Quantum Mechanical Transport in Submicron Electronic Devices

by

Philip Frederick Bagwell

S.M. Massachusetts Institute of Technology, 1988
B.S. Georgia Institute of Technology, 1984

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
December 1990

Signature of Author _____
Department of Electrical Engineering and Computer Science

Certified by _____
Professor Terry P. Orlando
Thesis Supervisor

Certified by _____
Professor Dimitri A. Antoniadis
Thesis Supervisor

Accepted by _____
Professor Arthur C. Smith
Chairman, Departmental Committee on Graduate Students

Quantum Mechanical Transport in Submicron Electronic Devices

by

Philip F. Bagwell

Submitted to the
Department of Electrical Engineering and Computer Science
on December 26, 1990 in partial fulfillment of the requirements
for the Degree of Doctor of Philosophy
in Electrical Engineering and Computer Science

Abstract

Electronic devices with characteristic dimensions of the order of 100 nm or less exhibit many novel quantum transport phenomena at low temperatures when the phase-breaking length becomes comparable to the device size. This thesis describes electron transport mechanisms and the resulting current-voltage relationships in quasi-one-dimensional wires, superlattices, and resonant tunneling devices. An intuitive 'convolution method' is developed to describe the energy averaging due to a finite bias voltage, finite temperature, disorder, and the influence of emitter dimensionality on these currents. We emphasize the dominant effect of evanescent or 'cutoff' electron waveguide modes in determining the shape of the electrical conductance versus Fermi energy in a confined geometry such as a quantum wire. Finally, we study experimentally the magnetoconductance of a novel Si 'grating gate' field effect transistor where the current path can be varied electrostatically in a single device from many narrow wires in parallel, to a modulated periodic potential, to a two-dimensional electron gas. Electron weak-localization, the classical Drude magnetoconductance, and the quantum Hall effect are modified by the periodic potential.

1 Introduction

This thesis describes phase-coherent electronic transport in small conductors. Quantum mechanical effects in the conductance of an electronic 'device' can be destroyed if the temperature is too high, if the applied voltage across the device is too large, if too many dopant impurities or geometrical irregularities are present in the structure, or if an unconstrained direction of free electron motion is permitted inside the device structure. We present a simple 'convolution method' which unifies the mathematical description of all these effects, and embodies in a very physical way some of the more important effects of temperature, voltage, disorder, and any additional directions of free motion in the device. We apply this 'convolution method' to describe the current-voltage characteristics of different device structures such as resonant tunneling devices, superlattices, and quasi-one-dimensional wires.

This 'convolution picture' shows that finite temperature, voltage, disorder, and any directions of free electron motion inside the device each serve to 'broaden' or 'energy average' the conductance in different ways. We find that when energy relaxation processes occur only in the contacts of the device, which often happens at low temperatures, the 'smearing' or energy averaging of the conductance due to a finite temperature is independent of the energy averaging due to the finite applied voltage. In this extreme coherent regime the applied voltage acts in some ways like an effective temperature, yet does not 'heat' the electrons. This difference between the energy averaging due to the finite voltage and finite temperature energy averaging has observable physical consequences in electrical transport. Incorporating additional directions of free electron motion inside the device increases the average conductance by classically adding many 'quantum contact conductances' in parallel, yet serves to 'smear' any quantum structure present in the conductance. In contrast to the thermal and voltage energy averaging, there is no characteristic energy scale associated with the 'dimensional energy averaging'.

The bulk of this thesis studies electron transport in disordered quasi-one-dimensional wires. Electrical transport in quantum wires is very similar to microwave transmission down a microwave waveguide, so that much of microwave physics can be applied to these 'electron waveguides'. It more difficult to fabricate a purely ballistic electron waveguide than a defect free microwave guide, so that studying how electrons scatter in an imperfect narrow wire is important to understanding the variation of the conductance with Fermi energy in the electron waveguide. The conductance of these narrow quantum wires is obtained by calculating how an electron scatters from obstacles in a low-dimensional structure and then applying Landauer's formula.

We find that electron scattering in quantum wires is dominated by the evanescent or 'cutoff' waveguide modes whenever the Fermi energy is near a subband minimum, so that transport can only be correctly described when these modes are properly considered. Our result is reasonable because the evanescent density of states diverges as the Fermi energy approaches a subband minimum from below, just as the propagating density of states diverges as the Fermi energy approaches a subband minimum from above. And even though the cutoff waveguide modes themselves carry no current, they strongly affect the scattering of propagating modes through the scattering boundary conditions when the

Fermi energy is near a subband minimum in a narrow wire. This effect is well known in microwave waveguides and is manifest as additional electromagnetic energy storage near any discontinuities or defects in a microwave guide. For electron waveguides, additional particles are 'stored' in these evanescent modes near scattering centers. One can physically picture these additional electrons as 'clinging' to the sides of the defect *in a classically allowed region.* The usual evanescent waves inside the classically forbidden region of a tunnel barrier can exist in any spatial dimension, but the type of evanescent or cutoff waveguide modes considered in this thesis can exist only in a low-dimensional structure such as a quantum wire.

The most striking effect of the evanescent waveguide modes is that, when an attractive scattering potential is present, bound or quasi-bound states composed of evanescent modes can form just below each new subband minimum in the wire. These 'quasi-donor levels' strongly reflect any carriers nearby in energy, leading to characteristic 'dips' in the conductance versus Fermi energy of the wire. Other scattering anomalies are present as a result of the evanescent modes, including 'perfect transparency' through a point scatterer whenever the Fermi energy aligns with a subband minimum. The electron moves as if the point defect were completely absent from the wire at these special energies. Anomalous behavior of the conductance versus Fermi energy due to the cutoff waveguide modes is also manifest in more complicated geometries, as shown in this thesis.

Finally, we experimentally study the two-terminal magnetoconductance of a Silicon 'grating gate' MOSFET at low temperatures. The device consists of a lower 'grating' gate and a continuous top gate. This dual-gate device structure permits one to gradually vary the inversion layer geometry electrostatically from many narrow wires in parallel, to a modulated potential, and to a two-dimensional electron gas. For low magnetic fields normal to the interface, weak localization becomes much more pronounced as the inversion layer is pinched from a two-dimensional electron gas into narrow wires. This proves the electrons are phase-coherence across the wire width. For intermediate magnetic fields of 1-10 Tesla, we also find a large drop in the device current of 90% or more and which persists to room temperature, as electrons are added to the device so that it opens electrostatically from many narrow wires in parallel to a two-dimensional electron gas. This large negative transconductance is a manifestation of the different boundary conditions which must be applied to the low-field Drude magnetoconductance tensor in a long and narrow versus a short and wide conductor. At high magnetic fields from 10-30 Tesla, the approach to the quantum Hall effect strongly and qualitatively depends on the inversion layer geometry. Furthermore, the conduction band degeneracy of (100) Silicon is not resolved in the narrow wires at high magnetic fields, leading to Hall steps of twice the expected size.

This thesis is organized as follows: The body of the thesis summarizes the main results obtained during this research. This introduction contains an even shorter summary of the main results. Detailed justification of these results can only be obtained in the appendices to the thesis, consisting of papers published and submitted for publication. The purpose in writing the new sections of this thesis is not to repeat material contained in the appendices, but to present it in an intuitive way so that the reader does not become distracted by additional details. Appendix C and J are in fact also reviews of material in the other appendices, and could easily have been substituted for portions of the remaining chapters of this thesis.

But where I have more to add, where a point could have been made more concisely, clearly, with increased physical understanding, or where connections between different concepts in these papers drawn, I have done so in writing the new sections of this thesis. I have arranged the appendices in logical order, rather than the chronological order of their publication. Quantum mechanical electron transport is already too large and too rapidly expanding a field to present a tutorial review in this thesis. I point to important books, papers, and review articles [1]-[10] to accomplish this task.

1.1 Experimental Background

Three of the best known quantum mechanical transport phenomena [1]-[10] involving normal electrons[1] in inversion layers are the quantum Hall effect [11], weak electron localization [12], and the 'universal' conductance fluctuations [13]-[15]. To this list could also be added the periodic magnetoconductance oscillations in long metallic cylinders [16]-[17], the Aharonov-Bohm effect in small metallic rings [18], the quantization of the ballistic conductance in GaAs/AlGaAs heterojunctions [19]-[20], modifications to the classical Hall effect in ballistic conductors at low magnetic fields [21]-[22], and electronic charging of a small and nearly isolated conducting region [23]-[24]. Highly nonlocal 'resistances' have also been observed in quantum conductors [25]-[26]. If one moves away from inversion layers or metals to vertical transport in GaAs/GaAlAs heterojunctions, the resonant tunneling diode [27] is the premier 'quantum device', with a record peak to valley current ratio of 30 to 1 at room temperature [28]. Undoubtedly, more exciting developments are yet to come. Other quantum mechanical effects in electronic transport have also possibly been observed.

Weak localization and the universal conductance fluctuations are both understood as wave interference effects arising from independent electrons scattering coherently from a 'random' impurity potential. The 'universal' conductance fluctuations are similar to optical speckle patterns, and are called 'universal' because they depend only weakly on the sample shape, dimensionality, and average value of the conductance itself (as long as the conductor is in the metallic limit where 'ensemble averaged' perturbation theory is valid). These 'universal' fluctuations give a specific 'fingerprint' for each individual sample. As the sample size becomes larger than the electron phase-coherence length L_ϕ, each individual phase-coherent segment acts like an independent sample, leading to the destruction of electron wave interference patterns in the conductance through 'self-averaging'. The universal conductance fluctuation pattern is altered in a small magnetic field, but is not destroyed until the magnetic field becomes large enough to force the conductor into the quantum Hall limit.

Weak electron localization, since it depends only on pairs of time reversed paths, is

[1] Another important class of 'quantum devices', not discussed in this thesis, arise when the charge carriers are superelectrons. The phase-coherence length in superconductors is believed to be infinite, circumventing the problem of maintaining phase-coherence in a quantum device when the current is carried by normal electrons. The superconducting input coil of a SQUID flux-to-voltage transducer can be made arbitrarily large without the device malfunctioning. Furthermore, superconducting devices, even of a finite size, are believed to be 'single channel', not 'multi-channel' as for small normal metal conductors.

not destroyed by this 'self-averaging'. So long as L_ϕ is larger than the size of the average diffusion path to return to the origin, weak localization persists even though the sample size L may be large so that $L \gg L_\phi$. Weak localization is destroyed by a large enough magnetic field, usually less than 1 Tesla, which breaks time reversal symmetry inside the conductor. Weak localization is related to the Aharonov-Bohm effect, and is in fact responsible for oscillations in the conductance of a long metallic cylinder in a magnetic field [16]. These 'Sharvin oscillations', which have a magnetic flux period of h/2e, were not initially believed in the West, even though they had been previously predicted [17]. The standard Aharonov-Bohm effect predicts h/e oscillations in the conductance of a small ring in a magnetic field, and has also been observed [18]. If an array of these rings are placed in parallel, the conductance oscillation period is found to be h/2e as for the Sharvin cylinder. A good review is found in Ref. [8].

The quantum Hall effect (and related experiments) has been beautifully explained by Büttiker [29] in terms of magnetic 'edge states' and the subsequent elimination of backwards scattering in a strong magnetic field. This simple one electron picture, when combined with a Landauer transport formalism, has brought great clarity to a previously confused research area. The 'bulk conduction picture' of the quantum Hall effect [30]-[31] has all but expired. If electronic currents really were flowing though the sample bulk in the quantum Hall limit, why does the measured two terminal resistance not depend on the size and shape of the conductor? [2] Electron 'localization' away from the center of a Landau level was formerly thought to be responsible for the quantum Hall effect [30]-[31], but to the mind of this author was never able to explain the Hall plateaus. The 'explanation' of the quantum Hall effect in terms of electron 'localization' between Landau bands also usually erroneously invoked current carrying states below the Fermi level. States below the Fermi energy cannot contribute to any net current flow at the device terminals. Of course, the 'extended' edge states actually carry the quantized Hall current at the Fermi energy. A strong piece of evidence in favor of the edge state picture is the breakdown of the quantum Hall effect in narrow samples, explained as tunneling between the sample edges [32].

A new type of symmetry for the electrical conductance in a magnetic field, when the conductor is connected to multiple current and voltage 'probes', has been predicted [3] [33] on the basis of the Landauer transport formalism. The same four-terminal phase-coherent conductance should be measured if the current and voltage leads are interchanged and the external magnetic field is reversed, even though the 'conductivity' may be non-local. This generalization of the Onsager reciprocity relations has been experimentally verified in multiprobe conductors [34]. Measured four-terminal voltages can have a sign opposite to that expected classically from the current flow pattern in this formalism, which is also experimentally observed. Power dissipation in the conductor as a whole is of course positive.

Two more recent experiments which have been extremely important in defining the field of quantum mechanical electronic transport in small devices are the 'quantization of the

[2] In the classical Hall effect both the two terminal and Hall resistances do depend on the size and shape of the conductor, even if the conductor is large. However, for a long and narrow 'Hall bar' geometry, the Hall voltage is given essentially by ρ_{xy} and the two-terminal voltage given by ρ_{xx} as explained in Appendix M.

ballistic point contact resistance' [19]-[20] and the 'quenching of the Hall effect' in small and nearly ballistic 'Hall crosses' [21]-[22]. The discovery of conductance quantization in a ballistic point contact was important because it clarified when different types of Landauer transport formulas could be applied to describe conduction and defined the extreme quantum ballistic limit to transport. It was also the first clear observation of quasi-one-dimensional subbands in electrical transport. The quenching of the Hall effect, and related low field magnetoresistance anomalies such as the 'negative bend resistance', were another success for the Landauer transport formalism. Discovery of the quenching of the Hall effect was also important because most of its dominant features could be explained by treating the electrons as classical billiard balls [22], simply ignoring their wave properties despite the extreme high mobility small samples held at milli-Kelvin temperatures.

1.2 Theoretical Background

The most important theoretical achievement in quantum electronic devices has been the development and understanding of the Landauer conductance formulae [1]-[6]. The discovery of the quantized ballistic conductance in point contacts, which could legitimately be said to have been 'predicted' by studies of the Landauer formulae [35]-[36], has greatly stimulated and clarified our understanding of quantum transport. By relating the conductance to the total transmission and reflection coefficients of the device [37], the Landauer conductance formulae have greatly simplified computations as well as providing a conceptual framework for thinking about conductance.

Two central insights were obtained in Landauer's early work. The first and most important is that any transport coefficient can be related to the transmission coefficients (or scattering matrix) of the device. While this insight seems evident today, like any important advance in science it is only immediately obvious in hindsight [38]. The scattering mechanism reflecting the carriers need not be purely elastic or involve only single-electron physics [37]. But Landauer's second and also crucial insight is that electrostatic potential is dropped locally around scattering centers in the device as follows: A steady incident current is applied from the left to an obstacle. The obstacle reflects some carriers leading to a charge build-up on the left a relative defecit on the right. This 'transport charge' should be screened self-consistently like any other embedded charge. When this is done, an electric dipole field builds up locally around the scattering obstacle and assists the incident applied current to flow around the obstacle. Integrating these 'residual resistivity dipole fields' across each scatterer gives the total voltage drop across the device.

There has already been a great deal of work done using the various Landauer conductance formulae, and its rate of growth is increasing rapidly as more people come to understand it. Like Feynman diagrams, Landauer's formulae have made quantum conductance calculations more tractable and understandable. It is somewhat surprising (and in the opinion of this author a major scandal) that the Landauer picture was first completely neglected for over two decades, and was then controvertial for almost another decade, before reaching common acceptance. During this period throughout the 1970s a conceptually identical conductance formula was in routine use in the research area of tunneling conductance, especially in the area of resonant tunneling, and was known at least since

the 1930's. Needless to say, its full implications were not appreciated [39]. A clear understanding of resonant tunneling theory during the 1970s would have forced any theorist to conclude that the one-dimensional ballistic conductance should be quantized. Instead, it took the development of microlithography and small electronic conductors to force a major rethinking of electrical transport theory, though the correct conceptual path should have been apparent long before the 1980's. Fully understanding the ballistic conductance quantization requires knowledge of how the voltage is dropped near the opening of the conductor [5]-[6].

Another large share of blame for the underdeveloped state of electrical conductance theory until the 1980's can be placed on the excessive authority and credibility placed in diagrammatic quantum field theory techniques. These techniques supposedly 'automate' physics calculations (meaning one doesn't have to think too deeply about the process), yet they usually obscure rather than assist physical understanding of electrical transport. Usually when one knows the final answer, it is at least possible to derive it again using diagrammatic field theory techniques. Yet Landauer's simple conductance formula has *still* not been convincingly derived from quantum field theory [5]-[6]. And discussions of quantum transport are still burdened with terminology tied to this one calculational technique. When electrical conduction is being described in the independent particle approximation, quantum field theory techniques seem both too obscure and unnecessary. In transport theory these techniques have been applied to produce notoriously wrong answers; conduction in a quasi-one-dimensional wire is a prominent specific example (see Chapter 3). Field theory techniques have also contributed little towards, and have done much to obfuscate [31], our understanding of the quantum Hall effect. Yet even the application of simple wave mechanics to electrical conduction has lagged until this decade. Quantum physics has been very successful in understanding the energy level spectra of atoms and molecules, the interaction of light and matter, etc. Phase-coherent electronic transport in solids can soon hopefully be added to its resume of successes.

1.3 Open Problems

The confusion previously surrounding the whole field of the quantized Hall effect sharply illustrates the previously confused thinking about conductance calculations in general. A theoretical situation similar to that which occurred in the quantum Hall effect is repeating itself for conductance calculations in quasi-one-dimensional wires (Chapter 3 and Appendix L). Magnetoconductance in bulk samples, especially the so-called 'longitudinal Shubnikov de-Haas oscillations' [40], have been described with a formalism almost identical to the erroneous one used to describe transport in narrow wires. Quantum Monte Carlo simulations of bulk semiconductors use essentially the same 'Golden-Rule' approach to describing transport [40]. Ref. [41] develops the Golden-Rule/Master Equation approach thoroughly and very physically, obtaining many important insights. Some elements of these transport calculations may eventually turn out to be correct, but until one understands which elements of the models can in fact be obtained from a more rigorous Landauer-like transport approach, these 'Golden-Rule' type transport calculations should be regarded as highly suspect. 'Golden-Rule' scattering rate transport calculations are especially unlikely (see

Chapter 3) to be correct near any scattering thresholds such as those arising from optical phonons, scattering into higher conduction band valleys in GaAs, and near confinement subbands in quantum wires, thin films, or inversion layers. The degree of discrepancy remains to be calculated for most of these systems, but for narrow wires the change of the scattering rate is in the wrong direction and the magnitude of the difference is quite large.

As far as we now understand, the Landauer approach is limited to time independent potentials. Generalization of the formula to the AC conductance is an outstanding theoretical problem [35], though some convincing attempts have been made [42]. Finite frequency effects are still usually computed using the Kubo formula [43], though physical interpretation of the results is usually unclear. Each new calculation using the Kubo formula constantly begs this question of interpretation: Why can the AC frequency be regarded as a quasi-chemical potential in the Kubo approach, since a more reasonable candidate for such a quasi-chemical potential is the rms AC voltage? How does the changing electrostatic potential distribution due to the AC voltage and inertia of the carriers affect the quantum AC conductance? Can the electron 'reservoirs' in fact be assumed to follow the applied AC voltage as in Ref. [42]? The research area of 'tunneling times' should make a strong connection with any AC Landauer conductance approach [44].

Another outstanding theoretical problem has been the inclusion of phase-breaking processes in electrical transport. Büttiker has proposed a simple theoretical model for phase-breaking scattering [45], consisting of placing an electron 'reservoir' at the scattering site. Carriers entering and exiting the reservoir can do so at a constant energy, but their phases are unrelated due to the long path lengths inside the reservoir. Presumably this model is related to low-energy acoustic phonon scattering, but its present use is mainly conceptual. Reference [46] distributes point oscillators everywhere in space and, after calculating how the electron couples to these oscillators in a first Born approximation (like a 'Golden-Rule' rate approach), finds that these oscillators act very much like tiny Büttiker reservoirs distributed in space. A viewpoint emphasizing the measurement of the particle by the 'environment' inducing dephasing is given in Ref. [47]. Reference [48] proposes an exactly solvable 'local phonon' model, in which the transmission amplitudes are convolution broadened by the inelastic scattering. This 'local phonon' model is presumably related to electron scattering from a barrier which oscillates in time [49], giving rise to phase-coherent 'sidebands' in the electron transmission.

Scattering from a magnetic impurity spin is the canonical example that inelastic scattering and phase-randomizing scattering are not necessarily the same. Imagine, for example, an Aharonov-Bohm loop with magnetic impurities sprinkled in one arm. If no magnetic field penetrates the arms of the interferometer, a magnetic impurity with spin up has the same energy as the conductor with the impurity spin down. Thermal agitations can (supposedly) induce impurity spin flips with no energy being dissipated by the conducting electrons. And the electron scattering elastically from each set of random impurity spin configurations has a different phase when it exits the arm of the Aharonov-Bohm loop. Therefore, averaging the conductance over each different spin configuration destroys the Aharonov-Bohm oscillations. This should lead to destruction of the Aharonov-Bohm effect when magnetic impurities are present in a metal loop and $\mu_B B < k_B T$. The Aharonov-Bohm effect would be restored if the magnetic field were allowed to penetrate the arms of the conductor and $\mu_B B > k_B T$. When a large enough magnetic field is present in the

arms, thermal agitations do not have enough energy to flip the magnetic impurity spins so that one no longer averages the conductance over the random phase from each spin configuration.

The preceeding paragraph really does not consider spin-flip scattering processes between the electron and the magnetic impurity. We emphasize that scattering from an individual spin is not necessarily an incoherent process. There are coherent spin-flip scatterings and incoherent (thermally induced) spin-flips. The same statement applies to phonons. Lattice vibrations can be coherently induced by electron motion as well as by incoherent thermal agitation from a heat bath. Electron scattering from an oscillating barrier is one example of coherent 'phonon' scattering [49]. The oscillating barrier is also an example that inelastic scattering and phase randomizing scattering are not necessarily the same, since any phase changes of the incident wave are faithfully reproduced in each emitted sideband.

What is the potential distribution around obstacles in a quantum conductor? How does electron screening occur when the conductor is phase coherent? Applying a local Thomas-Fermi screening approach is unlikely to be adequate under these conditions. The most systematic study of this question to date has been carried out in Ref. [50], though a fully self-consistent solution for any particular problem has yet to be carried out. Reference [50] argues that, for small applied voltages, these 'self-consistency' corrections to currents flowing out of the device terminal are likely to be negligible. Yet, as emphasized in Ref. [6] and Ref. [50], transport formalisms that ask questions only about what occurs at electron 'reservoirs' cannot begin to adequately answer how current is carried through the conductor, *i.e.* the transport mechanism. One must look inside the conductor and ask, 'What is the detailed current flow pattern and voltage drop around obstacles?' to understand transport processes [50].[3]

Much of the work attempting to find self-consistent solutions to electron transport problems involves quantum kinetic equations [50]-[52], an interesting area of study in its own right. Ref. [50] emphasizes that power dissipation does not necessarily occur at the same point in space as the voltage drop in a quantum conductor. Ref. [52] has emphasized, using the 'notch' state trapped in the accumulation layer of a resonant tunneling diode emitter, that the standard 'Golden Rule' scattering rate theory (or Pauli Master Equation) violates current conservation if the scattering states are not plane waves. Therefore, the Pauli Master Equation cannot describe scattering in a spatially inhomogeneous situation. Admitting off diagonal density matrix elements into the calculation to preserve current continuity then produces a full quantum kinetic theory.

Asking for the self-consistent electrostatic potential within the device requires a full understanding of the coulomb interactions between electrons. An interesting recent series of experiments [24], in which electrons are forced through a small region of conductor where 'charging' or 'interaction' effects are likely to be significant [53], has yet to be fully understood. Small normal metal tunnel junctions have shown 'coulomb blockade' effects [54], which have been explained semi-classically.

How is the noise current affected by quantization and phase-coherent transport? Surely

[3]Understanding screening in a phase-coherent conductor is increasingly important due to the recent observation that high electron density in a contact is much more important than high mobility to observe ballistic conductance quantization.

the Nyquist formula must suffer some modification [55]-[56]? One would hope that the noise could be expressed in terms of the static transmission properties of the conductor, possibly modified to account for individual electron transmission (shot noise). Perhaps the shot noise spectrum would also be modified by quantization and phase-coherent conduction [57]? Finally, there are 'classic' papers on electron transport [58]-[61], now decades old, whose insights and validity need to be re-examined in light of our current knowledge.

1.4 Guide to Appendices

Appendices A-C are concerned mainly with diffusive electron transport in quasi-one-dimensional wires and superlattices. The transport is treated 'semiclassically' in that the conductance is obtained as a solution to the semiclassical equations of motion for the electron wavepacket. Of course, this solution requires making a model for the scattering time (a point addressed in detail in Appendix L and Chapter 3).

A better intuitive feel for the semiclassical treatment in Appendices A-C is that the electron is assumed to move between scattering events at the Fermi velocity. For classical diffusive motion and a Drude-like conductance to apply, phase-coherence must be broken inside the conductor. Two alternative viewpoints are then possible: (1) Phase-coherence is assumed to be disrupted *at each scatterer.* (2) Phase-coherence is assumed to be disrupted *between scattering events.* The important question in both viewpoints is, 'What determines the scattering time τ_{tr} in the conductor? One should then have a reliable model to calculate τ_{tr} for a given scattering process.[4]

Consider the first viewpoint: When an electron scatters in the semiclassical model its phase-coherence is assumed to be broken. However, the electron motion between scattering events is phase-coherent and quantum-ballistic. Thus, all one needs to determine the conductance is the electron group velocity and scattering time as a function of Fermi energy. From this viewpoint the calculations in Appendices A-C really describe the conductance through a set of phase-breaking scatterers in a superlattice or quasi-one-dimensional wire. The transport lifetime τ_{tr} then describes phase-breaking scattering which also randomizes the momentum direction. [5]

The second viewpoint asserts the following: Phase-coherence of the electron is assumed to be broken between scattering events, but not by any process which significantly reflects the electron or alters its Fermi velocity. From this viewpoint the calculations in Appendices A-C describe conduction through a set of elastic scatterers with weak phase-randomizing processes between the elastic scatterers. The scattering time τ_{tr} is then determined solely by the momentum randomization from the elastic scatterers. The calculations in Appendices A-C were originally done with this second viewpoint in mind.

Appendices A-C also develop a 'convolution method' to include the effects of finite temperature and 'disorder' on the density of states and electrical conductivity. The 'thermodynamic density of states' $dn/d\mu$ is a well defined concept in Appendix B, explaining

[4] In either viewpoint the electron requires phase-coherence over a few superlattice periods for its velocity versus Fermi energy to have significant modulation. A different method must be used to calculate the conductance when the inelastic scattering length is shorter than the superlattice period.

[5] This viewpoint is developed somewhat in the discussion on classical diffusion in Appendix D.

thermal broadening of the density of states. However, the process leading to thermal broadening of the conductance is difficult to understand from the approach given in Appendix B. Thermal smearing of the conductance is shown to follow from a Landauer-type approach in Appendix D, if one assumes a carrier distribution incident on the conductor which is in thermal equilibrium with the device contacts. Appendix D therefore provides a qualitative understanding of why the conductance is said to 'thermally broaden'. Since the thermodynamic density of states applies even when no current flows in the device, it is an equilibrium property of a conductor. In contrast, thermal smearing of the conductance is a quasi-equilibrium property, since it depends on an imbalance of carrier velocities incident on the conductor.

'Disorder broadening' of the conductance and density of states in Appendices A-C is somewhat more problematic, but still probably correct and widely accepted as being correct. Although the standard quantum field theory calculation for this disorder broadening has been performed in detail [62]-[64], a good physical explanation for the broadening has never been given to my knowledge. I give one possible explanation for the disorder broadening in Chapter 2. The impurity broadening function is shown to be a histogram of the impurity potential if the potential varies slowly in space compared to the electron wavelength. It is still not completely clear if this 'impurity potential histogram' relates to the field theory discussion of Refs. [62]-[64], where the chosen scattering potentials are highly non-adiabatic. We give a possible relation between these two ideas in terms of the *non-local* influence of scattering potentials which are not adiabatic. The discussion in Chapter 2 recovers in a simple way the 'Golden-Rule' broadening of Refs. [62]-[64], yet still uses the key idea that, on average, the energy levels follow the background impurity potential for distances larger than the inelastic scattering length.

In Appendices A-C, convolution with the free electron density of states is shown to increase the 'dimensionality' of the density of states and conductivity. A good physical counting argument why this is true for the density of states is given in Appendix B. A similar result for the conductivity tensor was shown to follow from the standard semiclassical expression in Appendix B, but a physical explanation was not really attempted. This same result was also shown to hold for the tunneling current from Landauer's formula in Appendix D, providing a physical explanation for this result in terms of classically adding resistors in parallel. This physical interpretation is further extended and clarified in Appendix E (and Chapter 2). Each independent resistor, or current 'channel' to cross the conductor, has a value determined by the quantum mechanical transmission T of that 'channel', namely $R_{1D} = \pi\hbar/e^2T$. Classically adding these quantum mechanical resistors in parallel constitutes our dimensional 'convolution method'. This 'convolution method' for increasing the dimensionality of a conductor (or other system) sometimes appears in physics journals without an accompanying understanding of its physical origin or implications [65]-[66].

Appendix D studies the Landauer formula, which governs ballistic, diffusive, coherent, and incoherent transport. The 1D Landauer formula is

$$G_{1D}(E) = \frac{2e^2}{h}T(E) \quad , \tag{1}$$

where $T(E)$ is the quantum mechanical current transmission coefficient through the device

potential. The transmission coefficient $T(E)$ as a function of Fermi energy E can be evaluated for whatever scattering mechanisms exist: the gate potential, disorder due to impurities, acoustic or optical phonons, etc. [6] [37]. We emphasize here that Eq. (1) describes a 'two-terminal' conductance. The original Landauer formula in 1D is

$$G_{4p}(E) = \frac{2e^2}{h}\frac{T(E)}{R(E)} \quad , \tag{2}$$

where $R(E) = 1 - T(E)$ is the reflection coefficient, and describes a non-invasive 'four-terminal' measurement of the conductance [5]. But Eq. (landg1a) is not merely a formula of academic interest. A complete understanding of the ballistic conductance quantization, or indeed of any transport process, can be obtained only by using Eq. (landg1a) and its higher dimensional analogues of to see how the voltage is dropped near a constriction or other obstacle [5]-[6]. A large variety of other Landauer-type formulae are possible [67] depending on exactly what types of 'probes' are used to measure the voltage drop around the reflecting obstacle.

$T(E)$ can also be calculated classically through a chain of obstacles in 1D, as done in Appendix D. If the device has length L, the transmission coefficient is found to be $T(E) = l_{tr}/L$, where $l_{tr} = v_F \tau_{tr}$ is the transport mean free path. This recovers the standard Drude formula for the electrical conductance in 1D.[6] Viewed from this perspective, the conductance calculations in Appendices A-C are valid because the transmission coefficient is proportional to the Fermi velocity when the electron momentum direction is randomized on average every τ_{tr} seconds. Thus in Appendices A-C the transport must be regarded as semiclassical, that is partially phase-incoherent.

Appendix D also develops a novel 'voltage broadening' of the conductance. A source to drain bias is shown to energy average the transmission coefficient in much the same way as finite temperature averaging. Thermal and voltage energy averaging are shown to be statistically independent if the device is coherent, a surprising result since a drain voltage is generally regarded as 'heating' the electrons. 'Heating' due to power dissipation in the conductor and energy averaging due to a finite bias across the device are in fact quite different, with distinguishable physical consequences (Chapter 2). The independence of thermal and voltage broadening arises physically because no power dissipation occurs inside a coherent conductor. Power dissipation occurs only in the contacts to the device. If inelastic scattering does occur inside the device, so that the conductor is partially phase-incoherent, this picture is modified as described in Chapter 2.

The ideas in Appendix D were applied to resonant tunneling devices in Appendix E. The resonant tunneling diode currents are determined by adding many quantum contact resistances in parallel, so that the quantization of the ballistic resistance should have been predicted by people studying resonant tunneling diodes in the 1970's. We stress this point again in Chapter 4 using a simple example. The current versus base voltage, drain voltage, and emitter Fermi energy for one, two, and three dimensional double barrier devices are calculated in Appendix E using the 'convolution method'. All of the device I-V curves can be sketched very easily by hand once one understands how the different spatial dimensions

[6] This subject is not yet completely closed, since $\tau_{tr}/\tau = T/R$ from the Landauer approach, yet one obtains $\tau_{tr}/\tau = 1/2R$ from a 'Golden Rule' approach to the scattering time.

are related via convolutions, and if one also takes account of the energy averaging due to finite source-drain voltage. The shape of the conductance versus Fermi energy, versus drain voltage, or versus a 'base' voltage depends strongly on the emitter dimensionality as shown in Appendix E.

To describe a conductor of finite lateral size, a finite width quasi-1D, 2D, quasi-2D, or 3D conductor, one needs to consider a Landauer formula with many quantum 'channels' as

$$G(E) = \frac{2e^2}{h} \sum_{nm} T_{nm}(E) \quad . \tag{3}$$

The full power of a Landauer approach is needed in these 'multi-channel' scattering problems, or when the device has more than two terminals (a 'multi-probe' conductor). The resonant tunneling devices studied in Appendix E are a special case where T_{nm} is diagonal so that $T_{nm}(E) = \delta_{nm} T_n$, which just describes adding conductors $G_n = (2e^2/h)T_n$ classically in parallel. The transmission coefficient $T_{nm}(E) = \delta_{nm} T_n$ is diagonal in Appendix E due to the special shape of the double-barrier potential, which permits momentum conservation perpendicular to the tunneling direction. Again, each independent resistor may have a value determined by computing T_n quantum mechanically (as is the case for the resonant tunneling diode in Appendix E), or the resistor itself may be classical where the transmission coefficient T_n is determined using classical mechanics (as done in Appendix D). We may then sum either the classical or the quantum mechanical resistors in parallel using the 'convolution method'.

A simple model potential for which one can calculate T_{nm} exactly by hand is electron transmission through a point scatterer in a quasi-1D wire. This is done by a (brute force) wave-function matching method in Appendix F, and by a more 'elegant' Green function technique in Appendix G. Both approaches of course give the same answer for the transmission coefficients T_{nm}. The T_{nm}'s through these point scatterers are no longer diagonal since the scattering potential has no special Cartesian symmetry. The point scatterer is interesting in that, whenever the Fermi energy aligns with the bottom of a quasi-1D subband, the transmission is perfect as though no scatterer were present in the wire. This 'perfect transparency' effect is quite counterintuitive, and arises only because the evanescent wire modes are now properly included in the transmission calculation. This perfect transmission at a subband minimum is shown to be the only possible way to obey the scattering boundary conditions in Appendix I and Chapter 3.

Making the scatterer attractive leads to new 'quasi-donor levels' splitting off the confinement subbands. Donor levels below the conduction band minima of a doped semiconductor are quite well known, where the donor electrons are strictly bound to the parent atom at zero temperature. Requiring additional quantum confinement of the electronic motion to a narrow wire induces a 'remnant' of these donor levels below each subband minimum. These quasi-donor levels are therefore quite analogous to the original donor level in a bulk semiconductor, but are formed from the evanescent or 'cutoff' electron waveguide modes. The new quasi-bound states give rise to depressed transmission before a new confinement subband opens, where the energy separation between the 'dip' in conductance and the next subband is related to the binding energy of the state. Appendix F is the first published discussion relating these pronounced 'dips' in conductance to the 'quasi-donor' states.

Appendix F also uses a numerical scattering matrix technique to describe electron

transmission through a finite size rectangular box potential. Transmission is no longer perfect at a subband minimum when the scatterer is given a finite size. If the finite size scatterer is made attractive and large enough to admit several bound states, several conductance 'dips' appear before the opening of a subband. The appearance of multiple conductance 'dips' again confirms that the depressed conductance near these energies is related to bound states forming in the attractive potential.

An obvious question to ask once one has solved the multi-channel electron transmission through a point barrier is, 'What is the transmission through two such point barriers?'. The answer is given briefly in Appendix H, and in more detail in Appendix I. The interesting answer to emerge from studies of electron transmission through two point barriers is that the electron transmission is not always perfect despite that each individual barrier, taken by itself, is indeed perfectly transmitting. This result is in sharp contrast to the analogous situation in a strictly one-dimensional problem. It is found in Appendices H-I that wave interference between propagating modes, as in a Fabry-Perot resonant cavity, generally has little to do with the electrical conductance when the Fermi energy aligns with a subband minimum. Instead, the shape of the lowest evanescent waveguide mode determines the electron transmission so that the usual Fabry-Perot resonances are suppressed. A new series of resonances develop related to the evanescent modes, and these new resonances have completely different properties from the Fabry-Perot transmission oscillations. These new oscillations can be interpreted in terms of molecule-like 'bonding' and 'anti-bonding' states formed from the evanescent modes.

Appendices K and L are concerned with the multi-channel electron transmission through many point barriers in a quasi-one-dimensional wire. Appendix K criticizes an alternative view of the development of conductance 'dips' in a multi-channel wire before a subband opens. As alluded to previously, these drops in conductance before a new subband opens are due to 'quasi-donor-levels' splitting off a new confinement subband. Appendix L presents a positive calculation of the conductance in a quantum wire from the ballistic to the quantum diffusive regime, examining the conductance of both individual wires and the ensemble averaged conductance of a wire array.

A new subband structure fundamentally different from the quantized ballistic conductance steps is found for the conductance of a quantum wire array in Appendix L: The conductance falls rapidly after a subband opening when quantum diffusion is the dominant electron transport mechanism. This electron 'localization' subband structure manifests itself when the effects of the 'quasi-donor-levels' are carefully eliminated (by allowing only repulsive scatterers in the wire). Appendix L therefore proves that there is no unique subband structure of the conductance versus Fermi energy in a quantum wire. The depenedence of the conductance on Fermi energy in these narrow wires reflects the type of disorder present, even if one performs an ensemble average over different arrangements of the scatterers in an attempt to eliminate this dependence. If some fraction of the scatterers in the wire attract electrons, the quasi-donor levels dominate the shape of the conductance versus Fermi energy near subband crossings. The quasi-donor levels again produce large drops in the conductance before a new subband opens.

Appendix M describes measurements on Si MOSFET 'multiple parallel quantum wire' devices fabricated in the Submicron Structures Laboratory and Microsystems Technology Laboratories at MIT. The results shown have been reproduced on several different devices

at different times (four separate devices), and the measurements are shown to be consistent in several different regimes of temperature, gate biases, and magnetic field. A physical interpretation of the magnetoconductance is given which is consistent in all these different measurement regimes. It is shown that the 'grating gate' can actually pinch the two-dimensional electron gas into narrow wires. Interesting manifestations of electron weak-localization, the classical Drude magnetoconductance, and the quantum Hall effect are found.

Although our interpretation of the measurements in Appendix M require no startling new physical ideas, the combination of effects seen together in this multiple parallel wire device is unique and sharpens our understanding of magnetoconduction in inversion layers. Weak localization becomes much more pronounced as the device is pinched from a two-dimensional electron gas into narrow wires, proving that the electronic motion is phase-coherent over the width of the wire. This large modification to weak localization as one pinches the inversion layer into a narrow wire has not been previously confirmed experimentally to our knowledge. Certainly, measurements of weak localization have been performed on different two-dimensional and one-dimensional MOSFETs. However, direct comparison of the results is difficult because of the radically different scales of the conductance of a two dimensional inversion layer and a narrow wire. The 'grating gate' MOSFET allows a direct comparison of weak localization in one and two dimensions directly on the same device.

The 'grating gate' MOSFET also displays an unusual two terminal low-field Drude magnetoconductance, which has not been previously well understood. This can be confirmed by studying the references (which only peripherally discuss this subject) given in Appendix M. Indeed the standard review article on two-dimensional systems [69] cites a Japanese paper [70], which gives only the standard Drude magnetoconductance tensor in two-dimensions. Other workers [71]-[72] give general and formal 'solutions' for the two-terminal magnetoconductance by applying appropriate boundary conditions to a two-dimensional conductance tensor, but do not obtain the simple insights of Park [68] given in Appendix M.

In the low-field Drude magnetoconductance regime, we find a large (80-90%) drop in the two-terminal conductance as the inversion layer geometry is opened electrostatically from many narrow wires in parallel to a two-dimensional electron gas. This effect persists to room temperature, confirming its classical origin. We explain this large negative transconductance by requiring that the Drude conductance tensor properly satisfy the boundary conditions at the edges and contacts to the device, giving an intuitive picture of the effect. For a long and narrow conductor, the current must essentially continue to flow parallel to the sides of the device irrespective of the magnetic field. Therefore, a high aspect ratio conductor displays almost no magnetoresistance. For a short and wide conductor, however, the equipotential contours from the source and drain determine the current direction, so that the current must flow at the Hall angle from source to drain. Since the edges of the conductor are far away compared with the close proximity of source and drain contacts, the magnetic field has a large and dramatic effect on the transport in a low aspect ratio device. As the inversion layer geometry opens from many narrow wires in parallel to a two-dimensional electron gas in an applied magnetic field, the current path 'switches' from running parallel to the sides of the device to flowing at the Hall angle,

producing a large negative transconductance.

The high field magnetoconductance and approach to the quantum Hall effect in the grating gate device is also anomalous, with the two-terminal conductance characterized by a large rising background when the inversion layer geometry consists of many narrow wires in parallel. When the inversion layer forms a two-dimensional electron gas, the background two-terminal conductance continues to fall (as it did in the classical Drude regime). Thus, the magnetoresistance has large movements in opposite directions on different segments of the same device curve. The formation of quantum edge states in the wire array explains this result. Finally, when quantum Hall plateaus are well developed in the two-terminal conductance, we find the Hall steps are twice the expected size due to the conduction band valley of (100) Si not being resolved. The spin degeneracy is well resolved at the highest magnetic fields. The quantum Hall effect can also be used to count the number of conducting wires in parallel, giving additional confirmation of device integrity.

2 Convolution Picture of Quantum Transport

A simple 'energy averaging' picture can explain the lowest order effects of finite temperature, finite source-to-drain voltage, and 'disorder' in the conductance of quantum electronic devices. This 'convolution picture' of quantum transport is shown in Fig. 1, where one merely averages the transmission coefficient over an energy range kT, eV, and h/τ, respectively, to account for these effects. Each of these 'broadening mechanisms' is subject to limitations we discuss below, since temperature, voltage, and scattering do much more than simply broaden the energy distribution of conducting electrons at the Fermi level. However, in the lowest approximation, one can account for their effects by making the electron beam at the Fermi level non-monochromatic.

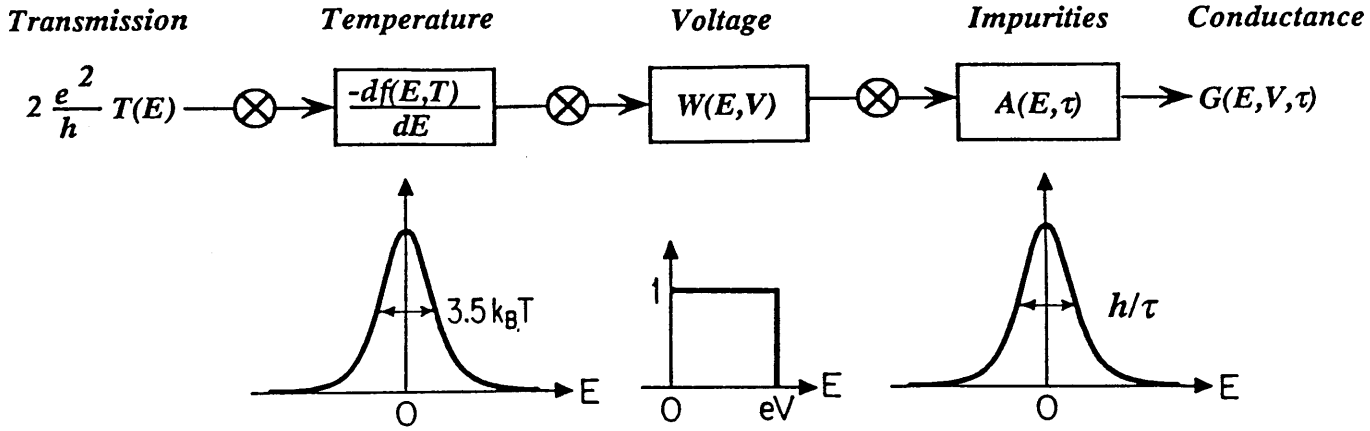


Figure 1: Energy level broadening due to finite temperature, finite source-to-drain voltage, and 'disorder' can be incorporated in the electrical conductance by convolving the transmission coefficient with functions of width $3.5kT$, eV, and h/τ, respectively.

The second aspect of our intuitive 'convolution method' is the incorporation of free electron motion perpendicular to the direction of transport shown in Fig. 2. As emphasized in Appendix E, this method is not a general way to increase the dimensionality of the conductance and density of states for any arbitrary device potential, but it works for an important class of electronic devices in addition to providing much physical insight into conduction. The method works if the total Hamiltonian is separable into a sum of components as $H = H_x(x) + H_y(y) + H_z(z)$ in Cartesian space. For any Hamiltonian having this form the transmission coefficient is diagonal as $T_{nm} = T_n \delta_{nm}$. For the tunneling diode, where $H_x(x) = p_x^2/2m$ and $H_y(y) = p_y^2/2m$ describing free electron motion perpendicular to the tunneling direction, the method is particularly simple: One merely convolves the transmission coefficient $T(E)$ with the free electron density of states for the other two directions of motion perpendicular to the tunneling barriers as shown in Fig. 2. The dimensional convolution method will also work if the Hamiltonian is separable in some other complicated coordinate space [65].

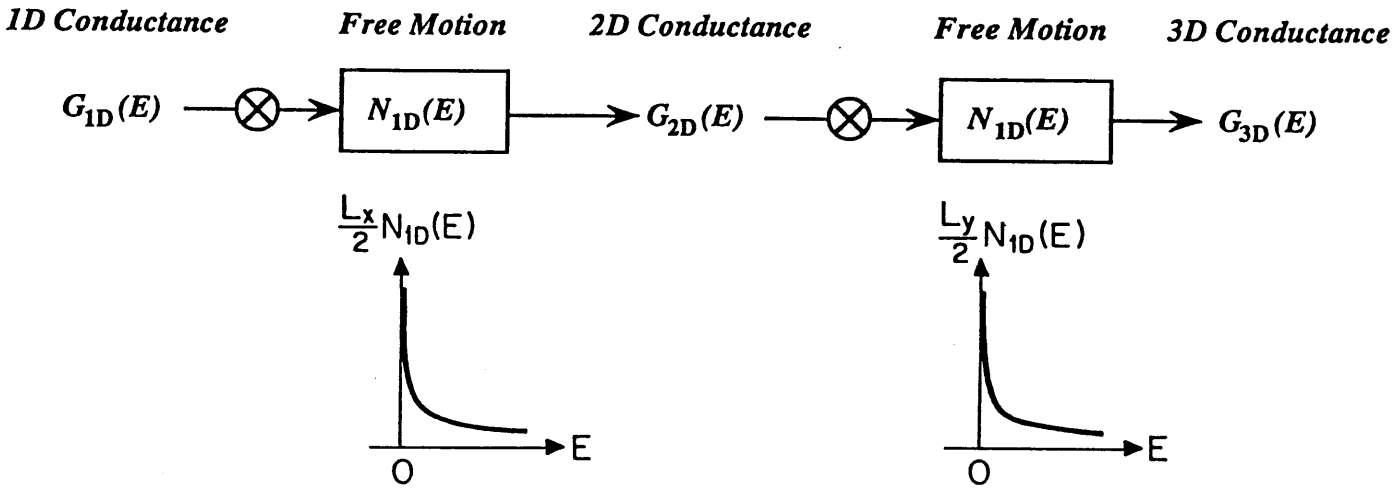


Figure 2: If the electrostatic potential is constant along the directions perpendicular to transport, one merely convolves the 1D transmission coefficient with the 'transverse' free electron density of states to obtain the total device current.

2.1 Disorder Broadening

Why are electron energy levels said to 'broaden' in the presence of a 'random' electrostatic scattering potential? Our simple contention is: (1) a scattering center which lowers the potential also lowers some of the energy levels on average, (2) a repulsive scattering center raises some of the energy levels on average, (3) the DC level of the potential 'shifts' the energy bands [73]. Of course, there will be fluctuations in the density of states in a phase-coherent sample due to electron wave-interference. Looking at the 'average' density of states is necessary in order to smooth out these interference patterns. *But the energy levels are still 'broadened' even in a specific sample which is phase-coherent.*

An Anderson model calculation [74] shows how this idea works. Fig. 3(a) shows the clean tight-binding periodic potential density of states. It has $1/\sqrt{E}$ singularities at both band edges. The 'on-site' disorder added in (b) and (c) has equal amounts of repulsive and attractive scatterers, so that there is no DC level shift of the states. The state density calculated in (b) and (c) is for a specific phase-coherent sample, so that wave-interference fluctuations can be seen in the density of states. However, the energy level spectrum is already 'broadening' in (b) and (c) even in this specific phase-coherent sample. One can easily imagine that averaging over many such calculated density of states curves, each with a slightly different arrangement of the disorder and therefore a different interference pattern, will produce a nice 'broadened' curve [75].

That attractive scatterers add states below the band edge and repulsive scatterers add states above the band edge can be seen in at least three limits: (1) the formation of 'bound' and 'anti-bound' states, (2) weakly varying 'adiabatic' potentials, and (3) in perturbation theory. Consider (1), the formation of bound and anti-bound states. It

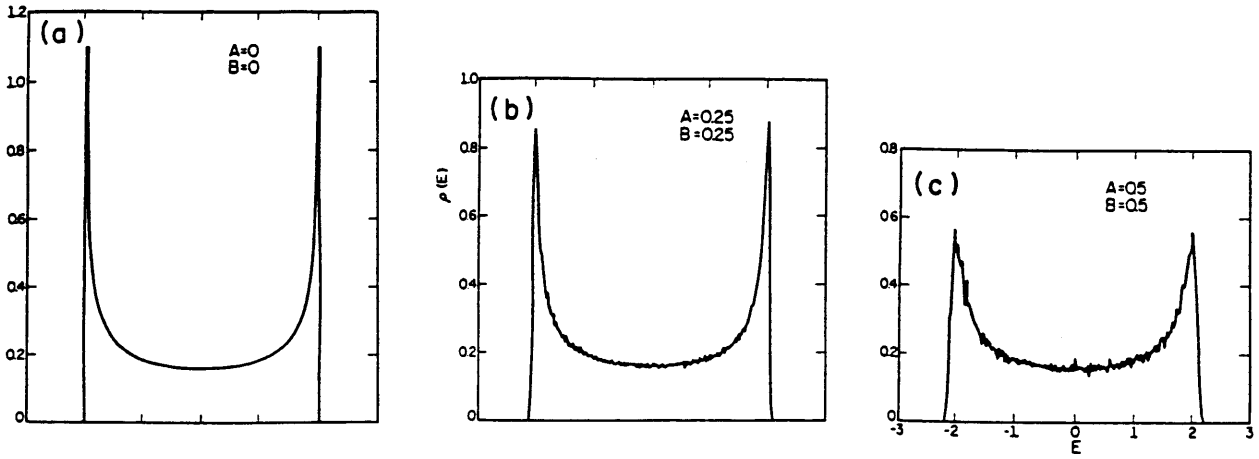


Figure 3: Calculated density of states *for a specific sample*, a tight-binding periodic potential with increasing amounts of 'on-site' disorder, showing the presence of energy level broadening. Attractive scatterers add states with $E \leq 2$, while repulsive scatterers add states with $E \geq 2$. The periodic potential is subject to (a) no disorder, (b) small disorder, and (c) medium disorder. From Ref. [74].

is well known that an attractive scattering center can produce a 'donor-level' below the conduction band edge in a doped semiconductor. But if the band has an upper limit of allowed energies as in Fig. 3, a repulsive scatterer can also produce an 'anti-bound' state above the band edge. The 'anti-bound' state is analogous to an acceptor level in a doped semiconductor. This 'anti-bound' state also illustrates a symmetry between the effects of repulsive and attractive scatterers when the energy spectrum is band-limited. Next consider (2), the limit of smoothly varying 'adiabatic' potentials. In this case the entire energy level spectrum follows the slowly varying background potential so that our contention in the first paragraph is trivially true. A first order non-degenerate perturbation theory calculation, in case (3), where an additional small scattering potential is added for example to a quantum well, also easily verifies these claims.

A similar view has been expressed in the context of the quantized Hall effect and the broadening of Landau levels in Ref. [30]: 'The various quantum states in each energy band can be divided into three general classes. The states near the bottom of each band, that is, those of lowest energy, are each localized in some small region of the sample. These low-energy localized states occur at "valleys" in the electrical potential energy, for example in the regions around impurity atoms that have an excess of positive charge. Similarly, near the top of the band are high-energy localized states. These are localized near "peaks" in the electric potential, perhaps in the regions around impurity atoms that have acquired electrons and so bear an excess of negative charge.'

Thus, the energy distribution is already 'broad' when scatterers are added, *even when the sample is phase-coherent.* The addition of many phase-coherent segments, to make up the incoherent sample of Fig. 4, is only necessary to 'smooth' the (already broadened)

density of states (by averaging out the quantum wave-interference fluctuations). The incoherent sample itself is not really necessary to produce energy level 'broadening', only level 'smoothing' of an already broadened spectrum.

How does a real macroscopic sample eliminate the wave-interference pattern which will manifest itself in the density of states? One imagines a situation like Fig. 4, in which phase-randomizing scattering effectively breaks the full sample up into a number of smaller coherent samples of size L_ϕ. Each phase-coherent segment of the conductor has a slightly different arrangement of impurity scatterers, so their quantum state densities will differ due to a complicated wave-interference pattern in each coherent region. One expects that the electrical conduction, density of states, optical absorption, etc., will simply be averages of the quantum properties over each phase coherent segment of the sample.

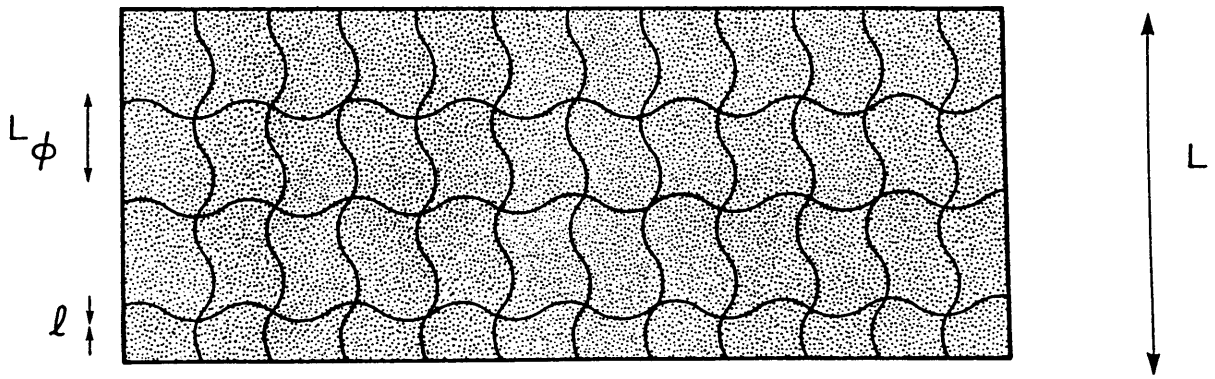


Figure 4: An incoherent macroscopic sample. Inelastic scattering inside the conductor breaks up the sample into many phase coherent segments of size L_ϕ.

Our discussion here relates to the one in Appendices A and B as follows: The impurity spectral function $A(E,\tau)$ tells us how much the energy levels move around from the extra disorder inside each phase coherent segment of the conductor. Inside each phase-coherent block $A(E,\tau)$ has peaky structure, due to electron wave-interference, which gets smoothed out if we look at the average spectrum. Thus, the 'broadening' due to elastic scattering is just a reflection of the range of states that are needed to form the eigenstates of the disordered system from those of the pure system. These newly formed states are eigenstates of the disordered potential inside each phase-coherent block, and again already have a broadened spectrum. The physics of energy level broadening is no more complicated than computing this spectrum.

'Ensemble averaging' over different arrangements of the scatterers, sometimes carried out in a type of perturbation theory called 'ensemble-averaged perturbation theory', mimics the incoherent macroscopic conductor of Fig. 4. When calculating the conductance of a narrow wire for example, 'ensemble averaging' corresponds to calculating the conductance for one specific arrangement of the static disorder in the wire, then for a second arrangement, etc. The conductances of the wires are first added classically like independent resistors in parallel, and then divided by the number of parallel wires to obtain the

'ensemble averaged' conductance. In ensemble averaged perturbation theory the conductance of the average sample, free of the wave-interference fluctuation pattern present in a single phase-coherent conductor, is calculated directly. However, the 'second moment' of the conductance can still be calculated (in a complicated way) in this special 'average' sample, which is how all the 'universal' conductance fluctuations results are obtained. Again, this type of 'self averaging' is (supposedly) carried out in large individual samples, like the macroscopic conductor of Fig. 4, due to phase-randomizing scattering. As more questions are asked about the nature of inelastic scattering itself, this assumption will suffer more scrutiny.

There are two common limits where energy level broadening can be explicitly calculated: (1) when the scattering potential is slowly varying in space, and (2) if the scattering potential is a delta function. We can easily translate the ideas from the previous paragraphs into mathematics if we assume that the impurity potential $V(x)$ varies slowly in space compared to the electron wavelength. We can then, to a good approximation, neglect the detailed wave interference pattern between impurities so that the Schrödinger equation applies locally at each point. The impurity potential $V(x)$ is then just a constant offset to the total energy at each point. Therefore, the average density of states $\bar{N}(E)$ is simply

$$\bar{N}(E) = \frac{1}{L}\int_0^L N_0(E - V(x))dx \quad , \tag{4}$$

where $N_0(E)$ is the density of states of the 'clean' system before the impurities are added. We can slightly rewrite Eq. (4) as

$$\bar{N}(E) = \int_{-\infty}^{\infty} dE' N_0(E - E') \int_0^L \delta(E' - V(x))\frac{dx}{L} \quad . \tag{5}$$

Eq. (5) can be expressed in a more transparent form

$$\bar{N}(E) = \int_{-\infty}^{\infty} dE' N_0(E - E')P(E') \quad , \tag{6}$$

where the probability density in energy $P(E)$ is found as

$$P(E) = \int_0^L \delta(E - V(x))\frac{dx}{L} \quad . \tag{7}$$

Eq. (7) makes it clear that the shape of the energy level 'broadening' function $P(E)$, sometimes called $A(E,\tau)$ in a different context, is simply a histogram of the impurity potential energy $V(x)$.

Note that Eq. (4) neglects wave interference between scattering events, thereby embodying the assumptions inherent in cases (2), adiabatic potentials which nearly eliminate reflections, and (3) first order non-degenerate perturbation theory. Eq. (4) in fact assumes that first order non-degenerate perturbation theory is valid locally at each point in space. These assumptions limit the validity of Eq. (4). For example, Eq. (4) cannot describe any impurity potential in which bound states can form. Also, due to electron wave-interference, inserting a repulsive potential can probably lower the energy of some states. But on average it is certainly true that an attractive potential lowers the energy of electron states and

a repulsive potential raises their energies. If the scattering potential is weak and varies slowly in space, Eq. (4) should provide a good approximation to the average density of states. Note also that because wave-interference is specifically neglected in arriving at Eq. (4), so that the 'ensemble-averaging' idea is not needed in this first approach.

The second model for energy level broadening consists of calculating $A(E,\tau)$ using ensemble averaged perturbation theory, and taking each impurity to be a point scatterer

$$V(x) = \sum_i \gamma \delta(x - x_i) \quad . \tag{8}$$

Then, as calculated in Refs. [62]-[64], the width in energy $\triangle E$ of $A(E,\tau)$ is given by the 'Golden-Rule' expression

$$\triangle E = 2\pi \frac{N_i}{L} \gamma^2 N_0(E_F) \equiv \frac{\hbar}{\tau} \quad . \tag{9}$$

The potential from Eq. (8) is highly non-adiabatic, so we cannot expect the level broadening from Eq. (7) to apply.

The usual argument justifying Eq. (9) runs as follows: The 'momentum spread' (*i.e.* the range of k states needed to form the eigenstates of the new system with scatterers from the 'clean' plane-wave states) around the Fermi surface introduced by scattering is $\triangle k$ and the corresponding energy spread is $\triangle E$. They must be related by

$$\frac{\hbar^2 (k_F + \triangle k)^2}{2m} = E_F + \triangle E \simeq E_F + \frac{\hbar^2 k_F \triangle k}{m} \quad . \tag{10}$$

One then argues the scatterers 'confine' the electrons in position on a scale of the mean free path so that $\triangle x = \ell$. Therefore, from $\triangle x \, \triangle k \simeq 1$, we obtain

$$\triangle E = \hbar \frac{v_F}{\ell} = \frac{\hbar}{\tau} \quad , \tag{11}$$

where τ is the mean free time. However, this does not explain why τ is determined from Eq. (9). One usually appeals to the 'Fermi Golden Rule' to justify why τ is given by Eq. (9). However, it is difficult to understand how this appeal to a 'scattering rate' can be related to the idea that each impurity locally 'shifts' the energy level spectrum around itself.

We give a different argument for Eq. (9) here: Suppose that the average spacing between the point scatterers is a, so that we can smear out each point scattering event over that spatial scale. Then, a naive application of Eq. (7) gives $\triangle E \simeq |\gamma|/a$, the same result one would obtain from first order non-degenerate perturbation theory. But this answer, $\triangle E \simeq |\gamma|/a$, is in conflict with the 'Golden-Rule' result of Refs. [62]-[64] and Eq. (9). However, because the scattering potential leading to Eq. (7) is assumed to be adiabatic, Eq. (7) requires that the impurity potential only *locally* influences the density of states. We can try a slightly different approach where the potential at a point x influences states at other points within a phase coherence length away from x. We also assume phase-coherence extends only over the spacing a to the next scatterer. By this assumption, we can still neglect wave interference between scattering events in calculating the average density of

states. These same assumptions are made implicitly in Refs. [62]-[64]. By again neglecting wave interference between scatterers in this second approach, we eliminate the need for 'ensemble averaging'.

