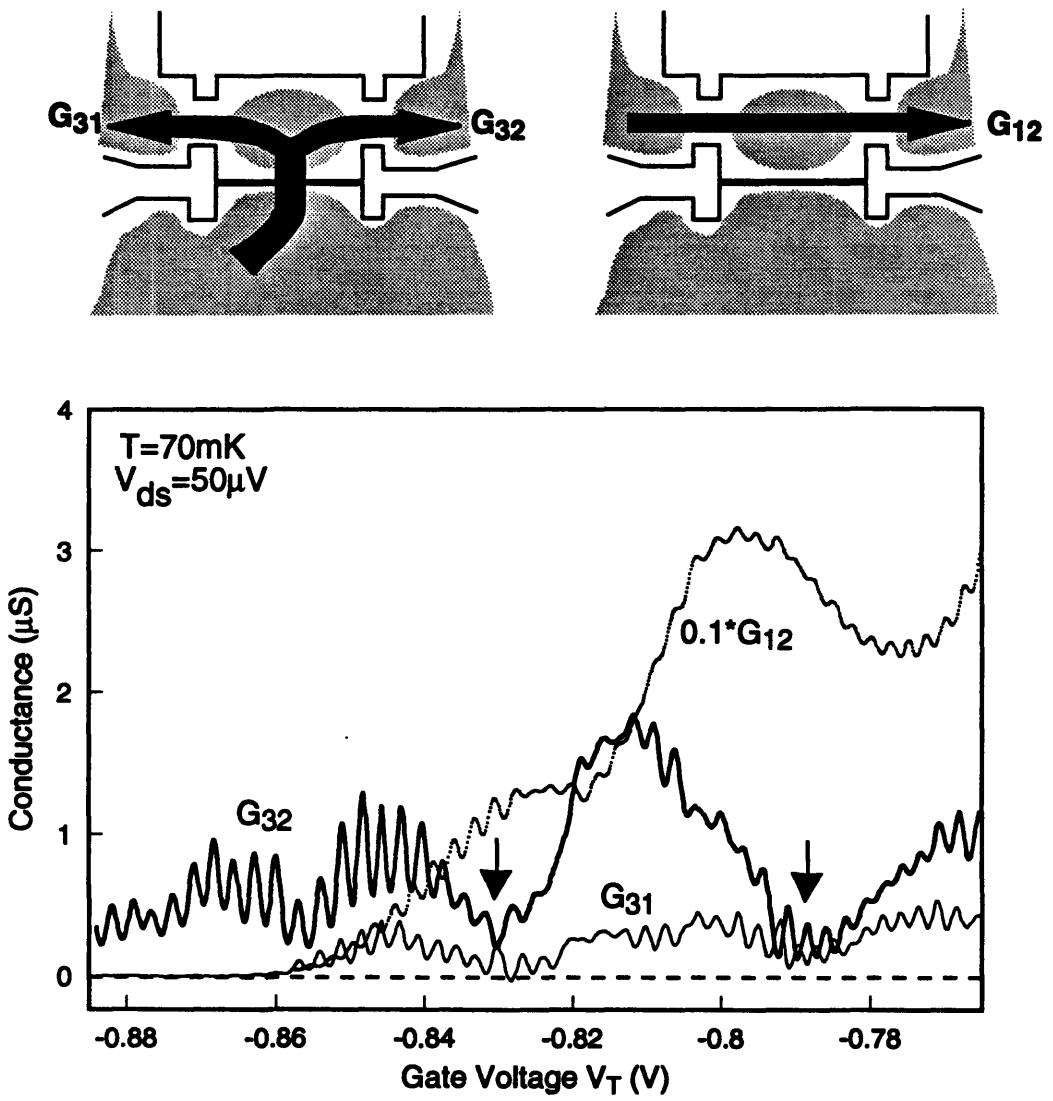


Figure 4-8: Conductances G_{31} and G_{32} for Device C. Regions of anti-correlated oscillations are marked by an arrow. For comparison, the conductance G_{12} (scaled down by a factor 10) is also plotted.

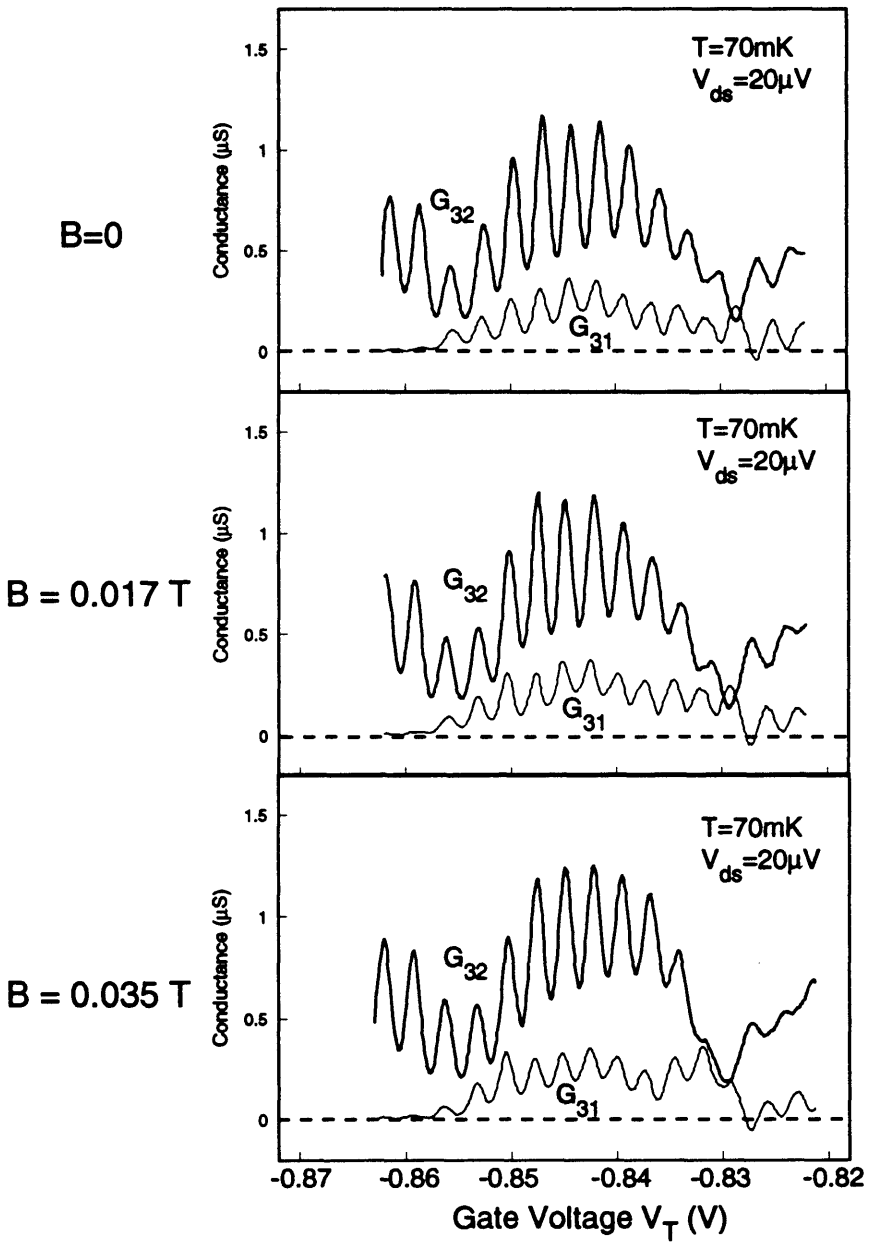


Figure 4-9: Conductances G_{31} and G_{32} for Device C in magnetic fields of $B = 0$, $B = 0.017$ T, and $B = 0.035$ T.

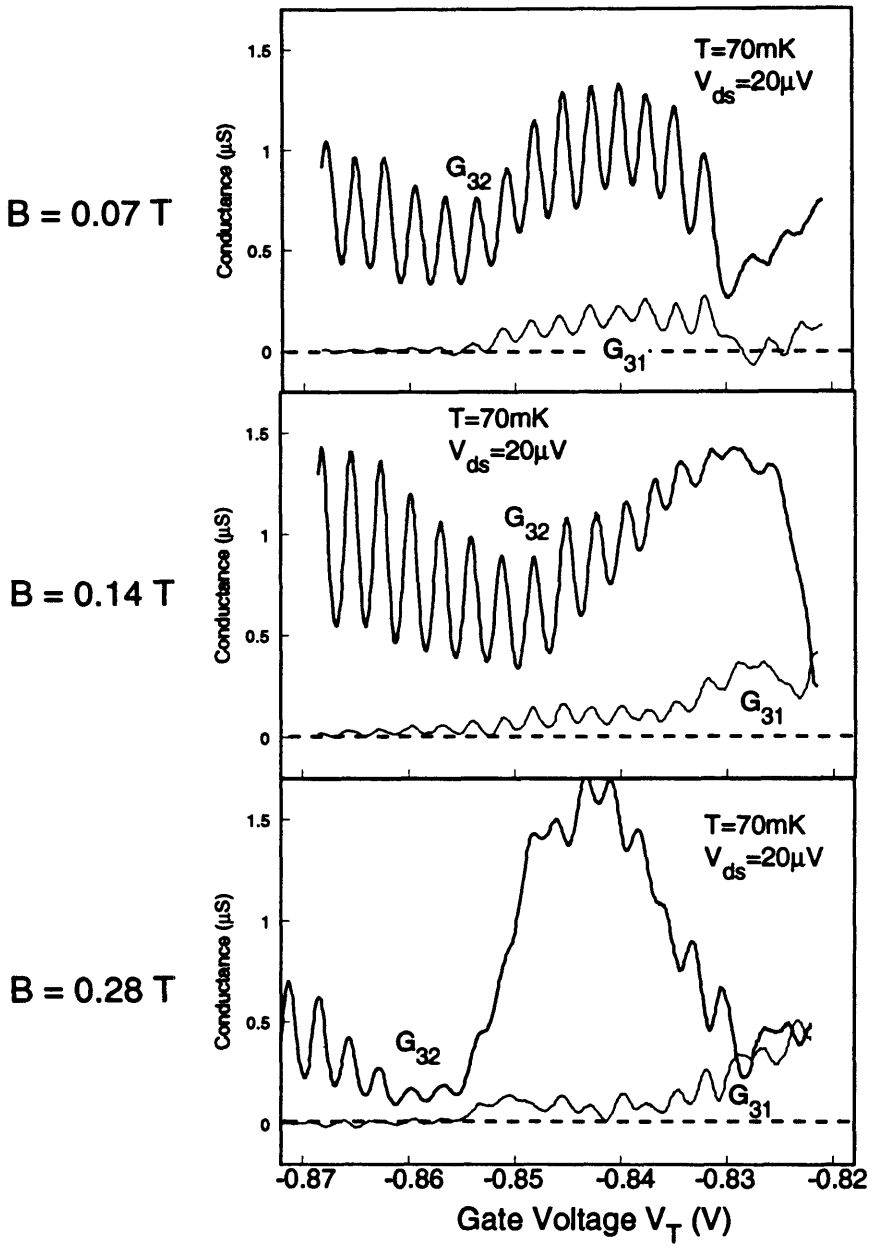


Figure 4-10: Conductances G_{31} and G_{32} for Device C in magnetic fields of $B = 0.07$, $B = 0.14$ T, and $B = 0.28$ T.

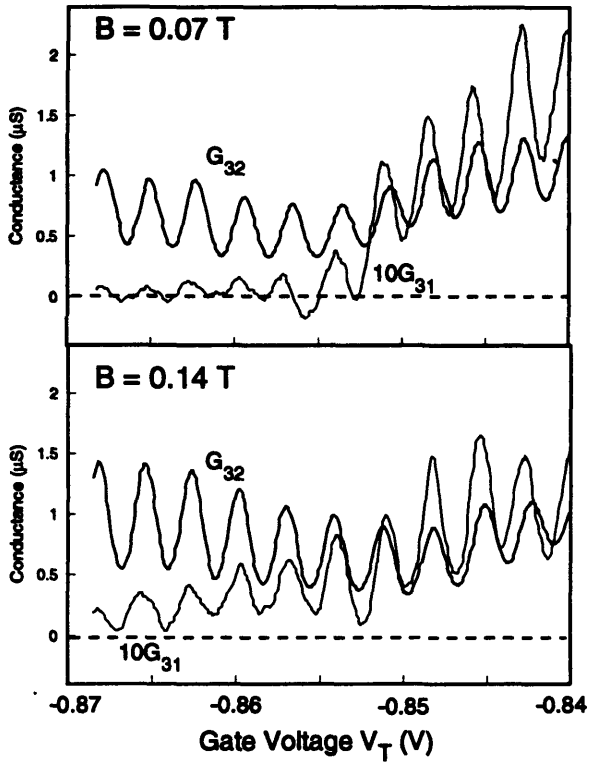


Figure 4-11: Magnified view of first few oscillations for $B = 0.07$ T and $B = 0.14$ T.

oscillations and a second valley is pulled negative. Increasing the field to $B = 0.14$ T brings G_{31} positive again.

Another change occurs in the first few oscillations, which are magnified in Fig. 4-11 for $B = 0.07$ T and $B = 0.14$ T. For $B = 0.07$ T, G_{31} dips negative in a region where the oscillations were previously well-correlated and positive definite. Again, raising the field to $B = 0.14$ T seems to restore the alignment of the oscillations as well as their positive definiteness. Finally, doubling the field again to $B = 0.28$ T causes large changes to both G_{31} and G_{32}.

Chapter 5

Theoretical Modeling

This chapter focuses on trying to understand the experimental results presented in Chapter 4. Section 5.1 discusses the weakly blockaded regime. Then two possible models are presented: Sect. 5.2 discusses the polarization model and Sect. 5.3 discusses the Ohmic model.

5.1 The Weakly Blockaded Regime

The data of Sects. 4.2 and 4.3 suggest that the anti-correlation effect occurs only when the dot has opened up significantly and is therefore in the weakly blockaded regime. Quantitative support for this hypothesis can be found by analyzing the lineshapes for the data of Device B, which showed clear regions of correlated and anti-correlated oscillations. Figure 5-1 shows a fit of the total conductance[1] $G_{31} + G_{32}$ to a thermally broadened lineshape

$$G(V_g) = \sum_i G_{ri} \cosh^{-2}\left(\frac{e\alpha(V_g - V_{ri})}{2.5k_BT}\right) \tag{5.1}$$

[1]The individual conductances G_{31} and G_{32} are not used for lineshape analysis because of the negative dips in G_{32} in the anti-correlated regime. In the correlated regime, the thermally broadened lineshape fits equally well to G_{31} and G_{32} as it does to $G_{31} + G_{32}$.

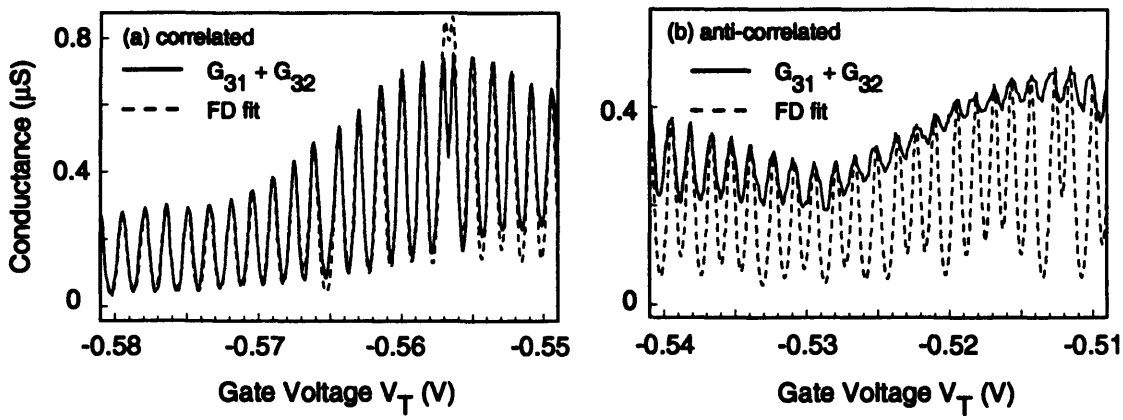


Figure 5-1: Fit of a thermally broadened lineshape to the data for Device B in the (a) correlated and (b) anti-correlated regimes.

in accordance with Eq. (2.14).[2] The sum is taken over all resonances, and α is a fitting parameter which physically represents the energy-to-voltage conversion ratio. In the correlated regime, the fit of the thermally broadened lineshape to the total dot conductance is excellent. However, in the anti-correlated regime, the fit of $G_{31}+G_{32}$ progressively worsens, predicting valleys much deeper and linewidths much narrower than given by the data.

Although the large valley conductance is a clear sign of the quenching of the Coulomb blockade, the precise manner in which the single electron tunneling (SET) picture breaks down has been a source of controversy. Foxman et al. [15] have found that the lineshapes can be fit more generally by a thermally-broadened Lorentzian parameterized by a FWHM of Γ, interpreted as the lifetime broadening of a resonance for a non-interacting system. In this case, both α and Γ are treated as fitting parameters, with $\alpha = C_g/C$ yielding directly the charging energy ($U = e^2/C$ in their notation). The results of the experiment of Foxman et al. [15] are shown in Fig. 5-2. As the gate voltage V_g is raised, the charging energy U falls rapidly while the lifetime broadening Γ increases exponentially. The charging energy

[2]It is more likely that we are in the classical regime because of the slowly varying envelope of the peak amplitudes, indicating no resonant structure due to individual energy levels. This is also reasonable considering the temperature of the experiment in relation to the expected energy level splitting.

U deduced from the lineshape agrees with that found from the sum of the gate and lead capacitances. The dramatic drop in the charging energy arises from the rapid increase of the capacitance C_r between the dot and one of the leads. Therefore the washout of the Coulomb blockade is viewed as a geometrical effect in which the divergence of the capacitance quenches the charging energy.

With finite dot-lead coupling, the first correction to the SET picture occurs by a *correlated* tunneling process called *co-tunneling,* in which current flows through the dot by simultaneous tunneling of an electron from source to dot and from dot to drain [51, 52]. Co-tunneling is a virtual process which does not require energy conservation in the dot. Pasquier et al. [49] have fit the finite off-resonance conductance to a theoretical inelastic co-tunneling current of the form [51]

$$I_{cot} = \frac{h}{6e^2}\sigma_\ell\sigma_r\left(\frac{1}{E_1}+\frac{1}{E_2}\right)^2\left[(k_BT)^2+\left(\frac{eV_{ds}}{2\pi}\right)^2\right]V_{ds}, \tag{5.2}$$

where $\sigma_{\ell,r}$ are the conductances of the tunnel barriers and $E_{1,2}$ are the changes in circuit energy for the two co-tunneling events. As the conductances $\sigma_{\ell,r}$ increase, the finite conductance at the Coulomb blockade valleys is dominated by co-tunneling with an algebraic dependence on temperature rather than by the simple thermal activation found in the SET picture.

In our data the valleys rise so rapidly that a lineshape analysis in the anti-correlated regime is difficult. We have attempted without success to explain the anti-correlation effect using the co-tunneling picture. However, the conductances in our system (e.g., G_{12} in Fig. 4-8) are large enough that co-tunneling and other corrections cannot be viewed in a perturbation approach, which is valid only when all conductances are small compared to $\sigma_Q = 2e^2/h$. Indeed, the very weak oscillations in G_{12} for Device C (Fig. 4-8) are reminiscent of the weak oscillations studied in Ref. [49] for the case of tunnel conductances comparable to σ_Q, when the dot is nearly in the Ohmic regime. A theoretical treatment in this regime

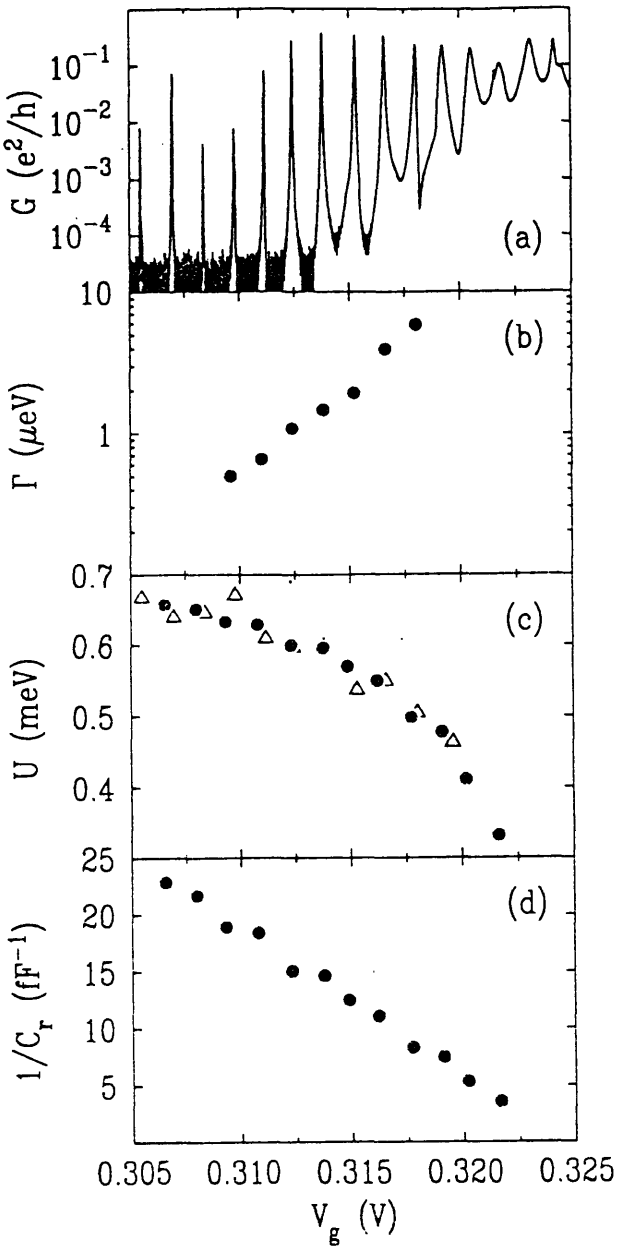


Figure 5-2: Results of the experiment of Foxman et al. [15] (a) As the gate voltage is increased, the fit of the resonances to the derivative of the Fermi-Dirac function breaks down in favor of a thermally broadened Lorentzian parameterized by a FWHM of Γ. (b) The fit value of Γ, the FWHM of the thermally broadened Lorentzian. (c) The charging energy, defined here to be $U = e^2/C$, deduced from the lineshape fit (triangles) and from summing the capacitances (filled circles). (d) The capacitance C_r between the dot and the right lead increases rapidly as the peak-to-valley ratio decreases.

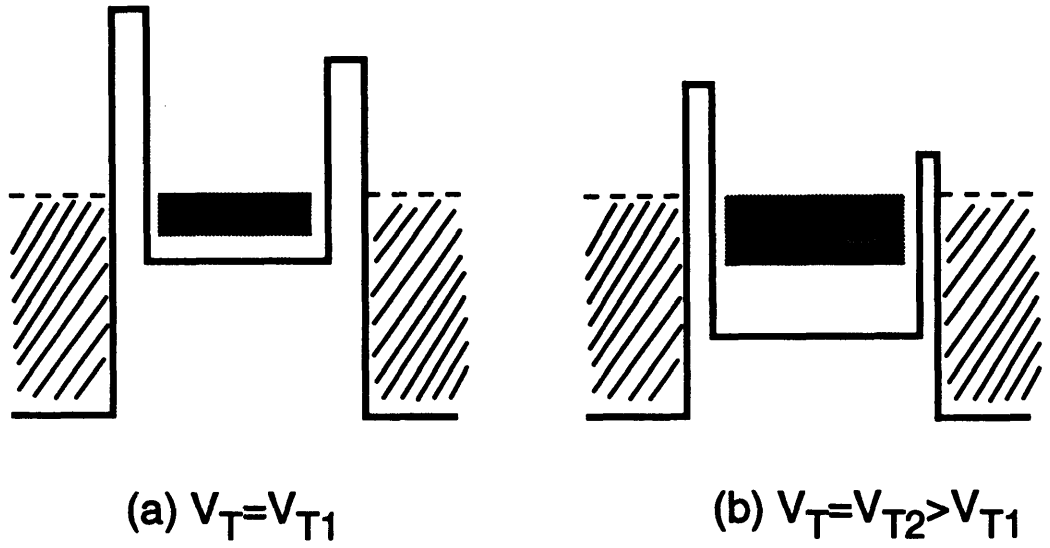


Figure 5-3: Energy bands for case of asymmetric barriers. As top gate voltage is increased from (a) to (b), the tunnel barriers become softer and the effective asymmetry is amplified. The shaded area represents filled states in the dot at a conductance maximum.

is lacking. In light of these difficulties, we choose in Sect. 5.2 to focus on a model based on classical capacitances rather than on a co-tunneling approach.

5.2 Polarization Model

The model presented in this section was proposed by Jari Kinaret of Nordita. The crucial ingredient of the model is that there is some intrinsic, unintentional asymmetry between the right and left tunnel barriers which causes the local chemical potentials to be different on the left and the right sides of the dot. We can associate an effective capacitance with each barrier, the value of which depends on the barrier thickness and height. As the barriers are made softer by increasing the top gate voltage V_T, the barrier capacitances increase and the relative barrier asymmetry becomes more important, as illustrated in Fig. 5-3. In this regime it is no longer obvious that the transitions necessary to produce current peaks at the left and right leads will resonate at the same gate voltage.

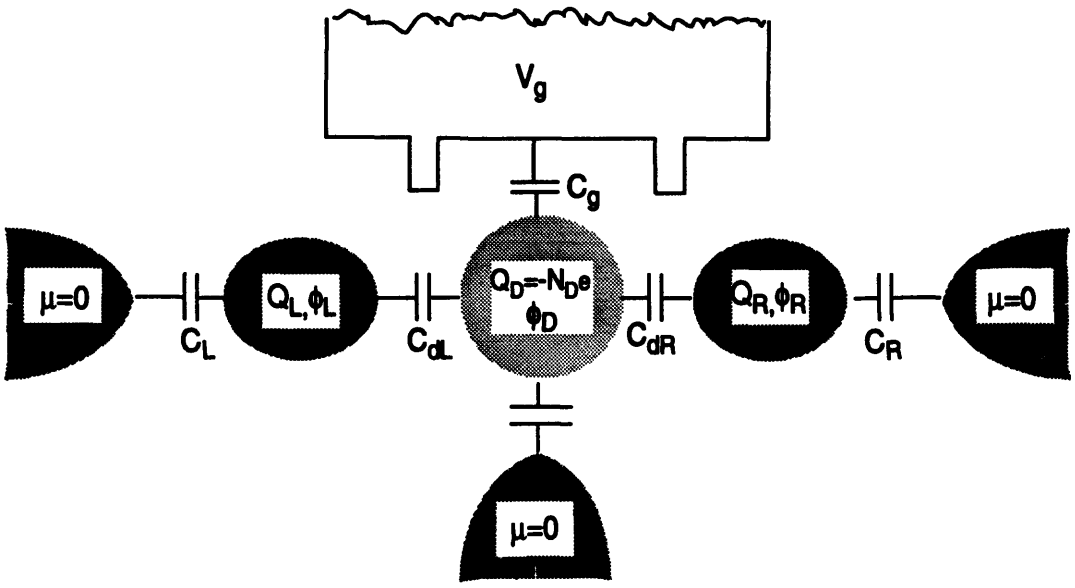


Figure 5-4: Lumped-element circuit containing dot and two quasi-reservoirs to model effect of unequal electrochemical potentials on left and right sides of the dot. The quasi-reservoirs are strongly coupled to the actual reservoirs by large capacitances C_L and C_R.

Figure 5-4 depicts a lumped-element circuit used as a first approximation to model the effect of unequal electrochemical potentials. In addition to the dot with quantized charge $Q_D = -N_D e$, the model contains two quasi-reservoirs which are coupled to the actual reservoirs by large capacitances C_L and C_R. Physically, if a test charge is placed in a quasi-reservoir, representing the part of the lead nearest the barrier, it will be partially imaged in the dot and partially imaged in the actual reservoir. The small capacitance to the injector reservoir is not included in the calculation for simplicity. The strategy is to calculate the resonant gate voltages for the transitions leading to peaks in the left and right currents.

Following the procedure used in Sect. 2.1, the potentials of the dots and quasi-reservoirs are expressed as functions of the charges and gate voltage:

$$\phi_D(Q_L, Q_D, Q_R; V_g) = \frac{1}{\tilde{C}_d}(Q_D + \gamma_L Q_L + \gamma_R Q_R + C_g V_g) \tag{5.3}$$

$$\phi_L(Q_L, Q_D, Q_R; V_g) = \frac{1}{\tilde{C}_L}(Q_L + C_{dL}\phi_D) \tag{5.4}$$

$$\phi_R(Q_L, Q_D, Q_R; V_g) = \frac{1}{\tilde{C}_R}(Q_R + C_{dR}\phi_D). \tag{5.5}$$

Here $\gamma_{L(R)} = C_{dL(dR)}/\tilde{C}_{L(R)}$, $\tilde{C}_{L(R)} = C_{L(R)} + C_{dL(dR)}$, and $\tilde{C}_d = C_g + (1-\gamma_L)C_{dL} + (1-\gamma_R)C_{dR}$. The electrostatic energy of the circuit in Fig. 5-4 is

$$\begin{aligned} W(Q_D, Q_L, Q_R; V_g) &= \int_0^{Q_D} \phi_D(0, Q'_D, 0; V_g)\, dQ'_D + \int_0^{Q_L} \phi_L(Q'_L, Q_D, 0; V_g)\, dQ'_L \\ &+ \int_0^{Q_R} \phi_R(Q_L, Q_D, Q'_R; V_g)\, dQ'_R \qquad (5.6) \\ &= \frac{1}{2\tilde{C}_d}(Q_D + \gamma_L Q_L + \gamma_R Q_R)^2 + \frac{1}{2\tilde{C}_L}Q_L^2 + \frac{1}{2\tilde{C}_R}Q_R^2 \\ &+ \frac{C_g}{\tilde{C}_d}V_g(Q_D + \gamma_L Q_L + \gamma_R Q_R). \qquad (5.7) \end{aligned}$$

The model assumes that the time-averaged charges $\overline{Q}_L$ and $\overline{Q}_R$ are given by the classical electrostatics relations $\partial W/\partial Q_{L(R)} = 0$, yielding

$$\overline{Q}_L(\overline{Q}_D; V_g) = -\frac{C_{dL}}{C_g + C_{dL} + C_{dR}}(\overline{Q}_D + C_g V_g) \tag{5.8}$$

$$\overline{Q}_R(\overline{Q}_D; V_g) = -\frac{C_{dR}}{C_g + C_{dL} + C_{dR}}(\overline{Q}_D + C_g V_g), \tag{5.9}$$

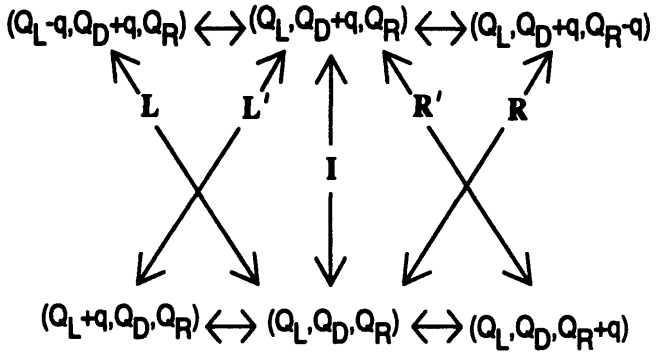
where $\overline{Q}_D$ is the time-averaged dot charge.

The next step is to choose a basis of states $(Q_L^{(i)}, Q_D^{(i)}, Q_R^{(i)})$ for the system consisting of the dot and the two quasi-reservoirs. The set of states used in the model is shown in Fig. 5-5(a) and includes the lowest-energy unit polarization fluctuations of the states (Q_L, Q_D, Q_R) and $(Q_L, Q_D + q, Q_R)$. The set of states is *minimal* because it includes only the crucial states[3] and *symmetric* because $q = e$ and $q = -e$ give the same result. Each state

[3]For example, the four excited states $(Q_L + q, Q_D - q, Q_R)$, $(Q_L, Q_D - q, Q_R + q)$,$(Q_L - q, Q_D + 2q, Q_R)$, and $(Q_L, Q_D + 2q, Q_R - q)$ are discarded even though they are unit polarization fluctuations.

i	P_i	$Q_L^{(i)}$	$Q_D^{(i)}$	$Q_R^{(i)}$
1	P_1	Q_L	Q_D	Q_R
2	P_2	Q_L	Q_D+q	Q_R
3	P_3	Q_L	Q_D+q	Q_R-q
4	P_4	Q_L	Q_D	Q_R+q
5	P_5	Q_L+q	Q_D	Q_R
6	P_6	Q_L-q	Q_D+q	Q_R

(a)



(b)

Figure 5-5: (a) Minimal, symmetric set of states used in polarization model. (b) Possible transitions between states.

i has a certain probability P_i of occupancy. In equilibrium (i.e., $V_{ds} = 0$), the occupancy probabilities are found in the grand canonical ensemble[4] so that

$$P_i = \frac{\exp(-W_i/k_BT)}{\sum_i \exp(-W_i/k_BT)}, \tag{5.10}$$

where W_i is the energy of state i calculated from Eq. (5.7).

The possible transitions between the states are shown in Fig. 5-5(b). The procedure for finding the gate voltages at which each of the 5 transitions I,L,L′,R,R′ are at resonance is as follows. The electrostatic energies of the two states i and j involved in each transition are equated, relating the resonant gate voltage V_g to the charges $(Q_L, Q_D = -N_De, Q_R)$:

$$W(Q_L^{(i)}, Q_D^{(i)}, Q_R^{(i)}; V_g) = W(Q_L^{(j)}, Q_D^{(j)}, Q_R^{(j)}; V_g). \tag{5.11}$$

The charges Q_L and Q_R can be related back to the gate voltage V_g and dot charge $Q_D = -N_De$ by equating the time-averaged charges $\overline{Q}_L$ and $\overline{Q}_R$ for the basis set with the classical electrostatics values given in Eqs. (5.8-5.9). Carrying out the time averages

[4]Recall that the zero of energy is taken to be the Fermi energy E_F in the reservoirs.

$\overline{Q}_{L(R)} = \sum_i P_i Q^{(i)}_{L(R)}$ for the basis set in Fig. 5-5(a) yields

$$Q_L = \overline{Q}_L(\overline{Q}_D; V_g) - \delta q_L \quad \text{with} \quad \delta q_L = (P_5 - P_6)q \tag{5.12}$$

$$Q_R = \overline{Q}_R(\overline{Q}_D; V_g) - \delta q_R \quad \text{with} \quad \delta q_R = (P_4 - P_3)q, \tag{5.13}$$

where the time-averaged charge $\overline{Q}_D = \sum_i P_i Q^{(i)}_D$ is similarly given by

$$\overline{Q}_D = Q_D + \delta q_D \quad \text{with} \quad \delta q_D = (P_2 + P_3 + P_6)q. \tag{5.14}$$

If Eq. (5.14) is substituted into Eqs. (5.8-5.9), Eqs. (5.12-5.13) and the resonant condition Eq. (5.11) yield three coupled transcendental equations in the unknowns V_g, Q_L, and Q_R for each of the 5 transitions shown in Fig. 5-5(b).

In general this system must be solved numerically, but let us illustrate some special limiting cases. In the limit $C_L, C_R \gg C_g, C_{dL}, C_{dR}$, we must recover the conventional SET theory result that all transitions are at resonance at the same gate voltage. In this limit, Eq. (5.7) reduces to

$$W(Q_D, Q_L, Q_R; V_g) = \frac{Q_D^2}{2C_g} + Q_D V_g, \tag{5.15}$$

which is the same as Eq. (2.3). All 5 transitions are found to resonate at the same gate voltage $V_g = e(N_D - \frac{1}{2})/C_g$ with q set to e.