We assume that each point scatterer influences the electron states only within a distance a between itself and the neighboring point scatterers. Then, the number of states which will be affected by the scatterer $V(x) = \gamma\delta(x)$ is approximately

$$\text{number of states} = \frac{\gamma}{a} N_0(E_F) a \quad . \tag{12}$$

Since these states are phase-coherent, they effectively 'see' the same point scatterer. Each state will therefore have its average energy moved by an amount γ/a. We can now treat the problem as if a number $(\gamma/a)N_0(E_F)a$ of *local* scatterers, each having a strength γ/a, are added to the conductor at each point. To avoid shifting the DC level of the potential, we assume for simplicity that half the scatterers are attractive and half are repulsive. These assumptions require

$$2P(E) \simeq \delta\left(E - (\gamma/a)^2 N_0(E_F)a\right) + \delta\left(E + (\gamma/a)^2 N_0(E_F)a\right) \quad . \tag{13}$$

Computing the variance of $P(E)$ from Eq. (13) recovers the 'Golden-Rule' broadening of Eq. (9). The argument in this paragraph is not completely rigorous, but may incorporate the basic physics of energy level broadening when the influence of the potential is *non-local.*

2.2 Thermal and Voltage Broadening

Thermal and voltage broadening follow directly from the calculation of tunneling currents in Appendix D, where these two broadening mechanisms are shown to be independent if the device is coherent. What limitations are there on this result? How is voltage broadening physically distinguishable from electron 'heating'?

If the source to drain voltage is large enough, electron heating certainly does occur inside a small conductor. A dilution refrigerator containing the sample can also be heated using this method if too little care is taken. But power dissipation need not occur inside the device where the voltage is dropped. It can occur in the large device contacts (electron 'reservoirs') which remain essentially in thermal equilibrium. As long as the demands placed on the device contacts to carry away the Joule heat produced are not excessive, conduction occurs at essentially the same 'temperature' as when the drain voltage is zero. It is in this sense that the energy averaging due to voltage and temperature are independent.

To illustrate the difference between Joule heating and energy averaging due to the finite source-drain voltage, consider the conductor in Fig. 5. Phase-randomizing (or inelastic?) scattering can be modeled as expanding the conductor out into an electron 'reservoir', by which we mean that it has a well defined temperature and chemical potential [45]. Geometrically diluting the current density by expanding the conductor out into a wide region ensures that the wide region is essentially in thermodynamic equilibrium so that $J(x,y) \simeq 0$ everywhere in the 'reservoir'. (A negligible amount of inelastic scattering inside each 'reservoir' is necessary to insure that thermodynamic equilibrium obtains in

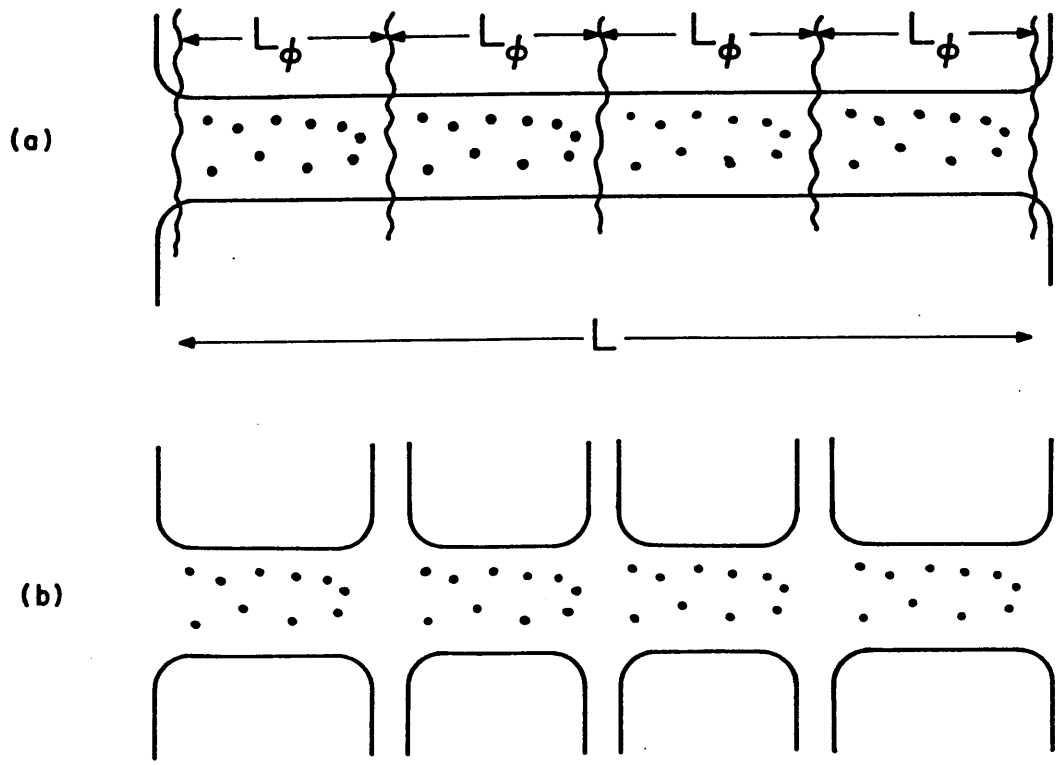


Figure 5: (a) A one-dimensional conductor much longer than an inelastic length, $L \gg L_\phi$. (b) Conceptual model of the same conductor as the addition of constriction resistances in series. Phase-randomizing scattering permits the classical addition of resistors in series. Each individual resistor, however, is phase-coherent. Adiabatically widening the conductor enforces thermal equilibrium after each inelastic scattering event.

each widened region, but we assume this additional scattering does not significantly reflect the carriers and therefore has little effect on the overall resistance.) Inelastic scattering inside the original conductor is therefore seen in the model of Fig. 5 as bringing the conductor locally into thermodynamic equilibrium.

Thermal averaging of the conductance will be $\sim kT$ in each coherent segment of the device in Fig. 5. But the voltage averaging of the conductance in each coherent device segment will be reduced to $\sim eV/n$, since only a fraction $\sim 1/n$ of the total voltage will be dropped across each phase-coherent segment if there are n inelastic scatterings between voltage probes. This difference offers the possibility of an experiment which can distinguish between electron 'heating' and energy averaging due to a finite source to drain voltage.

Consider again the conductor of Fig. 5(a), where there are four inelastic scattering lengths between the device contacts. Further, let us assume the characteristic energy scale E_c for structure in the conductance, *i.e.* the correlation energy E_c for 'universal' conductance fluctuations, is the same in each coherent segment. The wave-interference pattern in the current will then 'go away' at a temperature T^* determined as $kT^* \simeq E_c$,

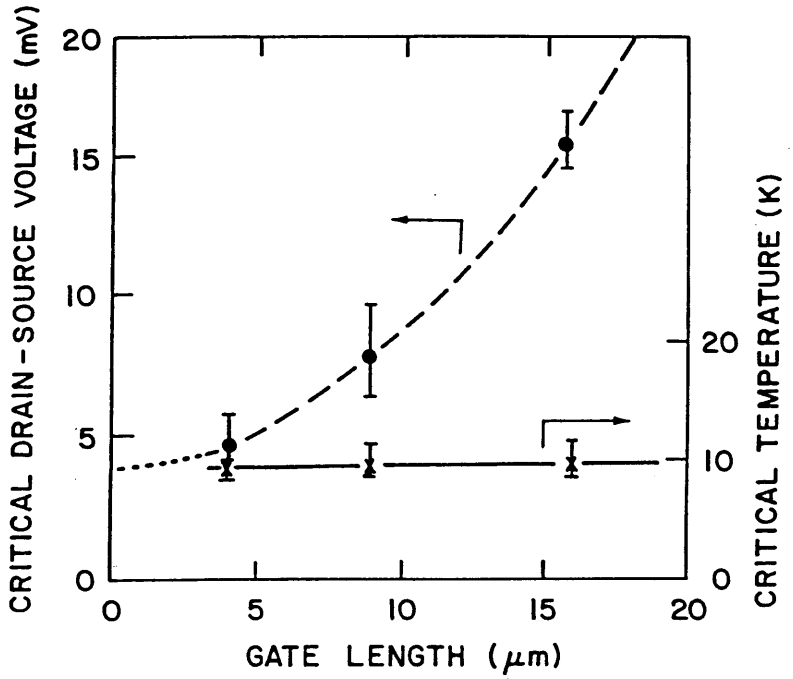


Figure 6: An experiment which could distinguish between electron 'heating' and energy averaging due to a finite source-drain bias voltage. One can also measure the phase-breaking length L_ϕ using this method, given by the intersection point of the solid and dotted lines. From Ref. [76].

independent of the number of inelastic scatterings between voltage probes. But the wave-interference pattern will persist up to a voltage V^* such that $eV^*/(L/L_\phi) \simeq E_c$, because only a fraction (L/L_ϕ) of the total voltage is dropped across each coherent segment of the conductor.

This viewpoint has been emphasized in Ref. [76] and in Appendix C, where V^* is plotted for different lengths L of the conductor. The oscillations persisting to higher drain-source voltages as the device is made longer would be clear evidence that energy averaging, not electron 'heating', is responsible for the structure washing out. If electron 'heating' were responsible, the criterion would be $eV^* = kT^* = E_c$ independent of the device length L. A drawback of this scheme is that the size of 'universal' conductance fluctuations naturally decreases anyway as the device is made longer, so one would have to correct for this power law effect. Furthermore, 'going away' would have to be better defined statistically. But such an experiment can qualitatively distinguish between energy averaging due to the finite device bias versus electron 'heating' even without these refinements. Additionally, one obtains an independent measurement of the phase-breaking length L_ϕ from the intersection point of the solid and dashed lines in Fig. 6.

2.3 Dimensional 'Broadening'

The simple dimensional convolution method shown for the conductance in Fig. 2, and which also applies to the density of states, can be understood in two dimensions as follows:

Since the Hamiltonian separates as $H = H_x(x) + H_z(z)$, the total energy can be written as $E = E_x + E_z$. We draw this constant energy line of constraint in Fig. 7.

The tunneling electron can have any combination of E_x and E_z such that the total energy E is conserved during tunneling. But let us choose a particular value of E_z and E. The current carried by the electron at energy E_z is simply

$$I_{1D}(E_z) = \frac{e^2}{\pi\hbar} T(E_z) \quad , \tag{14}$$

where $T(E_z)$ is the one-dimensional transmission coefficient through the tunneling potential described by $H_z(z)$. The number of electrons carrying this current in a small energy range dE_z is given by the density of transverse momentum states in this range, namely $L_x N_{1D}(E_x) dE_z/2$ where L_x is the size of the conductor along the x-direction. $N_{1D}(E_x)$ is the one-dimensional density of states from the Hamiltonian $H_x(x) = p_x^2/2m$, namely the free electron density of states.

Because the different transverse momentum directions do not scatter into each other, the currents in each channel can be classically added like resistors in parallel. The incremental device current dI at a particular value of E and E_z is therefore

$$dI_{2D}(E, E_z) = I_{1D}(E_z) L_x N_{1D}(E_x) dE_z/2 \quad . \tag{15}$$

When one allows the energy to be distributed between all values of E_z, the currents again add classically like resistors in parallel so that

$$I_{2D}(E) = \int dI_{2D}(E, E_z) = \int I_{1D}(E_z) L_x N_{1D}(E - E_z) dE_z/2 \quad . \tag{16}$$

This proves the convolution result shown in Fig. 2. A similar result holds for three-dimensional conductors and the electronic density of states in different spatial dimensions.

This 'dimensional convolution' method for the conductance simply corresponds to classically adding many quantum point contact resistances in parallel. This viewpoint is emphasized in Appendix E, and explained in a simpler way in Chapter 4. For the density of states, this 'dimensional convolution' method also holds. If particle motion is decoupled along the two orthogonal directions except for the constraint that the total energy of motion is a constant, the total number of available states is simply a product of the number of states along one direction times the number of states along the other. This is explained clearly in Appendix B.

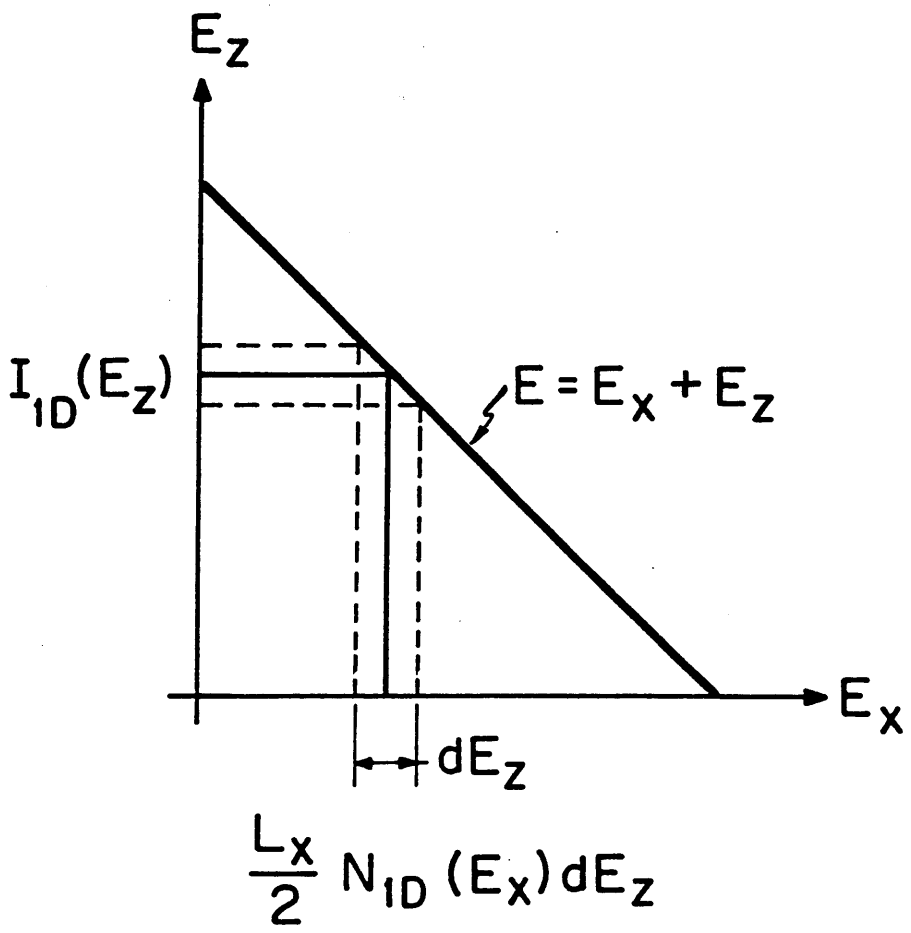


Figure 7: Currents added in parallel in energy space.

3 Conductance of Quantum Wires

A quantum wire or *electron waveguide*, where electrons are laterally confined to a few normal modes and transport occurs perpendicular to the confined direction, is shown in Fig. 8. The electron dispersion relation for these wires is shown in Fig. 9. The multiple channel Landauer formula is ideally suited for calculating the conductance of quasi-one-dimensional wires (see Appendices F through L). Each new normal mode, or electron transmission 'channel' to cross the conductor, carries a maximum current $I_{max} = e^2V/\pi\hbar$ of about 0.1 nA/μV. Equivalently, the ballistic resistance $h/2e^2$ of a single channel in the quantum wire is roughly 13 kΩ. Refs. [35]-[36] and [5]-[6] explain in detail why a ballistic conductor has a resistance of $h/2e^2$ per quantum channel.

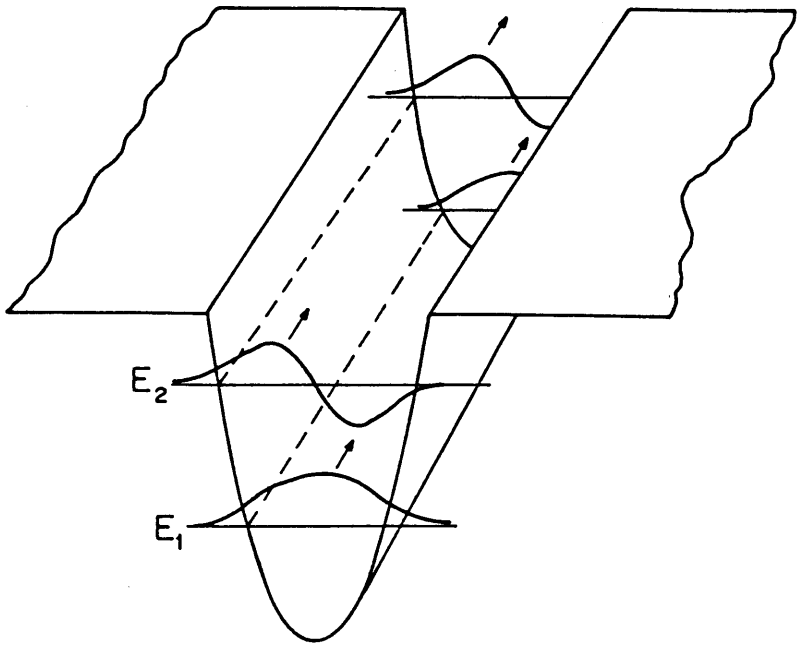


Figure 8: Transport modes in an electron waveguide.

The second and third sections of this chapter review rigorous quantum mechanical scattering calculations of the conductance in quantum wires done in Appendices F-L. We argue that these calculations are correct, give a consistent picture of electron scattering in low-dimensional geometries, and are on their way to becoming qualitatively understood. Evanescent or 'cutoff' electron waveguide modes shown in Fig. 9, and for which we later give a transparent physical picture, dominate electron scattering when the Fermi energy is near any subband minima in the wire.

However, to make connection with previous work and calculations done in my S.M. thesis [62], I first criticize previous approaches used to calculate conductance in these narrow wires. The criticism hinges mainly on the models used for the electron scattering time: the Fermi Golden-Rule approximation. This Born approximation scattering theory (or 'Golden Rule' approach) completely breaks down whenever the Fermi energy is near a

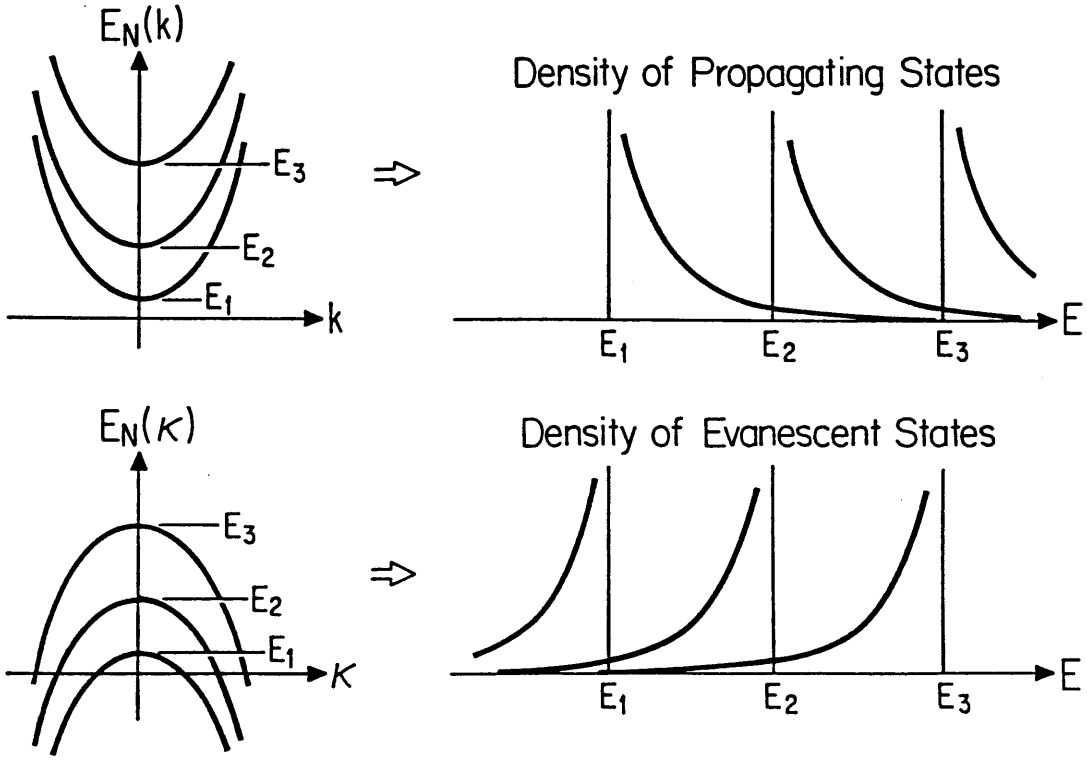


Figure 9: Dispersion relation for an electron waveguide. The top half indicates the propagating modes having dispersion $E = E_n + \hbar^2 k_n^2/2m$, while the bottom half shows 'negative dispersion' for the evanescent or 'cutoff' waveguide modes having $E = E_n - \hbar^2 \kappa_n^2/2m$ (where $\kappa_n = ik_n$). The corresponding evanescent density of states diverges *below* the subband minimum. From Appendix F.

subband minimum in a narrow wire. But because this particularly simple type of scattering theory breaks down near a subband minima does not imply that incoherent electron motion in a quantum wire cannot be classical and diffusive. The physical assumption of classical diffusive motion and the subsequent Drude-like conductivity can be valid for these wires in some circumstances. However, to properly implement the physical assumption of classical diffusive motion, the conductor must be 'assembled' by classically adding each coherent 'resistor' in series (see Fig. 5 in Chapter 2). To obtain the correct value of the total resistance, therefore, the scattering in each coherent segment of the conductor must be calculated properly. Equivalently, in the Drude-Sommerfeld language, the scattering times must be calculated properly. The Fermi-Golden Rule approach is incapable of properly computing these scattering times when confinement subbands exist in a narrow wire.

3.1 Fermi Golden-Rule Approach

To motivate this section, consider first a completely one-dimensional conductor where only the x-direction of space exists. The impurity potential we take to be a sequence of point scatterers. We assume the electron phase is randomized after each collision so the

resistances of each point barrier can be added classically in series. We can therefore calculate the resistance of each coherent segment of the conductor by quantum mechanically obtaining the electron transmission coefficient T through a single point barrier. The Landauer conductance of the chain, neglecting the contact resistance at the end for this long chain having low transmission, is therefore (see Appendix D)

$$G = \frac{e^2}{\pi\hbar}\frac{T}{R}\frac{1}{N_i} \quad . \tag{17}$$

Here N_i is the total number of barriers and $R = 1 - T$ the reflection coefficient of a single point barrier.

Each point scatterer in the one-dimensional wire has a potential $V(x) = \gamma\delta(x - x_i)$. The quantum mechanical transmission coefficient T through this potential is

$$T = \frac{1}{1 + \left(\frac{m\gamma}{\hbar^2 k}\right)^2} \quad , \tag{18}$$

where $v = \hbar k/m$ is the electron velocity. The resulting Landauer conductance of the impurity chain is

$$G = \frac{e^2}{\pi\hbar}\left(\frac{\hbar^2 k}{m\gamma}\right)^2 \frac{1}{N_i} = e^2\left\{v^2\left(\frac{\hbar^2 v}{2\gamma^2}\right)\frac{1}{N_i}\right\}N(E) \quad . \tag{19}$$

In writing this last expression we have introduced the density of states $N(E)$ in one-dimension, namely $N(E) = 2m/\pi\hbar^2 k = 2/\pi\hbar v$.

Eq. (19) can be directly compared with the Drude conductance of the one-dimensional chain, given by

$$G = \frac{\sigma}{L} = \frac{1}{L}e^2 D(E)N(E) = \frac{1}{L}e^2\left(v^2\tau_{tr}\right)N(E) \quad , \tag{20}$$

where $D(E)$ is the diffusion constant and τ_{tr} the momentum relaxation time or transport lifetime. For Eq. (19) and Eq. (20) to be equivalent we must have

$$\frac{\hbar}{\tau_{tr}} = \pi\frac{N_i}{L}\gamma^2 N(E) \quad . \tag{21}$$

Eq. (21) is one form of the well-known Fermi Golden Rule expression for the transport lifetime, specialized to a point scattering potential. The classical diffusive conductance obtained from the Landauer transport formula and the conventional 'Golden-Rule' scattering calculations give the same answer in this one-dimensional or single subband model. We find below that, due to the influence of evanescent modes on the transport, the Golden-Rule type scattering arguments do not predict the correct Drude conductance when the Fermi energy crosses a confinement subband.

Before the discovery of the quantized ballistic conductance, several calculations were made to determine the effect of quantum confinement on the transport properties of electrons. In most of these calculations, the 'ensemble averaged' conductance of diffusive

quasi-one-dimensional wires was obtained using the Kubo formula [77]-[82]. Similar calculations continue to appear [83]-[84]. A simplified discussion of these references is given in Appendices A-C. Here we repeat the most salient points of that discussion.

A sharp drop in the conductivity, due to rapid intersubband scattering, is obtained whenever the Fermi level passes into a new quasi-one-dimensional subband in the calculations of Refs. [77]-[82] and Appendices A-C. This is because the scattering rate in the wire is found to be proportional to the *total* electron density of states as in Eq. (21). When the density of states diverges above each subband, the scattering rate diverges and the conductivity falls to zero. The small number of additional carriers obtained when a new subband is populated cannot overcome the large increase of the scattering rate above a new subband in Refs. [77]-[82] and Appendices A-C. One small problem in these calculations is that the scattering rate depends on the density of states, but scattering changes the density of states, which again changes the scattering rate, etc. Recognizing this difficulty, most of Refs. [77]-[82] have applied the *self-consistent* Born approximation to calculate the conduction in narrow wires. (A self-consistent convolution broadening method should give similar results, if not identical ones.) However, the results even after this pseudo[7] 'self-consistency' is taken into account are qualitatively the same: electron scattering increases dramatically above each new subband minimum in the wire.

These calculations [77]-[82] are in contrast to the quantized conductance steps in a ballistic wire: the conductance increases whenever the Fermi energy passes through a new subband minimum in those experiments. Furthermore, the drop in conductance predicted in Refs. [77]-[82] and Appendices A-C sharply contrasts to the results shown in Appendices F-K for a small number of impurities. In those calculations the conductance always increases when a new subband is populated. This is because, in Appendices F-K the transmission and reflection at an individual impurity scatterer is correctly calculated. In addition, evanescent modes in the wire are not included in the calculations of Refs. [77]-[82] and Appendices A-C, and we know that such modes are very important in determining the subband crossing behavior of the conductance.

The Golden-Rule approximation of Refs. [77]-[82] and Appendices A-C is also questionable when more than one subband is occupied. Appendix G shows that the Born approximation breaks down completely at a subband minimum: It is found in Appendix G that the Born approximation gives the exact opposite answer as a rigorous scattering calculation. The Born approximation suggests a decrease in transmission when a subband is encountered, while the correct answer gives increasing transmission as a subband is crossed. This difficulty is not remedied in any self-consistent Born approximation scheme.[8] It is therefore difficult to understand from what perspective the detailed results for conduction in quasi-one-dimensional wires in Refs. [77]-[82] and Appendices A-C can be viewed as being valid.[9] I conclude that the Golden-Rule type scattering approach, touted in many

[7] All electrostatic charging effects are obviously neglected here and in the later discussions in this chapter. It is unknown at the present time how this affects the results presented.

[8] It has been suggested that a 'degenerate' Born scattering theory might work at a subband minimum. I do not believe this is possible, since the perturbation series in Appendix G diverges at a subband minimum. Furthermore it is not even clear how one might develop such a theory, since one is solving for the *scattering* of electrons from potential barriers rather than the stationary energy levels of a closed system.

[9] It is possible to obtain a sharp drop in the four probe conductance using the multiple channel Landauer

textbooks and papers on the subject of electrical conduction in solids, gives not just a quantitatively wrong answer, but the wrong qualitative dependence of electrical conductance on the Fermi energy in a narrow wire. It should be equally unreliable for discussing the Shubnikov-de Haas oscillations, scattering in a multi-valley semiconductor such as GaAs, and optical phonon limited mobility, since scattering thresholds similar to confinement subbands are encountered in all these cases.

Before leaving this section I would like to point out that none of the effects discussed above have any connection to the electron mobility enhancement in one-dimensional wires predicted by Sakaki [85]. Sakaki uses the same Golden-Rule type theory as above to support this claim. Reference [85] notes that the Fourier transform of the impurity scattering potential $V(q)$, which enters the general Golden Rule expression for the transport lifetime

$$\frac{\hbar}{\tau_{tr}(q)} = 2\pi \frac{N_i}{L} |V(q)|^2 N(E) \quad , \tag{22}$$

must be evaluated only at $q = 2k_F$ in one-dimension (since small angle scattering is not allowed in 1D). The usual Bessel function result for a screened Coulomb potential $V(q)$ falls off quite rapidly with increasing q, due to the finite size of scatterers in the wire, so Sakaki concludes that the electron mobility should be larger in a one-dimensional wire. However, for point scatterers $V(q) = \gamma$ independent of q, so point scatterers give none of Sakaki's mobility enhancement. Indeed, the Sakaki mobility enhancement should be negligible for scatterers having a size smaller than $1/2k_F$. It is only for very large scatterers that Sakaki's argument is relevant. I again emphasize that Sakaki's ideas say nothing about subband crossings, they are concerned only with conduction deep in the lowest subband of the wire. As we have seen, these 'Golden-Rule' type theories behave much better if the Fermi energy is not near a subband crossing, so the ideas in Ref. [85] may perhaps be correct.

3.2 Multi-mode Transmission Through a Single Barrier

To understand the scattering and subsequent electrical conductance in a quantum wire, recall the electron dispersion relation from Fig. 9. As the Fermi energy approaches a subband minimum from below, the evanescent density of states diverges and dominates the scattering properties of electrons from any obstacles in the wire. The shape of these states in real space is shown in Fig. 10, where the evanescent modes can be pictured as 'clinging' to the side of the barrier *in a classically allowed region of space.* These evanescent modes accumulate around any obstacles present in the wire, driven by the current incident on the obstacle and the mode coupling at the scatterer (as stressed in Appendix F). This phenomena is well known in the context of electromagnetic wave scattering in microwave waveguides. Even though these evanescent states do not carry any currents along the wire, they strongly influence the scattering boundary conditions when the Fermi energy is near

formula due to a screening effect [37]. This screening effect is also examined in Appendix F. However in the extreme diffusive limit of low transmission, the two-probe and four-probe Landauer conductances are almost identical. This screening effect therefore cannot rescue the Golden-Rule approach. Furthermore, none of the calculations I critize in this chapter considered screening of the scattering potential by reflected or transmitted electrons.

a subband minimum. Therefore, these exponentially decaying modes wind up *dominating* the scattering when the Fermi energy is near a subband minimum.

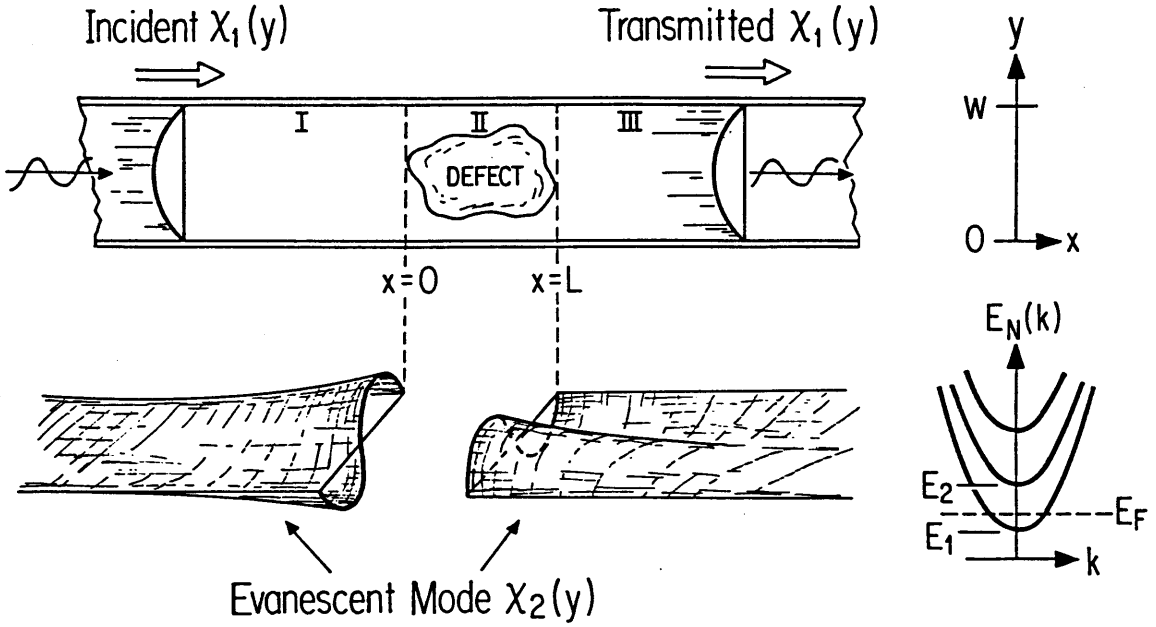


Figure 10: Due to electrons scattering from the current incident on an obstacle, evanescent modes accumulate around obstacles in the wire. This process becomes especially important in determining the dependence of electrical conductance on the Fermi energy near any subband crossings. From Appendix F.

A simple model scattering problem, where a single point scatterer is placed in a narrow wire, is sufficient to investigate the influence of these evanescent states on transport. The conductance versus Fermi energy through the point scatterer is shown in Fig. 11. A completely different qualitative dependence of the electrical conductance on the Fermi energy from that expected in Refs. [77]-[82] and Appendices A-C is found in Fig. 11: The conductance increases whenever the Fermi energy passes through a new subband in the wire. Cascading many of these defects by adding constriction resistances in series, as done in Fig. 5 and implicitly in the calculations of Refs. [77]-[82] and Appendices A-C, will also result in an increasing conductance whenever the Fermi energy crosses a new subband minimum in the wire.

Attractive scatterers strong enough to allow the formation of bound states result in a unique structure of the conductance versus Fermi energy in these narrow conductors. Just as an attractive impurity can form a donor level below the conduction band minimum in a semiconductor, so too can attractive defects placed in a narrow conductor form 'quasi-donor levels' below each subband in the wire. These new quasi-donor levels or 'quasi-bound states' manifest themselves as the drop in conductance just before the Fermi energy enters a new subband in the broken curve of Fig. 11. It is not obvious that these 'dips' in conductance before a new subband opens correspond to quasi-donor-states, as inspection of the References given in Appendix K indicates, but this matter is analyzed in some detail

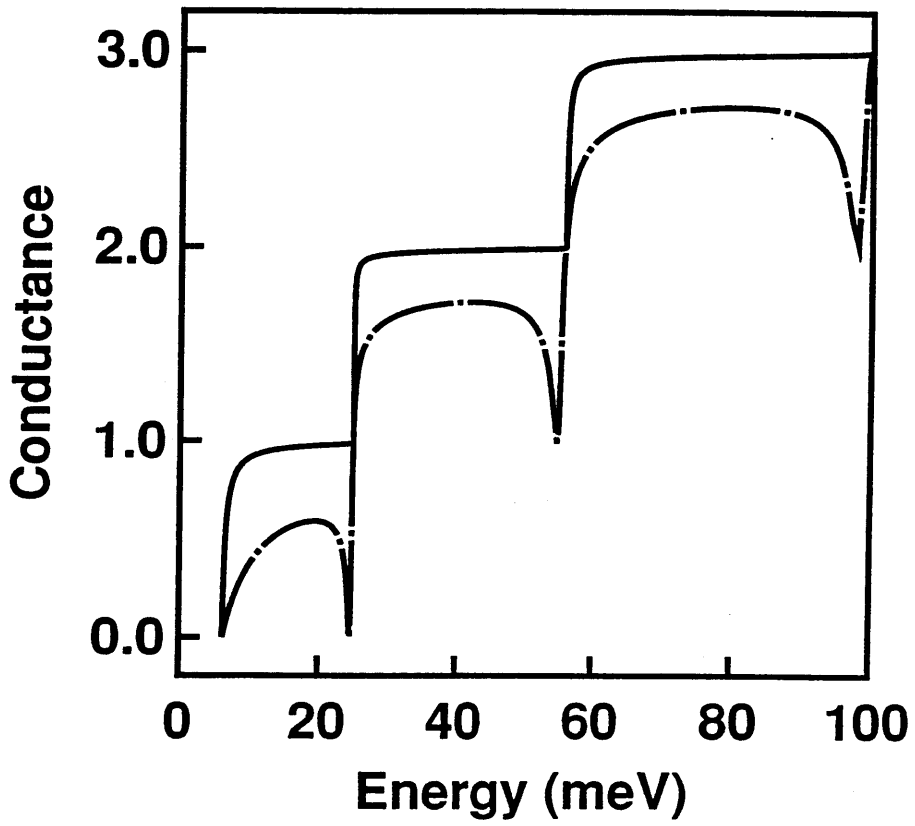


Figure 11: Landauer conductance through a single point defect in a narrow wire. The solid line is for a repulsive point defect, while the broken curve shows the conductance when the sign of the scattering potential is reversed. The dips in conductance before the opening of a new subband indicate 'quasi-donor levels' forming in the attractive scatterer. From Appendix F.

in Appendices F and I. Even though these conductance 'dips' are clearly associated with the quasi-donor levels in Appendices F and I, the detailed scattering mechanism leading to the depressed conductance is not clear.[10]

One puzzling feature of the electrical transmission through a point scatterer in a narrow wires is that the electron transmission is perfect whenever the Fermi energy aligns with a subband minimum regardless of the sign, strength, or location of the scatterer. Although recognized in Appendix F-G, this property was finally understood simply to be a consequence of the scattering boundary conditions in Appendix I. The only way in which all the boundary conditions can be obeyed is for the incident mode to propagate through the point defect without reflection. A more general arrangement of point scatterers also yeilding perfect transmission at a subband minimum was found in Appendix I.

Perfect transmission at a subband minimum results for the point scatterer simply by

[10]I thank Prof. Marc Kastner for pointing out the existence of these type of quasi-donor states when he saw the curves from Fig. 11 in an informal talk. At that time the origin of these conductance drops was not obvious to me.

imposing the boundary conditions at a defect having the form

$$V_d(x,y) = \gamma\delta(x)\delta(y-y_i) \quad . \tag{23}$$

The wavefunction $\psi(x,y)$ can be expanded in terms of the lateral normal modes of the wire $\chi_n(y)$ as

$$\psi(x,y) = \sum_n \alpha_n(x)\chi_n(y) \quad , \tag{24}$$

simply from the completeness property of the wavefunction. Integrating the Schrödinger equation gives

$$\left.\frac{d\alpha_n(x)}{dx}\right|_{x=0+} - \left.\frac{d\alpha_n(x)}{dx}\right|_{x=0-} = \sum_m \Gamma_{nm}\alpha_m(x=0) \quad , \tag{25}$$

where the mode coupling constants Γ_{nm} have a 'product form'

$$\Gamma_{nm} = \frac{2m\gamma}{\hbar^2}\chi_n(y=y_i)\chi_m(y=y_i) \quad . \tag{26}$$

Γ_{nm} factors in Eq. (25) due to the special shape of the point scatterer. Therefore, Eq. (25) can be rewritten using Eq. (26) as

$$\left.\frac{d\alpha_n(x)}{dx}\right|_{x=0+} - \left.\frac{d\alpha_n(x)}{dx}\right|_{x=0-} = \frac{2m\gamma}{\hbar^2}\chi_n(y=y_i)\left[\sum_m \chi_m(y=y_i)\alpha_m(x=0)\right] \quad , \tag{27}$$

Note that the term in square brackets on the right does not depend on mode n, so that the derivative jump of mode n at the defect is proportional to the value of the wavefunction $\chi_n(y=y_i)$ at the defect.

Eq. (27) implies that the derivative jump along the x-direction of propagation for mode n is proportional to the derivative jump of any other normal mode, since the $\sum_m$ in Eq. (27) is simply a number independent of n. Suppose we are at the second subband minimum, where the derivative jump of the second normal mode must be zero simply by wavefunction continuity, *i.e.* $(d\alpha_2/dx)|_{x=0+} - (d\alpha_2/dx)|_{x=0-} = 0$ when $E = E_2$. Then, from Eq. (27), the derivative jump of the the first normal at the scatterer must also be zero. Therefore, the lowest normal mode cannot reflect when the Fermi energy is equal to the second subband minimum.

Although this argument does not explain physically why the electron cannot reflect, it does prove that the scattering boundary conditions are not obeyed if any reflection occurs from the point scatterer at a subband minimum. If the scatterer is not a delta function, reflections can and do occur at a subband minimum without violating any scattering boundary conditions. However, this argument shows that the evanescent modes, which one might initially guess are unimportant in transport since they carry no current, actually dominate the scattering behavior of the conductor when the Fermi energy is near the openening of a new confinement subband. If the scatterer is not a point defect, but has some finite size which is small compared to the electron wavelength, we also expect the scattering to be reduced by this mechanism when the Fermi energy is near a subband minimum. But the dominant and most likely observable effect of evanescent modes, for any shape of an attractive scattering potential, is to form 'quasi-bound' or 'quasi-donor' levels below the subband minima.

3.3 Multi-mode Transmission Through Two Barriers

New physics emerges when an electron scatterers from two or more barriers placed in a narrow wire.[11] Figure 12 illustrates the basic idea leading to these new scattering properties: Evanescent waves now accumulate around both barriers in the wire, whereby some of the evanescent tail of the waves building up around one barrier can now extend to the second barrier. If the separation between the barriers along the x direction is d, and the decay length of the evanescent waves is κ, this effect is important when $d \leq \kappa$. However, besides simply determining κ, the position of the Fermi energy relative to a new confinement subband is also important. To produce a large effect on the transmission properties of the wire, it is not enough that the evanescent waves simply overlap. When the Fermi energy is away from a subband minimum, the modifications to electron scattering from the overlap of these evanescent waves are small because the evanescent density of states is also small. When the Fermi energy is near a subband minimum, however, the density of evanescent states is large so that the evanescent states dominate transport as they did for scattering from a single barrier. The decay length of the evanescent wave is also long near a subband minimum, so it is not difficult for the two tails of these evanescent waves to overlap.

Scattering from two barriers in a quasi-one-dimensional wire is calculated in Fig. 13 to illustrate these ideas. Two point scatterers, where the second scatterer is five times stronger than the first, are placed in a narrow wire and the resulting transmission coefficient T_{11} computed. When the Fermi energy is below the second subband minimum, $E = 0.9E_2$ in Fig. 13(a), the standard one-dimensional resonant tunneling picture generally holds. Only the first two transmission minimum in Fig. 13(a) are slightly altered due to the overlapping evanescent tails of the wavefunction. However, since the Fermi energy is not too near a subband minimum, the effect is small.

When the Fermi energy is placed on the second subband minimum, $E = E_2$ in Fig. 13(b), the transmission properties change drastically. There are three significant changes to usual resonant tunneling picture in Fig. 13(b) induced by the evanescent modes: (1) The transmission maxima are perfect even though the scatterer strengths are not equal. Furthermore, even though each barrier at $E = E_2$ taken alone is perfectly transmitting, the transmission coefficient when two such barriers are placed in the wire is perfect only at certain resonant values of the scatter separation. (2) The spacing between transmission maxima has changed from a half-wavelength to a full wavelength. (3) The transmission coefficient is not periodic, it approaches unity as the separation between the scatterers increases.

The essential conclusion to draw from Fig. 13(b) is that, when the Fermi energy aligns with a subband minimum in a narrow wire, the wave-interference between propagating modes no longer determines the electron transmission and subsequent electrical conductance of the wire. One would expect some type of Fabry-Perot resonances in Fig. 13(b) if this were not true, since both barriers are still coupled to the lowest propagating wavefunction in the wire. However, properties (1)-(3) of the preceeding paragraph show that

[11] All work in this section was performed in collaboration with Arvind Kumar.

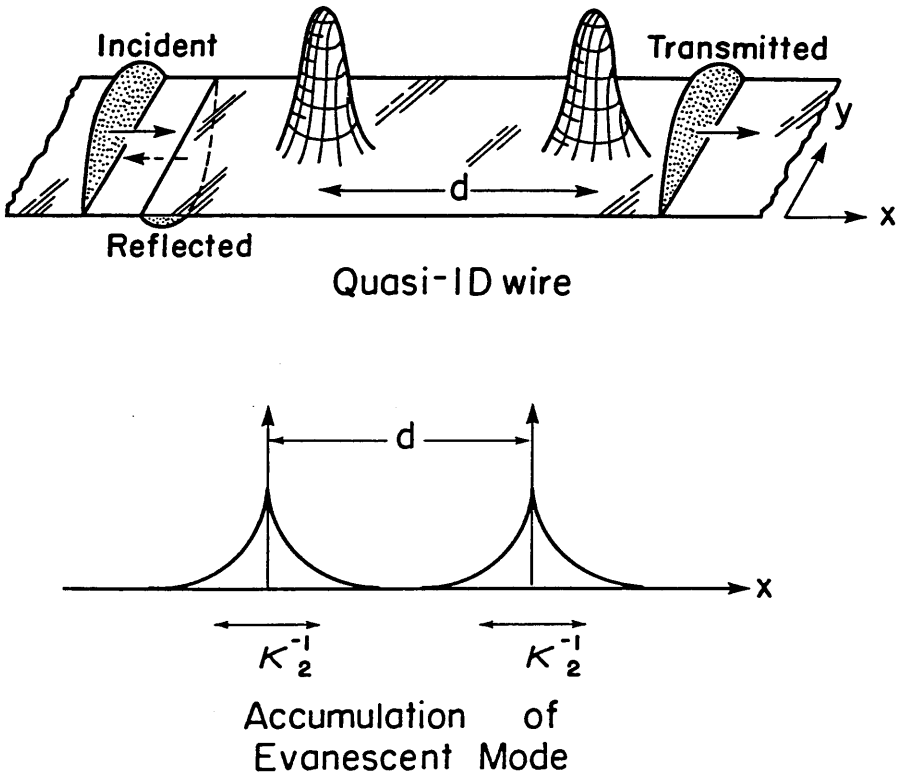


Figure 12: When an electron scatterers from two barriers in a narrow wire, evanescent waves accumulate around both of the barriers. When the tail of the evanescent wave accumulating around one of the barriers can reach to the second barrier, and the Fermi energy is near a subband minimum, the electron scattering and the resulting electrical conductance is strongly modified. From Appendix I.

the transmission oscillations when $E = E_2$ in Fig. 13(b) arise from a completely different mechanism, described in detail in Appendix I. The shape of the evanescent waves accumulating around the barriers determines transmission when $E = E_2$, rather than the usual wave interference between propagating modes, so that the usual Fabry-Perot resonances are completely suppressed and a new series of transmission oscillations with completely different properties emerges. This new series of conductance oscillations at $E = E_2$ can be understood by analogy to 'bonding' and 'anti-bonding' like states forming around the scatterers, and the interplay of the incident waveguide modes with these molecule-like states.

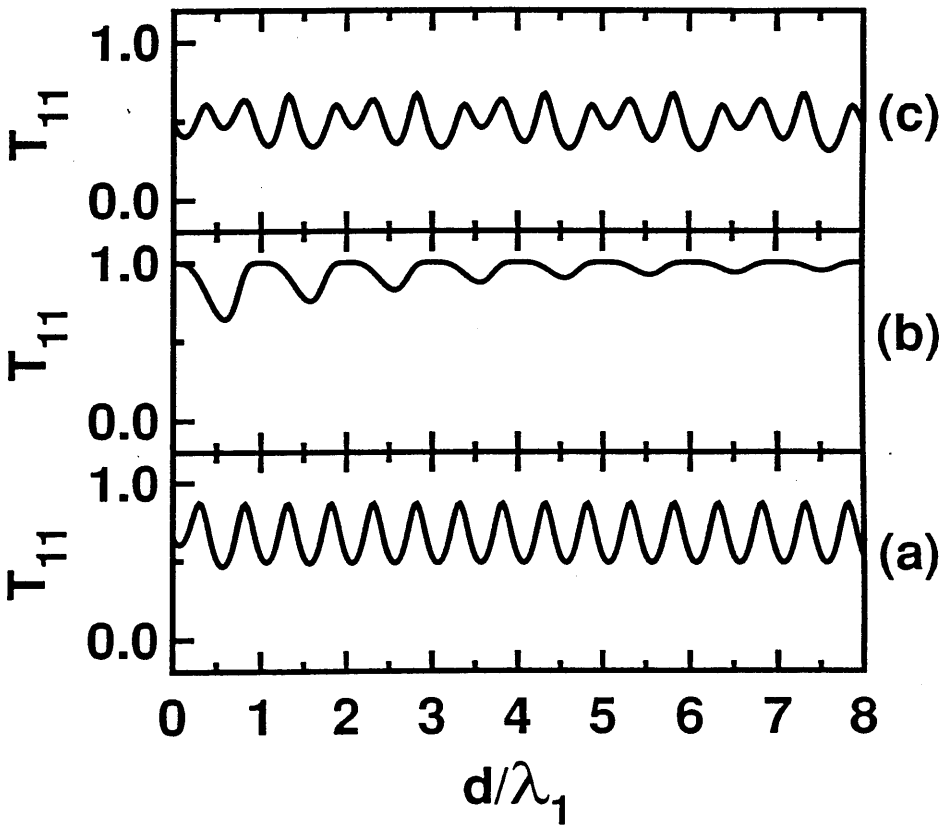


Figure 13: Transmission coefficient T_{11} versus distance d/λ_1 between two point scatterers in a narrow wire, where λ_1 is the electron wavelength of the lowest normal mode incident on the barriers. Curve (a) takes $E = 0.9E_2$, (b) sets $E = E_2$, and (c) sets $E = 1.1E_2$. From Appendix I.

4 Resonant Tunneling and Periodic Potentials

Resonant tunneling devices have received much attention in the past decade due to their large peak to valley current ratios at room temperature, their ability to operate at ultra-high frequency, their relatively large operating currents, and their nearly ideal current-voltage characteristics. All these properties combine to make a seemingly ideal quantum device with the ability to operate at room temperature. Several groups have explored the circuit possibilities of resonant tunneling three terminal devices, focusing primarily on multiple valued logic.

Yet the resonant tunneling diode shows ideal single electron behavior only on first glance. The resonant tunneling diode is in fact a non-ideal quantum device, with quite complicated internal dynamics. For example, the valley current is larger by a factor of 10-1000 in these devices than a simple quantum mechanical scattering calculation permits. It has been speculated that inelastic phonon scattering, elastic impurity scattering, and scattering through secondary conduction band valleys inside the tunneling barriers all play a significant role in increasing the valley current. Furthermore, the quantum well charges and discharges as the diode voltage is swept, giving rise to hysteresis or 'intrinsic bistability' in the I-V characteristics. If the transmission through the barriers is low, the intrinsic device bistability can be large.[12] Finally, the magnetoresistance of the tunneling diode, when a magnetic field is applied along the direction of the current, is completely non-ideal. An ideal device would show a dramatic magnetoresistance, where the triangle-like shape of the I-V characteristic would evolve into a large step. However, actual tunneling diodes show very little 'longitudinal' magnetoresistance even up to 30 Tesla fields. It is not known why the magnetoresistance displays such strong non-ideal behavior.

Despite these non-idealities in real resonant tunneling diodes, it is still useful to understand what an 'ideal' theory would predict for the I-V curves of different types of resonant tunneling devices. Resonant tunneling devices built in inversion layers or using lateral transport along GaAs/AlGaAs heterojunctions in a MODFET may also display closer to ideal behavior than those embodied using vertical transport through GaAs/AlGaAs heterojunctions. Also, for unknown reasons, the resonant tunneling diodes grown using molecular beam epitaxy display relatively close to ideal behavior when the magnetic field is zero. Figure 14 shows the expected ideal behavior of the currents in resonant tunneling devices having a (a) one, (b) two, or (c) three dimensional electron emitter, obtained using the 'convolution method' of the chapter 2.

The 1D device of Fig. 14(a) simply traces the transmission coefficient, displaying a single resonant peak at E_r and rising to unity at E_T as the Fermi energy passes above the tunneling barriers. The dependence of the 2D and 3D currents in Fig. 14 on the Fermi energy are obtained by classically adding many quantum point contact resistances in parallel. A simpler physical picture than that given in Appendix E can be used to understand this statement.[13] We understand that the 1D device current versus Fermi

[12] An additional 'bistability' can also arise if the external measurement circuit of the diode is improperly adjusted. But it is now fairly clear that intrinsic bistability of the tunneling diode is a real effect.

[13] I thank Ray Ghanbari for suggesting this picture.

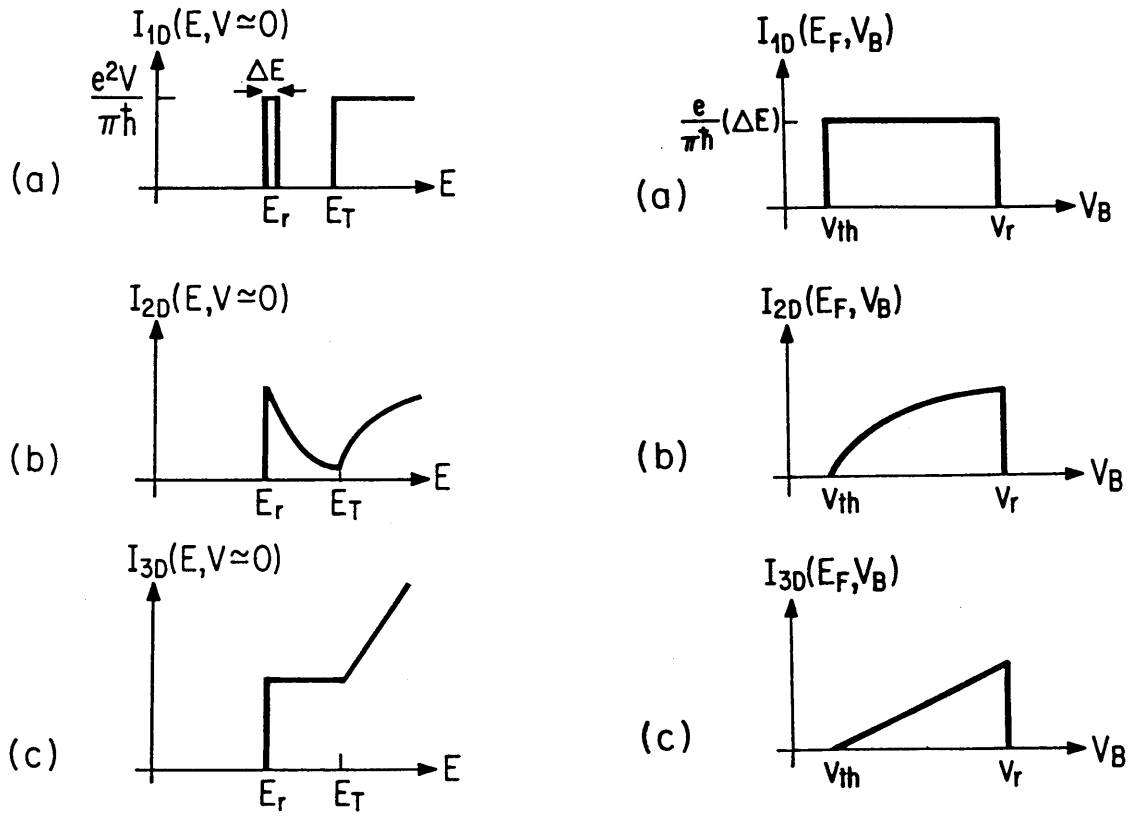


Figure 14: Ideal current versus emitter Fermi energy and current versus base voltage (or source to drain voltage) of (a) a 1D emitter, (b) a 2D emitter, and (c) a 3D emitter double barrier resonant tunneling device. From Appendix E.

energy simply traces the transmission coefficient through the double barriers. From this 1D transmission coefficient, we obtain the dependence of current on the Fermi energy for a 2D device as follows: Consider the 2D tunneling geometry of Fig. 15. The Hamiltonian

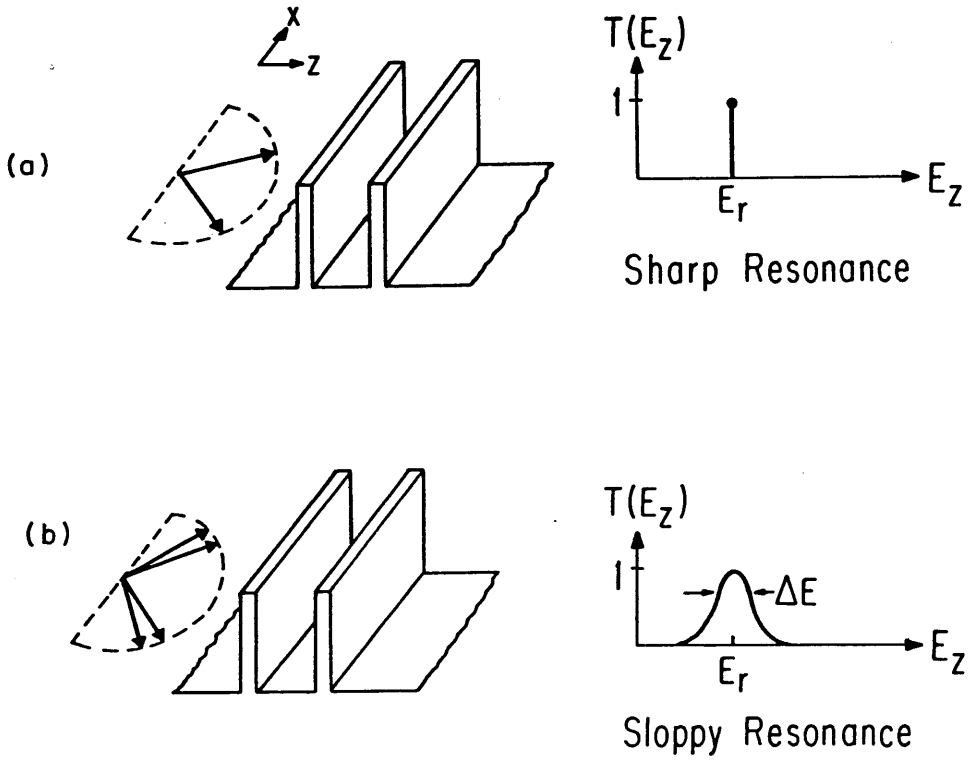


Figure 15: Transmission through a two-dimensional Fabry-Perot resonant cavity. The perfect cavity in (a) has a sharp momentum selection rule. If the cavity mirrors do not completely reflect electrons as in (b), there is some 'sloppiness' in the momentum selection, permitting a small range of k-vectors to pass through the barriers.

separates as described in Chapter 2 section 3, so we can consider the transmission coefficient as a function of E_z.

In Fig. 15(a) we assume each Fabry-Perot mirror is nearly perfectly reflecting, so that the resonant peak of the transmission coefficient is extremely sharp. We idealize it as letting through only electrons with a precisely defined wavevector $k_z = k_r$ as

$$T(E_z) = \delta_{E_z,E_r} \quad , \tag{28}$$

where δ is the Kroenecker delta function. This transmission coefficient has a resonant peak of zero width. Yet the barriers still permit electrons at the precisely defined wavevector k_r to pass, and only those electrons. The total current I for the sharply defined momentum 'filter' is

$$I(E) = \frac{e^2 V}{\pi\hbar}[2\delta_{E_z,E_r}] \quad . \tag{29}$$

The factor of two enters the conductance because there are two momentum directions allowed through by the mirrors. These two 'wavevector current channels' each contribute a conductance quantum to the total current, because they each propagate through the double barriers independently. This occurs because there are no terms in the 'separable' Hamiltonian which could cause one momentum direction to scatter into the other.

Next, we consider a Fabry-Perot cavity where the mirrors are no longer perfect. The resulting sloppiness in the momentum selection rule allows k-vectors having a small range of angles to pass through the barriers, as shown in In Fig. 15(b). This sloppiness is described by a transmission coefficient having a finite width, usually a Lorentzian of the form

$$T(E_z) = \frac{(\triangle E)^2}{(E - E_r)^2 + (\triangle E)^2} \quad . \tag{30}$$

For our purposes, the exact form of the sloppy transmission coefficient is unimportant; we simply approximate it by saying that electrons having E_z within an energy window $\triangle E$ near E_r transmit through the mirrors with probability one.

The number of electrons at an energy E transmitted through the imperfect cavity can now be simply estimated. In this two-dimensional problem the energy is determined as $E = E_z + E_x$. The electron kinetic energy available for motion transverse to the barriers is $E_x \simeq E - E_r$. The number of electron states available at this energy is simply the density of states for motion along the x direction. The energy range near $E_x \simeq E - E_r$ is $\triangle E$ due to the sloppy mirrors. And the number of 'point emitters' in space along the device can be simply obtained by multiplying the spatial density of point emitters by the length L_x of the device along the x-direction. These considerations give the number of transmitted electrons as

$$\text{number of electrons} = L_x(\triangle E)N_{1D}(E - E_r) \quad . \tag{31}$$

Each transmitted electron then carries a conductance quantum of current, so that the total current is

$$I_{2D}(E) = \frac{e^2 V}{\pi\hbar}[L_x(\triangle E)N_{1D}(E - E_r)/2] \quad . \tag{32}$$

We have simply classically added all the momentum channels in parallel, where each channel carries the quantum contact conductance of current since it is perfectly transmitting. Dividing by the extra factor of 2 in Eq. (32) is required so as not to double count the electron spin, a detail for our purpose here.

Eq. (32) describes only how the current carried by resonant electrons in Fig. 15 depends on the Fermi energy. We see from Eq. (32) that the dependence is $1/\sqrt{E - E_r}$ for the sloppy Fabry-Perot cavity, in contrast to the constant dependence of the current on Fermi energy from Eq. (29) when the cavity is nearly ideal. The sloppy mirrors are not choosy in deciding which k-vectors may pass, so that the current in Eq. (32) is much larger than in Eq. (29). On the other hand, the conductance versus Fermi energy for the sharp mirrors shows essentially one-dimensional behavior. One merely obtains the quantum contact resistance times the two possible independent momentum directions. There is also a nonresonant piece to the current on the left side of Fig. 15 which is examined in Appendix E.

The $1/\sqrt{E - E_r}$ decay of the 2D current on Fermi energy from Eq. (32) is easily pictured. As the incident electron energy E increases, the energy for motion of resonant

electrons perpendicular to the tunnel barriers E_z must also increase, since $E_z \simeq E - E_r$ for those electrons. Stated differently, resonant electrons must travel at a greater angle from normal incidence to satisfy the wavelength matching condition of the cavity. The number of available transverse momentum states for this motion decreases in one-dimension like $1/\sqrt{E}$, and is the origin of the decreasing current with Fermi energy above the resonant peak E_r in Fig. 14(b). Even though the resonant electrons must travel at an increasingly larger angle from normal incidence to the tunnel barriers as their emitted energy increases, the current carried by resonant electrons along the z direction is independent of this angle. This is easily seen when the tunnel barriers produce an extremely sharp transmission resonance, where one merely obtains twice the quantum contact conductance independent of energy.