As the tunnel barriers become softer with increasing V_g, the capacitances C_{dL},C_{dR} become comparable to C_L,C_R. As a simplifying example, let us consider the extreme asymmetric limit $C_{dL} = C_L = C_R \gg C_g$,$C_{dR} = 0$ so that $\gamma_L = 1/2$ and $\gamma_R = 0$. Physically, this corresponds to the case where the left tunnel barrier has essentially disappeared. Solving Eq. (5.11) for the resonant gate voltages for each transition yields (for the case $q = e$)

$$N_g^I = N_D + (\alpha_D + \alpha_L) - 1 \tag{5.16}$$

$$N_g^L = N_D - \frac{1}{2} \tag{5.17}$$

$$N_g^{L'} = N_D - \frac{1}{2} \tag{5.18}$$

$$N_g^R = N_D + (\alpha_D + \alpha_L) - \alpha_R - \frac{3}{2} \tag{5.19}$$

$$N_g^{R'} = N_D + (\alpha_D + \alpha_L) - \alpha_R - \frac{1}{2}, \tag{5.20}$$

where $N_g = C_g V_g/e$ and $\alpha_{L(D,R)} = \delta q_{L(D,R)}/e$. At the injector resonance, $P_1 = P_2 = P_5 = P_6$ and $P_3 = P_4$ so that $\alpha_D = 1/2, \alpha_L = 0$; thus we find that the transitions I,L,L′ are all at resonance when

$$N_g^I = N_g^L = N_g^{L'} = N_D - \frac{1}{2}. \tag{5.21}$$

The equations for the right resonances R,R′ must be solved numerically and are shown in Fig. 5-6 as a function of U/k_BT where $U = e^2/C_L$. The right resonances are strongly shifted; the temperature dependence can be understood by noting that for $U/k_BT \ll 1$ all states become equally populated, requiring $\alpha_L = \alpha_R = 0, \alpha_D = 1/2$. For $U/k_BT \leq 2$, the right resonances have moved *midway* between successive left/injector resonances, indicated by dashed lines.[5]

Figure 5-6 also indicates the most energetically favorable states of the system, which are degenerate at $N_g = N_D - \frac{1}{2}$. A curious consequence of the minus sign in front of the α_R term in Eqs. (5.19-5.20) is that the states that participate in the R and R′ transitions are *excited* states of the system over the gate voltage range for which the transitions are at resonance. Thus the right resonances always require thermal activation.

To study how the transitions evolve as the left and right tunnel barriers open up, we can fix T in the high temperature regime and vary the capacitances C_{dL}, C_{dR}. Figure 5-7 shows

[5]The low temperature ($U/k_BT \gg 1$) behavior should be interpreted with caution, since the right resonances begin to overlap with the neighboring left/injector resonances.

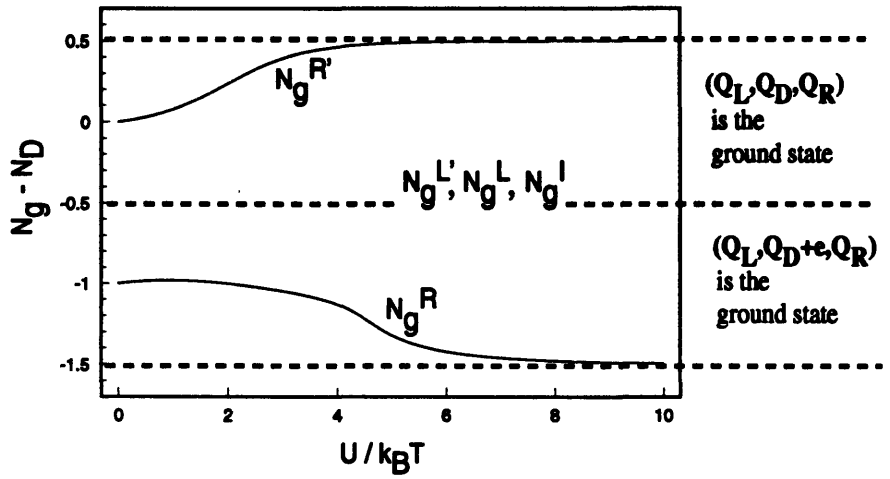


Figure 5-6: Gate voltage at which transitions R,R′ are at resonance as a function of U/k_BT where $U = e^2/C_L$, in the extreme asymmetric limit. The dashed lines indicate resonance positions for transitions I,L,L′.

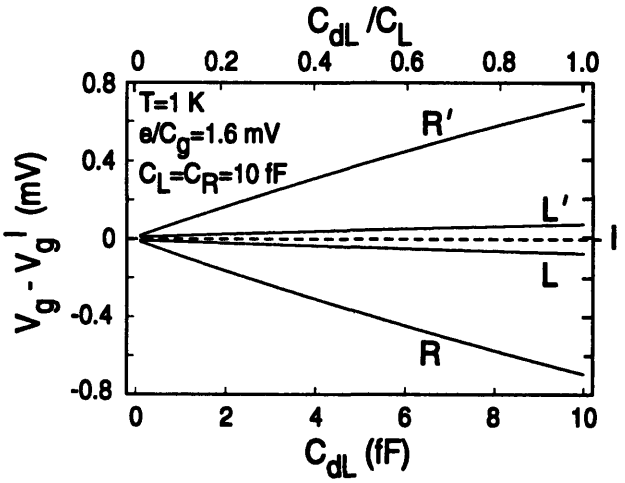


Figure 5-7: Gate voltages at which each of the transitions L,L′,R,R′ are at resonance (relative to the position of the I resonance) as capacitances C_{dL} and C_{dR} are increased, for the highly asymmetric case $C_{dL} = 10C_{dR}$.

the results of a calculation in which the resonant gate voltage positions for the transitions L,L′,R,R′ are plotted as a function of C_{dL} for the highly asymmetric case $C_{dL} = 10C_{dR}$. With increasing $C_{dL} = 10C_{dR}$, the injector resonance is unchanged and the left resonances are shifted only a little. However, the right resonances are strongly shifted; in the limiting case $C_{dL}/C_L \simeq 1$, when the left barrier has essentially disappeared, transitions R,R′ are at resonance almost midway between successive injector resonances, as in Fig. 5-6.

As in Sect. 2.1, the approach so far is sufficient for calculating the resonant gate voltages for each transition but not the currents at the left and right leads. A peak in the current measured at the left lead could occur by one of two possible chains of events:

$$\mathcal{L}' : (Q_L, Q_D, Q_R) \xrightarrow{I} (Q_L, Q_D + q, Q_R) \xrightarrow{L'} (Q_L + q, Q_D, Q_R) \rightarrow (Q_L, Q_D, Q_R) \quad (5.22)$$

$$\mathcal{L} : (Q_L, Q_D, Q_R) \xrightarrow{L} (Q_L - q, Q_D + q, Q_R) \rightarrow (Q_L, Q_D + q, Q_R) \xrightarrow{I} (Q_L, Q_D, Q_R). \quad (5.23)$$

The $\mathcal{L}'$ ($\mathcal{L}$) chain of events describes the transfer of a charge q ($-q$) from the injector to the left reservoir.[6] Analogous chains of events $\mathcal{R}'$,$\mathcal{R}$ give rise to current peaks at the right lead. In the extreme asymmetric limit, we have found that for $\mathcal{R}'$,$\mathcal{R}$, all transitions do not resonate at the same gate voltage and some transitions must be thermally activated. The currents can be found by using a rate-equation approach to determine the probabilities $\{P_i\}$ rather than the equilibrium distribution assumed in Eq. (5.10). Figure 5-8 shows the currents I_L,I_R, and $I_L + I_R$ obtained through such a calculation, in which 20 states are used instead of the 6 in the minimal, symmetric model.[7] When C_{dL},C_{dR} are still comparable to C_g, the current peaks are aligned. As C_{dL},C_{dR} are increased, the peaks evolve from being slightly separated in gate voltage to being completely anti-aligned when $C_{dL} \gg C_{dR} \gg C_g$. Over the same range, the total current $I_L + I_R$ evolves from being strongly modulated to being very weakly modulated.

[6] The $\mathcal{L}$ chain of events can be written in a form similar to that of $\mathcal{L}'$ by substituting $Q'_D = Q_D + q, q' = -q$.

[7] Actually, the minimal, symmetric model was conceived to physically understand these results.

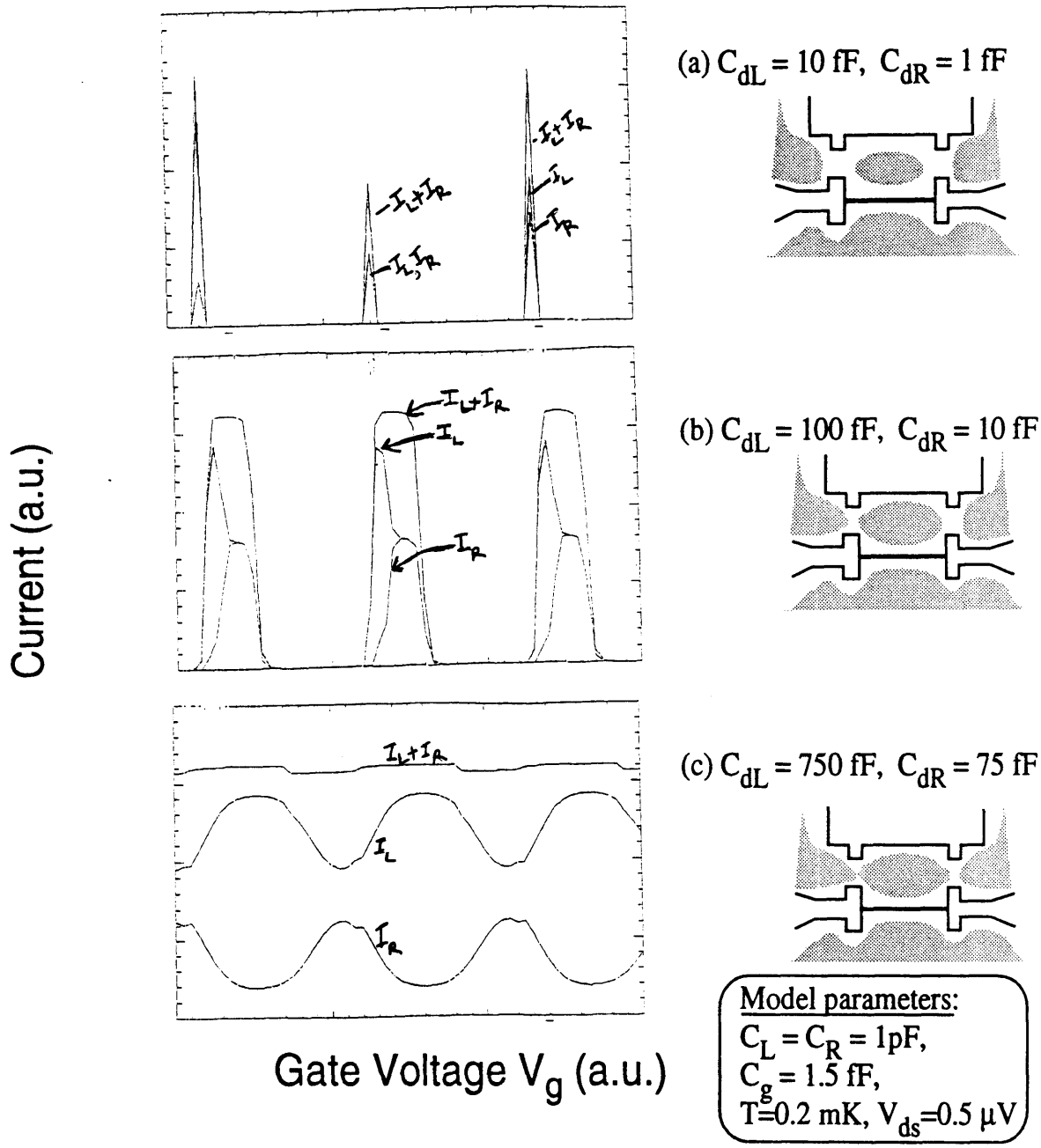


Figure 5-8: Results of numerical calculation showing separation of peaks in I_L and I_R as barrier capacitances are increased, leading to full anti-alignment.

5.3 Ohmic Model

The idea on which this model is based was suggested by David Carter of MIT. Suppose that one replaces each of the tunnel barriers in the three-terminal quantum dot by a simple resistor, as shown in the inset of Fig. 5-9. As the top gate voltage is increased, the resistances R_1 and R_2 will drop rapidly while R_3 will remain approximately constant. If $R_3 \gg R_1, R_2$, then R_3 will act as a current-limiting resistor so that the total current $I_{tot} = I_{31} + I_{32} \simeq V_{ds}/R_3$. If the resistances R_1, R_2 are changed in this current-biased regime, the currents I_{31} and I_{32} adjust themselves to keep the total current constant. Thus, even when R_2 decreases, I_{32} also decreases if R_1 falls more than R_2.

This increase in one current at the expense of the other is reminiscent of our conductance data in the regime where the oscillations are anti-correlated. To illustrate this, let us make an ad hoc assumption about the gate-voltage dependences of the conductances $G_1 \equiv (R_1)^{-1}$ and $G_2 \equiv (R_2)^{-1}$:

$$\begin{aligned} G_1 &= \sigma_Q \exp\left[-\alpha(N_g - N_{g1})\right] (1 - A \sin 2\pi N_g) \\ G_2 &= \sigma_Q \exp\left[-\alpha(N_g - N_{g2})\right] (1 - B \sin 2\pi N_g), \end{aligned} \tag{5.24}$$

where $N_g = C_g V_g/e$ is the normalized gate voltage. The exponential term represents the usual behavior of a QPC near turn-on while the sinusoid imposes some modulation with the characteristic e/C_g periodicity. To introduce some asymmetry between the QPCs, N_{g1} and N_{g2} are made unequal, as are A and B. Figure 5-9 shows the conductances $G_{31} = I_{31}/V_{ds}$, $G_{32} = I_{32}/V_{ds}$, and $G_{tot} = G_{31} + G_{32}$, as a function of V_g, for the choice of parameters $R_3 = 1$ MΩ, $\alpha = 0.2645$, $N_{g1} = 40$, $N_{g2} = 44.4$, $A = 0.55$, and $B = 0.45$. The oscillations evolve from being aligned to being anti-aligned, corresponding to the transition from a voltage-biased to a current-biased system.

The qualitative agreement with experiment is quite good. A fundamental question is

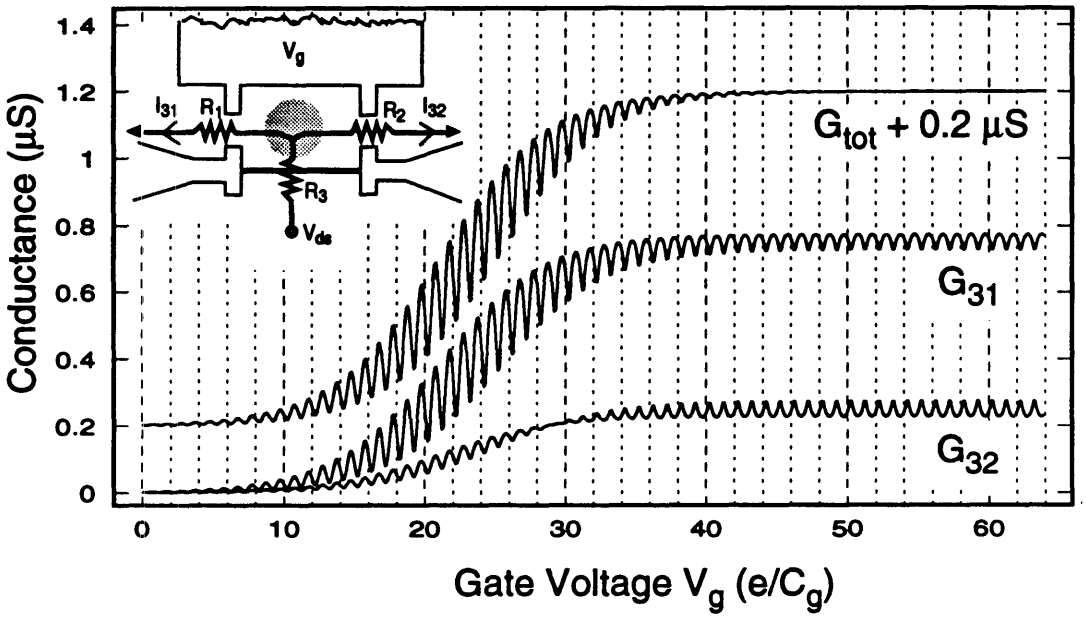


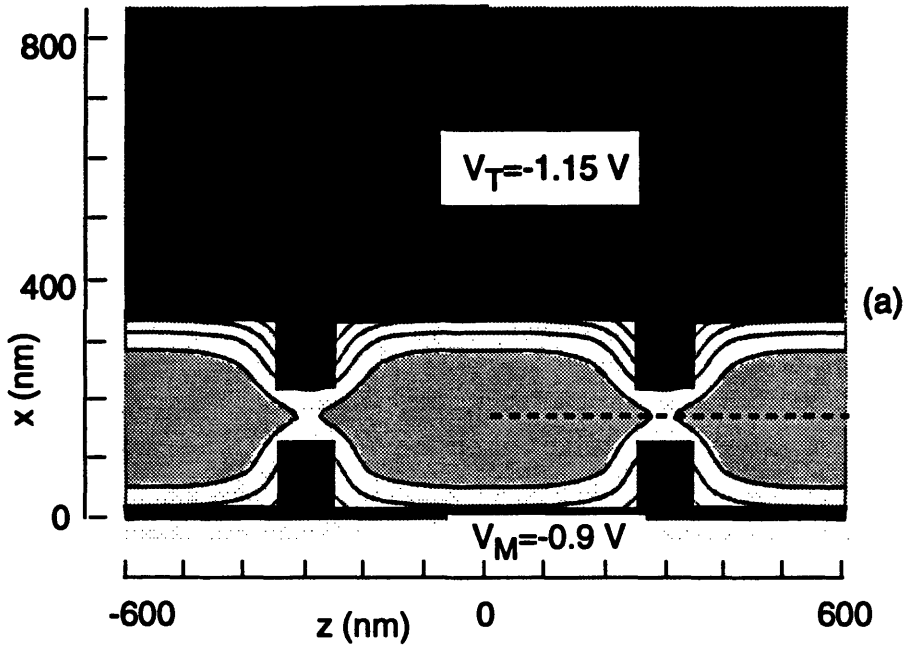
Figure 5-9: Conductances G_{31}, G_{32}, and $G_{tot} = G_{31} + G_{32}$ in Ohmic model. Inset shows Ohmic model of conduction in three-terminal quantum dot.

whether the ad hoc form in Eq. (5.24) has any justification. According to SET theory, the potential energy at the bottom of the dot undergoes a sawtooth variation with gate voltage V_g

$$-e\phi(N) = \frac{Ne^2}{C} + \frac{-eC_gV_g}{C} + \text{const} \tag{5.25}$$

as each new electron is added to the dot. This variation also affects the transmission of the tunnel barriers and could account for the periodic modulation central to the model.

To estimate the effect of this sawtooth variation on the tunnel barriers, we carried out a numerical simulation, using a three-dimensional Poisson solver similar to that used for the calculations described in Sect. 2.2. Figure 5-10(a) shows equipotential contours in the plane of maximum electron density, calculated for the device in Fig. 4-4 with $L = 0.6$ μm. Symmetry allows us to model only one half of the device structure ($z > 0$), reducing



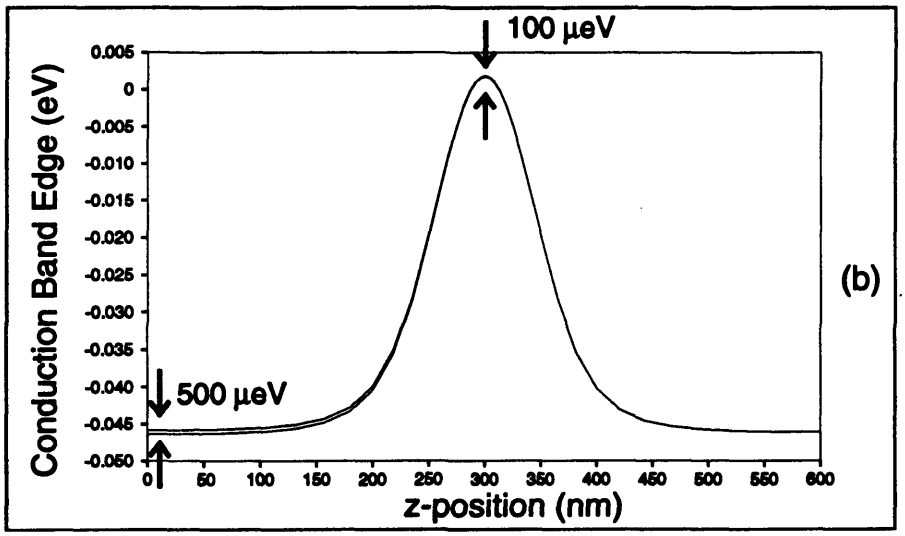


Figure 5-10: (a) Equipotential contours in the plane of maximum electron density, calculated for $V_T = -1.15$ V, $V_M = -0.9$ V. The area populated by electrons (where the conduction band edge $E_C < 0$) is shaded. (b) Conduction band edge along the cut indicated by dashed line in (a), showing the shift as the dot number changes from 172 to 173 electrons.

computation time. The gap of the split-gate is taken to be 100 nm instead of the lithographic value of 150 nm, as previous studies have found that the depletion gates have some effective extension into the open surface of the GaAs [65]. The rest of the open surface is modeled as an equipotential due to pinning of the Fermi level, as in Ref. [64]. The calculation is done in the semiclassical Thomas-Fermi approximation at temperature 4.2 K. However, the donor charge in the AlGaAs is assumed to be "locked" at the value attained at 100 K and thus does not change as the gate voltage is swept negative [71].

The standard Thomas-Fermi model neglects the quantization of the dot charge. [8] We can take the quantization into account by fixing the electron *number* in the dot, as in Ref. [21].[9] For the gate voltages used ($V_T = -1.15$ V, $V_M = -0.9$ V), the standard Thomas-Fermi model finds the dot to have approximately 172.5 electrons. Such a half-integer occupancy corresponds to a Coulomb blockade degeneracy where the bottom of the dot shifts upward as the electron number changes discretely from 172 to 173 electrons. We can evaluate the effect of this discrete jump on the tunnel barrier transmission by comparing the self-consistent potential for 172 electrons with that for 173 electrons, keeping the gate voltages fixed. Figure 5-10(b) shows the conduction band edge for the two cases along the cut indicated in Fig. 5-10(a). The addition of the 173rd electron raises the potential energy by about 500 μeV at the center of the dot while the top of the potential barrier shifts upward by about 100 μeV.

To evaluate the effect of this shift on the barrier transmission, we can model the conduction band edge in the vicinity of the QPC (at $\overline{x} = 0, \overline{z} = 0$) as a saddle-point of the form

$$E_c(\overline{x}, \overline{z}) = E_{cmax} + \frac{1}{2} m^* \omega_x^2 \overline{x}^2 - \frac{1}{2} m^* \omega_z^2 \overline{z}^2, \tag{5.26}$$

[8] Quantum-mechanically, the quantization can be included by evaluating the dot occupation probabilities in the canonical ensemble, where the number is fixed.

[9] Electrons in the dot still obey semiclassical statistics, but with a quasi-Fermi level which gives the specified electron charge for the self-consistent potential.

where E_{cmax} is the energy at the top of the barrier relative to the Fermi level in the reservoir and m^* is the electron effective mass. As discussed in Ref. [72], the conductance of a saddle-point constriction is

$$G_{saddle} = \sigma_Q \frac{1}{1 + \exp[(2\pi E_{cmax} + \pi\hbar\omega_x)/\hbar\omega_z]}. \tag{5.27}$$

As a simplifying approximation, we assume that effect of increasing the top gate voltage V_T is to change E_{cmax} linearly; the shape of the potential, parameterized by $\hbar\omega_x = 12$ meV and $\hbar\omega_z = 4.9$ meV, remains approximately constant. We thus assume that $E_{cmax}(V_T)$ depends on V_T in a sawtooth fashion, as illustrated in Fig. 5-11(a). The background slope and the period of the oscillation are extrapolated from the solution at another bias point close to $V_T = -1.15$ V,$V_M = -0.9$ V.

We can now use Eq. (5.27) to find the conductance as a function of gate voltage V_T. To introduce some asymmetry, we assume that the turn-on voltage of one QPC is shifted by 10 mV from that of the other QPC. Figure 5-11(b) shows the conductances G_{31} and G_{32}, where $R_3 = 1$ MΩ as in Fig. 5-9. Note that the conductance does not become appreciable until $E_{cmax} < 0$ – i.e., until the barrier region is no longer classically forbidden. This is a consequence of energy quantization in the lateral confinement, which is neglected in the semiclassical approximation. Unfortunately it is not possible to solve for the potential with a fixed dot charge using the method above in the regime $E_{cmax} < 0$ because the reservoir and the dot must be separated by a region depleted of carriers to enforce the condition of charge quantization; this is why we chose a voltage bias point such that $E_{cmax} > 0$ even though the transmission probability is very small.

The same *qualitative* trend leading to anti-aligned oscillations is observed in Fig. 5-11 as in Fig. 5-9; *quantitatively*, however, the amplitude of the oscillations is very small. Physically, the modulation is weak because the 100 μeV shift in the height of the tunnel barrier is small compared to the millivolt scale of $\hbar\omega_x$.

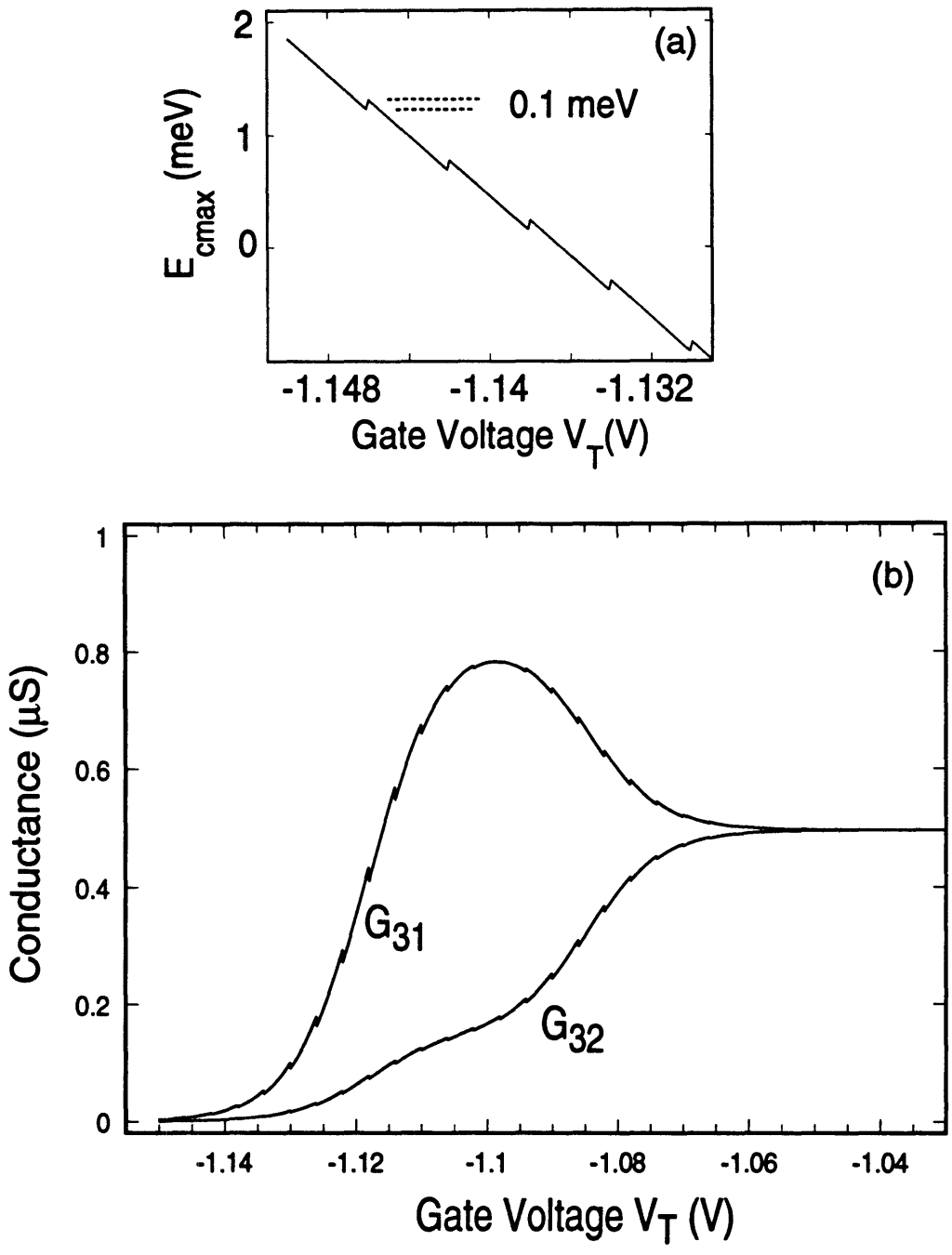


Figure 5-11: (a) Sawtooth variation of conduction band edge E_{cmax} with gate voltage V_T. (b) Conductances calculated in Ohmic model using saddle-point potentials.

Thus, the weak modulation of the tunnel barriers seems inadequate to explain the strong anti-aligned oscillations observed experimentally. Nevertheless, the uncanny resemblance of the experimental data to the simulated results in Fig. 5-9 precludes us from discarding the Ohmic model. It still seems plausible that the injector barrier limits the total current, even if a fundamental reason for the modulation with periodicity e/C_g is not evident.

Chapter 6

Conclusion and Future Work

This thesis has described the first realization and study of a device structure in which a quantum dot is contacted by *three* electron reservoirs. All previous transport studies had focused on quantum dots coupled to *two* electron reservoirs. When the coupled-dot structure was originally conceptualized, the three-terminal quantum dot was not at all envisioned. Rather, the biasing configuration used to realize the three-terminal dot was intended simply to evaluate the degree of asymmetry between the tunnel barriers.

As expected, conductance oscillations periodic in gate voltage were observed at both output leads of the three-terminal device. Our original expectation was that oscillations measured at the two leads would always align with each other in gate voltage, in accordance with conventional single electron tunneling (SET) theory. We found this prediction to hold well when the dot is in the strongly blockaded regime. However, as the tunnel barriers are made softer, we found that the oscillations are sometimes *anti-correlated* in gate voltage, with the current even changing sign at some of the deep valleys.

The anti-correlated oscillations were observed in devices with a continuous gate defining the injector barrier as well as with a split gate. The effect seems to be more robust in the former. Except in Ref. [73], the split-gate scheme has been almost universally used to form

tunnel barriers in planar devices. Our work confirms that a continuous gate is indeed a viable alternative to the split gate, although more challenging lithographically.

We established that the anti-correlated oscillations occur in the weakly blockaded regime. Although this regime is not well understood theoretically, we presented two models – a polarization model and an Ohmic model. Both models predict anti-correlated oscillations in the weakly blockaded regime. However, neither model is able to explain the sign change of the current at some of the deep valleys.

In Chapter 1, it was stated that the Coulomb blockade effect could be exploited to realize a "single electron transistor." By the same token, one could say that in the anti-correlated regime the three-terminal dot functions as a "single electron directional coupler," in which a gate is used to steer the current to one of two possible outputs. As such, the switching is based on a charging effect in which the gate voltage is required to change the dot occupancy rather than on a wave-interference effect in which the gate controls the transfer length between two waveguides [74].

Of course it is difficult to envision a practical application for any quantum-effect device unless the operating temperature can be raised. The observability temperature for charging effects could be raised by reducing the effective capacitance which sets the charging energy. Such a reduction in capacitance might be achieved by further reduction of device dimensions, using etching rather than gating to achieve confinement, or exploring different material systems. All of these are current areas of active research.

The present process has used electron beam lithography to pattern the fine features. As discussed in Sect. 3.3, the total number of steps in the process could be reduced if x-ray lithography were used in place of electron beam lithography. The risk of mobility degradation during processing is also reduced if x-ray lithography is used [75].

This thesis work was originally motivated by the prospect of realizing a coupled dot structure. A limitation of the original design consisting of three gates to define two dots is the inability to control independently the tunnel barriers which couple the dot to the

leads. In addition, it is desirable to have an additional gate to control the dot size. [1] Without independent control, the yield was found to be sufficiently low that it was nearly impossible to find a coupled dot device in which both dots showed robust Coulomb blockade oscillations. For this reason, Geometry C in Fig. 3-4 is a much more versatile design and will likely be a prototype for future work. Unfortunately, the samples where this design was implemented all suffered from various fabrication problems.

At the outset it was envisioned that the coupled dot structure could be used to study the interaction of two dots either in series or in parallel. The conductance through two dissimilar dots in series was predicted to show quasi-random oscillations [18]. The non-periodic peaks resulting from the interplay of a small dot with a larger one has quite possibly been observed recently [76, 77]. However, it would still be interesting to pursue this effect in our structure, where the series conductance through two dots can be compared directly with the conductance through a single dot. It is possible that such a system may hold a new surprise in the weakly blockaded regime.

Perhaps the most promising area for future work is the study of two interacting dots in parallel. In this configuration, one dot acts as a highly sensitive detector, responding to changes in the electrostatic potential of the other dot. It would be particularly interesting to investigate whether such a detector can function in the time domain. If so, it might be possible to detect the change in the potential arising from fluctuations in the number of electrons as current flows through the dot.

[1] Reference [55] has recently demonstrated that the electron density can be modulated by using the chip carrier as a back gate.

Appendix A

Sample Processing and Heliox Operation

Sample Processing

I. MESA ISOLATION (CW mask)

A. Solvent clean the sample:

1. Rinse in TCE.
2. Boil in TCE for 10 min (hotplate 2.5-3 or 120°C).
3. Ultrasound in acetone for 10 min.
4. Ultrasound in methanol for 10 min.
5. Blow dry (flush gun initially).

B. Spin and bake resist:

1. Spin Shipley 1813 resist at 4.5 krpm.
2. Bake at 90°C for 30 min.

C. Photolithography and development:

1. Expose resist (about 3.8 sec) at 400 nm.
2. Develop exposed resist with Shipley CD-30 developer.
3. Rinse in running DI water for about 3 times the development time.

4. Inspect in microscope.
5. Spray mask with acetone/methanol.
6. Blow dry mask.

D. Etch GaAs/AlGaAs:

1. Rinse in DI for 1 min. (optional)
2. Rinse in semico for 1 min. (optional)
3. Rinse in DI for 1-2 min.
4. Ammonium hydroxide/hydrogen peroxide etch:
   Etch 22 sec (=60-75 nm) in etchant 500 DI water: 10 NH_4OH: 3 H_2O_2.
   Sulfuric acid etch:
   Etch in 10:1 DI/etchant where etchant is H_2SO_4:H_2O_2:H_2O (1:10:10) (rate 11.5 nm/sec) for 8 sec.
5. Rinse in running DI water for 2 min.
6. Blow dry.
7. Inspect in microscope.

E. Strip resist:

1. Spray with acetone.
2. Spray with methanol.
3. Blow dry.
4. Inspect in microscope.
5. Measure mesa height.

II. OHMIC CONTACTS (CD mask)

A. Solvent clean the sample.

B. Spin and bake photoresist.

C. Chlorobenzene photolithography and development:

1. Expose resist (about 3.8 sec) at 400 nm.
2. Soak 3.5 min in chlorobenzene.
3. Blow dry.
4. Bake at 90 C for 5 min.
5. Develop exposed resist with Shipley CD-30 developer. Development time should be about 3 times longer than without chlorobenzene soak.
6. Rinse in running DI water for about 3 times the development time.
7. Inspect in microscope.
8. Spray mask with acetone/methanol.

9. Blow dry mask.

D. Surface clean the sample.

1. UV ozone for 30 sec. (Run uv beforehand for preclean.)
2. Rinse in DI water for 1 min.
3. Rinse in Semico for 1-2 min in ultrasound.
4. Rinse in DI water for 1 min.
5. Blow dry.

E. E-beam evaporate ohmic contacts: 5 nm Ni/5 nm Au/25 nm Ge/45 nm Au/10 nm Ni/50 nm Au.

F. Liftoff by soaking for 10 minutes in acetone followed by quick ultrasound.

G. Sinter ohmic contacts.

1. Solvent clean the sample in bldg 13.
2. Place sample face down on solvent-cleaned piece of SI GaAs. Sinter in RTA in forming gas at 430°C for 15 sec. (program AK430a).
3. Check tlm resistances on HP-4145. Contact resistance should be < 1 ohm-mm. Verify sheet resistance using 4-point vdp. Important: If in doubt about contact quality, check contacts at 77 K.

III. 1ST LEVEL GATE METALLIZATION (CP mask)

A. Solvent clean the sample.

B. Spin and bake photoresist.

C. Chlorobenzene photolithography and development.

D. Surface clean the sample.

E. Evaporate 15 nm Ti/140 nm Au(or 15 nm Ti/10 nm Pt /140 nm Au, in which case it is possible to reanneal after gates if necessary).

F. Liftoff metal.

IV. ELECTRON BEAM LITHOGRAPHY

A. Spin bilayer pmma, 2010/2041 (about 100 nm total).

B. Expose patterns.

C. Develop in 1:3 ipa:mibk at 21°C for 45 sec.

D. Evaporate 15-18 nm AuPd in thermal evaporator.

E. Liftoff in acetone/methlene chloride mixture (let soak 30 min) without ultrasound If this does not work, gently ultrasound, holding sample suspended in solution.

V. 2ND LEVEL GATE METALLIZATION (CM mask) – if no continuity between fine gates and 1st level gate metalization.

A. Solvent clean (no ultrasound, easy on boil).

B. Spin and bake photoresist.

C. Chlorobenzene photolithography and development.

D. Surface clean the sample (no ultrasound or uv ozone).

E. Evaporate Ti 20 nm Ti/230 nm Au.

F. Liftoff metal using spray bottles (no ultrasound).

Operation of the Heliox Insert

I. COOL TO 77 K

A. Open IVC.

1. Vent probe and remove IVC using sliding hammer, being careful not to scratch wires.
2. Clean both taper surfaces with lint-free wipes. Secure IVC to prevent it from falling, getting dirty, etc.

B. Load sample.

1. Mount sample holder, socket, and sample using Apiezon grease.
2. Apply silicone grease liberally to probe surface, but sparingly on interior of can.
3. If IVC has been vented several days, heat charcoal bean with heat gun (6" from nozzle to prevent overheat) to speed up outgas. Actually, the system should be kept under vacuum if not in use for more than 1 day.
4. Place a paper cylinder around wires to prevent thermal shorts to can).
5. Replace IVC. Do not push at top of tapered seal; the force of the vacuum will do this automatically. Be careful when sliding can on to avoid wires inside as well as to avoid removing vacuum grease.

C. Pump out IVC to 100 mTorr using rotary pump or leak detector. Tape down pickup tubes securely to the IVC.

D. Bleed in 2-10 Torr of He gas to the IVC using bladder. This may be done by filling the bladder with He gas, placing it on the nozzle of the IVC can while pinched off at the end, and very slowly opening the Speedivalve. Then place a blank over the IVC Speedivalve to prevent accidental venting of IVC.

E. To prevent freezeup, He gas must be flushed through pickup tubes before dipping the probe in LN_2. Bubbles should be visible in LN_2. Check flow through each pickup tube individually. Then close needle valves, 1K pot speedivalve valve and finally He gas valve to stop flow.

F. Leave probe in LN_2 until temperature reaches 77 K ; check temperature using sensor 3 (1K pot high temp) on ITC. When done, disconnect sensor 3 to achieve optimal base temperature. Check IVC pressure to be 77/300 of the pressure introduced at 300 K.

II. TRANSFER TO LIQUID HELIUM STORAGE VESSEL

A. Resume He gas flow through pickup tubes and transfer to the storage dewar. The flange on the probe should be as low as possible.

B. Once in storage dewar, close sorb and 1K pot needle valves and 1K pot Speedivalve and remove the helium gas lines. Slowly lower the probe into the dewar, pausing when the dewar popoff valve vents.

C. Connect the rotary pump to the 1K pot line and warm up the rotary pump (> 10 minutes). Open the 1K pot Speedivalve to pump out the 1K pot line.

D. Monitor the pressure in the IVC and wait until it drops to zero before proceeding. If the pressure does not fall to zero in a reasonable amount of time, either there is a leak or the charcoal bean is not working.

III. COOL TO 0.3K USING ITC TEMPERATURE CONTROLLER

There are three resistors. The sorb has a high temperature one. The 0.3K pot has both a high temperature one and a low temperature one. If cables A and B are connected to the repective ports on the probe, sensor 1 is the sorb resistor, sensor 2 is the low temperature 0.3K pot resistor, and sensor 3 is the high temperature 0.3K pot (should be disconnected for optimal base temperature). The temperature of the 1K pot can be monitored by interchanging the cables (see manual).

Open the 1K pot needle valve until the pressure needle is all the way full for 10 seconds. Close the 1K pot needle valve so that the LHe is pumped on. Open the needle valve again and condense more LHe if the pressure drops below 1 Torr.

Set the temperature on the sorb to be ~40 K and hit 'auto', adjusting the sorb needle valve as necessary to keep the 1K pot cold. Set the gain low (prop~180) to avoid overshooting the sorb temperature too much. When the 0.3K pot temperature reaches 1.1-1.3 K, all the ^{3}He is condensed. Turn off the sorb heater (by hitting 'man' and setting the voltage to zero) and watch the temperature drop rapidly to 0.30-0.34 K, adjusting the sorb needle valve if necessary. The 1K pot pickup tube must remain cracked open to allow a continuous fill pressure of ~1 Torr. The sorb pickup tube needle valve may be shut after the sorb temperature stabilizes to <4.6 K.

Appendix B

Reprints: Scattering in Quasi-1D Wires

This Appendix contains reprints of the following two papers written in collaboration with Philip F. Bagwell:

Arvind Kumar and Philip F. Bagwell, "Resonant Tunneling in a Quasi-One-Dimensional Wire: Influence of Evanescent Modes," Phys. Rev. B **43**, 9012 (1991).

Arvind Kumar and Philip F. Bagwell, "Evolution of the Quantized Ballistic Conductance with Increasing Disorder in Narrow-Wire Arrays," Phys. Rev. B **44**, 1747 (1991).

The interested reader is also referred to

Arvind Kumar and Philip F. Bagwell, "Resonant Tunneling in a Multi-Channel Wire," Solid State Comm. **75**, 949 (1990).

Philip F. Bagwell, Terry P. Orlando, and Arvind Kumar, "Low Dimensional Resonant Tunneling," in *Resonant Tunneling in Semiconductors*, ed. by L.L. Chang et al., Plenum Press, New York, 417 (1991).

Philip F. Bagwell, Arvind Kumar, and Roger Lake, "Scattering and Quantum Localization of Electrons in a Waveguide by Static and Time-Varying Potentials," in *Quantum Effect Physics, Electronics, and Applications*, ed. by K. Ismail, Adam Hilger, Bristol (1992).

PHYSICAL REVIEW B VOLUME 43, NUMBER 11 15 APRIL 1991-I

Resonant tunneling in a quasi-one-dimensional wire: Influence of evanescent modes

Arvind Kumar and Philip F. Bagwell
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 30 July 1990; revised manuscript received 30 November 1990)

We calculate the transmission coefficients through two point scatterers in a quasi-one-dimensional wire as the Fermi energy and distance between scatterers are varied. At a subband minimum the standard wave-interference pattern between propagating modes, characteristic of a Fabry-Perot interferometer, is completely suppressed. Instead, we find that the shape of the lowest evanescent waveguide mode determines the electron transmission so that a different pattern of conductance oscillations emerges. If either of the scatterers is attractive, the transmission is suppressed abruptly near the "quasidonor" levels formed in the scatterer below each subband minimum.

I. INTRODUCTION

Discovery of the quantized constriction resistance,[1,2] and its explanation in terms of the Landauer conductance formula,[3,4] has greatly stimulated interest in understanding scattering in quasi-one-dimensional conductors.[5–21] After first understanding the simplest case of scattering from a single barrier in a quasi-one-dimensional geometry,[5–7] one would then like to understand the scattering from two barriers,[13–16] and eventually the transport through many barriers.[18–21]

In this paper we attempt to qualitatively understand scattering from two barriers in a confined geometry. We use a point-defect model for the scatterers, which may be too idealized to describe present experiments on transport in quantized GaAs constrictions. The impurity potential in GaAs heterojunctions is believed to vary slowly (on the scale of 0.1 μm) compared with the electron wavelength. Nonetheless, the point-defect model is useful to obtain qualitative insights into electron scattering in low-dimensional geometries. The limit of slowly varying, or adiabatic, potentials has already been described in detail,[22,23] and so we choose to examine a highly *nonadiabatic* potential where interchannel scattering is significant.

If a current is flowing in a confined geometry, such as the quasi-one-dimensional wire illustrated in Fig. 1, the incident electrons can scatter into evanescent modes which accumulate locally around the scattering centers. When more than one barrier is present in a multimode wire, an electron scattered into the evanescent channel at one of the obstacles can again be scattered at another obstacle. This effect most strongly influences the transmission properties when the distance between the two barriers is less than decay length of the evanescent mode, so that the tail of the evanescent wave function accumulating around one of the scatterers will overlap with the other scatterer. If the Fermi energy is away from a subband minimum or quasibound state in a wire with two scatterers, we find that this effect results in only small deviations from one-dimensional resonant tunneling theory.[24–26]

However, as the Fermi energy approaches a subband minimum, where the decay length of the evanescent mode and the evanescent density of states both become infinite, the transmission properties are determined largely by the shape of the lowest evanescent mode. At a subband minimum the Fabry-Perot interferences between propagating modes are completely suppressed, and a new series of transmission resonances with completely different properties emerges due to the strong coupling to the evanescent mode. We use this result to obtain the most general arrangement of point scatterers which give perfect transmission at a subband minimum, extending the calculations of Refs. 5 and 6.

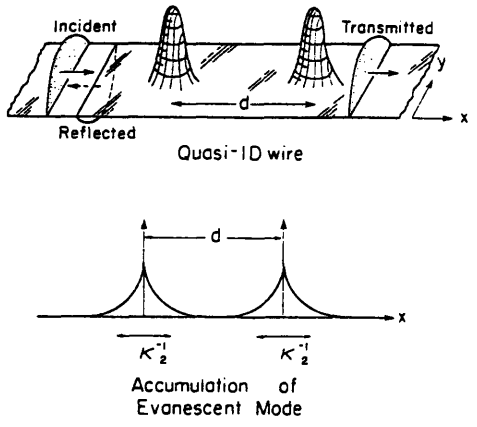


FIG. 1. An electron, incident from the left in the lowest normal mode, scattering from two barriers in a quasi-one-dimensional wire. Because of the confined geometry, evanescent modes accumulate locally around each tunnel barrier. If the barriers are separated by a distance $d \sim \kappa_2^{-1}$, the evanescent modes accumulating around each barrier begin to overlap, changing the conductance of the wire.

The Hamiltonian and transmission coefficients through the two-point scatterers are defined in Sec. II. In Sec. III we study the transmission coefficients when both scatterers in the wire are repulsive. Making one or both of the scatterers attractive introduces additional transmission properties, as investigated in Sec. IV. The Appendix describes how the transmission coefficients are calculated.[17]

II. SCATTERING IN A MULTIMODE WIRE

A quasi-one-dimensional wire, in which noninteracting electrons are free to move in the x direction but are confined in the y direction as shown in Fig. 1, is described by the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right]+V_c(y)+V_s(x,y)\right]\psi(x,y) = E\psi(x,y) . \quad (1)$$

The confinement potential $V_c(y)$ defines a basis set of normal modes $\{\chi_n(y)\}$ satisfying

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dy^2}+V_c(y)\right]\chi_n(y)=E_n\chi_n(y) , \quad (2)$$

where E_n is the subband energy for mode n. In the quasi-one-dimensional wire a transverse mode n is propagating if $E > E_n$ and is evanescent if $E \le E_n$, as shown in Fig. 2. The scattering potential $V_s(x,y)$ we take to be two-point scatterers separated along the x direction by a distance d as

$$V_s(x,y)=\gamma^{(1)}\delta(x)\delta(y-y_s^{(1)})+\gamma^{(2)}\delta(x-d)\delta(y-y_s^{(2)}) . \quad (3)$$

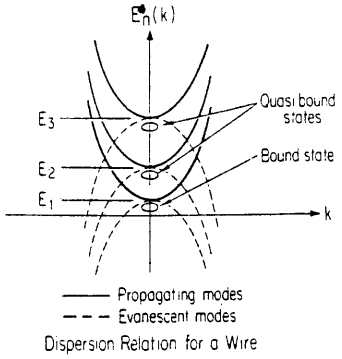


FIG. 2. Dispersion relation $E=E_n+\hbar k_n^2/2m$ for motion in mode n of a quasi-one-dimensional wire. The solid lines show the usual dispersion relation for the propagating modes ($E > E_n$). The confined geometry induces evanescent modes ($E \le E_n$), where $k_n=i\kappa_n$, having a dispersion relation given by the dashed lines. In addition, attractive scatterers give rise to bound or quasibound states associated with each subband formed from the evanescent-mode wave functions.

Since the wave function of the clean wire is separable, it can be expanded away from the scattering potential as follows:[17]

$$\psi(x,y)=\sum_n \alpha_n(x)\chi_n(y) , \quad (4)$$

where

$$\alpha_n(x)=\begin{cases} A_n e^{ik_n x}+B_n e^{-ik_n x}, & x<0 , \\ F_n e^{ik_n x}+G_n e^{-ik_n x}, & 0<x<d , \\ C_n e^{ik_n(x-d)}+D_n e^{-ik_n(x-d)}, & x>d . \end{cases} \quad (5)$$

Here the wave vector k_n is

$$k_n=\left[\frac{2m}{\hbar^2}(E-E_n)\right]^{1/2} , \quad (6)$$

and becomes imaginary if the mode n is evanescent so that $k_n=i\kappa_n$ where $\kappa_n>0$. The resulting dispersion relation for the wire modes is shown in Fig. 2. The boundary conditions require $A_n=D_n=0$ for evanescent modes so that there are no growing exponentials away from the scatterers. We take the particles to be incident from the left, so that $D_n=0$ in Eq. (5).

We define wave-function transmission and reflection amplitudes normalized to the amplitude of the incident mode A_m,

$$t_{nm}=\frac{C_n}{A_m}, \quad r_{nm}=\frac{B_n}{A_m} . \quad (7)$$

The boundary conditions for obtaining each t_{nm} and r_{nm} are given in the Appendix. Most of the following numerical results we present for t_{nm} and r_{nm} can be verified analytically (when the lowest two modes are present) by taking the appropriate limits of Eqs. (A6) and (A7) in the Appendix.

For propagating modes m and n the current transmission and reflection coefficients are given by

$$T_{nm}=\frac{k_n}{k_m}t_{nm}^* t_{nm}, \quad R_{nm}=\frac{k_n}{k_m}r_{nm}^* r_{nm} . \quad (8)$$

The normalized two-probe conductance g at zero temperature is then obtained from the multichannel Landauer formula[27–30]

$$g=\frac{G}{2e^2/h}=\sum_{n,m}T_{nm} , \quad (9)$$

where the sum in Eq. (9) runs over only the propagating modes of the wire. The remainder of the paper is concerned with analyzing the transmission coefficients T_{nm} through the scattering potential in Eq. (3). For definiteness throughout the remainder of this paper, we choose an infinite-square-well confining potential $V_c(y)$ having width $W=30$ nm and an electron mass $m=0.067m_e$ appropriate for GaAs heterojunctions.

III. TWO REPULSIVE SCATTERERS

We show in Fig. 3 the transmission coefficients and two-probe conductance as a function of energy E for $d=\lambda_1=34.64$ nm, where the wavelength λ_1 is evaluated at $E=E_2$. The scatterers are aligned at $y_s^{(1)}=y_s^{(2)}=13.3$ nm. We choose barriers having different strengths $\gamma^{(1)}=10$ feV cm^2 and $\gamma^{(2)}=5\gamma^{(1)}$ in Figs. 3 and 4. The lowest three modes have been kept for all numerical calculations in this section.

The two transmission resonances in Fig. 3 for $E_1<E<E_2$ arise from single-mode Fabry-Perot interference, and are less than unity because the barriers have unequal strengths. At the subband bottom ($E=E_2$), where the scattering is dominated by the evanescent mode, we observe that T_{11} is unity even though the Fermi energy is not near an expected Fabry-Perot resonance. When the Fermi energy rises into the second subband so that $E_2<E<E_3$, there is no remnant of Fabry-Perot-type resonances. The presence of two incident-electron wavelengths, combined with the intermode scattering, makes the Fabry-Perot resonance more difficult to achieve.

To obtain a better understanding of the transmission properties at a subband minimum in Fig. 3, we fix the probing-electron wavelength near the second subband minimum and vary the scatterer separation. In Figs. 4(a)–4(c) we study the variation of the transmission coefficient T_{11} versus d/λ_1 for three values of the energy: (a) $E=0.9E_2$, (b) $E=E_2$, and (c) $E=1.1E_2$. Note that the wavelength $\lambda_1=2\pi/k_1$ is different for each energy.

In Fig. 4(a) we observe the usual one-dimensional resonant tunneling behavior, subject to only small modifications. The transmission resonances are separated by $\lambda_1/2$ and the first maximum is offset from $d=0$ because of the phase shifts from each scatterer. Additionally, we note that T_{11} is no longer strictly periodic as a function of d, as can be seen from the first two minima. The scatterer separation for these two minima is in the range $d\sim 1/\kappa_2$ so that the evanescent tails of the second mode accumulating around each scatterer overlap significantly. The transmission coefficient is in general not a periodic function of scatterer separation d when evanescent modes are present.

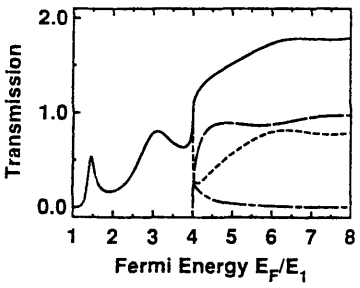


FIG. 3. Transmission coefficients T_{11} (short dashes), T_{22} (long dashes), $T_{12}=T_{21}$ (mixed dashes), and normalized conductance g (solid) through two-point barriers in a narrow wire. The two Fabry-Perot transmission resonances evident when $E<E_2$ are less than unity because the barriers have different scattering strengths, while the transmission is unity when $E=E_2$ despite the presence of scattering barriers.

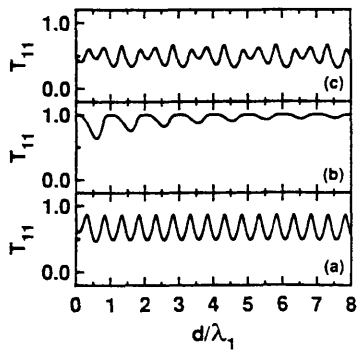


FIG. 4. Transmission coefficient T_{11} vs scatterer separation d/λ_1 for incident electron energies (a) $E=0.9E_2$, (b) $E=E_2$, and (c) $E=1.1E_2$. Spacing between transmission maxima changes from $\lambda_1/2$ to λ_1 as the Fermi energy aligns with the second subband in (b), so that the Fabry-Perot resonances in (a) are completely suppressed and a qualitatively different series of transmission resonances related to the evanescent modes emerges.

The variation of the transmission coefficient in Figs. 4(b) and 4(c) is markedly different from the standard one-dimensional resonant-tunneling result. In case (b), when the Fermi energy is aligned with the second subband, we see that successive transmission maxima are now separated by λ_1 and that the transmission maxima are unity even though the barriers have different strengths. The first maximum occurs at $d=0$ so that there is no scattering phase shift. Also, the minima of T_{11} tend asymptotically to unity as d is increased, so that the transmission is perfect when the scatterers have a large separation. The transmission coefficient T_{11} versus d/λ_1 in Fig. 4(b) is therefore highly aperiodic. These large qualitative differences between Figs. 4(a) and 4(b) demonstrate that the *Fabry-Perot wave interference between propagating modes is not the mechanism giving rise to conductance oscillations* when the Fermi energy is aligned with a subband minimum. In case (c) T_{11} has a beatlike pattern, due to the presence of two propagating wavelengths, and is no longer perfect anywhere.

Examining the shape of the evanescent mode around the scatterers will enable us to understand the change in transmission properties as the new subband becomes occupied. Figure 5(a) shows the lowest-evanescent-mode wave functions $\alpha_2(x)$ when the scatterers are separated by a half wavelength $d=\lambda_1/2$, while Fig. 5(b) shows the lowest-evanescent-mode wave function at a full wavelength separation $d=\lambda_1$. The important insight to gain from Fig. 5 is that, if the lowest evanescent mode is uniform along the length of the wire, perfect transmission re-

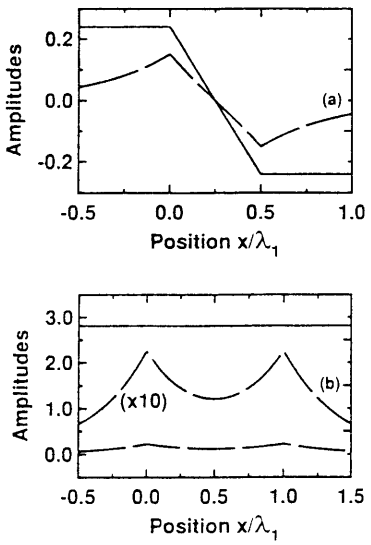


FIG. 5. Wave-function amplitudes of the second mode $[\alpha_2(x)]$ when $E=0.9E_2$ (dashed) and $E=E_2$ (solid) for scatterer separations (a) $d=\lambda_1/2$ and (b) $d=\lambda_1$. When the scatterers are separated by a half wavelength in (a), an "antibonding" state is formed. As the Fermi energy aligns with the second subband at $E=E_2$ the antibonding wave function has a discontinuous change in its derivative, leading to a reflected wave. Conversely, if the scatterers are separated by a full wavelength in (b), the corresponding "bonding" wave function is uniform along the length of the wire, forcing perfect transmission of the incident-mode wave function at $E=E_2$.

sults for the incident propagating wave function. The position of the scattering barriers and their coupling to the incident mode determine how the incident electrons scatter into the evanescent mode, the subsequent shape of the lowest-evanescent-mode wave function, and the resulting transmission properties. We have chosen the scatterer strengths to be equal in Fig. 5, $\gamma^{(1)}=\gamma^{(2)}=10$ feV cm^2, to illustrate the analogy between these evanescent states and "bonding" or "antibonding" orbitals in molecular physics.[31] The evanescent-mode wave function $\alpha_2(x)$ for the unequal barrier strengths studied in Figs. 3 and 4 is slightly more complex.

We can now given an intuitive argument to explain the electron transmission in Fig. 4(b). Because the evanescent-mode density of states diverges like κ_2^{-1} as the Fermi energy approaches the second subband minimum, we expect the evanescent mode to dominate the scattering at that energy. Suppose we place a point-defect barrier in the wire, which is known to give perfect transmission at a subband minimum,[5,6] and try to add a second barrier so that the transmission is still perfect at the second subband minimum. In general, the addition of the second scatterer will affect the evanescent-mode wave function at both scatterers and there will be some reflection in the incident mode. Therefore, to obtain perfect transmission, we must place the second scatterer in the wire in a way that does not change the evanescent-mode wave function at either scatterer.

For a single-point scatterer located at coordinates (x_s,y_s), the evanescent mode accumulates to a constant independent of x given by[5]

$$\alpha_2(x)=C_2=t_{21}e^{ik_1x_s}A_1\ , \tag{10}$$

where t_{21} is the transmission amplitude into the lowest evanescent mode when $x_s=0$,

$$t_{21}=r_{21}=-\Gamma_{12}/\Gamma_{22}=-\sin(\pi y_s/W)/\sin(2\pi y_s/W)\ . \tag{11}$$

Therefore, from Eqs. (10) and (11), the same value of the evanescent mode $\alpha_2(x)=C_2$ is obtained for different choices of the single-scatterer coordinates (x_s,y_s) only when

$$e^{ik_1x_s}\sin(\pi y_s/W)/\sin(2\pi y_s/W)=\text{const}\ . \tag{12}$$

Note that Eqs. (11) and (12) are independent of the scatterer strength γ.

Consider now the case of two-point scatterers in the wire when $E=E_2$. Fixing the position of the first scatterer at coordinates $(0,y_s^{(1)})$ determines the value of the constant in Eq. (12). Then from Eq. (12), a second scatterer aligned with the first can be placed at coordinates $(d,y_s^{(2)}=y_s^{(1)})$ and not affect the evanescent-mode wave function at either scatterer only when $d=j\lambda_1$, where $j=0,1,2,\ldots$. If the second scatterer is not an integer number of wavelengths from the first scatterer, the evanescent mode $\alpha_2(x)$ must change its value from one scatterer to the next. As explained in the Appendix, the resulting derivative jump in the evanescent mode must be met by a derivative jump in the first mode, giving rise to a reflected wave.

Therefore, perfect transmission is obtained only when the second barrier is placed an integer number of wavelengths from the first barrier, explaining the change in spacing between successive maxima of the transmission coefficient T_{11} from a half wavelength in Fig. 4(a) to a full wavelength in Fig. 4(b). Since Eq. (12) is independent of the scatterer strength, the barriers can have different strengths and perfect transmission still results at $E=E_2$. The approach of the transmission coefficient T_{11} to unity in Fig. 4(b) is also easily explained by noting that, if the positions of the scatterers force the evanescent-mode wave function to assume different values at each scatterer, the resulting discontinuity in the derivative of $\alpha_2(x)$ will become progressively smaller as the scatterer separation is increased.

If the two scatterers are disaligned, then in general it follows from Eq. (12) that the same value of the evanescent mode cannot be supported at both scatterers, and perfect transmission can no longer result at $E=E_2$.[32] We confirm this in Fig. 6 where T_{11} is shown as a function of d/λ_1 for three different y positions of the second

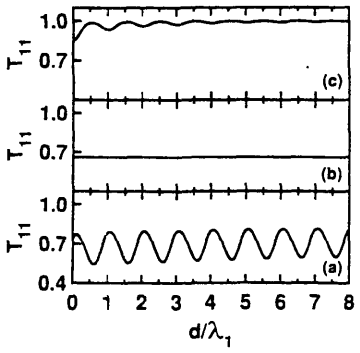


FIG. 6. Transmission coefficient T_{11} at an energy $E=E_2$ when the barriers are disaligned. In (c) we show the general transmission behavior for most lateral positions of the scatterers: the transmission resonances are hardly distinguishable from unity. Only when one of the barriers is placed (a) near or (b) on the node of the lateral wave function $\chi_2(y)$, are the resonances severely degraded. Note in (b) that, even though both barriers are still coupled to the incident propagating mode, there are no Fabry-Perot-type resonances.

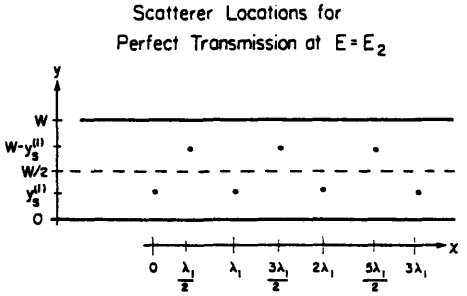


FIG. 7. Point scatterers of arbitrary signs and strengths can be placed at any combination of these positions in the wire, and perfect transmission $T_{11}=1$ still results when the Fermi energy is aligned with the second subband minimum at $E=E_2$. Any of the scatterers can be taken to have zero strength so that the scatterers do not necessarily form a periodic array.

scatterer: (a) $y_s^{(2)}=14.8$ nm, (b) $y_s^{(2)}=15.0$ nm, and (c) $y_s^{(2)}=23.7$ nm. We fix the energy at $E=E_2$ and the position of the first scatterer at $y_s^{(1)}=13.3$ nm as before. The barriers have the same strengths as in Figs. 3 and 4.

In Fig. 6(c) the transmission resonances are not exactly perfect, but the deviation from unity is not noticeable on the graph. We find this behavior of T_{11} generally holds for most lateral positions $y_s^{(2)}$. Only if the second scatterer is brought close to the node of $\chi_2(y)$, as in Figs. 6(a) and 6(b), does a different transmission behavior begin to emerge. When we place one of the scatterers directly on the node in $\chi_2(y)$, so that it does not couple to the lowest-evanescent-mode wave function, the transmission coefficient in Fig. 6(b) becomes essentially independent of separation d, *even though the barrier is still coupled to the incident mode.* The absence of transmission oscillations in Fig. 6(b) is easily understood, since the barrier in the node of $\chi_2(y)$ does not alter the shape of the evanescent modes needed to give perfect transmission[5,6] through the other barrier. Therefore, the transmission coefficient is independent of the separation d and limited only by the barrier in the node of $\chi_2(y)$. If one of the scatterers is placed very near but not exactly on the node in $\chi_2(y)$, as in Fig. 6(a), the resulting transmission resonances of period λ_1 are noticeably less than unity. However, the transmission still asymptotically approaches unity after many wavelengths, as in our previous analysis.

There is a special case of disaligning the two scatterers where perfect transmission still results at $E=E_2$. If $y_s^{(2)}=W-y_s^{(1)}$, so that the scatterer positions along y are mirrored about the axis of the wire, the mode-coupling constants between the incident mode and the lowest evanescent mode at the two scatterers are negatives of each other ($\Gamma_{21}^{(1)}=-\Gamma_{21}^{(2)}$). Then, if we displace the second scatterer by a half wavelength from the first along the x direction, the incident-mode wave function $e^{ik_1x}\sin(\pi y/W)$ changes sign at the second scatterer. Therefore, by Eq. (10), the same constant value of $\alpha_2(x)=C_2$ again satisfies the boundary conditions at both scatterers. By the arguments in the preceding paragraphs, perfect transmission is again obtained at $E=E_2$ if $y_s^{(2)}=W-y_s^{(1)}$ and $d=(j+1/2)\lambda_1$ for any sign or strength of the second scatterer.

To conclude this section, we use the argument from the preceding paragraph to consider adding more than two-point scatterers to the wire. By induction, we can continue to place additional scatterers having arbitrary signs and strengths on any subset of the grid points shown in Fig. 7 and still obtain perfect transmission when $E=E_2$. The perfect transmission through a single δ-function scatterer at a subband minimum[5,6] is a special case where only one site in the grid is occupied. Figure 7 represents the most general arrangement of δ function scatterers yielding perfect transmission $T_{11}=1$ when $E=E_2$.

IV. ATTRACTIVE SCATTERERS

If attractive scatterers are present in a multimode wire, quasibound or bound states made up of evanescent waves form below each subband minimum, as shown in Fig. 2. The quasibound state is localized around the attractive scatterer like a donor level below the conduction-band minimum of a semiconductor. Such a state is truly bound only if its energy lies below the bottom of the lowest subband; otherwise its energy is degenerate with that of a propagating mode and a particle in such a state has a finite lifetime. If one or both of the point scatterers in our two-barrier problem are made attractive, we expect zeros in the transmission coefficient in the vicinity of the quasibound states, similar to those in previous calculations.[5-7,19-21] We find that the qualitative behavior of the transmission coefficients is similar to that found for

repulsive scatterers, except when the Fermi energy is near a quasibound state.

In Fig. 8 we plot the zeros of the transmission coefficient T_{11} in the d-E plane, with the scatterer separation d normalized to the wavelength λ_1 at each energy E. The lowest two modes are retained in Figs. 8 and 9. Figure 8(a) shows one attractive scatterer and one repulsive scatterer, while Fig. 8(b) shows two attractive scatterers. The calculation in Fig. 8(b) can be viewed as transmission through a diatomic molecule attached to the wire. The figure insets show the position of the bound-state energies calculated by setting the coupling of the incident propagating mode to the evanescent mode to zero, mimicking a simple one-dimensional calculation of molecular binding energies. The scatterers are aligned at $y_i^{(1)}=y_i^{(2)}=13.3$ nm and have equal strength $|\gamma^{(1)}|=|\gamma^{(2)}|=25$ feV cm^2. The energy near the top of the figure where the zeros terminate abruptly is E_2, the bottom of the second subband.

Figure 8(a) displays a single bound state in the attractive scatterer for any value of the separation $d\neq 0$. As in elementary molecular physics, Fig. 8(b) displays two bound states when d is large, and only one bound state when d becomes small. When d becomes small in Fig. 8(b), the antibonding state is forced into the continuum above the second subband.

The zeros of T_{11} in Figs. 8(a) and 8(b) qualitatively follow the bound-state energies of the single-mode problem shown in the inset. These zeros in T_{11} exhibit some oscillatory structure of periodicity λ_1, and coincide with the bound-state energies shown in the inset when $d=j\lambda_1$. The oscillatory structure in T_{11} arises from the coupling of the propagating mode to the quasibound states, so that the zeros in T_{11} are sensitive to the relative phase of the incident mode at the two scatterers. We do not understand the precise relationship between the quasibound states of a multichannel system[33] and zeros in the transmission coefficient, though the two are clearly related. When both scatterers are made attractive we also observe, in Fig. 8(b), gaps of d for which there are no transmission zeros.

We plot the condutance versus Fermi energy in Fig. 9 when (a) both scatterers are repulsive, (b) the first scatterer is repulsive and the second is attractive, and (c) both scatterers are attractive. The strength and lateral positions of the scatterers are the same as in Fig. 8, and we choose the separation $d=90$ nm (or $d=2.6\lambda_1$ at $E=E_2$). There is one transmission zero immediately before the second subband opening in Fig. 9(b), two zeros in Fig. 9(c), and no transmission zeros in Fig. 9(a), as expected from Figs. 8(a) and 8(b).[34]

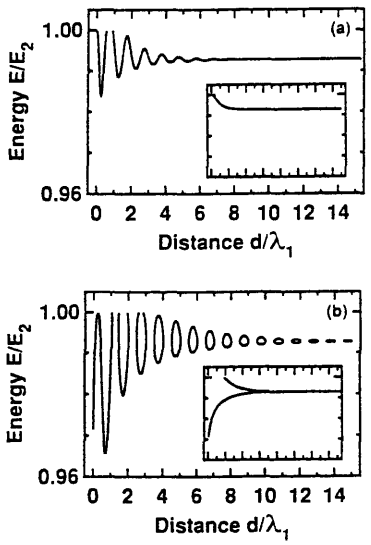


FIG. 8. Zeros in the transmission coefficient $T_{11}=0$ in the d-E plane for (a) one attractive and one repulsive scatterer and (b) two attractive scatterers. Zeros in T_{11} qualitatively follow the bound-state energies calculated in the insets. Transmission oscillations of period λ_1 indicate an interplay between the relative phase of the incident mode at each scatterer and the quasibound states.