The net resonant current depends only on the component of momentum directed along the tunnel barriers, not perpendicular to them. For resonant electrons, the velocity along the tunnel barriers is independent of the incident electron energy. Therefore, I emphasize again that, as the Fermi energy increases, the resulting increasing angle from normal incidence required to match the cavity resonance is *not* the mechanism leading to the decrease in current in Fig. 14(b). Instead, this negative transconductance arises from a reduction in the available number of transverse momemtum states to move perpendicular to the barriers at high energy.

The same arguments above also apply when the device emitter is three-dimensional. If the transmission coefficient is broad, as for the sloppy mirrors of Eq. (30), we find a total current

$$I_{3D}(E) = \frac{e^2 V}{\pi \hbar} [L_y L_x (\triangle E) N_{2D}(E - E_r)/2] \quad . \tag{33}$$

There are now two directions of free motion parallel to the tunnel barriers, which must be described by the two dimensional free electron density of states $N_{2D}(E - E_r)$. Since the 2D density of states is a constant independent of energy, the current versus Fermi energy in the 3D device of Fig. 14(c) is constant.[14]

The dependence of current on the device voltage can also be thought of as classically adding quantum point contact conductances in parallel in energy space. The applied voltage introduces an energy range eV into conduction, and all the carriers in that energy range must be considered. This is accomplished by convolving with the 'voltage box' or 'voltage window' of Chapter 2. Integrating the current under the sharp resonant peak gives the constant dependence of current on the 1D device voltage in Fig. 14(a). Similarly, the voltage dependence of the 2D and 3D devices (curves on the right of Fig. 14) are obtained by integrating over the range of Fermi energies eV (curves on the right of Fig. 14) filled by the device emitter. This is explained in detail in Appendix E.

Electrical conduction through periodic potentials has been discussed using many different methods, models, and approximations. There has also been a recent experimental resurgence of magnetoconductance studies through periodic potentials formed in inversion layers. The low-field structure of the magnetoconductance in these experiments however, is essentially classical in origin, being simply a classical geometrical resonance between the

[14]It is interesting to imagine what the analogy to Eq. (29) describing sharp mirrors would be if the device emitter is three-dimensional.

cyclotron orbit size and the periodic structure [86]. It does not demonstrate quantum conduction through a periodic potential. Furthermore, a recent experiment claiming to observe superlattice 'minibands' in a finite sample, may be in fact only electron 'charging' of the small constructed regions [87]. Clear 'quantum' superlattice effects in a two-dimensional electron gas have yet to be observed.

The elusive search for 'Bloch oscillations' has long been lurking in the background of studies of conduction in periodic potentials. These oscillations have never actually been observed to the knowledge of this author, and previous claims were undoubtedly overstated. They were originally sought in optical absorption near the band edge of bulk semiconductors [88] in the 1960's and 1970's, but the high electric fields required combined with broadening mechanisms made the search unsuccessful. The longer spatial periods in manmade structures, combined with long coherence lengths, make the search once again somewhat feasible.

An interesting real space picture of these Bloch oscillations, related to the Stark ladder states, has been proposed by Capasso and collaborators [89]. Assume two identical quantum wells are brought close together so that the levels split by an amount $\triangle E$. If an electron is initially localized to one of the wells, it will oscillate back and forth between the wells in time with a period $T = \hbar/\triangle E$. For the Stark ladder states, the separation between the levels is simply $\triangle E = eV_W = eV/N$, where V is the total voltage and N the number of superlattice periods. Therefore, the time for an electron to oscillate from one well in the superlattice to an adjacent well and back is simply $T = h/eV_W$.

In contrast, the Bloch oscillation period T_B can be derived from the semiclassical equations of motion for an electron wavepacket, in the limit where the applied voltage is low enough so that the energy band picture is still approximately valid. One straightforwardly obtains $T_B = h/(aeE)$ in this picture, where a is the superlattice period and E the uniform electric field. Of course $aE = V_W$ is the voltage drop across a single quantum well, so these two pictures predict an equivalent oscillation frequency. The oscillations between adjacent wells in the Stark ladder is asserted in Ref. [89] to be an equivalent real space picture of the Bloch oscillations.

Presumably, the DC current will reflect these Bloch oscillations by falling, perhaps to zero, when the oscillations begin. However, the appearance of Bloch oscillations is difficult to reconcile with a Landauer transport formalism. One can easily calculate the transmission coefficient through a periodic potential to obtain a DC current at all energies. Therefore, a finite DC current should exist at all energies. Yet, in the usual semiclassical Bloch oscillation picture, the DC current completely disappears when the oscillations begin and only an AC current survives. What additional physics needs to be considered in order to resolve this conflict is unclear to the present author. How to calculate the AC current response in a Landauer transport picture is also an open problem. One would also need to consider processes where one frequency applied to the conductor is converted into another.

5 Silicon Grating Gate MOSFET

Silicon 'grating gate' MOSFET devices were fabricated at M.I.T. in collaboration with Anthony Yen [90] as described below. The 'dual gate' structure of the device is shown below in Fig. 16, and permits easy control of the inversion layer geometry from many narrow wires in parallel, to a modulated potential, and to a two-dimensional electron gas. This control is achieved via electric field-effect by varying the voltages on the two gates shown in Fig. 16. The existence of the grating gate (but not its details) could be distinguished in an optical microscope, despite its period being smaller than the wavelength of light. See Figs. 3.13 and 3.14 of Ref. [90] for these pictures.

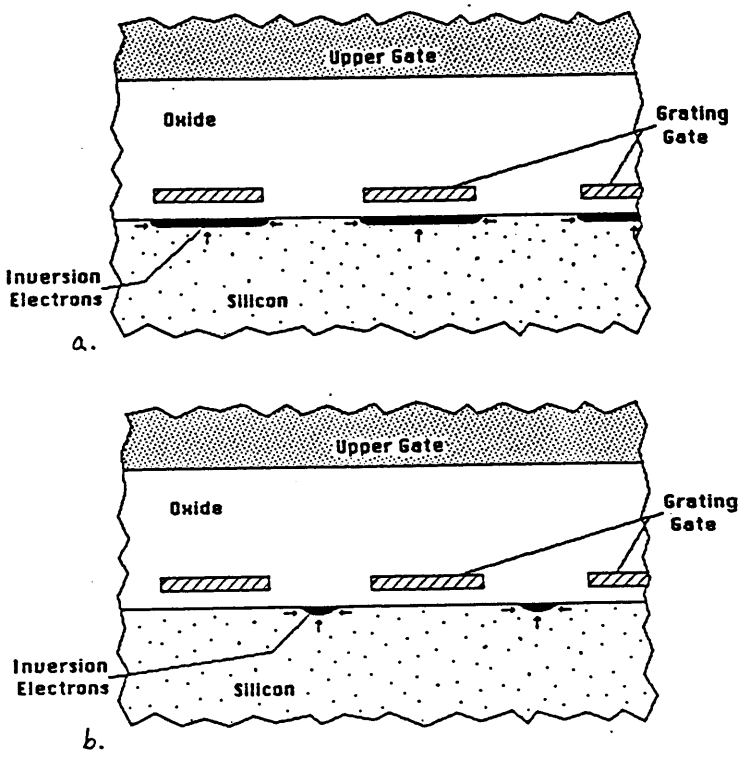


Figure 16: Cross section of the Silicon 'grating gate' MOSFET, showing the buried bottom gate and continuous top gate. Using field-effect control, the electron gas can be confined either underneath the gate wires or in the gap between wires. A modulated electron density or a two-dimensional electron gas can also be achieved in this geometry.

Electrical conductance of devices similar to the one in Fig. 18 was measured in temperatures ranging from 50 mK to 300 K, in magnetic fields ranging from 0 T to 30 T, and over the entire range of inversion layer geometry controlled by the gate bias. Several different devices were measured, all showing qualitatively the same dependence of the conductance on the various parameters. Measurements were done at low drain bias using an

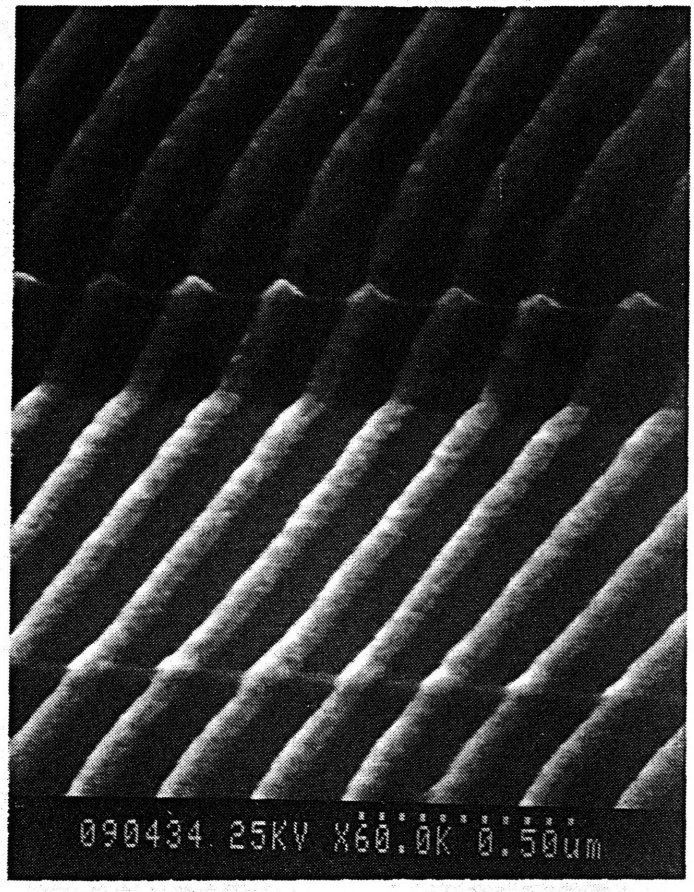


Figure 17: Grating gate running over the locally oxidized isolation areas (LOCOS). The grating lines proved quite robust over variations in surface topology. The small step in the surface near the bottom of the picture is where the gate oxide has grown at different rates over bare Si and a Phosphorous implanted region.

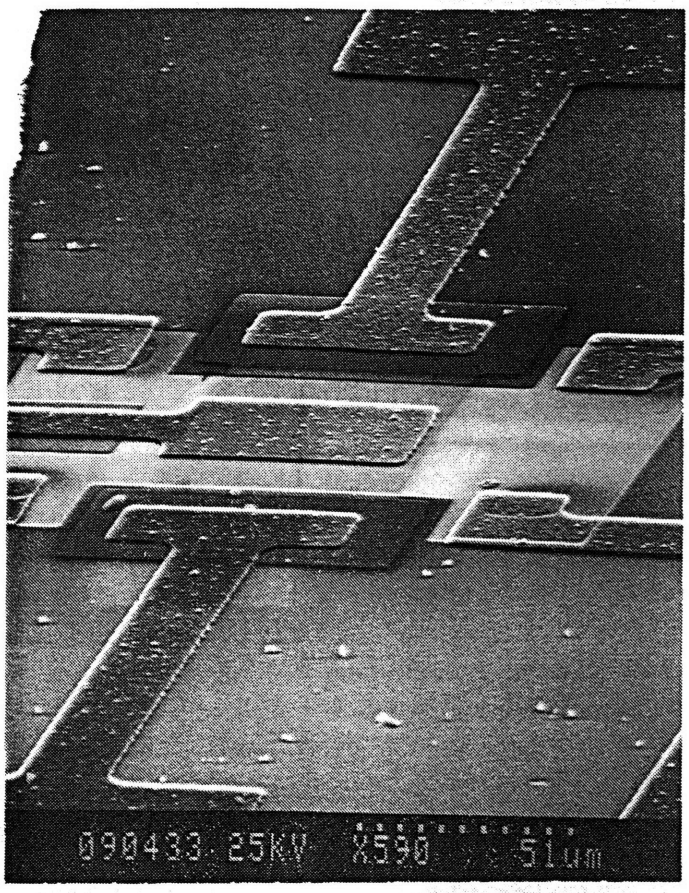


Figure 18: Scanning electron micrograph 'birds eye' view of the 'grating gate' MOSFET. The slightly discolored region running from source to drain is the grating gate.

AC lock-in amplifier and isolation transformer. Ultra-low temperature dilution refrigerator measurements were carried out in Prof. Kastner's laboratory at M.I.T, and additional measurements around 1.2 K but at higher magnetic fields were done at the Francis Bitter National Magnet Laboratory.

5.1 Physical Insights Gained

Several interesting magnetoresistance effects were observed in this device including modifications to electron weak localization (for low magnetic fields below 1 Tesla), an unusual manifestation of the classical Drude low-field magnetoconductance (for intermediate magnetic fields between 1-10 Tesla), and the approach to the quantum Hall effect (for high magnetic fields between 10-30 Tesla). We do not wish to repeat our detailed analysis of these effects, since that is done in Appendix M, but only to briefly illustrate the types of effects observed in this device.

Electron weak localization becomes much more pronounced as the inversion layer geometry is electrostatically pinched from a two- dimensional inversion layer into many narrow wires in parallel, proving that the wire width can be reduced below the electron phase coherence length. This is shown in Fig. 19 where one can see that, by pinching the electron gas into narrow wires, a pronounced depression in the magnetoconductance develops around $B = 0$. The classical diffusion paths of electrons, which originally were unconstrained in a two-dimensional electron gas, must now collide with the sides of the wire. Weak localization is extinguished approximately when the average size of a phase-coherent diffusion orbit encloses a magnetic flux equal to the flux quantum $\phi_0 = h/2e$. Weak localization is therefore destroyed at a magnetic field B such that $\phi_0 \simeq BL_\phi^2$ in the two-dimensional sample. In the reduced dimensional sample the size of the phase-coherent diffusive orbits is obviously much smaller, because the wire width W satisfies $W \leq L_\phi$, so weak localization persists to a much higher magnetic field such that $\phi_0 \simeq BWL_\phi$. In addition there will be flux cancellation, as in a SQUID magnetic field gradiometer, between 'figure eight' paths that cross each other. This additional flux cancellation serves to permit weak localization at even higher magnetic fields in a narrow wire.

In Fig. 19 we held the inversion layer geometry fixed (by holding the gate voltages constant) and varied the magnetic field. We can do the conductance measurement in a different way, where we fix the magnetic field and vary the inversion layer geometry. Typical conductance versus gate voltage data for a 15 Tesla magnetic field normal to the Si-SiO_2 interface is shown in Fig. 20. Each curve is for a different temperature ranging from 4.2 K to 300 K, with the larger conductance occurring at lower temperatures. It can be seen that the large negative transconductance persists to room temperature, indicating a classical origin. The dotted line is the conductance at 4.2 K in zero magnetic field. Consistent with Fig. 19, the magnetoconductance from Fig. 20 is relatively small when the gate voltage is less than zero volts, when conduction occurs through many narrow parallel wires. As the device geometry opens to a modulated potential above 0.0 V the conductance falls, reaches a minimum when the potential is an unmodulated two dimensional electron gas at around 10 V on the top gate, and rises again when the gate voltage reinduces a modulated potential as the gate voltage rises above 10 V.

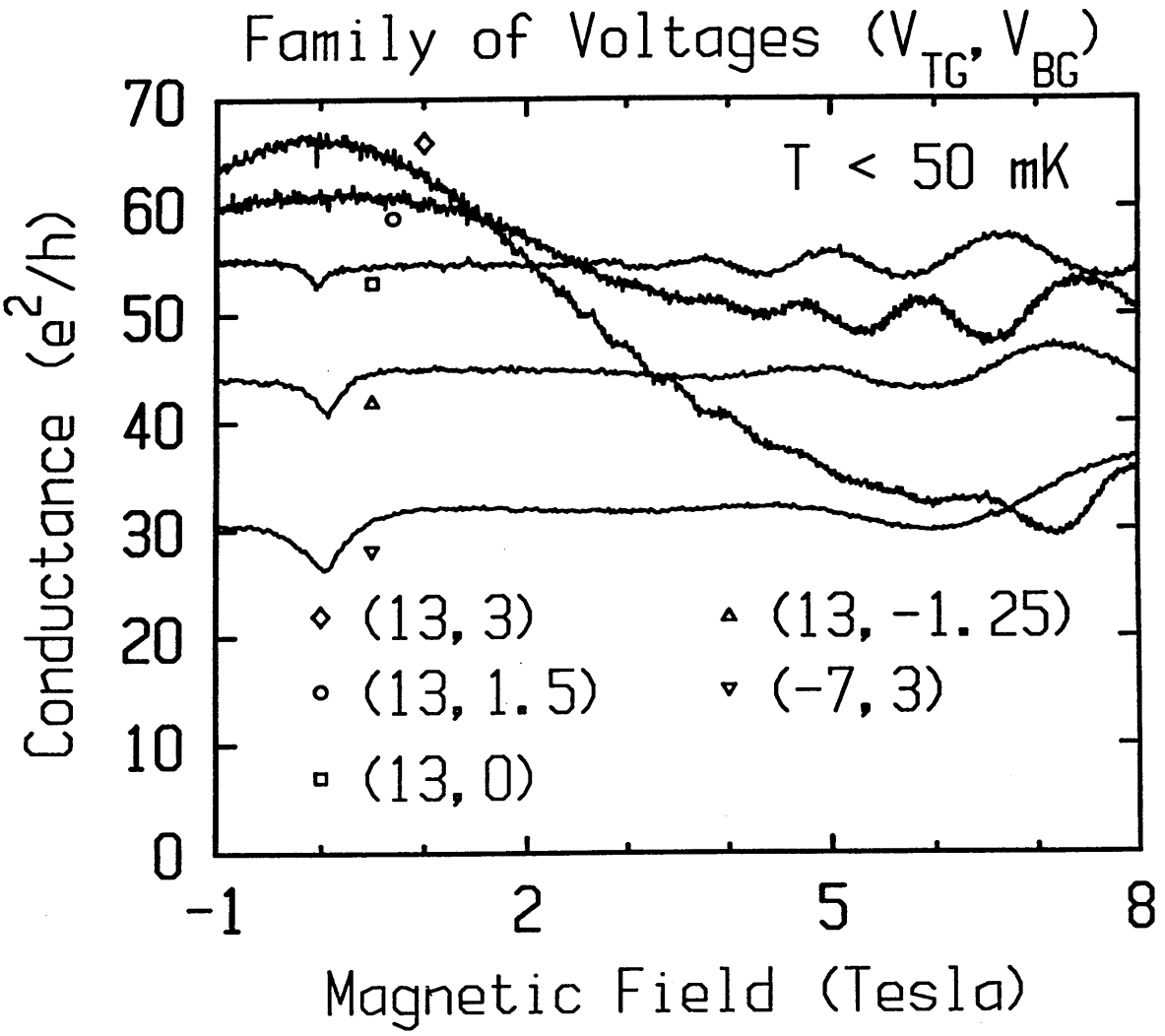


Figure 19: Two terminal magnetoconductance of the grating gate Si MOSFET showing an enhanced weak localization minimum around $B = 0$ as the inversion layer is pinched into narrow wires. Note the background magnetoconductance also changes as the inversion layer is pinched into narrow wires. Instead of decreasing with magnetic field as in a two-dimensional electron gas (top curve), the magnetoconductance remains essentially constant in a long and narrow wire (three bottom curves).

The significant observation in Fig. 20 is that the current decreases as more electrons are added to the inversion layer (in the presence of a large magnetic field). This is due to electrostatically changing the dominant two-terminal boundary condition on the classical Drude magnetoconductance tensor from that for a long and narrow to a short and wide geometry.[15] For a long and narrow conductor, the current must continue to flow essentially parallel to the sides of the wire, even at moderately high magnetic fields. Therefore, the two-terminal conductance is almost unaffected by the magnetic field. In contrast, for a short and wide device, the electric potential profile is fixed by the close proximity of source and drain, so that the current must flow across the device at the Hall angle. When our grating gate MOSFET is opened electrostatically from many wires in parallel to a two dimensional electron gas, the current path abruptly 'switches' to the Hall angle, giving the sharp drop in conductance seen in Fig. 20. The same argument also explains the change in background conductance from Fig. 19.

The conductance at high magnetic fields is described in Appendix M, but we list the qualitative features here. At high magnetic fields quantum edge states form in the wire array, giving opposite magnetoconductance for the parallel wires and wide electron gas. In other words, the approach to the quantum Hall effect is qualitatively different if the aspect ratio of the device is high or low. If the inversion layer is long and narrow, it will generally have a conductance much less than $4e^2/h$ so that the conductance must start to rise when the device begins to enter the quantum Hall regime. In contrast, a short and wide inversion layer will generally have a conductance much greater than $4e^2/h$, so that the conductance must continue to fall when the device enters the quantum Hall regime. At a magnetic field of 30 Tesla the conductance versus gate voltage of the narrow wires evolves into quantum Hall steps having a height of $4e^2/h$ multiplied by the number of wires in parallel. The evolution from Shubnikov-de Haas oscillations to the quantum Hall effect in these narrow wires qualitatively reproduces the 'anomalous magnetoresistance' described in Ref. [68] and [92].

5.2 Device Fabrication

I undertook the fabrication and measurement of Si 'grating gate' MOSFET devices jointly with Anthony Yen, as described in his S.M. Thesis [90]. A cross section of the device showing the two gates is given in Fig. 16. I concentrated on learning the more standard Si processing technology, while Yen's efforts were brought to bear on the novel grating-gate fabrication. After Anthony Yen's S.M. Thesis was written, the laboratory in which we were originally building the devices closed, and fabrication was moved to the M.I.T. Microsystems Technology Laboratories. The grating gate fabrication was still carried out in the Submicron Structures Laboratory using x-ray nano-lithography.

Moving device fabrication into the new laboratory resulted in a number of improvements which enabled the project to be carried successfully to completion. The lower gate oxide quality was measurably improved in the new furnaces. This thermally grown oxide was less leaky and its breakdown voltage higher than before. But more importantly, the

[15] Sam Park provided the correct interpretation for this part the experiment.

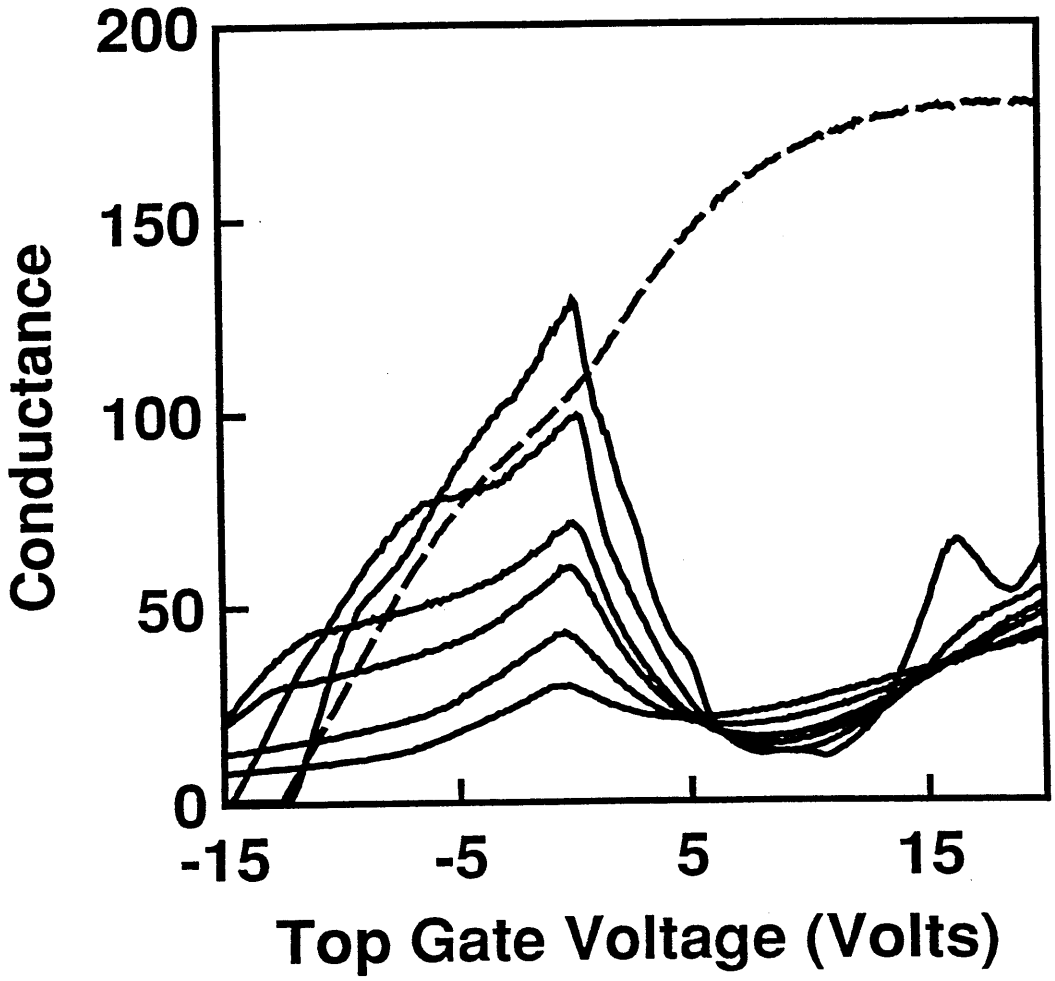


Figure 20: Two terminal magnetoconductance of the Si MOSFET for a 15 Tesla normal magnetic field at temperatures of 4.2 K, 20 K, 50 K, 100 K, 200 K, and 280 K. There is a sharp decrease in conductance, which persists to room temperature, as the device opens from many narrow wires in parallel to a two-dimensional electron gas. The electrostatic change in the aspect ratio of the inversion layer explains this classical phenomenon, where the current path 'switches' to flow at the Hall angle as the aspect ratio of the inversion layer changes.

top gate oxide could now be grown using high quality LPCVD oxide (low-pressure chemical vapor deposition), replacing the obsolete PYROX machine (thermal decomposition of silane gas in the presence of oxygen) grown oxide from the older laboratory. It is difficult to overemphasize how much the device characteristics improved because of this change. In addition, all the fabrication tools in the new laboratory were much better characterized, so that fabrication steps were more reproducible.

About this same time, a number of technological improvements in fabricating the grating-gate also occurred. A new deep-UV exposure tool was purchased for the Submicron Structures Laboratory, replacing the home built UV exposure tool we had previously used. Exposing the gate contact pads was now significantly easier, far less time consuming, more reliable, and uniform over a much larger area of the wafer than before. Also, a sequence of new x-ray masks was fabricated by Anthony Yen, eliminating the need to borrow a second hand x-ray mask and enabling excellent control over the grating line to space ratio. In addition, Yen developed an imaginative 'double shadowing' technique, which produced much more uniform grating lines. This technique can be appreciated by comparing Fig. 3.10 of Ref. [90] (before the double shadowing technique) with Fig. 17 above (using the new technique). The grating lines in Fig. 17 are noticeably smoother. Around this time feedback stabilization of the optical holographic technique, essential to fabricate x-ray masks, was also being developed. The x-ray masks could now be more reliably produced and were of higher quality. We also designed and installed a vacuum annealing furnace, evacuated to 10^{-7} torr by a turbomolecular pump, to anneal damage to the SiO_2 produced during x-ray lithography. In short, the grating gate fabrication was developing from a hit and miss operation into a reliable technology.

In the long run, grating fabrication proved the least of our difficulties in fabricating these devices. A large effort was already underway in the Submicron Structures Laboratory to improve x-ray mask fabrication and technology. Problems with the grating gate metallurgy proved more difficult. Anthony Yen's Cr-W-Cr metallization technology, combined with vacuum annealing, eventually made possible fair inversion layer mobilities. The idea behind high-temperature (900-1000 °C) vacuum annealing was to heat the gate oxide until it turned nearly liquid, thereby 'reflowing' the oxide and healing any defects. High temperature annealing in an inert gas would probably also have worked, but would have required more plumbing. We obtained useful assistance from John Scott-Thomas, who had run a matrix of times and temperatures for the anneal. If the grating metal was annealed too long or at too high a temperature, disasters of the type illustrated in Fig. 3.4 of Ref. [91] can occur. Various other disasters not shown in Ref. [91] also occur if the anneal is too aggressive. We did not stop to explore the materials science of these calamities. A lingering mystery is why the Chromium capping layer actually protects the grating lines from oxidation, since the lines are quite tall (27.5 nm) compared to the cap layer thickness (7.5 nm), so that the sides of the wires are exposed to oxygen.

As for the actual device fabrication steps themselves, the only major difference between the device runsheet given in the Appendix of Ref. [90] is that two masks were used to contact the final devices. The first masked the etching of source and drain contact holes. The second mask opened grating contact holes, but also had holes open for further overetching the source and drain contacts. Contacting the grating with HF etching was initially difficult. One did not wish to overetch and destroy the grating contact. (We solved the

problem of shorting the grating contact to the substrate by adapting a LOCOS geometry.) It is difficult to tell visually in a microscope if one has completely etched down to the grating gate contact or not. And a small oxide barrier left on top of the contact makes the device completely useless. Timing the etch is one solution, but not extremely reliable. Forming the device contact holes in two steps provided an unexpected solution to this problem. Usually the second mask was slightly misaligned from the first, so that a slightly misaligned source and drain contact was being etched on top of the original source and drain contact while the grating gate contact was being etched. When the 'lip' of the second and slightly misaligned source-drain contact was etched away, which could easily be determined by optical microscope, the etchant had now reliably uncovered the grating gate contact pad. Aluminum was then evaporated onto the both contacts simultaneously, patterned, and the device was sintered in forming gas. Many devices were also destroyed during bonding, both mechanically in the bonder and by electrostatic discharge through the gate oxide. To this author, device bonding still seems to be a black art.

6 Bibliography

References

[1] S. Datta and M.J. McLennan, 'A Review of Quantum Transport in Ultrasmall Electronic Devices', Reports of Progress in Physics, **53**, 1003 (1990).

[2] H. van Houten and C.W.J. Beenakker, 'Quantum Transport in Semiconductor Nanostructures' in *Solid State Physics*, H. Ehrenreich and D. Turnbull, eds., (Academic, New York, 1991).

[3] M. Büttiker, 'Symmetry of Electrical Conduction', IBM Journal of Research and Development, **32**, 317 (1988).

[4] M. Büttiker, 'Transmission, Reflection, and the Resistance of Small Conductors', in *Electronic Properties of Multilayers and Low-Dimensional Semiconductor Structures*, J.M. Chamberlain, L. Eaves, and J.C. Portal, eds., (Plenum, London, 1990).

[5] R. Landauer, 'Conductance Determined by Transmission: Probes and Quantized Constriction Resistance', Journal of Physics: Condensed Matter, **1**, 8099 (1989).

[6] R. Landauer, 'Electrons as Guided Waves in Laboratory Structures: Strengths and Problems', in *Analogies in Optics and Microelectronics*, W. van Haeringen and D. Lenstra, eds., (Kluwer, Academic Press, 1990).

[7] R.A. Webb and S. Washburn, 'Quantum Interference Fluctuations in Disordered Metals', Physics Today, **41**, 46 (1988).

[8] D.E. Prober, 'Quantum Transport in Microstructures', in *Microcircuit Engineering '86*, H.W. Lehmann and Ch. Bleiker, eds., (North-Holland, Amsterdam, 1986).

[9] M.A. Reed, ed., *Semiconductors and Semimetals*, (Academic Press, New York, 1990).

[10] M.A. Reed and W.P. Kirk, eds., *Nanostructure Physics and Fabrication*, (Academic, New York, 1989).

[11] K. von Klitzing, G. Dorda, and M. Pepper, 'New Method for High-Accuracy Determination of the Fine-Structure Constant Based on Quantized Hall Resistance', Physical Review Letters, **45**, 494 (1980).

[12] Gerd Bergman, 'Weak Localization in Thin Films: A Time of Flight Experiment With Conduction Electrons', Physics Reports, **107**, 1 (1984).

[13] P. A. Lee and A. D. Stone, 'Universal Conductance Fluctuations in Metals', Physical Review Letters, **55**, 1622 (1985).

[14] B. L. Al'tshuler, 'Fluctuations in the Extrinsic Conductivity of Disordered Conductors', Soviet Physics JETP Letters, **41**, 648 (1985).

[15] W. J. Skocpol, P. M. Mankiewich, R. E. Howard, L. D. Jackel, D. M. Tennant, and A. D. Stone, 'Universal Conductance Fluctuations in Silicon Inversion Layer Nanostructures', Physical Review Letters, **56**, 2865 (1986).

[16] D. Yu. Sharvin and Yu. V. Sharvin, 'Magnetic-Flux Quantization in a Cylindrical Film of a Normal Metal', JETP Letters, **34**, 272 (1981). A.G. Aronov and Yu. V. Sharvin, 'Magnetic Flux Effects in Disordered Conductors', Reviews of Modern Physics, **59**, 755 (1987).

[17] B.L. Al'tshuler, A.G. Aronov, and B.Z. Spivak, 'The Aaronov-Bohm Effect in Disordered Conductors', JETP Letters, **33**, 94 (1981).

[18] R.A. Webb, S. Washburn, C.P. Umbach, and R.B. Laibowitz, 'Observation of h/e Aharonov-Bohm Oscillations in Normal-Metal Rings', Physical Review Letters, **54**, 2696 (1985).

[19] B.J. van Wees, H. van Houten, C.W.J. Beenakker, J.G. Williamson, L.P. Kouwenhoven, D. van der Marel, C.T. Foxon, 'Quantized Conductance of Point Contacts in a Two Dimensional Electron Gas', Physical Review Letters, **60**, 848 (1988).

[20] 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, G.A.C. Jones, 'One Dimensional Transport and the Quantization of the Ballistic Resistance', Journal of Physics C: Solid State Physics, **21**, L209 (1988).

[21] M.L. Roukes, A. Scherer, S.J. Allen Jr., H.G. Craighead, R.M. Ruthen, E.D. Beebe, and J.P. Harbison, 'Quenching of the Hall Effect in a One-Dimensional Wire', Physical Review Letters, **59**, 3011 (1988).

[22] C.W.J. Beenakker and H. van Houten, 'Billiard Model of a Ballistic Multiprobe Conductor', Physical Review Letters, **63**, 1857 (1989).

[23] 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', Physical Review Letters, **62**, 583 (1989).

[24] U. Meirav, M.A. Kastner, and S.J. Wind, 'Single Electron Charging and Periodic Conductance Resonances in GaAs Nanostructures ', Physical Review Letters, **65**, 771 (1990).

[25] C.P. Umbach, P. Santhanam, C. van Haesendonck, and R.A. Webb, 'Nonlocal Electrical Properties in Mesoscopic Devices', Applied Physics Letters, **50**, 1289 (1987).

[26] W.J. Skocpol, P.M. Mankiewich, R.E. Howard, L.D. Jackel, D.M. Tennant, and A.D. Stone, 'Nonlocal Potential Measurements of Quantum Conductors', Physical Review Letters, **58**, 2347 (1987).

[27] L.L. Chang, L. Esaki, and R. Tsu, 'Resonant Tunneling in Semiconductor Double Barriers', Applied Physics Letters, **24**, 593 (1974).

[28] T.P.E. Broekaert, W. Lee, and C.G. Fonstad, 'Pseudomorphic $In_{0.53}Ga_{0.47}As/AlAs/InAs$ Resonant Tunneling Diodes With Peak to Valley Current Ratios of 30 at Room Temperature', Applied Physics Letters, **53**, 1545 (1988).

[29] M. Büttiker, 'Absence of Backscattering in the Quantum Hall Effect in Multiprobe Conductors', Physical Review B15, **38**, 9375 (1988).

[30] B.I. Halperin, 'The Quantized Hall Effect', Scientific American, **254-4**, 52 (1986).

[31] R.E. Prange and S.M. Girvin, eds., *The Quantum Hall Effect*, (Springer-Verlag, New York, 1987).

[32] A.M. Chang, G. Timp, T.Y. Chang, J.E. Cunningham, P.M. Mankiewich, R.E. Behringer, and R.E. Howard, 'Deviation of the Quantum Hall Effect From Exact Quantization in Narrow GaAs-$Al_xGa_{1-x}As$ Heterostructure Devices', Solid State Communications, **67**, 769 (1988).

[33] M. Büttiker, 'Four-Terminal Phase-Coherent Conductance', Physical Review Letters, **57**, 1761 (1986).

[34] A.D. Benoit, S. Washburn, C.P. Umbach, R.B. Laibowitz, and R.A. Webb, 'Asymmetry in the Magnetoconductance of Metal Wires and Loops', Physical Review Letters, **57**, 1765 (1986).

[35] Y. Imry, 'Physics of Mesoscopic Systems', in *Directions in Condensed Matter Physics*, G. Grinstein and G. Mazenko, Editors, (World Scientific Press, Singapore, 1986).

[36] R. Landauer, 'Electrical Transport in Open and Closed Systems', Zeitschrift fur Physik B, **68**, 217 (1987).

[37] M. Büttiker, Y. Imry, R. Landauer, S. Pinhas, 'Generalized Many-Channel Conductance Formula with Application to Small Rings', Physical Review B, **31**, 6207 (1985).

[38] It was extremely difficult for the present author to accept this result in the beginning. The early controversies in this field between prominent theorists did not help to make any claims plausible. Even the quantization of the ballistic conductance in a point contact did not gain my acceptance of this result, though in hindsight it should have. Understanding how the Drude-Sommerfeld conductance is obtained as the incoherent classical limit of Landauer's result eventually sealed the case for the present author. This illustrates the degree of trust the present author placed in the Drude-Sommerfeld formula when he erroneously believed it in possible conflict with Landauer's result. But, eventually, the many beautiful experiments which were explained using Landauer's transport formulae would have convinced even the most die-hard skeptic.

[39] W.A. Harrison, *Solid State Theory*, (Dover, New York, 1980). See section III.2.5 of this excellent book, especially following equation 3.25 where the resistance of a 1D

conductor is stated as being $\pi\hbar/(e^2T)$. When the author asked Dr. Harrison in March 1990 if he had ever thought about the limit $T \rightarrow 1$, which predicts a finite resistance even with no scattering, Harrison indicated that there were more important problems to work on at the time. The author personally thinks that early workers in the area of tunneling conductance did not have much confidence in these formulas, and believed them primarily in the limit of low electron transmission through a tunnel barrier. This situation should have been remedied after the discovery of resonant tunneling in semiconductors.

[40] B.K. Ridley, *Quantum Processes in Semiconductors*, (Clarendon Press, Oxford, 1988). This book is an excellent reference for 'Golden-Rule' and 'Boltzmann-like' approaches to transport calculations.

[41] S. Datta, *Quantum Phenomena*, (Addison-Wesley, Reading, Mass., 1989).

[42] H.C. Liu, 'Time-Dependent Approach to Double Barrier Quantum Well Oscillators', Applied Physics Letters, **52**, 453 (1988). H.C. Liu, 'High-Frequency Model of Double Barrier Resonant Tunneling', in *Resonant Tunneling in Semiconductors: Physics and Applications*, L.L. Chang and E.E. Mendez, eds., (Plenum, New York, in press).

[43] B. Kramer and J. Masek, 'Conductance of Finite Systems in the Ballistic Regime: The Dependence on Fermi Energy, Magnetic Field, and Frequency', Journal of Physics C: Solid State Physics, **21**, L1147 (1988). J. Masek and B. Kramer, 'Frequency Dependent Conductance of Linear Disordered Chain', Solid State Communications, **68**, 611 (1988).

[44] E.H. Hauge and J.A. Stovneng, 'Tunneling Times: A Critical Review', Reviews of Modern Physics, **61**, 917 (1989).

[45] M. Büttiker, 'Coherent and Sequential Tunneling in Series Barriers', IBM Journal of Research and Development, **32**, 63 (1988).

[46] S. Datta, 'Steady-State Quantum Kinetic Equation', Physical Review B, **40**, 5830 (1989).

[47] A. Stern, Y. Aharonov, and Y. Imry, 'Phase Uncertainty and Loss of Interference: A General Picture', (preprint).

[48] S. Feng, 'Quantum Transport in the Presence of Phase-Breaking Scattering: Generalized Landauer Formula', Physica A, **168**, 439 (1990).

[49] B.Y. Gelfand, S. Schmitt-Rink, and A.F.J. Levi, 'Tunneling in the Presence of Phonons: A Solvable Model', Physical Review Letters, **62**, 1683 (1989).

[50] M.J. McLennan, Y. Lee, and S. Datta, 'Where is the Voltage Drop? A Numerical Study Using a Quantum Kinetic Equation', (preprint).

[51] S. Datta, 'A Simple Kinetic Equation for Steady-State Quantum Transport', Journal of Physics: Condensed Matter, **2**, 8023 (1990).

[52] W. Frensley, 'Boundary Conditions for Open Quantum Systems Driven Far From Equilibrium', Reviews of Modern Physics, **62**, 745 (1990).

[53] Y. Meir, N.S. Wingreen, and P.A. Lee, 'Transport Through a Strongly Interacting Electron System: Theory of Periodic Conductance Oscillations', preprint.

[54] K.K. Likharev, 'Correlated Discrete Transfer of Single Electrons in Ultrasmall Tunnel Junctions', IBM Journal of Research and Development, **32**, 144 (1988).

[55] R. Landauer, 'Johnson-Nyquist Noise Derived From Quantum Mechanical Transmission', Physica D, **38**, 226 (1989).

[56] G.B. Lesovik, 'Excess Quantum Noise in 2D Ballistic Point Contacts', JETP Letters, **49**, 592 (1989).

[57] M. Büttiker, 'Scattering Theory of Thermal and Excess Noise in Open Conductors', Physical Review Letters, **65**, 2901 (1990).

[58] T. Holstein, 'Theory of Transport Phenomena in an Electron-Phonon Gas', Annals of Physics, **29**, 410 (1964). Linear response using the Fröhlich Hamiltonian.

[59] W. Kohn and J.M. Luttinger, 'Quantum Theory of Electrical Transport Phenomenon', Physical Review, **108**, 590 (1957). Perturbation theory in a periodic potential.

[60] C.B. Duke, *Tunneling in Solids*, (Academic Press, New York, 1969). Many particle treatment of electron tunneling using the Bardeen transfer Hamiltonian.

[61] Neil Ashby, 'Applications of Quantum Field Theory to Electrical Conductivity', PhD Thesis, Harvard University, 1961. (HU 90.7905.8) The conductivity of a two component gas reduces to the standard impurity averaging proceedure when one of the components becomes infinitely massive.

[62] P.F. Bagwell, 'Quantum Mechanical Transport Phenomenon in Nanostructured Inversion Layers', S.M. Thesis, Massachusetts Institute of Technology, (1988).

[63] S. Doniach and E.H. Sondheimer, *Green's Functions for Solid State Physicists*, (Benjamin, Reading, Mass., 1974).

[64] A.A. Abrikosov, L.P. Gorkov, and I.E. Dzyaloshinski. *Methods of Quantum Field Theory in Statistical Physics*, (Prentice-Hall, Englewood Cliffs, N.J., 1963).

[65] W.A. Schwalm and M.K. Schwalm, 'Extension Theory for Lattice Green Functions', Physical Review B, **37**, 9524 (1988). See Eq. (123).

[66] A.L. Yeyati, M. Weissmann, and E. Anda, 'Evaluation of the Kubo Formula for Conductivity Using the Recursion Method', Journal of Physics: Condensed Matter, **1**, 5429 (1989). See Eq. (21).

[67] M. Büttiker, 'Chemical Potential Oscillations Near a Barrier in the Presence of Transport', Physical Review B, **40**, 3409 (1989).

[68] S. Park, 'The Anomalous Magnetoresistance of the Electron Gas in a Restricted Geometry', Ph.D. Thesis, Massachusetts Institute of Technology, 1990.

[69] T. Ando, A.B. Fowler, and F. Stern, 'Electronic Properties of Two-Dimensional Systems', Reviews of Modern Physics, **54**, 437 (1982).

[70] T. Sugano, K. Hoh, H. Sakaki, T. Iizuka, K. Hirai, K. Kuroiwa, and K. Kakemoto, 'Quantum State and Electron Transport at Silicon-Silicon Dioxide Interface and MIS Device Technology', Journal of the Faculty of Engineering, The University of Tokyo (B), **32**, 159 (1973).

[71] R.W. Rendell and S.M. Girvin, 'Hall Voltage Dependence on Inversion Layer Geometry in the Quantum Hall Regime', Physical Review B, **23**, 6610 (1981).

[72] G.L.J.A. Rikken, J.A.M.M. van Haaren, W. van der Wel, A.P. van Gelder, H. van Kempen, P. Wyder, J.P. Andre, K. Ploog, and G. Weimann, 'Two-Terminal Resistance of Quantum Hall Devices', Physical Review B, **37**, 6181 (1988).

[73] The viewpoint we advance here is that changes in the *potential* energy of a state lead to energy level broadening. However, it is also possible for additional disorder to either reduce or add extra 'kinks' or oscillations in the wavefunction, thereby lowering or raising its *kinetic* energy. In the quantum Hall effect, where electron motion in the bulk follows equipotential contours, variations in the potential energy are primarily responsible for level broadening. One would expect that, unless wave interference plays a dominant role, energy level broadening is primarily due to changes in the potential energy of a quantum state.

[74] J.D. Dow, S.Y. Ren, and K. Hess, 'Random Superstructures', Physical Review B, **25**, 6218 (1982).

[75] Unfortunately, the histogram of the impurity potential $P(E)$ from Fig. 3 is not available to directly compare with any average calculated density of states $A(E,\tau)$. Indeed, the average density of states was not calculated in Fig. 3, though one can easily imagine its shape. Carrying out this calculation and comparing the average density of states $A(E,\tau)$ with the histogram of the potential energy $P(E)$ would make an interesting numerical experiment. As the impurity potential becomes increasingly non-adiabatic, one expects the agreement between the energy level broadening function $A(E,\tau)$ and the potential energy histogram $P(E)$ to worsen. The potential energy $V(x)$ in Fig. 3 is probably not adiabatic. For nearly adiabatic potentials we expect $A(E,\tau) \simeq P(E)$.

[76] K. Ismail, P.F. Bagwell, T.P. Orlando, D.A. Antoniadis, and H.I. Smith, 'Quantum Phenomena in Field-Effect-Controlled Semiconductor Nanostructures', (preprint).

[77] B.E. Sernelius, K.F. Berggren, M. Tomak, and C. McFadden, 'Effects of Quantum Confinement in a Special GaAs Field Effect Transistor: On the DC Conductance in the Regime of Metallic Transport', Journal of Physics C: Solid State **18**, 225 (1985).

[78] S. DasSarma and X.C. Xie, 'Calculated Transport Properties of Ultrasubmicrometer Quasi-One-Dimensional Inversion Lines', Physical Review B, **35**, 9875 (1987).

[79] M.J. Kearney and P.N. Butcher, 'The Effect of Lifetime Broadening on the Conductivity and Thermopower of a Quasi-1D Wire', J. Phys. C: Sol. State **20**, 47 (1987).

[80] N. Trivedi and N.W. Ashcroft, 'Quantum Size Effects in Transport Properties of Metallic Films', Phys. Rev. B **38**, 12298 (1989).

[81] P. Vasilopoulos and F.M. Peeters, 'Electrical Transport in a Quantum Wire: Influence of One- and Two- Body Interactions ', Phys. Rev B **40**, 10079 (1989).

[82] S.Y. Qiu and M.V. Jaric, Bull. APS **35**, 492 (1990). W. Bao and Z. Tesanovic, Bull. APS **35**, 492 (1990). K. Arya, J.P. Lu, and J.L. Birman, Bull. APS **34**, 459 (1989).

[83] M. Suhrke, S. Wilke, and R. Keiper, 'Conductance of Quantum Wires With Self-Consistent Level Broadening', Journal of Physics: Condensed Matter, **2**, 6743 (1990).

[84] J. Sone and S. Ishizaka, 'Effect of Diffusive Boundary Scattering in Quantized Conductance and Quenching of the Hall Resistance in Very Narrow Wires', Mat. Res. Soc. Extend. Abs., EA-26, 91, (1990).

[85] H. Sakaki, 'Scattering Suppression and High-Mobility Effect of Size-Quantized Electrons in Ultrafine Semiconductor Wire Structures', Japanese Journal of Applied Physics, **19**, L735 (1980).

[86] D. Weiss, M.L. Roukes, A. Menschig, P. Grambow, K. von Klitzing, G. Weimann, 'Electron Pinball and Commensurate Orbits in a Periodic Array of Scatterers', preprint.

[87] L.P. Kouwenhoven, F.W.J. Hekking, B.J. van Wees, C.J.P.M. Harmans, C.E. Timmering, C.T. Foxon, 'Transport Through a Finite One-Dimensional Crystal', Physical Review Letters, **65**, 361 (1990).

[88] J. Callaway, 'Optical Absorption in an Electric Field', Physical Review, **134**, A998 (1964).

[89] F. Beltram, F. Capasso, D.L. Sivco, A.L. Hutchinson, S.G. Chu, and A.Y. Cho, 'Scattering-Controlled Transmission Resonances and Negative Differential Conductance by Field-Induced Localization in Superlattices', Physical Review Letters, **64**, 3167 (1990).

[90] Anthony Yen, 'Grating Gate Si MOSFET for Study of Quantum Transport Effects', S.M. Thesis, Massachusetts Institute of Technology, 1987.

[91] John Scott-Thomas, 'Conductance Oscillations Periodic in the Charge Density of One Dimensional MOSFET Structures', Ph.D. Thesis, Massachusetts Institute of Technology, 1990.

[92] M.A. Kastner, S.B. Field, J.C. Licini, and S.L. Park, 'Anomalous Magnetoresistance of the Electron Gas in a Restricted Geometry', Physical Review Letters, 60, 2535 (1988)

Acknowledgements- Prof. Terry Orlando has both contributed to my understanding and stimulated my interest in quantum transport in too many ways to be written down in this small space. I could not have studied the topics in this thesis without his guidance. Prof. Dimitri Antoniadis and Prof. Henry I. Smith have always demanded that I make clearer explanations and better transistors, greatly improving the results presented here. Were it not for the influence of Prof. Marc Kastner, who encouraged me to more closely examine in his laboratory the Si grating gate transistors I had fabricated, I might have written them off as a failure. 'Anything which motivates an experiment' is indeed good.

I thank Arvind Kumar for being patient with me while he explained the volumes of work he produced. Keeping up with his methodical and unrelenting attack on transport problems made my last year at M.I.T. both hectic and enjoyable. During this same time period, I was eagerly pursuing some erroneous explanations for magnetoconductance of the grating gate MOSFET. Fortunately, Sam Park's insights brought me back to reality before I had wasted too much time. Soon we will undoubtedly resolve the energy level broadening controversy.

Tony Yen made possible the MOSFET work in this thesis with his fabrication skills and metallurgy experiments. I thank him for his enthusiasm and for not giving up when it looked like the whole project might be wasted. I also hope that he someday gets the full credit he deserves for making possible the devices built by John Scott-Thomas, in which single electron charging effects were possibly observed, since his Tony's experiments first demonstrated the feasibility of Cr-W-Cr metallurgy and vacuum annealing later refined by Scott-Thomas. I also thank Seiji Horiguchi from NTT Japan for working with some earlier (and unfortunately poorer) versions of these grating gate devices, and for his many theoretical insights which I should have better appreciated at the time. I regret that his early ideas have still not seen print.

I thank Khalid Ismail, Ray Ghanbari, John Scott-Thomas, Udi Meirav, Stuart Field, and Kevin Delin for constant discussions, debates, and experimental help. Norris Preyor, Jing Luo, John Kucera, Akira Toriumi, Mark Schattenburg, and Jimmy Carter made important contributions to the experimental work in this thesis. The staff of the MIT Microsystems Technology Laboratory were indespensible in fabricating high quality MOSFETs. I also thank Larry Rubin and Bruce Brandt of the Bitter Magnet Laboratory for help with high magnetic field measurements.

Tom Broekaert's and Prof. Clif Fonstad's achievements in building resonant tunneling diodes, along with Khalid Ismail's analogous work in a 2D electron gas, motivated me to try and better understand these devices. I thank them for their many insights and encouragement.

Ellen Spertus made me understand that recursion relations could be applied to classical diffusion, but I unforgivably neglected to acknowledge her contributions in the paper where I carried out those ideas. I apologize and thank her here.

I have greatly benefitted from discussions and correspondance with Supriyo Datta, Rolf Landauer, Markus Büttiker, Marc Cahay, Supriyo Bandyopadhyay, Bill Frensley, Charlie Kane, Xin Cheng Xie, Erkan Tekman, Chris Eugster, and Prof. Jesus del Alamo.

Lastly, I have been fortunate to have the continued love and support of my parents Jim and Eloise. Without their encouragement I could not have accomplished any of the work described herein.

7 Appendices: Publications

Appendix A

Philip F. Bagwell, Dimitri A. Antoniadis, and Terry P. Orlando, 'Quantum Mechanical and Nonstationary Transport Phenomena in Nanostructured Silicon Inversion Layers', in *Advanced MOS Device Physics*, edited by N. Einspruch and G. Gildenblat, (Academic, San Diego, 1989).

VLSI ELECTRONICS: MICROSTRUCTURE SCIENCE, VOL. 18

Chapter 8

Quantum Mechanical and Nonstationary Transport Phenomena in Nanostructured Silicon Inversion Layers

PHILIP F. BAGWELL
DIMITRI A. ANTONIADIS
TERRY P. ORLANDO

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

I. INTRODUCTION

This chapter discusses electron transport effects that are expected or have already been observed in ultrasmall three-terminal silicon MOS structures. These effects fall into three categories: quantum effects related to carrier confinement perpendicular to the direction of transport, quantum effects that result from carrier propagation along periodic potential variations with very fine spatial periodicities, and effects arising from nonstationary transport across very small distances.

We treat electrical conductivity throughout this chapter using the semiclassical Boltzmann transport theory in the relaxation time approximation. Sections II–IV deal with quantum mechanical effects generally observable at low temperature. There are several device structures that permit such observations. We refer to them collectively as "quantum MOS devices." Section II deals with the theory of conductivity in quantum MOS devices and derives necessary observability criteria for quantum effects in terms of characteristic energies, distances, or times. We apply this theory to treat the quasi-one-dimensional MOS transistor in Section III and the surface superlattice MOS transistor in Section IV.

In Sections II–IV, we assume that the carrier and lattice temperatures are the same. On the other hand, Section V deals with hot-carrier transport effects observable at room temperature. In this case the carrier temperature is typically higher than the lattice temperature. We show that the signature of quantum mechanical effects is conductance and transconductance oscillation with varying gate and drain bias conditions, while the signature of nonstationary transport is convective carrier velocities in excess of the saturation velocity.

II. SEMICLASSICAL CONDUCTIVITY OF QUANTUM MOS DEVICES

A semiclassical Boltzmann transport method of calculating electrical conductivity allows us to understand the current–voltage relationship in most silicon MOS devices. This semiclassical method uses quantum mechanics to calculate the allowed electron energy levels and the group velocity versus energy dispersion relationship of electron wave packets, but treats the electron as a point particle between collisions. Ignoring wave interference between collisions is the essence of a Boltzmann transport approximation. A full quantum mechanical treatment of electrical conductivity would include the possibility of electron waves interfering with each

other between collisions. Electron "localization" is one example of a non-Boltzmann transport process.[1] In silicon inversion layers, however, these bulk interference effects are generally weak, and the standard Boltzmann transport approximation is a good one. This particle-like approximation to the electron wave requires that we model the electron as a wave packet and use the semiclassical equations of motion for the wave packet.

Quantum mechanics also enters electrical conductivity in the calculation of a scattering rate. The effect of scattering on the momentum relaxation rate depends both on how often the particles scatter and on the probability of their being deflected through a given angle. Both of these are functions of the particle's energy and can be calculated in quantum mechanical scattering theory. Treating the electron as a wave during scattering events in no way conflicts with the standard Boltzmann transport approximation. However, if scattering events are closer together than a few electron wavelengths, the Boltzmann approximation begins to break down.

There are subtle relationships between scattering and electrical conductivity. Ohmic heating is a necessary consequence of power dissipation in a conductor. The electron gas must dissipate energy it gains from the electric field, and it cannot do so using purely energy conserving or elastic scattering processes. Inelastic or energy dissipating processes are therefore necessary to define resistance. But inelastic scattering disrupts the phase of the wave function, and is, thus, anathema to electronic devices that depend on the wave nature of the electron to give interesting structure to the conductivity.

It is also important to understand the effects of temperature, elastic scattering, and inelastic scattering in quantum MOS devices. So called "elastic" broadening of electron energy levels can average away interesting features in electrical conductivity. Yet the term "elastic" broadening is a misnomer. Inelastic processes are required for the incoherent averaging necessary to produce "elastic" energy-level broadening. Finite temperature also leads to an incoherent averaging of energy levels. The resulting temperature broadening of energy levels can average away interesting structure in the electrical conductivity.

To understand how these issues relate to quantum MOS devices in detail, we study a semiclassical model of electron transport that is equivalent to the Boltzmann transport equation in the relaxation time approximation. We consider the role of elastic and inelastic scattering in this model, before discussing the semiclassical calculation of electrical conductivity in a quantum MOS device. We consider the cases of a quasi-one-dimensional inversion layer MOSFET and a superlattice transistor as examples of this

[1] It is possible to modify the Boltzmann equation by adding a term that is nonlocal in time to include these memory effects [1].

semiclassical method. Incoherent processes such as inelastic scattering and finite temperature are seen to impose the operating limits on quantum MOS devices.

A. Scattering Mechanisms

1. Elastic Scattering

Consider a particle scattering elastically with an elastic mean free time τ. At every collision the momentum direction of the electron is altered. In the simplest approximation the electron scattering is isotropic. When it hits a scatterer, the electron has equal probability to go in any direction, as shown in Fig. 1. These scatterings tend to randomize the electron's momentum direction.[2] This is the mechanism responsible for finite electrical conductivity.

If a group of electrons are injected with some initial momentum, then their momentum direction at times greater than τ is totally random.[3] Elastic

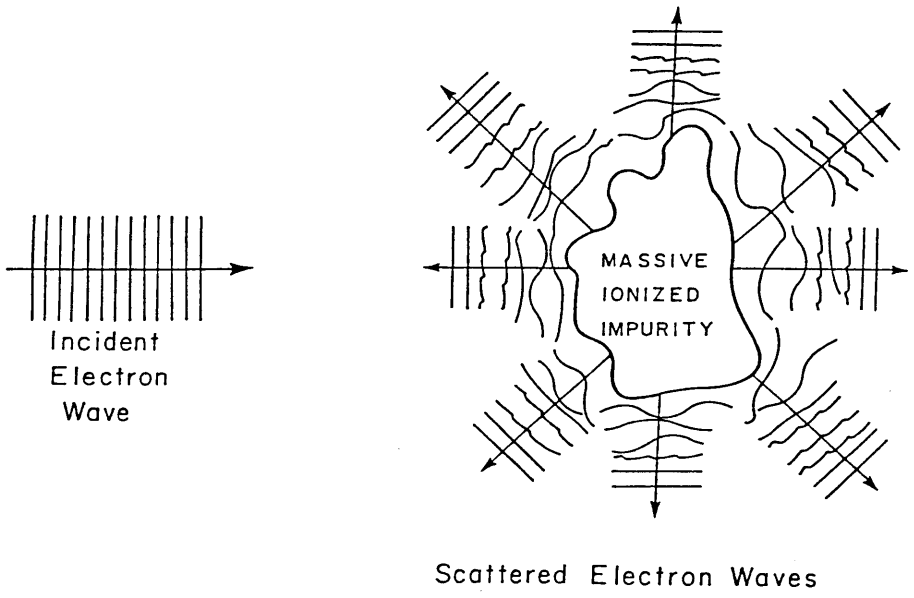


Fig. 1. Elastic scattering from a defect.

[2] For a thorough quantum mechanical description, a sum over scattering histories approach, describing a general scattering cross section see Ref. [2].

[3] Note that, for an isotropic scattering cross section, the elastic mean free time is equal to the momentum relaxation time.

scattering also requires that the magnitude of each electron's momentum before and after scattering are equal. It may seem strange that no energy or momentum is imparted to the scattering obstacle in the collision. This is, of course, an idealization that holds approximately true if the scattering obstacle is much heavier than the electron. Electron-ionized impurity collisions in an inversion layer are only approximately elastic, but we will treat them as such.

The scattering rate depends on electron energy if the density of states (i.e., the number of states available to scatter into) varies with energy. Let the quantum number k label a wave packet made up of either free electron eigenstates $e^{ik\cdot r}$ or Bloch states of a crystal and centered about momentum $\hbar k$. The electron energy E_k depends on k. The scattering rate from the state k to the state k' is then

$$\hbar/\tau_{k',k} = 2\pi n_{\text{imp}} |\langle k'|V|k\rangle|^2 N(E_k) \tag{1}$$

with $E_k = E_{k'}$. Here n_{imp} is the density of impurity scatterers and $N(E)$ the electronic density of states. Equation (1) is called the Fermi Golden Rule approximation for the scattering rate [3]. The scattering matrix element $\langle k'|V|k\rangle$ is simply the Fourier transform of the impurity scattering potential $V(r)$. Our assumption of an isotropic scattering requires that the probability for elastic scattering from one k state to another be independent of the direction of k, but still conserve energy. We further assume a white-noise power spectrum for the scattering potential, so that $\langle k'|V|k'\rangle = |V|^2 =$ constant. Other choices for the impurity power spectrum (e.g., Gaussian randomness) are possible but lead to more complicated results. Our expression for the elastic scattering time becomes

$$\frac{\hbar}{\tau(E_k)} = 2\pi n_{\text{imp}} |V|^2 N(E_k) \tag{2}$$

We now discuss inelastic scattering in order to develop a specific calculational scheme for the electrical conductivity.

2. Inelastic Scattering

Two main types of collisions that change the electron energy in silicon inversion layers are electron–electron inelastic scattering and electron–phonon inelastic scattering. Electron–electron inelastic scattering is basically the same coulombic type of scattering we studied in the Section II.A.1. However the masses of both particles are now comparable, and we can no longer neglect the recoil of the "target." The energy and momentum of both electrons change during collision, which serves to equilibrate electron temperature within the inversion layer. This type of inelastic scattering dominates in silicon inversion layers below about 10 K [4].

Electron-phonon inelastic scattering serves to equilibrate the electron gas temperature with the lattice temperature. This scattering mechanism is responsible for ohmic heating in a conductor. An applied electric field raises the temperature of the electron gas, which must then dissipate its excess energy. If the average length for an electron-phonon inelastic event is longer than the sample size, the electron gas begins to heat. This has been observed in silicon inversion layers [5]. To avoid all ohmic heating effects, the voltage drop across the sample (or one inelastic block of the sample) should be less than $k_B T/q$. It is also observed experimentally that quantum oscillations in MOSFET I-V curves disappear at high drain biases [6]. When electron heating begins to occur, the range of validity of Ohm's law in electric field also shrinks to zero, since electrical "conductivity" changes as the electron gas continues to heat during a constant applied bias.

Experimentally, the inelastic scattering rate depends on a power law of the temperature as $\tau_\phi^{-1} \propto T^P$, and becomes arbitrarily long at low temperature. For electron-phonon inelastic scattering $P=3$ since the number of phonons at low temperature is proportional to T^3. For electron-electron inelastic scattering $P=1$. Approximate values for τ_ϕ in two-dimensional MOS inversion layers at low temperatures as determined from low field magnetoresistance measurements [4] are that $\tau_{e\text{-}ph}^{-1} \simeq 10^7 T^3$ and $\tau_{e\text{-}e}^{-1} \simeq 10^8 R_\square T$. Here the times are in inverse seconds, temperature in Kelvins, and the sheet resistance in ohms per square. Reviews of weak localization are found in Refs. [7-9].

B. Broadening of Energy Levels

1. "Elastic" Broadening

If inelastic scattering is indeed necessary to dissipate the energy given to the electrons between elastic collisions, how is it possible to derive an electrical conductivity without considering inelastic scattering? The implicit assumption in all such calculations is that inelastic scattering happens much less frequently than elastic scattering, and can thus be neglected in calculating conductivity. Yet enough inelastic scattering occurs to equilibrate the electron temperature with the lattice.

To quantify this assumption, let τ be the mean elastic time and τ_ϕ be the mean inelastic time. Because the electron scatters elastically many times before traversing the sample or undergoing an inelastic scattering event, the motion of the electron will appear to be diffusive, with a diffusion constant D. Let τ_s be the time to diffuse across the sample. The average

distances an electron can diffuse in these times are

$$l = \sqrt{D\tau} \tag{3}$$

$$L_\phi = \sqrt{D\tau_\phi} \tag{4}$$

$$L = \sqrt{D\tau_s} \tag{5}$$

where l is the mean free path or elastic scattering length, L_ϕ the inelastic scattering length, and L the sample size. Therefore, the usual assumption is that $L \gg L_\phi \gg l$.

Electron phase coherence relates to inelastic scattering in the following way: The time-independent solutions to Schrödinger's equation have a fixed energy. Each energy eigenvalue corresponds to a single electron wave function extending over the entire sample. Each eigenstate m is given by

$$\psi_m(\vec{r}, t) = \psi_m(\vec{r})\, e^{-iE_m t/\hbar} \tag{6}$$

and has a definite energy E_m. If there are many elastic scattering sites, then $\psi_m(\vec{r})$ can still be found and the energy levels will still be definite. However, as soon as the electron inelastically scatters, its energy must change, forcing it to assume a different wave function, which will have a different phase. Consequently, on the average, only regions of length smaller than L_ϕ can be described by a coherent wave function that has a definite phase or energy. The sample as a whole may be pictured as broken up into many coherent blocks of length L_ϕ, each block with a single electron wave function valid throughout the block. In Fig. 2 we show the case $L > L_\phi > l$. This sample as a whole has no well-defined phase because the wave function in one coherent region is uncorrelated in phase to another neighboring coherent region. We call the sample incoherent.

In an incoherent macroscopic sample we cannot measure the properties of each individual coherent region, we can only measure average properties. It is reasonable to assert that macroscopic properties like density of states, electrical conductivity, and optical absorption are simply averages of these

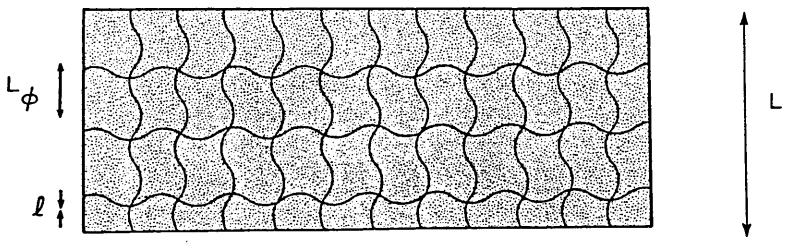


Fig. 2. An incoherent macroscopic sample.

properties over all the coherent regions. The electrons inside each coherent region of the sample solve a different Schrödinger equation in that region, since the arrangement of impurities varies from region to region. A fixed impurity configuration in each region leads to a well-defined eigenvalue spectrum (i.e., the energy levels have no width). However, any given impurity configuration will have only a small probability of occurrence when we average over the configurations. A given energy eigenvalue associated with some specific arrangement of impurities will now also only have a certain probability of occurrence.

For a random impurity potential with a white-noise power spectrum in a sample with $L > L_\phi$, the probability of an electron having some energy E away from the energy level E_m, where E_m is an energy level when there is no scattering, is given by a Lorentzian line shape [2, 10]. We denote the normalized Lorentzian probability density, shown in Fig. 3, as the impurity broadening function,

$$A(E - E_m, \hbar/(2\tau)) = \frac{1}{\pi} \frac{\hbar/2\tau}{(E - E_m)^2 + (\hbar/2\tau)^2} \tag{7}$$

which has a full width at half maximum (FWHM) of $\hbar/\tau$. The Lorentzian shape arises from an incoherent average of the solutions to Schrödinger's equation in all the coherent regions of the sample. If the sample size becomes of the order of L_ϕ, the energy eigenvalues will again become discrete. The exact shape of $A(E - E_m, \hbar/2\tau)$ depends on the assumption for the impurity power spectrum. For Gaussian randomness the impurity broadening function would not be a Lorentzian.

If the sample were to have no scattering at all it could be described by its energy levels as given by the density of states $N(E)$. The density of states with scattering, $N(E, \hbar/2\tau)$, in the sample with $L > L_\phi > l$ is then

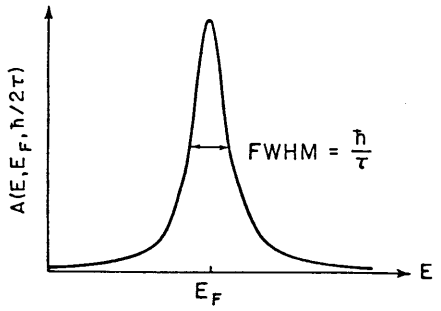


Fig. 3. Impurity broadening function $A(E, \hbar/2\tau)$.

given by finding all the states that have energy E, namely,

$$N(E, \hbar/(2\tau)) = \int_{-\infty}^{\infty} N(E')A(E-E', \hbar/(2\tau))\, dE'$$
$$= N(E) \otimes A(E, \hbar/2\tau) \tag{8}$$

Here $\otimes$ denotes a convolution. Consequently, the broadened density of states is the convolution of the unbroadened density of states with the impurity broadening function $A(E, \hbar/2\tau)$. *Although $A(E, \hbar/2\tau)$ has a width set by the elastic scattering, it is inelastic scattering that leads to the incoherent averaging process necessary to produce "elastic" broadening.*

If the inelastic length becomes of the order of the sample size, however, we must consider another mechanism for energy-level broadening: broadening due to a finite sample size. When we allow currents to flow in a finite device, each electron spends a finite amount of time in the sample. The energy eigenvalues have additional broadening from the uncertainty principle $\Delta E\, \Delta t > \hbar$, as well as from the inelastic processes, we have discussed previously. In a diffusive sample $\Delta t = \tau_s$, where τ_s is the time to diffuse across the sample given by $L = \sqrt{D\tau_s}$. In this simplistic argument we must ask, however, what defines the size of the "sample." The sample size L is the distance between two inelastic contacts (i.e., the inelastic length for a coherent sample). In this case the sample size plays the role of the inelastic length, the average distance a particle must travel before suffering an inelastic collision. Inelastic processes are again necessary to produce an energy-level broadening, entering in this context through the guise of a finite sample size. We can again represent the energy-level broadening due to a finite sample size by a broadening function of width $\hbar/\tau_s$, and convolve it, instead of the impurity broadening function, with the density of states.[4] If we make the sample size too small, energy-level broadening due to a finite time to diffuse across the sample can actually be worse than "elastic" broadening. Operating MOS devices in the regime $L < L_\phi$ also introduces sample-specific fluctuations in the conductance of size e^2/h, the so-called universal conductance fluctuations, which we discuss in Section III. If we attempt to average away the conductance fiuctuations by placing many "identical" devices in parallel, the energy-level broadening again becomes $\hbar/\tau$, since we must average over different impurity configurations inside each device. This is the same incoherent averaging over impurity positions that lead to "elastic" broadening of energy levels. The usual regime of device operation, however, is that energy-level broadening due to a finite

[4] Energy levels in an isolated finite sample (i.e., one not connected to a temperature or particle bath) also have zero width. Since we cannot contact these types of samples with wires to carry current, we are not interested in them here.

sample size is negligible compared to "elastic" broadening, which is true in a typical device where $L > L_\phi > l$. We, therefore, concentrate our discussion primarily on this regime.

2. Temperature Broadening

The occupation of electron energy levels at finite temperature is determined by the Fermi–Dirac probability distribution function:

$$f_0(E - E_F, T) = \frac{1}{1 + \exp(E - E_F)/k_B T} \tag{9}$$

The derivative of this function at finite temperature gives the probability density of an electron being some energy away from E_F

$$-\frac{\partial f_0(E - E_F, T)}{\partial E} = \frac{1}{4k_B T} \operatorname{sech}^2\left(\frac{E - E_F}{2k_B T}\right), \tag{10}$$

and is known as the thermal broadening function. It is a normalized probability density function that has a FWHM of $3.5k_B T$. We display the thermal broadening function graphically in Fig. 4.

We count the total density of particles in a quantum system at finite temperature by

$$n(E_F, T) = \int_{-\infty}^{\infty} N(E', T = 0) f_0(E' - E_F, T)\, dE' \tag{11}$$

We can express this result more usefully by differentiating with respect to E_F,

$$\frac{dn(E, T)}{dE} \equiv N(E, T) = N(E, T = 0) \otimes \left(-\frac{\partial f_0(E, T)}{\partial E}\right) \tag{12}$$

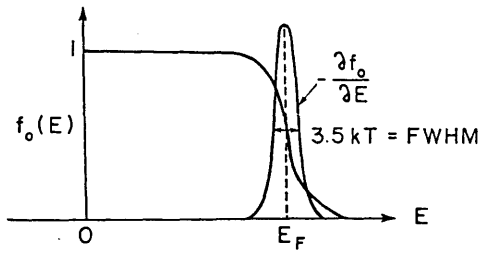


Fig. 4. Thermal broadening function $-\partial f_0/\partial E$ and the Fermi–Dirac probability distribution function $f_0(E - E_F, T)$.

The electron density of states at zero temperature is convolved with its probability density of being a certain energy away from the Fermi energy at zero temperature to obtain the effective thermodynamic density of states $N(E, T)$ at finite temperature. We can calculate the electron density by

$$n(E_F, T) = \int_{-\infty}^{E_F} N(E, T)\, dE \tag{13}$$

Temperature broadening of energy levels can be connected to distance and time scales in the following way: Consider a beam of electrons in the solid at the Fermi energy. Because it is connected to a temperature bath, the electron beam at the Fermi energy has a spread in energy $k_B T$. Because it is at a finite temperature, the electron beam is not monochromatic. Suppose the electron beam initially has all its frequency components in phase. How far can this beam of electrons travel before the probability distribution of measuring a certain phase is nearly uniform? In other words, how far can the electron beam travel before we must say it really has no well-defined phase at all?

To estimate this distance we ask how far the electron beam must travel before the phase at its center frequency is π away from the phase at a frequency approximately one standard deviation away. This is the condition that the center frequency interferes destructively with the frequency one standard deviation away, namely, $-\omega_F t = \pi - [\omega_F + (\Delta\omega/2)]t$. With $\Delta E = k_B T = \hbar\, \Delta\omega$, this corresponds to a thermal time $\tau_T = \hbar / k_B T$, and in a diffusive sample to a distance scale

$$L_T = \sqrt{D\tau_T} = \sqrt{D\frac{\hbar}{k_B T}} \tag{14}$$

called the thermal diffusion length. We cannot conduct electron interference experiments in a solid on distance scales longer than L_T because the electron has no well-defined phase beyond this distance.

The thermal diffusion length L_T is analogous to the mean inelastic length L_ϕ. This is because "temperature" is a macroscopic parameter used to describe inelastic scattering between electrons and a temperature bath. The difference in these length scales depends on which of the inelastic scattering processes connect electrons to the heat bath. For example, L_T might include only a fraction of the electron-phonon inelastic scattering, while L_ϕ might include all electron-phonon inelastic scattering and electron-electron inelastic scattering. In this case we would have $L_\phi < L_T$. If all inelastic scattering processes are between the electrons and the heat bath we have $L_\phi = L_T$. Note also that L_T can create the so-called "elastic" broadening of energy levels if L_T is much smaller than the sample size.

3. Total Energy-Level Broadening

Energy-level broadening due to "elastic" scattering is statistically independent of finite temperature broadening. The total probability density function is then a convolution of the ones for these two independent random processes. This leads to an expression for the effective density of states

$$N(E, T, \hbar/(2\tau)) = N(E, T=0) \otimes \left(-\frac{\partial f_0(E, T)}{\partial E} \right) \otimes A(E, \hbar/(2\tau)) \quad (15)$$

at finite temperatures with the electron inelastic scattering length short compared to the device size. As before, the total electron density is given by

$$n(E_F, T, \hbar/(2\tau)) = \int_{-\infty}^{E_F} N(E, T, \hbar/(2\tau))\, dE \quad (16)$$

If the mean inelastic length is long compared to the device size, the electron distribution function f may deviate from an equilibrium Fermi-Dirac distribution function f_0 inside the device. Thus, temperature is not well defined inside the device. Assuming the deviation from a Fermi-Dirac distribution is not too severe, we can model this regime of device operation by replacing the mean elastic time τ in the impurity spectral function by the time to diffuse across the sample τ_s. Recall that sample-specific fluctuations in conductance of size e^2/h appear in this regime.

C. Broadening of the Conductivity

If we neglect electron wave interference between collisions, then the Boltzmann transport equation is adequate. We consider here an infinitely long device (i.e., a free electron gas or superlattice in any dimension) where the distribution function $f(r, p, t)$ is only a function of p. We determine the current from

$$j_i = e \int v_i(k) f(k) \cdot 2 \frac{d^d k}{(2\pi)^d} \quad (17)$$

where d is the dimension of the system and e is the electron charge, which is negative. In the MOSFET inversion layer, $d = 2$. For a periodic system or a free-electron gas the electron group velocity can be calculated as

$$v_i(k) = \frac{1}{\hbar} \frac{\partial}{\partial k_i} E(k) \quad (18)$$

The semiclassical equation of motion leads to the following expression for

the conductivity tensor [11]

$$\sigma_{ij} = e^2 \int \left(\frac{1}{\hbar}\frac{\partial E_{\vec{k}}}{\partial k_i}\right)\left(\frac{1}{\hbar}\frac{\partial E_{\vec{k}}}{\partial k_j}\right)\tau(k_j)\left[-\frac{\partial f_0(E_{\vec{k}} - E_F, T)}{\partial E_{\vec{k}}}\right] 2\frac{d^d k}{(2\pi)^d} \quad (19)$$

At zero temperature

$$-\frac{\partial f_0(E_{\vec{k}}, E_F, T=0)}{\partial E_{\vec{k}}} = \delta(E_{\vec{k}} - E_F) \quad (20)$$

so that

$$\sigma_{ij}(E_F, T=0) = e^2 \int \delta(E_{\vec{k}} - E_F)\left(\frac{1}{\hbar}\frac{\partial E_{\vec{k}}}{\partial k_i}\right)\left(\frac{1}{\hbar}\frac{\partial E_{\vec{k}}}{\partial k_j}\right)\tau(E_F) 2\frac{d^d k}{(2\pi)^d} \quad (21)$$

At finite temperatures, Eq. (19) can be written as

$$\sigma_{ij}(E, T) = \sigma_{ij}(E, T=0) \otimes \left(-\frac{\partial f_0(E, T)}{\partial E}\right) \quad (22)$$

We merely convolve the zero temperature conductance with the thermal smearing function to obtain the finite temperature conductivity. This is the usual semiclassical scheme for calculating the conductivity [11, 12].

A complete quantum mechanical calculation [2] of the electrical conductivity with "elastic" broadening and with temperature broadening gives

$$\sigma(E, T, \hbar/(2\tau)) = \sigma(E, T=0) \otimes \left(-\frac{\partial f_0}{\partial E}(E, T)\right) \otimes A(E, \hbar/(2\tau)) \quad (23)$$

The temperature broadening follows directly from our standard semiclassical treatment. Just as with the density of states, we merely convolve the zero-temperature conductance with the thermal broadening function and the impurity broadening function for random impurities to obtain the full conductivity.

Now we have a full semiclassical algorithm to calculate the current versus voltage relationship in a quantum MOS device.[5] In MOS devices, terminal voltages really control the electron density, not the Fermi energy. Therefore, what we want to graph is $\sigma(E_F, T, \hbar/2\tau)$ versus $n(E_F, T, \hbar/2\tau)$ with both variables described parametrically by E_F for a given T and $\hbar/2\tau$. First, for a given temperature, elastic scattering time (i.e., device mobility), and density, we calculate the Fermi energy by demanding that the Fermi energy satisfy

$$n = n(E_F, T, \hbar/(2\tau)) = \int_{-\infty}^{E_F} N(E, T, \hbar/(2\tau))\, dE \quad (24)$$

[5] A fully self-consistent algorithm would solve the coupled Poisson and Schrödinger equations.

where

$$N(E, T, \hbar/(2\tau)) = N(E, T=0) \otimes \left(-\frac{\partial f_0}{\partial E}(E, T) \right) \otimes A(E, \hbar/(2\tau)) \qquad (25)$$

Next, we calculate $\sigma(E, T=0, \hbar/2\tau=0)$ as in Eq. (21), then broaden the conductivity with temperature and elastic scattering as

$$\sigma(E, T, \hbar/(2\tau)) = \sigma(E, T=0) \otimes \left(-\frac{\partial f_0}{\partial E}(E, T) \right) \otimes A(E, \hbar/(2\tau)) \quad (26)$$

We can now plot the conductivity versus density parametrically with Fermi energy at finite temperature and in the presence of elastic and inelastic scattering. Since electron density is proportional to gate voltage in MOS devices, this completely specifies the MOSFET current versus gate voltage relationship.[6]

D. Observability Criteria

To see if any MOS device structure will have an interesting conductivity versus density relationship, we must carry out the algorithm previously outlined. Before proceeding with this algorithm, we first find some necessary criteria that must be satisfied in order to observe structure in the conductivity from quantum mechanical effects due to the device geometry. It is reasonable to assume that if the effective density of states $N(E, T, \hbar/2\tau)$ has its structure smeared out, then the conductivity will also lose its structure. Let ΔE be the spacing between the structure in the density of states $N(E)$. For the structure to remain in $N(E, T, \hbar/2\tau)$ after convolution with the temperature broadening function we must have $\Delta E > k_B T$. For the structure to remain in the presence of elastic and inelastic scattering, there are two possibilities. One possibility is to make the sample size smaller than the inelastic length so that incoherent averaging is less severe. In this case the wave function is coherent in phase and depends on the exact type and location of elastic scattering sites. However, since the exact location of scattering sites is unknown, the density of states and the conductivity cannot be predicted. Furthermore, Lee and Stone [13] have shown that the differences in conductivity between two samples with slightly different locations of scattering sites can be very large. This phenomenon is known as universal conductance fluctuations and will be discussed in Section III. In order to have predictive power, we must have $L > L_\phi$, so that we can calculate the effective density

[6] If electron density is not proportional to a terminal voltage, we require a semiclassical solution to Poisson's equation in the device to determine this relationship.

of states, while at the same time have the elastic broadening small enough so as not to smear out any structure in the density of states. In other words, we must demand that $L > L_\phi$ and $\Delta E > \hbar/\tau$.

As an example of these two criteria, consider an ultranarrow MOSFET of width W. The spacing between the nth and $(n+1)$th energy levels due to quantization in the width direction for this system is

$$\Delta E = \frac{\hbar}{2m^*}\frac{\pi^2}{W^2}(2n+1) \tag{27}$$

where m^* is the carrier effective mass in the semiconductor. The temperature criterion becomes

$$k_B T < \frac{\hbar}{2m^*}\frac{\pi^2}{W^2}(2n+1) \tag{28}$$

and using the mobility $\mu = q\tau/m^*$, the elastic broadening criterion becomes

$$\mu > \frac{2q}{\hbar}\frac{W^2}{\pi^2}\frac{1}{(2n+1)} \tag{29}$$

These criteria are satisfied between the first and second energy levels for a 75-nm wide channel if the device mobility is greater than 6000 cm^2/V sec and the temperature is below 8 K[14]. These are typical numbers for exploratory silicon MOSFETs. We should note again that these observability criteria are only necessary conditions. Even if they are fulfilled one must still calculate the conductivity to see if it actually has any structure.

Energy scales for confining potentials on the order of 100 nm are in the range of one-tenth of a milli-electron-volt. For example, in a finite square well of width 100 nm and height 1 meV, the three bound state energy levels (as measured from the bottom of the well) are at 0.114, 0.444, and 0.909 meV. For a reasonably high mobility silicon device, 15,000 cm^2/V sec, on a (100) surface (which corresponds to an electron effective mass $m^* = 0.2m_0$) we obtain a mean free time, from $\mu = q\tau/m^*$, of 1.7×10^{-12} sec. This corresponds to an energy-level broadening, $\Delta E = \hbar/\tau$, of 0.386 meV. The temperature broadening, $\Delta E = k_B T$, at 1 K is 0.086 meV. At liquid helium temperature, 4.2 K, temperature broadening is 0.362 meV, approximately the same as elastic broadening. In contrast, a GaAs heterostructure having a mobility of 200,000 cm^2/V sec with an electron effective mass $m^* = 0.067m_0$ has a mean free time of 7.6×10^{-12} sec and an elastic broadening of 0.086 meV. This roughly corresponds to 1 K.

A good approximate relation for the total energy-level broadening (exactly valid if both the thermal smearing and elastic broadening functions are Gaussian) is that

$$(\Delta E_{tot})^2 = (\hbar/\tau)^2 + (3.5k_B T)^2 \tag{30}$$

III. QUASI-ONE-DIMENSIONAL MOSFETs

The experimental work on one-dimensional conductors began as an effort to see the effects of one-dimensional subbands in the electrical conductivity. When structure was observed in conductivity versus gate voltage (density) of quasi-one-dimensional MOSFETs, experimentalists were initially very excited about having seen the one-dimensional subbands. As more experiments were done, however, it was found that the structure varied from sample to sample even though reproducible within a given sample, that it sometimes changed if the sample was heated to room temperature and then recooled, and that the structure was much denser in gate voltage than one would expect from the number and spacings of the one-dimensional subbands. These results suggest a connection to random disorder.

A clear picture of the fluctuations has recently emerged [15]. In the low gate voltage or strongly disordered regime, the fluctuations in conductivity are very large and result from resonant tunneling (at $T=0$) or phonon-assisted hopping ($T\neq 0$) between localized levels. In the high gate voltage or weakly disordered regime, the rms conductance fluctuations have an universal size e^2/h [13], independent of the sample size (for a coherent sample) and the degree of disorder (as long as the sample remains metallic).

We begin this section by studying the regular features in conductivity one would expect in the presence of one-dimensional subbands, then return to the subject of conductance fluctuations in the regime of weak disorder. The conductance fluctuations are more pronounced in lower dimensional systems, so we study them along with the one-dimensional conductivity.

A. Conductivity of Quasi-One-Dimensional Inversion Layers

A quasi-one-dimensional conductor is an one-dimensional conductor of finite width. The crossover from two dimensions to one occurs when the width of the device becomes narrower than the inelastic diffusion length. A single coherent wave function then extends across the entire width of the device. Schrödinger's equation can be solved by the separation of variables method in this case and predicts energy eigenvalues [16]

$$\frac{2mE}{\hbar^2}=k_x^2+\left(\frac{\pi n}{W}\right)^2, \qquad n=1,2,3,\ldots \tag{31}$$

for an infinite square-well confining potential. Here we take the x direction to be along the length of the wire. The wave function has a product form

$$\psi(x,y)=\sqrt{\frac{2}{W}}\sin\left(\frac{\pi n}{W}y\right)\exp(ik_x x) \tag{32}$$

Consider the two constant energy surfaces shown in Fig. 5. If the Fermi level lies in a higher subband (the third subband in Fig. 5), it becomes possible to scatter between different transverse eigenstates on the constant energy surface. Even though we discuss an infinite square well here for simplicity, our conclusions about conduction in quasi-one-dimensional wires are independent of the shape of the confining potential. It is desirable, of course, to have the confinement as strong as possible so that the subbands are widely spaced in energy.

To understand the motivation for making quasi-one-dimensional conductors, we investigate what the semiclassical conductivity formula would predict for such structures. Consider first the case of an one-dimensional conductor without subbands (i.e., only one spatial dimension exists, so that $E = \hbar^2 k_x^2/2m$). The group velocity v_g, density of states $N(E)$, density, and conductivity at a given Fermi energy are, respectively,

$$v_g = \frac{1}{\hbar}\frac{dE}{dk} = \sqrt{\frac{2E}{m}} \tag{33}$$

$$N(E) = \frac{1}{\pi}\sqrt{\frac{2m}{\hbar^2 E}} \tag{34}$$

$$n(E) = \frac{2}{\pi}\sqrt{\frac{2mE}{\hbar^2}} \tag{35}$$

$$\sigma(E_F, T=0) = e^2(v_F^2 \tau_F) N(E_F) \tag{36}$$

Equation (36) follows from direct integration of Eq. (21). We show the results for the approximation when $\tau =$ constant in Fig. 6. The conductivity rises as the square root of Fermi energy and is linear in the electron density.

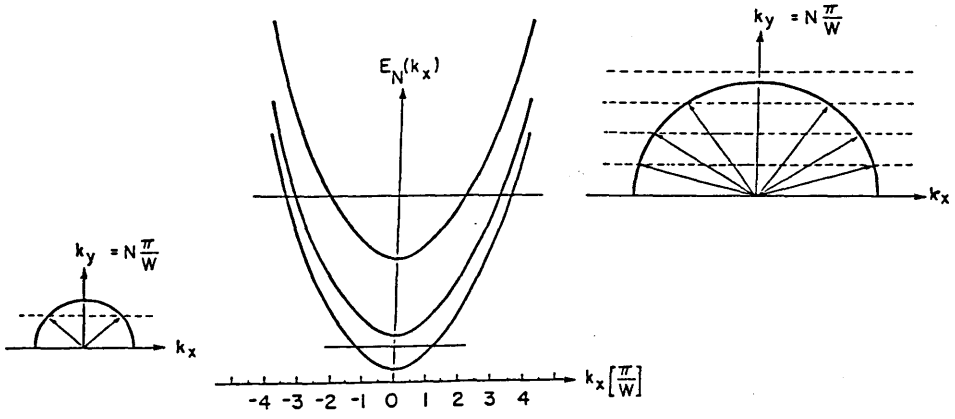


Fig. 5. Energy dispersion and constant energy surfaces for a quasi-one-dimensional conductor.

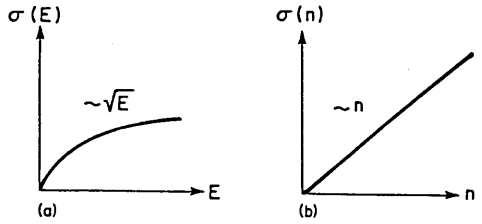


Fig. 6. One-dimensional conductivity for $\tau =$ constant as a function of (a) energy and (b) density.

Using the same equations, if we choose an energy-dependent scattering rate, as discussed in Section II.A; namely,

$$\frac{\hbar}{\tau(E)} = 2\pi n_{\text{imp}} |V|^2 N(E) \tag{37}$$

we obtain a conductivity linear in the Fermi energy and parabolic in electron density, as in Fig. 7.

These relations assume that only one physical dimension exists. To make quasi-one-dimensional conductors, we must begin in three dimensions and make two of the dimensions smaller than the inelastic length, thus creating one-dimensional subbands [17]. Each subband is an additional channel to carry current; their conductivities add like resistors in parallel if we ignore wave interference between different subbands. Then

$$\sigma(E) = \sum_i \sigma_i(E) = e^2 \sum_i v_{F,i}^2 N_i(E) \tau_i(E) \tag{38}$$

where we label the sum over all *occupied* subbands with the index i. The density of states adds similarly,

$$N(E) = \sum_i N_i(E) \tag{39}$$

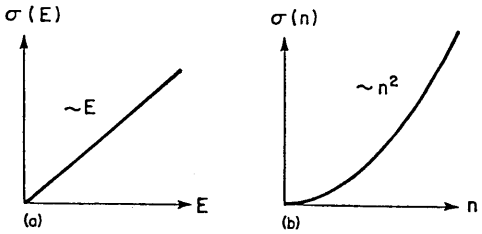


Fig. 7. One-dimensional conductivity for $\tau^{-1}(E) \approx N(E)$ as a function of (a) energy and (b) density.

We consider the extreme quantum limit of widely spaced subbands at $T = 0$ with negligible "elastic" broadening in order to see clearly the effect of the one-dimensional subbands. We take the density of states for the case of two such one-dimensional subbands as shown in Fig. 8.

There are three possibilities for the scattering time in a quasi-one-dimensional conductor. One case is when $\tau =$ constant. The results for $\sigma(E)$ and $\sigma(n)$ are shown in Fig. 9. The constant τ case is interesting because,

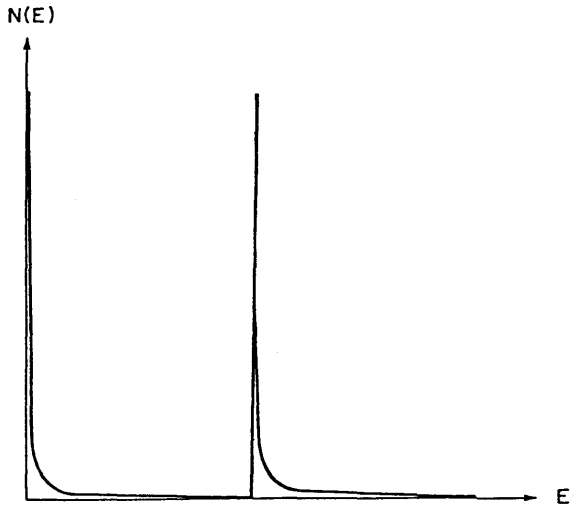


Fig. 8. Quasi-one-dimensional density of states for two subbands.

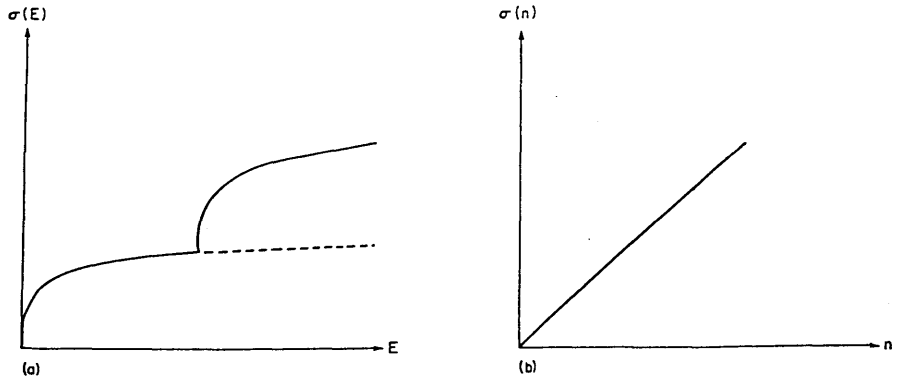


Fig. 9. Quasi-one-dimensional conductivity for constant τ as a function of (a) energy and (b) density.

even though conductivity versus energy shows structure, the physically important plot, conductivity versus density, shows none.

The second case we consider allows only intrasubband scattering, with no intersubband scattering. Mathematically this means

$$\frac{\hbar}{\tau_i(E)} = 2\pi n_{\text{imp}} |V|^2 N_i(E) \tag{40}$$

The results are shown in Fig. 10. The sharply decreasing derivative of $\sigma(n)$ when the second subband begins to be populated is because almost all the carriers are being added to the lower velocity subband, due to its very large density of states.

There is no *a priori* reason why electrons cannot scatter from one subband to another. This is the third case to consider, and the one we believe actually corresponds to experiment. Allowing both inter- and intrasubband scattering, our model for the mean free time is

$$\frac{\hbar}{\tau(E)} = 2\pi n_{\text{imp}} |V|^2 \sum_j N_j(E) \tag{41}$$

We present the results graphically in Fig. 11. The dramatic increase in scattering rate as the second subband becomes populated overwhelms the modest increase in the additional number of low-velocity carriers, and makes the conductivity drop to zero.

If there is "elastic" or finite temperature broadening, the conductivity will not drop to zero as the Fermi level crosses into a new subband, since the scattering rate remains finite. We expect the resulting $\sigma(n)$ to look

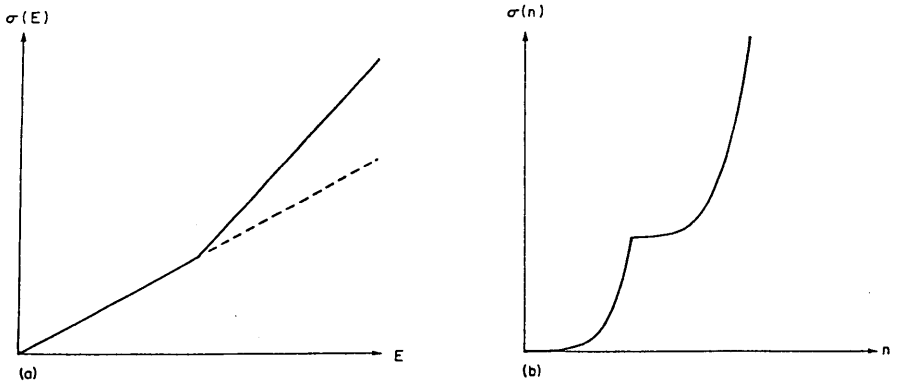


Fig. 10. Quasi-one-dimensional conductivity allowing only intrasubband scattering as a function of (a) energy and (b) density.

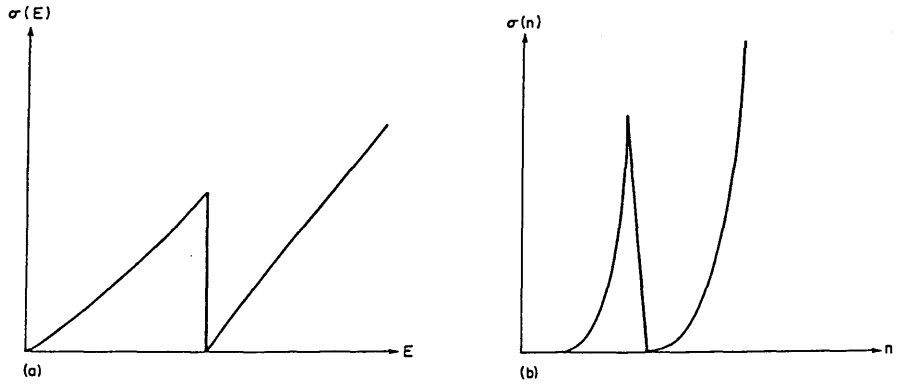


Fig. 11. Quasi-one-dimensional conductivity with both intra- and inter-subband scattering as a function of (a) energy and (b) density.

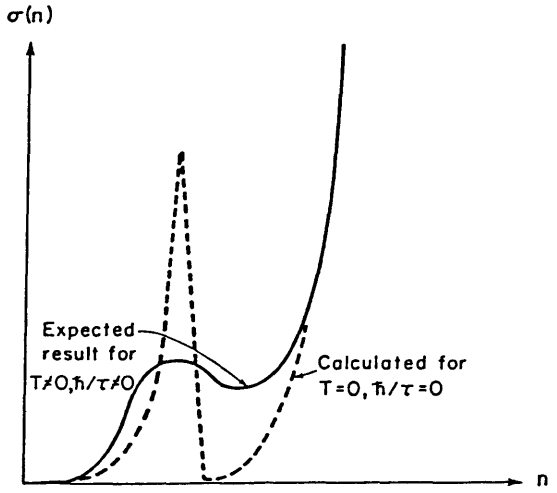


Fig. 12. Broadened quasi-one-dimensional conductivity with both intra- and intersubband scattering.

something like Fig. 12. This result qualitatively agrees with the Green function calculation of DasSarma and Xie [18]. We show their results in Fig. 13.

There has been one experiment [19] that has claimed to observe the effects of one-dimensional subbands in the electrical conductivity. This experiment used an array of hundreds of quasi-one-dimensional MOS

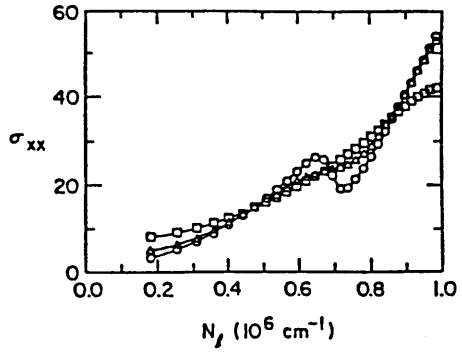


Fig. 13. Broadened quasi-one-dimensional conductivity with intra- and intersubband scattering, assuming $n_{imp} = 10^{11}$ cm^{-2}, at $T = 1.2$ K (○), 4.2 K (△), and 10 K (□). After DasSarma and Xie [18].)

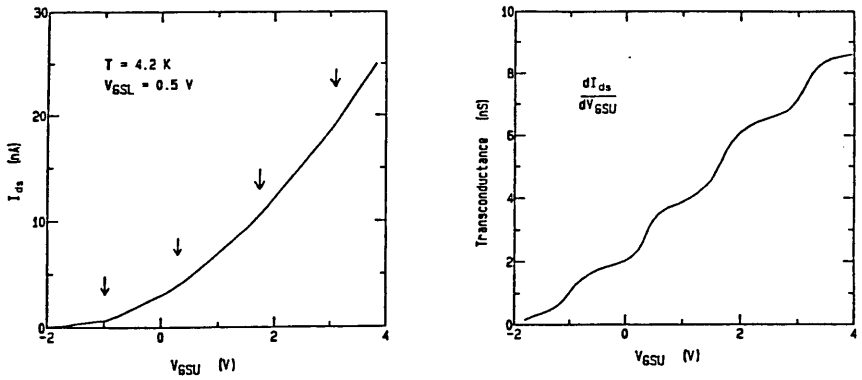


Fig. 14. Measured quasi-one-dimensional conductivity of MOS wire array. I_{ds} vs. V_{GSU} is equivalent to σ_{xx} vs. n. (After Warren [19].)

inversion lines in parallel, rather than a single wire. We discuss the advantages and disadvantages of using arrays of wires after examining universal conductance fluctuations in the next section. The measured results of Warren *et al.* [19] are shown in Fig. 14. We also show quantitatively how closely such a periodic array of wires should approximate a quasi-one-dimensional conductor in our discussion of the two-dimensional superlattice transistor.

An important test to determine whether or not any structure seen in the conductivity of quasi-one-dimensional MOS devices is due to one-dimensional subbands is the magnetoresistance. Assuming the effect of the magnetic field is only to shift the subband energies, the dips in conductance should follow the so-called fan diagrams of Ref. [20].

B. Universal Conductance Fluctuations

Rapid progress has been made in the study of electronic conduction in disordered metals. A good review of this subject and its relevance to silicon MOSFETs is the work of Kastner *et al.* [15]. For a general review, see Ref. [21]. One striking result of this work is the prediction that all one-dimensional conductors become insulators at $T=0$. A new length scale, the localization length, emerges in this theory. For a coherent sample in the "localized" regime, for which $k_F l < 1$, the fluctuations in conductance with density or gate voltage can be large. We do not discuss the localization length or the physics of the "localized" regime, since the useful regime of MOSFET operation is the high density or "extended" regime. In a two-dimensional MOSFET with a 100-Å mean free path, the device enters the "extended" regime, for which $k_F l > 1$,[7] at electron densities of around 1×10^{11} cm^{-2}. Gate voltage breakdown limits the maximum electron density to about 2×10^{13} cm^{-2}.

For a coherent sample in the "extended" regime, the normal MOSFET operating regime, the rms value of conductance fluctuations were predicted to have the universal value [13],

$$\delta G_{\text{rms}} = \frac{e^2}{h} \simeq \frac{1}{25813\ \Omega} \simeq 38.7\ \mu\text{S} \tag{42}$$

Many experiments helped lead to [22] and later confirm [23, 24] this prediction. Intuitive accounts of why the universal conductance fluctuations are "universal" and have size e^2/h are found in Ref. [25]. The basic mechanism of the universal conductance fluctuations is that, as the sample becomes coherent, we can no longer neglect the wave interference of electrons scattering from imperfections. The constructive and destructive interference of the many electronic scattering paths depends strongly on the exact number and location of scattering centers and on the Fermi energy. It is the relative changes in the interference of electron scattering paths that give rise to universal fluctuations in conductivity. Universal conductance fluctuations are, thus, a manifestation of true "quantum transport" processes, like weak electron localization, and unlike all the device examples we calculated using the semiclassical conductivity formula.

The fluctuations are called "universal" in that they are independent of the sample size (as long as the sample remains coherent), degree of disorder (as long as the sample remains metallic), and only weakly dependent on dimensionality. They are perfectly reproducible with gate voltage and are, thus, not "noise" [17]. We show an example of these fluctuations and their

[7] $k_F l > 1$ is also a necessary condition for validity of the Boltzmann transport approximation.

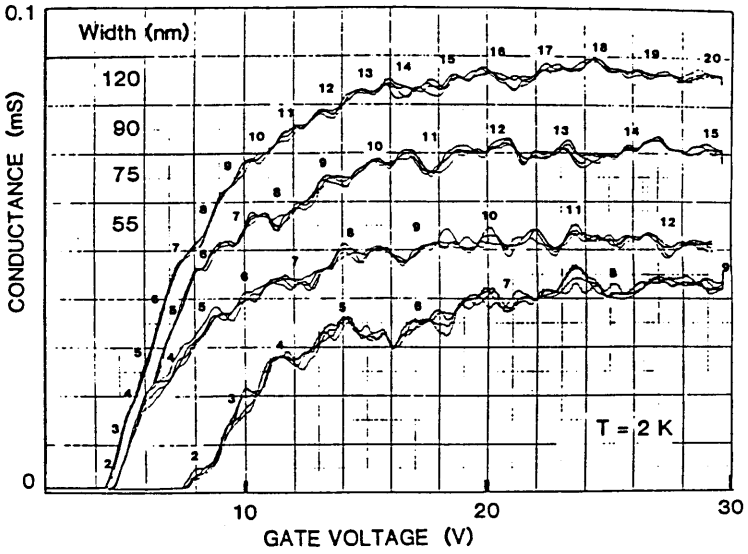


Fig. 15. Universal conductance fluctuations in a quasi-one-dimensional MOS wire. Zero field structure in four devices of different widths on the same chip. The small numbers indicate the pattern of the structure expected from simple "particle in the box" quantization, based on the measured device widths. (From Skocpol *et al.* [17].)

reproducibility in Fig. 15, after Skocpol *et al.* [17]. Further indication that the universal conductance fluctuations are a true manifestation of "quantum transport" has come from the experimental observation of voltage fluctuations outside the classical current path in MOS inversion layers [26]. We expect nonlocal effects in wires with multiple voltage probes spaced closer together than an inelastic length, since several of the probes are coherent.

If the device is incoherent, we know that eventually all the quantum fluctuations in conductance will average to zero. Consider a sample N

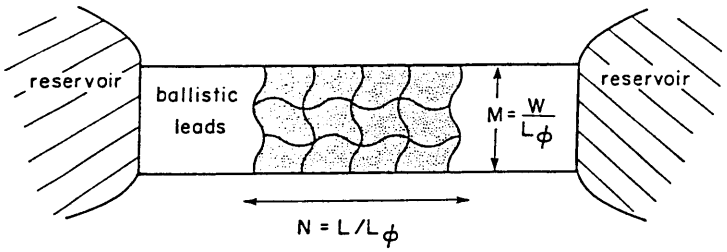


Fig. 16. Incoherent averaging of universal conductance fluctuations.

inelastic blocks long and M inelastic blocks wide connected to perfectly ballistic leads and nonreflecting but phase randomizing reservoirs, as in Fig. 16.

The relative fluctuation $\delta G/G$, where δG is the rms fluctuation in the conductance and G the average conductance, will average away like the square root of the number of coherent regions in the device [24], as one would expect from standard statistical averaging;

$$\frac{\delta G}{G} = \frac{1}{\sqrt{MN}} \frac{\delta G_\phi}{G_\phi} \tag{43}$$

The relationship between the total conductance G and the conductance of an inelastic block G_ϕ is

$$G = (M/N) G_\phi \tag{44}$$

so that

$$\delta G = \frac{1}{\sqrt{MN}} \frac{M}{N} \delta G_\phi \tag{45}$$

or

$$\delta G = \left(\frac{W}{L_\phi}\right)^{1/2} \left(\frac{L_\phi}{L}\right)^{3/2} \frac{e^2}{h} \tag{46}$$

If the sample is quasi-one-dimensional (i.e., $W < L_\phi$), then

$$\delta G = \left(\frac{L_\phi}{L}\right)^{3/2} \frac{e^2}{h} \tag{47}$$

In addition to the average size of the universal conductance fluctuations, it is also possible to predict their average spacing in Fermi energy or magnetic field [13].

C. Single Wires versus Wire Arrays

In considering whether the subbands of a quasi-one-dimensional wire can be more easily seen in the conductivity of single wires or arrays of wires in parallel, we must examine several cases. We first re-examine our observability criterion for a single wire and then consider wire arrays.

The simplest case for a single wire is when the wire becomes long compared with the mean inelastic length. In that case, conductance fluctuations are negligible, but the subband spacing must be larger than $\hbar/\tau$. This constrains the wire width to some maximum value. If, in addition, the wire

has some statistical variation in its width, the energy levels in each coherent region of the wire will be slightly different due to their different widths. Since the spacings of energy eigenvalues go as $E \approx W^{-2}$, this introduces a spread in energy eigenvalues,

$$-\Delta E = E\frac{\Delta W}{W} \tag{48}$$

Recalling our statistical averaging ideas of Section II, we can model the effects of random variations in the wire width by convolving the density of states and the conductivity with a new stochastic broadening function. The exact shape of the broadening function depends on the statistics of the random width variations along the wire, but we can estimate that its variance will be about ΔE if the fractional variance of wire width is $\Delta W/W$. Our observability criterion suggests that the random spread of energy eigenvalues along the wire must be smaller than the intersubband spacings to observe one-dimensional subbands in the conductivity. Notice that the broadening becomes worse at higher energies.

If a single wire has a length less than or comparable to the inelastic diffusion length, we reduce "elastic" broadening from $\hbar/\tau$ to $\hbar/\tau_s$ and can ignore the energy-level broadening from statistical variations in the wire width. The question in this regime is whether or not we can impose a regular modulation in the conductance whose size is much larger than e^2/h, in order that the regular features dominate over conductance fluctuations. Due to the extremely large conductance modulation we expect from intersubband scattering, we might safely assume, in the absence of numerical simulations, to observe one-dimensional subbands in this regime as well.

We turn now to the possibilities for observing one-dimensional subbands in wire arrays. In wire arrays, we have two additional possibilities: the spacing between wires in the array is either larger than the mean inelastic length or it is smaller. We also have our two previous possibilities: that the length of wires in the array is either longer or shorter than the inelastic diffusion length. This makes a total of four cases to examine.

We consider first the case where the length of wires in the array is much longer than the inelastic length and the spacing between wires is greater than an inelastic length (i.e., each wire is independent). The same broadening mechanisms that affect single wires also limit the wire array; however, they are no worse in the array. It is important in this case that all wires in the array have a random variation in their nominal widths that is no worse than for a single wire. The wire array may also offer a practical advantage of larger currents, hence a larger signal-to-noise ratio. It is not difficult, however, to measure the conductance of single wires. We conclude in this limit that neither the single wire nor the wire array shows any advantage.

If the length of the wires becomes shorter than the inelastic length but the spacing between wires remains larger than an inelastic length, the relative conductance fluctuations average away like the square root of the number of wires. Energy-level broadening in the wire array, however, is set by the elastic mean free time $\hbar/\tau$, while the broadening in single wires is set by the time to diffuse across the sample $\hbar/\tau_s$. This is because each wire has a slightly different arrangement of impurities, so that peaks and valleys in the average conductance occur at slightly different values of the electron density. This is really the same procedure as for averaging over impurities in an incoherent macroscopic sample we discussed in Section II. We conclude that there is no advantage in making each wire in the array shorter than an inelastic diffusion length. The wire array also has the possible disadvantage of introducing additional energy broadening due to random width variations between wires.

The possibility of making several quasi-one-dimensional channels closer together than an inelastic length is an intriguing one, but more complicated to analyze. To begin with, it is not clear that the mean inelastic length between wires will be the same as the one along the wire. Electrostatic control may offer one possibility to vary this length [18]. Several wires may be coherent at once, and the energy eigenvalues are determined by the properties of these several wires.

If the length of the wires is long compared to the inelastic diffusion length, having several wires coherent with each other will lead to a larger modulation in conductance, due to additional scattering between wires. This assumes the transport lifetime is still governed by the Fermi Golden Rule approximation. We consider the extreme case, where all wires in the array are coherent, in our discussion of the two-dimensional superlattice transistor. If the confining potential of the wire is large compared to the Fermi energy, a coherent array of wires can quite closely resemble a quasi-one-dimensional conductor, but can have a much larger modulation in its scattering time due to much greater variations in the electron density of states. This is analogous to the intersubband scattering mechanism, which lead to the large regular conductance variation we saw earlier. If the one-dimensional channels are incoherent, we can do no worse than add their separate conductances in parallel. If instead we allow some scattering between the channels, we suppress the conductance even more strongly when the Fermi level enters a new subband. The wire array may show an advantage in this regime, provided the energy-level broadening from random width variations is no worse in the wire array than for a single wire.

If the wire length is shorter than the mean inelastic length and the spacing between wires is smaller than the inelastic diffusion length, the conductance fluctuations average away as the square root of the number of coherent

regions of the sample, not as the square root of the number of wires. The scattering between quasi-one-dimensional channels will again lead to a larger regular modulation in conductance than the single wire case. The wire array, however, again has the disadvantage of a larger energy-level broadening, $\hbar/\tau$ for the array versus $\hbar/\tau_s$ for the single wire. If the entire sample is coherent, the wire array and the single wire both have the same energy-level broadening $\hbar/\tau_s$. The wire array does not reduce universal conductance fluctuations, but may show larger regular variations in its conductance due to scattering between wires.

In summary, for cases where the spacing between wires in a wire array is long compared to the inelastic diffusion length between wires, the wire arrays and long single wires offer equal possibility for observing one-dimensional subbands in the conductance. When the length of a single wire becomes short compared to the inelastic diffusion length, we can impose a larger regular modulation in the conductance, but conductance fluctuations of size e^2/h begin to appear. Having many wires in an array coherent with each other improves the chance of observing one-dimensional subbands. In terms of fabrication, very narrow single wires are easily fabricated using shadowing techniques [15]. Making large-area, uniform, wire arrays can be difficult. New fabrication techniques for wire arrays have promise to eliminate this problem [27], yielding almost atomically smooth and very uniform wire arrays.

IV. SURFACE SUPERLATTICE TRANSISTORS

The idea of an electrically controllable superlattice transistor dates from the late 1970s [28]. After the invention of molecular-beam epitaxy and the tiny vertical dimensions it made possible, experimentalists began to explore fixed potential superlattices [29]. Extremely high-quality GaAs/GaAlAs heterojunctions combined with vertical spacings on the angstrom scale spurred these investigations, leading as well to the development of a resonant tunneling diode and the proposed resonant tunneling transistor. For a review of these topics see Ref. [30]. Device research on the surface superlattice transistor, which depends on fine lateral patterning, had to wait for the development of new lithography techniques.

In sharp contrast to systems emulating static crystal properties, such as superlattices grown using molecular-beam epitaxy, a surface superlattice transistor would function as a true voltage controlled crystal. In the ideal case, device electrodes would allow independent control of the average electron density as well as the density modulation.

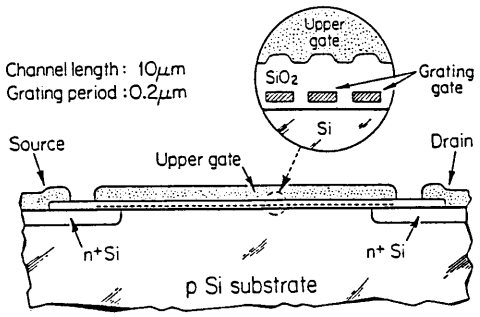


Fig. 17. A silicon surface superlattice MOS transistor.

Figure 17 shows the cross section of a silicon surface superlattice device that meets this criterion. The grating gate (bottom gate) electrode controls electron density in the channel directly beneath the grating. The continuous gate (top gate) electrode and the grating gate compete for control of the electron density between grating wires via bottom gate fringing fields. From a practical viewpoint, it is extremely important in fabricating this device to make both insulators as thin as possible and, if possible, to use a high dielectric constant material for the insulators. This ensures that the top gate will have adequate control over the device operation.

To understand the operation of this device we first consider the simplest case, a periodic potential in one dimension. We study both the energy dispersion and conductivity in one dimension to gain insight into the problem. We then derive analytic expressions for the energy eigenvalues, density of states, and conductivity in a two-dimensional superlattice device, given those in an one-dimensional superlattice.

A. Energy Momentum Dispersion in a Surface Superlattice

1. One Dimension

An electron in a periodic potential $V(x) = V(x+a)$ can form standing waves when an integer multiple of its wavelength is twice the superlattice period

$$n\lambda = 2a, \qquad n = 1, 2, 3, \ldots \tag{49}$$

A factor of two enters the Bragg reflection condition because the electron density, not the probability amplitude, must have the superlattice periodicity.

We must solve Schrödinger's equation

$$\left[-\frac{\hbar^2}{2m^*}\frac{d^2}{dx^2}+V(x)\right]\psi(x)=E\psi(x) \tag{50}$$

for the energy eigenvalues of the problem. Using the Fourier transform and the condition that $V(x)$ is a periodic function, Eq. (50) can be transformed into a matrix equation, the so-called central equation [12], for the wave function and energy eigenvalues.

We can label all the eigenvalues of the central equation by a band index N and a crystal momentum k [12]. The label k in the eigenvalues $E_N(k)$ is no longer the electron's momentum. The quantum number k is restricted to the range $|k|<(\pi/a)$. The value $k=\pm(\pi/a)$ defines the Brillouin zone [12] boundary.

The group velocity of an electron wave packet, the quantity of interest to us, is still given by Eq. (18) [12].

2. Two Dimensions

For a two-dimensional surface superlattice, we must solve Schrödinger's equation written as

$$\left[-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial x^2}-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial y^2}+V(x)\right]\psi(x,y)=E\psi(x,y) \tag{51}$$

This has solution

$$\psi_{\bar{k},N}(x,y)=e^{ik_y y}\psi_{k_x,N}(x) \tag{52}$$

where $\psi_{k_x,N}(x)$ is a solution to the one-dimensional Schrödinger equation in a periodic potential $V(x)$ as given in the previous section. The energy eigenvalues in two dimensions are

$$E_N(k_x,k_y)=E_{N,x}(k_x)+\frac{\hbar^2k_y^2}{2m^*}=E_{N,x}(k_x)+E_y(k_y) \tag{53}$$

where $E_{N,x}(k_x)$ are the eigenvalues of the one-dimensional superlattice problem. The group velocity of an electron wave packet is given again by Eq. (18) [12].

We show an example of energy dispersion for a two-dimensional superlattice in Fig. 18 [31]. Surfaces of constant energy, shown in Fig. 19, are circles if $V(x)=0$, deforming to ellipsoidal shapes in a periodic potential along the x direction [31]. The ellipsoids open at the Brillouin zone boundary $k_x=\pm\pi/a$, or at the zone center $k=0$, when electrons form standing waves.

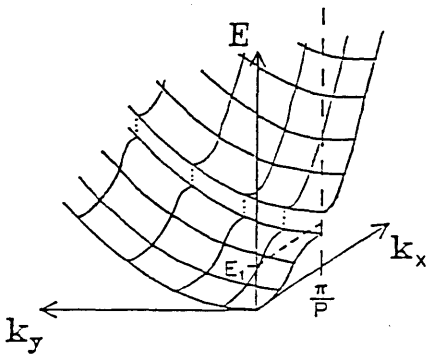


Fig. 18. Energy dispersion in a two-dimensional superlattice. (After Warren [31].)

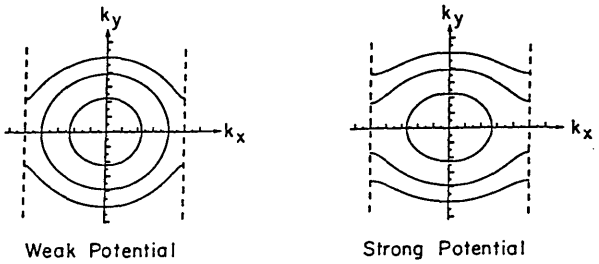


Fig. 19. Constant energy surfaces in a two-dimensional superlattice. (After Warren [31].)

To find the total number of electrons below the Fermi energy in this two-dimensional $\vec{k}$ space, we must integrate the density of electron states up to E_F:

$$n_{2\text{-D}}(E_F) = \int_{-\infty}^{E_F} N_{2\text{-D}}(E)\, dE \tag{54}$$

For potentials of the form $V(x, y) = V_x(x) + V_y(y)$, the x and y motions of electrons are independent except for the constraint that the total energy of motion is a constant. We can then express the two-dimensional density of states as a convolution,

$$N_{2\text{-D}}(E) = \tfrac{1}{2} N_x(E) \otimes N_y(E) \tag{55}$$

where we calculate $N_x(E)$ and $N_y(E)$ from solutions to the two one-dimensional Schrödinger equations with potentials $V(x)$ and $V(y)$, respectively.

For a two-dimensional surface superlattice device, $N_x(E)$ is the density of states we calculate numerically from the solution to the one-dimensional

central equation, and $N_y(E_y)$ is the free-electron density of states for motion perpendicular to the superlattice given in Eq. (34). Equation (55) is a very intuitive one. Electron motion in the x direction is independent of motion along y, except for the constraint that the total energy $E = E_{N,x}(k_x) + E_y(k_y)$ is a constant. Thus, for some density of available states along x with a contribution to the total energy $E_{N,x}$, the total number of available states at a fixed energy $E = E_{N,x} + E_y$ is simply the product $N_x(E_x)N_y(E - E_x)$. To find the total number of available states at energy E, we must sum over all the states E_x obeying the energy conservation condition.

$$N_{2\text{-D}}(E) = \frac{1}{2}\int_{-\infty}^{\infty} N_x(E_x)N_y(E - E_x)\,dE_x \tag{56}$$

The factor $\frac{1}{2}$ in this convolution compensates for overcounting electron spin states in the one-dimensional state densities. Any gaps in the energy spectrum in the one-dimensional superlattice are removed by convolution with $N_y(E)$. There are no true energy gaps in a two-dimensional superlattice transistor.

To determine the energy scales and electron densities in a two-dimensional silicon superlattice device, we consider first the two-dimensional free-electron gas with $m^* = 0.2m_0$ and an imposed periodicity $a = 2000$ Å. In two dimensions for a free-electron gas of density n, the Fermi wave vector and Fermi energy are, respectively,

$$k_F = (2\pi n)^{1/2}, \qquad E_F = \frac{\hbar^2 k_F^2}{2m^*} = \frac{\pi\hbar^2}{m^*}\,n \tag{57}$$

Typical numbers are given in Table I.

If we apply a periodic modulation to the free-electron gas, the values of E_F for a given k_F (and, thus, a given density) will change, but the relation between k_F and the electron density will be the same. Thus, the band number N at a given density is $k_F/(\pi/a)$, and the density wavelength at a given density, $\lambda_D = \lambda_F/2$, remain unaffected by the periodic potential. In a periodic

TABLE I
Parameters of a Two-Dimensional Silicon Superlattice Device

$n(\text{cm}^{-2})$	10^{10}	10^{11}	10^{12}
$E_F(\text{meV})$	0.1	1	10
$k_F(\text{cm}^{-1})$	2.5×10^5	8.0×10^5	2.5×10^5
$\pi/a(\text{cm}^{-1})$	1.6×10^5	1.6×10^5	1.6×10^5
$k_F/(\pi/a)$	2	5	15
$\lambda_D = \frac{1}{2}\lambda_F = \pi/k_F$ (Å)	1200	400	120

potential, the electron density n refers to the average electron density in the channel. These numbers suggest, for example, that for an electron density of 10^{11} cm^{-2}, we must have an appreciable energy gap between the fifth and sixth energy bands to see structure in the conductivity. The fifth harmonic of the electron density wavelength must make a standing wave with the superlattice. Contrary to our intuition from perturbation theory, this is not difficult to do if the periodic potential is large, even for a potential having only the lowest spatial harmonic (i.e., a cosine potential).

The Fermi energy can have a modulation on the order of 10 meV in a superlattice. But the Fermi energy (i.e., the electrochemical potential) is the sum of the electrostatic potential and the chemical potential. The electrostatic potential $V(x)$ is the potential in Schrödinger's equation. Simulations [32] of this superlattice configuration have shown that modulation of the electrostatic potential can be of the order of 100 meV.

The one-dimensional energy eigenvalues for 0.1 and 10 meV peak-to-peak modulation of the electrostatic potential at the fundamental wavelength of 2000 Å are shown graphically[8] in Figs. 20 and 21. These solutions are enough to qualitatively understand the superlattice transistor. Note that almost all

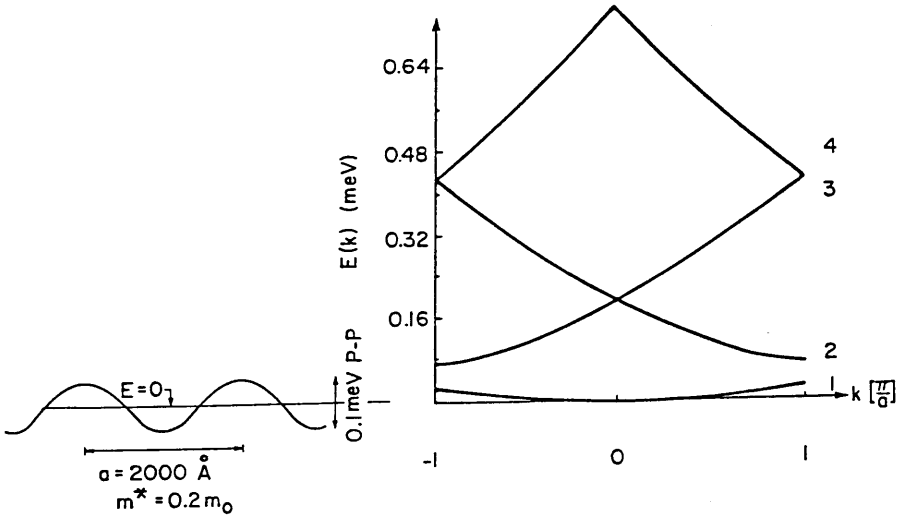


Fig. 20. Energy eigenvalues for $V(x) = \frac{1}{2} V_{p\text{-}p} \cos(2\pi x/a)$ where $V_{p\text{-}p} = 0.1$ meV. $V(x)$ is shown to the left of the dispersion relationship and the band index N is shown to the right.