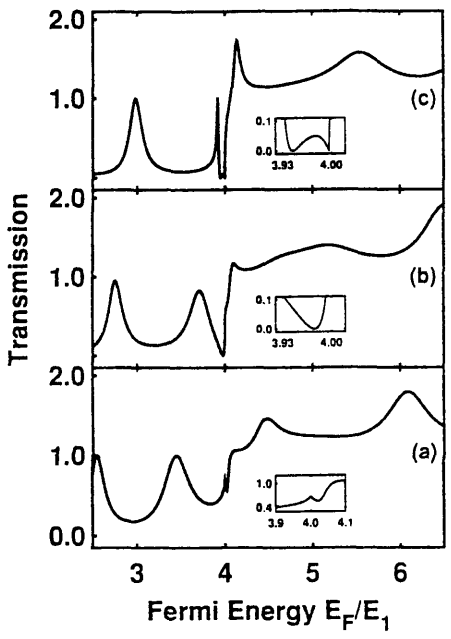


FIG. 9. Normalized conductance g vs Fermi energy for (a) two repulsive scatterers, (b) one attractive and one repulsive scatterer, and (c) two attractive scatterers. The insets show more detailed behavior near the second subband minimum. Conductance dips below the second subband minimum in (b) and (c) are associated with "molecular" bound states.

V. CONCLUSION

We have calculated the transmission coefficients for an electron scattering from two-point barriers in a quasi-one-dimensional wire. Our main conclusion is that, if the Fermi energy is near a subband minimum in a quasi-one-dimensional wire, the scattering properties of the wire are determined primarily by the shape of the lowest evanescent mode around the scattering centers. Wave interference between propagating modes in the wire, which normally produces an interference pattern of oscillations in the conductance, is no longer the dominant mechanism giving rise to structure in the conductance when the Fermi energy is near a subband minimum.

All of the unusual scattering properties we find can occur only for electrical conduction in a low-dimensional structure, where electrons accumulate in evanescent or cutoff waveguide modes around each scattering center. For example, in a one-dimensional scattering problem, perfect transmission through a single barrier implies perfect transmission through two such barriers. Contrary to our original expectations, when the Fermi energy coincides with a subband minimum, the multimode transmission through two-point barriers is perfect only for certain resonant values of the scatterer separation. Even though each single barrier in the multimode wire, taken by itself, appears completely transparent, an incident electron can be reflected from two such perfectly transmitting barriers in series.

Although our model-scattering potential is highly idealized, and probably not directly applicable to present experiments on GaAs constrictions, we expect the same general transmission properties to hold for transport in real low-dimensional systems. In particular, the recurrence of transmission resonances every wavelength of the incident electron should be insensitive to the exact choice of the scattering potential. If electron charging effects[35] or inelastic scattering between the barriers becomes dominant, the transmission properties studied in this paper may be suppressed. Similar transmission properties would result from two dielectric posts placed in a microwave waveguide.

ACKNOWLEDGMENTS

We thank Terry P. Orlando, both for contributing his ideas and for his suggested improvements of this manuscript. We are also indebted to Ray Ghanbari, Boris Altshuler, Peter Hagelstein, Marc Cahay, Marc Kastner, Dimitri A. Antoniadis, and Henry I. Smith for helpful discussions. This work was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-88-0304. One of us (A.K.) gratefully acknowledges partial support from the Semiconductor Research Corporation.

APPENDIX: SCATTERING BOUNDARY CONDITIONS

In this appendix we give the method used to obtain the transmission and reflection amplitudes following Ref. 17. Numerical calculations in the body of the paper were performed by cascading scattering matrices (because they proved to be numerically stable near a subband minimum), while in this appendix we give an analytical result for t_{11} obtained by multiplying transfer matrices.

The wave function in Eq. (4) must be continuous. After multiplying Eq. (4) by $\chi_m(y)$ and integrating over all y, the orthogonality property of the $\{\chi_n(y)\}$ demands that each individual mode be continuous so that $\alpha_n(0^-)=\alpha_n(0^-)$ and $\alpha_n(d^-)=\alpha_n(d^-)$. Therefore, at $x=0$,

$$A_n+B_n=F_n+G_n \tag{A1}$$

holds for all n.

The boundary condition on the derivatives of the $\{\alpha_n(x)\}$ follows similarly by inserting Eq. (4) into Eq. (1). The orthogonality property of the $\{\chi_n(y)\}$ can then be used to show that

$$\left\{\frac{d^2}{dx^2}+k_n^2\right\}\alpha_n(x)=\sum_n[\Gamma_{nm}^{1,i}\delta(x)+\Gamma_{nm}^{2,i}\delta(x-d)]\alpha_m(x)\ . \tag{A2}$$

Here the mode-coupling constants for barrier (i) are given by

$$\Gamma_{nm}^{i,i}=\frac{2m\gamma^{(i)}}{\hbar^2}\chi_n(y_s^{(i)})\chi_m(y_s^{(i)})\ ,\quad i=1,2\ . \tag{A3}$$

At $x=0$ the discontinuity in the first derivative of each mode, by integrating Eq. (A2), must satisfy

$$\begin{aligned}\frac{d\alpha_n(x)}{dx}\Big|_{0^-}-\frac{d\alpha_n(x)}{dx}\Big|_{0^-}&=\sum_m\Gamma_{nm}^{1,i}\alpha_m(0)\\&=\chi_n(y_s^{(1)})\left\{\frac{2m\gamma^{(1)}}{\hbar^2}\sum_m\chi_m(y_s^{(1)})\alpha_m(0)\right\}\ .\end{aligned} \tag{A4}$$

Since Γ_{nm} from Eq. (A3) factors into a product form, the right-hand side of Eq. (A4) is simply $\chi_n(y_s^{(1)})$ times a constant independent of n. If the right-hand side of Eq. (A4) is not zero so that mode n has a derivative jump, Eq. (A4) implies that every other mode with a nonzero wave-function amplitude $\chi_n(y_s^{(1)})$ at the scatterer must also have a proportional derivative jump.

When the Fermi energy coincides with a subband minimum for the case of a single δ-function scatterer, wave-function continuity from Eq. (A1) requires that the lowest-evanescent-mode wave function $\alpha(x)$ must be constant everywhere in space. The term in large parentheses in Eq. (A4) must then be zero if the lowest evanescent mode is populated. Therefore, the incident mode must have a continuous first derivative at the scatterer because the evanescent mode has a continuous first derivative. Since for a point barrier there can be no reflected wave without a derivative jump in the incident mode, it follows that the transmission must be perfect at each subband minimum when only a single-point barrier is present.

Inserting Eq. (5) in Eq. (A4) yields

$$ik_n(F_n-G_n)-ik_n(A_n-B_n)=\sum_m \Gamma^{(1)}_{nm}(A_m+B_m)\ . \tag{A5}$$

Equation (A5) couples mode n to all other modes in the problem. For simplicity, we cut off the sum in Eq. (A5) at a finite number of modes. Likewise, equations analogous to Eqs. (A1) and (A5) enforcing the boundary conditions at $x=d$ can also be written down.

For the case of two modes, multiplying the individual transfer matrices by hand yields t' and r' of Ref. 17 directly. We then interchange $\Gamma^{(1)}$ and $\Gamma^{(2)}$ to obtain t_{11} as

$$t_{11}=\frac{1}{\mathcal{D}_2}\left[\frac{i}{2k_2}\sum_{n=1}^{2}\frac{\Gamma^{(1)}_{n2}\Gamma^{(2)}_{2n}}{k_n}\sin(k_nd)+\left[1+\frac{i(\Gamma^{(1)}_{22}+\Gamma^{(2)}_{22})}{2k_2}\right]e^{-ik_2d}\right], \tag{A6}$$

$$\begin{aligned}\mathcal{D}_2=&\frac{ie^{-ik_2d}}{2k_1}\sum_{n=1}^{2}\frac{\Gamma^{(1)}_{n1}\Gamma^{(2)}_{1n}}{k_n}\sin(k_nd)+\frac{ie^{-ik_1d}}{2k_2}\sum_{n=1}^{2}\frac{\Gamma^{(1)}_{n2}\Gamma^{(2)}_{2n}}{k_n}\sin(k_nd)\\&+\left[1+\frac{i(\Gamma^{(1)}_{11}+\Gamma^{(2)}_{11})}{2k_1}+\frac{i(\Gamma^{(1)}_{22}+\Gamma^{(2)}_{22})}{2k_2}-\frac{\Gamma^{(1)}_{11}\Gamma^{(2)}_{22}+\Gamma^{(1)}_{22}\Gamma^{(2)}_{11}}{4k_1k_2}\right]e^{-i(k_1+k_2)d}+\frac{\Gamma^{(1)}_{12}\Gamma^{(2)}_{21}}{4k_1k_2}e^{-i2k_1d}+\frac{\Gamma^{(1)}_{21}\Gamma^{(2)}_{12}}{4k_1k_2}e^{-i2k_2d}\ .\end{aligned} \tag{A7}$$

Most of the results of this paper can be obtained by taking appropriate limits of Eqs. (A6) and (A7). To obtain the curves in Fig. 8 we have set $t_{11}=0$ from Eq. (A6). The resulting expression is

$$\frac{1}{2}\left[\frac{\Gamma^{(1)}_{12}\Gamma^{(2)}_{21}}{k_1\kappa_2}\sin(k_1d)+\frac{\Gamma^{(1)}_{22}\Gamma^{(2)}_{22}}{\kappa_2^2}\sinh(\kappa_2d)\right]+\left[1+\frac{\Gamma^{(1)}_{22}+\Gamma^{(2)}_{22}}{2\kappa_2}\right]e^{\kappa_2d}=0\ , \tag{A8}$$

where we have written $\kappa_2=-ik_2$ since the second mode is evanescent in the energy range over which quasibound states occur. The graphs in the inset of Fig. 8 are obtained by setting the intermode coupling to zero, namely $\Gamma^{(1)}_{12}=\Gamma^{(2)}_{21}=0$ in Eq. (A8).

One can also obtain t_{21} and r_{21} by the same procedure used to obtain t_{11} of Eq. (A6). Taking appropriate limits of these expressions for t_{21} and r_{21} recovers the results shown in Fig. 5, thus verifying Fig. 5 analytically when two modes are present.

[1]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).

[2]D. A. Wharam, T. J. Thornton, R. Newbury, M. Pepper, H. Ahmed, J. E. F. Frost, D. G. Hasko, D. C. Peacock, D. A. Ritchie, and G. A. C. Jones, J. Phys. C **21**, L209 (1988).

[3]Y. Imry, in *Directions in Condensed Matter Physics*, edited by G. Grinstein and G. Mazenko (World Scientific, Singapore, 1986).

[4]R. Landauer, Z. Phys. B **68**, 217 (1987).

[5]P. F. Bagwell, Phys. Rev. B **41**, 10 354 (1990); P. F. Bagwell, J. Phys. Condens. Matter. **2**, 6179 (1990).

[6]C. S. Chu and R. S. Sorbello, Phys. Rev. B **40**, 5941 (1989); C. S. Chu and R. S. Sorbello, *ibid.* **38**, 7260 (1988).

[7]E. Tekman and S. Ciraci, Phys. Rev. B **42**, 9098 (1990).

[8]F. M. Peeters, in *Science and Engineering of 1- and 0-Dimensional Semiconductors*, edited by S. Beaumont and C. Sotomayor-Torres (Plenum, New York, 1990).

[9]C. W. J. Beenakker and H. van Houten, Phys. Rev. Lett. **63**, 1857 (1989); in *Electronic Properties of Multilayers and Low-Dimensional Structures*, edited by J. M. Chamberlain, L. Eaves, and J. C. Portal (Plenum, New York, 1991).

[10]R. L. Schult, H. W. Wyld, and D. G. Ravenhall, Phys. Rev. B **41**, 12 760 (1990).

[11]H. U. Baranger, Phys. Rev. B **42**, 11 479 (1990).

[12]A. Szafer and A. D. Stone, Phys. Rev. Lett. **62**, 300 (1989).

[13]G. W. Bryant, Phys. Rev. B **39**, 3145 (1989).

[14]Y. Avishai, M. Kaveh, and Y. B. Band, Phys. Rev. B **42**, 5867 (1990).

[15]A. Kumar and P. F. Bagwell, Solid State Commun. **75**, 949 (1990).

[16]P. F. Bagwell, T. P. Orlando, and A. Kumar, in *Resonant Tunneling in Semiconductors: Physics and Applications*, edited by E. E. Mendez and L. L. Chang (Plenum, New York, in press).

[17]M. Cahay, M. McLennan, and S. Datta, Phys. Rev. B **37**, 10 125 (1988).

[18]M. Cahay, S. Bandyopadhyay, M. A. Osman, and H. L. Grubin, Surf. Sci. **228**, 301 (1990); see also S. Bandyopadhyay and M. Cahay, *Computational Electronics: Semiconductor Transport and Device Simulation*, edited by K. Hess, J. P. Leburton, and U. Ravaioli (Kluwer-Academic, Boston, 1990), p. 223.

[19]J. Masek, P. Lipavsky, and B. Kramer, J. Phys. Condensed Matter **1**, 6395 (1989).

[20]I. Kander, Y. Imry, and U. Sivan, Phys. Rev. B **41**, 12 941 (1990).

[21]A. Kumar and P. F. Bagwell (unpublished).

[22]L. I. Glazman, G. B. Lesovik, D. E. Khmel'nitskii, and R. I. Shekhter, Pis'ma Zh. Eksp. Teor. Fiz. **48**, 218 (1988) [JETP Lett. **48**, 238 (1988)].

[23]A. Yacoby and Y. Imry, Phys. Rev. B **41**, 5341 (1990).

[24]G. A. Toombs and F. W. Sheard, in *Electronic Properties of Multilayers and Low-Dimensional Structures*, edited by J. M.

Chamberlain, L. Eaves, and J. C. Portal (Plenum, New York, 1991).
[25]P. J. Price, Phys. Rev. B 38, 1994 (1988).
[26]P. F. Bagwell, T. P. E. Broekaert, T. P. Orlando, and C. G. Fonstad, J. Appl. Phys. 68, 4634 (1990).
[27]M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Phys. Rev. B 31, 6207 (1985).
[28]R. Landauer, J. Phys. Condens. Matter 1, 8099 (1989); in *Analogies in Optics and Mirco-Electronics*, edited by W. van Haerington and D. Lenstra (Kluwer-Academic, Dordrecht, 1990).
[29]M. Büttiker, IBM J. Res. Dev. 32, 317 (1988).
[30]S. Datta and M. McLennan, Rep. Prog. Phys. 53, 1003 (1990).
[31]It is interesting to note that, when the second mode is in an "antibonding state" at $E=E_2$, there is also a buildup of the third and higher modes in an "antibonding" configuration. This is true even though, for a single-point scatterer, there is zero probability amplitude in the third and higher modes when $E=E_2$. In contrast, there is no buildup of the third and higher modes when the second mode is in a "bonding" state at $E=E_2$. It is not immediately obvious how to explain these results as a sum over Feynman path histories.
[32]It is also interesting that, if we fix $d=0$, the reflection from two disaligned scatterers is more severe if the scatterers are on opposite sides of the node in $\chi_2(y)$ as expected from Eq. (11).
[33]A. I. Baz', Y. B. Zel'dovich, and A. M. Perelomov, *Scattering, Reactions, and Decay in Nonrelativistic Quantum Mechanics* (Israel Program for Scientific Translations, Jersualem, 1969).
[34]In Fig. 9(b) the Fabry-Perot resonances in the lowest subband are clearly not perfect, even though they are unity in the absence of mode coupling. This anomalous effect extends well below the second subband minimum in Fig. 9(b), although not present in Figs. 9(a) and 9(c). We also note a drop in transmission immediately after the second subband is populated in Fig. 9(a) for large d, even though no such drop is seen in Fig. 3 when d is small.
[35]U. Meriav, M. A. Kastner, and S. J. Wind, Phys. Rev. Lett. 65, 771 (1990).

PHYSICAL REVIEW B VOLUME 44, NUMBER 4 15 JULY 1991-II

Evolution of the quantized ballistic conductance with increasing disorder in narrow-wire arrays

Arvind Kumar and Philip F. Bagwell*
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 4 February 1991)

We study the two-terminal Landauer conductance averaged over a parallel array of disordered narrow wires as the Fermi energy and length of the disordered region are varied. As disorder in the wires is increased, so that quantum diffusion becomes the dominant electron-transport mechanism, we find numerically that the quantized conductance steps characteristic of ballistic transport evolve into conductance drops after each new subband is populated. Consistent with this result, the electron localization length decreases above each new subband. Adding attractive scatterers to the wires strongly modifies these results due to "quasidonor levels" forming in the impurities.

I. INTRODUCTION

Discovery of the quantized ballistic conductance through a point contact[1,2] has greatly stimulated theoretical studies on the effect of impurity scattering in nearly ballistic quantum wires.[3–17] When only a few impurities are present in a wire, these studies have predicted that the average conductance should rise after the opening of each new subband channel, although structure in the conductance of a single wire may be obscured by wave-interference fluctuations.[17,18] If some of the scatterers are attractive, pronounced conductance drops before the opening of each new subband were also found to occur,[3–9] due to the formation of "quasidonor levels" in the impurities.[4–7] Thus repulsive and attractive scatterers result in a very different subband structure for the conductance versus Fermi energy in a narrow wire.

If disorder in the wire is increased, so that the transport becomes diffusive rather than nearly ballistic, it might be expected that a fundamentally different subband structure should be observed experimentally[19–22] in narrow quantum wires. Indeed, based on a modified Drude model and assuming that electrons scatter in the Born approximation, Refs. 23–31 conclude that electron scattering increases whenever a new subband becomes occupied, leading to a drop in conductance versus Fermi energy *after* each new subband opening. In contrast, Refs. 3–9 have argued that pronounced drops in conductance should occur *before* each new subband opening. Some structure in the conductance versus electron density has possibly been observed in arrays of narrow wires,[19–21] but it is unclear what physics this structure might represent or where it occurs in relation to subband minima in the wires.

In this paper we calculate the two-terminal Landauer conductance[32–39] of a parallel array of disordered quantum wires to obtain the "ensemble-averaged" conductance. A "point-scatterer" model[16] is used to describe the disorder in each wire. We show that there is a clear transition in this model between a quantum ballistic regime, marked by increasing conductance after each new subband channel opens, and a quantum diffusive regime, marked by a sharp drop in conductance when a new subband is populated. We associate this conductance drop with a decrease in the electron localization length immediately above a subband, so that the drop in conductance after each subband opening in the diffusive regime depends on *quantum diffusion*, rather than the classical diffusion of the Drude model. Additionally, we show that the standard "golden-rule" or "Born-approximation" scattering theory is invalid near a subband minimum, so that the scattering at each individual impurity must be properly calculated to obtain the correct dependence of the conductance versus Fermi energy. When this is done, we do not find conductance drops after the opening of a new subband channel if the electron diffusion is classical. Finally, both for quantum and classical electron diffusion, we find conductance drops before the opening of a new subband channel only when attractive scatterers are present in the wires.

II. MODEL FOR A DISORDERED QUANTUM WIRE

We choose a model Hamiltonian describing electrons free to move along the x direction and confined along the y direction:

$$\left[-\frac{\hbar^2}{2m}\left[\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right]+V_c(y)+V_d(x,y)\right]\psi(x,y) = E\psi(x,y)\ . \quad (1)$$

The confinement potential $V_c(y)$ gives rise to confinement subbands E_n such that

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dy^2}+V_c(y)\right]\chi_n(y)=E_n\chi_n(y)\ . \quad (2)$$

We choose the impurity potential to be a sequence of point scatterers

$$V_d(x,y)=\sum_i \gamma_i\delta(x-x_i)\delta(y-y_i)\ , \quad (3)$$

where the ith scatterer is located at position (x_i,y_i) and has strength γ_i. The conductance is obtained from the

two-probe "Landauer formula"[32-39]

$$G=\frac{I}{V}=\frac{e^2}{\pi\hbar}\sum_{m,n}T_{mn}\ , \tag{4}$$

where T_{mn} denotes the transmission coefficient from mode n to mode m. Many individuals have contributed to our understanding of Eq. (4), as discussed in Refs. 32–39 and citations therein.

The transmission coefficients T_{mn} in Eq. (4) are found numerically by cascading together the individual scattering matrices for each point defect and each intermediate region of free propagation between defects.[16] We include the lowest five modes in our calculations, enough to understand the qualitative features of the conductance, although we expect the inclusion of higher modes to have quantitative influence on our results. This numerical model and our particular implementation of it are both very reliable, since they agree with analytical results obtained for electron transmission through both one[4,5] and two[7] point defects in a narrow wire. Current conservation is also numerically well satisfied in these simulations, giving additional confidence in the reliability of our results.

For a single wire with a disordered region of length L along the x direction, we randomly position the scatterers with a uniform probability density over the ranges $[0,W]$ across the channel and $[0,L]$ along the wire. We choose a fraction f of the scatterers to be attractive ($\gamma_i<0$). All the scatterers have equal strengths $|\gamma_i|=10$ feV cm^2. We choose the mean spacing between impurities along the x direction of the wire to be 10 nm, so that there are five impurities in the wire when $L=50$ nm and 50 impurities when $L=500$ nm. We model the confinement using an infinite square-well potential of width $W=30$ nm, and take the electron mass to be 0.067 times the free mass. This choice of parameters is consistent with experiments on GaAs/Al_xGa_{1-x}As heterostructures. In a more realistic model, where the wire consists of a potential well having finite depth, it will probably be necessary to incorporate evanescent modes from the continuum in order to obtain proper convergence of the calculation.

Impurity potential fluctuations in good-quality GaAs heterostructures are believed to be mainly weak and smooth.[40] Electron transport through such a smooth potential can be nearly "adiabatic," producing a nearly exact quantization of the ballistic conductance.[41-43] It has also been argued that such a smooth scattering potential should give rise to enhanced electron mobility in a narrow wire when electron transport is restricted to the lowest subband,[44] since any carrier deflections are primarily due to small-angle forward scattering, which does not substantially degrade the electron mobility. Roughness along the channel edges in GaAs heterostructures, a type of disorder neglected in this work, is also believed to vary smoothly compared to the electron wavelength.[45] Conversely, the potential in Eq. (3) is quite rough and irregular, and in fact produces s-wave scattering[3] when embedded in a two-dimensional plane. In most calculations using a tight-binding Anderson model the potential is also rough and irregular,[8,9] since the "on-site energies" are random.

However, fair conductance quantization is still produced for electron transmission through a point-scattering potential in a narrow channel.[3-5] The point-scatterer model of disorder may therefore be a reasonable one to understand the effects of wave interference between scattering events in a low-dimensional conductor,[7] even though it is a highly nonadiabatic scattering potential. The point-scatterer model also gives rise to "universal conductance fluctuations"[17] in agreement with experiment. In addition, electron transmission through a finite-size rectangular barrier in a wire is qualitatively similar to transmission through a point scatterer,[4] indicating that shrinking the scatterer to zero size may not be a serious limitation. Edge roughness has not been included explicitly, but some roughness is simulated because some of the scatterers lie near the channel edges. For the case of a single scatterer, the qualitative dependence of the electron transmission on Fermi energy for an edge defect is no different from that of a defect located in the middle of the channel.[4] Comparing our present results using the scattering potential in Eq. (3) with transmission through a smooth disorder potential is left to a future study. We caution that the scattering effects near a subband crossing studied in this paper may possibly be exaggerated or even qualitatively different when compared with transmission through a smoother impurity potential.

III. CALCULATED CONDUCTANCE OF A NARROW-WIRE ARRAY

In Fig. 1(a) we show the conductance as a function of Fermi energy for a single wire in the ensemble having $L=50$ nm for both $f=0.0$ and 0.5. The conductance is seen to rise after the opening of each new subband whether all the scatterers are repulsive ($f=0.0$) or half the scatterers are attractive ($f=0.5$). However, the introduction of attractive scatterers gives rise to pronounced dips in conductance below each subband minimum, near the energies of quasidonor levels splitting off from the confinement subbands. The spacing ΔE from these quasidonor levels to the next subband is of order $\Delta E\simeq(m/\hbar^2)(\gamma/W)^2$, close to the binding energy of a state trapped in the point defect.[4-7] If the length of the disordered region is increased to $L=500$ nm, as shown in Fig. 1(b) for $f=0.0$, the resulting electron wave-interference pattern obscures any regular structure in the conductance. When $f=0.5$ and $L=500$ nm as in Fig. 1(c), the conductance exhibits similar fluctuations that obscure any underlying subband structure. Some unusual properties of these fluctuations have been noted in Ref. 17. Interestingly, the lowest subband is still approximately discernible in Fig. 1(c) due to the depressed transmission associated with the quasibound states.

To manifest the underlying subband structure of the conductance versus Fermi energy, we plot in Fig. 2 the conductance from Fig. 1 averaged over an array of 100 independent wires in parallel. Each wire has a different random arrangement of the scatterers, but the length of

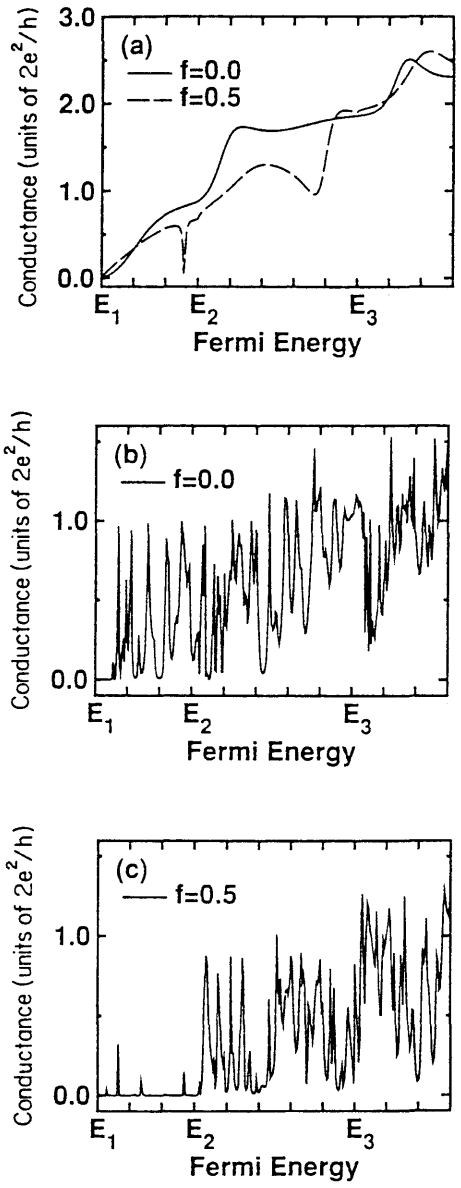


FIG. 1. Landauer conductance vs Fermi energy for a single quasi-one-dimensional wire having (a) $L=50$ nm and (b) and (c) $L=500$ nm. When only a few scatterers are present (a), the conductance varies smoothly with Fermi energy and, if some attractive scatterers are present ($f=0.5$), dips abruptly near the "quasidonor levels" below each subband. As more scatterers are added, wave-interference conductance fluctuations in (b) and (c) obscure the underlying regular structure due to confinement subbands.

the disordered region is kept fixed. In Fig. 2(a) each wire has $L=50$ nm. Some broad resonances present in the single wire are eliminated after averaging when all the scatterers are repulsive ($f=0.0$). When half the scatterers are made attractive ($f=0.5$), the quasibound-state energies vary from wire to wire,[7] resulting in a broadened dip in the average conductance before the opening of each new channel.

If we increase the length of the disordered region to $L=500$ nm, as in Fig. 2(b), a fundamentally different subband structure of the conductance versus Fermi energy emerges. The conductance now *drops* abruptly *after* the opening of each new subband channel when all the scatterers are repulsive ($f=0.0$, top curve). When half the scatterers are made attractive ($f=0.5$, bottom curve),

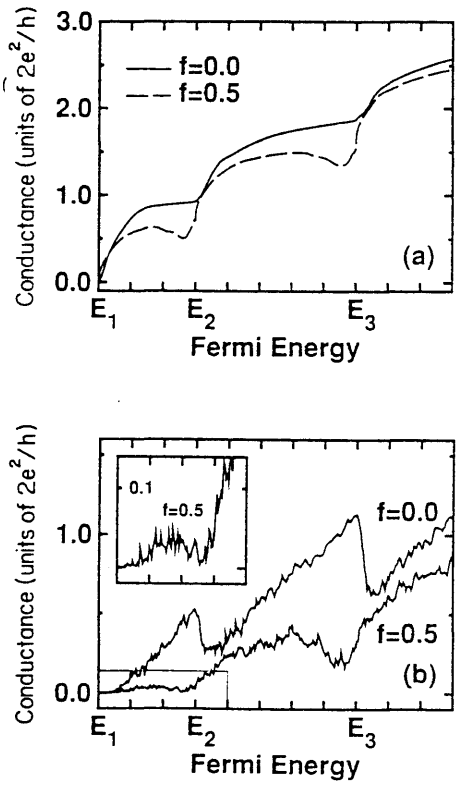


FIG. 2. Landauer conductance averaged over an array of 100 parallel wires, each having length (a) $L=50$ nm and (b) $L=500$ nm. The ballistic conductance steps are rounded for the short quantum wires in (a). A new "diffusive" subband structure emerges for the long quantum wires [(b), top curve] having $f=0.0$: The conductance falls after each new subband opens. Quasidonor states are still observed for either short [(a) bottom curve] or long [(b) bottom curve] quantum wires when $f=0.5$.

the quasidonor states still give rise to a broadened conductance dip before the new channel opens. The net effect of this broadened conductance dip for $f=0.5$ is that the average conductance is so suppressed before the opening of each new channel that conductance drops after the new channel opens are not observed. The inset of Fig. 2(b), an expanded view of the lightly boxed region, shows clearly the drop in conductance for $f=0.5$ before the second subband channel opens.

To better understand this transition from the nearly ballistic conductance in Fig. 2(a) to the diffusive subband structure in Fig. 2(b), we examined the variation of the conductance with the length of the disordered region. The average conductance of 100 parallel wires containing only repulsive scatterers ($f=0.0$) is plotted versus L in Fig. 3(a). The Fermi level is placed at energies just below (dashed), directly on (solid), and just above (dot-dashed) the second and third subband minima. The average conductance decreases roughly exponentially with length at each value of the Fermi energy, as in one-dimensional electron localization theory.[46,47] For short disordered segments L, the average conductance is seen always to increase with Fermi energy. However, as the disordered region is made longer, a crossover length L_c is found such that when $L > L_c$ the average conductance falls after the Fermi energy passes through a new subband minimum. Consistent with this result, the electron localization length η, found from $\langle G \rangle \simeq \exp(-L/\eta)$, is appreciably shorter just after the Fermi energy moves into a new quasi-one-dimensional subband. This decrease in localization length is systematic and repeats around each new subband minimum in Fig. 3(a). If the scatterers are made stronger by increasing γ_i, the point at which the curves "cross over" occurs for a shorter length L_c of the disordered region.

In Fig. 3(b) we plot the ensemble-averaged conductance versus length when $f=0.5$ for the same energies as in Fig. 3(a). Quantum diffusion is still evident, since the conductance still decreases roughly exponentially with L. But in contrast to Fig. 3(a), there is no crossing over of the conductance versus length curves so that η is roughly the same (or increases) for increasing values of the Fermi energy. The conductance drops versus Fermi energy when $f=0.5$ in Figs. 2(a) and 2(b), therefore, do not depend on electron "localization" phenomena, and can be seen in the conductance of each individual ensemble member when only a few scatterers are present. Indeed, conductance drops of this sort occur if only *one* attractive scatterer is present in a single wire,[3-6] and therefore clearly do not depend on multiple reflections between different scatterers. In contrast, when all of the scatterers are repulsive in Fig. 2(b), the conductance drops after a subband opens occur only if the electron motion is phase coherent over a long enough segment of the conductor.[48]

A drop in conductance due to enhanced scattering after the opening of a subband channel has been argued previously,[23-31] based on a Drude model in which the collision time is modified to account for scattering between quantum channels. Calculating the electron transmission along the conductor semiclassically, by exactly treating the quantum-mechanical scattering at each impurity but neglecting the wave interference between scattering events, is a valid approximation in many circumstances. However, Refs. 23-31 also assume that the electron scattering at each individual impurity can be calculated in the Born approximation. Although the Born approximation adequately describes single subband transport,[49] we feel that the Born approximation (or "golden-rule" scattering rate) probably cannot be used to describe multiple subband transport along a quantum wire.

Because wave interference between different scattering events is neglected in the Drude approximation, a drop in

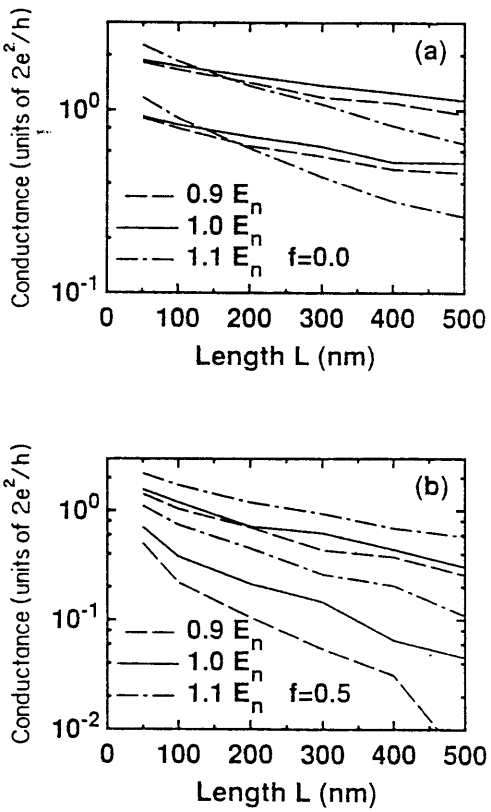


FIG. 3. Average conductance vs length L for an array of 100 parallel wires having (a) $f=0.0$ and (b) $f=0.5$. Six values of the Fermi energy are shown: $0.9E_2$ and $0.9E_3$ (dashed lined), E_2 and E_3 (solid line), and $1.1E_2$ and $1.1E_3$ (dot-dashed line). The "crossing over" of the conductance curves in (a) indicates a much shorter localization length when the Fermi energy moves into a new subband. Quasidonor states present in (b) depress the conductance below each subband, so there is no crossing over of the conductance curves.

the Drude conductance after a subband crossing should occur only if the transmission at each individual scatterer decreases as a new channel is opened. However, as shown, for example, in Refs. 3–6, the electron transmission through a single impurity increases when a new subband channel is opened. Furthermore, the golden-rule (or first-Born-approximation) approach to calculating scatteirng rates is an approximation depending only on the square magnitude of the scattering potential, so that attractive and repulsive scatterers erroneously give the same subband structure of the conductance versus Fermi energy in the Born approximation. Finally, the Born approximation explicitly breaks down at a subband minimum, as each term in the Born series becomes infinite.[5] Treating the first term in this series in a "self-consistent Born approximation" does not remedy the problem, since the change in transmission at a subband crossing is still in the wrong direction.

The transition from quantum diffusion to classical diffusion must be made by introducing additional phase-breaking scattering in the conductor.[49] If this is done using the method of Büttiker,[50] and if the electron suffers a phase-breaking event with probability 1 between each elastic-scattering event, the resulting classical diffusive transport is equivalent to adding many constriction conductances in series (where each constriction also contains an elastic scatterer). The conductance versus Fermi energy of such a classical diffusive wire will then resemble Fig. 2(a), i.e., it makes little qualitative difference in the shape of the conductance versus Fermi energy whether the ensemble-averaged conductance is obtained by adding the conductors classically in series or in parallel. Other assumptions about how the electron phase is broken inside the conductor may lead to different subband structures for the conductance versus Fermi energy when the transport is classical and diffusive. If the electron phase could be disrupted without obtaining any extra resistance, such that one would add the multichannel "R/T" Landauer resistances[35] in series to approach classical diffusive behavior,[49,51] the ensemble-averaged conductance would then resemble the "four-probe" result shown in Fig. 8 of Ref. 4. There should be little qualitative difference in the dependence of conductance versus Fermi energy between quantum diffusion and classical diffusion if some fraction of the scatterers in the wire are attractive, since the transmission through each individual attractive scatterer qualitatively resembles the coherent ensemble-averaged transmission through the wire array in this case.