[8] All numerical calculations of energy dispersion relations diagonalize a 32×32 Hamiltonian matrix with 64 samples of the potential $V(x)$. They assume $m^* = 0.2m_0$ and $a = 2000$ Å. Electrical conductivities were calculated numerically using these solutions.

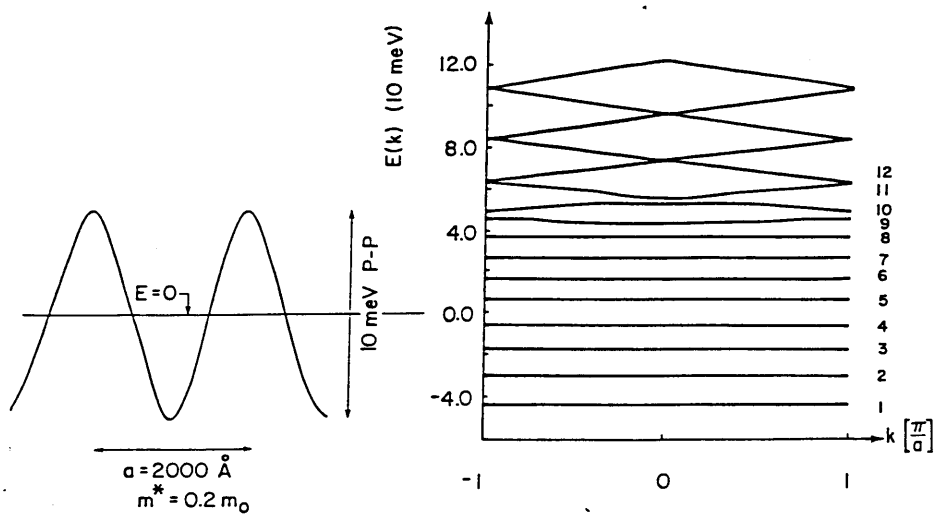


Fig. 21. Energy eigenvalues for $V(x) = \frac{1}{2} V_{p\text{-}p} \cos(2\pi x/a)$ where $V_{p\text{-}p} = 10$ meV.

the energy gaps lie below the top of each cosine potential.[9] This is not the case if $V(x)$ has higher spatial harmonic components.

B. Conductivity of a Surface Superlattice

We begin our discussion of conductivity in the surface superlattice in the simplest case of one spatial dimension. We assume one-dimensional subbands are either not present or lie too high in energy to enter into the problem. After analyzing one dimension, we obtain analytic expressions for the conductivity of a two-dimensional surface superlattice transistor, given the conductivity of an one-dimensional superlattice. These analytic expressions permit us to qualitatively explain how features in conductivity will change in the two-dimensional surface superlattice.

1. One Dimension

Take first the case $\tau =$ constant. The conductivity versus density and versus Fermi energy relationships for the two cosine potentials are shown in Figs.

[9] There is an easy way to estimate the number and average spacings of the bound energy levels in a cosine potential. Match the cosine to a harmonic oscillator potential at the bottom of each cosine well. Assume the last bound level is when the harmonic oscillator eigenvalue has a larger energy than the top of the cosine. The average spacing in energy is simply the strength of the cosine divided by the number of bound levels. This estimation method agrees very well with the exact solutions.

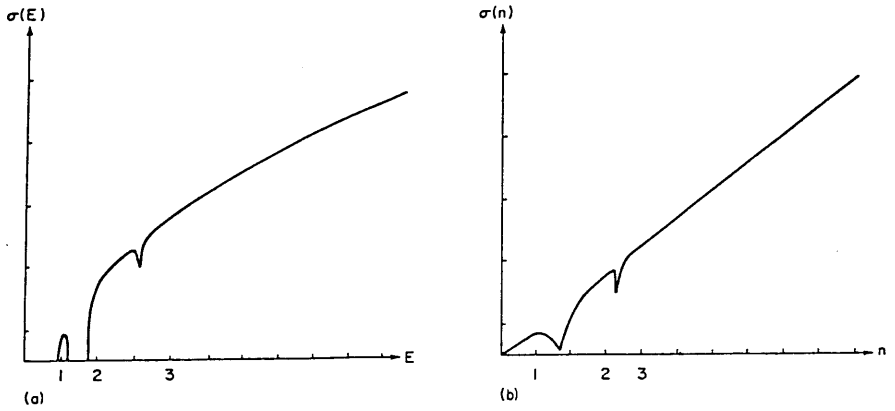


Fig. 22. Conductivity versus (a) Fermi energy and (b) density for $V_{p\text{-}p} = 0.1$ meV and constant τ. Numbers along the x-axis denote the energy band associated with a feature in the conductance.

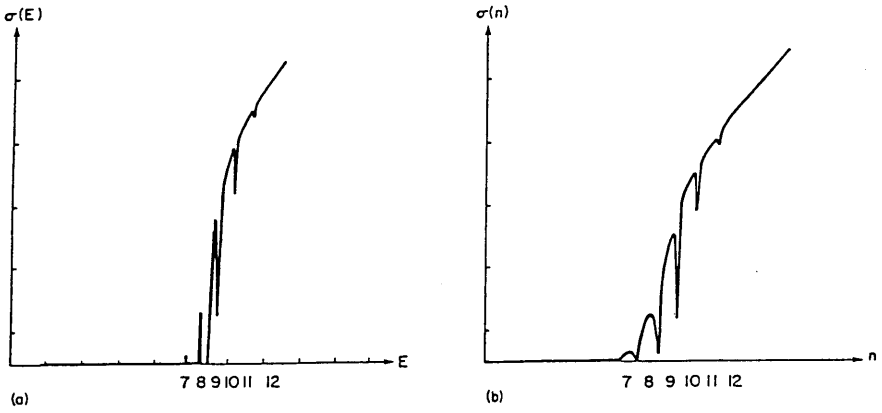


Fig. 23. Conductivity versus (a) Fermi energy and (b) density for $V_{p\text{-}p} = 10$ meV and constant τ. Band numbers are shown along the x-axis.

22 and 23. The conductance modulation is solely a result of the Fermi group velocity tending to zero at a Brillouin zone boundary ($k = \pm\pi/a$) and at the zone center ($k = 0$). Notice how rapidly the conductance drops to zero in the strong cosine potential of Fig. 23, compared to the free-electron case, once the Fermi energy falls into the cosine potential. This is because the Fermi group velocity for energy levels bound deeply in the cosine well is nearly zero. These strongly bound energy levels can be important in optical experiments, but not in transport. The envelope of conductivity versus

density is linear in n for large densities, while the envelope of conductivity versus Fermi energy goes as the square root of energy, as we showed in the discussion of one-dimensional conductors.

We also examine the case of $\tau^{-1} \sim N(E)$. Conductivity versus density and versus Fermi energy relationships for the cosine potential 0.1 and 10 meV are shown in Fig. 24 and 25, respectively. Conductivity is suppressed significantly over the constant τ case. Some features in conductance that

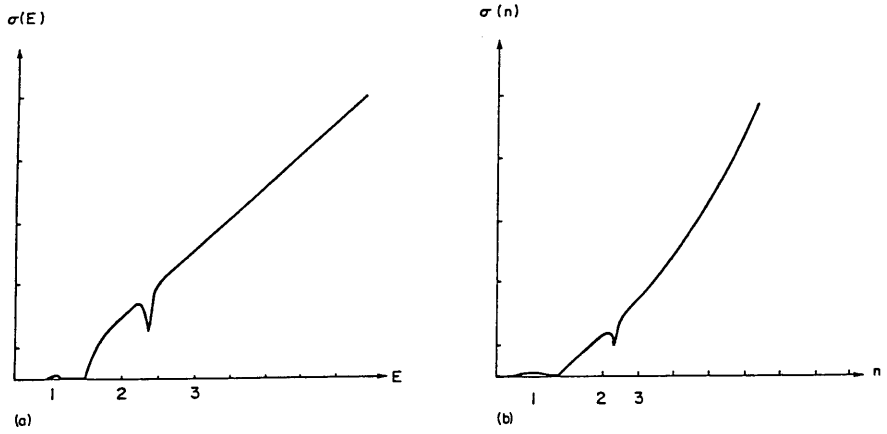


Fig. 24. Conductivity versus (a) Fermi energy and (b) density for $V_{p-p} = 0.1$ meV and $\tau^{-1} \sim N(E)$. Band numbers are shown along the x-axis.

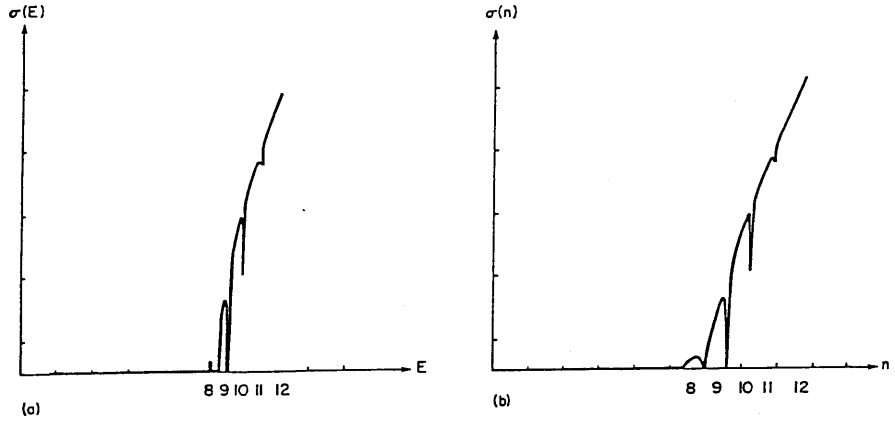


Fig. 25. Conductivity versus (a) Fermi energy and (b) density for $V_{p-p} = 10$ meV and $\tau^{-1} \sim N(E)$. Band numbers are shown along the x-axis.

were present in the constant τ case have now disappeared. This is because, in addition to the Fermi group velocity approaching zero at a Brillouin zone boundary or zone center, the scattering rate now becomes very high as well. This suppresses conductance still further over the constant τ case. Again, as for an one-dimensional conductor, the envelope of conductivity versus density is parabolic in density for high densities, while the envelope of conductivity versus energy is linear in energy.

Substantial energy-level broadening in the one-dimensional case will alter our qualitative results for conductivity. For small periodic potentials, conductivity in the bound energy levels should rise and approach the free-electron gas result as we increase energy level broadening.

2. Two Dimensions

The qualitative structure of conductivity versus density will change once we add free motion perpendicular to the superlattice. The additional y degree of freedom in the electron motion will broaden the structure in the superlattice or x degree of freedom, resulting in a weaker modulation of the conductance. This has been explored numerically in the conductance versus Fermi energy as in Fig. 26, after Kelly [33]. Here $\sigma_{xx}(E)$ does not fall to zero, since electron motion perpendicular to the superlattice is forbidden only for a given k_x, not a given E. For any given Fermi energy, it is always possible to choose many k_y's such that k_x does not match the Bragg reflection condition. Some current leaks through because there are no true gaps in the energy spectrum.

Using arguments analogous to those for computing the two-dimensional density of states in a superlattice, we can express the two-dimensional conductivity as a convolution:

$$\sigma_{xx,2\text{-D}}(E) = \tfrac{1}{2}\sigma_{xx,1\text{-D}}(E) \otimes N_y(E) \tag{58}$$

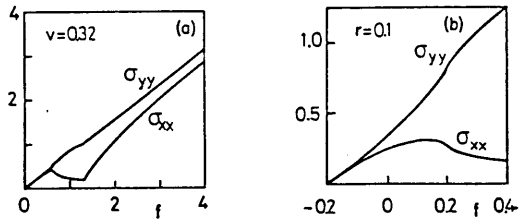


Fig. 26. Conductivity versus Fermi energy for a two-dimensional superlattice. The conductivity σ_{xx} as a function of reduced Fermi energy in the presence of (a) a weak and (b) strong periodic potential, both in units of $e^2\tau/m^*a^2$. (After Kelly [33].)

Here $\sigma_{xx,1\text{-D}}(E)$ is the conductivity of a one-dimensional superlattice, calculated in the previous section. $N_y(E)$ is the one-dimensional free-electron density of states of Eq. (34). The same physical interpretation of this convolution we made when discussing two-dimensional density of states in the superlattice is again valid for the conductivity. Free motion parallel to the superlattice broadens features in the superlattice density of states and conductivity by convolution with the free-electron density of states. Conductivity perpendicular to the superlattice does not fall to zero, qualitatively explaining the results shown in Fig. 26 in which $\sigma(E)$ has roughly a $1/\sqrt{E}$ dependence in the gap regions. If τ is not a constant in the two-dimensional superlattice, we must first carry out the calculation assuming it is constant, then scale conductivity by the variable scattering time afterwards.[10]

Consider now conduction parallel to the superlattice, again for a constant scattering time. Arguments similar to those above for $\sigma_{xx,2\text{-D}}(E)$ yield the relation

$$\sigma_{yy,2\text{-D}}(E) = \tfrac{1}{2}\sigma_{yy,1\text{-D}}(E) \otimes N_x(E) \tag{59}$$

Here $N_x(E)$ is the one-dimensional superlattice density of states calculated from solution of the central equation. Here $\sigma_{yy,1\text{-D}}(E)$ is the one-dimensional conductivity for a constant scattering time, as given in Section III. If the mean free time is a function of the particle's energy, we scale the conductance by this variable scattering time after carrying out the convolution.

The strongly bound energy levels, which carry no current perpendicular to the superlattice, approximate quite well a quasi-one-dimensional conductor for transport parallel to the superlattice. Convolution with the one-dimensional superlattice density of states, $N_x(E)$, determines how closely $\sigma_{yy,2\text{-D}}$ approximates the quasi-one-dimensional conductor. In the limit that $N_x(E)$ is a sum of delta functions, the approximation is exact. This will be roughly true in the strongly bound energy levels. As the Fermi level sweeps above the top of the cosine potential, there should be a transition from one-dimensional to two-dimensional behavior in the conductance.

One experiment that has attempted to probe the current–voltage characteristics of the superlattice transistor is shown in Fig. 17. Full details of this work by Warren *et al.* are found in Ref. [34]. We reproduce the experimental curves[11] in Fig. 27. Ismail *et al.* [35] have seen similar results in a GaAs

[10] A much stronger modulation in conductivity perpendicular to the superlattice, similar to the one-dimensional superlattice, is possible if the superlattice potential is periodic in both the x direction and y direction. To calculate the conductivity in these proposed devices, it is both necessary to generalize the central equation to two dimensions to find the energy dispersion and to use the more difficult method of integrating over the Fermi line to calculate conductivity.

[11] The transconductance plot is versus current, not voltage. The rationale is that current is proportional to density, which is proportional to Fermi level in two dimensions. This argument is true only for a free-electron gas in the case of constant τ.

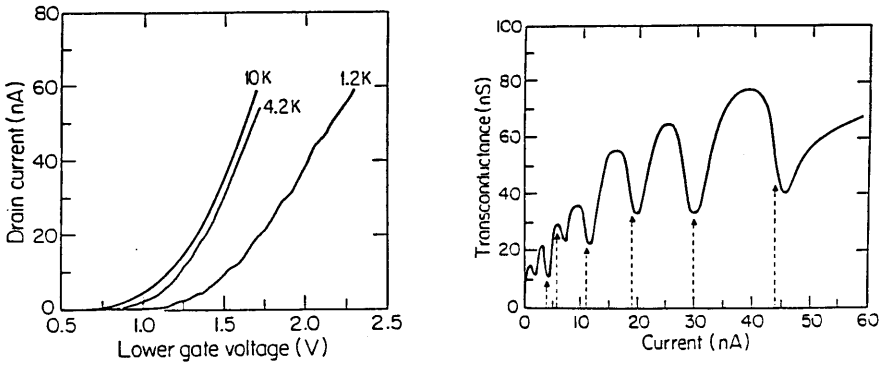


Fig. 27. Measured characteristics of superlattice transistor. (After Warren *et al.* [34].)

surface superlattice. This device structure has been simulated by Tokura and Tsubaki [36].

While our calculations of conductivity assume an infinite superlattice, the superlattice transistor has finite extent. It may be important that the electron coherence length limits the number of superlattice periods seen by the wave function. This constraint may, in fact, play some role in the observability criterion for a superlattice transistor. The finite device extent can be modeled in several ways. We can broaden the energy eigenvalues by the time to diffuse one inelastic length (i.e., by $\hbar/\tau_\phi$) as we discussed in Section II. This is not a severe constraint for a diffusive superlattice, since we know "elastic" energy-level broadening smears the energy levels by $\hbar/\tau$, a much more severe limitation as $\tau < \tau_\phi$. It is also possible to explore the effect of a finite device size on the conductance using transmission coefficient.

V. DYNAMICS OF ELECTRON TRANSPORT IN HIGH ELECTRIC FIELDS

A. Boltzmann Transport Theory

The steady-state terminal I-V characteristics of long-channel MOSFETs can be accurately calculated by describing the flow of carriers in the channel by the classical "drift-diffusion" hydrodynamic equation, written below for electrons:

$$j_n = q\mu n\mathscr{E} + qD\nabla n \tag{60}$$

where j_n is the electron current density, q is the magnitude of the electron charge, μ the electron mobility, n the electron concentration, $\mathscr{E}$ is the electric field, and D is the electron diffusivity. Assuming little or no impact ionization due to the primary current in the channel, Eq. (60) along with the following two equations model accurately (NMOS) devices under steady-state low-field conditions:

$$\nabla \cdot j_n = 0 \tag{61}$$

$$\nabla^2 \psi = \frac{q}{\varepsilon_{si}} (n - N_D^+ + N_A^-) \tag{62}$$

Equation (61) is an expression of carrier conservation, and Eq. (62) is the Poisson equation with ψ the electrostatic potential, ε_{si} the silicon dielectric constant, and N_D^+ and N_A^- the ionized donor and acceptor concentrations, respectively. Equations (61) and (62) are derived from Boltzmann's equation for electrons in steady-state under the assumption that the distribution function of electrons is Maxwellian with a characteristic width determined by the lattice temperature T_0. In other words, it is assumed that the electron temperature T_e is equal to T_0, which is equivalent to saying that the electric field is low enough so that the drift velocity of electrons $\mathbf{v}$ is much smaller than their thermal velocity, for example,

$$|\mathbf{v}| \ll \sqrt{\frac{3k_B T_e}{m^*}} \tag{63}$$

At room temperature, $T_e = 300$ K, Eq. (63) becomes $|\mathbf{v}| \ll 2.4 \times 10^7$ cm/sec.

Equation (60) can also be considered as an expression of momentum conservation while Eq. (61) is a statement of mass conservation, both in steady-state under the condition that the distribution function is only weakly perturbed from its thermal equilibrium state. Since the energy of the electrons is assumed in equilibrium with the lattice there is no need for an expression of energy conservation, so Eqs. (60) and (61) are a complete set of equations of electron motion. For most practical applications of modern MOSFETs the "weak perturbation" conditions just described are readily violated. For example, in a MOSFET in saturation, electrons move with drift velocity equal to or exceeding the value of the so-called electron saturation velocity v_{sat}, which is around 10^7 cm/sec at room temperature [37]. Similarly, the electron temperature T_e rises above 2000 or 3000 K [38]. Thus, both the thermal equilibrium condition, $T_e \approx T_0$, and the condition in Eq. (63) are violated.

To properly model MOSFETs under these conditions the Boltzmann equation must be solved directly to derive the distribution function with no *a priori* assumptions whatsoever about its form. This is typically done

by Monte Carlo techniques and has the distinct advantage that the collisional terms can be properly taken into account [39]. Since accurate direct integration of the Boltzmann equation is computationally expensive and does not lend itself to increased physical insight, a popular alternative solution is to assume a displaced Maxwellian form for the distribution function and to directly derive from the Boltzmann equation expressions for the conservation of electrons, momentum, and energy [40, 41]:

$$\frac{\partial n}{\partial t}+\nabla\cdot(n\mathbf{v})=G \tag{64}$$

$$\frac{\partial m^*\mathbf{v}}{\partial t}+\mathbf{v}\cdot\nabla(m^*\mathbf{v})=-q\varepsilon-\frac{k_B}{n}\nabla(nT_e)-\frac{m^*\mathbf{v}}{\tau} \tag{65}$$

$$\frac{\partial w}{\partial t}+\mathbf{v}\cdot\nabla w=-q\varepsilon\cdot\mathbf{v}-\frac{k_B}{n}\nabla\cdot(nT_e\mathbf{v})-\frac{1}{n}\nabla\cdot(\kappa\nabla T_e)-\frac{w-w_0}{\tau_\phi} \tag{66}$$

where G is the net generation rate, τ is the momentum relaxation time, τ_ϕ is the energy relaxation time, κ is the electron thermal conductivity, and w is the electron kinetic energy given by

$$w=\tfrac{3}{2}k_B T_e+\frac{m^*v^2}{2}=\begin{pmatrix}\text{energy of}\\ \text{random motion}\end{pmatrix}+\begin{pmatrix}\text{energy of}\\ \text{convective motion}\end{pmatrix} \tag{67}$$

and w_0 is the equilibrium energy, $\frac{3}{2}k_B T_0$. Note that $v=|\mathbf{v}|$. Equations (64)–(66) have simple physical interpretations. Equation (64) is the particle conservation equation under nonsteady-state conditions and in the presence of net generation, for example, due to impact ionization. Equation (65) is the electron momentum conservation equation. The left-hand side is the sum of the inertial and momentum outflow terms, set equal to the sum of the driving forces consisting of the electric field and temperature gradient (pressure) terms minus the rate of momentum lost to collisions expressed in the simplified momentum relaxation form.

Equation (66) is the electron energy conservation equation. Similar to Eq. (65), the left-hand side is the sum of the rate of change and energy outflow terms, and it is equal to the sum of the energy rate supplied by the electric field, the work rate performed by the electron pressure, the divergence of heat flux, and the rate of energy lost to collisions expressed in the simplified energy relaxation time form.

The conservation equations along with the Poisson equation [Eq. (62)] have been solved to various degrees of approximation depending on the number of terms that are neglected. The two relaxation times are generally complicated functions of electron energy [37, 42, 43]. For silicon the energy relaxation time for room lattice temperature is a slowly increasing function

of electron energy with values somewhat less than 10^{-12} sec. On the other hand, the momentum relaxation time decreases sharply with energy because of increasing optical phonon scattering. It is worthwhile mentioning here that under steady-state conditions and provided the electric field does not change very abruptly in space, the convective energy term in Eq. (67), $m^*v^2/2$, is always much smaller than the temperature term [43, 44]. Often the convective term is neglected, and, therefore, T_e is assumed to represent directly the electron energy. So, often τ and τ_ϕ are given as function of T_e.

B. Electron Velocity Overshoot

Carrier velocity overshoot has been much discussed recently, primarily because of its promise to obtain currents from short-channel compound semiconductor FET devices exceeding those predicted by simple device scaling [44, 45]. Velocity overshoot means that in some region of the device the carrier velocity exceeds the saturation velocity. The phenomenon occurs when the electric field in the device changes rapidly in space or in time, and it is a direct result of the dynamic nature of the conservation Eqs. (65) and (66). Here we will restrict our discussion to silicon with an eye on its applicability to short-channel MOSFETs.

The simplest model for a short-channel FET is to imagine that electrons are continuously emitted from a cold reservoir (i.e., the source), which is in equilibrium with the lattice with $T_e = T_0$, into a region of uniform high field (i.e., the channel) where they proceed to accelerate and heat up according to Eqs. (64)-(66). They transit the channel and are collected at the drain. Since this is a one-dimensional steady-state problem, Eqs. (64)-(66) can be written as

$$\frac{d(nv)}{dx} = G \tag{68}$$

$$m^* v \frac{dv}{dx} = -q\mathscr{E} - \frac{k_B}{n}\frac{d(nT_e)}{dx} - \frac{m^* v}{\tau} \tag{69}$$

$$v\frac{dw}{dx} = -qv\mathscr{E} - \frac{k_B}{n}\frac{d(nvT_e)}{dx} - \frac{1}{n}\frac{d}{dx}\left(\kappa\frac{dT_e}{dx}\right) - \frac{w - w_0}{\tau_\phi} \tag{70}$$

Correct treatment of even this simplified problem requires simultaneous numerical solution of the above equations. Several simplifications have been done in the past by omitting various terms, for example, all terms with differentials involving T_e [46]. However, Baccarani and Wordeman [41] have shown that it is not safe to neglect these terms under general conditions, and that indeed they might have profound effect in the solution.

Before proceeding further we should discuss briefly the two relaxation times. These can be obtained as functions of T_e under steady-state homogeneous conditions from Eqs. (69) and (70). Noting that under these conditions we have $w \approx \frac{3}{2} k_B T_e$, we obtain

$$\tau(T_e) = \frac{m^*}{q} \frac{v}{|\mathscr{E}|} \tag{71}$$

$$\tau_\phi(T_e) = \frac{3}{2} k_B \frac{T_e - T_0}{q\mathscr{E}^2} \frac{m^*}{q\tau(T_e)} \tag{72}$$

under steady-state homogeneous conditions T_e is a direct function of $\mathscr{E}$. It is interesting to note that Eq. (72) states that all the joule heating from the electric field relaxes like the electron energy difference from thermal equilibrium:

$$\sigma_0 \mathscr{E}^2 = \frac{n(w - w_0)}{\tau_\phi} \tag{73}$$

where $\sigma_0 = ne^2\tau/m^*$ is the conductivity due to electrons. To eliminate $\mathscr{E}$ from Eqs. (71) and (72), we note first that

$$v/\mathscr{E} = \mu(T_e) = \mu(\mathscr{E}) \tag{74}$$

where μ is the chordal mobility of the electrons, which is experimentally available as a function of $\mathscr{E}$. A reasonable fit to the data is [47]

$$\mu(\mathscr{E}) = \mu_0[1 + (\mu_0 \mathscr{E}/v_{sat})^2]^{-1/2} \tag{75}$$

where μ_0 is the low-field ($T_e = T_0$) mobility of electrons. Resorting now to an additional experimental observation, which is also confirmed by Monte Carlo calculations, namely, that the diffusivity of electrons is nearly independent of w and, therefore, of T_e [37], we obtain from the generalized Einstein relation [41]

$$\mu(T_e) = \frac{T_0}{T_e} \mu_0 \tag{76}$$

Therefore, Eq. (71) becomes

$$\tau(T_e) = \frac{m^* \mu_0}{q} \frac{T_0}{T_e} \tag{77}$$

and by using Eqs. (75) and (77) in Eq. (72)

$$\tau_\phi = \frac{3}{2} \frac{k_B T_0}{q} \frac{\mu_0}{v_{sat}^2} \frac{1}{1 + (T_0/T_e)} \tag{78}$$

which is an increasing function of the electron temperature. It is interesting to note that at a given temperature T_0 the higher the low-field mobility μ_0,

the higher are both relaxation times and, therefore, the more important are the dynamical effects of carrier transport under inhomogeneous conditions.

To demonstrate the effect of μ_0 and to illustrate clearly the phenomenon of velocity overshoot, we simplify the set of Eqs. (69) and (70) by ignoring all terms involving T_e, as already done by Huang and Ladbrooke [46]. The reader is cautioned here that this simplification cannot be justified and is allowed here only to provide a qualitative description of nonstationary, or dynamic, transport effects. Also, omission of the continuity equation is unjustifiable. Our simplified momentum and energy conservation equations become

$$m^* v \frac{dv}{dx} + \frac{qv}{\mu_n(T_e)} = -q\mathscr{E} \tag{79}$$

$$\frac{d(T_e - T_0)}{dx} + \frac{T_e - T_0}{\tau_\phi \mu(T_e)\mathscr{E}} = -\frac{2}{3}\frac{q}{k_B}\mathscr{E} \tag{80}$$

where it was assumed that $w \approx \frac{3}{2} k_B T_e$. These two equations are decoupled. Equation (80) is solved for a given $\mathscr{E}$ using Eqs. (76) and (78). The resulting $T_e(x)$ is then substituted into Eq. (76), and Eq. (79) is solved for $v(x)$. The results are shown in Fig. 28 for $\mathscr{E} = 50$ kV/cm and various values of μ_0. As can be seen in Fig. 28a, at a short distance from the source, electrons can accelerate to velocities higher than v_{sat} when μ_0 is greater than about 500 cm^2/V sec. The velocity overshoot is a direct consequence of the slower rise of $T_e(x)$ as shown in Fig. 28b. It is also worthwhile to note that the higher the mobility, the higher the final T_e, because the energy loss rate is decreasing as mobility increases.

Proper solution of the system of Eqs. (68)–(70) with $G = 0$ (i.e., no net carrier generation) has been performed by Baccarani and Wordeman [41]. We refer here to their results in order to discuss the salient features of the velocity overshoot effect. Figure 29 shows the distribution of electron concentration, electron velocity, and electron temperature along with the ratio of the convective energy to the total kinetic energy of the electrons. In this hypothetical device, electrons with concentration $n = 10^{16}\ cm^{-3}$ and temperature $T_e = 300$ K are assumed to be presented at $x = 0$ to a low electric field region where $\mathscr{E}_l = 10^3$ V/cm. At $x = 0.5\ \mu m$, it is assumed that the field changes abruptly to $\mathscr{E}_h = 10^5$ V/cm. This low-to-high field transition corresponds somewhat to a MOSFET channel in saturation. In Figs. 29a and 29b the distribution of concentration and of velocity calculated from the classical drift-diffusion (DD) model [i.e., from Eqs. (60) and (61)] is also shown (dotted lines). The only difference is at the field discontinuity. There the carrier velocity overshoots the saturation velocity, $v_{sat} = 10^7$ cm/sec, by a factor of almost two, while the DD model simply predicts $v = v_{sat}$. Close

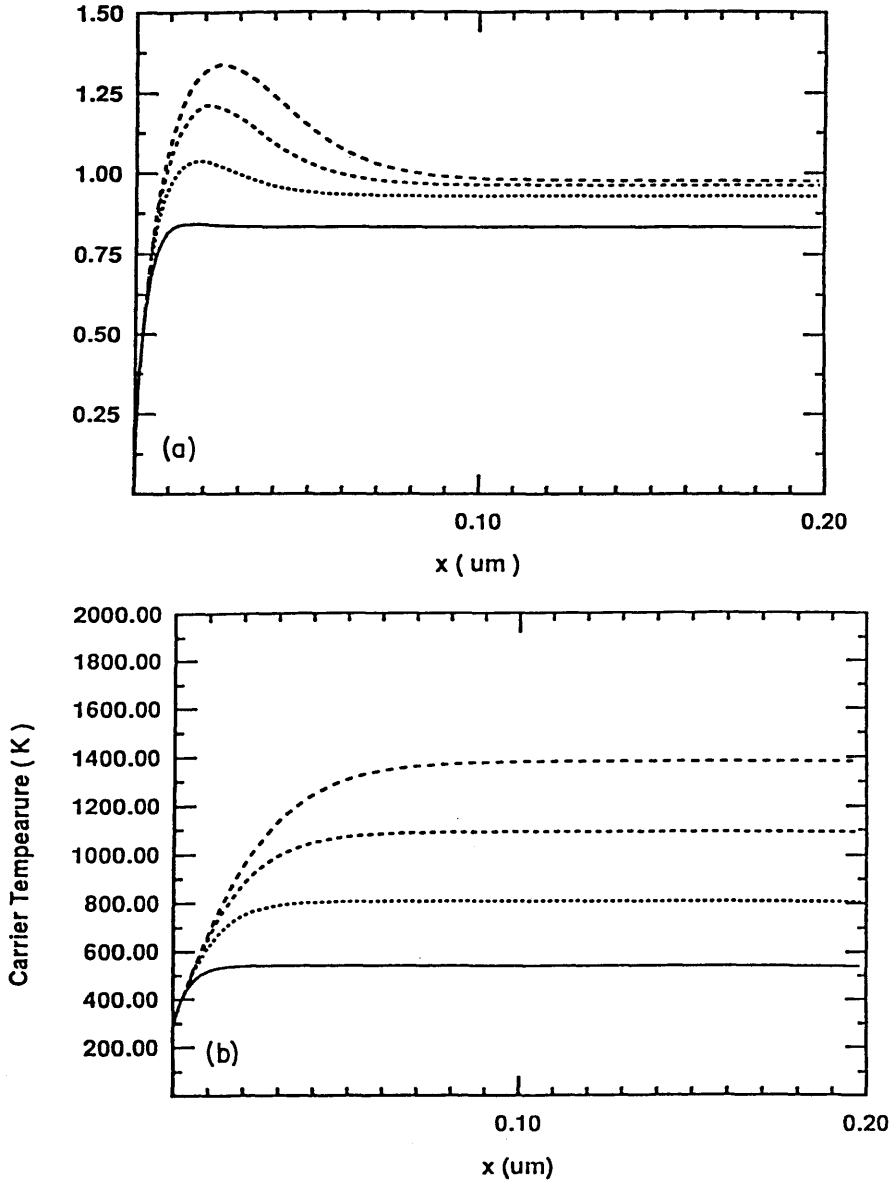


Fig. 28. (a) Electron mean velocity versus distance from the "cold" source injection point into a constant field ($T_0 = 300$ K, $\mathscr{E} = 5 \times 10^4$ V/cm). (b) The corresponding electron temperature ($T_0 = 300$ K, $\mathscr{E} = 5 \times 10^4$ V/cm). Lines correspond to mobilities from 300 (bottom) to 900 cm^2/V sec (top) in increments of 200 cm^2/V sec.

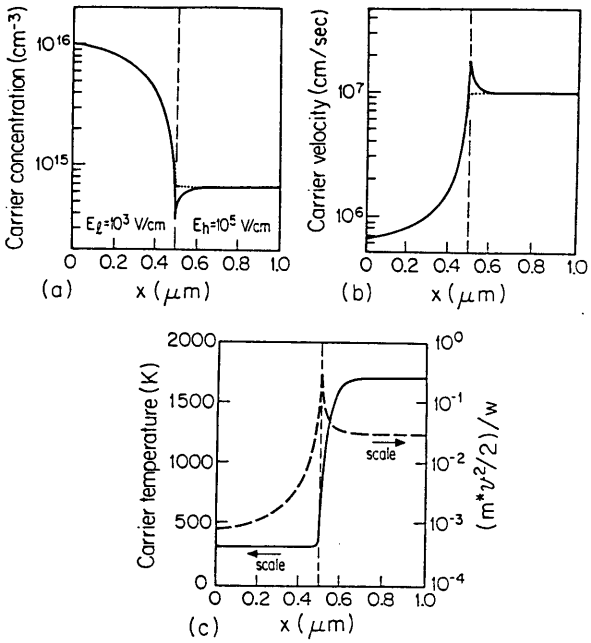


Fig. 29. (a) Electron concentration versus distance for a low-to-high step field distribution. Electron mobility is assumed to be 500 cm^2/V sec. Solid line is the ME solution and dotted line is the DD solution. (b) Corresponding electron velocity versus distance. (c) Corresponding electron temperature and ratio of convective to total energy versus distance. (After Baccarani and Wordeman [41].)

examination of Fig. 29c reveals that, as expected, the overshoot is the result of the slow rise of T_e. It is also interesting to note that at the field discontinuity, the convective energy is about 30% of the total kinetic energy and, therefore, nonnegligible. It is also important to note that, as can be seen from Figs. 29a and 29b, the current density, $j = -qnv$, predicted by the momentum-energy (ME) model and the DD model are identical. In other words, velocity may be overshooting inside a device at a low-to-high field transition without the effect being evident in the I-V characteristics. Indeed several MOSFET simulations based on the ME model have demonstrated this for devices with channel lengths around 0.5 to 1.0 μm operating in the saturation region [48]. Under such conditions the only benefit of velocity overshoot is a slightly reduced carrier transit time.

In order to take advantage of velocity overshoot to increase the current of a device the electrons must be injected cold into a high-field region. This situation can be approximated in extreme submicron-channel-length

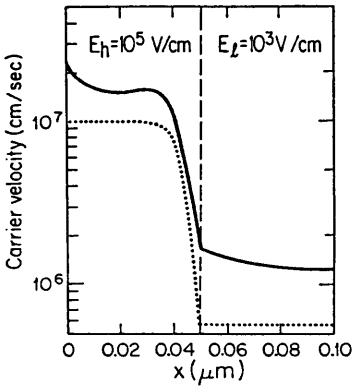


Fig. 30. Calculated electron velocity versus distance for a high-to-low step field distribution. Electron mobility is assumed to be 500 cm^2/V sec. (After Baccarani and Wordeman [41].)

MOSFETs. A hypothetical example of this effect is demonstrated in Fig. 30, after Baccarani and Wordeman [41], which shows the velocity distribution in a device with a high-low field discontinuity at $x = 0.05$ μm. In this example electrons at $n = 10^{16}$ cm^{-3} are presented cold to a 10^5 V/cm field at $x = 0$. As can be seen at the "source," $v \approx 2 \times 10^7$ cm/sec, which means that there is a current enhancement by a factor of two in the ME solution (solid line) compared to the DD solution (dotted line).

C. Extreme Submicron-Channel-Length MOSFETs

As we discussed previously, the benefit of velocity overshoot in the channel of a MOSFET cannot be realized until the carriers are overshooting very near their injection point from the source into the channel. The expected benefit is increased drain current and transconductance in the saturation region of operation.

The drain current I_d can be written as

$$I_d = WQ_{I0}v_0 \tag{81}$$

where W is the channel width, Q_{I0} is the channel charge density at the source end ($x = 0$), and v_0 is the carrier velocity at $x = 0$. Assuming no parasitic series resistances, $Q_{I0} = C_g(V_{gs} - V_T)$, where C_g is the gate capacitance per unit area, and V_{gs} and V_T have their usual meaning; Eq. (81) becomes

$$I_d = WC_g(V_{gs} - V_T)v_0 \tag{82}$$

Without velocity overshoot $v_0 \leq v_{sat}$, and, therefore, for a given C_g and V_T there exists an upper limit, I_{dl}, for I_d:

$$I_{dl} \leq WC_g(V_{gs} - V_T)v_{sat} \tag{83}$$

On the other hand, if velocity overshoot does take place at the source end of the channel, v_0 can exceed v_{sat}, and, therefore, I_d can exceed I_{dl}.

Because of uncertainties in defining V_T experimentally it is safer to study v_0 by means of the device intrinsic transconductance:

$$g_{mi} = \left.\frac{\partial I_d}{\partial V_{gs}}\right|_{V_{ds}} = WC_g v_0 \tag{84}$$

g_{mi} can be obtained easily from the device transconductance after corrections for the parasitic series resistances. On the other hand, C_g cannot be obtained very accurately. Large-area capacitance measurement yields the static gate capacitance C_g', which may differ from the dynamic capacitance C_g. Calculations [49] and simulation [50] show that C_g at $v_0 \approx v_{sat}$ is only slightly smaller than C_g'. Keeping this in mind, a lower bound for v_0 can be obtained from experiment:

$$v_0 \geq \frac{g_{mi}}{WC_g'} \tag{85}$$

The first reported experimental observation of velocity overshoot was in sub-100-nm channel-length MOSFETs at 4.2 K [51]. No overshoot was observed at 77 or 300 K, according to that report. Subsequently Shahidi *et al.* [50] reported velocity overshoot at 77 and 300 K in MOSFETs with similar channel lengths. The progress from the first report to the second is attributed to an increase in the low-field mobility μ_0 due to deliberate reduction in the channel doping. The mobility in the second report was

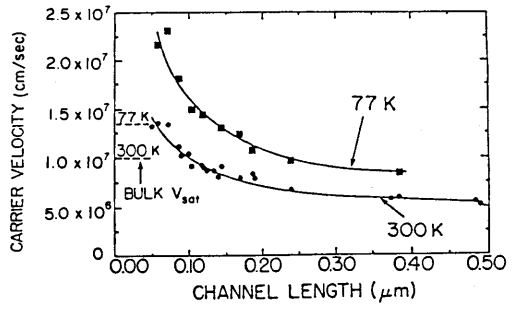


Fig. 31. Electron velocity versus channel length at room temperature and 77 K. (After Shahidi *et al.* [50].)

450 $cm^2/V\,sec$ with a surface concentration of dopants of $5 \times 10^{16}\,cm^{-3}$, while in the first report mobility was 300 $cm^2/V\,sec$ with surface concentration of $10^{18}\,cm^{-3}$. This serves to illustrate the significance of low-field mobility in the attainment of velocity overshoot, as discussed in Section V.B. Calculated v_0 according to Eq. (85) versus channel length at 77 K and at 300 K, from Ref. [50], is shown in Fig. 31. As can be seen, overshoot is observed at somewhat longer channels at 77 K as compared to 300 K. This is expected because the energy relaxation time increases somewhat with decreasing lattice temperature [37, 42] even though the approximate expression given by Eq. (78) does not reflect this fact.

ACKNOWLEDGMENTS

We acknowledge useful discussions with C. Kane, X. C. Xie, H. I. Smith, and M. A. Kastner. We gratefully acknowledge financial support from an IBM fellowship (P.F.B.), a MIT-EECS Departmental grant (T.P.O. and P.F.B.), and the support of U.S. Air Force contract AFOSR-85-0376.

REFERENCES

1. S. Hershfield and V. Ambegaokar, *Phys. Rev. B: Condens. Matter* [3] **34,** 2147 (1986).
2. P. F. Bagwell, Quantum mechanical transport phenomena in nanostructured inversion layers. S.M. Thesis, MIT, Cambridge, Massachusetts (1988).
3. O. Madelung, "Introduction to Solid State Theory." Springer-Verlag, Berlin and New York, 1978.
4. D. E. Prober, Quantum transport in microstructures. *In* "Microcircuit Engineering '86" (H. W. Lehmann and Ch. Bleiker, eds.). North-Holland, Amsterdam, 1986, and references therein.
5. Short discussions of electron heating are found in R. A. Webb, A. Hartstein, J. J. Wainer, and A. B. Fowler, *Phys. Rev. Lett.* **54,** 1577 (1985); see also A. B. Fowler, A. Harstein, and R. A. Webb, *ibid.* **48,** 196 (1982).
6. S. Y. Chou, D. A. Antoniadis, H. I. Smith, and M. A. Kastner, *Solid State Commun.* **61,** 571 (1987); see also A. Yagi, *Jpn J. Appl. Phys.* **20,** 2255 (1981).
7. G. Bergman, *Phys. Rep.* **107,** 1 (1984).
8. B. L. Al'tshuler, D. Khmel'nitskii, A. I. Larkin, and P. A. Lee, *Phys. Rev. B* **22,** 5142 (1980).
9. D. E. Khmel'nitskii, *Physica B (Amsterdam)* **126B,** 235 (1984); see also A. I. Larkin and D. E. Khmel'nitskii, *Sov. Phys.—Usp. (Engl. Transl.)* **25**(3), 185 (1982).
10. A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics.*" Dover, New York, 1975.
11. N. W. Ashcroft and N. D. Mermin, "Solid State Physics." Holt, New York, 1976.
12. C. Kittel, "Introduction to Solid State Physics." Wiley, New York, 1986.
13. P. A. Lee and A. D. Stone, *Phys. Rev. Lett.* **55,** 1622 (1985). The same effect was predicted independently by B. L. Al'tshuler, *Sov. Phys. JETP Lett. (Engl. Transl.)* **41,** 648 (1985).

14. D. A. Antoniadis, A. C. Warren, and H. I. Smith, *Tech. Dig.—Int. Electron Devices Meet.*, p. 558 (1985).
15. M. A. Kastner, R. F. Kwasnik, and J. C. Licini, *Phys. Rev. B: Condens. Matter* [3] **36,** 8015 (1987).
16. R. L. Liboff, "Introductory Quantum Mechanics." Holden-Day, Oakland, Caliifornia, 1980.
17. W. J. Skocpol, L. D. Jackel, R. E. Howard, H. G. Craighead, L. A. Fetter, P. M. Mankiewich, P. Grabbe, and D. M. Tennant, *Surf. Sci.* **142,** 14 (1984); see also R. G. Wheeler, K. K. Choi, and R. Wisnieff, *ibid.*, 19.
18. S. DasSarma and X. C. Xie, *Phys. Rev. B* **35,** 9875 (1987).
19. A. C. Warren, D. A. Antoniadis, and H. I. Smith, *Phys. Rev. Lett.* **56,** 1858 (1986).
20. S. B. Kaplan and A. C. Warren, *Phys. Rev. B: Condens. Matter* [3] **34,** 1346 (1986).
21. D. Vollhardt, *in* "Solid State Physics, Lecture Notes of the NORDITA Spring School, Vol. 1, p. 73, Tvärmine, Finland, 1984.
22. C. P. Umbach, S. Washburn, R. B. Laibowitz, and R. A. Webb, *LT-17 Contrib. Pap.* 717 (1984); see also C. P. Umbach, S. Washburn, R. B. Laibowitz, and R. A. Webb, *Phys. Rev. B: Condens. Matter* [3] **30,** 4048 (1984); see also S. Washburn, C. P. Umbach, R. B. Laibowitz, and R. A. Webb, *ibid.* **32,** 4789 (1985); see also R. A. Webb, S. Washburn, C. P. Umbach, and R. B. Laibowitz, *Phys. Rev. Lett.* **54,** 2696 (1985).
23. J. C. Licini, D. J. Bishop, M. A. Kastner, and J. Melngailis, *Phys. Rev. Lett.* **55,** 2987 (1985).
24. W. J. Skocpol, P. M. Mankiewich, R. H. Howard, L. D. Jackel, D. M. Tennant, and A. D. Stone, *Phys. Rev. Lett.* **56,** 2865 (1986).
25. P. A. Lee, "Proceedings of STATPHYS-16." North-Holland Publ., Amsterdam, 1986; see also Y. Imry, *Europhys. Lett.* **1,** 249 (1986); see also B. L. Al'tshuler and B. I. Shklovskii, *Sov. Phys. JETP* (Engl. transl.) **64,** 127 (1986). See also A. D. Stone and Y. Imry, *Phys. Rev. Lett.* **56,** 189 (1986).
26. W. J. Skocpol, P. M. Mankiewich, R. E. Howard, L. D. Jackel, and D. M. Tennant, *Phys. Rev. Lett.* **58,** 2347 (1987); see also A. Benoit, C. P. Umbach, R. B. Laibowitz, and R. A. Webb, *ibid.*, p. 2343; see also C. P. Umbach, P. Santhanam, C. van Haesendonck, and R. A. Webb, *Appl. Phys. Lett.* **50**(18), 1289 (1987).
27. H. I. Smith, M.I.T., private communication.
28. R. T. Bate, *Bull. Am. Phys. Soc.* **22,** 407 (1977).
29. L. Esaki and R. Tsu, *IBM J. Res. Dev.* **14,** 61 (1970); see also L. Esaki, L. L. Chang, W. E. Howard, and V. L. Rideout, *Proc. Int. Conf. Phys. Semicond. 11th, 1972,* p. 431 (1972); see also L. L. Chang, L. Esaki, and R. Tsu, *Appl. Phys. Lett.* **24**(12), 593 (1974). This has recently been observed in silicon; see S. Miyazaki, Y. Ihara, and M. Hirose, *Phys. Rev. Lett.* **59,** 125 (1987).
30. T. C. L. G. Sollner, E. R. Brown, and H. Q. Lee, *in* "Physics of Quantum Electronic Devices" (F. Capasso, ed.). Springer-Verlag, Berlin and New York (in press).
31. A. C. Warren, Surface superlattice and quasi-one-dimensional conduction in silicon inversion layers. Ph.D. Thesis, M.I.T., Cambridge, Massachusetts (1985).
32. A. C. Warren, D. A. Antoniadis, and H. I. Smith, *IEEE Electron Device Lett.* **EDL-7,** 413 (1986).
33. M. J. Kelly, *Surf. Sci.* **170,** 49 (1986).
34. A. C. Warren, D. A. Antoniadis, H. I. Smith, and J. Melngailis, *IEEE Electron Device Lett.* **EDL-6,** 294 (1985); see also A. C. Warren, I. Plotnik, E. H. Anderson, M. L. Schttenberg, D. A. Antoniadis, and H. I. Smith, *J. Vac. Sci. Technol., B* [2] **4,** 365 (1986).
35. K. Ismail, W. Chu, D. A. Antoniadis, and H. I. Smith, *Appl. Phys. Lett.* **52,** 1071 (1988).
36. Y. Tokura and K. Tsubaki, *Appl. Phys. Lett.* **51,** 1807 (1987).
37. C. Jacoboni, C. Canali, G. Ottaviani, and A. Alberigi Quararta, *Solid-State Electron.* **20,** 77 (1977).

38. A. Toriumi, M. Yoshimi, M. Iwase, Y. Akiyama, and K. Taniguchi, *IEEE Trans. Electron. Devices* **ED-34,** 1501 (1987).
39. H. D. Rees, *J. Phys. Chem. Solids* **30,** 643 (1969).
40. K. Blotekjaer, *IEEE Trans. Electron Devices* **ED-17,** 38 (1970).
41. G. Baccarani and M. R. Wordeman, *Solid-State Electron.* **28,** 407 (1985).
42. E. Rosencher, *Solid State Commun.* **38,** 1293 (1981).
43. J. P. Nougier, J. C. Vaissiere, D. Gasquet, J. Zimmermann, and C. Constant, *J. Appl. Phys.* **52,** 825 (1981).
44. J. G. Ruch, *IEEE Trans. Electron Devices* **ED-19,** 652 (1972).
45. T. J. Maloney and J. Frey, *IEEE Trans. Electron Devices* **ED-23,** 519 (1976).
46. R. S. Huang and P. H. Ladbrooke, *J. Appl. Phys.* **48,** 4791 (1977).
47. D. M. Caughey and R. E. Thomas, *IEEE Proc.* **55,** 2192 (1967).
48. T. Kobayashi and K. Saito, *IEEE Trans. Electron Devices* **ED-32,** 788 (1985).
49. G. W. Taylor, *Bell Lab. Tech. J.* **63,** 1325 (1984).
50. G. G. Shahidi, D. A. Antoniadis, and H. I. Smith, *IEEE Electron Device Lett.* **9,** 94 (1988).
51. S. Y. Chou, D. A. Antoniadis, and H. I. Smith, *IEEE Electron Device Lett.* **EDL-6,** 665 (1985).

Appendix B

Philip F. Bagwell and Terry P. Orlando, 'Broadened Conductivity Tensor and Density of States for a Superlattice Potential in One, Two, and Three Dimensions', Physical Review B, 40, 3735 (1989).

PHYSICAL REVIEW B VOLUME 40, NUMBER 6 15 AUGUST 1989-II

Broadened conductivity tensor and density of states for a superlattice potential in one, two, and three dimensions

Philip F. Bagwell and Terry P. Orlando
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 7 November 1988; revised manuscript received 27 February 1989)

We use a semiclassical method to calculate the conductivity tensor and density of states for independent electrons subject to a periodic potential along one direction. We incorporate into our model collision broadening due to impurities, temperature broadening, and an increase in spatial dimension due to free-electron motion perpendicular to the superlattice. We show that these three effects can all be put on an equal basis by convolving the one-dimensional unbroadened conductivity and density of states with the impurity spectral function, the Fermi-Dirac probability density function, and the one-dimensional free-electron density of states, respectively. We discuss the implications of our results for superlattice and quasi-one-dimensional field-effect transistors.

I. INTRODUCTION

Recent low-temperature conductance measurements on Si metal-oxide-semiconductor field-effect transistors (MOSFET's) and GaAs modulation-doped FET's (MODFET's) which impose a periodic modulation on electrons in a quasi-two-dimensional inversion layer,[1-4] or electrostatically confine electrons to quasi-one-dimension,[5-7] have demonstrated weak structure in the device conductance versus gate voltage attributed to Bragg diffraction from the surface superlattice and the filling of quasi-one-dimensional subbands, respectively. These devices all operate in the diffusive regime of transport in which the inelastic diffusion length is much longer than the elastic mean free path. The devices also operate in the typical incoherent regime of metallic conduction, for which the sample size is much longer than the inelastic diffusion length. The semiclassical equations of motion for an electron wave packet work well in this diffusive, incoherent regime of transport.

Some incoherence in the device is an advantage in that it allows measurements between different devices to be reproducible, i.e., independent of the specific arrangement of defects in the device. However, incoherence also imposes the operating limits through inelastic scattering and finite temperature on these types of quantum-electronic devices. In this paper we show these limits arise naturally from convolving the conductivity and density of states with the impurity spectral function, which includes elastic and inelastic scattering, and the Fermi-Dirac probability density function, to account for finite temperature. The effect of these two convolutions is to average the conductivity and density of states over an energy range of the inverse elastic mean free time and the temperature, respectively.

We use the semiclassical equations of motion together with the electronic band structure to calculate conductivity versus electron density and conductivity versus Fermi energy assuming a model Hamiltonian for the device. We emphasize that, throughout this paper, the potential will be periodic only in one direction regardless of whether the Hamiltonian has one, two, or three kinetic degrees of freedom. Because we assume a potential periodic only along the x direction, the free-electron motion along y and z decouples from motion along x except for the constraint that the total energy of motion is a constant. Increasing the dimensionality of space from the one-dimensional periodic potential along x means first adding free-electron motion perpendicular to the superlattice along the y direction to obtain two spatial dimensions, then adding free-electron motion perpendicular to the superlattice along the z direction to arrive at three spatial dimensions. Increasing the dimensionality of space in this manner allows the calculation of conductivity and density of states in two and three dimensions by first solving the problem in one dimension, then convolving with the one-dimensional free-electron density of states for each additional spatial dimension. This convolution provides an intuitive picture of the effect of free-electron motion perpendicular to the superlattice on the conductivity and density of states. Unlike the convolutions used to incorporate broadening effects, increasing the dimensionality of space via convolution does not lead to any well-defined "observability criteria" as to which spatial dimension a feature in conductivity or density of states will become extinct, since the one-dimensional free-electron density of states has no well-defined variance.

By restricting ourselves in this paper to devices which are incoherent, we can neglect so-called "mesoscopic" effects, i.e., wave-interference effects destroyed by self-averaging in the sample such as the universal fluctuations in conductance[8,9] or a nonzero off-diagonal element of the conductivity tensor due to a fixed impurity arrangement. Our results are also applicable to experiments on quasi-one-dimensional wires, since conduction perpendicular to the superlattice reduces to the same result as many independent wires in parallel when the Fermi energy is much less than the superlattice potential. Our method may also be used to calculate the conductivity of

many quasi-two-dimensional planes in parallel. A more detailed discussion of our calculational method may be found elsewhere.[10,11]

The superlattices incorporated as a gate into field effect transistors[1-4] offer several unique features as compared to standard compositional or doping superlattices. Both the Fermi level and strength of the periodic potential can be varied separately with dc voltage sources, thus the transistor functions as a voltage-controlled crystal. In addition, because the superlattice period is on the order of a thousand times the periodicity of the underlying crystal and the superlattice potential changes gradually over that distance, the effective-mass theorem should hold. There are no ambiguities as to the form of the Hamiltonian or boundary conditions on the envelope function. For these reasons, the superlattices incorporated into field-effect transistors offer a nearly ideal testing ground for the theory of electrical conduction in a periodic potential. The goal of this paper is to use a semiclassical theory of electrical conductance to gain insight into the current voltage characteristics of superlattice field-effect transistors[1-4] and quasi-one-dimensional field-effect transistors[5-7] at low temperatures. Elastic impurity scattering is the dominant velocity relaxation mechanism in these devices at low temperatures, so we design our computational method to account primarily for elastic scattering. Our method to calculate electrical conductivity must be altered if inelastic scattering comes to dominate velocity relaxation in the device, as is the case at higher temperature. We specifically do not account for phonon scattering except as it contributes to energy level broadening as discussed in Sec. II, and that we assume it maintains the electron gas temperature in equilibrium with the phonons. We also assume the electric fields are low so that nonparabolicities and subsidiary minima of the bulk band structure can be neglected.

II. IMPURITY BROADENING

Let τ be the elastic mean free scattering time and τ_ϕ be the mean inelastic scattering time. Electron motion is assumed diffusive so that the characteristic distance over which the electron's phase is coherent is the inelastic scattering length $L_\phi=\sqrt{D\tau_\phi}$, where D is the diffusion constant. The sample length L is assumed much larger than L_ϕ so that in the sample as a whole the electron's phase is incoherent. Therefore, the sample as a whole may be pictured as broken up into many coherent blocks of length L_ϕ, each block with a single-electron wave function valid throughout the block as in Fig. 1.

In an incoherent macroscopic sample we cannot measure the properties of each individual coherent region, we can only measure average properties. It is reasonable to assert that macroscopic properties like density of states, electrical conductivity, optical absorption, etc. are simply averages of these properties over all the coherent regions. Only the incoherence brought about by inelastic scattering allows this type of averaging over the different arrangements of impurities in each coherent region. This type of sample is also called "self-averaging", since it mimics an ensemble average over all possible arrange-

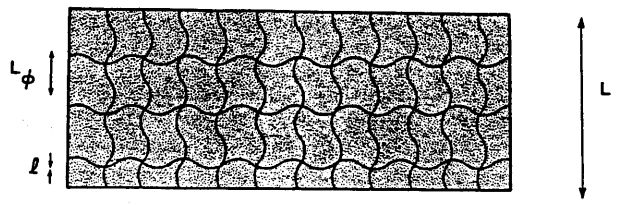


FIG. 1. An incoherent macroscopic sample. Here l is the elastic mean free path, L_ϕ the inelastic diffusion length, and L the sample size.

For such an ensemble averaged random impurity potential, the probability of an electron having some energy E away from the energy level E_m, where E_m is an energy level when the impurity potenial is zero, is given by the impurity spectral function A.[11,12] For the white-noise impurity power spectrum described below, A has a Lorentzian shape

$$A(E-E_m,\hbar/2\tau)=\frac{1}{\pi}\frac{\hbar/2\tau}{(E-E_m)^2+(\hbar/2\tau)^2} \tag{1}$$

and is a normalized probability density function

$$\int_{-\infty}^{\infty} A(E-E_m,\hbar/2\tau)dE=1 \tag{2}$$

with a full width at half maximum of $\hbar/\tau$.

If the sample were to have no scattering at all it could be described by its energy levels as given by the density of states $N(E)$. The density of states with scattering $N(E,\hbar/2\tau)$ is then given by finding all the states which have energy E, namely

$$N(E,\hbar/2\tau)=\int_{-\infty}^{\infty} N(E')A(E-E',\hbar/2\tau)dE' =N(E)\otimes A(E,\hbar/2\tau)\ . \tag{3}$$

Here $\otimes$ denotes a convolution in energy. Although $A(E,\hbar/2\tau)$ has a width set by the elastic scattering, it is inelastic scattering that leads to the incoherent averaging process necessary to produce "elastic" broadening.

The elastic scattering rate from the state k to the state k' using the Fermi golden rule is[12,13]

$$\frac{\hbar}{\tau_{k',k}}=2\pi n_{\text{imp}}|\langle k'|V|k\rangle|^2N(E_k) \tag{4}$$

with $E_k=E_{k'}$. Here n_{imp} is the density of impurity scatterers. We assume the impurities scatter incident electrons isotropically and with a white-noise power spectrum so that

$$|\langle k'|V|k\rangle|^2=|V(E_k)|^2=|V|^2=\text{const}\ . \tag{5}$$

Equation (5) is a good approximation if the scattering potential $V(r)$ for a single defect is taken to be short ranged.

III. TEMPERATURE BROADENING

The occupation of electron energy levels at finite temperature is determined by the Fermi-Dirac probability

$$f_0(E-\mu,k_BT)=\frac{1}{1+e^{(E-\mu)/k_BT}} . \quad (6)$$

The derivative of this function at finite temperature is known as the Fermi-Dirac probability density function or the thermal smearing function and is given by

$$-\frac{\partial f_0(E-\mu,k_BT)}{\partial E}=\frac{1}{4k_BT}\mathrm{sech}^2\left[\frac{E-\mu}{2k_BT}\right] . \quad (7)$$

The deviation of $-\partial f_0/\partial E$ from a δ function describes nonmonochromatic electrons at finite temperatures. Since $-\partial f_0/\partial E$ is a probability density function it satisfies

$$\int_{-\infty}^{\infty}\left[-\frac{\partial f_0(E-\mu,k_BT)}{\partial E}\right]dE=1 . \quad (8)$$

Also, $-\partial f_0/\partial E$ has a full width at half maximum of $3.5k_BT$.

We count the average density of particles in a quantum system at finite temperature by

$$n(\mu,k_BT)=\int_{-\infty}^{\infty}N(E)f_0(E-\mu,k_BT)dE . \quad (9)$$

We can express this result more usefully by differentiating Eq. (9) with respect to μ to get

$$\frac{dn(\mu,k_BT)}{d\mu}\equiv N(\mu,k_BT) , \quad (10)$$

so that $N(\mu,k_BT)$ is the thermodynamic density of states. Using $\partial f_0/\partial\mu=-\partial f_0/\partial E$ we find an expression for the thermodynamic density of states in terms of the zero-temperature density of states:

$$N(E,k_BT)=N(E)\otimes\left[-\frac{\partial f_0}{\partial E}(E,k_BT)\right] . \quad (11)$$

The electron density of states at zero temperature is convolved with its probability density of contributing a certain energy away from the Fermi level to obtain the thermodynamic density of states $N(E,k_BT)$ at finite temperature. We calculate the electron density by

$$n(\mu,k_BT)=\int_{-\infty}^{\mu}N(E,k_BT)dE . \quad (12)$$

In addition to broadening the zero-temperature density of states, finite temperature also introduces another important length scale. The thermal diffusion length,

$$L_T=\sqrt{D\tau_T}=\sqrt{D\hbar/k_BT} , \quad (13)$$

set approximately by the variance of Eq. (7), gives the mean distance for the nonmonochromatic electron beam, with all its frequency components initially in phase, to dephase. We cannot conduct electron interference experiments in a solid on distance scales longer than L_T because the electron has no well-defined phase beyond this distance.