Many of the issues raised in this paper have also been considered in connection with the quantized Hall effect. Differences between the electrical transmission through repulsive versus attractive scatterers have been of some interest in the quantized Hall effect.[53] Similarly, states analogous to the "quasidonor levels" of this paper can occur due to the formation of "quasibound states" at a local widening of the conductor in mesoscopic Hall crosses.[54] Electron "localization" in a narrow wire subject to a magnetic field has also been studied.[55]

IV. CONCLUSION

We find that the electrical conductance versus Fermi energy in an array of disordered quasi-one-dimensional wires can indeed evolve from the ballistic conductance steps into conductance drops after a new subband opens, but for reasons totally different from those given in Refs. 23–31. In the Drude model, the conductance decreases linearly with the length of the wire so that the electrons are delocalized. In our calculation, we find that the electrical conductance in a disordered quasi-one-dimensional wire decreases roughly exponentially with the length of the disordered region. The localization length associated with this decay is reduced when the Fermi energy crosses a confinement subband, leading to a new quantum diffusive subband structure in arrays of long quantum wires. Therefore, we find that the occurrence of conductance drops after a subband channel opening depends on quantum diffusion, and does not occur if the diffusion is classical (as in the Drude model). Similar results should follow from tight-binding transmission calculations[8,9] if care is taken to exclude the formation of quasidonor levels in a locally attractive impurity.[4–7]

These quasibound states in the attractive impurities are manifest as pronounced conductance dips below each subband minimum,[3–9] as reconfirmed for the ensemble-averaged conductance[8,9] in Fig. 2 of this work. These same conductance drops below a new subband[3–9] also arise if only one or a few attractive impurities are present in the wire. Therefore, they are essentially due to changes in transmission occurring at a single attractive impurity site and do not depend on either quantum or classical diffusion to occur, although their exact position in Fermi energy can be somewhat modified due to electron wave interference.[7] If the potential energy is locally attractive, so that quasidonor levels can form in an impurity, the new subband structure arising from localization phenomena studied in this paper is modified by the depressed transmission near the quasibound states. In conclusion, our study shows that there is no unique "ensemble-averaged conductance" of a disordered quasi-one-dimensional wire. Instead, the ensemble-averaged conductance of such a wire depends strongly and qualitatively on the type of disorder present.

Note added in proof. Since submission of this paper, we have become aware of several studies treating the conductance of disordered narrow wires.[56–65]

ACKNOWLEDGMENTS

We thank Terry P. Orlando, Dimitri A. Antoniadis, Henry I. Smith, and Kevin Delin for useful discussions. This work was sponsored by the U.S. Air Force Office of Scientific Research under Grant No. AFOSR-88-0304. A.K. gratefully acknowledges support from the Semiconductor Research Corporation.

*Present address: School of Electrical Engineering, Purdue University, West Lafayette, IN 47907.

[1]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).

[2]D. A. Wharam, T. J. Thornton, R. Newbury, M. Pepper, H. Ahmed, J. E. F. Frost, D. G. Hasko, D. C. Peacock, D. A. Ritchie, and G. A. C. Jones, J. Phys. C 21, L209 (1988).

[3]C. S. Chu and R. S. Sorbello, Phys. Rev. B 40, 5941 (1989); 38, 7260 (1988).

[4]P. F. Bagwell, Phys. Rev. B 41, 10 354 (1990).

[5]P. F. Bagwell, J. Phys. Condens. Matter 2, 6179 (1990).

[6]E. Tekman and S. Ciraci, Phys. Rev. B 42, 9098 (1990).

[7]A. Kumar and P. F. Bagwell, Phys. Rev. B 43, 9012 (1991); A. Kumar and P. F. Bagwell, Solid State Commun. 75, 949 (1990).

[8]J. Masek, P. Lipavsky, and B. Kramer, J. Phys. Condens. Matter 1, 6395 (1989).

[9]I. Kander, Y. Imry, and U. Sivan, Phys. Rev. B 41, 12 941 (1990).

[10]F. M. Peeters, in *Science and Engineering of 1- and 0- Dimensional Semiconductors*, edited by S. Beaumont and C. Sotomayor-Torres (Plenum, New York, 1990).

[11]D. van der Marel and E. G. Haanappel, Phys. Rev. B 39, 7811 (1989).

[12]E. Tekman and S. Ciraci, Phys. Rev. B 40, 8559 (1989).

[13]S. He and S. Das Sarma, Phys. Rev. B 40, 3379 (1989).

[14]A. Szafer and A. D. Stone, Phys. Rev. Lett. 62, 300 (1989).

[15]E. Castaño and G. Kirczenow, Solid State Commun. 70, 801 (1989).

[16]M. Cahay, M. McLennan, and S. Datta, Phys. Rev. B 37, 10 125 (1988); S. Datta, M. Cahay, and M. McLennan, *ibid.* 36, 5655 (1987).

[17]M. Cahay, S. Bandyopadhyay, M. A. Osman, and H. L. Grubin, Surf. Sci. 228, 301 (1990); M. Cahay, P. Marzolf, and S. Bandyopadhyay, in *Computational Electronics: Semiconductor and Device Simulation*, edited by K. Hess, J. P. Leburton, and U. Ravioli (Kluwer Academic, Boston, 1990), p. 263; S. Bandyopadhyay and M. Cahay, in *ibid.*, p. 223.

[18]P. A. Lee and A. D. Stone, Phys. Rev. Lett. 55, 1622 (1985).

[19]A. C. Warren, D. A. Antoniadis, and H. I. Smith, Phys. Rev. Lett. 56, 1858 (1986).

[20]K. Ismail, D. A. Antoniadis, and H. I. Smith, Appl. Phys. Lett. 54, 1130 (1989).

[21]J. R. Gao, C. de Graaf, J. Caro, S. Radelaar, M. Offenberg, V. Lauer, J. Singleton, T. J. B. M. Janssen, and J. A. A. J. Perenboom, Phys. Rev. B 41, 12 315 (1990).

[22]J. Faist, P. Gueret, and H. Rothuizen, Phys. Rev. B 42, 3217 (1990).

[23]B. E. Sernelius, K. F. Berggren, M. Tomak, and C. McFadden, J. Phys. C 18, 225 (1985).

[24]S. Das Sarma and X. C. Xie, Phys. Rev. B 35, 9875 (1987).

[25]M. J. Kearney and P. N. Butcher, J. Phys. C 20, 47 (1987).

[26]N. Trivedi and N. W. Ashcroft, Phys. Rev. B 38, 12 298 (1989).

[27]P. F. Bagwell and T. P. Orlando, Phys. Rev. B 40, 3735 (1989).

[28]P. Vasilopoulos and F. M. Peeters, Phys. Rev. B 40, 10 079 (1989).

[29]T. P. Orlando, P. F. Bagwell, R. A. Ghanbari, and K. Ismail, in *Electronic Properties of Multilayers and Low-Dimensional Semiconductor Structures*, edited by J. M. Chamberlain, L. Eaves, and J. C. Portal (Plenum, London, 1990); P. F. Bagwell, D. A. Antoniadis, and T. P. Orlando, Quantum Mechanical and Nonstationary Transport Phenomenon in Nanostructured Silicon Inversion Layers, in *Advanced MOS Device Physics*, edited by N. Einspruch and G. Gildenblat (Academic, San Diego, 1989), Vol. 18.

[30]S. Y. Qiu and M. V. Jaric, Bull. Am. Phys. Soc. 35, 492 (1990); W. Bao and Z. Tesanovic, *ibid.* 35, 492 (1990).

[31]M. Suhrke, W. Wilke, and R. Keiper, J. Phys. Condens. Matter 2, 6743 (1990).

[32]Y. Imry, in *Directions in Condensed Matter Physics*, edited by G. Grinstein and G. Mazenko (World Scientific, Singapore, 1986).

[33]M. Büttiker, Phys. Rev. Lett. 57, 1761 (1986).

[34]R. Landauer, Z. Phys. B 68, 217 (1987).

[35]M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Phys. Rev. B 31, 6207 (1985).

[36]M. Büttiker, IBM J. Res. Dev. 32, 317 (1988).

[37]M. Büttiker, in *Electronic Properties of Multilayers and Low-Dimensional Semiconductor Structures* (Ref. 29).

[38]R. Landauer, J. Phys. Condens. Matter 1, 8099 (1989).

[39]R. Landauer, in *Analogies in Optics and Micro-Electronics*, edited by W. van Haeringen and D. Lenstra (Kluwer Academic, Dordrecht, 1990).

[40]J. G. Williamson, C. E. Timmering, C.J.P.M. Harmans, J. J. Harris, and C. T. Foxon, Phys. Rev. B 42, 7675 (1990).

[41]L. I. Glazman, G. B. Lesovik, D. E. Khmel'nitskii, and R. I. Shekhter, Pis'ma Zh. Eksp. Teor. Fiz. 48, 218 (1988) [JETP Lett. 48, 238 (1988)].

[42]A. Yacoby and Y. Imry, Phys. Rev. B 41, 5341 (1990).

[43]M. Büttiker, Phys. Rev. B 41, 7906 (1990).

[44]H. Sakaki, Jpn. J. Appl. Phys. 19, L305 (1980).

[45]A. Kumar, S. E. Laux, and F. Stern, Appl. Phys. Lett. 54, 1270 (1989).

[46]R. Landauer, Philos. Mag. 21, 863 (1970); P. W. Anderson, D. J. Thouless, E. Abrahams, and D. S. Fisher, Phys. Rev. B 22, 3519 (1980).

[47]The exponential decrease of conductance with L at $E=E_2$ and $E=E_3$ we find surprising, since Ref. 7 shows that wave interference between propagating modes does not determine the electron transmission properties at a subband minimum. The "localization" must therefore arise from a different mechanism.

[48]We note that more ensemble members are needed to average over fluctuations if attractive scatterers are present, so that the curves in Fig. 3(b) are not as smooth as the ones in Fig. 3(a). We have also explored making all the scatterers attractive ($f=1.0$), and find behavior qualitatively similar to the $f=0.5$ case. However, the average transmission is generally lowered as a greater fraction of the scatterers are made attractive.

[49]R. Landauer, Philos. Mag. 21, 863 (1970).

[50]M. Büttiker, Phys. Rev. B 33, 3020 (1986).

[51]P. F. Bagwell and T. P. Orlando, Phys. Rev. B 40, 1456 (1989).

[52]P. F. Bagwell, Ph.D. thesis, Massachusetts Institute of Technology, 1991.

[53]R. J. Haug, R. R. Gerhardts, K. v. Klitzing, and K. Ploog, Phys. Rev. Lett. 59, 1349 (1987).

[54]R. L. Schult, H. W. Wyld, and D. G. Ravenhall, Phys. Rev. B 41, 12 760 (1990).

[55]J. M. Kinaret and P. A. Lee, Phys. Rev. B 43, 3847 (1991).

[56]Y. Joe and S. E. Ulloa, in *Nanostructures: Fabrications in Physics*, edited by T. P. Smith III, S. D. Burger, D. Kern, and H. Craighead, MRS Extended Abstract No. 26 (Materials Research Society, Pittsburgh, 1990), p. 47, and unpublished.

[57]H. U. Baranger, Phys. Rev. B 42, 11 479 (1990).

[58]G. Y. Hu and R. F. O'Connell, Phys. Rev. B 43, 12 341 (1991).

[59]J. A. Nixon, J. H. Davies, and H. U. Baranger, Phys. Rev. B 43, 12 638 (1991).
[60]L. Wang, S. Feng, and Y. Zhu, Bull. Am. Phys. Soc. 36, 357 (1991).
[61]V. Mitin, Bull. Am. Phys. Soc. 36, 412 (1991).
[62]S. Choudhuri, S. Bandyopadhyay, and M. Cahay, in *Proceedings of the Second International Symposium on Nanostructures and Mesoscopic Systems*, edited by W. P. Kirk (Academic, New York, in press).
[63]H. Kasai, K. Mitsutake, and A. Okiji, J. Phys. Soc. Jpn. (to be published).
[64]H. Tamura and T. Ando (unpublished).
[65]M. K. Schwalm and W. A. Schwalm (unpublished).

Appendix C

Reprint: Anti-Correlated Oscillations in a Three-Lead Quantum Dot

This Appendix contains a reprint of the following paper:

A. Kumar, J. Kinaret, C. Eugster, T. Orlando, D. Antoniadis, M. Rooks, and M. Melloch, "Anti-Correlated Oscillations in a Three-Lead Quantum Dot," presented at *Rencontres de Moriond: Coulomb Blockade and Interference Effects in Small Electronic Structures*, Villars-sur-Ollon, Switzerland, 22-29 January 1994 (to be published in the Proceedings).

Manuscript for RENCONTRES DE MORIOND: COULOMB AND INTERFERENCE EFFECTS IN SMALL ELECTRONIC STRUCTURES, *Villars-sur-Ollon, Switzerland, 22-29 January 1994*

ANTI-CORRELATED OSCILLATIONS IN A THREE-LEAD QUANTUM DOT

A. Kumar, J. Kinaret*, C. Eugster, T. Orlando, D. Antoniadis, M. Rooks†, M. Melloch‡
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, U.S.A.

We present results of transport measurements on a quantum dot in which a novel gate geometry allows the dot to be contacted by *three*, rather than two, leads. When the dot charge is well-confined, periodic conductance oscillations due to Coulomb charging are observed in-phase with each other at two of the leads in response to a small excitation voltage at the third. As the tunnel barriers are made softer by changing the gate voltage, a strikingly different phenomenon is observed: conductance peaks at the two output leads evolve from perfect *correlation* to perfect *anti-correlation* with each other. A simple model incorporating polarization states of the dot-lead system is presented as a possible explanation.

Transport measurements have found that the conductance of a quantum dot structure is a periodic function of an external gate voltage.[1)] This striking modulation of the conductance results from the condition that the dot charge is an integer multiple of the electron charge. If the gate voltage is such that the electrochemical potential of the dot lies between the quasi-Fermi levels of the leads, the number of electrons in the dot can fluctuate classically by one and hence the current is at maximum. For all other gate voltages, there is a Coulomb charging energy associated with the tunneling of an additional electron from the input lead, and current flow is suppressed.

The above Single Electron Tunneling (SET) picture of electron transport has been highly successful in explaining the majority of conductance experiments to date[1)] in which a quantum dot is coupled to *two* electron reservoirs. In this paper, we present results of transport

*Nordita, Blegdamsvej 17, DK-2100, Copenhagen Ø, Denmark.

†National Nanofabrication Facility, Knight Laboratory, Cornell University, Ithaca, New York 14853-5403.

‡School of Electrical Engineering, Purdue University, West Lafayette, Indiana 47907.

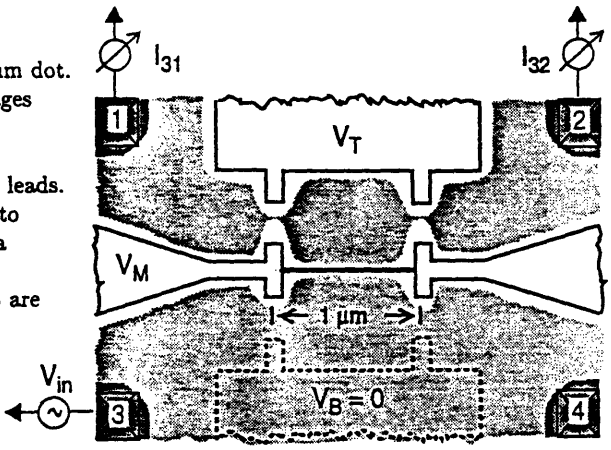


Figure 1: Schematic of the gate geometry of our three-lead quantum dot. Application of negative gate voltages V_T and V_M depletes the 2DEG underneath, leaving behind an electron "island" coupled to three leads. Four ohmic contacts allow access to the 2DEG. In our measurements a voltage V_{in} is applied at contact 3 and currents I_{31} and I_{32} are measured simultaneously.

measurements on a structure in which a novel gate geometry allows us to study a quantum dot coupled via tunnel barriers to *three* leads. A small voltage excitation at one lead results in currents at the other two leads, which can be measured simultaneously. We find that the first several Coulomb blockade oscillations in the two currents line up with each other as the gate voltage is swept, as expected from the simple SET picture. However, the main finding of this work is that, as the gate voltage is increased further, the oscillations undergo a striking transition from being lined up with each other to being perfectly *out-of-phase* with each other.

Figure 1 shows a schematic of our gate geometry, fabricated over a standard GaAs/AlGaAs heterostructure with mobility 25 $m^2\ V^{-1}\ s^{-1}$ and carrier density $3\times10^{15}\ m^{-2}$ at 4.2 K. Application of a negative bias voltage to the top gate, V_T, and thin ($\sim$ 40 nm linewidth) middle gate, V_M, depletes the high-mobility 2DEG formed in a GaAs/AlGaAs modulation-doped heterostructure, resulting in an "island" of electrons coupled to narrow channels on the left and right, and to a semi-infinite 2DEG on the bottom. The currents at the output leads (lead 1 and lead 2) are measured in response to a small ac voltage V_{in} applied at the input lead (lead 3). Two low noise current amplifiers are used to measure simultaneously the currents $I_{31} = G_{31}V_{in}$ and $I_{32} = G_{32}V_{in}$ at the output leads. Each current amplifier outputs a voltage which is measured using an ac lock-in technique at frequency 7.7 Hz. Blocking capacitors are used at the inputs of the current amplifiers to shield the device from offset voltages; each output lead is therefore a virtual ground. As a check that the offset voltages are properly compensated, essentially no change is observed in I_{31} and I_{32} if the two amplifiers are interchanged. All measurements are carried out at a base cryostat temperature of 300 mK.

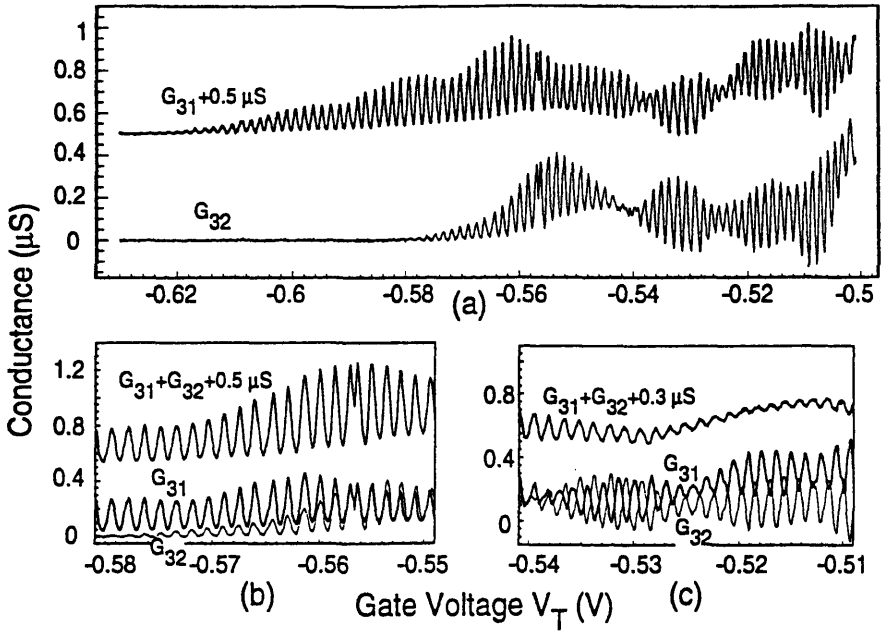


Figure 2: (a) Conductances G_{31} (offset by 0.5μS) and G_{32} as the top gate voltage V_T is swept. The middle gate voltage $V_M = -0.7$V is fixed in the tunnel regime. (b) Detailed plot of the first several resonances in (a), indicating that they are in-phase with each other. (c) Detailed plot of resonances at higher V_T, indicating that they have evolved from almost perfect *correlation* to almost perfect *anti-correlation* in gate voltage.

Figure 2(a) shows the conductances G_{31} and G_{32}, measured concurrently, as the top gate voltage V_T is swept. The middle gate voltage $V_M = -0.7$V is kept fixed in the tunnel regime, as determined by an exponential tail in its pinchoff characteristic.[2] Due to some intrinsic, unintentional asymmetry in our structure, the quantum point contact near lead 2 has a turn-on voltage about 40 mV higher than the one near lead 1.

The first several conductance oscillations in G_{31} and G_{32}, along with the total conductance $G_{31} + G_{32}$ through the dot, are shown in detail in Fig. 2(b). The perfect alignment of the oscillations in G_{31} and G_{32} for $V_T <\sim -0.555$V can be understood simply from standard SET theory. At a conductance maximum, an electron which has tunneled into the dot from the input lead has some probability of being transmitted through either one of the two output leads. As the gate voltage is increased in Fig. 2(c), the peak-to-valley ratio in the total dot conductance $G_{31} + G_{32}$ drops markedly. However, instead of broadening accordingly, resonances in the individual conductances G_{31} and G_{32} evolve from being perfectly correlated to being perfectly anti-correlated with each other. Another striking feature of the anti-correlated regime is that

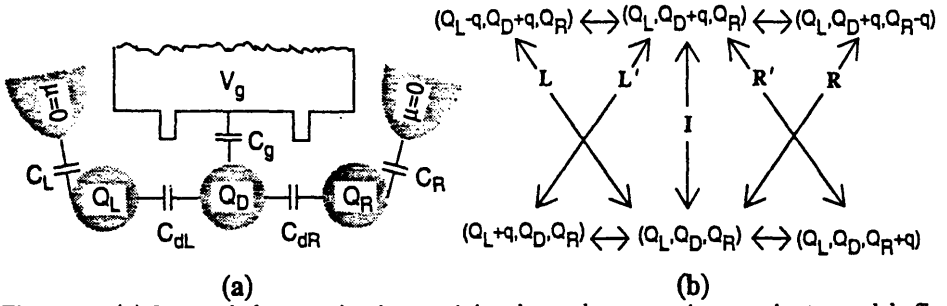


Figure 3: (a) Lumped element circuit containing dot and two quasi-reservoirs to model effect of unequal electrochemical potentials on left and right sides of the dot. (b) Set of states used in calculation and possible transitions between them. The set of states includes the lowest-energy unit polarization fluctuations of the states (Q_L, Q_D, Q_R) and $(Q_L, Q_D + q, Q_R)$. The unlabeled transitions show no resonant structure because of the strong coupling between the quasi-reservoirs and the reservoirs.

G_{32} actually changes sign at some of the deep valleys. The total conductance through the dot $G_{31} + G_{32}$ is, however, always positive.

We have fitted the lineshapes of $G_{31}+G_{32}$. In the correlated regime we find an excellent fit of the lineshapes to the derivative of the Fermi-Dirac function. As the gate voltage is increased so that the oscillations become anti-correlated, we find that the Fermi fit to $G_{31}+G_{32}$ progressively worsens, predicting valleys much deeper and linewidths much narrower than given by the data. Foxman et al.[3] have found that such a transition from thermally to intrinsically broadened resonances is accompanied by a rapid increase in the capacitance between the dot and one of the leads.

We now turn to a preliminary model to account for the separation of the peaks in the left and right currents with gate voltage. The crucial ingredient of the model is that there is some intrinsic, unintentional asymmetry between the right and left tunnel barriers which causes the local chemical potentials to be different on the left and the right sides of the dot. We can associate an effective capacitance with each barrier, the value of which depends on the barrier thickness and height. As the barriers are made softer by increasing the top gate voltage V_T, the barrier capacitances increase, and the relative barrier asymmetry becomes more important. In this regime it is no longer obvious that the transitions necessary to produce current peaks at the left and right leads will resonate at the same gate voltage.

Figure 3(a) depicts a lumped-element circuit used as a first approximation to model the effect of unequal electrochemical potentials. In addition to the dot with quantized charge $Q_D = -N_D e$, the model contains two quasi-reservoirs which are coupled to the actual reservoirs

by large capacitances C_L and C_R. Physically, we expect that a test charge placed in a quasi-reservoir, representing the part of the lead nearest the barrier, will be partially imaged in the dot and partially imaged in the actual reservoir.

The electrostatic energy of the circuit in Fig. 3(a) is

$$W(Q_D, Q_L, Q_R, V_g) = \frac{1}{2\tilde{C}_d}(Q_D + \gamma_L Q_L + \gamma_R Q_R)^2 + \frac{1}{2\tilde{C}_L}Q_L^2 + \frac{1}{2\tilde{C}_R}Q_R^2 + \frac{C_g}{\tilde{C}_d}V_g(Q_D + \gamma_L Q_L + \gamma_R Q_R), \quad (1)$$

where $\tilde{C}_{L(R)} = C_{L(R)} + C_{dL(dR)}$, $\gamma_{L(R)} = C_{dL(dR)}/\tilde{C}_{L(R)}$, and $\tilde{C}_d = C_g + (1-\gamma_L)C_{dL} + (1-\gamma_R)C_{dR}$. Figure 3(b) shows the set of states used in the model and the possible transitions between them. The set of states includes the lowest-energy unit polarization fluctuations of the states (Q_L, Q_D, Q_R) and $(Q_L, Q_D + q, Q_R)$. The procedure for finding the resonant gate voltages for each of the 5 transitions I,L,L′,R,R′ is as follows. The electrostatic energies of the two states involved in each transition are equated, relating the resonant gate voltage V_g to the charges (Q_L, Q_D, Q_R). To find the charges, it is assumed that the time-average charges $\overline{Q}_{L(R)} = \sum_i P_i Q_{L(R)}^{(i)}$ are given by the classical electrostatics relations $\partial W/\partial Q_{L(R)} = 0$, where the occupancy probability P_i for state i is taken in the grand canonical ensemble.

This procedure yields a set of coupled transcendental equations. In the limit $C_L, C_R >> C_g, C_{dL}, C_{dR}$, we recover the conventional SET theory result that all transitions are at resonance at the same gate voltage $V_g = e(N_D + 1/2)/C_g$. As the tunnel barriers become softer with increasing gate voltage, the capacitances C_{dL}, C_{dR} become comparable to C_L, C_R, in which case the equations must be solved numerically. Figure 4 shows the results of such a calculation, in which the resonant gate voltage positions for the transitions L,L′,R,R′ are plotted as a function of C_{dL} for the highly asymmetric case $C_{dL} = 10C_{dR}$. With increasing $C_{dL} = 10C_{dR}$, the injector transition I is unchanged and the transitions L,L′ are shifted only a little. However, the transitions R,R′ are strongly shifted; in the limiting case $C_{dL}/C_L \simeq 1$, when the left barrier has nearly disappeared, transitions R,R′ are at resonance almost midway between successive resonances of the injector transition I.

This polarization model suggests a mechanism by which a phase shift in the resonances can occur. For a peak to occur in the current measured at the right lead, there must be a chain of events involving transition I and either transition R or R′, which no longer occur at the same gate voltage. The currents I_{31}, I_{32} versus gate voltage can be found numerically by replacing the equilibrium probabilities P_i with a non-equilibrium distribution function determined by a rate equation.[1)] Such a calculation has been carried out[4)] and indeed yields anti-aligned oscillations

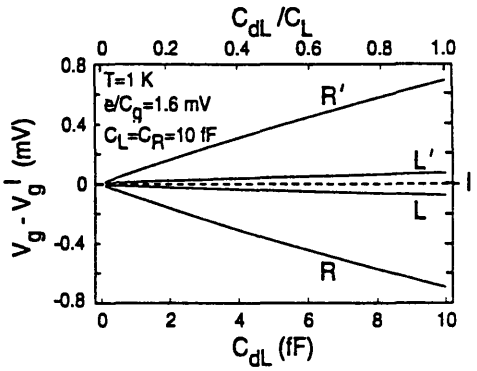


Figure 4: Gate voltages at which each of the transitions L,L',R,R' are at resonance (relative to the position of the I resonance) as capacitances C_{dL} and C_{dR} are increased, for the highly asymmetric case $C_{dL} = 10C_{dR}$. For $C_{dL} = C_{dR} = 0$, we recover the standard SET result that all transitions occur at the same gate voltage. As C_{dL}/C_L approaches 1, so that the left barrier has nearly disappeared, transitions R,R' occur almost midway between neighboring I transitions.

in the limit that C_{dL} and C_{dR} are large but asymmetric. However, the currents do not change sign.

In summary, we have fabricated and studied a three-lead quantum dot. When the dot confinement is strong, the two output leads behave as two independent, parallel channels for current flow, resulting in Coulomb blockade oscillations which are aligned with each other in gate voltage. As the leak rates to the output leads increase, peaks in the total incident current broaden accordingly, but peaks in the two output currents evolve from being correlated with each other to being anti-correlated with each other. In this regime the simple single electron tunneling picture is inadequate to explain our results, and we suggested a simple circuit model which reproduces many of the observed features.

The authors wish to thank N.R. Belk, M. Burkhardt, D.J. Carter, M.K. Mondol, R.R. Perilli, and G.E. Rittenhouse for help at various stages of this work. We have also benefited from discussions with P. Debray, E.B. Foxman, D.C. Glattli, and H.I. Smith. This work was sponsored by the U.S. Air Force Office of Scientific Research under agreement F49620-92-J-0064. This work was performed in part at the National Nanofabrication Facility which is supported by the National Science Foundation under Grant ECS-8619049, Cornell University, and industrial affiliates. One of us (A.K.) gratefully acknowledges support from a SRC fellowship.

1) For a review, see H. van Houten, C.W.J. Beenakker, and A.A.M. Staring, in *Single Charge Tunneling*, ed. by H. Grabert and M.H. Devoret (Plenum, New York, 1991).

2) C.C. Eugster and J.A. del Alamo, Phys. Rev. Lett. **67**, 3586 (1991).

3) 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).

4) J. Kinaret (private communication).

Appendix D

Reprints: Numerical Modeling of Quantum Dot Structures

This Appendix contains reprints of the following four papers:

Arvind Kumar, Steven E. Laux, and Frank Stern, "Channel Sensitivity to Gate Roughness in a Split-Gate GaAs/AlGaAs Heterostructure," Appl. Phys. Lett. **54**, 1270 (1989).

Arvind Kumar, Steven E. Laux, and Frank Stern, "Electron States in a GaAs Quantum Dot in a Magnetic Field," Phys. Rev. B **42**, 5166 (1990).

Arvind Kumar, "Self-Consistent Calculations on Confined Electrons in Three-Dimensional Geometries," Surf. Sci. **263**, 335 (1992) (presented at *Ninth International Conference on Electronic Properties of Two-Dimensional Systems*, Nara, Japan, 1991).

Arvind Kumar, Steven E. Laux, Frank Stern, A. Zaslavsky, J.M. Hong, and T.P. Smith, "Effect of Nonequilibrium Deep Donors in Heterostructure Modeling," Phys. Rev. B **48**, 4899 (1993).

The interested reader is also referred to

U. Meirav, P.L. McEuen, M.A. Kastner, E. B. Foxman, A. Kumar, and S.J. Wind, "Conductance Oscillations and Transport Spectroscopy of a Quantum Dot," Zeitschrift für Physik B **85**, 357 (1991).

Channel sensitivity to gate roughness in a split-gate GaAs-AlGaAs heterostructure

Arvind Kumar,[a] Steven E. Laux, and Frank Stern
IBM Research Division, T. J. Watson Research Center, Yorktown Heights, New York 10598

(Received 14 November; accepted for publication 13 January 1989)

The response of the width of the electron channel at a GaAs-AlGaAs heterointerface to variations in the gate opening of a split-gate structure is calculated using a three-dimensional solution of the Poisson equation in the continuum approximation and is analyzed in terms of the Fourier components of the perturbation. It is found that the effective potential well for the channel electron gas attenuates high wave vector components of the gate roughness.

The split-gate GaAs/AlGaAs heterostructure has been widely used as a device configuration to achieve electron confinement in two dimensions.[1,2] Laux *et al.*[3] have calculated the potential and electron distributions for such a device, with a split gate over a conventional GaAs-based heterostructure, using a self-consistent numerical solution of the Poisson and Schrödinger equations. Davies[4] has given an analytic treatment valid for cases in which there is little or no charge in the conducting electron channel in the GaAs. These treatments assumed that the gate edges are smooth, something not easy to achieve in real structures.[5] In this letter, we investigate the response of the electron channel to a perturbation in the width of the gate slit using a semi-classical approximation.

The structure we consider has the same cross section as the device in Fig. 1 of Ref. 3. However, our gate has a periodic square-wave modulation, illustrated in Fig. 1, such that the slit width averages to 400 nm, the value used in Ref. 3. From our calculation, the abrupt discontinuity in slit width shown in Fig. 1 results in a rounded contour of the conduction band edge at the Fermi energy in the plane of the GaAs/AlGaAs interface, as illustrated in Fig. 2.

To obtain the potential distribution, we use a three-dimensional extension of the Poisson equation solver described in Ref. 6, with provisions for carrier freezeout and discontinuous band gaps. The Poisson equation is discretized using a standard seven-point finite-difference scheme and is solved with a conjugate gradient method employing polynomial preconditioning.[7] For a given device geometry and set of bias conditions, the computation time of the matrix solver grows with the number of nodes N as N^{α}. We find α to range between 1.2 and 1.3, based on runs with values of N from $\sim 3\times 10^3$ to $\sim 3\times 10^5$. About 80% of the total computation time is spent in vector mode on an IBM 3090 computer.

We restrict our solution to a classical calculation because of the computational difficulty of dealing with continuum eigenstates in a quantum mechanical calculation. Zero-gradient boundary conditions at the edges of the mesh allow us to model only the half-period $|x| < L/4$ and to obtain the solution in the rest of the period by symmetry. All calculations were done at a temperature of 4.2 K.

We simulated structures having perturbations $2d = 80$ nm in slit width and periods ranging from 0.2 to 20.0 μm. To analyze our results, we decompose the conduction band edge contour[8] $y = f(x)$ at the Fermi energy in the plane $z = 0$ into its discrete Fourier components:

$$f(x) = b_0 + \sum_{m=1}^{\infty} a_m \sin\frac{2\pi m x}{L} + \sum_{m=1}^{\infty} b_m \cos\frac{2\pi m x}{L}, \quad (1)$$

where L is the period of the perturbation in gate slit width. For a pure square wave with the origin chosen as indicated in Fig. 1, we have $a_m = 0$ for even m and $b_m = 0$ for $m \neq 0$. For

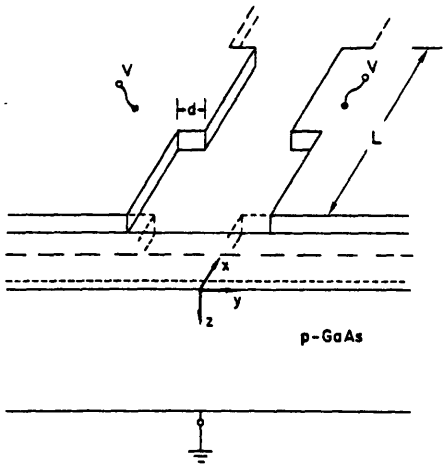


FIG. 1. Device geometry used in this work. A split gate having a square-wave variation of the slit width is placed on a GaAs-AlGaAs heterojunction. The layer thicknesses are 24 nm for the undoped GaAs top layer, 36 nm for the doped AlGaAs layer with 6×10^{17} donors per cm^3, 10 nm for the undoped ALGaAs spacer layer, and 0.93 μm for the GaAs substrate with 10^{14} acceptors per cm^3. The donor binding energy and heterojunction conduction band offset are taken to be 0.05 and 0.3 eV, respectively. Our calculation assumes a uniform surface charge density in the gate opening, ignoring a region of positive surface charge near the edge of the gate. The gate slit has an average width of 400 nm.

[a] Present address: Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139.