Energy-level broadening due to "elastic" scattering is statistically independent of finite temperature broadening. Therefore, the total probability density function is a convolution of the broadening functions for these two independent random processes. This leads to an expression for the thermodynamic density of states

$$N(E,k_BT,\hbar/2\tau)=N(E)\otimes\left[-\frac{\partial f_0}{\partial E}(E,k_BT)\right]\otimes A(E,\hbar/2\tau) . \quad (14)$$

Writing this convolution explicitly gives

$$N(E,k_BT,\hbar/2\tau)=\int_{-\infty}^{\infty}dE'\int_{-\infty}^{\infty}dE''\,N(E-E'-E'')\left[-\frac{\partial f_0}{\partial E}(E',k_BT)\right]A(E'',\hbar/2\tau) . \quad (15)$$

The total electron density is still given by

$$n(\mu,k_BT,\hbar/2\tau)=\int_{-\infty}^{\mu}N(E,k_BT,\hbar/2\tau)dE . \quad (16)$$

IV. BROADENING OF THE CONDUCTIVITY

We determine the current from the semiclassical expression[14]

$$j_i=e\int v_i(\mathbf{k})f(\mathbf{k})2\frac{d^dk}{(2\pi)^d} , \quad (17)$$

where the factor 2 is due to the spin of the electron and d is the dimensionality of the system. For a periodic system the electron group velocity can be calculated as

$$v_i(\mathbf{k})=\frac{1}{\hbar}\frac{\partial}{\partial k_i}E(\mathbf{k}) , \quad (18)$$

from a solution of the central equation.[14] The dynamics of the wave vector $\mathbf{k}$ are determined by the semiclassical equations of motion for an electron wave packet,[14]

$$\frac{d}{dt}(\hbar\mathbf{k})=e\mathcal{E}-\frac{\hbar\mathbf{k}}{\tau} , \quad (19)$$

where the crystal momentum's time rate of change equals the electric force plus the average random force. In the time-independent case, the equilibrium $\mathbf{k}$ values are shifted by an amount

$$\Delta\mathbf{k}=\frac{e\mathcal{E}\tau}{\hbar} \quad (20)$$

leading to a change in the electron distribution function

$$f(\mathbf{k})=f_0(\mathbf{k}-\Delta\mathbf{k})\simeq f_0-\frac{\partial f_0}{\partial E_\mathbf{k}}\frac{\partial E_\mathbf{k}}{\partial\mathbf{k}}\cdot\Delta\mathbf{k}+\cdots , \quad (21)$$

$$j_i=e^2\int\left[\frac{1}{\hbar}\frac{\partial E_{\mathbf{k}}}{\partial k_i}\right]\left[\frac{1}{\hbar}\frac{\partial E_{\mathbf{k}}}{\partial k_j}\right]\tau(k_j)\times\left[-\frac{\partial f_0}{\partial E_{\mathbf{k}}}(E_{\mathbf{k}}-\mu,k_BT)\right]2\frac{d^dk}{(2\pi)^d}\mathcal{E}_j\ . \quad (22)$$

At zero temperature, where $\mu=E_F$ we have

$$-\frac{\partial f_0}{\partial E_{\mathbf{k}}}(E_{\mathbf{k}}-\mu,k_BT=0)=\delta(E_{\mathbf{k}}-E_F) \quad (23)$$

so that, if we assume the scattering time is only a function of the particle's energy and is independent of the direction of $\mathbf{k}$, we obtain a zero-temperature conductivity

$$\sigma_{ij}(E_F)=e^2\tau(E_F)\int\delta(E_{\mathbf{k}}-E_F)\left[\frac{1}{\hbar}\frac{\partial E_{\mathbf{k}}}{\partial k_i}\right]\left[\frac{1}{\hbar}\frac{\partial E_{\mathbf{k}}}{\partial k_j}\right]2\frac{d^dk}{(2\pi)^d}\ . \quad (24)$$

In general there is a different relaxation time for each k vector on the Fermi surface so that the scattering time cannot be removed from the integral over $\mathbf{k}$. Equation (24) assumes it is a good approximation to replace this k-dependent scattering rate by an average relaxation rate for all the k vectors on the Fermi surface. The average relaxation time $\tau(E_F)$ depends only on the Fermi energy and not on each individual k vector. This is a standard approximation in transport theory which is exact only for elastic scattering and isotropic constant energy surfaces.[13] At finite temperatures Eq. (24) becomes

$$\sigma_{ij}(E)=\sigma_{ij}(E)\otimes\left[-\frac{\partial f_0}{\partial E}(E,k_BT)\right]\ . \quad (25)$$

We merely convolve the zero-temperature conductivity with the thermal smearing function to obtain the finite-temperature conductivity. This is the usual semiclassical scheme for calculating the conductivity.[14]

A complete quantum-mechanical calculation[11] of the electrical conductivity with "elastic" broadening and with temperature broadening gives

$$\sigma(E,k_BT,\hbar/2\tau)=\sigma(E)\otimes\left[-\frac{\partial f_0}{\partial E}(E,k_BT)\right]\otimes A(E,\hbar/2\tau)\ . \quad (26)$$

The temperature smearing follows directly from our standard semiclassical treatment. Just as with the density of states, we merely convolve the zero-temperature conductivity with the Fermi-Dirac probability density function and the spectral density function for random impurities to obtain the full conductivity. After carrying out the convolutions we set $E=\mu$.

In GaAs and MOS devices, terminal voltages are usually proportional to the electron density and have a more complicated relationship to the Fermi energy or chemical we can now plot the conductivity versus density parametrically with the chemical potential at finite temperature and in the presence of elastic and inelastic scattering. Such a graph should closely resemble the device current-voltage characteristics, and will be presented in detail in Sec. VI.

As a crude observability criteria for quantum effects in the device current-voltage characteristics, it is reasonable to assume that if the thermodynamic density of states $N(E,k_BT,\hbar/2\tau)$ has its structure (whose characteristic energy scale is ΔE) smeared out, then the conductivity will also lose its structure. For the structure to remain in $N(E,k_BT,\hbar/2\tau)$ after convolution with the temperature broadening function we must have $\Delta E>k_BT$. For the structure to remain in the presence of elastic and inelastic scattering and be independent of the exact arrangement of impurities in the device we must demand that $L>L_\phi$ and $\Delta E>\hbar/\tau$, leading to the observability criteria of Ref. 16.

V. EFFECT OF DIMENSIONALITY

An electron in a superlattice potential with periodicity a along the x direction, namely $V(x,y,z)=V(x)=V(x+a)$, can form standing waves when an integer multiple of its wavelength is twice the superlattice period. Since the periodic potential depends only on the x coordinate, the total effect of dimensionality appears in the kientic energy term of the Hamiltonian H, where

$$H=-\frac{\hbar^2}{2m^*}\sum_{i=1}^{d}\frac{d^2}{dx_i^2}+V(x)\ . \quad (27)$$

Here m^* is the effective mass of the crystal before the superlattice potential is applied.

We must solve the Schrödinger equation[14]

$$H\psi_{\mathbf{k},N}(x,y,z)=E_N(\mathbf{k})\psi_{\mathbf{k},N}(x,y,z) \quad (28)$$

for the energy eigenvalues of this problem. The wave functions describing free-electron motion perpendicular to the superlattice along the y and z directions are plane waves, so we need only solve the motion along the x direction. This motion we calculate using a Fourier decomposition of the potential and the central equation, as in standard band-structure calculations.[14] The eigenvectors and eigenvalues are

$$\psi_{\mathbf{k},N}(x,y,z)=\psi_{k_x,N}(x)e^{ik_yy}e^{ik_zz}\ , \quad (29)$$

$$E_N(k_x,k_y,k_z)=E_{N,x}(k_x)+\frac{\hbar^2k_y^2}{2m^*}+\frac{\hbar^2k_z^2}{2m^*}=E_{N,x}(k_x)+E_y(k_y)+E_z(k_z)\ , \quad (30)$$

where $\psi_{k_x,N}(x)$ and $E_{N,x}(k_x)$ are the solution to the one-dimensional Schrödinger equation in a periodic potential $V(x)$. The one-dimensional density of states along the x direction $N_x(E_x)$ can be obtained from

$$N_x(E_x)=\frac{2}{\pi}\frac{1}{|dE_x(k_x)/dk_x|}\ , \quad (31)$$

calculated as in the previous section.

Consider now the case of two dimensions at zero temperature. To find the total number of electrons below the Fermi energy in this two-dimensional **k** space, we must integrate the density of electron states up to E_F, namely

$$n_{2D}(E_F)=\int_0^{k(E_F)} 2\frac{d^2k}{(2\pi)^2}=\int_{-\infty}^{E_F} N_{2D}(E)dE \ . \quad (32)$$

We have, therefore, an expression for the density of states:

$$N_{2D}(E)=\int \delta(E-E_N(k))2\frac{d^2k}{(2\pi)^2}$$
$$=\frac{1}{2}\int \delta(E-E_{N,x}(k_x)-E_y(k_y))2\frac{dk_x}{2\pi}2\frac{dk_y}{2\pi} \ . \quad (33)$$

We can express this in terms of the one-dimensional density of states, recognizing the one dimensional identities

$$2\frac{dk_x}{2\pi}=N_x(E_x)dE_x \quad (34)$$

and

$$2\frac{dk_y}{2\pi}=N_y(E_y)dE_y \ , \quad (35)$$

where $N_x(E_x)$ and $N_y(E_y)$ are the 1D density of states including spin. Thus

$$N_{2D}(E)=\tfrac{1}{2}\int \delta(E-E_x-E_y)N_x(E_x)dE_xN_y(E_y)dE_y \ . \quad (36)$$

Integrating over the δ function leads to our final result

$$N_{2D}(E)=\tfrac{1}{2}\int N_y(E-E_x)N_x(E_x)dE_x$$
$$=\tfrac{1}{2}N_x(E)\otimes N_y(E) \ . \quad (37)$$

Our result, Eq. (37), is a very intuitive one. Electron motion in the x direction is independent of motion along y, except for the constraint that the total energy $E=E_{N,x}(k_x)+E_y(k_y)$ is a constant. Thus for some density of available states along x with a contribution $E_{N,x}$ to the total energy E, the total number of available states at a fixed energy $E=E_{N,x}+E_y$ is simply the product $N_x(E_x)N_y(E-E_x)$. To find the total number of available states at an energy E, we must sum over all the states E_x obeying the energy conservation condition as in (37). The factor 1/2 in this convolution compensates for overcounting electron spin states in the one-dimension state densities. $N_y(E_y)$ is the free-electron density of states for motion perpendicular to the superlattice given by

$$N_y(E_y)=\frac{2}{\pi}\frac{1}{|dE_y(k_y)/dk_y|}$$
$$=\frac{1}{\pi}\left[\frac{2m^*}{\hbar^2}\right]^{1/2}\frac{1}{\sqrt{E_y}}\Theta(E_y)$$
$$\equiv N_{1D}^{\text{free}}(E_y) \ , \quad (38)$$

ergy spectrum of the one-dimensional superlattice are removed by convolution with $N_y(E)$, since $N_y(E)$ has a finite value for all $E\geq 0$. It is also important to note that, since it is impossible to calculate the mean or variance of Eq. (38), the convolutions used to incorporate free electron motion have no natural energy scale associated with them.

We can also find an expression for the conductivity in two dimensions at zero temperature given the results in one dimension, analogous to Eq. (37). Assuming a constant scattering time, Eq. (24) for the conductivity can be rewritten using arguments similar to the ones above as

$$\sigma_{2D}^{xx}(E)=e^2\tau\int \delta(E-E_x-E_y)\left[\frac{1}{\hbar}\frac{\partial E_x}{\partial k_x}(E_x)\right]$$
$$\times\left[\frac{1}{\hbar}\frac{\partial E_x}{\partial k_x}(E_x)\right]\frac{1}{2}N_x(E_x)dE_x$$
$$\times N_y(E_y)dE_y \ , \quad (39)$$

to give the result analogous to Eq. (37)

$$\sigma_{2D}^{xx}(E)=\tfrac{1}{2}\int \sigma_{1D}^{xx}(E_x)N_y(E-E_x)dE_x$$
$$=\tfrac{1}{2}\sigma_{1D}^{xx}(E)\otimes N_y(E) \ . \quad (40)$$

If the scattering time τ varies with Fermi energy, we first carry out the convolution assuming it is a constant, then scale the conductivity by the variable scattering time afterwards. The tensor component $\sigma_{2D}^{yy}(E)$ can be easily calculated using the convolution

$$\sigma_{2D}^{yy}(E)=\tfrac{1}{2}\sigma_{1D}^{yy}(E)\otimes N_x(E) \ , \quad (41)$$

where $\sigma_{1D}^{yy}(E)$ is the free-electron conductivity in one dimension for constant scattering time given by

$$\sigma_{1D}^{yy}(E)=e^2\tau v_y^2(E)N_y(E)=\frac{e^2\tau}{m^*}\frac{2}{\pi}\left[\frac{2m^*E}{\hbar^2}\right]^{1/2} \ . \quad (42)$$

The off-diagonal tensor component $\sigma_{2D}^{xy}(E)$ is identically zero as can be confirmed from Eq. (24). This scheme is easily generalized to three dimensions by convolving Eqs. (37), (40), and (41) with one-half of the one-dimensional free-electron density of states, $N_{1D}^{\text{free}}(E)/2$, to account for additional free-electron motion in the z direction.

Although we originally thought of $V(x)$ as a periodic potential, we could just as easily have taken $V(x)=0$ for a free electron. We then obtain well-known analytic results for the free-electron density of states in two and three dimensions, given the free-electron density of states in one dimension $N_{1D}^{\text{free}}(E)$. Using the identities

$$\frac{1}{\sqrt{E}}\Theta(E)\otimes\frac{1}{\sqrt{E}}\Theta(E)=\pi\Theta(E) \quad (43)$$

and

$$\frac{1}{\sqrt{E}}\Theta(E)\otimes\frac{1}{\sqrt{E}}\Theta(E)\otimes\frac{1}{\sqrt{E}}\Theta(E)=2\pi\sqrt{E}\,\Theta(E) \ , \quad (44)$$

it is possible to obtain the standard free-electron density

$$N_{2D}^{free}(E)=\tfrac{1}{2}N_{1D}^{free}(E)\otimes N_{1D}^{free}(E)=\frac{m^*}{\pi\hbar^2}\Theta(E) \quad (45)$$

and

$$N_{3D}^{free}(E)=\tfrac{1}{2}N_{2D}^{free}(E)\otimes N_{1D}^{free}(E)$$
$$=\frac{1}{2\pi^2}\left[\frac{2m^*}{\hbar^2}\right]^{3/2}\sqrt{E}\,\Theta(E)\,, \quad (46)$$

where $N_{1D}^{free}(E)$ is defined in Eq. (38). Similar relationships hold for the free-electron conductivity in one, two, and three dimensions assuming a constant scattering time.

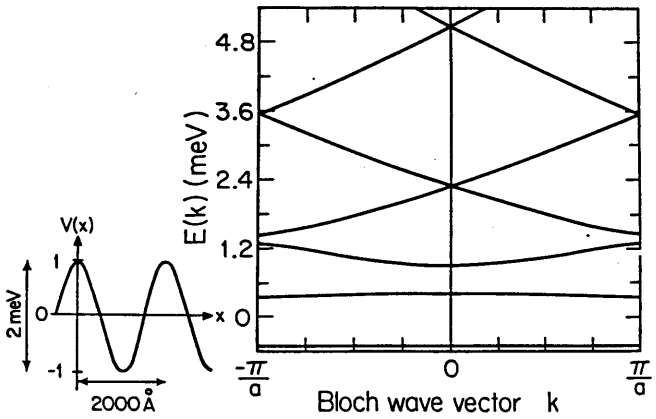


FIG. 3. Enery bands for $V(x)=(1\text{ meV})\cos(2\pi x/a)$ with $a=2000$ Å and a mass of $m^*=0.067m_0$.

VI. CALCULATED RESULTS

The main motivation behind this work is to understand the current-voltage characteristics of the electron devices studied in Ref. 1, in which a grating is incorporated as a Schottky gate onto a standard GaAs MODFET structure. We show this geometry in Fig. 2, which is adapted from Fig. 1 of Ref. 1. Electrons move from source to drain, which we choose to be the x direction as in Fig. 2. The carriers form a quasi-two-dimensional electron gas modulated by the periodic potential imposed from the gate. The average electron density is controlled from a backside contact. The x direction would correspond to motion normal to the grown layer in standard compositional superlattices. The plane of the grown layer would correspond to the y-z plane in standard compositional superlattices.

Following Ref. 1, we choose $V(x)=V_0\cos(2\pi x/a)$ with $V_0=1$ meV, $m^*=0.067m_0$, and $a=2000$ Å. All conductivity calculations we present assume this potential except for Fig. 6. This is a conservative approximation for the potential in these devices. We show the energy-band structure for this potential in Fig. 3. Notice that this potential, while still relatively weak for a GaAs device in which $V_0=250$ meV is possible,[1] is already well outside the limit of first-order degenerate perturbation theory.[14] We expect in first-order degenerate perturbation theory to have only one energy gap opening at the first spatial harmonic of the potential, yet significant energy gaps are already opening at the second and third spatial harmonic of the potential. It is also interesting that the energy gap at the third harmonic is above the top of the cosine potential. For stronger cosine potentials almost all the energy gaps lie below the top of the cosine potential. This will not be true if the potential has higher spatial Fourier components, such as a square-wave potential.

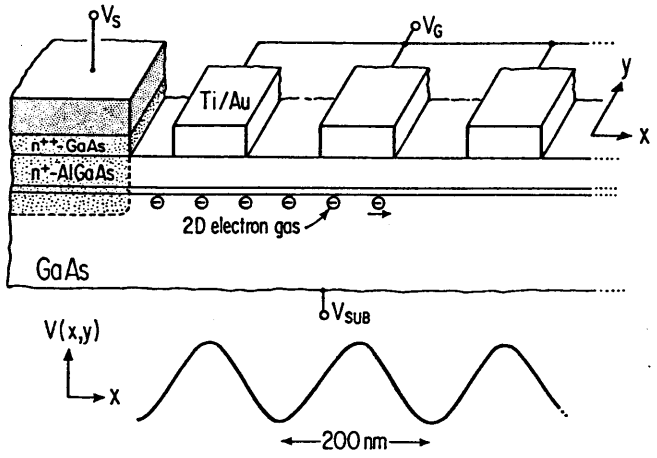


FIG. 2. Electrons moving in a two-dimensional gas subject to an artificially imposed periodic potential from a grating gate.

Figure 4(a) shows the one-dimensional thermodynamic density of states versus energy calculated for the same potential. The lowest energy band, being almost flat, contributes a δ-function-like feature to the density of states. At higher energies the curve behaves asymptotically like the one-dimensional free-electron density of states, having an inverse square root of energy dependence. The thermal broadening full width at half maximum for the temperatures 50 mK, 0.6 K, and 1.2 K is drawn on the figure for comparison with the energy gaps. As can be verified from the figure, the lowest temperature is comparable in size to the fourth energy gap and allows the lower energy gaps to be well resolved, while the highest temperature is roughly comparable in size to the second energy gap.[17]

Disorder broadening qualitatively resembles thermal broadening. The broadening of the density of states in a one-dimensional disordered superlattice has been calculated for a given arrangement of disorder in Ref. 18. From the figures in Ref. 18, one can easily see how the density of states over an ensemble average of many different configurations of the disorder can be calculated via convolution with the impurity spectral function.

We now consider the conductivity in one spatial di-

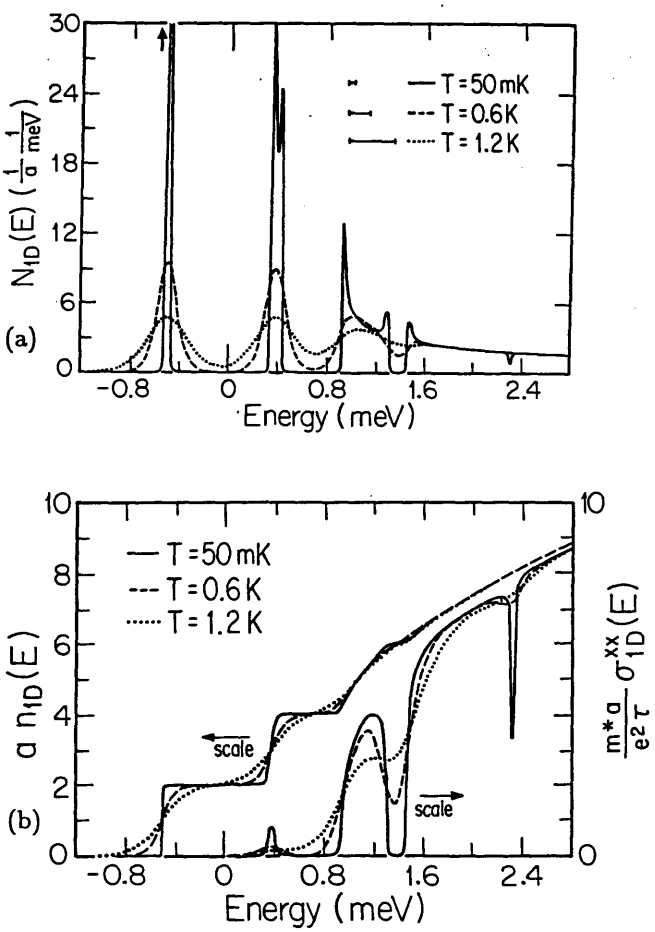


FIG. 4. (a) Thermodynamic density of states $N_{1D}(E,k_BT)$ vs energy in one dimension. The bars beside each temperature indicate the scale of the energy averaging. (b) Conductivity $\sigma^{xx}_{1D}(E,k_BT)$ assuming a constant scattering time and electron density $n_{1D}(E,k_BT)$ vs energy in one dimension.

time. When the Fermi energy becomes much larger than the superlattice potential, we expect the conductivity and electron density to be proportional from the free-electron conductivity formula $\sigma^{free}=ne^2\tau/m^*$. Figure 4(b) confirms this is true in one spatial dimension, showing the conductivity versus Fermi energy for a constant scattering time as well as the density versus Fermi energy on the same graph. These two curves approach each other at high energies. The lowest energy band has a negligible group velocity and makes practically no contribution to the conductivity. The conductivity also falls to zero when the Fermi energy lies in an energy gap at zero temperature. The conductivity modulation is solely a result of the Fermi group velocity tending to zero at a Brillouin zone boundary ($k=\pm\pi/a$) and at the zone center ($k=0$). Also of interest is the structure in conductivity associated with the fourth spatial harmonic of the potential. Despite the fourth energy gap being negligible compared with k_BT, the structure remains in conductivity at $T=0.6$ K. This is because the spacing between structure in the conductivity versus Fermi energy is much larger than the energy gap for this feature. The electron group velocity, whose maximum is associated with an inflection point in the energy band diagram, tends to zero before reaching the band gap. Thus it is the inflection points in $E_N(k)$ which determine the characteristic energy scale for the conductivity and not the band gaps. At lower energies the structure in conductivity more closely parallels the band gaps, but can never be smaller than them. The band gaps are thus an overly conservative estimate for the characteristic energy scale in conductivity. The electron density in Fig. 4(b) first undergoes a series of steps corresponding to the flat bands. Then the density and conductivity asymptotically approach a square root of energy at higher energies typical of free electrons. Another interesting feature in the electron density versus Fermi energy is that the lowest two energy bands, which are almost flat, contribute a shape which is the mirror image of a Fermi distribution function.

The final goal of our procedure is to plot conductivity versus electron density, since it should closely resemble the actual current-voltage relationship of a superlattice field effect transistor. We graph conductivity versus electron density in one dimension again assuming a constant scattering time in Fig. 5. Since the conductivity in a periodic potential in one dimension asymptotically approaches the free-electron conductibity at large densities, the free-electron conductivity is shown for reference with unit slope in this figure. Again note that the lowest enerby band makes a negligible contribution to the conductivity due to its small Fermi group velocity. It is interesting to compare the conductivity versus density for this relatively weak cosine potential, having $V_0=1$ meV, with a stronger potential. In Fig. 6 we show the conductivity versus electron density for a stronger cosine potential having $V_0=25$ meV but with the same period and effective mass. The scattering time is again assumed to be constant. Notice how rapidly the conductivity drops to zero, compared to the free-electron case, once the Fermi energy falls below the top of the strong cosine potential. For this stronger cosine potential, we can see from

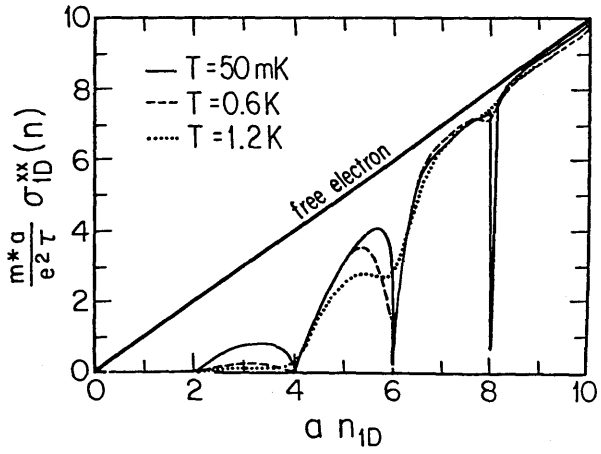


FIG. 5. Conductivity $\sigma^{xx}_{1D}(n)$ versus electron density in one dimension assuming a constant scattering time. The one-dimensional free-electron conductivity having unit slope is shown for reference.

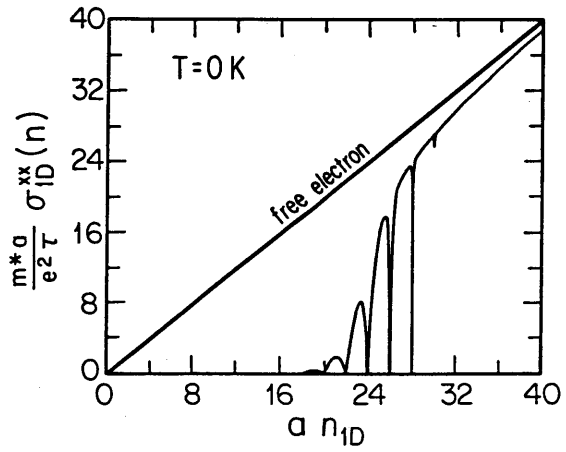


FIG. 6. Conductivity $\sigma_{1D}^{xx}(n)$ vs electron density in one dimension for the case of a strong potential having $V_0=25$ meV and assuming a constant scattering time. The free-electron conductivity is shown having unit slope.

Fig. 6 that the first nine energy bands make a negligible contributon to the conductivity. This is because the Fermi group velocity for energy levels bound deeply in the cosine well is nearly zero. These levels can be important in optical experiments, but not in transport parallel to the superlattice.

Matching the cosine well to a harmonic oscillator potential shows us that the energy spacings of these strongly bound energy levels are approximately

$$\Delta E \simeq 2\pi\sqrt{\hbar^2 V_0/m^* a^2} \ . \tag{47}$$

This ratio should determine how varying the parameters m^*, a, and V_0 improve the observability of quantum oscillations in the conductivity versus density of the device. For other shapes of the periodic potential, the characteristic energy scale ΔE would be different and would lead to different relations analogous to Eq. (47). In rough agreement with the energy spacings for a cosine potential, Eq. (47), the structure in conductivity between Figs. 5 and 6 disappears at a temperature around $\sqrt{25}=5$ times higher for the stronger potential, or about 10 K, though this is not shown in the figure. Varying the parameters m^*, a, and V_0 affect the observability of quantum oscillations in general as follows: It is easy to show from the differential equation of motion in the x direction for any periodic potential $V(x)=V_0 h(x)$, where $h(x)=h(x+a)$ that, for any value of the product $m^* a^2 V_0=$const, the normalized eigenenergies $E_{N,x}(k_x a)/V_0$ or $m^* a^2 E_{N,x}(k_x a)$ are also a constant. Thus, the solution of a single one-dimensional problem also gives the solution for a family of values of the parameters m^*, a^2, and V_0.[19] Equation (47) for a cosine potential is consistent with this general theorem.

The qualitative structure of conductivity versus density changes once we add free-electron motion perpendicular to the superlattice. The additional y degree of freedom in the electron motion will smear the structure in the super-

tion with the one-dimensional free-electron density of states, resulting in a weaker modulation of the conductivity. However, since the one-dimensional free-electron density of states in Eq. (38) has no well-defined variance, it is not possible to assign a well-defined observability criterion as to which spatial dimension a feature in the conductivity will become extinct. This observation is motivated by analogy with the convolutions used to incorporate thermal broadening and disorder broadening. Equation (7) for thermal broadening has a variance of approximately $k_B T$ and Eq. (1) for disorder broadening has a variance of approximately $\hbar/\tau$. If these variances are comparable with some characteristic energy scale in the conductivity, then structure in the conductivity is no longer observable at that particular temperature or electron mobility. However, it is not possible to calculate the variance of the free-electron density of states in Eq. (38). Even though disorder, temperature, and increasing the dimension of space all enter through convolutions with known functions, the convolution used to increase the spatial dimension is different and cannot be thought of as broadening the conductivity. If Eq. (38) had a well-defined variance, then it would be possible to say, for example, that a certain feature in the conductivity should be observable in one dimension but not two, just as we say that a certain feature in the conductivity should be observable at a temperature of 1 K but not at 50 K.

The two-dimensional density of states is shown in Fig. 7(a). We obtained it by convolving Fig. 4(a) with the free-electron density of states in Eq. (38). Notice the inverse square root of energy singularities typical of a quasi-one-dimensional system eventually approach a constant density of states characteristic of free electons in two dimensions at high energy. This is the correct asymptotic behavior in two dimensions as given by Eq. (45). The feature associated with the fourth energy gap, while visible in the one-dimensional density of states, is no longer visible in two dimensions.

Figure 7(b) displays the conductivity and electron density versus energy in two dimensions, again for a constant scattering time. The conductivity $\sigma_{2D}^{xx}(E)$ was obtained from $\sigma_{1D}^{xx}(E)$ in Fig. 4(b) by convolution. Conductivity versus Fermi energy in two dimensions has been previously explored numerically.[20] $\sigma_{2D}^{xx}(E)$ no longer falls to zero, since electron motion parallel to the superlattice is forbidden only for a given k_x, not a given E. For any given energy, it is always posible to choose many k_y's such that k_x does not match the Bragg reflection condition. The current is finite because there are no true gaps in the energy spectrum. Notice the conductivity has an inverse square root of energy behavior for energies lying in the forbidden energy gaps in one dimension. This is characteristic of the convolution of broadened δ-function-like features in the one-dimensional conductivity with the one-dimensional free-electron density of states. An inverse square root function convolved with a smeared δ function looks like a slightly smeared inverse square root of energy. The feature associated with the fourth energy gap, a strong feature in the one-dimensional conductivity σ^{xx} is greatly reduced in the

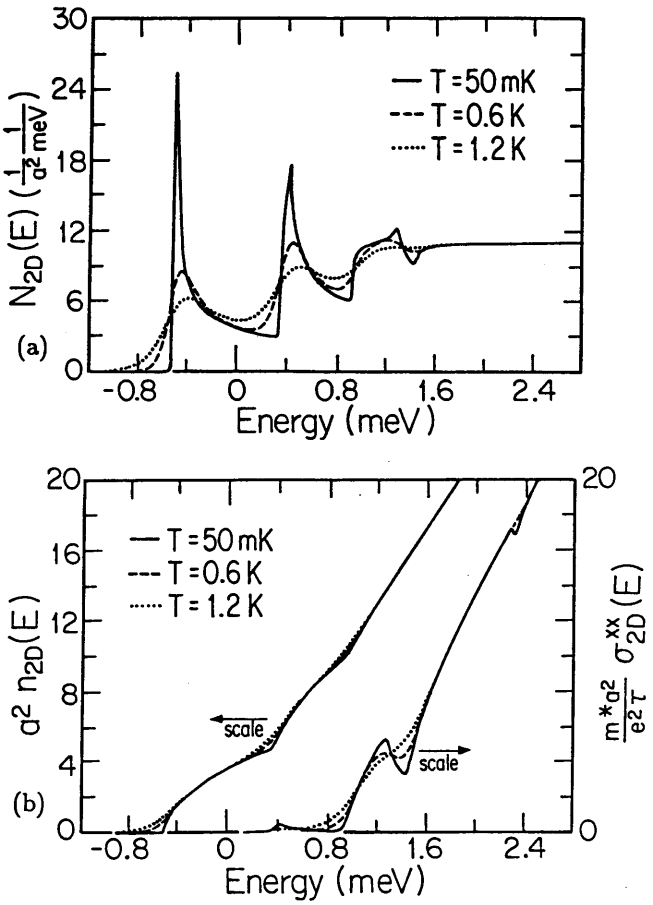


FIG. 7. (a) Thermodynamic density of states $N_{2D}(E,k_BT)$ vs energy in two dimensions. (b) Conductivity $\sigma_{2D}^{xx}(E,k_BT)$ assuming a constant scattering time and electron density $n_{2D}(E,k_BT)$ vs energy in two dimensions.

ty and electron density versus energy have a linear dependence on energy in the limit of large energies.

The density versus energy in Fig. 7(b) can also be interpreted as σ_{2D}^{yy} at constant scattering time. The strongly bound energy levels, which carry virtually no current parallel to the superlattice, approximate quite well a quasi-one-dimensional conductor for transport perpendicular to the superlattice. This is easily seen in Fig. 7(b). In the first energy band σ_{2D}^{yy} is nonzero while σ_{2D}^{xx} is virtually zero. Convolution of σ_{1D}^{yy} with the one-dimensional superlattice density of states, $N_x(E)$ in Eq. (41), determines how closely σ_{2D}^{yy} approximates a quasi-one-dimensional conductor. In the limit that $N_x(E)$ is a sum of δ functions, the approximation is exact. This will be roughly true in the strongly bound energy levels, leading the first two energy bands in Fig. 7(b) to display approximately a square root of energy dependence as in Eq. (42). As the Fermi level sweeps above the top of the cosine potential, there is a transition from quasi-one-dimensional to two-dimensional behavior in the conductivity and density of states.

Again, the goal of our procedure is to find the conduc-

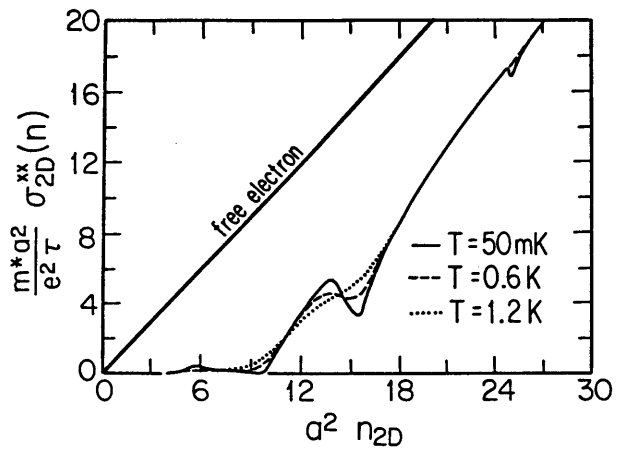


FIG. 8. Conductivity $\sigma_{2D}^{xx}(n)$ vs electron density in two dimensions assuming a constant scattering time. The free-electron conductivity in two dimensions is shown for reference with unit slope.

closely resemble the actual current-voltage characteristics of the superlattice field-effect transistor in Ref. 1. Figure 8 shows the conductivity σ_{2D}^{xx} versus density for constant scattering time in two dimensions, with the free-electron conductivity or σ_{2D}^{yy} plotted on the same graph and normalized to unit slope. Note that $\sigma_{2D}^{yy}(n)$ shows no structure when plotted versus electron density if the scattering time is a constant. The conductivity σ_{2D}^{xx} again asymptotically approaches the free-electron value for large densities, though this cannot be seen in the figure.

Consider now the three-dimensional case. The density of states versus energy in Fig. 9(a) shows a series of steps characteristic of quasi-two-dimensional behavior asymptotically approaching the square root of energy dependence of the free-electron density of states in three dimensions given in Eq. (46). The conductivity σ_{3D}^{xx} versus energy and electron density versus energy follow in Fig. 9(b). Their asymptotic behavior is proportional to the three halves power of energy. The plateaulike features in σ_{3D}^{xx} when the energy lies in the forbidden gaps in one dimension arise from the step function in Eq. (45) being convolved with the smeared δ function-like features in σ_{1D}^{xx} to generate σ_{3D}^{xx}. A step convolved with a smeared δ function looks like a slightly smeared step. The weak feature in σ_{2D}^{xx} associated with the fourth spatial harmonic of the periodic potential is no longer visible in three dimensions. The electron density versus Fermi energy, shown also in Fig. 9(b), again closely approximates σ_{3D}^{yy} for constant scattering time.

In contrast to the one- and two-dimensional cases, the conductivity σ_{3D}^{xx} will not asymptotically approach the free-electron value at high energies, but will asymptotically approach the free-electron value minus a constant $\Delta\sigma_{3D}^{xx}$, determined by

$$\Delta\sigma_{3D}^{xx}=\lim_{E\to\infty}\left[[\sigma_{1D}^{\text{free}}(E)-\sigma_{1D}^{xx}(E)]\otimes\frac{1}{2}\frac{m^*}{\pi\hbar^2}\Theta(E)\right]$$
$$\neq 0\ . \qquad (48)$$

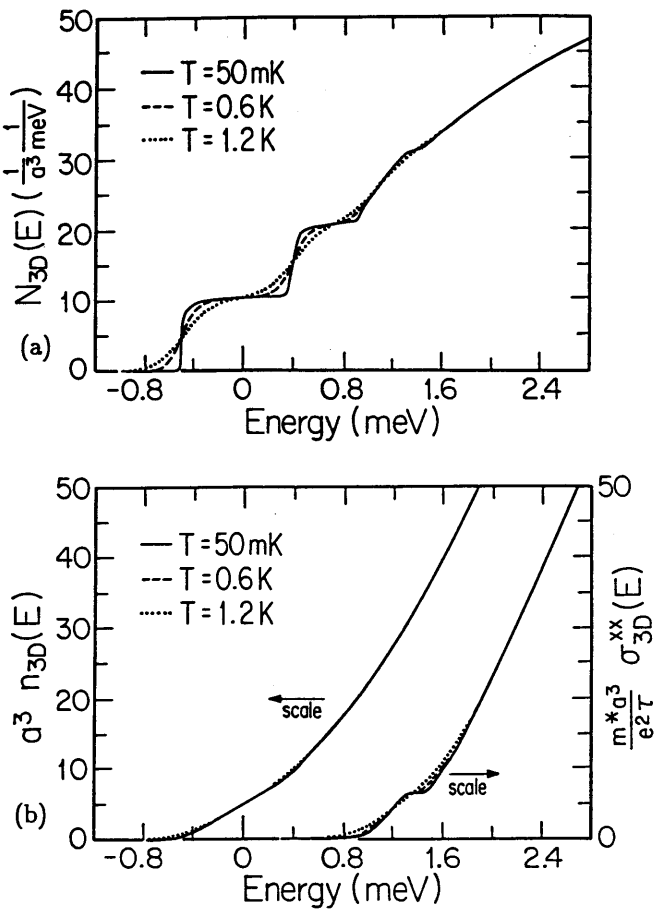


FIG. 9. (a) Three-dimensional thermodynamic density of states $N_{3D}(E,k_BT)$ vs energy. (b) Three-dimensional conductivity $\sigma_{3D}^{xx}(E,k_BT)$ assuming a constant scattering time and electron density $n_{3D}(E,k_BT)$ vs energy.

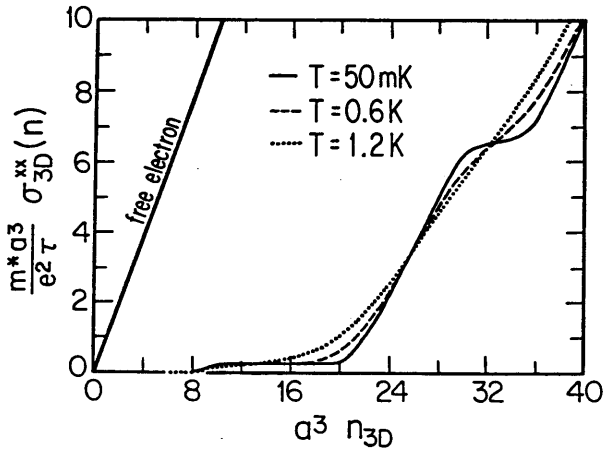


FIG. 10. Three-dimensional conductivity $\sigma_{3D}^{xx}(n)$ vs electron density assuming a constant scattering time. The free-electron conductivity in three dimensions is shown having a unit slope.

proaches the free-electron conductivity at high energy. Two dimensions is the marginal case, where the conductivity in a superlattice approaches the free-electron conductivity very slowly because the free-electron density of states in one dimension decays slowly. The two-dimensional free-electron density of states, however, does not decay at high energies. Consequently, any residual area remaining after subtracting $\sigma_{1D}^{free}(E)-\sigma_{1D}^{xx}(E)$ will not decay as $E\rightarrow\infty$ after performing the convolution in Eq. (48). The conductivity σ_{3D}^{xx} versus density in three dimensions is shown in Fig. 10 with the free-electron conductivity, or equivalently $\sigma_{3D}^{yy}=\sigma_{3D}^{zz}$, also shown with a unit slope for comparison. The features in conductivity σ_{3D}^{xx} parallel to the superlattice as a function of density are plateaulike.

The previous conductivity calculations all assume a constant scattering time. We have also examined the case of an energy-dependent scattering time $\tau^{-1}(E)\sim N(E)$, as in Eqs. (4) and (5).[21] The conductivity tensor component σ_{2D}^{yy} plotted versus density is a straight line if the scattering time is a constant independent of more striking behavior if we consider the scattering time to vary as a function of Fermi energy as shown in Fig. 11. The electron density of states versus energy for this potential has already been shown in Fig. 7(a). σ_{2D}^{yy} shows a minimum whenever the density of states becomes large, since the scattering rate in Fermi's golden rule becomes very high. This intersubband scattering is the dominant mechanism for Drude conductivity oscillations in quasi-one-dimensional wires, as was pointed out in a self-consistent Green's function calculation of the conductivity in narrow wires Das Sarma and Xie,[22] and by Kearney and Butcher.[23] Figure 12 shows σ_{2D}^{yy} plotted along with the free-electron conductivity as a function of electron density, again with a variable scattering time given by Fermi's golden rule. Figure 12 should closely resemble

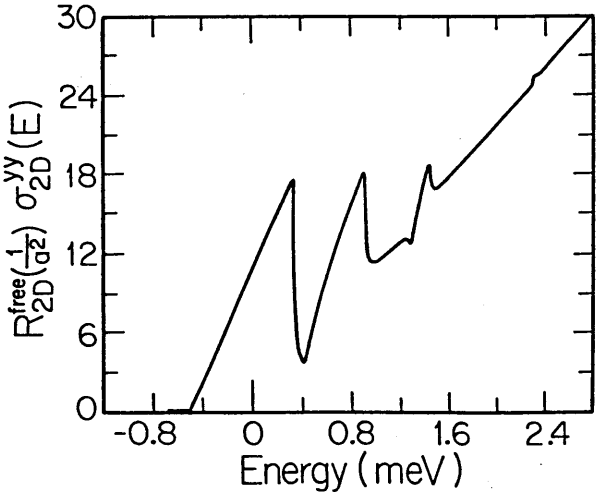


FIG. 11. Conductivity $\sigma_{2D}^{yy}(E,T=50\text{ mK})$ vs Fermi energy with a variable scattering time given by the Fermi golden rule. The drop in conductivity as the Fermi level enters a new subband is due to intersubband scattering. The curve is not

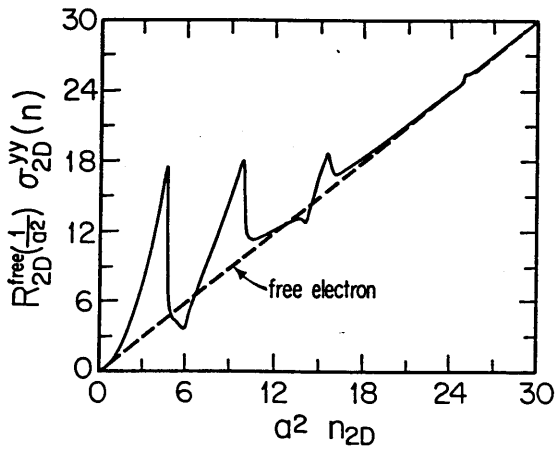


FIG. 12. Conductivity σ_{2D}^{yy} vs electron density with a variable scattering time given by the Fermi golden rule at $T=50$ mK. The free-electron conductivity in two dimensions with a variable scattering time is the straight line with unit slope shown on this graph.

the I-V characteristic of a quasi-one-dimensional wire or quasi-one-dimensional grating gate transistor. This conductivity can be broadened with either a constant or a variable scattering time, with the qualitative effect being similar to thermal broadening. The differential mobility $d\sigma_{2D}^{yy}/dn$, is higher than $d\sigma_{2D}^{free}/dn$ when the density of states available to scatter into is less than the free two-dimensional density of states $N_{2D}<N_{2D}^{free}$, and vice versa.

Sakaki[24] observes that all physical scattering centers have a finite spatial extent, and that these finite-range scatterers cannot deflect electrons as effectively in a quasi-one-dimensional wire as they can in two and three dimensions. This leads to an additional suppression of all scattering in the wire. But Sakaki's argument will not drastically change the shape of $\sigma_{2D}^{yy}(n)$ in Fig. 12.

To obtain quantitative values for the electrical conductivity in two dimensions, it is necessary to extract the magnitude of the impurity scattering length $n_{imp}|V|^2$, which has the dimensions of energy squared times area, from the mobility of a standard quasi-two-dimensional field effect transistor. Equation (4) implies for the standard field-effect transistor that

$$n_{imp}|V|^2=\frac{e\hbar^3}{2\mu(m^*)^2}\ . \tag{49}$$

Therefore, the resistance of a two-dimensional free-electron gas at a density of one electron per area a^2 is

$$1/R_{2D}^{free}(n=1/a^2)\equiv\sigma_{2D}^{free}(n=1/a^2)$$
$$=e\mu\frac{1}{a^2}=\frac{e^2\hbar^3}{2n_{imp}|V|^2(m^*)^2a^2}\ . \tag{50}$$

Consider now a GaAs MODFET with a mobility of 100 000 cm^2/V s and a mass of $m^*=0.067m^*$. We obtain from Eq. (49) $n_{imp}|V|^2=9800$ (meV Å)^2=2450 (μeVa)^2, where $a=2000$ Å. The two-dimensional free-electron density of states, obtained from either Eq.

meVa^2. Therefore, a typical energy level broadening obtained from Eq. (4) would be $\hbar/\tau=0.170$ meV. This broadening is comparable to a temperature of 2 K, so collision broadening will dominate below that temperature for a GaAs device having a mobility of 100 000 cm^2/V s. For this case we have $R_{2D}^{free}(n=1/a^2)=25\,000\ \Omega$. This resistance sets the scale of the conductance axis in Figs. 11 and 12.

It is also possible to obtain an analytic expression for the conductivity and density versus Fermi energy in quasi-one-dimensional wires with intersubband scattering.[10,11] The results look very much like the ones presented in Figs. 11 and 12 when the Fermi energy is much smaller than the periodic potential. Consider a quasi-one-dimensional wire with an energy dispersion $E_i(k)=E_i+\hbar^2k^2/2m^*$. Each subband in a quasi-one-dimensional wire is an additional channel to carry current; their conductivities add like resistors in parallel if we neglect interference between the subbands. Then

$$\sigma(E)=\sum_i\sigma_i(E)=e^2\sum_i v_i^2(E)N_i(E)\tau_i(E)\ , \tag{51}$$

where we label the sum over all occupied subbands with the index i. The density of states adds similarly

$$N(E)=\sum_i N_i(E)\ . \tag{52}$$

Allowing for both inter- and intrasubband scattering, our model for the mean free time in the quasi-one-dimensional wire is

$$\frac{\hbar}{\tau(E)}=\frac{\hbar}{\tau_i(E)}=2\pi n_{imp}|V|^2N(E)$$
$$=2\pi n_{imp}|V|^2\sum_j N_j(E)\ , \tag{53}$$

which is independent of the subband index i. Note that the scattering rate into each subband simply adds as in Matthiessen's rule, there is no wave interference between different scattering events. This form for the scattering rate gives rise to a conductivity

$$\sigma(E)=e^2\frac{\hbar}{2\pi n_{imp}|V|^2}\left[\frac{\hbar}{m^*}\right]^2$$
$$\times\left[\sum_j\left[\frac{\hbar^2}{2m(E-E_j)}\right]^{1/2}\right]^{-1}$$
$$\times\sum_i\left[\frac{2m(E-E_i)}{\hbar^2}\right]^{1/2}\ , \tag{54}$$

and an electron density

$$n(E)=\frac{2}{\pi}\sum_i\left[\frac{2m(E-E_i)}{\hbar^2}\right]^{1/2}\ , \tag{55}$$

where we again restrict the sum to include only occupied subbands. The dramatic increase in scattering rate as the higher subbands become populated overwhelms the modest increase in the additional number of low-velocity carriers, and makes the conductivity drop to zero when the Fermi level passes into a new subband in the wire. In ad-

for this drop in conductivity due to intersubband scattering in quasi-one-dimensional wires, Büttiker *et al.*[25] also argued for a drop in the conductivity in quasi-one-dimensional wires as a new subband becomes populated based on the Landauer formula.

Analogous to our discussion of intersubband scattering in the quasi-one-dimensional wire and σ^{yy}_{2D} for the periodic potential in two dimensions, the tensor component σ^{yy}_{3D} also shows an interesting structure if we allow the scattering time to vary with Fermi energy. The conductivity tensor component σ^{yy}_{3D} will behave like the conductivity in a quasi-two-dimensional plane, i.e., a thin film with normal subbands, for Fermi energies small compared to the periodic potential. The conductivity σ^{yy}_{3D} drops where the density of states shows a maximum due to the increase in scattering rate, exactly analogous to the conductivity tensor component σ^{yy}_{2D}. This has also been pointed out by Trivedi.[26] The relative conductivity modulation, however, is much weaker than in two dimensions. It is also possible to derive formulas analogous to Eqs. (54) and (55) above for conductivity in a quasi-two-dimensional plane.

We now describe the effect of a variable scattering time on σ^{xx}. While the asymptotic behavior of the conductivity at large energies or densities is different in one and three dimensions for the variable scattering time, the qualitative features in σ^{xx} are very similar to the case of constant τ. The variable scattering time has the effect of forcing the conductivity to fall even faster at a Brillouin zone boundary or zone center. This is because, in addition to the Fermi group velocity approaching zero at a Brillouin zone boundary, the scattering rate now becomes higher as well. The energy level broadening also becomes larger around a Brillouin zone boundary or zone center if the conductivity is averaged in energy using a variable τ.[21]

We now wish to give a qualitative argument for the density values at which features appear in the conductivity versus electron density in different spatial dimensions. We have computed the conductivity versus electron density n in the dimensionless units of na^d where $d=1,2,3$ is the dimension of space. If we had considered the case where the potential was periodic with periodicity a along the x, y, and z directions (which we did not do in this paper), and if the periodic potential were strong enough, then structure would appear in the conductivity at values $na^d=2,4,6,\ldots$ corresponding to putting one electron of each spin into a length a in one spatial dimension, a square of area a^2 in two dimensions, and a cube of volume a^3 in three dimensions, thereby filling up a Brillouin zone. In the case of a periodic potential with the same periodicity along all directions, there is a single characteristic distance in the problem. Consequently, structure appears at the same value of the normalized density na^d in all three dimensions.

Consider now the case which we have analyzed in this paper, namely, where there is a periodic potential with period a only along the x direction. In one spatial dimension, the density values at which structure appears are given by the above counting argument. Next, we added density of states is built up by convolution, given by Eq. (37), but there is a way to estimate these density values. The density of states for motion along the x direction in the lowest two energy bands is approximately two δ functions having an energy separation of $\Delta E \simeq 0.85$ meV. Because the density of states is a convolution, the change in electron density between features is approximately a product

$$\Delta n_{2D} a^2 \simeq \tfrac{1}{2}\Delta n_x a \Delta n_y a = \tfrac{1}{2} 2a \int_0^{\Delta E} N^{\text{free}}_{1D}(E_y) dE_y \ . \tag{56}$$

Numerically Eq. (56) gives $\Delta n_{2D} a^2 \simeq 5$, approximately the density spacing to the first feature in two dimensions in Figs. 8 and 12. The density spacings between other features can be estimated in a similar manner. Notice that the density spacings between feaures in two dimensions are constrained by the density of states for motion along x, meaning that they are constrained by the strength and shape of the periodic potential along x as well as by the free electron density of states. A stronger periodic potential corresponds to a wider spacing ΔE between allowed energy bands in the one-dimensional problem. Thus, a larger number of carriers are required before the second allowed energy band can be occupied in higher dimensions. The three-dimensional analog to Eq. (56) is

$$\Delta n_{3D} a^3 \simeq \tfrac{1}{2}\Delta n_x a \Delta n_{yz} a^2 = \tfrac{1}{2} 2a^2 \int_0^{\Delta E} N^{\text{free}}_{2D}(E_{yz}) dE_{yz} \ . \tag{57}$$

Numerically Eq. (57) gives $\Delta n_{3D} a^3 \simeq 9$, approximately the density spacing to the first feature in three dimensions in Fig. 10.

It is possible to extend our model for the conductivity in several ways. The scattering matrix element $\langle k|V(r)|k'\rangle$ in general is dependent on both k and k'. It is necessary to include this momentum dependence in the conductivity Eq. (24). Finally, nothing has been said about possible high field effects. We can include the effects of moderate electric fields by averaging the conductance and density of states over an energy range eV^*.[27] We implement this mathematically by convolving with the function $[\Theta(E+eV^*)-\Theta(E)]/eV^*$, where V^* is an appropriate voltage defined in Ref. 27, which describes the nonmonochromatic electron beam introduced by the electron reservoirs (contacts) being at different electrochemical potentials. Electrical conduction now occurs in a finite-energy window of width eV^*. Note that this is not a hot-electron effect, but an effect of energy averaging in the spirit of our previous treatment of temperature broadening and disorder broadening. The effects of higher fields may be even more interesting. It may be possible to observe a negative differential conductance in quasi one dimensional wires analogous to the intervalley transfer in bulk GaAs responsible for the Gunn oscillations at high electric fields. Consider electrons occupying only the lowest subband of a wire. An electric field may raise their energy to the point where a fraction of the electrons may scatter into the second subband. Their velocity in the second subband at the same energy

negative differential conductance should also be possible in a quasi-two-dimensional sheet from transfer of energetic electrons between subbands normal to the surface, i.e., in a standard GaAs MODFET transistor. Büttiker and Thomas[28] have pointed to the possibility that instabilities similar to the Gunn oscillations may form in superlattices due to their negative differential conductance.

Another important extension of our model is to truly two-dimensional periodic potentials. A much stronger modulation in σ_{2D}^{xx}, similar to the one-dimensional superlattice conductivity σ_{1D}^{xx}, is posible if the superlattice potential is periodic in both the x and y directions. This has been observed in the recent measurements of Ismail *et al.*[1] by comparing the I-V characteristics of grating versus grid gate devices. The features in conductivity in a periodic potential along both x and y directions should have roughly the same strength and extinguish at roughly the same temperatures as the one-dimensional conductivity in a periodic potential. To calculate this precisely requires a more computationally difficult method because the Hamiltonian of the system is no longer separable.

Finally, an important observability criteria for oscillations in σ^{xx} is that the electron coherence length limits the number of superlattice periods seen by the wave function. We need $L_\phi > a$ for our calculations to apply to real devices. Exactly how much their ratio must be cannot be determined by our method. A finite superlattice may also play some role in limiting the observability of quantum oscillations in superlattice devices, regardless of the coherence length. Horiguchi[19] has investigated both of these questions using the transmission coefficient in a periodic potential, tentatively concluding that not very many periods must be coherent before standing waves due to quantum interference can be observed in a superlattice and that the superlattice itself need not consist of very many periods.

VII. CONCLUSIONS

We have presented a semiclassical model for the conductivity of electrons in a periodic potential which incorporates elastic and inelastic scattering, finite temperature, and the effect of increasing spatial dimension. All three of these effects are incorporated into the model by first solving for the one-dimensional zero-temperature conductivity and density of states without scattering, then convolving with various known functions. The main effect of elastic and inelastic scattering is to average the zero-temperature conductivity and density of states over an energy range $\hbar/\tau$, where τ is the elastic mean free time. Inclusion of finite temperature averages the conductivity and density of states over an energy range $k_B T$. This energy averaging takes the mathematical form of convolving the conductivity and density of states with the impurity spectral function and the Fermi-Dirac probability function to include scattering and finite temperature, respectively. Since Schrödinger's equation can be separated for a superlatice potential along only one direction, increasing the dimensionality of space only involves convolving the electron density of states and conductivity with the one-dimensional free-electron density of states. This is both mathematically and conceptually much simpler than the standard method of integrating over the Fermi line in two dimensions, or Fermi surface in three dimensions, to obtain the density of states and conductivity in higher dimensions.

Many arguments have been made about the shape of conductivity in a quasi-one-dimensional wire or a periodic potential based on the electron density of states. Even though conductivity falls when the density of states goes through a maximum in either the periodic potential or quasi-one-dimensional wire, totally different physical mechanisms are responsible. In the quasi-one-dimensional wires, intersubband scattering is the mechanism responsible for the drop in conductivity as the Fermi energy enters a new subband. In a superlattice the Fermi group velocity falls to zero as the electrons form a standing wave with the superlattice. And for the conduction of a free-electron gas in a two-dimensional plane with a magnetic field applied perpendicular to the plane, just the opposite is the case.[29] The conductance goes through a maximum when the density of states reaches a maximum. The conclusion we must draw is that, in general, nothing can be said about the shape of conductivity in any electron device by considering solely the electron density of states. Conduction processes depend on many factors other than the electron state density.

Conductance versus electron density will closely resemble the actual current-voltage characteristics of field-effect transistors, so we emphasize the importance of obtaining $\sigma(n)$ in this paper. The terminal voltages of a field-effect transistor have a complicated functional relationship with Fermi energy, but in many cases the gate voltage of a field-effect transistor is directly proportional to the electron density. Also, semiclassical electron device simulators can always determine the average electron density, but may not be able to determine the Fermi energy in a complex device structure. If we wish to determine the actual current-voltage characteristic of a field-effect transistor using semiclassical device simulators, obtaining the conductance versus electron density is an imporant advantage.

ACKNOWLEDGMENTS

We thank D. A. Antoniadis, Henry I. Smith, and M. A. Kastner for many useful discussions. We also thank X. C. Xie, K. Ismail, A. Yen, W. Chu, C. L. Kane, and S. Horiguchi for contributing to our understanding of this subject. We gratefully acknowledge financial support from IBM, an MIT-EECS Departmental grant, and the support of U. S. Air Force Contract No. AFOSR-88-0304. We also acknowledge partial support from Joint Services Electronics Program Contract No. DAALO3-89-C-0001.

[1]K. Ismail, W. Chu, D. A. Antoniadis, and Henry I. Smith, Appl. Phys. Lett. **52**, 1071 (1988). Figure 1 of this paper shows the experimental geometry we are discussing in two dimensions. K. Ismail, W. Chu, D. A. Antoniadis, and Henry I. Smith, J. Vac. Sci. Technol. B **6**, 1824 (1988); K. Ismail, W. Chu, A. Yen, D. A. Antoniadis, and Henry I. Smith, Appl. Phys. Lett. **54**, 460 (1989); K. Ismail, W. Chu, D. A. Antoniadis, and Henry I. Smith, GaAs and Related Compounds (to be published), K. Ismail, Ph.D. thesis, Massachusetts Institute of Technology, 1989.

[2]Y. Tokura and K. Tsubaki, Appl. Phys. Lett. **51**, 1807 (1987).

[3]G. Bernstein and D. K. Ferry, J. Vac. Sci. Technol. B **5**, 964 (1987).

[4]A. C. Warren, D. A. Antoniadis, H. I. Smith, and J. Melngailis, IEEE Electron Device Lett. **EDL-6**, 294 (1985). See also A. C. Warren, I. Plotnik, E. H. Anderson, M. L. Schattenberg, D. A. Antoniadis, and Henry I. Smith, J. Vac. Sci. Technol. B **4**, 365 (1986).

[5]J. H. F. Scott-Thomas, M. A. Kastner, D. A. Antoniadis, Henry I. Smith, and Stuart Field, J. Vac. Sci. Technol. B **6**, 1841 (1988).

[6]K. Ismail, W. Chu, D. A. Antoniadis, and Henry I. Smith, Appl. Phys. Lett. **54**, 1130 (1989).

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

[8]P. A. Lee and A. D. Stone, Phys. Rev. Lett. **55**, 1622 (1985). See also B. L. Al'tshuler, Pis'ma Zh. Eksp. Teor. Fiz. **41**, 530 (1985) [JETP Lett. **41**, 648 (1985)].

[9]W. J. Skoçpol, P. M. Mankiewich, R. E. Howard, L. D. Jackel, D. M. Tennant, and A. D. Stone, Phys. Rev. Lett. **56**, 2865 (1986).

[10]P. F. Bagwell, D. A. Antoniadis, and T. P. Orlando, in *Advanced MOS Device Physics*, edited by N. Einspruch (Academic, San Diego, 1989).

[11]P. F. Bagwell, M.S. thesis, Massachusetts Institute of Technology, 1988.

[12]A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, *Methods of Quantum Field Theory in Statistical Physics* (Prentice-Hall, Englewood Cliffs, NJ, 1963); G. Rickayzen, *Green's Functions and Condensed Matter* (Academic, New York, 1980).

[13]O. Madelung, *Introduction to Solid State Theory* (Springer-Verlag, New York, 1978).

[14]C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1986). See also N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976).

[15]If the electron density is not proportional to a terminal voltage, we require a semiclassical solution to Poisson's equation in the device to determine this relationship.

[16]D. A. Antoniadis, A. C. Warren, and Henry I. Smith, in *International Electron Device Meeting Technical Digest, 1985* (Electrical and Electronic Engineers, Princeton, NJ, 1985), Vol. 85, p. 558.

[17]The energy gaps are 0.849, 0.493, 0.150, and 0.011 meV. The thermal energies $k_B T$ corresponding to the temperatures 50 mK, 0.6 K, and 1.2 K shown in the figure are 0.004, 0.052, and 0.103 meV, respectively.

[18]John D. Dow, Shang Yuan Ren, and Karl Hess, Phys. Rev. B **25**, 10 (1982).

[19]S. Horiguchi (private communication).

[20]M. J. Kelley, Surf. Sci. **170**, 49 (1986).

[21]Broadening using a variable scattering time can be done using equations such as

$$N(E,\hbar/2\tau(E))=\int \frac{1}{\pi}\frac{\hbar/2\tau(E')}{(E-E')^2+[\hbar/2\tau(E')]^2}N(E')dE' .$$

[22]S. DasSarma and X. C. Xie, Phys. Rev. B **35**, 9875 (1987).

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

[24]Hiroyuki Sakaki, Jpn. J. Appl. Phys. **19**, L735 (1980).

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

[26]N. Trivedi, Ph.D. thesis, Cornell University, 1987.

[27]Philip F. Bagwell and Terry P. Orlando, Phys. Rev. B **40**, 1456 (1989).

[28]M. Büttiker and H. Thomas, Z. Phys. B **33**, 275 (1979).

[29]H. Aoki, Rep. Prog. Phys. **50**, 655 (1987); **50**, 681 (1987).

Appendix C

Terry P. Orlando, Philip F. Bagwell, Khalid Ismail, and Ray Ghanbari, 'Quantum Device Modeling with the Convolution Method', in *Electronic Properties of Multilayers and Low-Dimensional Semiconductor Structures*, J.M. Chamberlain, L. Eaves, and J.C. Portal, eds., (Plenum, London, 1990).

To appear in: "Electronic Properties of Multilayers and Low-Dimensional Semiconductor Structures", J.M. Chamberlain, L. Eaves, and J.C. Portal, editors, (Plenum Press, London, 1990).

QUANTUM DEVICE MODELING WITH THE CONVOLUTION METHOD

T. P. Orlando, P. F. Bagwell, R. A. Ghanbari, and K. Ismail

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

1 Introduction

Recent advances in materials fabrication and nanolithography have made possible a generation of semiconducting structures whose conductance is governed by quantum mechanical phenomena. In particular, nanostructures on Si MOSFETs and GaAs MODFETs have shown modulations in their conductance versus gate voltage characteristics that have been attributed to quantum mechanical effects. In this paper, we review a modeling scheme which gives a unified way of understanding how these quantum effects are affected by temperature, mobility, voltage, and the structure of the device. This model provides not only a qualitative understanding of the various quantum phenomena, but also a basis for developing efficient computational algorithms for modeling specific devices. We have called this scheme the *convolution method* because most of the calculations can be written in terms of separate convolutions involving the individual phenomena of temperature, mobility, voltage, and structure.

In §2 specific quantum devices and their characteristics will be briefly described. The main part of the paper in §3 will present the convolution method and apply it to understanding the devices presented in §2.

2 Fabricated Devices

Quantum effects have been seen in the conductance of two principle types of semiconducting devices. In the first type, the electrons are electrostatically confined so that subbands are formed in one dimension, while the electrons remain free to move in the other dimension. The structure in the conductance of such quasi-one dimensional (Q1D) devices has been attributed to the filling of the Q1D subbands in both Si MOSFET and GaAs MODFET devices (Scott-Thomas et al., 1988; Ismail et al., 1989b; Warren et al, 1986). In the second type of structure, a periodic superlattice is imposed on the electrons in a two dimensional electron gas (2DEG). The modulation in the conductance of these devices is attributed mainly to Bragg diffraction from the potential induced by the lateral surface superlattice (LSSL) structure (Ismail

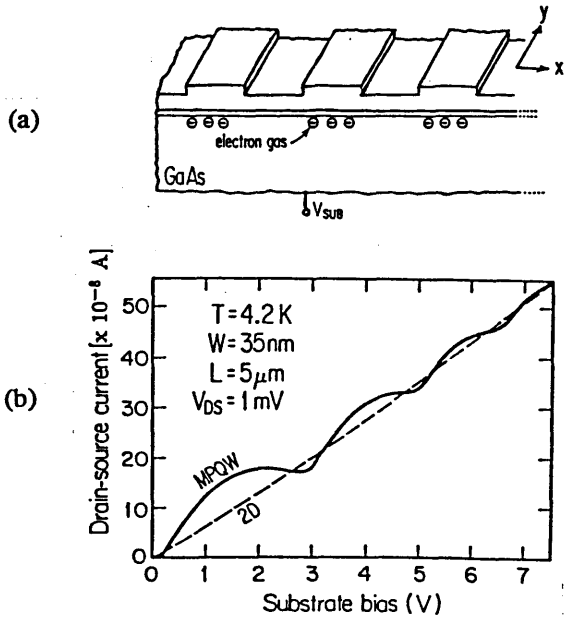


Figure 1: (a) Schematic of three multiple parallel quantum wires (MPQW) in a GaAs MODFET. (b) Drain-source current as a function of the substrate bias for a device with 100 parallel wires. The dashed line is for a scaled down 2D device. From Ismail et al. (1989b).

et al., 1988; Ismail et al., 1989a; Tokura and Tsubaki, 1987; Bernstein and Ferry, 1987; Warren et al., 1985). We now discuss three devices that illustrate the kinds of observed quantum effects.

Figure 1a depicts three parallel Q1D wires fabricated in a GaAs MODFET. The electrons are confined in the x-direction to a region about 40 nm wide, and are free to move in the y-direction. The actual device (Ismail et al., 1989b) consists of 100 parallel wires which were fabricated by ion milling a shallow grating into the doped AlGaAs layer through a mask produced by x-ray lithography. The period of the lines is 200 nm. The drain-source current along the y-direction is shown in Figure 1b as a function of back gate bias. As the back gate bias is increased the density of the electrons in the inversion layers increases, causing successive Q1D subbands to be occupied. The conductance decreases as electrons occupy each new subband. The structure in Figure 1b washes away with increased temperature and is less pronounced for devices with lower mobility.

The same x-ray mask used to make the Q1D conductors in Figure 1 can be rotated as shown in Figure 2a so that the electrons now flow perpendicular to the grating lines. In this device (Ismail et al., 1988) a 200 nm-period Schottky barrier grating gate of Ti/Au replaces the usual continuous gate in the MODFET configuration. As the gate-source potential is changed, the drain-source current along the x-direction is modulated as shown in Figure 2b. This modulation is believed to be caused by

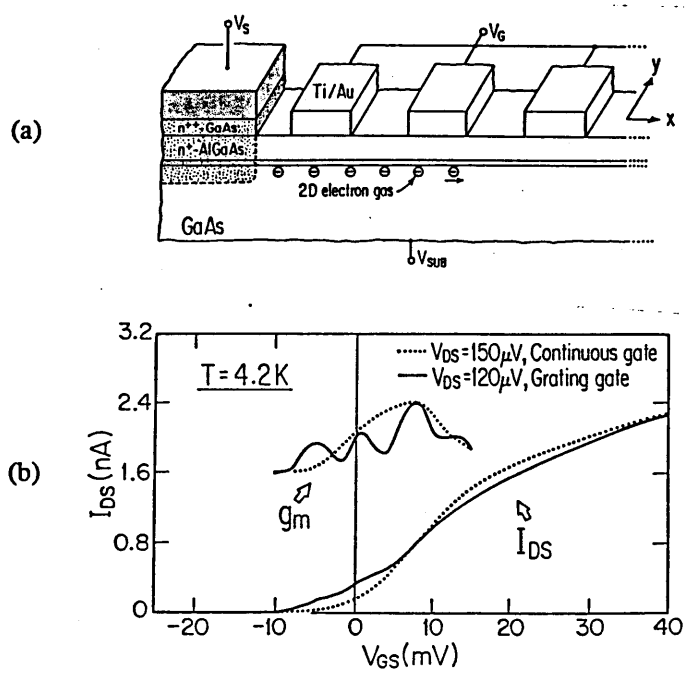


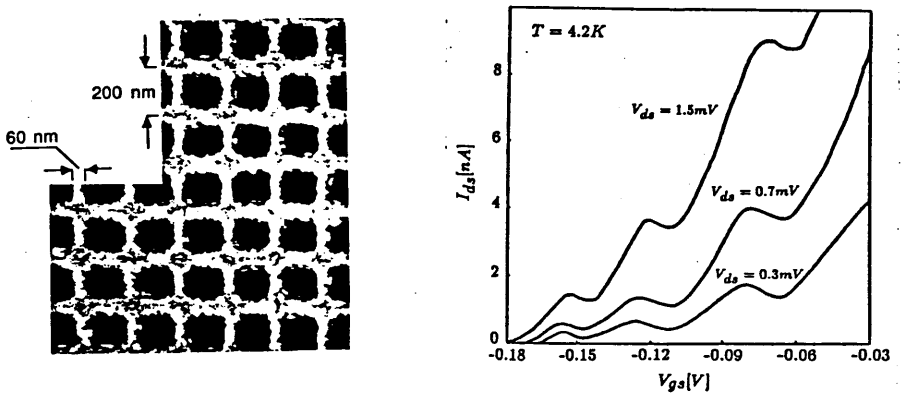
Figure 2: (a) Schematic view of a lateral surface superlattice (LSSL) device. (b) Drain-source current as a function of the gate-source voltage. The dotted curve is for a continuous gate device. The insert shows the transconductance g_m for the LSSL device. From Bagwell and Orlando (1989b) and Ismail et al. (1988).

Bragg diffraction of the electrons from the periodic potential. The structure in this device also washes away at higher temperatures and for devices with lower mobility. In addition, the structure diminishes with increasing drain-source voltage.

The structure in the conductance for the grating-gate LSSL is weak because the electrons experience a periodic potential in only one direction. A truly two dimensional periodic potential was fabricated in the grid-gate configuration as shown in Figure 3a. The period of the grid-gate is 200 nm and the linewidth is about 60 nm (Ismail et al., 1989a). A much larger modulation of the conductance is seen for this device in Figure 3b than in the grating-gate LSSL because the Bragg diffraction condition can be satisfied such that true minigaps are possible. As in the other two devices, the modulation of the conductances washes away as the temperature and drain-source voltage are increased, and as the mobility is decreased.

3 Device Modeling

There are three characteristic lengths that determine the type of electrical transport. The first is the length of the device L. The second is the average distance between elastic scattering events ℓ, which is known as the mean free path. The conduction electron has an average speed given by the Fermi velocity v_F, so that the



(b)

Figure 3: (a) SEM micrograph of a grid-gate with 200 nm period and a 60 nm line width. (b) Drain-source current as a function of the gate-source voltage for various drain-source voltages. From Ismail et al. (1989a).

mean free path can be related to a mean time τ between elastic scattering events by

$$\ell = v_F \tau . \tag{1}$$

The third characteristic length is L_ϕ, the phase breaking length. This is the length over which the phase of the quantum mechanical wavefunction of the conducting electron is correlated. The phase of the wavefunction is randomized, on average, each time τ_ϕ that the electron undergoes an inelastic scattering event which changes its energy. For this reason, we refer to L_ϕ as the inelastic scattering length. Depending on the relationship between these three lengths, the type of electrical conduction can range from diffusive transport where many scattering events occur, to ballistic transport where no scattering events occur. We now discuss these two limiting forms of electrical transport.

If $L \gg \ell$, then the electron will undergo many elastic scattering events before it traverses the device. Classically, the electron undergoes a random walk as it scatters, making its motion diffusive. The classical diffusion constant D is given by

$$D = \frac{1}{d} v_F^2 \tau = \frac{1}{d} \ell v_F, \tag{2}$$

where d is the dimensionality of the transport ($d = 2$ for a 2DEG). When electron motion is diffusive, its characteristic lengths and times are related via the diffusion constant. For example, the time τ_L for an electron to diffuse across the device and the phase breaking time τ_ϕ are given by

$$\tau_L = \frac{L^2}{D} \qquad \text{and} \qquad \tau_\phi = \frac{L_\phi^2}{D}, \tag{3}$$

as in Table 1.

The semiclassical Boltzmann transport equation gives a method for combining the concepts of classical diffusion with quantum mechanical properties. In the semiclassical method quantum mechanics is used to calculate the energy levels of the electrons.

Table 1: Length Scales and Time Constants

Diffusive $L \gg L_\phi \gg \ell$	Ballistic $L_\phi \gg \ell \gg L$
$\ell \approx \sqrt{D\tau} = v_F\tau$ $L_\phi = \sqrt{D\tau_\phi}$ $L = \sqrt{D\tau_L}$	$\ell = v_F\tau$ $L_\phi = v_F\tau_\phi$ $L = v_F\tau_L$

The electrons are then treated as wavepackets with their group velocities given by the dispersion relationship of the energy levels. This is the standard method used to calculate the conductivity of metals and semiconductors (Kittel, 1986; Ashcroft and Mermin, 1976). For the semiclassical method to be valid, there must be many inelastic scattering events in the sample such that $L \gg L_\phi \gg \ell$. Because the phase of the wavefunction is randomized at each inelastic event, the overall properties of the sample will be an ensemble average of all the possible scattering configurations. It is this averaging process that justifies the semiclassical method. Note that the wavefunction still has its phase correlated over the distance L_ϕ.

To see why there must be many inelastic scattering events to use the semiclassical method, let us suppose, on the contrary, that there are no inelastic scatterers in the sample. Quantum mechanically, the elastic scatterers could then be described by some scattering potential $V(\mathrm{r})$ which depends only on position. Therefore, the eigenstates for this potential could be found. These eigenstates have a discrete energy E so that the phase ϕ of the wavefunction would increase as $\phi = Et/\hbar$ and would be correlated for all times. Indeed, if $L_\phi \gg L$ the conduction electron would always have its phase correlated and the resulting conduction would be sensitive to the exact positions of the elastic scatterers. For such a regime of lengths $L_\phi \gg L \gg \ell$, the conduction would be dominated by unpredictable but repeatable fluctuations known as "universal conduction fluctuations" (Lee and Stone, 1985; Al'tshuler, 1985; Skocpol et al., 1986). To have diffusive transport, the phase of the wavefunction must be randomized many times during the transport across the sample, that is, $L \gg L_\phi \gg \ell$.

In the opposite limit to diffusive transport, there are no elastic scatterers in the sample so that the transport is ballistic ($\ell \gg L$). We also assume that the inelastic length is larger than the other two lengths so that $L_\phi \gg \ell \gg L$. In this case the electron moves without scattering at the Fermi velocity v_F. Therefore, the mean free path is given by Equation 1. The inelastic scattering length is related to the scattering time by

$$L_\phi = v_F\tau_\phi \tag{4}$$

and the time τ_L for the electron to ballistically traverse the sample is

$$\tau_L = \frac{L}{v_F}. \tag{5}$$

Table 1 lists these lengths and times.† We now discuss our method to calculate the transport in these two limiting regimes.

†In Table 1, L_ϕ, ℓ, τ_ϕ, and τ are defined as if the material between the contacts were infinite and disregards any influence of the contacts on the device. If the contacts are considered in the definitions

3.1 Diffusive transport

In the diffusive limit of transport where $L \gg L_\phi \gg \ell$, the calculation of the conductivity and the density of electrons can be done using the semiclassical method (Kittel, 1986; Ashcroft and Mermin, 1976). We will recast the usual expression for the density and the conductivity in terms of convolutions as the first example of our convolution method. Let $N(E)$ be the density of states for electrons in the material. The density of electrons n at a temperature T is given by

$$n(E_F,T) = \int_{-\infty}^{\infty} N(E')f(E'-E_F,T)\,dE', \tag{6}$$

where E_F is the chemical potential and $f(E,T)$ is the Fermi-Dirac distribution function. The thermodynamic density of states $N(E,T)$ is defined by

$$N(E,T) \equiv \frac{\partial n(E,T)}{\partial E} = \int_{-\infty}^{\infty} N(E')f'(E'-E,T)\,dE'. \tag{7}$$

Here $f'(E,T)$ is the negative derivative of $f(E,T)$ with respect to E and is given by

$$f'(E,T) = \frac{1}{4kT}\operatorname{sech}^2\left(\frac{E}{2kT}\right). \tag{8}$$

This bell-shaped function has a full width at half maximum of $3.5kT$ and is symmetric in E. Due to this symmetry of $f'(E)$, the thermodynamic density of states in Equation 7 can also be written as

$$N(E,T) = N(E) \otimes f'(E,T). \tag{9}$$

The symbol $\otimes$ denotes a convolution which is given by

$$\begin{aligned} C(E) \otimes D(E) &= \int_{-\infty}^{\infty} C(E')D(E-E')\,dE' \\ &= \int_{-\infty}^{\infty} D(E')C(E-E')\,dE' \end{aligned} \tag{10}$$

Because the convolution tends to broaden the first function by the width of the second function, $f'(E,T)$ will be referred to as the thermal broadening function.

In the semiclassical method the conductivity tensor is given by (Kittel, 1986; Ashcroft and Mermin, 1976).

$$\sigma^{ij}(E_F,T) = e^2 \int_{-\infty}^{\infty} v_i(E')v_j(E')\tau(E')N(E')\left[-\frac{\partial f(E'-E_F,T)}{\partial E'}\right] dE'. \tag{11}$$

At zero temperature the derivative of the Fermi function is a delta function so that the conductivity is simply

$$\sigma^{ij}(E) \equiv e^2 v_i(E)v_j(E)\tau(E)N(E). \tag{12}$$

Furthermore, the derivative of the Fermi function in Equation 11 is just $f'(E_F-E',T)$ since this function is symmetric about the Fermi energy. Therefore, the conductivity

of these lengths and times, it is then possible to interpret the contacts as a source of inelastic (phase-breaking) scattering such that $L_\phi = L$ and $\tau_\phi = \tau_L$. Nevertheless, ballistic transport still means that no scattering occurs in the material between the contacts.

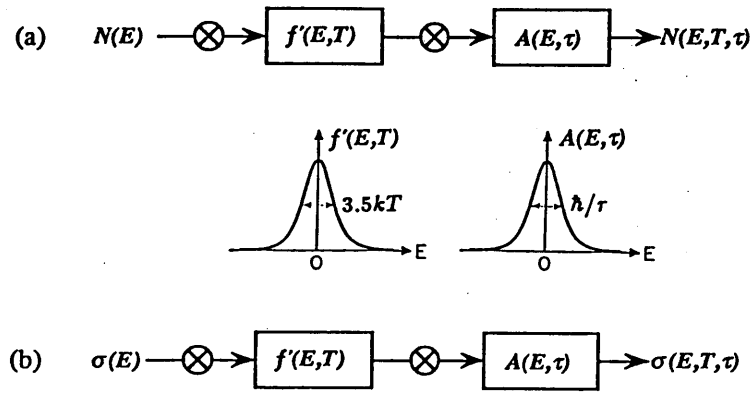


Figure 4: The convolution method for (a) the thermodynamic density of states $N(E,T,\tau)$ and (b) the conductivity $\sigma^{ij}(E,T,\tau)$. The thermal broadening function $f'(E,T)$ and the impurity broadening function $A(E,\tau)$ are shown graphically along with their widths. The symbol $\otimes$ denotes a convolution in energy.

as a function of temperature in Equation 11 can also be written as a convolution; namely,

$$\sigma^{ij}(E,T) = \sigma^{ij}(E) \otimes f'(E,T)\,. \tag{13}$$

We see that if the density of states and the conductivity are known at zero temperature, then their values at finite temperatures can be found by a simple convolution with the thermal broadening function $f'(E,T)$.

In the diffusive transport limit, the ensemble averaging (due to the phase randomization from inelastic scattering) leads to an additional energy level broadening (whose width is set by the elastic scattering). This additional broadening is characterized by the impurity broadening function $A(E,\tau)$ given by (Abrikosov et al., 1963; Rickayzen, 1980; Bagwell, 1988)

$$A(E,\tau) = \frac{1}{\pi}\frac{\hbar/2\tau}{E^2 + (\hbar/2\tau)^2}\,. \tag{14}$$

The impurity broadening function has a Lorentzian shape with a full width at half maximum of $\hbar/\tau$. The thermodynamic density of states including the broadening from both temperature and impurities is given by (Bagwell et al., 1989; Bagwell and Orlando, 1989b)

$$N(E,T,\tau) = N(E) \otimes f'(E,T) \otimes A(E,\tau)\,. \tag{15}$$

Likewise the conductivity is given by (Bagwell et al., 1989; Bagwell and Orlando, 1989b)

$$\sigma^{ij}(E,T,\tau) = \sigma^{ij}(E) \otimes f'(E,T) \otimes A(E,\tau)\,. \tag{16}$$

Figure 4 summarizes the convolution method for diffusive transport and shows the shapes of the two broadening function. Because the convolution with $f'(E,T)$ depends only on the temperature while the convolution with $A(E,\tau)$ depends only on

the impurity scattering, we see that the broadening effects from temperature and impurities are independent effects.

This convolution method readily implies that, if the conductance has a feature which occurs over some energy range ΔE, then to see that structure the widths of the broadening functions have to be smaller than ΔE. For temperature and impurity broadening, this means that

$$\Delta E > 3.5kT \qquad \text{and} \qquad \Delta E > \frac{\hbar}{\tau}. \tag{17}$$

These two restrictions lead to the observability criteria (Antoniadis et al., 1985) that to see quantum effects, the electron should have a small effective mass, high mobility, low temperature, and be confined to a small length scale. Measurements clearly show the washing away of structure in device I-V characteristics as the temperature is increased and as the mobility is decreased (Ismail, 1989).

The convolution method clearly demonstrates why the structure in the conductivity tends to wash away at high temperatures and low mobilities. However, this method can also be used to understand device characteristics. In most field effect devices, the conductivity (of the source-drain) is measured as a function of gate-source voltage. We will assume here that the gate-source voltage is proportional to the density of electrons in the inversion layer. (If the density is related to the gate-source voltage in a more complicated fashion, then this more complicated relationship must be used.) Therefore, we seek in our convolution model to plot the conductivity versus density n. From the definition of the thermodynamic density of states in Equation 7, the density is

$$n(E_F,T) = \int_{-\infty}^{E_F} N(E,T,\tau)\,dE\,. \tag{18}$$

Another piece of information needed to model a specific device is the scattering time τ that enters the conductivity $\sigma^{ij}(E)$ and the impurity broadening function $A(E,\tau)$. This scattering time can be calculated from the quantum mechanical expression for the scattering rate τ^{-1} given by Fermi's golden rule; namely,

$$\frac{1}{\tau} = \frac{2\pi n_{\text{imp}}\,|V|^2}{\hbar} N(E)\,. \tag{19}$$

Here n_{imp} is the density of impurities and $|V|^2$ is the square of the matrix element between scattering states, which is assumed constant for the average of random impurities (Madelung, 1978). Hence, we see that τ is inversely proportional to the density of states. This is plausible because more states to scatter into implies a higher scattering rate and therefore a lower mobility. The constant of proportionality between τ and $N(E)$ can be found if the mobility $\mu(E) = e\tau(E)/m$ is known for a given energy E_o (or equivalent density n_o). Then,

$$\tau(E) = \frac{m\mu(E_o)}{e}\frac{N(E_o)}{N(E)} = \tau_o \frac{N(E_o)}{N(E)}\,. \tag{20}$$

Equations 15, 16, 18, and 19 are the central results of the convolution method needed to calculate the diffusive conductivity for a given device. We will now use these results to find the conductivity for a few devices. The first example will be for a one dimensional (1D) conductor. We will then show how the conductivity for a

Q1D device and the LSSL device can be calculated by means of yet another simple convolution.

As the first example of using the convolution method, we consider a 1D conductor in which the electrons are free to move in the y-direction. The energy-momentum dispersion relationship is

$$E_{1D}^{y} = \frac{\hbar^2 k_y^2}{2m}. \tag{21}$$

The group velocity is then

$$v^{y}(E) = \frac{1}{\hbar}\frac{\partial}{\partial k_y} E_{1D}^{y} = \sqrt{\frac{2E}{m}}. \tag{22}$$

The density of states for the electrons (including both spin states) is the usual 1D result (Kittel, 1986; Ashcroft and Mermin, 1976),

$$N_{1D}^{y}(E) = \sqrt{\frac{2m}{\pi^2\hbar^2 E}}, \tag{23}$$

which integrates to give the density

$$n_{1D} = \sqrt{\frac{8mE}{\pi^2\hbar^2}}. \tag{24}$$

The conductivity is non-zero only in the y-direction so that

$$\sigma_{1D}^{yy} = e^2 v_y^2(E)\tau(E)N_{1D}^{y}(E). \tag{25}$$

These zero temperature results are displayed in the first column of Figure 5 where the scattering time has been assumed to be a constant for all energies. The density of states $N_{1D}^{y}(E)$ displays the inverse square root singularity in E, while the density $n_{1D}(E)$ has the resulting $\sqrt{E}$ dependence. The conductivity $\sigma_{1D}^{yy}(E)$ also has a $\sqrt{E}$ dependence because the energy dependence of the density of states cancels the energy dependence of one factor of velocity in Equation 25. Therefore, the conductivity for a constant scattering time has the same energy dependence as the velocity. Conductivity versus density is obtained by parametrically combining the conductivity-vs-energy and the density-vs-energy plots. This combination results in a conductivity versus density is that is linear, as shown in the bottom plot in the first column. This reflects the fact that $\sigma = en\mu$ and the mobility is a constant. This result also holds at non-zero temperatures.

The importance of taking τ to be inversely proportional to the density of states as in Equation 20 is illustrated in Figure 6. For zero temperature and for no impurity broadening, the density of states and the density are the same as in Figure 5. The linear increase of σ_{1D}^{yy} with energy is a direct result of τ being inversely proportional to $N_{1D}^{y}(E)$ so that these two factors cancel in Equation 25. The conductivity then has the energy dependence of v_y^2 which is proportional to E. Therefore, since n_{1D} depends on the square root of energy, the conductivity depends on the square of the density as shown in the bottom figure in the first column. The effect of temperature broadening is included by simply convolving the zero temperature results for $N_{1D}^{y}(E)$ and $\sigma_{1D}^{yy}(E)$ with $f'(E,T)$. These results are shown as the dashed lines in Figure 6. Although we have only shown the results for temperature broadening, the

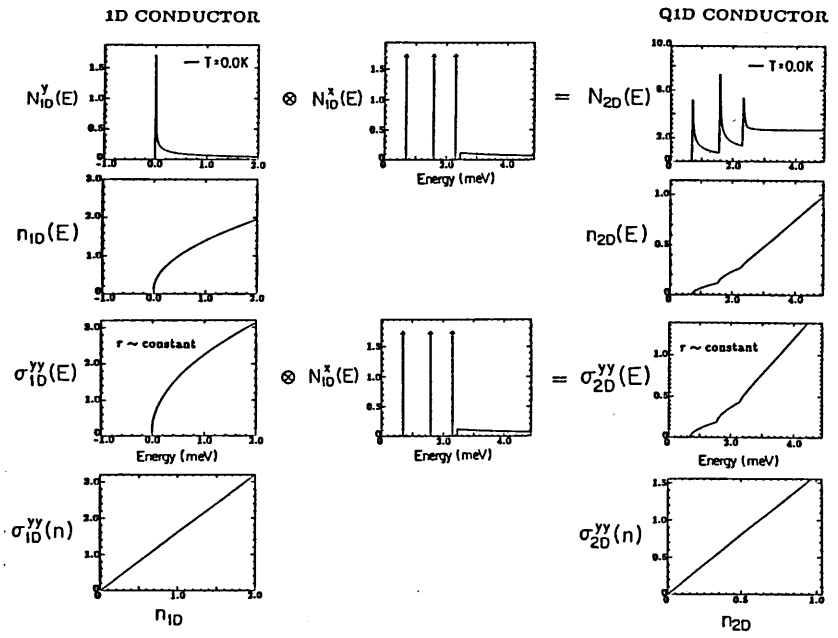


Figure 5: Calculated diffusive conductance for 1D and Q1D wires. The first column shows the density of states, density, and conductivity as a function of energy and the conductivity versus density for a 1D conductor where the electrons are free to move in the y-direction and τ *is taken to be a constant.* The second column is the density of states in the x-direction for a Q1D conductor. The third column shows the full results for the Q1D conductor. The symbol $\otimes$ denotes a convolution in energy. Units are arbitrary.

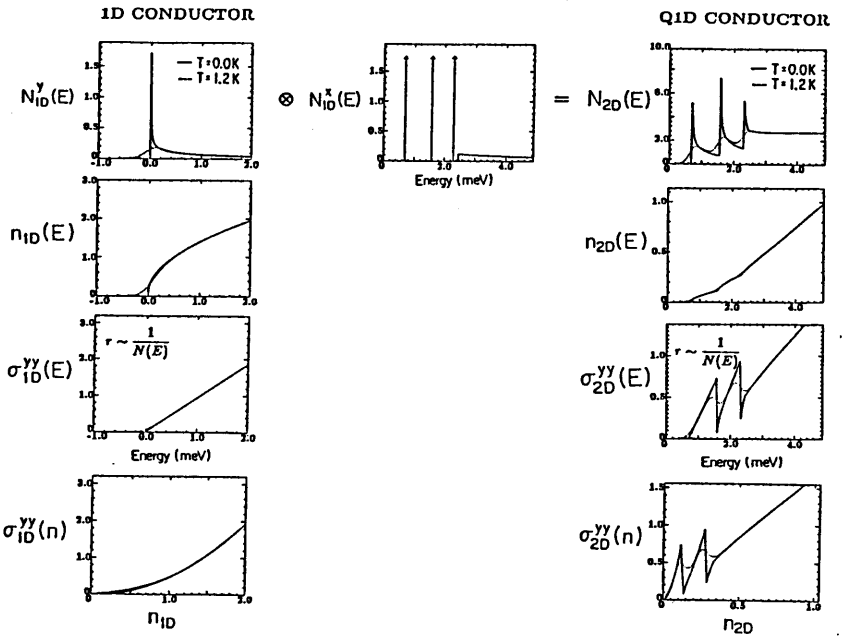


Figure 6: Calculated diffusive conductance of 1D and Q1D wires as Figure 5, but with the scattering time τ *taken to be inversely proportional to* $N_{2D}(E)$. The solid curves are for $T = 0$, and the dashed curves are for $T = 1.2\,\mathrm{K}$. Note that $\sigma^{yy}_{2D}(E) \neq \sigma^{yy}_{1D}(E) \otimes \frac{1}{2}N^{x}_{1D}(E)$ since the scattering time is now a function of energy.

results for impurity broadening are similar because the resulting convolutions will give qualitatively similar plots.

To find the conductivity for the Q1D and grating-gate LSSL devices (but not the grid-gate), the two dimensional problem must be solved. In both of these devices the total potential is of the separable form

$$V(x,y) = V_x(x) + V_y(y)\,. \tag{26}$$

With this potential the total energy E of the 2D system is

$$E = E^x_{1D} + E^y_{1D} \tag{27}$$

where E^x_{1D} and E^y_{1D} are the eigenenergies for the corresponding 1D problems with potential $V_x(x)$ and $V_y(y)$ respectively. Let $N^x_{1D}(E^x)$ and $N^y_{1D}(E^y)$ be the corresponding density of states for the 1D problem. Then the total density of states $N_{2D}(E)$ for E consistent with Equation 27 is simply

$$N_{2D}(E) = \frac{1}{2}\int_{-\infty}^{\infty} N^x_{1D}(E_x) N^y_{1D}(E - E_x)\, dE_x\,. \tag{28}$$

The factor of $1/2$ compensates for the overcounting due to the spin states in the 1D densities of states. Equation 28 is a convolution so that it can be written as

$$N_{2D}(E) = N^y_{1D}(E) \otimes \frac{1}{2} N^x_{1D}(E)\,. \tag{29}$$

Figure 7a shows the diagram for this convolution.

The convolutions for the conductivity are not as straightforward as for the density of states. This is because the scattering time depends on the total density of states. Nevertheless, the components of the conductivity tensor in 2D for a separable potential can be written as (Bagwell et al., 1989; Bagwell and Orlando, 1989b)

$$\frac{\sigma^{yy}_{2D}(E)}{\tau_{2D}(E)} = \frac{\sigma^{yy}_{1D}(E)}{\tau_{1D}(E)} \otimes \frac{1}{2} N^x_{1D}(E) \tag{30}$$

and

$$\frac{\sigma^{zz}_{2D}(E)}{\tau_{2D}(E)} = \frac{\sigma^{zz}_{1D}(E)}{\tau_{1D}(E)} \otimes \frac{1}{2} N^y_{1D}(E) \tag{31}$$

and

$$\sigma^{xy}_{2D}(E) = \sigma^{yx}_{2D}(E) = 0\,. \tag{32}$$

A convenient algorithm for doing the convolutions for the conductivity is to first do the problem for the case when the scattering times are constant and equal. Then Equations 30 and 31 are simply convolutions with the 1D conductivities. That result can be multiplied by the energy dependent scattering time (which is proportional to the inverse of the density of states) to give the conductivity. Figure 7b shows the diagram for the convolution in Equation 30 and Equation 31.

To see how the convolutions for going from 1D to 2D devices work, consider the case of the Q1D conductors shown in Figure 1. Each one of the 1D parallel conductors can be considered to have a number of subbands at discrete energy levels. For simplicity we take the number of subbands to be three. The corresponding density of states $N^x_{1D}(E)$ describing the subbands in the x-direction is shown in the second

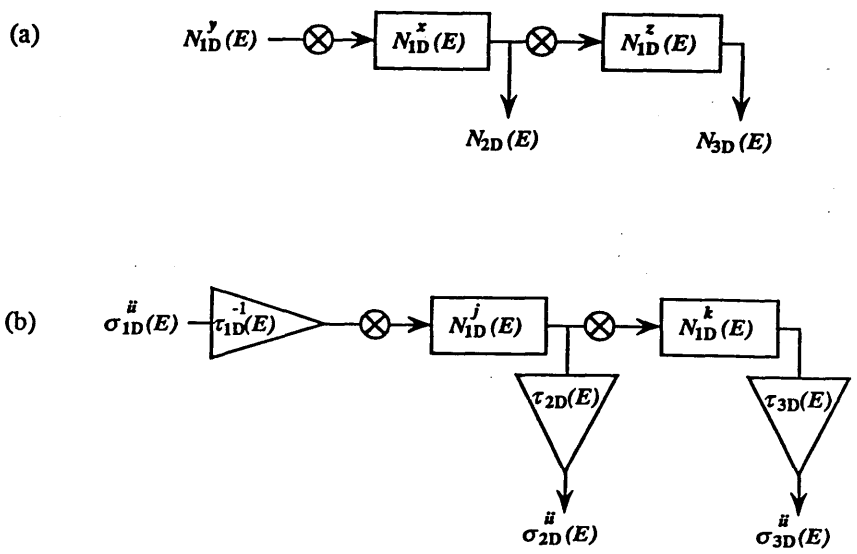


Figure 7: The diagrams for calculating (a) the density of states and (b) the conductivity when the potential is separable. These diagrams assume zero temperature and no impurity broadening. The symbol ⊗ denotes a convolution in energy.

column of Figure 5. The three delta functions correspond to the three energy levels of the subbands for a model potential. We have also included a portion of the 1D density of states going as $E^{-1/2}$ to describe the density of states when the electron's energy exceeds the confining potential, since at those high energies the electron is virtually free. Of course, this is only a model density of states. Calculations based on solving Schrödinger's equation for the confining potential give similar results (Bagwell and Orlando, 1989b). The density of states N_{2D} is calculated by convolving the two 1D density of states as is shown in the top row of Figure 5. The density $n_{2D}(E)$ is just the integral of the density of states. The conductivity $\sigma_{2D}^{yy}(E)$ is a convolution of the 1D conductivity with the density of states if the scattering time is taken to be a constant. The result of that convolution is shown in the third row of Figure 5. Note as a function of density that $\sigma_{2D}^{yy}(n)$ is linear even though the density of states and the conductivity both have structure in energy. This must be so for a constant scattering time, since the conductivity is given by $\sigma = en\mu$ and μ is a constant for a constant τ.

Taking $\tau(E)$ to be inversely proportional to $N_{2D}(E)$ gives $\sigma_{2D}^{yy}(E)$ as shown in Figure 6. $\sigma_{2D}^{yy}(E)$ in Figure 6 follows from multiplying $\sigma_{2D}^{yy}(E)$ in Figure 5 by the ratio $\tau(E)/\tau_o$ from Equation 20, where τ_o is the constant scattering time in Figure 5. Now when we plot $\sigma_{2D}^{yy}(E)$ versus density in Figure 6 the structure remains. The conductivity for a Q1D device is a minimum each time a new subband begins to be filled, because the scattering time (and mobility) is a minimum there due to the sudden increase in the density of states. Hence, we see that it is the modulation of the scattering time (mobility) that gives rise to the structure in the Q1D device.

By including temperature broadening and impurity broadening, we have been

able to model the conductivity of the Q1D device shown in Figure 1. The model calculation semi-quantitatively reproduces the features in the measured conductance.

The convolution method can also be applied to find σ_{2D}^{zz}. If we consider a periodic array of wires, then this configuration is equivalent to finding the conductivity for the grating-gate LSSL. The model for this device has been discussed in (Bagwell and Orlando, 1989b).

3.2 Ballistic transport

In the ballistic limit of transport where $L_\phi \gg \ell \gg L$, the conductivity can be calculated using Landauer's formula. Landauer's formula connects the current in the device to the quantum mechanical transmission coefficient $T(E,V)$ through the device (Landauer, 1957; Landauer, 1970). In 1D the current flowing in the y-direction from a left contact, which is at a voltage V greater than the right contact, is given by (Wolf, 1985)

$$I_{1D}(E,V,T) = e\int v_{1D}^{+}(E')T_{1D}(E',V)N_{1D}^{+}(E') \\ [f(E'-E,T) - f(E'-(E-eV),T)]\,dE'. \tag{33}$$

Here $v_{1D}^{+}(E)$ is the group velocity for electrons moving in the positive y-direction, $N_{1D}^{+}(E)$ the density of states for electrons of both spins moving in the positive y-direction in a one dimensional free electron gas, $f(E,T)$ is the Fermi-Dirac distribution function, T the temperature of both contacts, and $T_{1D}(E,V)$ the transmission coefficient. Here the energy E is measured from the bottom of the band of electrons in the left contact.

The product of the group velocity and the electron density of states in one dimension is a constant given by

$$v_{1D}^{+}(E)N_{1D}^{+}(E) = \frac{1}{\pi\hbar}. \tag{34}$$

Note that this is a half of what Equations 22 and 23 imply since we are only interested in the electrons which are traveling in the positive direction, which is half the total density of states at a given energy.

The Fermi-Dirac function can also be expressed as the convolution

$$f(E,T) = [1-\theta(E)] \otimes f'(E,T) \tag{35}$$

where θ is the unit step function. Equation 33 can therefore be rewritten as (Bagwell and Orlando, 1989a)

$$I_{1D}(E,V,T) = \frac{e}{\pi\hbar}T_{1D}(E,V) \otimes W(E,V) \otimes f'(E,T). \tag{36}$$

Here

$$W(E,V) = [\theta(E) - \theta(E-eV)] \tag{37}$$

is the voltage broadening function.[‡] Figure 8 depicts the convolution method for the current in 1D as given by Equation 36. Figure 8 is valid in the limit of ballistic transport.

[‡]If the applied voltage is greater than the Fermi energy then $W(E,V)$ must be cut off at the Fermi energy, namely

$$W(E,V) = \begin{cases} [\theta(E) - \theta(E-eV)] & eV < \mu \\ [\theta(E) - \theta(E-\mu)] & eV \geq \mu \end{cases}. \tag{38}$$

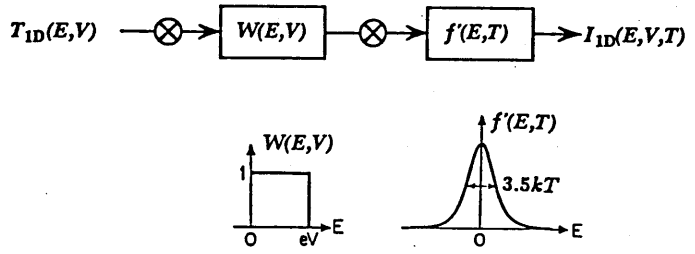


Figure 8: The convolution method for the current in 1D ballistic transport. The thermal broadening function $f'(E,T)$ and the voltage broadening function $W(E,V)$ are shown graphically along with their widths.

As an example of the convolution method, consider a ballistic Q1D conductor of uniform width. If E_1 is the energy of the first subband, then, neglecting all the higher subbands, the Q1D conductor has unity transmission coefficient for all energies greater than E_1 such that

$$T_{1D}(E,V) = \theta(E - E_1)\,. \tag{39}$$

For small voltages the zero temperature current becomes

$$I_{1D}(E,V) = \frac{e^2}{\pi\hbar} V\theta(E - E_1)\,, \tag{40}$$

where structures of order eV have been neglected. Hence, we see that a Q1D conductor with perfect transmission is characterized by a resistance $\pi\hbar/e^2$ for all energies greater than E_1. This is the quantum contact resistance (van Wees et al., 1988; Wharam et al., 1988). If there are two 1D subbands at E_1 and E_2 in a Q1D device, then for perfect transmission the contact conductance would be $e^2/\pi\hbar$ for each subband. The convolution broadening of Equation 40 with temperature and voltage, similar to Figure 9 below, is shown in Bagwell and Orlando (1989a).

Next, consider the case where the 1D subbands are formed in a constriction between two wider regions. This geometry describes the experiments of van Wees et al. (1988) and Wharam et al. (1988). For this case, there is an electostatic potential difference between the wider regions and the constriction (Payne, 1989). As the applied voltage increases, the 1D subbands formed in the constriction shift to a lower energy with respect to the incident electron distribution in the wide regions. The transmission coefficient becomes approximately

$$T_{1D}(E,V) = \theta(E - E_1 + meV)\,, \tag{41}$$

where $m = 1/2$ for the case of a symmetrical constriction. In general m is a phenomenological parameter where $0 \leq m \leq 1$ (Kouwenhoven et al., 1989). For two Q1D subbands at E_1 and E_2, the conductance when V is small is given by the solid line in Figure 9. We consider a symmetrical constriction so that, when the applied voltage is not negligible, the broadening of the conductance $G = I/V$ due to finite voltage is shown by the dashed line in Figure 9. The effect of temperature broadening

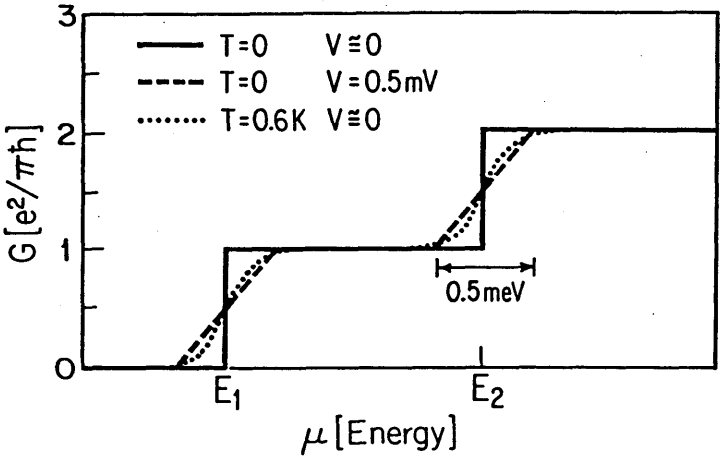


Figure 9: The ballistic conductance versus emitter Fermi energy for two 1D subbands at E_1 and E_2 formed in a constriction between two wider regions. The two subbands have perfect transmission as shown by the solid line. The energy of the left contact is assumed to be at a voltage V above the right contact. The dashed curved shows the broadening due to voltage and the dotted curve the broadening due to temperature following Bagwell and Orlando (1989a).

is also shown in the figure. Note that the shapes of the two broadenings are qualitatively different. Results similar to Figure 9 have also been obtained by Glazman and Khaetskii (1988, 1989). The nonlinear conductance of these ballistic devices has been measured and analyzed (Kouwenhoven et al., 1989). In addition, Landauer (1987, 1989) has addressed many fundamental issues relating to the conductance of these constrictions.

Many types of ballistic devices can be made in two and three dimensions. Again, if the transmission coefficient is known for the device, then the Landauer formula can be used to calculate the conductance. Unfortunately, the Landauer formula for higher dimensions, known as the multi-channel Landauer formula (Büttiker et al., 1985; Fischer and Lee, 1981), is in general quite complicated to evaluate. However, for separable potentials of the form given by Equation 26, the multichannel formula is equivalent to adding in parallel all the possible channels opened up by the second dimension. This parallel combination is equivalent to a convolution with the free electron density of states and hence, is similar to the diffusive result of Equations 29 and 30.

If the 1D transmission coefficient is known, then the convolution method will give the current for a 1D conductor. For a 2D conductor described by a separable potential as in Equation 26, the current density can be written as (Bagwell and Orlando, 1989a)

$$J_{2D}(E,V,T) = I_{1D}(E,V,T) \otimes \frac{1}{2} N^z_{1D}(E), \tag{42}$$

where $I_{1D}(E,V,T)$ is known from Equation 36. Likewise, the current density in 3D

can be written for a separable potential in 3D as (Bagwell and Orlando, 1989a)

$$J_{3D}(E,V,T) = I_{1D}(E,V,T) \otimes \frac{1}{2}N^s_{1D}(E) \otimes \frac{1}{2}N^s_{1D}(E)\,. \tag{43}$$

These two equations have been used to correctly give the I-V characteristics for 3D resonant tunneling devices as well as for 2D resonant tunneling devices.

We have restricted our calculations to only the diffusive and ballistic limits. In the region between these two limits the effects of inelastic and elastic scattering must be treated at the same time. At present this has been done for a few isolated examples because of the complexity of the general problem. However, a rough estimate of how the transport is affected by energy level broadening in between these two limits can be given as follows: Divide the sample into blocks which have a linear dimension of the average L_ϕ. The conductivity can be calculated in each of these blocks by considering the transport in each block to be ballistic. The conductances of each block can then be added up as a network of resistors (Büttiker, 1988). The broadening due to temperature would be the same for each block. Hence, if there was structure on the order of ΔE in each block, then the temperature of about $\Delta E = 3.5kT$ would wash the structure away. However, the voltage broadening would only be due to that part of the voltage that was dropped across each block. Therefore, the voltage broadening would be much less if there were more inelastic lengths in the sample. For example, a 1D conductor which has $L = mL_\phi$ has a voltage broadening of width eV/m where V is the total voltage drop across the sample. Therefore, a voltage such that $eV = m\Delta E$ is necessary to wash away structure of the order ΔE. This dependence of voltage broadening on the number of inelastic lengths has been observed in grid gate LSSL devices and has been used to find L_ϕ (Ismail, 1989) for devices where m is small. As the number of inelastic blocks gets large, we expect the result from adding up the ballistic resistances of the blocks to merge into the result calculated from the diffusive limit when $L \gg L_\phi \gg \ell$.

4 Summary

In this paper we have shown how the convolution method provides a convenient framework for understanding and calculating the conductivity in the diffusive and ballistic regimes of transport. In the diffusive regime the effects of temperature and impurity broadening can be written as convolutions. In the ballistic regime the effects of temperature and finite voltage can also be written as convolutions. We have also seen that for separable potentials, which describe many but not all devices, that the density of states and conductivity for 2D and 3D structures can be written as simple convolutions with the results for the 1D structure. Hence, we have used the convolution method in two ways: (1) to include physical broadening effects, and (2) to include dimensional effects.

Acknowledgements

We thank D. A. Antoniadis and Henry I. Smith for useful discussions. We gratefully acknowledge financial support from an IBM fellowship, an Office of Naval Research Fellowship, an MIT-EECS Departmental special projects grant, and the support of U. S. Air Force contract AFOSR-88-0304. We also acknowledge partial support from Joint Services Electronics Program contract DAALO3-89-C-0001.

References

Abrikosov, A. A., Gorkov, L. P., and Dzyaloshinski, I. E., 1963, "Methods of Quantum Field Theory in Statistical Physics," Prentice-Hall, Englewood Cliffs, N.J.

Al'tshuler, B. L., 1985, Fluctuations in the Extrinsic Conductivity of Disordered Conductors, *Sov. Phys. JETP Lett.*, 41:648.

Antoniadis, D. A., Warren, A. C., and Smith, H. I., 1985, Quantum Mechanical Effects in Very Short and Very Narrow Channel MOSFETs, *IEDM Tech. Dig.*, 562.

Ashcroft, N. W. and Mermin, N. D., 1976, "Solid State Physics," Holt, Rinehart, and Winston, New York.

Bagwell, P. F., 1988, "Quantum Mechanical Transport Phenomenon in Nanostructured Inversion Layers," S. M. Thesis, MIT.

Bagwell, P. F., Antoniadis, D. A., and Orlando, T. P., 1989, Quantum Mechanical and Non-Stationary Transport Phenomenon in Nanostructured Silicon Inversion Layers, *in*: "Advanced MOS Device Physics," N. Einspruch and G. Gildenblat, ed., Academic Press, San Diego.

Bagwell, Phillip F., and Orlando, Terry P., 1989a, Landauer's Conductance Formula and its Generalization to Finite Voltages, *Phys. Rev. B*, 40:1456.

Bagwell, P. F., and Orlando, T. P., 1989b, Broadened Conductivity Tensor and Density of States for a Superlattice Potential in One, Two, and Three Dimensions, *Phys. Rev. B*, 40:3735.

Bernstein, G., and Ferry, D. K., 1987, Negative Differential Conductivity in Lateral Surface Superlattices, *J. Vac. Sci. Technol. B*, 5:964.

Büttiker, M., Imry, Y., Landauer, R., and Pinhas, S., 1985, Generalized Many-Channel Conductance Formula with Application to Small Rings, *Phys. Rev. B*, 31:6207.

Büttiker, M., 1986, Role of Quantum Coherence in Resistors, *Phys. Rev. B*, 33:3020. See also Büttiker, M., 1988, Coherent and Sequential Tuneling in Series Barriers, *IBM J. Res. Dev.*, 32:63.

Fischer, Daniel S., and Lee, Patrick A., 1981, Relation Between Conductivity and Transmission Matrix, *Phys. Rev. B*, 23:6851.

Glazman, L. I., and Khaetskii, A. V., 1988, Nonlinear Quantum Conductance of a Point Contact, *JETP Lett.*, 48:591.

Glazman, L. I., and Khaetskii, A. V., 1989, Nonlinear Quantum Conductance of a Lateral Microconstraint in a Heterostructure, *Europhys. Lett.*, 9:263.

Ismail, K., Chu, W., Antoniadis, D. A., and Smith, Henry I., 1988, Surface-Superlattice Effects in a Grating-Gate GaAs/GaAlAs Modulation-Doped Field-Effect-Transistor, *Appl. Phys. Lett.*, 52:1071.

Ismail, K., Chu, W., Yen, A., Antoniadis, D. A., and Smith, Henry I., 1989a, Negative Transconductance and Negative Differential Resistance in a Grid-Gate Modulation Doped Field-Effect Transistor, *Appl. Phys. Lett.*, 54:460.

Ismail, K., Chu, W., Antoniadis, D. A., and Smith, Henry I., 1989b, One Dimensional Subbands and Mobility Modulation in GaAs/AlGaAs Quantum Wires, *Appl.*

Phys. Lett., 54:1130.

Ismail, Khalid, 1989, "The Study of Electron Transport in Field-Induced Quantum Wells on GaAs/AlAs", Ph. D. Thesis, MIT.

Kittel, C., 1986, "Introduction to Solid State Physics", Wiley, New York.

Kouwenhoven, L. P., van Wees, B. J., Harmans, C. J. P. M., Williamson, J. G., van Houten, H., Beenakker, C. W. J., Foxon, C. T., and Harris, J. J., 1989, Nonlinear Conductance of Quantum Point Contacts, *Phys. Rev. B.*, 39:8040.

Landauer, R., 1957, Spatial Variation of Currents and Fields Due to Localized Scatterers in Metallic Conduction, *IBM J. Res. Dev.*, 1:223.

Landauer, Rolf, 1970, Electrical Resistance of Disordered One Dimensional Lattices, *Phil. Mag.*, 21:683.

Landauer, Rolf, 1987, Electrical Transport in Open and Closed Systems, *Z. Phys. B.*, 68:217.

Landauer, Rolf, 1989, Conductance Determined by Transmission: Probes and Quantized Constriction Resistance, to appear in *J. Phys. Cond. Matt.*.

Lee, P. A., and Stone, A. D., 1985, Universal Conductance Fluctuations in Metals, *Phys. Rev. Lett.*, 55:1622.

Madelung, O., 1978, "Introduction to Solid State Theory," Springer-Verlag, New York.

Payne, M. C., 1989, Electrostatic and Electrochemical Potentials in Quantum Transport, *J. Phys. Cond. Matt.*, 1:4931.

Rickayzen, G., 1980, "Green's Functions and Condensed Matter," Academic Press, New York.

Scott-Thomas, J. H. F., Kastner, M. A., Antoniadis, D. A., Smith, Henry I., and Field, Stuart, 1988, Si MOSFETs with 70nm Slotted Gates for Study of Quasi One Dimensional Quantum Transport, *J. Vac. Sci. Technol. B*, 6:1841.

Skocpol, W. J., Mankiewich, P. M., Howard, R. E., Jackel, L. D., Tennant, D. M., and Stone, A. D., 1986, Universal Conductance Fluctuations in Silicon Inversion Layer Nanostructures, *Phys. Rev. Lett.*, 56:2865.

Tokura, Y., and Tsubaki, K., 1987, Conductivity Oscillation due to Quantum Interference in a Proposed Washboard Transistor, *Appl. Phys. Lett.*, 51:1807.

van Wees, B. J., van Houten, H., Beenakker, C. W. J., Williamson, J. G., Kouwenhoven, L. P., van der Marel, D., and Foxon, C. T., 1988, Quantized Conductance of Point Contacts in a Two Dimensional Electron Gas, *Phys. Rev. Lett.*, 60:848.

Warren, A. C., Antoniadis, D. A., Smith, H. I., and Melngailis, J., 1985, Surface Superlattice Formation in Silicon Inversion Layers Using 0.2-μm Period Grating-Gate Electrodes, *IEEE Electron Dev. Lett.*, EDL-6:294.

Warren, A. C., Antoniadis, D. A., and Smith, H. I., 1986, Quasi One-Dimensional Conduction in Multiple, Parallel Inversion Lines, *Phys. Rev. Lett.*, 56:1858.

Wharam, D. A., Thornton, T. J., Newbury, R., Pepper, M., Ahmed, H., Frost, J. E. F., Hasko, D. G., Peacock, D. C., Ritchie, D. A., and Jones, G. A. C., 1988, One Dimensional Transport and the Quantization of the Ballistic Resistance, *J. Phys. C: Solid State Phys.*, 21:L209.

Wolf, E. L., 1985, "Principles of Electron Tunnelling Spectroscopy", Oxford University Press, New York. See equations (2.5) and (2.6).

Appendix D

Philip F. Bagwell and Terry P. Orlando, 'Landauer's Conductance Formula and its Generalization to Finite Voltages', Physical Review B, 40, 1456 (1989).

Landauer's conductance formula and its generalization to finite voltages

Philip F. Bagwell and Terry P. Orlando
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 16 December 1988; revised manuscript received 14 March 1989)

We recast the expression for tunneling current to show that the energy averaging due to a finite voltage is statistically independent from the energy averaging due to finite temperature. The energy averaging takes the mathematical form of convolutions. Convolving the standard Landauer conductance formula in energy with a voltage windowing function averages the conductance over an energy range equal to the applied voltage, just as the independent convolution with the Fermi-Dirac probability density function averages the conductance over an energy range equal to the temperature. We illustrate the effects of voltage broadening versus ordinary thermal broadening of the conductance using a quasi-one-dimensional ballistic conductor as a model system, as in the recent experiments of van Wees *et al.* [Phys. Rev. Lett. **60**, 848 (1988)] and Wharam *et al.* [J. Phys. C **21**, L209 (1988)].

I. INTRODUCTION

Since Landauer's original work[1,2] relating electrical conductance to the transmission probability through a region of elastic scatterers between two temperature baths, his formula has been extended to the case of multiple quantum channels,[3,4] extended to include inelastic scattering within the sample,[5] applied to many different experimental geometries,[6,7] and derived from linear-response theory.[4,8] A review of the two types of Landauer formulas, discussing the different measurement conditions under which each applies, is given in Refs. 6, 8, 9, and 10. References 6, 8, and 10 review the higher-dimensional analogs of these two types of Landauer formulas. The extension of Landauer's formula to finite voltages has also been investigated.[11,12]

The main goal of this paper is to express the effects of energy averaging due to finite voltage and finite temperature on the electrical conductance mathematically in terms of convolutions. This objective is important because convolutions are a very natural way to incorporate random processes leading to a broadening of energy levels, as is the case with the energy level broadening due to elastic and inelastic scattering in a normal dirty metal.[13] We show in Sec. II that to generalize the zero-voltage and zero-temperature Landauer conductance formula to finite voltage and finite temperature, one merely convolves it in energy with both a voltage windowing function and a thermal smearing function, respectively. These two convolutions show that the effect of finite temperature is to average the conductance over a region $k_B T$ near the Fermi level, while one effect of finite voltage is to average the conductance over an energy range eV near the Fermi level, as previously mentioned in Refs. 11 and 12. By proving that the difference in Fermi functions characterizing electrical conduction can be written as a convolution of two separate functions, one function depending only on the applied voltage and the other depending only on temperature, we show in this paper that voltage broadening and temperature broadening are statistically independent. We apply this convolution broadening method to a quasi-one-dimensional ballistic conductor to understand the qualitative similarities and differences between voltage broadening and ordinary thermal broadening of the conductance. By emphasizing that electrical conduction necessarily occurs at finite voltages, we gain additional insight into the physical mechanism responsible for the quantum contact resistance in one-dimensional ballistic conductors. In Sec. III we examine Landauer's formula in the classical diffusive limit in one dimension, showing that it reduces to the standard Drude conductance result. We show that one also obtains the quantum contact resistance from the Drude-Sommerfeld conductance formula in the limit that the mean free path grows to its upper bound of the device size.

A subsidiary goal of this paper is to examine the relationship of Landauer's conductance formula to previous treatments of electron tunneling at finite voltages. In the tunneling limit, by which we mean many parallel quantum channels with no scattering between them, we show in Sec. IV that the finite-voltage Landauer formula reduces to many well-known treatments of electron tunneling at finite voltages.[12,14,15] However, even in this well-known tunneling limit, we can gain physical insight into the shape of the current density versus Fermi energy in different spatial dimensions by noting that, because the tunneling Hamiltonian is separable, we can obtain the tunneling current density in two spatial dimensions by convolving the one-dimensional tunneling current with the one-dimensional free-electron density of states, repeating this process to obtain the three-dimensional tunneling current density.

II. LANDAUER'S FORMULA IN ONE DIMENSION

Our model for a quantum conductor in one dimension is a potential $U(x,V)$ between two thermal reservoirs as in Fig. 1, where U depends both on position x and the applied voltage V. The potential $U(x,V)$ depends on the applied voltage because electrons reflected from a scattering obstacle pile up on one side of the obstacle, as emphasized in Ref. 1, leading to a density gradient and nonuniform electric field concentrated around the obstacle. These density and potential gradients in turn influence the tunneling potential so that $U(x,V)$ must be determined self-consistently, and our treatment totally neglects such effects. We emphasize that the self-consistent determination of $U(x,V)$ is the central issue in any specific calculation of tunneling currents. An approximate method to self-consistently determine the tunneling potential at finite voltages, but which still assumes a uniform electric field, is given in Ref. 11. In this paper we will seek results which will not depend strongly on the exact shape of the tunneling potential, as long as the electron tunnels through some $U(x,V)$.

Electrons deep inside the two reservoirs are assumed to have a Fermi-Dirac distribution shown in Fig. 1. We assume the effect of an applied voltage between the two reservoirs is to create an imbalance in the electrochemical potentials deep inside the reservoirs equal to the applied voltage. This results in a current[14]

$$I(\mu,V,T)=e\int v^{+}(E)T(E,V)N^{+}(E) \times[f(E-(\mu+eV),T) -f(E-\mu,T)]dE\ , \tag{1}$$

where $v^{+}(E)$ is the group velocity for electrons moving in the positive x direction, $N^{+}(E)$ the density of states for electrons of both spins moving in the positive x direction in a one-dimensional free-electron gas, f is the Fermi function, T the temperature of both reservoirs, and

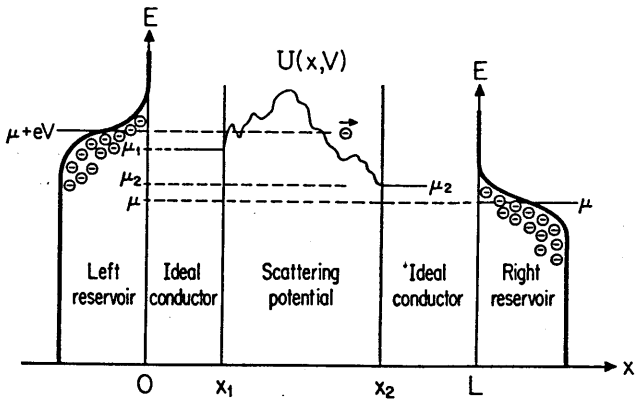


FIG. 1. Model for a one-dimensional quantum conductor: An elastic scattering potential $U(x,V)$ located between two thermal reservoirs. μ_1 and μ_2 are the electrochemical potentials in the left and right ideal conductors, respectively. An applied voltage V creates an imbalance between the electrochemical potentials $\mu_1=\mu+eV$ in the left reservoir and $\mu_2=\mu$ in the right

$T(E,V)$ the transmission coefficient through the potential $U(x,V)$. By enforcing time reversal symmetry and current conservation during tunneling, it is possible to show that the current transmission probability through the potential $U(x,V)$ is the same from left to right as from right to left, i.e., $T(E,V)=T_{lr}(E,V)=T_{rl}(E,V)$. While Eq. (1) is intuitive, it is also possible to derive it using the transfer Hamiltonian method and Fermi's golden rule.[12,14]

The product of the group velocity and the electron density of states in one dimension is a constant given by

$$v^{+}(E)N^{+}(E)=\left[\frac{1}{\hbar}\left|\frac{dE}{dk}\right|\right]\left[\frac{1}{\pi}\frac{1}{|dE/dk|}\right]=\frac{1}{\pi\hbar}\ . \tag{2}$$

The Fermi function in Eq. (1) can itself be expressed as a composite probability function

$$f(E-\mu,T)=[1-\theta(E-\mu)]\otimes\left[-\frac{df}{dE}(E,T)\right]\ , \tag{3}$$

where $\otimes$ denotes a convolution in energy, θ is the unit step function, and

$$-\frac{df}{dE}(E,T)=\frac{1}{4k_BT}\mathrm{sech}^2\left[\frac{E}{2k_BT}\right] \tag{4}$$

is the thermal smearing function or Fermi-Dirac probability density function. The convolution $\otimes$ of two function $A(E)$ and $B(E)$ has its usual meaning,

$$A(E)\otimes B(E)=\int_{-\infty}^{\infty}A(E-E')B(E')dE' =\int_{-\infty}^{\infty}A(E')B(E-E')dE'\ . \tag{5}$$

The difference of Fermi functions in Eq. (1) can therefore be written as

$$f(E-(\mu+eV),T)-f(E-\mu,T) =\{\theta(E-\mu)-\theta[E-(\mu+eV)]\}\otimes\left[-\frac{df}{dE}(E,T)\right]\ . \tag{6}$$

Equation (6) shows that the difference of Fermi functions in Eq. (1) can be expressed as the convolution in energy of two independent functions, one of which depends only on the applied voltage and the other which depends only on the temperature. Equation (6) therefore separates the effect of finite voltage on the electrical current from the effect of finite temperature.