 0003-6951/89/131270-02$01.00

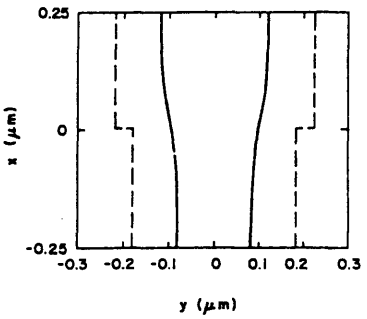


FIG. 2. Contour lines in the plane $z = 0$ of Fig. 1, showing where the conduction band edge of an electron channel crosses the Fermi level. Electrons occupy the area between the full curves. The parameters of the slit width modulation are $L = 1.0\ \mu$m and $2d = 80$ nm. The gate voltage is 0.3 V above the threshold for appearance of an electron channel at the GaAs/AlGaAs interface. The dashed line is the projection of the gate in this plane.

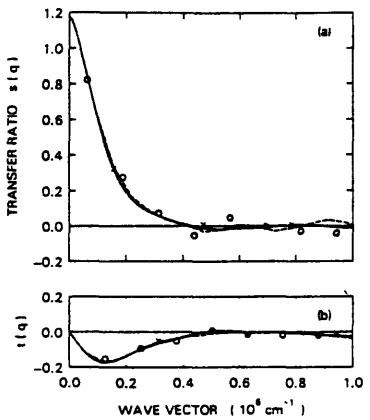


FIG. 3. (a) Transfer ratio $s(q)$ at a gate voltage 0.3 V above threshold, for gate slit width perturbation with amplitude $2d = 80$ nm and with four periods L: 0.4 μm, crosses; 1.0 μm, circles; 2.0 μm, dashed curve; and 20 μm, full curve. The full curve was calculated with more mesh points than the other three cases, and is considered to be more accurate. (b) "Forbidden" transfer ratio $t(q)$ for the same parameters as in (a).

the conduction band edge contour, we find a_m to be essentially zero for even m. However, we find b_m for even m to be nonzero, although small compared to a_m for adjacent odd m.

We divide the amplitudes of the sine terms in the expansion of the conduction band edge contour by the corresponding amplitudes of the Fourier components of the square wave modulation in the gate slit width to obtain the "transfer ratio" at wave vector $q = 2\pi m/L$:

$$s(q) = \frac{a_m}{2d/\pi m}, \quad m = 1,3,5,... \tag{2}$$

Figure 3(a) shows the transfer ratio $s(q)$ for four values of the gate slit modulation period L at a gate voltage $V = -1.2$ V, about 0.3 V above the threshold for appearance of charge in the channel. All simulations reveal a qualitatively similar behavior characterized by the filtering of high q components of the perturbation. We note that the region between $q = 5\times10^4$ cm^{-1} and $q = 3.5\times10^5$ cm^{-1} is given approximately by

$$s(q) \simeq C\exp(-\beta q), \tag{3}$$

where $C = 1.88$ and $\beta = 1.14\times10^{-5}$ cm.

For the coefficients of the cosine terms in Eq. (1) we define a "forbidden" transfer ratio at wave vector $q = 2\pi m/L$:

$$t(q) = \frac{b_m}{2d/\pi m}, \quad m = 2,4,6,... \tag{4}$$

Values of $t(q)$ are shown in Fig. 3(b).

Our calculations show that for a small range of gate voltages near threshold the square-wave slit width modulation used here leads to pockets of electrons under the wide slit areas, with carriers excluded under the narrow slit areas. As the gate voltage becomes more positive, these pockets of electrons expand to form continuous contours, as shown in Fig. 2.

A real gate can be expected to have random roughness, whose Fourier expansion—like that of the square wave assumed here—will include a wide range of wave vectors. Our results indicate that the large wave vector components, including those whose magnitude is comparable to the Fermi wave vector of the electrons (expected to be of order 10^6 cm^{-1} for a quasi-one-dimensional electron gas[3,6]), will be substantially attenuated for structures similar to the one considered here, reducing their effectiveness in scattering.

Among problems left for future work are the behavior of the transfer ratio as the distance between the gate and the channel is reduced and the implementation of a fully quantum-mechanical self-consistent treatment. The latter problem is complicated by the need to deal with continuum eigenstates and by the long computation times involved in three-dimensional calculations.

We would like to thank S. Furkay, F. Pileggi, M. Amidon, and L. Borucki for graphics and data base support. One of us (A. K.) is grateful for the opportunity to work at IBM Research during the summer.

[1] T. J. Thornton, M. Pepper, H. Ahmed, D. Andrews, and G. J. Davies, Phys. Rev. Lett. 56, 1198 (1986).
[2] H. Z. Zheng, H. P. Wei, D. C. Tsui, and G. Weimann, Phys. Rev. B 34, 5635 (1986).
[3] S. E. Laux, D. J. Frank, and F. Stern, Surf. Sci. 196, 101 (1988).
[4] J. H. Davies, Semicond. Sci. Technol. 3, 995 (1988). We are indebted to Dr. Davies for a preprint.
[5] See, for example, U. Merkt, Ch. Sikorski, and J. P. Kotthaus, Superlatt. Microstruct. 3, 679 (1987), Fig. 1(b), and the papers on nanostructure technology in Vol. 32, No. 4 of the IBM Journal of Research and Development.
[6] S. E. Laux and F. Stern, Appl. Phys. Lett. 49, 91 (1986).
[7] O. G. Johnson, C. A. Micchelli, and G. Paul, SIAM J. Numer. Anal. 20, 362 (1984).
[8] There are two contours, symmetric about the line $y = 0$. We use the one with $y > 0$.

PHYSICAL REVIEW B VOLUME 42, NUMBER 8 15 SEPTEMBER 1990-I

Electron states in a GaAs quantum dot in a magnetic field

Arvind Kumar
*IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598 and Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139**

Steven E. Laux and Frank Stern
IBM Research Division, Thomas J. Watson Research Center, P.O. Box 218, Yorktown Heights, New York 10598
(Received 12 March 1990; revised manuscript received 6 June 1990)

Self-consistent numerical solutions of the Poisson and Schrödinger equations have been obtained for electron states in a GaAs/$Al_xGa_{1-x}As$ heterostructure with confinement in all three spatial dimensions. The equations are solved in the Hartree approximation, omitting exchange and correlation effects. Potential profiles, energy levels, and the charge in the quantum dot are obtained as functions of the applied gate voltage and magnetic field. First, the zero-magnetic-field case is considered, and the quantum-dot charge is allowed to vary continuously as the gate voltage is swept. Then, in connection with the phenomenon of Coulomb blockade, the number of electrons in the quantum dot is constrained to integer values. Finally, the calculation is extended to examine the evolution of levels in a magnetic field applied perpendicular to the heterojunction. Our results indicate that the confining potential has nearly circular symmetry despite the square geometry of the gate, that the energy levels are quite insensitive to the charge in the quantum dot at a fixed gate voltage, and that the evolution of levels with increasing magnetic field is similar to that found for a parabolic potential.

I. INTRODUCTION

Modern fabrication techniques have made possible confinement of a two-dimensional layer of electrons into wires, grids, or dots where quantum-mechanical effects are strongly manifested. Of particular current interest are quasi-zero-dimensional systems, which have been made by selective etching of a GaAs cap on a GaAs-$Al_xGa_{1-x}As$ heterostructure,[1] by depositing a cross-grid gate structure on a GaAs heterostructure[2] or on Si,[3] by using crossed holographically defined gratings,[4] and by using an array of small Latex particles as an etch mask,[5] to cite some recent examples. Such quantum-dot structures offer a dispersionless system with an electron-energy spectrum that can be modulated either by varying gate bias voltage or by applying an external magnetic field. Smith *et al.*[1] have reported oscillatory structure in capacitance versus gate voltage in zero magnetic field and have attributed it to the discrete energy states of a quantum dot. Recently, Hansen *et al.*[6] have reported observing Zeeman splitting of quantum-dot capacitance features, as expected when a magnetic field is applied perpendicular to the heterojunction. There are many papers that treat the energy-level structure of related systems, including the paper by Darwin that treats a two-dimensional harmonic-oscillator potential in the presence of a normal magnetic field,[7] Robnik's paper on a disk in a magnetic field,[8] and the recent calculations by Brum and co-authors on a model quantum dot.[6,9] Sivan and Imry[10] have described the evolution of states in a quantum dot versus magnetic field in relation to magnetization and persistent currents, which are not considered here. In this paper we present numerical self-consistent results in the Hartree approximation for potential profiles, energy levels, envelope wave functions, and charge distributions for quantum dots like those studied by Hansen *et al.*[6]

Self-consistent numerical treatments of electron states in quasi-one-dimensional systems in the absence of a magnetic field have been carried out for a narrow channel in silicon by Laux and Stern[11] and for a split-gate GaAs/$Al_xGa_{1-x}As$ heterostructure by Laux *et al.*[12] Numerical methods for such systems, which have quantities that vary in two spatial dimensions, have also been used by Kojima *et al.*[13] and by Kerkhoven *et al.*[14] However, the analogous calculation for a totally confined system requires a coupled solution of Poisson's equation and Schrödinger's equation in three spatial dimensions, increasing the computation requirements significantly. Application of a magnetic field, which leads to complex wave functions and a Hermitian rather than a real symmetric eigenvalue problem, also adds to the computational burden. In Sec. II we discuss our formulation of this problem. In Sec. III we discuss results of such a calculation on a GaAs/$Al_xGa_{1-x}As$ quantum-dot structure in zero applied magnetic field. In particular, we discuss the effect of varying the charge in the quantum dot on the energy-level structure and the quasi-Fermi-level, and its relation to the Coulomb blockade.[15]

In Sec. IV we extend our self-consistent calculation to include the effect of an applied magnetic field perpendicular to the heterojunction on the potential, charge density, and electron states of the quantum dot. We find good

qualitative agreement with the results of earlier calculations for model potentials.

II. FORMULATION OF THE PROBLEM

The structure we consider is a model of a single quantum dot from the array of dots used in the experiments of Hansen *et al.*,[6] described above. It is based on a heterostructure with an *n*-type GaAs substrate layer with a net ionized donor concentration of 10^{18} cm^{-3}, an 80-nm layer of undoped GaAs (a background acceptor concentration of 10^{14} cm^{-3} is assumed throughout and diffusion of donors from the substrate is ignored), a 20-nm layer of undoped $Al_{0.4}Ga_{0.6}As$, a 20-nm layer of the same material with a donor concentration of 1.5×10^{18} cm^{-3}, and a 30-nm GaAs cap layer. The cap is etched away, except in the central 300 nm$\times$300 nm portion of a 500 nm$\times$500 nm area, and the structure repeats on a square lattice. Finally, a metal gate is deposited over the entire top surface. A negative voltage on the gate depletes the charges in the GaAs channel, except under the remaining GaAs cap, and this three-dimensionally confined "puddle" of electrons is the quantum dot being studied. The *n*-type GaAs substrate allows a low-impedance capacitative contact to the dot. Figure 1 shows the conduction-band edge in the structure versus vertical distance along a line through the center of the dot, and in the inset is sketched the semiconductor region included in the model.

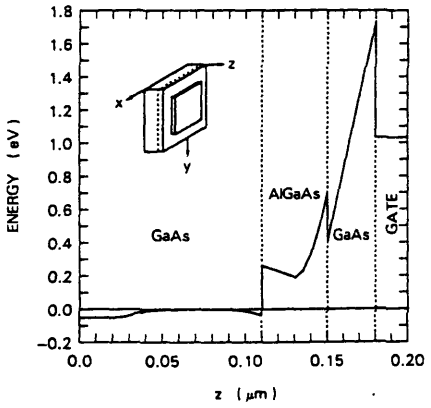


FIG. 1. Conduction-band edge along a vertical line through the center of the GaAs-$Al_xGa_{1-x}As$ structure considered here, for a gate voltage of -1.03 V. The layers of the structure, from left to right, are 30 nm of n^+-GaAs, 80 nm of undoped GaAs, 20 nm of undoped $Al_{0.4}Ga_{0.6}As$, 20 nm of $Al_{0.4}Ga_{0.6}As$ with $N_D=1.5\times10^{18}$ cm^{-3}, and a 30 nm GaAs cap. The repeating unit is 500 nm square, and the GaAs cap layer is etched away, except under a central 300-nm square mesa. A metallic gate is then deposited over the structure. The Schottky barrier associated with the gate suppresses induced charge in the GaAs, except under the central portion of the mesa. All calculated results are for 4.2 K and the zero of energy is taken at the Fermi level in the substrate.

For the GaAs, we use an electron effective mass of $0.07m_0$ and a dielectric constant of 13; for the $Al_xGa_{1-x}As$, we use $0.11m_0$ and 11.8, respectively, corresponding to an AlAs mole fraction $x=0.4$. The conduction-band offset is taken to be 0.3 eV. The binding energy of the deep donor in the $Al_{0.4}Ga_{0.6}As$ is taken to be 0.15 eV, and the effective Schottky-barrier heights of the gate electrode to the GaAs and the $Al_{0.4}Ga_{0.6}As$ are taken to be 0.7 and 0.95 eV, respectively.

Although the present calculations deal with the structure of Hansen *et al.*,[6] the methods to be described in this paper can be used for a wide class of structures in which three-dimensional confinement of electrons is achieved by a combination of band offsets and electrostatic means.

We solve the Schrödinger and Poisson equations self-consistently. Image effects[16] in the Schrödinger equation are ignored and we use the Hartree approximation, ignoring exchange and correlation effects. Bryant[17] showed that many-electron interactions can have significant quantitative and qualitative influence on the energy spectrum of a quantum dot with a small number of electrons. Similar effects are expected for the structures studied here, but have not been included in our calculation.

The electrostatic potential ϕ is governed by the Poisson equation

$$\nabla\cdot[\epsilon(x,y,z)\nabla\phi(x,y,z)]=-\rho(x,y,z)\ , \tag{1}$$

where ϵ is the permittivity (in the present case it depends only on the z coordinate), with boundary conditions determined by voltages applied at the contacts. At boundaries where there are no contacts, the normal derivative of the potential is taken to be zero. The total charge density ρ in Eq. (1) includes the charge in quantum states, calculated as described below, as well as the contribution from ionized impurities in the $Al_xGa_{1-x}As$, and of any electrons outside the Schrödinger domain. In particular, any electrons in the cap layer are treated classically.

In a magnetic field (B_x,B_y,B_z) the three-dimensional Schrödinger equation for the electron envelope function (in the effective-mass approximation) becomes

$$\sum_{j=1}^{3}\left[\frac{\hbar}{i}\frac{\partial}{\partial x_j}+eA_j\right]\frac{1}{2m_j}\left[\frac{\hbar}{i}\frac{\partial}{\partial x_j}+eA_j\right]\zeta_n +[U(x,y,z)-E_n]\zeta_n=0\ , \tag{2}$$

where m_j is the electron effective mass in the jth direction and the electron charge is $-e$. We choose the symmetric gauge

$$A_x=(B_yz-B_zy)/2\ , \tag{3}$$

and cyclic permutations. In the present case $m_x=m_y=m_z$ and $B_x=B_y=0$.

The electron charge density in the quantum dot is

$$\rho_{\rm inv}(x,y,z)=-2e\sum_n\zeta_n^*(x,y,z)\zeta_n(x,y,z) \times f((E_{qF}-E_n)/k_BT))\ , \tag{4}$$

where the sum is over all states n, the factor 2 is for spin degeneracy (spin splitting is ignored in this calculation), E_{qF} is the quasi-Fermi-energy, and f is the Fermi-Dirac occupation function at temperature T.

If the Fermi energy in the quantum dot is equal to the Fermi energy in the n-type substrate, then the calculated charge in the dot will be a continuous function of the gate voltage. The charge per quantum dot will, in general, be a nonintegral multiple of the electron charge, and will represent the average for a large ensemble of dots. Physically, however, the charge in an isolated dot should be an integral multiple of the electron charge. If we constrain the charge in the dot to be an integral multiple of the electron charge, then we apply Fermi-Dirac statistics to determine the quasi-Fermi-level that gives the prescribed charge from the calculated energy levels.

Among many simplifying assumptions in our calculation is the neglect of the interface image potential and of many-electron contributions to the potential. Then the potential energy is $U=-e\phi+\Delta E_c$, where the second term is the position-dependent conduction-band offset relative to the bottom of the conduction band in the GaAs. Level broadening has not been included explicitly, but some broadening, small compared to typical level spacings, is simulated because we carry out the calculations at $T=4.2$ K.

Both the Poisson and Schrödinger equations are cast into discrete form on a nonuniformly graded, tensor-product (finite-difference) mesh, with no interior mesh-line terminations,[18] and the resultant matrix equations are solved numerically. The Schrödinger mesh includes only the region of significant dot charge; elsewhere electrons are treated semiclassically. Electrostatic potential, envelope functions, and charge-density values are defined at mesh nodes, whereas material properties such as dielectric constant, effective mass, and effective band-edge shift ΔE_c are piecewise constant in the individual rectangular parallelepiped elements defined by the mesh. Equations for the potential and envelope function at each node are obtained by integrating Eqs. (1) and (2) over the box defined by the six planes bisecting the lines connecting the node to its nearest neighbors (for nodes on the boundary, only the volume inside the boundary is included). For the Poisson equation, this results in a real symmetric matrix problem $\bar{L}\phi=-\Omega\rho$, where $\bar{L}$ is the operator $\nabla\cdot\epsilon\nabla$ integrated over the boxes, ϕ and ρ are vectors of the nodal potentials and charge densities, and Ω is a diagonal matrix of the nodal box volumes. For the Schrödinger equation, one similarly obtains $\bar{H}\zeta_n=E_n\Omega\zeta_n$, where $\bar{H}$ is the Hamiltonian integrated over the boxes and ζ_n is the complex vector of the envelope function for state n at each node. This equation is readily transformed into a standard matrix eigenvalue problem by premultiplying both sides by $\Omega^{-1/2}$ and substituting $\mathbf{I}=\Omega^{-1/2}\Omega^{1/2}$ to give $(\Omega^{-1/2}\bar{H}\Omega^{-1/2})(\Omega^{1/2}\zeta_n)=E_n(\Omega^{1/2}\zeta_n)$, or simply $\mathbf{H}\mathbf{y}_n=E_n\mathbf{y}_n$, where $\mathbf{H}=\Omega^{-1/2}\bar{H}\Omega^{-1/2}$ is still an Hermitian matrix and $\mathbf{y}_n=\Omega^{1/2}\zeta_n$. In zero magnetic field, the Hermitian matrix reverts to a real symmetric matrix.

The Poisson equation in discrete form is nonlinear, since the charge density depends on the potential. The solution to this nonlinear problem constitutes the search for self-consistency between the charge and the potential. The solution to the Schrödinger equation enters as part of the evaluation of the total charge density in the device, for a given potential. We linearize the Poisson equation via Newton's method. The vector ϕ which is the zero of the function $\mathbf{F}(\phi)=\bar{L}\phi+\Omega\rho(\phi)$ is sought by iterating

$$\mathbf{F}'(\phi^l)\bar{\phi}^l=-\mathbf{F}(\phi^l)\,, \tag{5a}$$

$$\phi^{l+1}=\phi^l+t\bar{\phi}^l\,, \tag{5b}$$

until convergence is obtained. Here, l is the iteration index and the scalar damping factor t is selected according to a modified Bank-Rose damping scheme as discussed in Ref. 19. The evaluation of the Jacobian matrix $\mathbf{F}'$ is a possible stumbling block because the dependence on ϕ of the charge density in the channel given in Eq. (4) is nonlocal, which would destroy the seven-diagonal structure of the Jacobian, rendering the matrix solution significantly more difficult. Instead, as discussed in Ref. 19, a rather crude approximation to the dependence of the channel charge on local potential is made for purposes of calculating $\mathbf{F}'$ only, in order to circumvent this difficulty. While this precludes a second-order convergence rate of the Newton iteration, converged solutions can still be obtained in an acceptable number of iterations. The linear matrix equations in (5a) above are solved via a conjugate-gradient method. Such methods require a preconditioner to accelerate convergence; we have selected a polynomial preconditioner,[20] as it has proven robust and highly vectorizable.

The discrete Schrödinger equation is solved by one of two methods. Far away from self-consistency between charge and potential, a Lanczos method is employed.[21] This method forms an approximate tridiagonalization $\mathbf{T}$ of the matrix $\mathbf{H}$. No reorthogonalization is used in this process. Then, the eigenvalues of $\mathbf{T}$ are found in a specified energy interval (from the minimum of the quantum dot potential to 5–10 meV above the Fermi energy) by a bisection search together with Sturm sequencing.[21] Care must be taken in discarding potentially "spurious" eigenvalues of $\mathbf{T}$, that is, eigenvalues of $\mathbf{T}$ which are not good approximations to true eigenvalues of $\mathbf{H}$.[21] Finally, inverse iteration is used to find the associated eigenvectors. Gaussian elimination is used to solve the tridiagonal matrix equations involved in inverse iteration.

Near self-consistency between the charge and the potential, a simple Rayleigh quotient-iteration algorithm[22] is used to solve the eigensystem. This algorithm requires an initial guess for the eigenfunctions, and can be summarized as follows. Let $\sigma(\mathbf{y})=(\mathbf{y}^H\mathbf{H}\mathbf{y})/(\mathbf{y}^H\mathbf{y})$ be the usual Rayleigh quotient (superscript H denotes Hermitian conjugate) and let $\mathbf{y}^0$ be an initial guess for the nth eigenfunction. Then solve

$$[\mathbf{H}-\sigma(\mathbf{y}^l)\mathbf{I}]\mathbf{x}^{l+1}=\mathbf{y}^l\,, \tag{6a}$$

$$\mathbf{y}^{l+1}=\mathbf{x}^{l+1}/\|\mathbf{x}^{l+1}\|\,, \tag{6b}$$

$$\mathbf{y}^{l+1}=\mathbf{y}^{l+1}-\sum_{i=0}^{n-1}(\mathbf{y}_i^H\mathbf{y}^{l+1})\mathbf{y}_i \quad \text{(this step only if } n>0)\,; \tag{6c}$$

if $\|\mathbf{H}\mathbf{y}^{l+1}-\sigma(\mathbf{y}^{l+1})\mathbf{y}^{l+1}\|<\epsilon$, then done . (6d)

The solution becomes $\mathbf{y}_n=\mathbf{y}^{l+1}$ and $E_n=\sigma(\mathbf{y}^{l+1})$.

Note that the step (6c) above is not a part of the standard Rayleigh quotient iteration. This step serves to remove components related to previously determined eigenfunctions $\mathbf{y}_i$, $i=0,1,2,\ldots,n-1$, from the vector $\mathbf{y}^{l+1}$ which is evolving into the eigenfunction $\mathbf{y}_n$. In practice, this orthogonalization step has an important benefit: In solving the equations for a series of gate voltages or magnetic fields, the time-consuming Lanczos method can be omitted from the iteration for self-consistency between charge and potential, provided the new solution is not too distant from the previous solution. The orthogonalization ensures that energy levels which are "close" at some initial solution do not converge to the same level at a later step. This procedure gives correct eigenstates using significantly less computation time, but may eventually miss some intermediate eigenvalues if extended over too wide a range of gate voltage or magnetic field without an intervening Lanczos solution.

The Hermitian matrix system in (6a) above is solved with the polynomial-preconditioned conjugate-gradient method.[20] Although the matrix $\mathbf{H}-\sigma\mathbf{I}$ in (6a) is not positive definite, this method of solution has always been robust for the class of problems we have encountered.

The boundary condition used in our numerical method is that the normal derivative of ζ_n vanish on the Schrödinger mesh boundary. In the lateral (x and y) directions this condition occurs sufficiently far from the region of induced charge that it has no appreciable effect on the results. In the direction normal to the inversion layer, we truncate the Schrödinger mesh 36 nm below the GaAs/$Al_xGa_{1-x}As$ interface to avoid the quasicontinuum of eigenstates arising from the heavily n-type doped substrate. This may lead to significant errors in the values of energy levels and thresholds. In particular, some of the qualitative results for the present structure may not apply to a dot with stronger vertical confinement, as could be obtained if a p-type substrate were used.

The convergence criterion for self-consistency is that the nodal potential energies of successive iterations differ by no more than 0.01 meV anywhere on the Poisson mesh. The necessarily limited mesh size ($51\times51\times35$ for the Poisson mesh and $43\times43\times18$ for the Schrödinger mesh, in the x, y, and z directions, respectively) and other approximations made in the calculation will lead to errors that are larger than this convergence criterion. An IBM 3090 computer with vector processor was used for these calculations. A single Newton's loop, in which the Poisson and Schrödinger equations are each solved once, required approximately 15 min of computation time for $B=0$ and 45 min for $B\neq0$ if the Lanczos recursion was used. If the Rayleigh quotient algorithm was used in place of the Lanczos method, the solution of the Schrödinger equation (nearly all the computation time) ranged from 5 to 50 times faster, depending on the quality of the initial guess. A typical bias point required 4–20 Newton's loops to converge.

III. ZERO MAGNETIC FIELD

Potential contours in a plane 8 nm below the GaAs-$Al_xGa_{1-x}As$ interface, near the maximum of the electron charge distribution, are shown in Fig. 2. Note that the potential contours are nearly circular, especially at the lower energies, although the defining gate geometry is a square. That follows from the attenuation of higher Fourier components of the potential in regions some distance from the gate, as found previously for fluctuations in the width of a gate opening.[23] Also, the effective size of the quantum dot, given by the contour at the Fermi level, is considerably smaller than the size of the defining structure in the gate. In Fig. 2 and throughout this paper we make cuts in representative planes or along representative lines to display functions of three spatial coordinates. The figures are intended to indicate the main features of the calculated results, but should not be considered to be complete. The raggedness of some of the later curves is a consequence of the necessarily coarse mesh used in the discretization.

Figure 3 shows the number of electrons in the quantum dot, the lowest-energy levels, and the quasi-Fermi-level as functions of the voltage on the gate at 4.2 K for zero magnetic field. The notion of "quasi-Fermi-level" does not arise in calculations for quasi-one-dimensional wires, for which the charge can be considered to vary continuously, provided a suitable means of equilibrating with an adjacent gate or contact exists. For the very small structures considered here, where a dot may contain only a

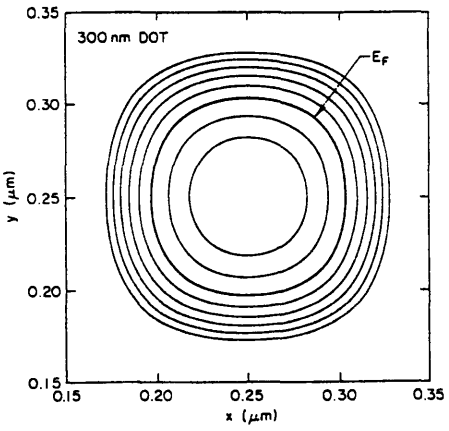


FIG. 2. Lateral potential contours in the plane 8 nm below the GaAs/$Al_xGa_{1-x}As$ interface, near the peak of the vertical charge density, for a gate voltage of -1.03 V. The innermost contour is 15 meV below the Fermi level, which is indicated by the heavy line, and the remaining contours are at 10-meV intervals from -10 to $+50$ meV. Note the nearly circular symmetry despite the square geometry of the cap. The effective quantum dot size, with a diameter of about 100 nm, is considerably smaller than the 300-nm square mesa in the GaAs cap layer.

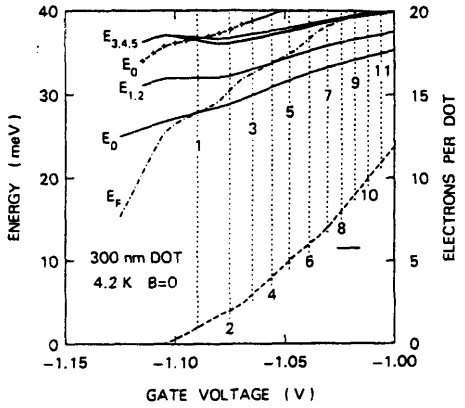


FIG. 3. Energy levels (solid lines) and Fermi level (dotted-dashed line) relative to the bottom of the potential well, and number of electrons in the quantum dot (dashed line), vs gate voltage. The second level is doubly degenerate, and the next three levels lie very close to each other. In addition, each level has a twofold spin degeneracy. These energies reflect the combined effect of vertical and lateral confinement. The energy of the lowest state with a node in the z direction is indicated by the plus signs. Only integer electron occupations, indicated by the vertical dotted lines, correspond to physically realizable states of an isolated quantum dot.

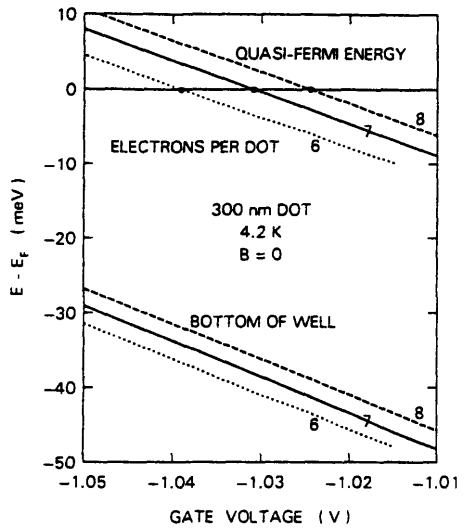


FIG. 4. Quasi-Fermi-level and energy of the bottom of the well, vs gate voltage for six, seven, and eight electrons in the quantum dot. The energy difference between the quasi-Fermi-level and the Fermi level gives a driving force for electrons to move between the dot and the substrate. The circles correspond to gate voltages for which the dot is in equilibrium with the substrate for an integer electron occupation.

few electrons, discontinuities can arise because transfer of just one electron can have a significant effect on the energies in the problem. This gives rise to the Coulomb blockade, as found in many experiments.[15] Because of this effect, the curves in Fig. 3 have no physical signficance for an isolated quantum dot at points where the number of electrons in the dot is different from an integer.

Some of the energy levels we calculate are degenerate (apart from the spin degeneracy, which applies to all levels in our calculation) and others are nearly so. For example, the second and third levels are exactly degenerate at zero magnetic field because of the square symmetry of the structure we consider. The fourth and fifth levels would be degenerate at $B=0$ if our system had circular symmetry. The small splitting results from the weak remnant of the square symmetry of the cap. Finally, the sixth level, which is close to the fourth and fifth, would be exactly degenerate with them if the system had circular symmetry and had the perfectly parabolic potential treated by Darwin.[7] Similar considerations apply for higher-lying levels, except that they are increasingly influenced by the deviations from a circularly symmetric potential.

Figure 4 shows the quasi-Fermi-energy and the bottom of the potential relative to the Fermi energy in the GaAs substrate when the number of electrons in the quantum dot is six, seven, or eight. At a given gate voltage, several different charge states of the dot are possible, although the state with the quasi-Fermi-level closest to the Fermi level in the substrate contact is the one most likely to be observed. The buildup of potential difference before a charge transfer occurs is a signal of the Coulomb blockade.[15]

One measure of capacitance of our structure is obtained by using the lower curve in Fig. 3 to calculate a gate-to-dot capacitance $C_g = dQ_{QD}/dV_g$. That capacitance varies from about 1×10^{-17} F at small values of dot charge to about 3×10^{-17} F when there are about 12 electrons per dot. More directly relevant to the experiment of Hansen *et al.*[6] is an effective substrate-to-dot capacitance C_s, which we obtain by dividing the electron charge by the vertical separation between successive quasi-Fermi-level curves in Fig. 4, to obtain a value of about 6×10^{-17} F for a dot occupation between seven and eight electrons. Both of these effective capacitances will increase with increasing dot charge.

The dynamical behavior of this system depends on charge-transfer rates between the quantum dot and adjacent electrodes, a problem which is outside the scope of the present static calculation.[15] Note that the barrier between the dot and the substrate is very small, as indicated in Fig. 1 for a line through the center of the dot. This barrier would have been larger had we used a larger value than 10^{14} cm^{-3} for the net acceptor doping in the nominally undoped GaAs.

Figure 5 shows a few of the lowest-energy levels versus gate voltage when the number of electrons in the quan-

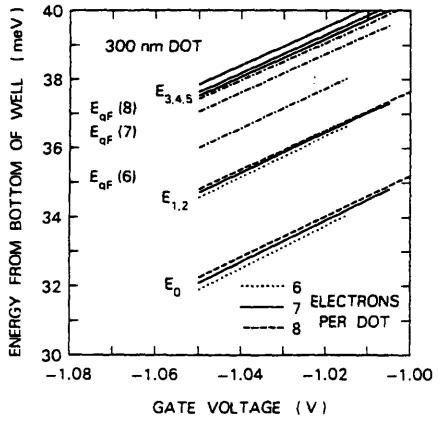


FIG. 5. The six lowest-energy levels (note that the second and third levels are degenerate), and the quasi-Fermi-energies, for six, seven, and eight electrons per quantum dot, vs gate voltage. Each level has a twofold spin degeneracy. For the upper three levels only the values for seven electrons per dot are shown; the results for six and eight electrons per dot almost coincide. These results are for 4.2 K and $B=0$.

tum dot is fixed at six, seven, or eight with zero magnetic field. The energy levels depend remarkably little on the charge state, but are quite sensitive to gate voltage.

Some details of potential and charge density are given in Figs. 6–8, both as functions of charge in the quantum dot at $B=0$ and as functions of magnetic field (as discussed in the next section) for fixed charge in the dot. Figure 6 shows the charge density along a vertical line through the center of the dot. The charge density peaks about 8 nm below the GaAs/$Al_xGa_{1-x}As$ interface, but is truncated—as described above—before the rise of charge density in the substrate begins. A lateral cut through the charge density near the peak in Fig. 6 is shown in Fig. 7. Finally, Fig. 8 shows the variation of the conduction-band edge in the x direction, in the same plane as in Fig. 7. The effective size of the dot is about 100 nm, considerably smaller than the 300-nm square mesa in the GaAs cap layer. The potential somewhat resembles the truncated parabola found previously for n-i-p-i doping superlattices[24] and for wires in Si (Ref. 11) and GaAs,[12] but with more structure, which can be attributed to the small number of discrete states that contribute to the charge in the cases shown.

IV. NONZERO MAGNETIC FIELDS

When a magnetic field is applied normal to the surface, the Schrödinger equation, Eq. (2), becomes complex, and

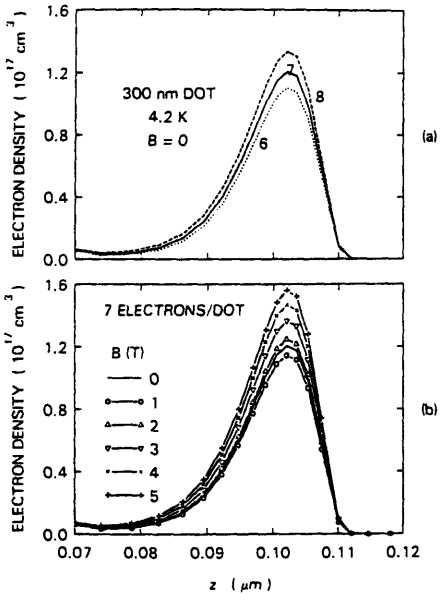


FIG. 6. Total charge density in the vertical direction along a line through the quantum dot center for (a) six, seven, and eight electrons per dot, with $B=0$, and (b) $B=0$, 1, 2, 3, 4, and 5 T, with seven electrons per dot. The z coordinate and the gate voltage are the same as in Fig. 1.

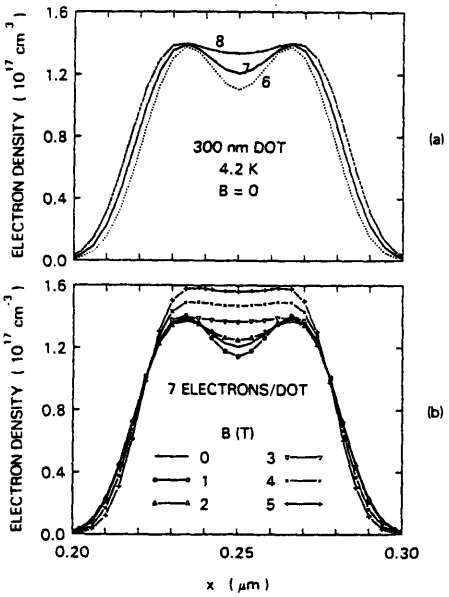


FIG. 7. Lateral cut of total charge density in a plane 8 nm below the GaAs/$Al_xGa_{1-x}As$ interface. The cut is taken through the center of the quantum dot. Other quantities as in Fig. 6.

TABLE I. Calculated quantities for states of the quantum dot at 4.2 K for $B = 5$ T and a gate voltage of -1.03 V, with seven electrons in the dot. The energy is relative to the bottom of the potential well in the dot, $\langle l_z \rangle$ is the expectation value of the z component of the "canonical" angular momentum $\mathbf{r} \times \mathbf{p}$, $\langle R \rangle$ is the expectation value of the two-dimensional radial distance from a vertical axis through the center of the dot, δR is its standard deviation, and $\langle L_z \rangle$ is the expectation value of the z component of the total angular momentum $\mathbf{r} \times m\mathbf{v}$. The 0' state is the lowest state with a node in the z direction.

State	Energy (meV)	$\langle l_z \rangle/\hbar$	$\langle R \rangle$ (nm)	δR (nm)	$\langle L_z \rangle/\hbar$
0	35.4	0.07	14.3	7.5	1.06
1	35.5	−1.05	21.2	7.5	0.87
2	35.7	−1.98	25.9	7.3	0.76
3	36.2	−2.91	29.4	7.1	0.57
4	37.0	−3.83	32.3	6.9	0.32
5	38.1	−4.75	34.8	6.8	0.03
6	39.5	−5.62	37.1	6.6	−0.24
7	41.1	−6.46	39.1	6.6	−0.49
8	42.9	−7.32	41.0	6.5	−0.78
9	43.6	0.98	21.0	7.4	2.87
0'	43.7	0.00	15.9	9.4	1.30

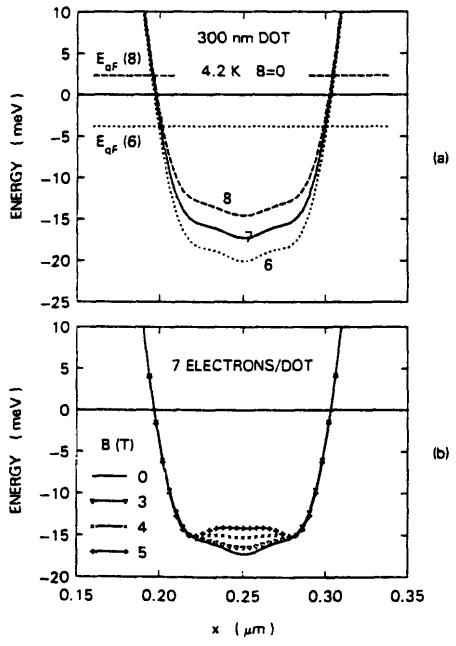


FIG. 8. Potentials along the same line as in Fig. 7. The Fermi energy is at zero. The quasi-Fermi-energies for six and eight electrons per quantum dot are indicated in (a). The quasi-Fermi-energy is within 1 meV of the Fermi energy for seven electrons per dot for the range of magnetic fields shown, and has been omitted. Also omitted in (b) are the curves for $B = 1$ and 2 T, which lie very close to the curve for $B = 0$.

its discretized form leads to an Hermitian matrix. In our case, this matrix has about 30 000 rows and columns, and a corresponding number of eigenstates, but we typically look for only the ~20 eigenstates with the lowest energy. Nevertheless the calculation, as described above, is very time consuming. We show in Fig. 9 the energy levels for the case of seven electrons per quantum dot, with a gate

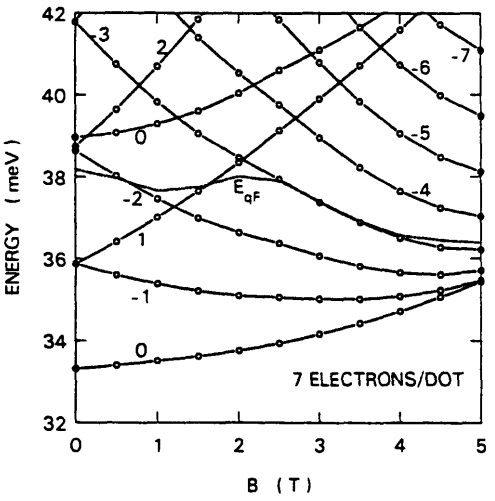


FIG. 9. Energy levels vs magnetic field for a quantum dot with seven electrons and a gate voltage of -1.03 V. The labels give *approximate* values of the z component of the canonical angular momentum $\mathbf{r} \times \mathbf{p}$ in units of $\hbar$.

voltage of -1.03 V. As shown in Fig. 2, the potential has nearly circular symmetry, and therefore angular momentum is approximately a good quantum number. The curves are labeled with an integer to represent the approximate z component of angular momentum (in units of $\hbar$), but the calculated expectation values for the points shown differ from an integer by up to 10%, and by less than 0.1 for the zero-angular-momentum states. These labels should therefore be considered to have only qualitative significance. At $B=0$, where the envelope eigenfunctions are real, the angular momentum is zero for all the states. The calculated curves are in good qualitative agreement with the results found by Darwin[7] for states in a two-dimensional harmonic-oscillator potential in a magnetic field. The curves are shown to cross, as would apply for states with different angular momentum in a circularly symmetric potential, although we expect that small anticrossing gaps would appear if the calculation were carried out with greater resolution. The difference between the crossing behavior in a circularly symmetric case and the anticrossing for positive-parity states in a rectangular box is nicely illustrated in Figs. 1 and 2 of the paper by Robnik.[8]

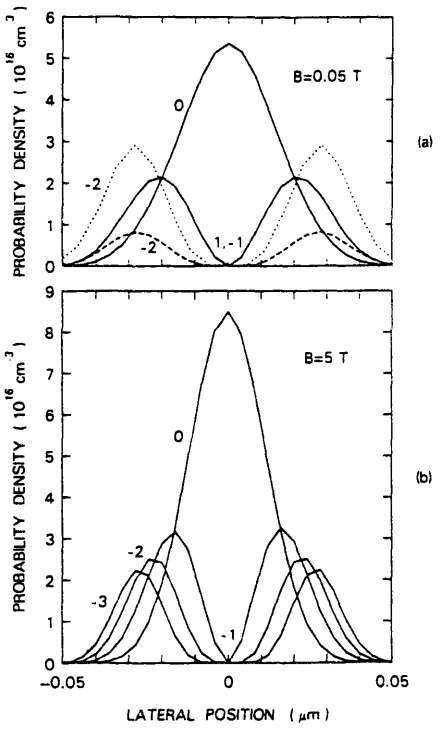


FIG. 10. Probability densities (absolute squares of the normalized envelope wave functions) for the four lowest eigenstates, in a plane 8 nm below the $GaAs/Al_xGa_{1-x}As$ interface. The gate voltage is -1.03 V and there are seven electrons in a quantum dot. Results are shown for (a) $B=0.05$ T and (b) $B=5$ T. The labels give *the approximate value of the* z component of the canonical angular momentum in units of $\hbar$. The states labeled $1,-1$ for $B=0.05$ T have probability densities which are almost the same (they correspond to opposite angular-momentum combinations of the x- and y-like degenerate solutions for $B=0$). The probability densities for the states shown in both (a) and (b) are approximately circularly symmetric, except for the state labeled -2 in (a), for which cuts along the x direction (dashed line) and along the diagonal $x=y$ (dotted line) are shown.

The angular momentum referred to in the preceding paragraph is what Van Vleck[25] has called the canonical angular momentum. It is the expectation value of $l=\mathbf{r}\times\mathbf{p}$, where p is the operator $-i\hbar\nabla$. The "true" angular momentum, $\mathbf{L}=\mathbf{r}\times m\mathbf{v}$, has an additional term[25,26] $(e/2)\mathbf{r}\times(\mathbf{B}\times\mathbf{r})$, analogous to the additional term in the Hamiltonian in the presence of a magnetic field. The angular-momentum quantum number associated with the z component of the canonical angular momentum is the integer l that appears in the angular factor $\exp(il\phi)$ in the wave function in a circularly symmetric potential.

Table I gives some additional information for the lowest states for $B=5$ T. We show the expectation value of the energy relative to the bottom of the well, of the z component of the canonical angular momentum, of $R=(x^2+y^2)^{1/2}$, with lateral position measured relative to a vertical axis through the center of the quantum dot, of $\delta R=(\langle R^2\rangle-\langle R\rangle^2)^{1/2}$, and of the z component of the "true" angular momentum, $\langle L_z\rangle=\langle l_z\rangle+(eB/2)\langle R^2\rangle$. The last state in the table is the lowest state with a node in the z direction.

The expectation value of the true angular momentum for a one-electron problem is related to the magnetic moment μ by[25] $\mu_z=-dE/dB=-(e/2m)\langle L_z\rangle$, where we assume the magnetic field to be in the z direction, as in the example treated in this paper. Our numerical results deviate somewhat from this relation, a difference which we attribute to the inclusion of the Hartree terms for the electron-electron interaction in the potential energy.

The energy levels in Fig. 9 are all associated with states that have no nodes in the z direction. States with such nodes, which would correspond to the first excited subband in a two-dimensional electron gas in an unpatterned GaAs heterojunction, appear at energies above 42 meV.

As already shown in Figs. 6–8, the character of the solution changes with increasing magnetic field. The radial wings of the charge density contract, with a corresponding increase in charge density near the center of the quantum dot and a change in the shape of the bottom of the potential well. The shape of the charge density of the four lowest states in a dot with seven electrons is shown in Fig. 10 for magnetic fields of 0.05 and 5 T. Even at 5 T, for which the magnetic length, $(\hbar/eB)^{1/2}=11$ nm, is considerably smaller than the effective dot radius, about 50 nm, a distinction between bulklike and edgelike states is not obvious from the charge densities or angular momenta of the occupied states. Note that spin splittings, which we have ignored, will become significant at the upper end of the magnetic field range that we use.

V. DISCUSSION

As noted earlier, many approximations have been made in these calculations. In particular, the substrate structure of the sample we have modeled required truncation of the Schrödinger mesh on a plane where the wave functions had not yet decayed to zero. There must be another, for the present not well understood, approximation in our description of the sample, because the calculated voltage threshold is about -1 V, while the measured threshold is about -0.2 V.[6] The measurements are made in the dark, and the calculations use a deep-donor binding energy consistent with that condition. The large discrepancy between calculated and measured threshold voltages may be due to changes in the properties of the top layers and of the interfaces caused by the processing steps used in defining the lateral sample geometry. The neglect of many-body interactions is also significant. We expect, however, that many of the qualitative results for the internal structure of the quantum dot remain valid.

We found that the energy-level structure can be considered to be a perturbation of the states of a parabolic potential in a magnetic field, with angular momentum a rough guide to the properties of the states. We also found, in contrast to our original expectations, that the energy levels measured from the bottom of the potential well are quite insensitive to the number of electrons in the quantum dot, for a fixed gate voltage. A weak dependence of level separations on electron population was obtained theoretically by Chaplik.[27] A number of authors have found theoretically that optical transitions for a parabolic potential in superlattices,[28] quantum wells,[29] wires,[30] and dots[31] reflect the underlying structure of the bare harmonic-oscillator potential and are unaffected by electron-electron interactions. Experiments on quantum wires[32] and quantum dots[2,4] are consistent with this result.

We have shown how the quasi-Fermi-level in the quantum dot depends on gate voltage for different charge states of the dot. As the gate voltage changes from a value corresponding to an integer electron occupation, the difference between the quasi-Fermi-level in the dot and in the adjacent substrate electrode increases, related to the Coulomb blockade. The gate voltage at which the charge changes discretely is not considered here. Finally, we gave some pictures of energy levels and wave functions, with approximate values of angular momentum, for a range of values of gate voltage, charge in the dot, and magnetic field.

At least one of the authors began this work expecting to find a clear qualitative distinction between bulklike and edgelike states. Our computed envelope wave functions do not show any abrupt qualitative differences, which can be considered to be a consequence of the rather soft potential at the walls of the quantum dot.

Note added in proof. Since completion of this work we have become aware of two related publications. The eigenfunctions and eigenvalues of the two-dimensional harmonic oscillator in a magnetic field were obtained by Fock[33] three years before the paper by Darwin.[7] Maksym and Chakraborty[34] have treated the energy levels of quantum dots with three and four electrons moving in a two-dimensional harmonic oscillator potential with an applied magnetic field, including effects of electron-electron interaction.

ACKNOWLEDGMENTS

We are indebted to Ralph Willoughby and Jane Cullum for access to the Lanczos eigenvalue program used here, to Mark Amidon, Len Borucki, Orest Bula, Steve Furkay, and Fred Pileggi for graphics and database support in earlier stages of our work, to Trey Smith for information about the samples and for discussions of the experiments and their interpretation, to Jose Brum for discussions of his calculations, and to Boris Alt'shuler, Pradip Bakshi, Gottfried Döhler, Alan Fowler, Wolfgang Hansen, Joe Imry, Khalid Ismail, Rolf Landauer, Uri Sivan, and Phil Stiles for helpful discussions and comments.

---

*Present address.

[1]T. P. Smith III, K. Y. Lee, C. M. Knoedler, J. M. Hong, and D. P. Kern, Phys. Rev. B **38**, 2172 (1988).

[2]C. T. Liu, K. Nakamura, D. C. Tsui, K. Ismail, D. A. Antoniadis, and H. I. Smith, Appl. Phys. Lett. **55**, 168 (1989).

[3]J. Alsmeier, E. Batke, and J. P. Kotthaus, Phys. Rev. B **41**, 1699 (1990).

[4]U. Sikorski and Ch. Merkt, Phys. Rev. Lett. **62**, 2164 (1989).

[5]H. Fang, R. Zeller, and P. J. Stiles, Appl. Phys. Lett. **55**, 1433 (1989).

[6]W. Hansen, T. P. Smith III, K. Y. Lee, J. A. Brum, C. Knoedler, D. Kern, and J. M. Hong, Phys. Rev. Lett. **62**, 2168 (1989).

[7]C. G. Darwin, Proc. Cambridge Philos. Soc. **27**, 86 (1931).

[8]M. Robnik, J. Phys. A **19**, 3619 (1986).

[9]J. A. Brum and G. Bastard, in *Science and Engineering of One- and Zero-Dimensional Semiconductors*, edited by S. P. Beaumont and C. M. Sotomayor-Torres (Plenum, New York, in press).

[10]U. Sivan and Y. Imry, Phys. Rev. Lett. **61**, 1001 (1988).

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

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

[13]K. Kojima, K. Mitsunaga, and K. Kyuma, Appl. Phys. Lett. **55**, 882 (1989).

[14]T. Kerkhoven, A. T. Galick, J. H. Arends, U. Ravaioli, and Y. Asad, J. Appl. Phys. (to be published).

[15]See, for example, R. Wilkins, E. Ben-Jacob, and R. C. Jaklevic, Phys. Rev. Lett. **63**, 801 (1989), and references therein to earlier work.

[16]The image potential associated with the different dielectric constants of GaAs and $Al_xGa_{1-x}As$ has been shown to have a very small effect on energy levels in a heterojunction [F. Stern and S. Das Sarma, Phys. Rev. B **30**, 840 (1984)] and is expected to have an insignificant effect on the results of the present calculation.

[17]G. Bryant, Phys. Rev. Lett. **59**, 1140 (1987).

[18]See, for example, G. E. Forsythe and W. R. Wasow, *Finite*

Difference Methods for Partial Differential Equations (Wiley, New York, 1960).

[19]S. E. Laux, in *Proceedings of the Fifth International Conference on the Numerical Analysis of Semiconductor Devices and Integrated Circuits (NASECODE V)*, edited by J. J. H. Miller (Boole, Dun Laoghaire, Ireland, 1987), pp. 270–275.

[20]O. G. Johnson, C. A Micchelli, and G. Paul, SIAM J. Numer. Anal. **20**, 362 (1983).

[21]J. K. Cullum and R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations* (Birkhäuser, Boston, 1985), Vols. I *(Theory)* and II *(Programs)*.

[22]B. N. Parlett, *The Symmetric Eigenvalue Problem* (Prentice-Hall, Englewood Cliffs, NJ, 1980).

[23]A. Kumar, S. E. Laux, and F. Stern, Appl. Phys. Lett. **54**, 1270 (1989).

[24]P. Ruden and G. H. Döhler, Phys. Rev. B **27**, 3538 (1983).

[25]J. H. Van Vleck, *The Theory of Electric and Magnetic Susceptibilities* (Oxford University Press, London, 1932), Secs. 7 and 36.

[26]H. L. Zhao, Y. Zhu, and S. Feng, Phys. Rev. B **40**, 8107 (1989).

[27]A. V. Chaplik, Pis'ma Zh. Eksp. Teor. Fiz. **50**, 38 (1989) [JETP Lett. **50**, 44 (1989)].

[28]P. Ruden and G. H. Döhler, Phys. Rev. B **27**, 3547 (1983).

[29]L. Brey, N. F. Johnson, and B. I. Halperin, Phys. Rev. B **40**, 10 647 (1989).

[30]V. B. Shikin, T. Demel, and D. Heitmann, Zh. Eksp. Teor. Fiz. **96**, 1406 (1989) [Sov. Phys.—JETP **69**, 797 (1989)].

[31]K. Kempa, D. A. Broido, and P. Bakshi, Bull. Am. Phys. Soc. **35**, 768 (1990); P. Bakshi, D. A. Broido, and K. Kempa, Phys. Rev. B (to be published).

[32]W. Hansen, M. Horst, J. P. Kotthaus, U. Merkt, Ch. Sikorski, and K. Ploog, Phys. Rev. Lett. **58**, 2586 (1987).

[33]V. Fock, Z. Phys. **47**, 446 (1928).

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

Surface Science 263 (1992) 335–340
North-Holland

surface science

Self-consistent calculations on confined electrons in three-dimensional geometries

Arvind Kumar

Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 22 May 1991; accepted for publication 26 August 1991

We review modeling of structures with confined geometries and describe our numerical solutions of the coupled Poisson and Schrödinger equations in three spatial dimensions. We discuss the computational issues involved and present results of modeling on three structures: confining potential of a quantum dot, charge density in a quasi-periodic structure, and gate capacitance of a "Coulomb island" of electrons.

1. Introduction

The physics of quantum-effect device structures can be better understood by modeling the electrostatic potential, charge density, and quantum energy levels resulting from confinement. To find the single-electron states due to confinement, the electrostatic potential governed by the Poisson equation must be solved self-consistently with the quantum charge obtained from the Schrödinger equation for the electron envelope function. This paper first reviews approaches to modeling quantum-effect device structures and then describes our numerical solutions of the coupled Poisson and Schrödinger equations in three spatial dimensions.

Purely analytic approaches to modeling quantum structures have been followed by Davies [1], who calculated the electron states of a narrow wire near threshold; by Sherwin and Drummond [2], who self-consistently treated GaAs/AlGaAs wires with cylindrical symmetry; and by Chaplik [3], who modeled multi-electron quantum dots in the Thomas–Fermi approximation. Another common approach involves solving the Schrödinger equation numerically in a fixed model potential so that self-consistency is not required. Bryant [4] has studied the effects of electron–electron interactions in a hard-walled quantum box, Maksym and Chakraborty [5] have studied interacting electrons in a parabolic quantum dot in a magnetic field, and Lent [6] has considered edge state currents in a circular dot in a magnetic field, to cite a few examples.

The advantage of a self-consistent numerical approach is the flexibility to specify arbitrary layer structures, device geometries, and bias conditions. Self-consistent numerical solutions of the Poisson and Schrödinger equations were first carried out by Laux et al., who calculated electron states in narrow Si [7] and GaAs [8] wires. A two-dimensional solver employing a number of numerically efficient algorithms has been implemented by Kerkhoven et al. [9]. Two-dimensional solvers have also been implemented by Smoliner et al. [10] to study etched quantum wires and by Stopa [11] to study planar triangular quantum dots. Because of much more demanding computational requirements, however, self-consistent calculations on three-dimensional structures have been reported only recently [12].

After summarizing in section 2 the numerical methods used to solve the self-consistent problem, we present in section 3 three examples of

structures modeled using our numerical solver. We first consider the quantum dot used in the experiments of Smith et al. [13] and Hansen et al. [14]. In an earlier paper [12] we found that the energy levels depend very weakly on the number of electrons in the dot; here we examine the effect of electron number on the size of the confining potential. As a second example, we calculate the electron density in the structure of Haug et al. [15], where a gate is used to impose a quasi-periodic potential. Finally, we model (semiclassically) a structure fabricated by Meirav et al. [16], in which transport through a constricted "island" of electrons is regulated by Coulomb charging.

2. Numerical solution of self-consistent problem

Details of the numerical methods used are presented in refs. [12.17]; we present only a summary here. Both the Poisson and Schrödinger equations are discretized using a seven-point finite-difference approximation with a box-integration scheme. For the Poisson equation, we use Newton's method to solve $F(\phi) = -\nabla \cdot (\epsilon \nabla \phi) - \rho(\phi) = 0$ iteratively, where $\phi(x, y, z)$ is the electrostatic potential, ϵ is the dielectric permittivity (taken to be piecewise constant), and the charge density $\rho(\phi)$ includes contributions from electrons, holes, and ionized dopants. After linearization we obtain

$$F'(\phi^{(n)})\delta\phi^{(n+1)} = -F(\phi^{(n)}). \qquad (1)$$

where $\delta\phi^{(n+1)} = \phi^{(n+1)} - \phi^{(n)}$ is the correction to the nth guess for the potential [18]. Eq. (1) becomes a linear matrix equation which we solve using an inverse iteration conjugate gradient method employing polynomial preconditioning [19], which is well suited for sparse matrices.

Evaluation of the Jacobian matrix $F'(\phi)$ in eq. (1) is a possible stumbling block because the quantum charge contributing to $\rho(\phi)$ depends *nonlocally* on the potential ϕ. To avoid the nonlocality in differentiating $\rho(\phi)$ as well as the resulting destruction of our sparse seven-diagonal matrix structure, we make the crude approximation of treating the charge semiclassically to evaluate the Jacobian matrix only, as discussed in ref. [17]. One of two schemes [20,21] is used in conjunction with the Newton's iteration to accelerate convergence, but the approximation made in the Jacobian precludes the quadratic convergence rate normally achieved with Newton's method. An alternate method of dealing with this difficulty is described in ref. [9].

After each correction to the potential is found, the quantum charge contributing to $\rho(\phi)$ must be evaluated. The electron charge is evaluated quantum-mechanically only inside the Schrödinger mesh, which is chosen to extend over the region of significant inversion charge. In the Schrödinger domain the electron density is given by $n_{\text{inv}}(x, y, z) = 2\Sigma_l |\zeta_l(x, y, z)|^2 f(E_l)$, where $f(E_l)$ is the temperature-dependent occupancy of the twofold spin degenerate level at energy E_l. The envelope wavefunctions are found by solving the effective-mass Schrödinger equation

$$\sum_{j=1}^{3} \left(\frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial x_j}\right) \frac{1}{2m_j} \left(\frac{\hbar}{\mathrm{i}} \frac{\partial}{\partial x_j}\right) \zeta_l(x, y, z) + [U(x, y, z) - E_l]\zeta_l(x, y, z) = 0, \qquad (2)$$

where m_j is the electron effective mass in the jth direction. We use the Hartree approximation so that the electron potential energy is given simply by $U(x, y, z) = -e\phi + \Delta E_c$, where the second term is the position-dependent conduction band offset.

After discretization, eq. (2) can be transformed into a real symmetric matrix eigenvalue problem which we solve by a Lanczos algorithm [22] far away from self-consistency between charge and potential and by a Rayleigh quotient algorithm [23] close to convergence. The Lanczos recursion method has the advantage of solving for only the few lowest eigenvalues of a very large matrix, while the Rayleigh quotient method is usually much faster but can only be used when a good initial guess to the eigensolution is available. Nearly all the computation time is spent in the eigensystem solvers.

Dirichlet boundary conditions are enforced at contacts. At mesh boundaries zero-gradient Neumann boundary conditions are used except in structures that repeat periodically along one di-

rection. In this case we apply the boundary conditions

$$\phi(x+a, y, z) = \phi(x, y, z), \quad (3)$$

$$\zeta_l(x+a, y, z) = \zeta_l(x, y, z) \exp(ik_x a). \quad (4)$$

for a structure periodic in the x direction with period a. We choose values for the wavevector k_x discretely, and a plot of energy eigenvalues obtained at each k_x yields the dispersion relation for motion along the periodic direction. These boundary conditions also introduce an additional off-diagonal element in the equations for nodes on the boundaries of the periodic direction, making the quantum eigenvalue problem Hermitian rather than real symmetric.

Many simplifying approximations have been made to keep the calculation at a tractable level. As discussed by Bryant [4], many-electron interactions strongly affect the energy spectrum of a quantum dot so that the extent to which the Hartree approximation can be used to treat interactions is unclear. Level broadening is not included explicitly, but some broadening is simulated because we usually carry out the calculations at a temperature of 4.2 K. Finally, the threshold voltages we calculate usually do not agree with those measured experimentally. We believe that this discrepancy may be due to non-equilibrium charge on deep donors when gate voltage is applied with the sample at low temperature, as well as to possible effects of processing damage.

3. Representative results

3.1. Confining potential of a quantum dot

We consider a structure modeled after the quantum dots used in the recent experiments of Smith et al. [13] and Hansen et al. [14]. As shown in the inset of fig. 1, a Schottky gate is used to create a small "puddle" of electrons underneath a 300 nm × 300 nm square of the GaAs cap layer which is left unetched. From our simulation we find the confining potential to be rounded, despite the square geometry of the gate. Moreover, the diameter of the potential well at the Fermi level is about 100 nm, considerably smaller than the 300 nm cap dimension.

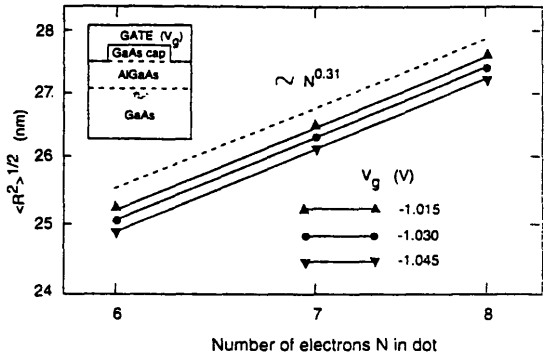


Fig. 1. Log–log plot of one measure of effective dot size, $\langle R^2 \rangle^{1/2}$, with number of electrons N in the dot, for three values of the gate voltage V_g. The structure used in the calculation is modeled after the quantum dots in the experiments of Smith et al. [13] and Hansen et al. [14]. A cross-section through the dot is shown in the inset.

In our model we constrain the number of electrons in the dot to be an integer and sweep the gate voltage, solving self-consistently for the energy levels and quasi-Fermi level corresponding to the dot occupancy. For a given gate voltage, more than one charge state of the dot is possible. In an earlier work [12] we found that the energy levels relative to the bottom of the potential well are insensitive to the number of electrons for dot occupancies from six to eight electrons as gate voltage is swept. This weak dependence may be roughly explained by noting that the effective dot size increases slightly with the addition of each new electron. To measure the dot size we calculate the expectation value

$$\langle R^2 \rangle = \frac{\int n_{\mathrm{inv}}(x, y, z) R^2 \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z}{\int n_{\mathrm{inv}}(x, y, z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z}, \quad (5)$$

where $R = (x^2 + y^2)^{1/2}$ is the radial distance measured from the center (z) axis of the dot. In fig. 1 we plot $\langle R^2 \rangle^{1/2}$ as a function of the number of electrons N in the dot for three values of the gate voltage V_g. We find that the dot radius increases as $N^{0.31 \pm 0.1}$ with the number of electrons in the dot. Our results are in good agree-

ment with the theoretical calculation by Chaplik [3], who also found the energy levels to be nearly independent of dot occupancy and obtained an $N^{1/3}$ dependence of radius on electron number for parameters like those we consider.

3.2. Charge density in a quasi-periodic structure

We can also model quasi-periodic structures such as the device fabricated by Haug et al. [15]. In this structure the two Schottky gates depicted at the top of fig. 2 impose a periodic potential over some finite range in x. We specify only a single 0.1 μm period of the structure and apply the boundary conditions of eqs. (3) and (4). To reduce computation time, we first find a converged solution using a small number of discrete values for k_x and then use this as an initial guess for a more accurate solution.

We show in fig. 2 contours of the electron density integrated perpendicular to the interface for voltages on the two confining gates of (a) $V_g = -0.25$ V and (b) $V_g = -0.40$ V. Because of uncertainties in Schottky barrier height and other assumptions of the calculation, the gate voltages cannot be compared directly to the experimental values, and have only relative significance. For a gate voltage of $V_g = -0.25$ V, there is an appreciable penetration of electrons into the gate opening. As the voltage on the confining gates is made more negative to $V_g = -0.40$ V, the electron channel becomes nearly uniform, as in a quantum wire. Only for a very small range of gate voltages near threshold is it possible to achieve isolated pockets of electron charge. This inability of the electron channel to follow the rapid spatial variation of the patterned gate can be understood in terms of the attenuation of high wavevector Fourier components of the potential, as discussed in an earlier paper [24].

3.3. Gate capacitance of a "Coulomb island" of electrons

Recently Meirav et al. [16] have reported that the conductance of a structure with two Schottky

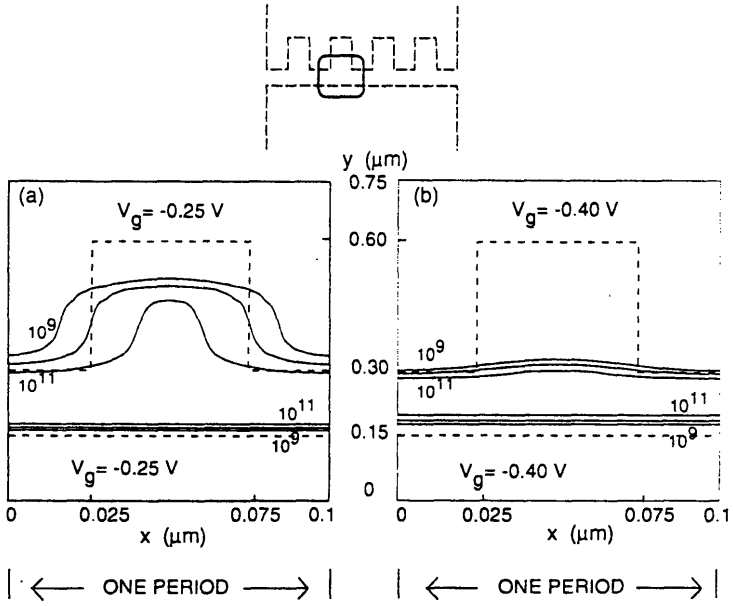


Fig. 2. Logarithmic contours of constant electron density (in cm^{-2}) integrated over the z-direction for gate voltages (a) $V_g = -0.25$ V and (b) $V_g = -0.40$ V in the quasi-periodic structure of Haug et al. [15]. The full gate structure is indicated at the top, with the single 0.1 μm period used in the simulation outlined. For clarity the length scales along the x- and y-directions are different. The dashed lines indicate the projection of the two confining Schottky gates. As the gate voltage becomes more negative in (b), the electron channel becomes uniform, as in a quantum wire.

gates as outlined in fig. 3 exhibits striking oscillations that are almost perfectly periodic in the voltage of a bottom gate used to control the electron density in the channel. We modeled this structure in the Thomas–Fermi approximation by using only the Poisson part of the solver. We expect that the semiclassical capacitance we calculate will differ little from that in a full quantum-mechanical treatment, as shown in ref. [25] for the case of a narrow wire. Including quantum mechanics would be very time-consuming and probably give little additional information when a large number of electron states are populated.

From our simulation a negative voltage on the two Schottky gates results in a "Coulomb island" of electrons between the constrictions, as shown in fig. 3. It is believed that a conductance maximum occurs when the electrostatic energy of the island remains the same if one more electron is added, whereas a conductance minimum occurs when the energy cost of adding another electron to the island is highest [26]. Our objective in modeling this structure is to compare the measured period of the oscillations to the calculated bottom gate voltage required to add a single electron to the island.

We calculated the integrated electron charge between the constrictions as a function of bottom gate voltage for three devices with different values of the barrier separation L. The variation of integrated charge with bottom gate voltage is approximately linear after more than about 50 electrons populate the island, as would be consistent with oscillations periodic in electron density. For samples with $L = 1$, 0.8, and 0.6 μm having successive conductance peaks separated experimentally by 0.6, 1.0, and 1.8 mV, respectively, we calculated the bottom gate voltage required to add a single electron to the island to be 0.65, 0.9, and 1.2 mV in the linear regime. The agreement is very good, especially for the first two samples, and would tend to support the theory that each conductance peak corresponds to the addition of a single electron to the island.

There are many interesting problems remaining in modeling quantum structures. Faster numerical algorithms would make the use of three-dimensional simulations more practical. The approximations made in these calculations can also be improved. Many-electron interactions are significant [4] and their influence on self-consistent calculations remains to be investigated. Another problem for future work is the inclusion of current flow, particularly with the discovery of the quantized constriction resistance and recent interest in quasi-ballistic transport. Finally, our

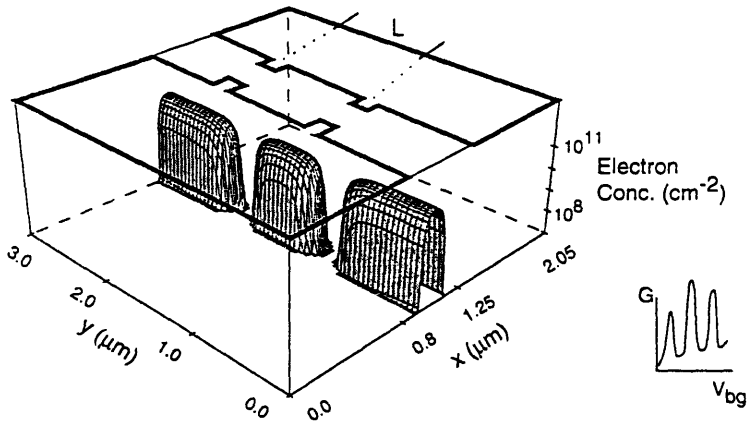


Fig. 3. Electron density integrated over the z-direction in the structure of Meirav et al. [16] (note the logarithmic scale). The two Schottky gates outlined in bold on the top surface are used to isolate the puddle of electrons in the middle, through which transport is regulated by Coulomb charging. For the case shown the bottom gate voltage is 50 mV higher than the voltage at which electrons first appear in the puddle. Typical oscillations of the conductance G in bottom gate voltage V_{bg} are shown at the right.

work has been limited to modeling structures in rectilinear coordinates; the use of a finite element method, for example, would allow more arbitrary geometries to be modeled.

Acknowledgments

I am deeply indebted to Frank Stern and Steve Laux for guidance and support throughout this work. I also wish to thank Dimitri Antoniadis, Terry Orlando, and Gee Rittenhouse for helpful discussions; Trey Smith, Rolf Haug, and Udi Meirav for information about their samples; Ralph Willoughby and Jane Cullum for their Lanczos program; and Ray Ghanbari for help in generating figures. I gratefully acknowledge use of the computer facilities at the IBM Thomas J. Watson Research Center and financial support from IBM, SRC, and AFOSR.

References

[1] J.H. Davies, Semicond. Sci. Technol. 3 (1988) 995.
[2] M.E. Sherwin and T.J. Drummond, J. Appl. Phys. 66 (1989) 5444.
[3] A.V. Chaplik, Pis'ma Zh. Eksp. Teor. Fiz. 50 (1989) 38 [JETP Lett. 50 (1989) 44].
[4] G.W. Bryant, Phys. Rev. Lett. 59 (1987) 1140.
[5] P.A. Maksym and T. Chakraborty, Phys. Rev. Lett. 65 (1990) 108.
[6] C.S. Lent, Phys. Rev. B 43 (1991) 4179.
[7] S.E. Laux and F. Stern, Appl. Phys. Lett. 59 (1986) 91.
[8] S.E. Laux, D.J. Frank and F. Stern, Surf. Sci. 196 (1988) 101.
[9] T. Kerkhoven, A.T Galick, U. Ravaioli, J.H. Arends and Y. Saad, J. Appl. Phys. 68 (1990) 3461.
[10] J. Smoliner, G. Berthold, F. Hirler and N. Reinacher, Semicond. Sci. Technol. 6 (1991) 642.
[11] M. Stopa, presented at the Fifth Int. Conf. on Modula Semiconductor Structures, Nara, Japan, 1991.
[12] A. Kumar, S.E. Laux and F. Stern, Phys. Rev. B 42 (1990) 5166.
[13] T.P. Smith III, K.Y. Lee, C.M. Knoedler, J.M. Hong and D.P. Kern, Phys. Rev. B 38 (1988) 2172.
[14] W. Hansen, T.P. Smith III, K.Y. Lee, J.A. Brum, C.M. Knoedler, J.M. Hong and D.P. Kern, Phys. Rev. Lett. 62 (1989) 2168.
[15] R. Haug, K.Y. Lee, T.P. Smith III and J.M. Hong, in: The Physics of Semiconductors, Vol. 3, Eds. E.M. Anastassakis and J.D. Joannopoulos (World Scientific, Singapore, 1990) p. 2443.
[16] U. Meirav, M.A. Kastner and S.J. Wind, Phys. Rev. Lett. 65 (1990) 771.
[17] S.E. Laux, in: Proc. Fifth Int. Conf. on the Numerical Analysis of Semiconductor Devices and Integrated Circuits (NASECODE V), Ed. J.J.H. Miller (Boole, Dun Laoghaire, Ireland, 1987) p. 270.
[18] The first guess to the potential $\phi^{(0)}$ can come from either a classical calculation or from the solution at a previous bias point.
[19] O.G. Johnson, C.A. Micchelli and G. Paul, SIAM J. Numer. Anal. 20 (1983) 362.
[20] R.E. Bank and D.J. Rose, SIAM J. Numer. Anal. 17 (1980) 806.
[21] R. Guerrieri, M. Rudan, G. Baccarani and P. Ciampolini, in: Proc. Fourth Int. Conf. on the Numerical Analysis of Semiconductor Devices and Integrated Circuits (NASECODE IV), Ed. J.J.H. Miller (Boole, Dun Laoghaire, Ireland, 1985) p. 293.
[22] J.K. Cullum and R.A. Willoughby, Lanczos Algorithms for Large Symmetric Eigenvalue Problems, Vols. 1, Theory, and 2, Programs (Birkhauser, Boston, 1985).
[23] B.N. Partlett, The Symmetric Eigenvalue Problem (Prentice-Hall, Englewood Cliffs, NJ, 1980).
[24] A. Kumar, S.E. Laux and F. Stern, Appl. Phys. Lett. 54 (1989) 1270.
[25] S.E. Laux and A.C. Warren, in: Technical Digest, 1986 IEDM, Los Angeles, CA (1986) p. 567.
[26] Y. Meir, N.S. Wingreen and P.A. Lee, Phys. Rev. Lett. 66 (1991) 3048;
C.W.J. Beenakker, Phys. Rev. B 44 (1991) 1646.

PHYSICAL REVIEW B VOLUME 48, NUMBER 7 15 AUGUST 1993-I

Effect of nonequilibrium deep donors in heterostructure modeling

Arvind Kumar
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Steven E. Laux, Frank Stern, A. Zaslavsky, J. M. Hong, and T. P. Smith III
IBM Research Division, Thomas J. Watson Research Center, Yorktown Heights, New York 10598
(Received 21 April 1993)

The deep donors (*DX* centers) which supply carriers in many GaAs-$Al_xGa_{1-x}As$ heterostructures are known to be metastable at low temperatures, maintaining a nonequilibrium state because of a microscopic barrier to recombination if ionized. Earlier modeling by three of us [Kumar, Laux, and Stern, Phys. Rev. B 42, 5166 (1990)] incorrectly assumed the donors in the $Al_xGa_{1-x}As$ to be in equilibrium with the electrons in the GaAs channel. We present data on the threshold voltage of heterostructures cooled under bias and results of calculations which assume that the deep donor charge is locked at the value attained at 100 K when the sample is cooled with a bias voltage applied. The revised calculations account for part of the discrepancy between the calculated threshold voltage for the quantum dot structure used in experiments by Hansen *et al.* [Phys. Rev. Lett. 62, 2168 (1989)] and the observed value. Part of the remaining discrepancy may be due to processing damage.

I. INTRODUCTION

Electrons are commonly introduced in GaAs-$Al_xGa_{1-x}As$ heterostructures by adding donors to the barrier material, usually with an undoped spacer layer separating the doped region from the GaAs region in which the carriers are found. Such donors, often called *DX* centers, are known to be metastable. If they are ionized, they do not return to the ground state at low temperatures because of a microscopic barrier that inhibits recombination. This field has a vast literature, and we refer only to a few articles that summarize relevant work.[1-3] If the donor doping density exceeds the density required to supply carriers to the channel and to compensate the charge associated with surface states, the remaining donors will be neutral.[4]

The density of electrons in the conducting channel is modulated by the gate voltage applied in an actual device structure, and therefore the charge on the deep donors might be expected to change with gate voltage if the donors were in equilibrium with the channel, as assumed in an earlier calculation by some of us.[5] Because of the nonequilibrium behavior of the deep donors, a more realistic model would fix the charge on the deep donors at the value attained at the temperature where they go out of equilibrium, of order 100–150 K for the situation considered here. Gate voltage changes applied when the sample is held at lower temperatures have little or no effect on the deep donor charge. The result can be a dramatic dependence of the low-temperature threshold gate voltage on the gate voltage applied as the sample is cooled.[6,7]

In this paper we present experimental results on a particular heterostructure sample to show the effect of cooling under different gate voltages on the threshold voltage for the appearance of charge in the channel. The measurements were taken both on the patterned quantum dot samples previously studied by Hansen *et al.*[8] (see also Smith *et al.*[9]) and on unpatterned control samples made from the same heterostructure material. We also present modeling results to show the effect of the nonequilibrium nature of the deep donors on electron confinement in the quantum dot.

II. THRESHOLD VOLTAGE FOR SAMPLES COOLED UNDER BIAS

The GaAs-$Al_{0.4}Ga_{0.6}As$ heterostructure material employed in these measurements is the same as used in earlier experiments[8,9] and simulations.[5] It is based on an n-type GaAs substrate layer with a net ionized donor concentration of 10^{18} cm^{-3}, an 80-nm layer of undoped GaAs (a background acceptor concentration of 10^{14} cm^{-3} is assumed), a 20-nm layer of undoped $Al_{0.4}Ga_{0.6}As$, a 20-nm layer of the same material with a donor concentration of 1.5×10^{18} cm^{-3}, and a 30-nm GaAs cap layer. On unpatterned samples a circular metal gate of 250-μm diameter was deposited directly on the cap layer. On patterned quantum dot samples the cap layer was etched away except in the central 300×300 nm^2 portion of a 500×500 nm^2 area, with the structure repeated on a square lattice over a 600×600 μm^2 field, which was subsequently covered with a metal gate. The samples were cooled from room temperature to $T=4.2$ K with a constant dc gate bias V_{gc} applied to the gate, $-0.3\le V_{gc}\le0.3$ V (higher values of $|V_{gc}|$ were avoided due to large leakage currents at room temperature). The gate voltage derivative dC/dV_g of the capacitance between the top metal gate and the substrate was then measured by superimposing a small signal $V_{ac}=10$ mV at a frequency $f=10$ kHz to the slowly swept gate bias and measuring the signal at $2f$ by lock-in detection. A typical dC/dV_g curve of an unpatterned sample cooled with the gate grounded ($V_{gc}=0$) is shown in the inset of Fig. 1, while the quantum dot sample traces were published previously.[8,9]

0163-1829/93/48(7)/4899(4)/$06.00

As the samples were cooled with a constant, nonzero gate bias V_{gc}, the overall dC/dV_g line shape did not change. However, as shown in Fig. 1, the capacitance threshold (defined as the gate voltage at which dC/dV_g peaks) shifted with V_{gc}. When the gate bias during cooling was positive, $V_{gc} \geq 0$, the threshold shifted approximately linearly with V_{gc} in both patterned and unpatterned samples, showing a weak tendency to saturate for $V_{gc} \geq 0.3$ V. For V_{gc} negative during cooling, the threshold shift of the patterned sample saturated for $V_{gc} \leq -0.1$ V (see Fig. 1), while the unpatterned sample's threshold continued to shift weakly with V_{gc}. In the gate voltage range where the patterned and unpatterned sample thresholds track each other, there is an overall shift of $\simeq 0.15$ V between the two. Finally, the absolute magnitude of the peak in dC/dV_g (see inset of Fig. 1) normalized to unit area was larger in unpatterned samples than in patterned samples by a factor of ~ 150.

III. MODELING RESULTS

Our original calculations,[5] done in the Hartree approximation, assumed the donors to be in equilibrium with the channel. As noted above, a more realistic model assumes that the donor charge is "locked" when the sample is cooled through the temperature range 100–150 K. We have modified our modeling program to record the donor charge distribution at equilibrium at 100 K (calculated classically) with the gate voltage equal to the gate voltage V_{gc} applied during cool down, and have used this locked donor charge distribution in our quantum-mechanical self-consistent calculations at 4.2 K.

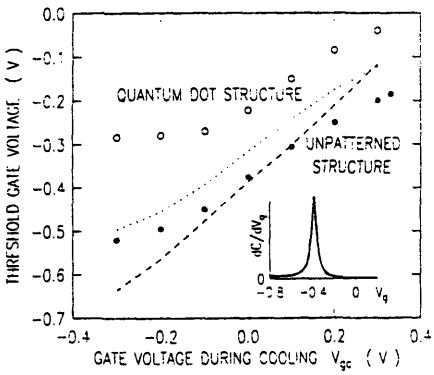


FIG. 1. Measured threshold voltage at 4.2 K for onset of channel charge in a heterostructure cooled with a gate voltage applied. The sample had an $Al_xGa_{1-x}As$ layer with AlAs fraction $x=0.4$ and nominal Si donor density of 1.5×10^{18} cm^{-3} separated from the GaAs channel by an undoped spacer layer 20 nm thick. The full circles are for unpatterned devices, and the open circles are for the quantum dot structures of Hansen *et al.* (Ref. 8). The dashed and dotted curves show the corresponding modeling results assuming the donors (with density taken to be 1×10^{18} cm^{-3} and effective thickness 23 nm) for the unpatterned and patterned structures, respectively, to be locked at the values attained at 100 K during cool down. The threshold densities for the unpatterned and patterned structures are 10^{10} electrons per cm^2 and $\frac{1}{6}$ electrons per dot, respectively. The corresponding calculated threshold voltages with the donors assumed to be in equilibrium with the channel electrons are -0.67 and -0.50 V, respectively. The inset shows a representative trace of the derivative of the capacitance with respect to gate voltage vs gate voltage. The measured gate voltage at the peak gives the ordinate for the plotted points.

Shown in Fig. 1 are calculated threshold values for the patterned and unpatterned cases. The deep-donor density used in these calculations is 1×10^{18} cm^{-3}, smaller than the nominal density of 1.5×10^{18} cm^{-3}. The threshold voltage in the unpatterned cases is set at a channel electron density of 1×10^{10} cm^{-2}, and the threshold in the patterned case is set at $\frac{1}{6}$ electron per dot, consistent with the 500-nm repeat distance of the structure and with the observed difference in peak height.

The trend of the calculated curves follows that of the measured points, although the saturation effects are not as pronounced. The difference in calculated thresholds between the patterned and unpatterned samples is smaller than the observed difference, but that may result from the somewhat arbitrary choice of threshold for the calculated curves. The calculated curves assume a Schottky barrier height of 0.7 eV for the gate on GaAs and 1 eV on $Al_{0.4}Ga_{0.6}As$, and the measured and calculated voltage scales may be offset somewhat.

Figure 2 shows potential contours at the Fermi level in a plane near the peak of the induced electron charge in the GaAs channel, calculated using the original model and the revised model with locked charge. In each case the dot charge is fixed at seven electrons, but the gate

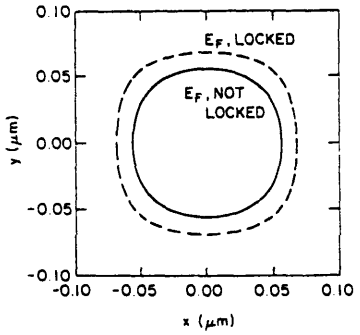


FIG. 2 Calculated contours showing the Fermi level in the plane 9.5 nm below the GaAs-$Al_xGa_{1-x}As$ interface, near the peak of the vertical electron density, for the quantum dot structure of Hansen *et al.* (Ref. 8). The dashed contour is for a gate voltage of -0.45 V when the deep donors are assumed to be locked in the charge state corresponding to zero gate voltage during cool down. The solid contour is for a gate voltage of -0.26 V when the deep donors are assumed to be in equilibrium with the channel electrons. The electron channel and the n^+ GaAs substrate are taken to be at ground throughout. Both cases correspond to a charge of seven electrons in the quantum dot.

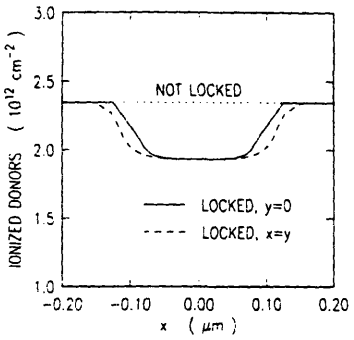


FIG. 3. Integrated sheet density of ionized donors, for the same conditions as in Fig. 2. The dotted curve, with all donors ionized, corresponds to donors in equilibrium with the channel electrons. The full and dashed curves correspond to donors locked at the charge they have at 100 K during cool down with zero gate voltage applied, and reflect horizontal and diagonal slices, respectively.

voltage is 0.19 V larger for the locked charge case because of the higher threshold voltage. Locking the charge leads to a larger dot and correspondingly larger gate-dot capacitance. Note that here, as in the earlier calculation, a small neck of charge connects the quantum dot to the n^- substrate, leading to some arbitrariness in the magnitude of the charge on the dot. Figure 3 shows the lateral variation of integrated sheet charge on the deep donors for the two cases of Fig. 2. Locking the charge reduces the effect of the donors in screening the gate potential, and therefore results in a potential that more closely reflects the square shape of the overlying structure.

IV. DISCUSSION

The discrepancy between calculated and measured threshold voltage in our earlier calculation for quantum wires[10] and quantum dots[5] was attributed to the effect of processing damage. There is voluminous literature showing that certain kinds of patterning, including the reactive ion etching used to fabricate the samples used in the experiments of Refs. 8 and 9, introduce significant changes in transport and other properties.[11–14] The effects of processing damage are not amenable to simple modeling but presumably affect the samples considered here and need further investigation.

Although calibrations give information about impurity concentrations in relation to growth conditions, some effects such as compensation and impurity motion during and after sample growth can complicate the doping structure in a given sample. Thus, there is not always definitive information about the donor concentration and its spatial distribution, which is an additional unknown in modeling. Our results suggest that the net donor concentration in the samples considered here is somewhat less than the nominal value.

Despite these uncertainties, the results presented here show that the nonequilibrium behavior of deep donors is one source for the difference between calculated and measured threshold voltages in the samples considered here and presumably in other samples. Careful comparison between experiments and realistic models for patterned and unpatterned structures, perhaps also including the effect of persistent light-induced changes in carrier concentration, should make it possible to narrow the uncertainties in sample parameters and attain satisfactory agreement between models and experiments.

We have considered a particular doping structure here. The extent to which nonequilibrium effects enter in other samples depends strongly on the excess donor concentration.

ACKNOWLEDGMENTS

One of us (F.S.) is indebted to T. N. Theis, P. M. Mooney, and K. Ensslin for discussions of properties of heterostructures with deep donors. He is also indebted to E. L. Hu, A. R. Long, P. M. Petroff, M. Rahman, M. Roukes, and C. D. W. Wilkinson for discussions of processing damage and to P. M. Mooney for comments on the manuscript.

[1]P. M. Mooney, J. Appl. Phys. 67, R1 (1990).
[2]P. M. Mooney and T. N. Theis, Comments Condens. Matter Phys. 16, 167 (1992).
[3]D. V. Lang, in *Deep Centers in Semiconductors*, 2nd ed., edited by S. T. Pantelides (Gordon and Breach, Yverdon, 1992), pp. 591–641.
[4]There is good evidence (see, for example, Ref. 2) that deep donors in $Al_xGa_{1-x}As$ are "negative-U" centers, preferring to be either positively or negatively charged rather than neutral, but we ignore that aspect here since it complicates the discussion without changing the qualitative aspects of the threshold behavior considered here.
[5]A. Kumar, S. E. Laux, and F. Stern, Phys. Rev. B 42, 5166 (1990).
[6]P. M. Mooney, P. M. Solomon, and T. N. Theis, in *Gallium Arsenide and Related Compounds 1984*, edited by B. de Cremoux, IOP Conf. Ser. No. 74 (Institute of Physics, and Physical Society, London, 1985), p. 617; P. M. Mooney, N. S. Caswell, and S. L. Wright, J. Appl. Phys. 62, 4786 (1987); T. N. Theis and B. D. Parker, Appl. Surf. Sci. 30, 52 (1991).
[7]C. Ghezzi, E. Gombia, and R. Mosca, Semicond. Sci. Technol. 6, B31 (1991).
[8]W. Hansen, T. P. Smith III, K. Y. Lee, J. A. Brum, C. Knoedler, D. Kern, and J. M. Hong, Phys. Rev. Lett. 62, 2168 (1989).
[9]T. P. Smith III, K. Y. Lee, C. M. Knoedler, J. M. Hong, and D. P. Kern, Phys. Rev. B 38, 2172 (1988).
[10]S. E. Laux, D. J. Frank, and F. Stern, Surf. Sci. 196, 101 (1988).
[11]H. F. Wong, D. L. Green, T. Y. Liu, D. G. Lishan, M. Bellis, E. L. Hu, P. M. Petroff, P. O. Holtz, and J. L. Merz, J. Vac. Sci. Technol. B 6, 1906 (1988).

[12]W. Beinstingl, R. Christanell, J. Smoliner, C. Wirner, E. Gornik, G. Weimann, and W. Schlapp, Appl. Phys. Lett. 57, 177 (1990).
[13]D. Lootens, P. Van Daele, P. Demeester, and P. Clauws, J. Appl. Phys. 70, 221 (1991).
[14]M. Rahman, N. P. Johnson, M. A. Foad, A. R. Long, M. C. Holland, and C. D. W. Wilkinson, Appl. Phys. Lett. 61, 2335 (1992).

Bibliography

[1] H.I. Smith and D.A. Antoniadis, "Seeking a Radically New Electronics," Tech. Rev. **93**, 26 (1990).

[2] R.T. Bate, Sci. Am. **258**, 78 (1988).

[3] K.K. Likharev, "Correlated Discrete Transfer of Single Electrons in Ultrasmall Tunnel Junctions," IBM J. Res. Dev. **32**, 144 (1988).

[4] R. Landauer, Phys. Today **42**, 119 (1989).

[5] For a review of localization phenomena, see A.B. Fowler, J.J. Wainer, and R.A. Webb, "Electronic Transport in Small Strongly Localized Structures," IBM J. Res. Dev. **32**, 372 (1988), and references therein.

[6] 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).

[7] D.A. Wharam, T.J. Thornton, R. Newbury, M. Pepper, H. Ahmed, J.E.F. Frost, D.G. Hasko, D.C. Peacock, D.A. Ritchie, and G.A.C. Jones, J. Phys. C **21**, L209 (1988).

[8] K.E. Ismail, P.F. Bagwell, T.P.Orlando, D.A. Antoniadis, and H.I. Smith, "Quantum Phenomena in Field-Effect-Controlled Semiconductor Nanostructures," Proc. IEEE **79**, 1106 (1991).

[9] J.H.F. Scott-Thomas, S.B. Field, M.A. Kastner, H.I. Smith, and D.A. Antoniadis, "Conductance Oscillations Periodic in the Density of a One-Dimensional Electron Gas," Phys. Rev. Lett. **62**, 583 (1989); S.B. Field, M.A. Kastner, U. Meirav, J.H.F. Scott-Thomas, D.A. Antoniadis, H.I. Smith, and S.J. Wind, "Conductance Oscillations Periodic in the Density of One-Dimensional Electron Gases," Phys. Rev. B **42**, 3523 (1990).

[10] H. van Houten and C. W. J. Beenakker, "Comment on 'Conductance Oscillations Periodic in the Density of a One-Dimensional Electron Gas,"' Phys. Rev. Lett. **63**, 1893 (1989).

[11] U. Meirav, M.A. Kastner, and S.J. Wind, "Single-Electron Charging and Periodic Conductance Resonances in GaAs Nanostructures," Phys. Rev. Lett. **65**, 771 (1990).

[12] M. Field, C.G. Smith, M. Pepper, D.A. Ritchie, J.E.F. Frost, G.A.C. Jones, D.G. Hasko, "Measurements of Coulomb Blockade with a Noninvasive Voltage Probe," Phys. Rev. Lett. **70**, 1311 (1992).

[13] C.W.J. Beenakker, "Theory of Coulomb-Blockade Oscillations in the Conductance of a Quantum Dot," Phys. Rev. B **44**, 1646 (1991).

[14] M.P. Stopa and Y. Tokura, "Capacitance in Single Electron Tunneling," in *Science and Technology of Mesoscopic Structures* (Springer-Verlag, Tokyo, 1992).

[15] 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, "Effect of Quantum Levels on Transport through a Coulomb Island," Phys. Rev. B **47**, 10020 (1993).

[16] E.B. Foxman, "Single Electron Charging and Quantum Effects in Semiconductor Nanostructures," PhD thesis, Massachusetts Institute of Technology (1993).

[17] Y. Meir, N.S. Wingreen, and P.A. Lee, "Transport through a Strongly Interacting Electron System: Theory of Periodic Conductance Oscillations," Phys. Rev. Lett. **66**, 3048 (1991).

[18] L.I.Glazman and V. Chandrasekhar, "Coulomb Blockade Oscillations in a Double-Dot System," Europhys. Lett. **19**, 623 (1992); I.M. Ruzin, V. Chandrasekhar, E.I. Levin, and L.I. Glazman, "Stochastic Coulomb Blockade in a Double-Dot System," Phys. Rev. B **45**, 13469 (1992).

[19] I.O. Kulik and R.I. Shekhter, "Kinetic Phenomena and Charge Discreteness in Granular Media," Sov. Phys. JETP **68**, 623 (1975).

[20] A. Kumar, S.E. Laux, and F. Stern, "Channel Sensitivity to Gate Roughness in a Split-Gate GaAs-AlGaAs Heterostructure," Appl. Phys. Lett. **54**, 1270 (1989).

[21] A. Kumar, S.E. Laux, and F. Stern, "Electron States in a GaAs Quantum Dot in a Magnetic Field," Phys. Rev. B **42**, 5166 (1990).

[22] A. Kumar, "Self-Consistent Calculations on Confined Electrons in Three-Dimensional Geometries," Surf. Sci. **263**, 335 (1992).

[23] A. Kumar, "Electron States and Potentials in Quantum Dot Structures," APS Bull. **37**, 429 (1992).

[24] U. Meirav, P.L. McEuen, M.A. Kastner, E.B. Foxman, A. Kumar, and S.J. Wind, "Conductance Oscillations and Transport Spectroscopy of a Quantum Dot," Z. Phys. B **85**, 357 (1991).

[25] M.A. Kastner, S.B. Field, U. Meirav, J.H.F. Scott-Thomas, D.A. Antoniadis, and H.I. Smith, "Comment on 'Conductance Oscillations Periodic in the Density of a One-Dimensional Electron Gas,"' Phys. Rev. Lett. **63**, 1894 (1989).

[26] F.M. de Aguiar and D.A. Wharam, "Transport through One-Dimensional Channels," Phys. Rev. B **43**, 9984 (1991).

[27] W. Hansen, T.P. Smith III, K.Y. Lee, J.A. Brum, C. M. Knoedler, J.M. Hong, and D.P. Kern, "Zeeman Bifurcation of Quantum-Dot Spectra," Phys. Rev. Lett. **62**, 2168 (1989); T.P. Smith III, K.Y. Lee, C.M. Knoedler, J.M. Hong, and D.P. Kern, "Electronic Spectroscopy of Zero-Dimensional Systems," Phys. Rev. B **38**, 2172 (1988).

[28] M. Stopa, "Coulomb Oscillation Amplitudes and Semiconductor Quantum-Dot Self-Consistent Level Structure," Phys. Rev. B **48**, 18340 (1993).

[29] N.W. Ashcroft and N.D. Mermin, *Solid State Physics*, Holt, Rinehart, and Winston, Philadelphia (1976).

[30] G.W. Bryant, "Electronic Structure of Ultrasmall Quantum-Well Boxes," Phys. Rev. Lett. **59**, 1140 (1987).

[31] P.A. Maksym and T. Chakraborty, "Quantum Dots in a Magnetic Field: Role of Electron-Electron Interactions," Phys. Rev. Lett. **65**, 108 (1990).

[32] U. Merkt, J. Huser, and M. Wagner, "Energy Spectra of Two Electrons in a Harmonic Quantum Dot," Phys. Rev. B **43**, 7320 (1991).

[33] N.F. Johnson and M.C. Payne, "Many-Body Effects in Resonant Tunneling through Quantum Dots," Phys. Rev. B **45**, 3819 (1992).

[34] D. Pfannkuche, V. Gudmundsson, and P.A. Maksym, "Comparison of a Hartree, Hartree-Fock, and an Exact Treatment of Quantum-Dot Helium," Phys. Rev. B **47**, 2244 (1993).

[35] F. Stern and S.D. Sarma, "Electron Energy Levels in GaAs-$Ga_{1-x}Al_xAs$ Heterojunctions," Phys. Rev. B **30**, 840 (1984).

[36] U. Meirav, M.A. Kastner, M. Heiblum, and S.J. Wind, "One-Dimensional Electron Gas in GaAs: Periodic Conductance Oscillations as a Function of Density," Phys. Rev. B **40**, 5871 (1989).

[37] A.A.M. Staring, H. van Houten, C.W.J. Beenakker, and C.T. Foxon, in *High Magnetic Fields in Semiconductor Physics III*, ed. by G. Landwehr, Springer, Berlin (1991).

[38] C. de Graaf, J. Caro, S, Radelaar, V. Lauer, and K. Heyers, "Coulomb Blockade Oscillations in the Conductance of a Silicon Metal-Oxide-Semiconductor Field-Effect-Transistor Point Contact," Phys. Rev. B **44**, 9072 (1991).

[39] V. Chandrasekhar, Z. Ovadyahu, and R.A. Webb, "Single-Electron Charging Effects in Insulating Wires," Phys. Rev. Lett. **67**, 2862 (1991).

[40] J. Weis, R.J. Haug, K. von Klitzing, and K. Ploog, "Magnetotransport Investigations of a Quantum Dot with a Small Number of Electrons," Physica B **189**, 111 (1993); J. Weis, R.J. Haug, K. von Klitzing, and K. Ploog, Phys. Rev. B **46** 12837 (1992).

[41] H. van Houten, C.W.J. Beenakker, and A.A.M. Staring, "Coulomb-Blockade Oscillations in Semiconductor Nanostructures," in *Single Charge Tunneling*, ed. by H. Grabert and M.H. Devoret, Plenum, New York (1991).

[42] M.A. Kastner, "Artificial Atoms," Phys. Today **46**, 24 (1993).

[43] C.S. Lent, "Edge States in a Circular Quantum Dot," Phys. Rev. B **43**, 4179 (1991).

[44] P.L. McEuen, E.B. Foxman , U. Meirav, M.A. Kastner, Y. Meir, N.S. Wingreen, and S.J. Wind, "Transport Spectroscopy of a Coulomb Island in the Quantum Hall Regime," Phys. Rev. Lett. **66**, 1926 (1991).

[45] P.L. McEuen, E.B. Foxman. J. Kinaret, U. Meirav, M.A. Kastner, N.S. Wingreen, and S.J. Wind, "Electron-Electron Interactions and the Addition Spectrum of a Coulomb Island in the Quantum Hall Regime," Phys. Rev. B **45**, 11419 (1992).

[46] U. Meirav, "Single-Electron Charging and Periodic Conductance Oscillations in Gallium Arsenide Nanostructures," PhD thesis, Massachusetts Institute of Technology (1990).

[47] U. Meirav, M. Heiblum, and F. Stern, "High-Mobility Variable-Density Two-Dimensional Electron Gas in Inverted GaAs-AlGaAs Heterojunctions," Appl. Phys. Lett. **52**, 1268 (1988).

[48] L.P. Kouwenhoven, N.C. van der Waart, A.T. Johnson, W. Kool, C.J.P.M. Harmans, J.G. Williamson, A.A.M. Staring, and C.T. Foxon, "Single Electron Charging Effects in Semiconductor Quantum Dots," Z. Phys. B **85**, 367 (1991).

[49] C. Pasquier, U. Meirav, F.I.B. Williams, D.C. Glattli, Y. Jin, and B. Etienne, " Quantum Limitation on Coulomb Blockade Observed in a 2D Electron System," Phys. Rev. Lett. **70**, 69 (1993);

[50] L.P. Kouwenhoven, A.T. Johnson, N.C. van der Waart, C.J.P.M. Harmans, and C.T. Foxon, "Quantized Current in a Quantum-Dot Turnstile Using Oscillating Barriers," Phys. Rev. Lett. **67**, 1626 (1991).

[51] D.V. Averin and Yu. V. Nazarov, Phys. Rev. Lett. **65**, 2446 (1990).

[52] D.C. Glattli, "Coulomb Blockade and Off-Resonance Tunneling in Small Electronic Systems," Physica B **189**, 88 (1993).

[53] A.T. Johnson, L.P. Kouwenhoven, W. de Jong, N.C. van der Waart, C.J.P.M. Harmans, and C.T. Foxon, "Zero-Dimensional States asnd Single Electron Charging in Quantum Dots," Phys. Rev. Lett. **69**, 1592 (1992).

[54] B.W. Alphenaar, A.A.M. Staring, H. van Houten, M.A.A. Mabesoone, O.J.A. Buyk, and C.T. Foxon, "Influence of Adiabatically Transmitted Edge Channels on Single-Electron Tunneling through a Quantum Dot," Phys. Rev. B **46**, 7236 (1992).

[55] J.Weis, R.J. Haug, K. v. Klitzing, and K. Ploog, "Competing Channels in Single-Electron Tunneling through a Quantum Dot," Phys. Rev. Lett. **71**, 4019 (1993).

[56] R.C. Ashoori and R.H. Silsbee, "Capacitance Study of the Single Electron Addition Spectrum of Quantum Dot Arrays," in *Proceedings of the International Conference of Nanostructures and Mesoscopic Systems* (Academic, Santa Fe, 1991); R.C. Ashoori, H.L. Stormer, J.S. Weiner, L.N. Pfeiffer, S.J. Pearton, K.W. Baldwin, and K.W. West, "Single Electron Capacitance Spectroscopy of Discrete Quantum Levels," Phys. Rev. Lett. **68**, 3088 (1992).

[57] B. Meurer, D. Heitmann, and K. Ploog, "Single-Electron Charging of Quantum-Dot Atoms," Phys. Rev. Lett. **68**, 1371 (1992).

[58] C.J. Gorter, Physica **17**, 777 (1951).

[59] T.A. Fulton and G.J. Dolan, "Observation of Single-Electron Charging Effects in Small Tunnel Junctions," Phys. Rev. Lett. **59**, 109 (1987).

[60] R. Wilkins, E. Ben-Jacob, and R.C. Jaklevic, "Scanning-Tunneling-Microscope Observation of Coulomb Blockade and Oxide Polarization in Small Metal Droplets," Phys. Rev. Lett. **63**, 801 (1989).

[61] E.R. Williams, R.N. Ghosh, and J.M. Martinis, "Measuring the Electron's Charge and the Fine-Structure Constant by Counting Electrons on a Capacitor," J. Res. Natl. Inst. Stand. Technol. **97**, 299 (1992).

[62] T. Ando, A.B. Fowler, and F. Stern, "Electronic Properties of Two-Dimensional Systems," Rev. Mod. Phys. **54** (1982).

[63] C.C. Eugster, "Electron Waveguide Devices," PhD thesis, Massachusetts Institute of Technology (1993).

[64] J. Davies, "Electronic States in Narrow Semiconducting Wires near Threshold," Semicond. Sci. Tech. **3**, 995 (1988).

[65] C.C. Eugster, J.A. del Alamo, P.A. Belk, and M.J. Rooks, "Criteria for the Observation of One-Dimensional Transport in Split-Gate Field-Effect Quantum Wires," Appl. Phys. Lett. **58**, 2966 (1991).

[66] C.C. Eugster and J.A. del Alamo, "Tunneling Spectroscopy of an Electron Waveguide," Phys. Rev. Lett. **67**, 3586 (1991); C.C. Eugster, J.A. del Alamo, M.R. Melloch, and M.J. Rooks, Phys. Rev. B **48**, 15057 (1993).

[67] M.J. Rooks, C.C. Eugster, J.A. del Alamo, G.L. Snider, and E.L. Hu, "Split-Gate Electron Waveguide Fabrication Using Multilayer Poly(methylmethacrylate)," J. Vac. Sci. Technol. B **9**, 2856 (1991).

[68] F.F.Y. Wang, *Introduction to Solid State Electronics*, North-Holland, New York (1980).

[69] Khalid Ismail, "The Study of Electron Transport in Field-Effect-Induced Quantum Wells on GaAs/AlGaAs," PhD thesis, Massachusetts Institute of Technology (1989).

[70] O.V. Lounasmaa, *Experimental Principles and Methods below 1K*, Academic, New York (1974).

[71] A. Kumar, S.E. Laux, F. Stern, A. Zaslavsky, J.M. Hong, and T.P. Smith III, "Effect of Nonequilibrium Deep Donors in Heterostructure Modeling," Phys. Rev. B **48**, 4899 (1993).

[72] M. Büttiker, "Quantized Transmission of a Saddle-Point Constriction," Phys. Rev. B **41**, 7906 (1990).

[73] L.P. Kouwenhoven, "Transport of Electron Waves and Single-Charges in Semiconductor Nanostructures," PhD thesis, Delft University of Technology (1992).

[74] J.A. del Alamo and C.C. Eugster, "Quantum Field-Effect Directional Coupler," Appl. Phys. Lett. **56**, 78 (1990).

[75] R.A. Ghanbari, M. Burkhardt, D.A. Antoniadis, H.I. Smith, M.R. Melloch, K.W. Rhee, and M.C. Peckerar, "Comparative Mobility Degradation in Modulation-Doped GaAs devices after E-Beam and X-Ray Irradiation," J. Vac. Sci. Tech. B **10**, 2890 (1992).

[76] M. Burkhardt, private communication.

[77] F.R. Waugh, M.J. Berry, D.J. Mar, R.M. Westervelt, A.C. and A.C. Gossard, "Single-Electron Charging for Double Quantum Dots," APS Bull. **39**, 887 (1994).