Our expression for the current now becomes

$$I(E,V,T)=\frac{e}{\pi\hbar}T(E,V)\otimes W(E,V)\otimes\left[-\frac{df}{dE}(E,T)\right] \equiv\frac{e^2V}{\pi\hbar}\bar{T}(E,V)\ . \tag{7}$$

Equation (7), where $W(E,V)=[\theta(E+eV)-\theta(E)]$, is our interpretation of the finite-voltage Landauer formula in one dimension in terms of convolutions. $\bar{T}(E,V)$ is the energy averaged transmission coefficient. After carrying

graphically in Fig. 2. At this point we ignore the last two convolutions shown in the figure. If both eV and k_BT are small compared with any structure in the transmission coefficient $T(E,V)$, Eq. (7) reduces to the well-known result

$$I(E)=\frac{e^2}{\pi\hbar}T(E)V=\frac{V}{R_t} \tag{8}$$

which is Landauer's conductance formula in one dimension. Here R_t is the total resistance of the electron moving from one contact to the other. This formula was also obtained explicitly in Ref. 12 in a discussion of tunneling in one dimension between two thermal reservoirs.

Equation (7) has a simple physical interpretation: The convolution with the voltage window $W(E,V)=[\theta(E+eV)-\theta(E)]$ implies that the electron beam near the Fermi surface contributing to conduction is not monochromatic; it has a width in energy of eV. The box-like shape of $W(E,V)$ as shown in Fig. 2 argues that the electron beam near the Fermi surface has a uniform probability to have any energy between the electrochemical potential in the right contact and that energy plus eV. Equivalently, the average tunneling electron near the Fermi energy has a variance about its mean energy of $\Delta E\simeq eV$ due to the applied voltage.[16] The convolution with $-df/dE$ is just the standard thermal broadening of the conductance. Just as thermal broadening of the conductance makes it necessary to average the zero-temperature conductance over an energy range k_BT near the Fermi level, finite voltages require averaging the conductance over an energy scale eV near the Fermi level. The separate convolutions imply that thermal broadening and voltage broadening are statistically independent. Note also that the phase coherence time $\tau_V=\hbar/eV$ introduced by the nonmonochromatic electron beam at finite voltages is analogous to the better-known phase coherence time at finite temperatures $\tau_T=\hbar/k_BT$, both times

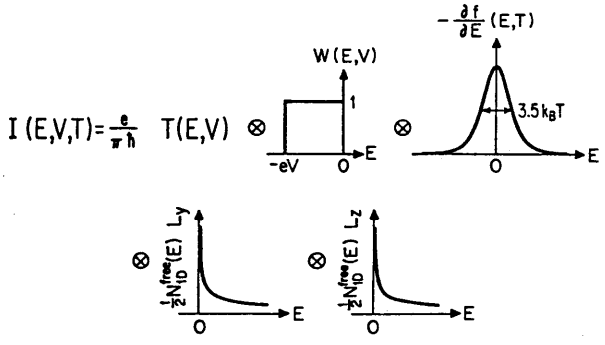


FIG. 2. Effect of finite temperature and finite voltage on the tunneling current displayed graphically as independent convolutions. Energy averaging due to finite voltages is included in the voltage window $W(E,V)$, while the energy averaging due to finite temperature results from the thermal smearing function $-df(E,T)/dE$. The two convolutions with the one-dimensional free-electron density of states relate tunneling current density in three dimensions to the one-dimensional tunneling being set approximately by the variances of $W(E,V)$ and $-df(E,T)/dE$, respectively.

Equation (8) predicts that a conductor having perfect transmission $T(E)=1$ has a contact resistance of $R_c=h/2e^2$. This result seems contrary to one's intuition from the Drude formula as well as from Landauer's original formula,[17] which gives the device resistance R_d as

$$R_d=\frac{h}{2e^2}\frac{R(E)}{T(E)}\ . \tag{9}$$

Here $R(E)$ is the reflection coefficient so that $T(E)+R(E)=1$. Büttiker[6] discusses the assumptions necessary to obtain Eq. (9). Both the Drude formula as well as Landauer's original formula, Eq. (9), assert that a length of perfect conductor has zero resistance.[17] It is now understood[9,10] that whether one obtains Eq. (8) or (9) in an actual experiment depends on how the resistance is measured. Connecting two weakly coupled voltage probes to the ideal conductors in Fig. 1, which measure the electrochemical potentials μ_1 and μ_2 as defined in Ref. 18, one obtains the original Landauer formula

$$V_d=(\mu_1-\mu_2)/e=IR_d=I(h/2e^2)(R/T)\ ,$$

which is Eq. (9). Landauer's original formula for R_d corresponds to the standard four-probe geometry for measuring resistance, in which the voltmeter is only weakly coupled to the sample one wishes to measure and no current flows into the voltmeter. If, on the other hand, one wishes to obtain the resistance using a two-probe geometry, which measures the electrochemical potentials $\mu+eV=\mu_l$ and $\mu=\mu_r$, one obtains

$$V_t=V=(\mu_l-\mu_r)/e=IR_t=I(h/2e^2)(1/T)$$

which is Eq. (8). In this case the voltmeter is the reservoir, which is strongly coupled to the system one wishes to measure. Due to the incoherence introduced by the reservoirs[5] in Fig. 1, the total resistance R_t is simply the sum of the device resistance R_d plus the contact resistance R_c, i.e., $R_t=R_d+R_c$. That R_c is in fact a contact resistance associated with the introduction of the two reservoirs in Fig. 1 was explained by Imry[10] in terms of the effusion of a gas through a small hole, and by Büttiker[9,6] who emphasized that carrier motion from one reservoir to the other was essential to obtain the contact resistance.

The contact resistance implied by Eq. (8) follows from the argument in Eq. (1) that the current is simply a product of the carrier charge, velocity, and density. For small temperatures and voltages this gives

$$I=ev_F^+n^+(E_F)=ev_F^+[N^+(E_F)eV]=\frac{e^2}{\pi\hbar}V\ ,$$

where $n^+(E_F)=N^+(E_F)eV$ is the carrier density near the Fermi level. As the Fermi level is raised, the Fermi velocity increases but the carrier density decreases. Equation (2) tells us that, in one dimension, these two factors exactly cancel each other for all energies. The contact resistance in Eq. (8) is then a consequence of the seemingly fortuitous cancellation in Eq. (2). We wish to

not rely on Eq. (2), and which emphasizes the way in which current is drawn out at a contact.

To gain further insight into the contact resistance and make a plausibility argument for Eq. (8), introduce the times

$$I=\frac{2e}{\tau_c+\tau_d}=\frac{e^2}{\pi\hbar}T(E)V \text{ and } I=\frac{2e}{\tau_c}=\frac{e^2}{\pi\hbar}V \tag{10}$$

which define the effective time of an electron in the device τ_d and the contact τ_c. One obtains for these times

$$\tau_c=\frac{2\pi\hbar}{eV}\simeq\tau_V \text{ and } \tau_d=\frac{2\pi\hbar}{eV}\frac{1-T(E)}{T(E)} . \tag{11}$$

The time τ_d corresponds to the original Landauer formula R_d while the time τ_c corresponds to the contact resistance R_c. If the conductor is in the classical diffusive limit, the time τ_d is proportional to the time for a classical particle to diffuse across the conductor.[19]

We can now make a plausibility argument for Eq. (8). We consider a single quantum ballistic channel having unity transmission so that the resistance R_d is negligible, i.e., $\tau_d\to 0$. That the conductance of the perfect wire by itself is infinite one can understand simply from the Drude formula, in which an impulse function electric field excites a current which flows forever without decaying. Our emphasis must be on understanding the resistance R_c, which we do by arguing for the time τ_c as follows: Consider the case of zero temperature. Far away from the interface between the device and the reservoir, the right contact must be described by a Fermi function having a single current-carrying state at the electrochemical potential μ_r. As the right contact rapidly draws out an electron at energy μ_r, another electron can occupy that energy level. It cannot, however, do so instantaneously. When the average tunneling electron near the Fermi surface reaches the right contact, it must dissipate on average an amount of energy $\Delta E=eV$ in that contact. Even if inelastic scattering processes occur on very rapid time scales in the contact, there is a limit imposed by the uncertainly principle for the rate at which energy can be dissipated into the thermal bath, namely $1/\Delta\tau=\Delta E/\hbar$. The time $\Delta\tau$ to dissipate an amount of energy $\Delta E=eV$ into the contact should be the same order of magnitude as the time τ_c to equilibrate with the measurement reservoir. This equilibration step will limit the current flow because, if the electron is out of equilibrium with the measurement reservoir, the reservoir cannot detect the electron. Therefore, only one electron of charge e can flow across the device every $\hbar/eV$ sec.[20] Imposing this rate-limiting process in the contacts, we argue that the electrical current even in this ballistic case is

$$I\simeq\frac{e}{\Delta\tau}\simeq\frac{e}{\hbar}\Delta E\simeq\frac{e^2}{\hbar}V , \tag{12}$$

which is the same order of magnitude as Eq. (8). This contact resistance corresponds to energy dissipation in the reservoirs, since no energy can be dissipated by elastic scattering in the sample. For the electron to dissipate an amount of energy eV in the reservoirs requires some time be transported into the right contact even in a perfectly ballistic device. Since the rate-limiting step for the current occurs in the contact, not in the device itself, the conductance formula (7) is independent of the length of the device as $T(E)\to 1$. The contact resistance is also plausible from the perspective that the voltage sources, which influence only the electrochemical potential deep inside the reservoirs, have no way of knowing that an electron has, in fact, transited the device until it comes into equilibrium with a Fermi function inside the reservoir. While the above argument is only heuristic, we believe it can be placed on firmer mathematical grounds by analogy with the onset of resistance in narrow superconductors via "phase slips."

The factor of $T(E)$ which enters into Eq. (8) can also be understood from an energy dissipation viewpoint: Only those electrons which in fact transit the device dissipate energy and wind up contributing to the resistance. We also note that the analog of Eq. (9) at finite voltages has been suggested in Ref. 11 to be $I/V_d=(2e^2/\hbar)(\bar{T}/\bar{R})$, where $\bar{T}$ is simply the energy averaged transmission coefficient defined in Eq. (7) and $\bar{T}+\bar{R}=1$.

Consider now a quasi-one-dimensional ballistic wire described by the separable potential $U(x,y,V)=U_x(x,V)+U_y(y)$, where $U_y(y)$ is a confining potential which quantizes the electron energies perpendicular to the direction of transport. The separable Hamiltonian implies that there is no scattering between the different subbands created by the confining potential $U_y(y)$. Equation (7) therefore generalizes, as we show in Sec. IV, to simply a sum over all the possible paths for electrons to transmit between the left and right contacts:

$$I(E,V,T)=\frac{e}{\pi\hbar}\sum_i T_i(E,V)\otimes W(E,V)\otimes\left[-\frac{df}{dE}(E,T)\right] , \tag{13}$$

where the sum is over all occupied subbands. For a ballistic quasi-one-dimensional wire with $U_x(x,V)\simeq 0$, we have

$$T_i(E,V)\simeq\theta(E-E_i) \tag{14}$$

at moderate voltages. Here E_i is the subband energy for quantized motion perpendicular to the direction of transport. The limiting case of this formula at zero temperature is when the thermal smearing function becomes a δ function so that

$$I(E,V,T=0)=\frac{e}{\pi\hbar}\sum_i\theta(E-E_i)\otimes W(E,V) . \tag{15}$$

We can carry out the convolution to obtain

$$\theta(E-E_i)\otimes W(E,V)=\begin{cases}0, & E<E_i-eV\\ E-(E_i-eV), & E_i-eV<E<E_i\\ eV, & E>E_i\end{cases} . \tag{16}$$

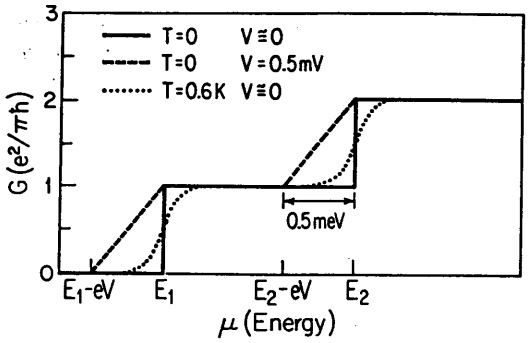


FIG. 3. Ballistic contact resistance in a quasi-one-dimensional quantum wire illustrating the effects of finite voltages and finite temperature. The conductance is shown versus the electrochemical potential μ in the right contact. The finite width of the steps in conductance is inversely proportional to the phase coherence time introduced by finite voltage or finite temperature.

dimensional wire with several subbands versus the electrochemical potential μ in the right contact in Fig. 3. The conductance G is defined as the current divided by voltage, even at finite voltages. These steps in the ballistic conductance have been observed[21,22] and explained using formalisms similar to the one in this paper as quantum contact resistances. The conductance at finite voltages, given by the dashed line in Fig. 3, shows initially a linear dependence versus Fermi energy instead of a step at zero temperature. We also show the effect of finite temperature on the conductance in Fig. 3 for comparison. Both finite temperature and finite voltage smear out the sharp steps in the zero-temperature conductance. The maximum slope of these steps in the conductance versus Fermi energy occurs at conductance values half-way between the plateaus, and is given by

$$\left.\frac{dG}{d\mu}\right|_{\max}=\frac{e^2}{\pi\hbar}\frac{1}{4k_BT}\simeq\frac{e^2}{\pi\hbar}\frac{\tau_T}{\hbar}\quad\text{for } V\simeq 0 \tag{17}$$

at finite temperature and

$$\left.\frac{dG}{d\mu}\right|_{\max}=\frac{e^2}{\pi\hbar}\frac{1}{eV}\simeq\frac{e^2}{\pi\hbar}\frac{\tau_V}{\hbar}\quad\text{for } T=0 \tag{18}$$

at finite voltages. The maximum slope of the steps in conductance versus Fermi energy is limited by the phase coherence times τ_T and τ_V, and is proportional to the phase coherence time. In a field effect transistor where the chemical potential is proportional to the gate bias voltage V_g, and if the chemical potential were controlled by such a gate, one could interpret $g_m=dI/dV_g=eV\,dG/d\mu$ as the transconductance of the device. For such a structure, the transconductance would have a finite maximum value even in the ballistic transport regime. When the applied voltage is much greater than the temperature, this maximum transconductance in the quasi-one-dimensional ballistic wire be-

III. CLASSICAL DIFFUSIVE LIMIT OF LANDAUER'S FORMULA

We are interested in the case with there are n scatterers randomly distributed in one dimension between two reservoirs as in Fig. 4. We consider the diffusion of classical particles down the chain of scatterers. Evaluating Landauer's formula for this case requires evaluating the transmission probability of a classical particle down a chain of n identical scatterers. This can be treated as a classical random walk down the n scattering sites. The classical transmission probability T_n down the total chain of n scatterers can be obtained in terms of a recursion relation

$$T_n=\frac{T_1T_{n-1}}{1-R_1R_{n-1}}\quad\text{where } T_n+R_n=1\ . \tag{19}$$

The solution of this recursion relation, which can be motivated from physical arguments involving the addition of classical resistors in series,[2] is

$$T_n=\frac{1}{(R_1/T_1)n+1}\ , \tag{20}$$

where T_1 is the probability of a classical particle being transmitted by a single scatterer. This result has been obtained in Ref. 2, and the special case with $T=R=\frac{1}{2}$ treated in Ref. 23 using a different method. Substituting this classical transmission coefficient into Landauer's conductance formula in one dimension, Eq. (8), gives the diffusive conductance

$$G_{1D}=\frac{e^2}{\pi\hbar}\frac{1}{(R_1/T_1)n+1}\ . \tag{21}$$

If we chose the scattering centers to be different, with scatterer a having a transmission coefficient T_a, etc., we obtain for the resistance R_{1D} after solving a recursion relation similar to Eq. (19)

$$\frac{e^2}{\pi\hbar}R_{1D}=\frac{R_a}{T_a}+\frac{R_b}{T_b}+\frac{R_c}{T_c}+\cdots+1\ , \tag{22}$$

which is just the classical sum of the original R/T Landauer resistance from each scatterer plus the contact resistance.

If we choose $T_a=R_a=\frac{1}{2}$ for each scatterer, corresponding to the assumption of isotropic scattering in the

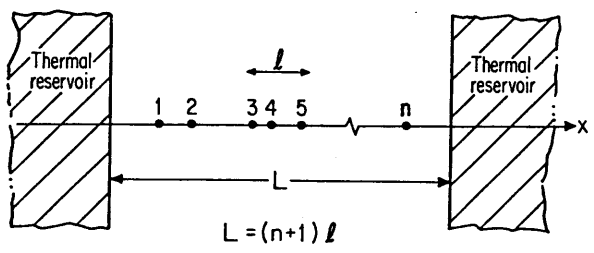


FIG. 4. n randomly positioned scatterers between two thermal reservoirs in one dimension. The two reservoirs are also regarded as scatterers so that the mean separation l between scattering centers is $L=(n+1)l$, where L is the sample

conductor, each elastic collision with a scatterer completely randomizes the particle's momentum direction. Our form for the classical diffusive conductance, Eq. (22), suggests that we interpret the contact resistance on the right-hand side as stating that the thermal reservoirs completely randomize the momentum direction of the electron when it enters the reservoir. The reservoir captures all particles incident upon it, then immediately randomizes their momentum direction due to rapid scattering inside the reservoir. This assumption is included in our derivation of the conductance formula (8) by assuming the reservoirs are described by a Fermi distribution.

We now wish to compare Eq. (21) to the standard Drude-Sommerfeld conductivity expression in one dimension,

$$\sigma_{1D}=e^2N(E_F)D(E_F)=e^2\frac{2}{\pi\hbar}v_F\tau_{tr}=2\frac{e^2}{\pi\hbar}l_{tr}\ , \qquad (23)$$

where $l_{tr}=v_F\tau_{tr}$ is the transport mean free path at the Fermi level and $D(E_F)=v_F^2\tau_{tr}$ is the diffusion coefficient. To obtain the conductance in one dimension, we merely divide by the length of the sample:

$$G_{1D}=\frac{\sigma_{1D}}{L}=2\frac{e^2}{\pi\hbar}\frac{l_{tr}}{L}\ . \qquad (24)$$

In order for the diffusive conductance from Landauer's formula, Eq. (21), to equal the Drude conductance in Eq. (24), we must conclude that

$$2\frac{l_{tr}}{L}=\frac{1}{(R_1/T_1)n+1}\ . \qquad (25)$$

To show that Eq. (25) is correct, consider the special case of isotropic scattering in the conductor. Here $T_1=R_1=\frac{1}{2}$, so that the transport mean free path is equal to the mean distance between the scatterers, i.e., $l_{tr}=l$. In that case Eq. (25) becomes

$$2\frac{l}{L}=\frac{1}{n+1}\ . \qquad (26)$$

In the limit that n becomes large, Eq. (26) implies that the mean free path is half the average spacing between scattering centers. We would obtain this result for the mean free path in one dimension if we average the distance to collide with a scatterer over a uniform distribution of all possible starting positions for the random walk. This statistical average is consistent with the standard assumption used in deriving the Drude conductivity, in which we consider an electron executing a random walk inside an "average" conductor that is spatially uniform. We can either keep the scatterers fixed and average over all possible starting positions for the random walk, or we can keep the origin of the random walk fixed and average over all possible impurity locations. Either way we obtain that the mean free path which should appear in the Drude formula, Eq. (24), is half the average spacing between scatterers. Consequently, for large n, Eq. (26) is correct, so that the Drude conductance is equivalent to the diffusive limit of the Landauer conductance. Moreover, even for small n, Eq. (26) can be shown the arrangement of the n scatterers between two thermal reservoirs as shown in Fig. 4. The reservoirs are regarded as being two additional scatterers which absorb all carriers incident upon them while simultaneously randomizing the carrier's momentum direction, so they must be included in computing the mean free path l. From the geometry of the scatterers we therefore have for a given scattering configuration that $L=(n+1)l$. The mean free path in the average configuration of the scatterers is therefore $L=2(n+1)l$ as in Eq. (26). We see again that the Drude conductance, given by Eq. (24), and Landauer's conductance formula, Eq. (8), in the classical diffusive limit are the same.

The Drude conductance formula and the Landauer conductance formula can also be made equivalent in the ballistic limit of no scattering between the thermal reservoirs ($n=0$), if we reinterpret the meaning of the transport mean free path which enters the Drude conductance in Eq. (24). The standard interpretation of the transport mean free path $l_{tr}=v_F\tau_{tr}$ is that τ_{tr} is the scattering time in an infinite system. In this standard interpretation l_{tr} can increase without bound. If we now reinterpret l_{tr} as the average distance for the conduction electron to randomize its momentum direction in a finite sample, we see that the longest possible transport mean free path for any given sample is just the sample length. For the average sample described by Eq. (24), this argument gives $2l_{tr}=2l=L$, in which case the ballistic conductance is

$$G_{1D}=\frac{e^2}{\pi\hbar} \qquad (27)$$

as before. The contact resistance can be thought of from this viewpoint as a consequence of the electron's transport mean free path being limited to at most the sample length.

We can gain further insight into the Drude-Sommerfeld conductance by considering a quantum particle moving in a set of potential barriers in one dimension, with barrier a having a quantum-mechanical transmission probability T_a, etc. In computing the total transmission probability of a quantum particle down this chain of scatterers one cannot ignore the phase of the particle. The total transmission coefficient down the chain will therefore be much more complicated than in Eq. (22). Büttiker[5] has worked out the general case for the addition of quantum resistors in series. In the case where the particle inelastically scatters with probability one between each elastic scatterer, his results can be written very simply as

$$\frac{e^2}{\pi\hbar}R_{1D}=\left[\frac{R_a}{T_a}+1\right]+\left[\frac{R_b}{T_b}+1\right]+\left[\frac{R_c}{T_c}+1\right]+\cdots+1\ . \qquad (28)$$

The quantum particle picks up the contact resistance each time it inelastically scatters. Equation (28) was derived for zero applied voltage. In the case of a finite voltage, we must average each transmission and reflection coefficient over the energy dissipated into the inelastic

the applied voltage divided by the number of additional inelastic scatterings inside the sample. If the number of inelastic scattering events increases with temperature, then the energy averaging due to voltage and temperature will be related, in contrast to the case where there is no inelastic scattering inside the device. Another interesting property of Eq. (28) is that the total transmission coefficient decreases linearly with the length of the 1D chain as expected if the motion is diffusive. It is generally believed that, for most types of disorder, the resistance of a 1D chain without inelastic scattering increases exponentially with the length of the chain.[24] Therefore, it appears that inelastic scattering processes are necessary to obtain diffusion in 1D. Equation (28), if compared with the Drude conductance in Eq. (24), gives Matthiessen's rule $l_{\rm tr}^{-1}=l_{\rm el}^{-1}+l_{\rm in}^{-1}$, where $L/2l_{\rm el}=(R_1/T_1)n$ for the case of identical barriers and $L/2l_{\rm in}=n+1$.

IV. TUNNELING LIMIT OF LANDAUER'S FORMULA IN TWO AND THREE DIMENSIONS

We consider here the generalization of Landauer's conductance formula to include many quantum channels or sets of good quantum numbers in the ideal device leads as described in Refs. 3 and 10. Define the current transmission probability T_{ij} from channel j in the left reservoir to channel i in the right reservoir evaluated at a given energy as

$$T_{ij}=\frac{J_{i,\rm right}}{J_{j,\rm left}}=t_{ij}t_{ij}^{*} \tag{29}$$

when all other incoming currents are zero. Here t_{ij} is the current transmission amplitude as defined in the scattering matrix.[10] The current from left to right, in the spirit of the derivation in Ref. 14, is therefore obtained as

$$I_{\rm lr}=\sum_{i,j} e\int v_j^{+}(E)N_j^{+}(E)T_{ij}(E,V) \times f_{\rm l}(E-(\mu+eV))[1-f_{\rm r}(E-\mu)]dE\ . \tag{30}$$

The current from right to left is similarly

$$I_{\rm rl}=\sum_{i,j} e\int v_i^{-}(E)N_i^{-}(E)T_{ji}(E,V)f_{\rm r}(E-\mu) \times[1-f_{\rm l}(E-(\mu+eV))]dE\ . \tag{31}$$

Here $v^{+}=v^{-}$ and $N^{+}=N^{-}=N/2$.

Current conservation along with time reversal symmetry give the constraint $t_{ij}=t_{ji}$, implying that $T_{ij}=T_{ji}$. The net current is therefore

$$I=I_{\rm lr}-I_{\rm rl} =\sum_{i,j}\frac{e}{\pi\hbar}\int T_{ij}(E,V)[f_{\rm l}(E-(\mu+eV))-f_{\rm r}(E-\mu)]dE\ . \tag{32}$$

Equation (32) reduces to the general result

$$I=\frac{e}{\pi\hbar}\sum_{i,j} T_{ij}(E,V)\otimes W(E,V)\otimes\left[-\frac{df(E,T)}{dE}\right], \tag{33}$$

where $W(E,V)$ is again the voltage window given by

$$W(E,V)=[\theta(E+eV)-\theta(E)]\ .$$

Equation (33) shows that temperature broadening and voltage broadening are statistically independent even in this multichannel case. Equation (33) reduces at low temperatures and small voltages to the well-known result[3,4,8]

$$I=\frac{e^2V}{2\pi\hbar}\sum_{i,j} T_{ij}=\frac{e^2V}{2\pi\hbar}\sum_{i,j} t_{ij}t_{ij}^{*}=\frac{e^2V}{2\pi\hbar}{\rm Tr}(tt^{\dagger})\ , \tag{34}$$

where the sum in Eq. (34) also runs over the individual spin quantum numbers, thus differing by a factor of 2 from Eq. (33). Equation (34) is analogous to the Landauer formula proportional to T, Eq. (8) describing a two-probe measurement. References 3, 6, and 10 give the multiple channel Landauer formula which describes a four-probe measurement, analogous to the T/R Landauer formula in Eq. (9).

For a general T_{ij}, Eq. (34) allows the possibility that electrons can scatter between different quantum channels. We are interested in the tunneling limit of this formula, for which $T_{ij}=T_i\delta_{ij}$. The current transmission probability takes this form only if the Hamiltonian is separable as $\hat{H}_{\rm total}=\hat{H}_x(x)+\hat{H}_y(y)+\hat{H}_z(z)$, thus justifying Eq. (13) in our earlier treatment of the ballistic quasi-one-dimensional wire. This tunneling limit corresponds to the classical addition of resistors in parallel, i.e., there is no communication between the parallel resistors. Because there is no scattering between the different conducting channels inside the sample, we only have to sum over one channel index in Eq. (34), namely,

$$I=\frac{e^2V}{2\pi\hbar}\sum_{i,j} T_i\delta_{ij}=\frac{e^2V}{2\pi\hbar}\sum_{i} T_i\ . \tag{35}$$

This limit also corresponds to the usual treatment of electron tunneling, for which we assume a static potential having spatial variation only along the x direction.[12,14,15] To prove this, note first that the Hamiltonian is separable so that $E=E_x+E_y+E_z$. Furthermore, the eigenfunctions along y and z are plane waves so that $E_y=\hbar^2k_y^2/2m$ and $E_z=\hbar^2k_z^2/2m$. The transverse channel index i can be labeled by $i=(k_y,k_z)$. Translational invariance along the y and z directions implies there is no scattering potential to couple the transverse channels. We enforce conservation of energy by writing

$$T_i(E_x)=T(E-E_y(k_y)-E_z(k_z))\ .$$

Expressing the multichannel Landauer formula, Eq. (35), in terms of a current density then gives

$$J_{\rm 3D}=\frac{I}{L_yL_z} =\frac{e^2V}{2\pi\hbar}\frac{1}{L_yL_z}\sum_{k_y,k_z,\rm spin} T(E-E_y(k_y)-E_z(k_z))\ . \tag{36}$$

Recognizing the one-dimensional identity

$$2\frac{1}{L_y}\sum_{k_y}\rightarrow\int N_{1D}^{free}(E_y)dE_y\ , \tag{37}$$

where $N_{1D}^{free}(E)$ is the one-dimensional free-electron density of states including spin

$$N_{1D}^{free}(E)=\frac{1}{\pi}\left[\frac{2m}{\hbar^2}\right]^{1/2}\left[\frac{1}{E}\right]^{1/2}\theta(E)\ , \tag{38}$$

the current density now becomes

$$J_{3D}=\frac{e^2V}{2\pi\hbar}\sum_{spin}\int\int\tfrac{1}{2}N_{1D}^{free}(E_y)\tfrac{1}{2}N_{1D}^{free}(E_z)\times T(E-E_y(k_y)-E_z(k_z))dE_y\,dE_z\ . \tag{39}$$

Equation (39) can be written as a convolution

$$J_{3D}=\frac{I}{L_yL_z}=\frac{e^2V}{\pi\hbar}T(E)\otimes\tfrac{1}{2}N_{1D}^{free}(E)\otimes\tfrac{1}{2}N_{1D}^{free}(E)\ . \tag{40}$$

If we had started from a standard tunneling formalism,[12,14] we would have obtained

$$J_{3D}(E,V,T)=\frac{e}{\pi\hbar}T(E,V)\otimes W(E,V)\otimes\left[-\frac{df}{dE}(E,T)\right]\otimes\tfrac{1}{2}N_{1D}^{free}(E)\otimes\tfrac{1}{2}N_{1D}^{free}(E) \tag{41}$$

as the current density at finite voltage and finite temperature. Equation (41) is shown graphically in Fig. 2. From Eq. (41) one can obtain any tunneling current at finite voltage and temperature found in Refs. 12, 14, and 15. A helpful identity in obtaining these tunneling currents is

$$\tfrac{1}{2}N_{1D}^{free}(E)\otimes\tfrac{1}{2}N_{1D}^{free}(E)=\tfrac{1}{2}N_{2D}^{free}(E)=\frac{1}{2}\frac{m}{2\pi\hbar^2}\theta(E)\ . \tag{42}$$

Results analogous to Eqs. (42) and (40) for the density of states and electrical currents follow for any separable potential.[13] Using Eqs. (42) and (40), we see that in the standard type of quantum-mechanical tunneling calculations in three dimensions, the quantum contact resistance never appears. The reason is, of course, that in going over to a continuum of states we add an infinite number of conducting channels in parallel. For tunneling at finite voltages, Eq. (33) reduces to the standard tunneling current density at finite voltages, Eq. (41), thus having the correct behavior in the tunneling limit.

By analogy with our consideration of classical diffusion in the previous section, we note that the multiple-channel Landauer formula has been applied to consider voltage fluctuations in small diffusive conductors.[25]

V. CONCLUSIONS

Using the Landauer formalism we have examined the effects of finite voltages on quantum transport, showing that the finite voltage acts to average the electrical conductance over an energy range equal to the applied voltage. This voltage broadening is similar to ordinary thermal broadening of the conductance, which averages the conductance over an energy range equal to the temperature. We showed that both thermal and voltage broadening of the conductance can be expressed as independent convolutions with known functions, proving that voltage broadening and ordinary temperature broadening are statistically independent. Motivated by the recent experiments of van Wees *et al.*[21] and Wharam *et al.*,[22] we examined thermal and voltage broadening of the conductance using a quasi-one-dimensional ballistic conductor as a model system. The finite width of the rises in the quantized steps in conductance versus electrochemical potential in these experiments is inversely proportional to the phase coherence time of the electron introduced by either finite voltage or finite temperature. The quantum contact resistances can be thought of either as a consequence of the rate limitation imposed by the uncertainty principle in the device leads or as a consequence that the mean free path is limited at most to the device size.

We also examined both the diffusive and tunneling limits of Landauer's conductance formula. In the classical diffusive limit the one-dimensional Landauer formula reduces to the Drude-Sommerfeld conductance in one dimension. In the tunneling limit, the finite-voltage Landauer formula is equivalent to many previous treatments of electron tunneling at finite voltages. In this tunneling limit, we have shown that simply by repeatedly convolving the tunneling current in one dimension with the one-dimensional free electron density of states, one obtains the tunneling current density in two and three spatial dimensions.

ACKNOWLEDGMENTS

We thank Dimitri A. Antoniadis, Henry I. Smith, and Khalid Ismail for useful discussions. We gratefully acknowledge financial support from U.S. Air Force of Scientific Research Contract No. AFOSR-88-0304.

[1]R. Landauer, IBM J. Res. Dev. **1**, 223 (1957).
[2]Rolf Landauer, Philos. Mag. **21**, 683 (1970).
[3]M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Phys. Rev. B **31**, 6207 (1985).
[4]Daniel S. Fischer and Patrick A. Lee, Phys. Rev. B **23**, 6851 (1981).
[5]M. Büttiker, Phys. Rev. B **33**, 3020 (1986). See also M. Büttiker, IBM J. Res. Dev. **32**, 63 (1988). Büttiker, IBM J. Res. Dev. **32**, 317 (1988). This last reference points out that Eq. (9) is not the most general four-probe resistance, even in the case of weak coupling to the voltage-measurement probes. A more general formula given by Büttiker considers possible wave-interference effects in the probes, as well as the possibility that the voltage probes couple differently to left- and right-moving carriers. In a four-probe–measurement geometry it is even possible to obtain

voltage probes.

[7]S. Datta, M. Cahay, and M. McLennan, Phys. Rev. B 36, 5655 (1987).

[8]A. Douglas Stone and Aaron Szafer, IBM J. Res. Dev. 32, 384 (1988).

[9]Rolf Landauer, Z. Phys. B 68, 217 (1987).

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

[11]D. Lenstra and R. T. M. Smokers, Phys. Rev. B 38, 6452 (1988).

[12]Walter A. Harrison, *Solid State Theory* (McGraw-Hill, New York, 1970). See Eqs. (3.23)–(3.25).

[13]Philip F. Bagwell, Dimitri A. Antoniadis, and Terry P. Orlando, in *Advanced MOS Device Physics*, edited by N. G. Einspruch and G. Gildenblat (Academic, New York, 1989). See also P. F. Bagwell and T. P. Orlando, Phys. Rev. B (to be published).

[14]E. L. Wolf, *Principles of Electron Tunneling Spectroscopy* (Oxford University Press, New York, 1985). See Eqs. (2.5) and (2.6).

[15]R. Tsu and L. Easaki, Appl. Phys. Lett. 22, 562 (1973).

[16]If the applied voltage is greater than the Fermi energy, then $W(E,V)$ must be cut off at the Fermi energy, namely,

$$W(E,V)=\begin{cases}[\theta(E+eV)-\theta(E)], & eV<\mu \quad (43a)\\ [\theta(E+\mu)-\theta(E)], & eV\geq\mu\ . \quad (43b)\end{cases}$$

[17]Rolf Landauer, Phys. Lett. 85A, 91 (1981).

[18]H. L. Engquist and P. W. Anderson, Phys. Rev. B 24, 1151 (1981).

[19]Use the Drude-Sommerfeld formula to describe the conductance of one electron of each spin in a device of length L_d by

$$I=\frac{2e}{\tau_d}=e^2D(E_F)N(E_F)\frac{V}{L_d}\ , \quad (44)$$

where V is the applied voltage and $D(E_F)=v_F^2\tau_F$ is determined by the Fermi velocity and momentum relaxation time. We later show that Eq. (44) is equivalent to Eq. (9) in the classical diffusive limit. The time τ_d is then given as

$$\frac{2}{\tau_d}=D\frac{N(E_F)eV}{L_d}=\left[\frac{D}{L_d^2}\right][N(E_F)eVL_d] \quad (45)$$

describing an electron gas in one dimension of arbitrary density $n=N(E_F)eV$. Thus τ_d is the time for one electron to diffuse across the device divided by the number of electrons in the device.

[20]This argument makes one wonder if it would be possible to see any indication of this frequency eV/h, which is half the Josephson frequency, in the current of a ballistic quasi-one-dimensional conductor as a function of time. Because of the spin degeneracy two electrons cross the conductor every h/eV sec.

[21]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).

[22]D. A. Wharam, T. J. Thorton, 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).

[23]R. Landauer and M. Büttiker, Phys. Rev. B 36, 6255 (1987).

[24]P. W. Anderson, D. J. Thouless, E. Abrahams, and D. S. Fischer, Phys. Rev. B 22, 3519 (1980).

[25]M. Büttiker, Phys. Rev. B 35, 4123 (1987).

Appendix E

Philip F. Bagwell, Tom P.E. Broekaert, Terry P. Orlando, and Clifton G. Fonstad, 'Resonant Tunneling Diodes and Transistors with a One, Two, or Three Dimensional Electron Emitter', Journal of Applied Physics, 68 4634, (1990).

Resonant tunneling diodes and transistors with a one-, two-, or three-dimensional electron emitter

Philip F. Bagwell, Tom P. E. Broekaert, T. P. Orlando, and Clifton G. Fonstad
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

(Received 25 January 1990; accepted for publication 22 June 1990)

We calculate the current versus emitter to collector voltage, current versus emitter Fermi energy, and current versus potential energy in the quantum well for double-barrier resonant tunneling devices having a one-, two-, or three-dimensional electron emitter. We consider both transistor and diode operation of the devices. For each device, the current is obtained using a method which describes the effects of finite temperature, finite voltage, and free-electron motion perpendicular to the tunneling direction as independent convolutions.

I. INTRODUCTION

Resonant tunneling diodes have improved steadily since the 1970's,[1] attaining frequency responses in the terahertz range[2] and peak-to-valley current ratios of over 30 at room temperature.[3] The development of fine lithography has led to other types of resonant tunneling devices, both vertical[4] and in the plane of a GaAs/GaAlAs heterojunction layer.[5–7] Many of these new types of electron devices based on resonant tunneling are designed to operate as a transistor,[8,9] and not merely as a diode, as was the case in earlier work.

Luryi[10] developed one way to understand the currents in resonant tunneling devices. He considered a three-dimensional electron emitter and, by enforcing momentum conservation perpendicular to the tunneling direction, concluded that only electrons lying on a disk in momentum space defined by the resonant wavelength contribute to the current. Liu and Aers[11] extended this picture to explain the current-voltage characteristics for electrons tunneling from a three-dimensional emitter (3DE) into a quantum well, quantum wire, and quantum dot. But recently fabricated resonant tunneling devices contain a two-dimensional emitter (2DE)[5–7] as well as a one-dimensional emitter (1DE).[4] While the standard picture of Luryi and its extensions can undoubtedly be used to understand the current-voltage characteristics of these devices, we believe that a complementary graphical method based on convolutions[12] is a powerful method for understanding these new types of resonant tunneling devices.

The goal of this paper is to apply our convolution method[12] to calculate the currents in resonant tunneling devices having a one-, two-, or three-dimensional electron emitter shown schematically in Fig. 1. In Sec. II we rewrite the standard formula for tunneling current[1] to clearly separate the effects of finite temperature, finite voltage, and an increase in the spatial dimension of the electron emitter in these devices. Finite voltage, finite temperature, and increasing the emitter spatial dimension enter the formula for tunneling current as a convolution in energy with a voltage broadening function, a thermal broadening function which is the derivative of the Fermi function, and the free-electron density of states, respectively. The effects of temperature, voltage, and the emitter dimensionality on the tunneling current can therefore be considered independently and in any order.

The separation of the effects of finite voltage, finite temperature, and the spatial dimension of the emitter in the formula for tunneling current, and the resulting intuitive picture of tunneling currents, is the main advantage of our method. The voltage convolution averages the device current over the electron energy distribution in the emitter, accounting for the emitter launching electrons over a range of different speeds. The thermal convolution accounts for the variance in these speeds. And the convolution with the free-electron density of states merely counts the number of resonant electrons, given that the emitter can launch electrons in many possible directions.

We evaluate these convolutions graphically to obtain the current as a function of Fermi energy in the emitter, as a function of bias across the device, and as a function of potential energy in the quantum well. In Sec. III we consider the resonant tunneling devices of Fig. 1 operated as transistors by sweeping the Fermi energy at low source-to-drain voltage. This simulates field-effect transistor operation. We

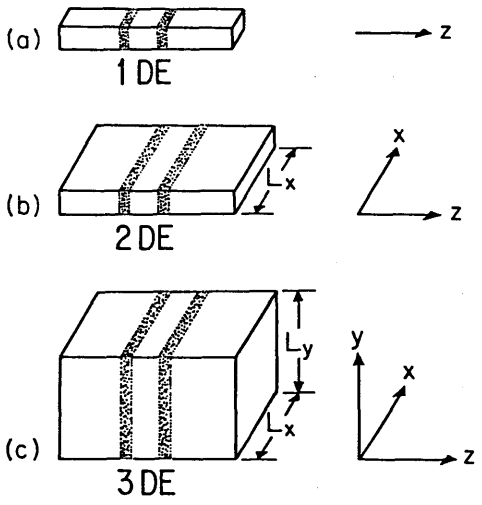


FIG. 1. Resonant tunneling from (a) a one-dimensional emitter (1DE), (b) a two-dimensional emitter (2DE), and (c) a three-dimensional emitter (3DE). The subbands created when confining the electrons are assumed to be widely separated compared with any other energy in the problem.

compare the transconductance of devices having different emitter dimensionality. In Sec. IV we consider the three resonant tunneling devices operated as diodes. The resonant tunneling peak has a shape characteristic of the emitter spatial dimension. In Sec. IV we also consider transistor operation resulting from modulating the potential energy of the quantum well. This simulates adding a base contact to the device. Our goal in this paper is not to obtain quantitative *I-V* curves for a particular device structure, but to understand how the general shape of the curves change for each different spatial dimension of the device emitter. In Sec. V we show how the currents in other possible resonant tunneling device geometries are related via our convolution method.

II. CONVOLUTION METHOD

In this section we give an intuitive introduction to our convolution method[12] using the devices in Fig. 1 as an illustrative example. The potential $U(z,V)$ in these devices depends only on the z coordinate as in Fig. 2, so that momentum perpendicular to the tunneling direction is conserved. The essential physical difference between the three devices in Fig. 1 is shown in Fig. 3. The 1DE can launch electrons only in the z direction, the 2DE can launch electrons in any direction with equal probability through an angle of π in the z-x half plane, and the 3DE can launch electrons in any direction with equal probability through a solid angle of 2π in the $z-x-y$ half space. We also display this as a graph in momentum space on the left-hand side of Figs. 4(a), 4(b), and 4(c), respectively.

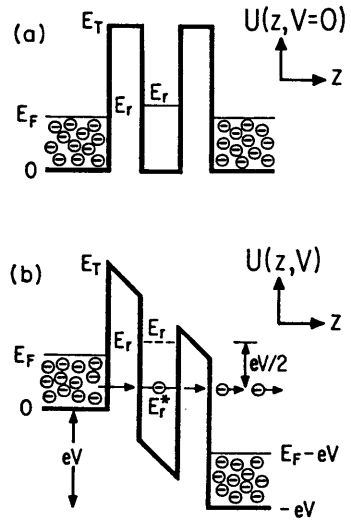


FIG. 2. Device potential $U(z,V)$ (a) under zero source to drain bias ($V \simeq 0$) and (b) under an applied bias V. The applied voltage creates an imbalance between the electrochemical potentials in the contacts equal to the applied voltage. The position of the resonant level E_r can be either above or below the Fermi level E_F in the emitter when $V = 0$. In (b) we have neglected the accumulation layer in the emitter and depletion layer in the collector which would be present in an actual device.

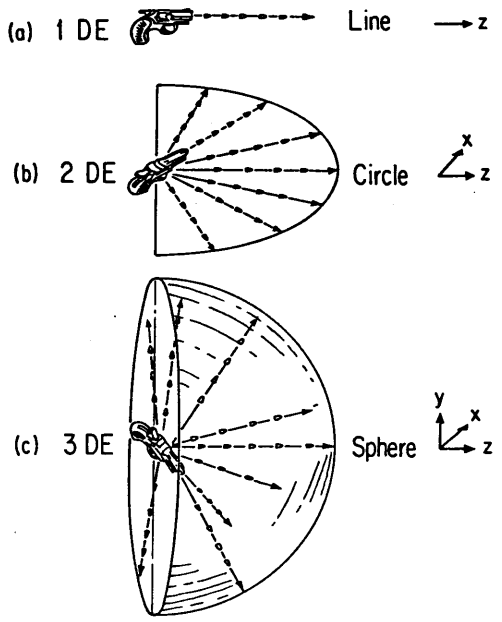


FIG. 3. Basic physical difference between (a) a one-dimensional electron emitter, (b) a two-dimensional emitter, and (c) a three-dimensional emitter. The 1DE can only launch electrons in the z direction. The 2DE can launch electrons at any angle with equal probability in the z-x half plane. The 3DE can launch electrons at any angle with equal probability in the z-x-y half space.

The current carried by each electron on the left-hand side of Fig. 4 depends only on its wave vector k_z along the z direction, and only resonant electrons with $k_z = k_r$ can contribute to the device current. Following Luryi,[10] we identify

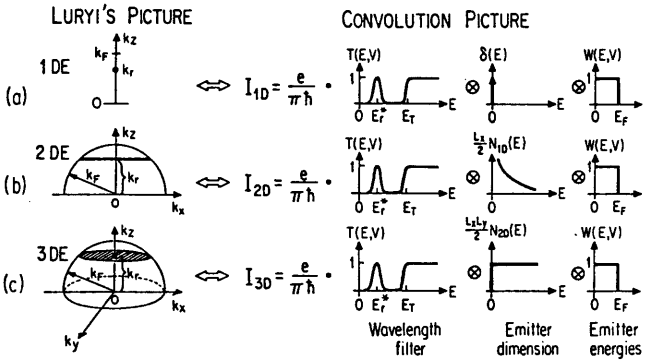


FIG. 4. Luryi's picture of resonant tunneling (left-hand side) compared with the convolution picture (right-hand side) for a resonant tunneling device with (a) a one-dimensional emitter (1DE), (b) two-dimensional emitter (2DE), and (c) a three-dimensional emitter (3DE). This figure is valid only for the case $E_r > E_F$. For the convolution method on the right-hand side of the figure, the resonance condition is contained in the transmission coefficient $T(E,V)$ in the first column. Convolution ($\otimes$) with the free-electron density of states in the second column, $\delta(E)$, $N_{1D}(E)$, and $N_{2D}(E)$, in (a), (b), and (c), accounts for the emitter launching electrons in all possible directions in one, two, and three dimensions. Convolution with $W(E,V)$ in the third column averages the conductance over the energy distribution of these electrons, that is, it accounts for the different possible energies with which the emitter can launch electrons in all the different directions. The independent convolutions in each of the three columns separates these three physical effects. Finite temperature is described by another convolution not shown in the figure.

these resonant electrons as a dot, a line, and a disk on the left-hand side of Fig. 4. The device current is simply proportional to the number of resonant electrons, so the problem of calculating the current reduces to computing the number of resonant electrons and their velocity along z given the specific electron distribution launched by the emitter. The standard method[1,10] for counting the number of resonant electrons requires computing the volume of the dot, line, and disk on the left-hand side of Fig. 4. Note that on the left-hand side of Fig. 4, the emitter is not merely launching electrons in all possible directions for each spatial dimension, but over the entire range of emitter energies from zero to the Fermi energy.

In the convolution method, shown graphically on the right-hand side of Fig. 4, we obtain the number of resonant electrons as follows: The current transmission coefficient $T(E,V)$ through the potential $U(z,V)$ contains the wavelength matching condition $k_z = k_r$. We display $T(E,V)$ in the first column on the right-hand side of Fig. 4. Convolution with the free-electron density of states in the second column on the right-hand side of Figs. 4(a), 4(b), and 4(c), either $\delta(E)$, $N_{1D}(E)$, or $N_{2D}(E)$, accounts for the emitter's ability to launch electrons at any angle in one, two, and three dimensions with equal probability, i.e., it accounts for the spatial distribution of the emitted electrons. Since this spatial distribution is different for each of the three devices in Fig. 1, only this second column is different for each device. The convolution with the energy distribution in the emitter, $W(E,V)$ in the third column, accounts for the emitter's launching electrons over a range of different energies or speeds. In the convolution method, the geometrical problem of counting the number of resonant electrons is handled automatically by convolving the transmission coefficient both with the energy distribution in the emitter and with either the one- or two-dimensional free-electron density of states.

Our convolution method relates the currents in each of the three devices in Fig. 1. The 1DE device current of Fig. 4(a) is obtained in the usual manner by evaluating the transmission coefficient $T(E,V)$. Since the 1DE device of Fig. 4(a) can only launch electrons in one direction, i.e., there is no "transverse energy," the second column is just a convolution with a delta function. The 1DE device current is found by convolving $T(E,V)$ with $W(E,V)$. Figure 4(b) shows that one may obtain the current of a 2DE device by convolving the 1DE device current with the one dimensional free electron density of states $N_{1D}(E)$. Convolution with $N_{1D}(E)$ accounts for the 2DE launching electrons at all possible angles in the $z-x$ half plane. Similarly, Fig. 4(c) asserts that one obtains the 3DE device current by convolving the 1DE device current with the two-dimensional free-electron density of states $N_{2D}(E)$. Convolution with $N_{2D}(E)$ accounts for the 3DE launching electrons at all possible angles in the $z-x-y$ half space. Although it is not shown in Fig. 4, the 2DE and 3DE device currents are also related via convolution with $N_{1D}(E)$.

We now summarize the equations to calculate tunneling current[12] for the devices of Fig. 1. After writing down these equations, we give an intuitive argument for them in the remainder of this section. Following Fig. 2(b), we assume the left- and right-hand device contacts are described by a Fermi function, and that the applied voltage induces a difference in electrochemical potentials between the two contacts by an amount equal to the applied voltage. We then compute the current as the product of charge, density, and velocity for each electron. These assumptions allow us to obtain the current in the 1DE device of Fig. 1(a) from a generalization of one of Landauer's[13] formulas

$$I_{1D}(E,V,T) = \frac{e}{\pi\hbar} T(E,V) \otimes W(E,V) \otimes \left(-\frac{df}{dE}(E,T) \right), \tag{1}$$

shown in Fig. 4(a). For the 2DE device of Fig. 1(b) we obtain the current from

$$I_{2D}(E,V,T) = I_{1D}(E,V,T) \otimes (L_x/2) N_{1D}(E), \tag{2}$$

shown in Fig. 4(b). Current in the 3DE device of Fig. 1(c) is

$$I_{3D}(E,V,T) = I_{2D}(E,V,T) \otimes (L_y/2) N_{1D}(E). \tag{3}$$

We argue below that Eq. (3) is the same as the equation shown graphically in Fig. 4(c).

We display Eq. (3) graphically in Fig. 5. In Fig. 5 and Eq. (3), $\otimes$ denotes a convolution in energy,[14] $T(E,V)$ is the current transmission coefficient through the one-dimensional potential $U(z,V)$ of Fig. 2, and $W(E,V)$ is the voltage broadening function:

$$W(E,V) = \begin{cases} \theta(E) - \theta(E-eV), & eV < E_F, \\ \theta(E) - \theta(E-E_F), & eV > E_F. \end{cases} \tag{4}$$

$\theta(E)$ in Eq. (4) is the unit step function.

The convolution with $W(E,V)$ simply averages the conductance over the energy distribution in the emitter. It describes the emitter launching electrons over a range of different speeds. Thus $W(E,V)$ has a maximum width set by the emitter Fermi level E_F as in Eq. (4), since E_F is the highest-energy electron available in the emitter. The third column on the right-hand side of Fig. 4 shows $W(E,V)$ for the case of $eV > E_F$, while Fig. 5 shows $W(E,V)$ when $eV < E_F$. The function $W(E,V)$ is slightly different from the one used in our earlier work,[12] since we take the zero of energy to lie at

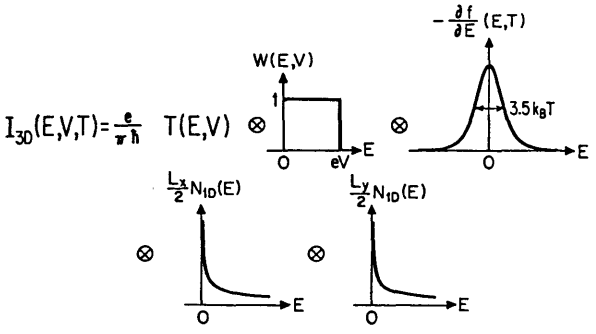


FIG. 5. Convolution method for calculating the 3DE device current corresponding to Eq. (3). The transmission coefficient $T(E,V)$ is successively convolved ($\otimes$) with $W(E,V)$, $-df(E,T)/dE$, and $N_{1D}(E)$. $W(E,V)$ is the emitter energy distribution at zero temperature. The separate convolution with $-df/dE$ includes the effects of finite temperature. The convolutions with the free-electron density of states relate current in the 1DE device to the 2DE and 3DE devices. $W(E,V)$ is shown for the condition $eV < E_F$.

the bottom of the zero temperature Fermi distribution in the emitter as shown in Fig. 2. This is a more convenient choice for calculations at large voltages.[15]

The thermal broadening function $-df/dE$ in Eq. (3) and Fig. 5 is the negative derivative of the Fermi function

$$-\frac{df}{dE}(E,T)=\frac{1}{4k_BT}\operatorname{sech}^2\left(\frac{E}{2k_BT}\right), \tag{5}$$

and describes the thermal variation in the speeds with which the emitter can launch electrons. Figure 4 is drawn for the case of zero temperature, so that the convolution with $-df/dE$ does not appear on the right-hand side of the figure. For the case of finite temperature the convolution with $-df/dE$ would enter as a fourth column on the right-hand side of Fig. 4, and would be the same for all three devices. In Eq. (3) and Fig. 5, $N_{1D}(E)$ is the free-electron density of states in one dimension

$$N_{1D}(E)=\frac{1}{\pi}\sqrt{\frac{2m}{\hbar^2}}\sqrt{\frac{1}{E}}\,\theta(E), \tag{6}$$

and L_x and L_y are the size of the device along the x and y directions drawn in Fig. 1.

We have previously shown[12] that Eq. (3) and Fig. 5 are equivalent to the standard methods used to calculate tunneling currents.[1] But Eq. (3) is conceptually and mathematically easier than the standard method since it separates the contributions to the device current from several different physical effects. *The convolutions are independent and can be done in any order, and each convolution incorporates different physics.* Free-electron motion perpendicular to the tunneling direction is incorporated through the convolutions with $N_{1D}(E)$, accounting for the spatial distribution of emitted electrons. Thermal broadening is included in the convolution with $-df/dE$. And voltage broadening effects are described through convolution with $W(E,V)$. The convolutions with $W(E,V)$ and $-df/dE$ together account for the total energy distribution of emitted electrons. These separate convolutions prove that thermal broadening is independent of voltage broadening, that the effect of adding free-electron motion along x is independent from adding free-electron motion along y, and that adding free-electron motion is independent of any voltage or temperature broadening. These conclusions are of course subject to the assumptions under which Eq. (3) is valid as we now discuss.

Equations (1), (2), and (3) describe electron motion through a separable potential $V(x,y,z)=V_x(x)+V_y(y)+V_z(z)$. This separable potential describes many but not all types of electron devices. For the devices of Fig. 1 the potential depends only on the z coordinate as in Fig. 2, so a separable potential with $V_x(x)=V_y(y)=0$ describes these devices. Equations (1), (2), and (3) also assume that the electron coherence length is long compared to the size of the confining well, although it is possible to remove this restriction.[16,17] The convolution with the free-electron density of states in Eqs. (2) and (3) is equivalent to the sum over electron channels in the multichannel Landauer formula.[18] The different channels correspond to the different spatial directions through which the emitter can launch electrons.[19]

The two convolutions describing free-electron motion perpendicular to the tunneling direction for the 3DE device of Eq. (3) can be evaluated by means of a simple physical argument.[20] Consider the case of free electrons in two dimensions. Since the potential is separable, electron motion along the z direction is independent of motion along x except for the constraint that the total energy $E=E_z+E_x$ is constant. The Schrödinger equation for electron motion along x can be solved for the density of states $N(E_x)$ available for motion along the x direction, and similarly for motion along the z direction. Therefore, the total density of states at a fixed energy E for any given energy E_z is just the product of the number of states available for motion along x times the number of states available for motion along z as

$$N_{1D}(E_z)N_{1D}(E_x)dE_z=N_{1D}(E_z)N_{1D}(E-E_z)dE_z.$$

But there are an infinite number of ways that the total energy E can be distributed between E_z and E_x. Thus, to find the total density of states at a fixed E, we must sum over all these possibilities as

$$N_{2D}(E)=\frac{1}{2}\int N_{1D}(E_z)N_{1D}(E-E_z)dE_z. \tag{7}$$

The extra factor of $\frac{1}{2}$ in the convolution compensates for counting the electron spin twice in N_{1D}. We can rewrite Eq. (7) as

$$\tfrac{1}{2}N_{1D}(E)\otimes\tfrac{1}{2}N_{1D}(E)=\tfrac{1}{2}N_{2D}(E)=\tfrac{1}{2}(m/\pi\hbar)\theta(E). \tag{8}$$

By a similar argument for free electrons in three dimensions,

$$\tfrac{1}{2}N_{2D}(E)\otimes\tfrac{1}{2}N_{1D}(E)=\tfrac{1}{2}N_{3D}(E)$$
$$=\frac{1}{2}\frac{1}{2\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\sqrt{E}\,\theta(E). \tag{9}$$

Equations (8) and (9) will allow us to easily evaluate the convolutions required to compute the currents for the devices in Fig. 1. We will also use the result[21]

$$\theta(E)\otimes\theta(E)=E\theta(E). \tag{10}$$

Using Eq. (8), one can thus proceed directly from the 1DE device current to the 3DE device current, as shown in Fig. 4(c), by convolving $I_{1D}(E,V,T)$ with the two-dimensional free-electron density of states $N_{2D}(E)$. Equation (3) can therefore be rewritten as

$$I_{3D}(E,V,T)$$
$$=I_{1D}(E,V,T)\otimes(L_x/2)N_{1D}(E)\otimes(L_y/2)N_{1D}(E)$$
$$=I_{1D}(E,V,T)\otimes(L_xL_y/2)N_{2D}(E), \tag{11}$$

as in Fig. 4(c).

The arguments used above for the density of states in each dimension are equally valid for the current, and can be used to argue for Eqs. (2) and (3). Consider the current in two dimensions. As above, $E=E_z+E_x$ since the potential is separable. Furthermore, the current carried by each electron depends only on its wave vector k_z, that is, it depends only on energy E_z for motion along the tunneling direction. Let the current carried by electrons having energy E_z in one dimension be $I_{1D}(E_z)$. The number of parallel one-dimensional channels which have this current in two dimensions is $L_xN_{1D}(E_x)dE_z/2$. To find the total current in two dimensions we must add all the possible ways to distribute the total energy E between motion along the tunneling direction and

motion perpendicular to the tunneling direction. We must also take into account the different number of parallel channels and different values of the current in each channel. Therefore, the current in two dimensions is

$$I_{2D}(E) = \int I_{1D}(E_z)\frac{L_x}{2} N_{1D}(E - E_z)dE_z, \qquad (12)$$

which is Eq. (2) and Fig. 4(b). A similar argument can be made for the current in three dimensions giving

$$I_{3D}(E) = \int I_{1D}(E_z)\frac{L_x L_y}{2} N_{2D}(E - E_z)dE_z, \qquad (13)$$

as in either Eq. (3), Eq. (11), or Fig. 4(c).

If the electron distribution were monochromatic, so that the width in energy of $W(E,V)$ was small, electrons would be launched at a point of constant energy in momentum space for a 1DE device, a half circular ring of constant energy in momentum space for a 2DE device, and a half-spherical shell of constant energy in momentum space for the 3DE device. This is also shown on the left-hand side of Fig. 4, if we consider only the electrons at a radius k_F. These constant energy "shells" at k_F correspond to the emitter launching electrons with a constant velocity, but in all possible directions with equal probability. Therefore, convolution with the free-electron density of states accounts for launching electrons at a constant energy but in all possible directions with equal probability. Convolution with the voltage broadening function $W(E,V)$ then adds up all the different energy shells or velocity shells, accounting for the emitter's launching of electrons over a range of energies set by the width of $W(E,V)$.

Having outlined our convolution method, the only remaining task is to obtain the transmission coefficient $T(E,V)$ as a function of energy and voltage for the double-barrier diode structure in Fig. 1. Several papers review the physics of resonant tunneling and determine the transmission coefficient.[16,22–24] We review the energy dependence of the transmission coefficient in Sec. III and the approximate voltage dependence in Sec. IV.

III. TRANSISTOR OPERATION

Most ideas for operating the resonant tunneling devices of Fig. 1 as transistors focus on the possibility of making direct contact to the well between the two tunnel barriers, as in Refs. 8, 9, and 25. We consider this transistor mode of operation after treating the operation of resonant tunneling devices as diodes in the next section. In this section we consider another possible way to operate the devices of Fig. 1 as transistors. Motivated by the field-effect transistor experiments in Ref. 5, we calculate the current as a function of emitter Fermi energy assuming the emitter to collector voltage V is small. This type of transistor operation can be realized either by incorporating a substrate gate into high electron mobility transistor (HEMT) devices, or by varying the Fermi energy with light-emitting diode (LED) light as was done in Ref. 5. In the 3DE device it would be difficult to control the Fermi energy with a gate, though one can imagine varying it by shining light on the device.

Our purpose in this section is to examine only the effect of the emitter spatial dimension on the tunneling current. We wish to exclude any effects due to a finite width in energy of the emitted electrons. We can consider these effects separately as explained in Sec. II. Thus, in this section we consider the dependence of the current on emitter Fermi energy assuming the applied voltage is small. In Sec. IV we can convolve the device currents in each spatial dimension obtained in this section with the electron distribution in the emitter, thereby calculating the diode currents when the voltage is large. We also work at zero temperature so that we can omit the convolution with $-df/dE$.

The energy diagram for the resonant tunneling device under zero bias is shown in Fig. 2(a). For simplicity we assume that only a single resonant level exists in the well. Under zero bias we take the resonant level to lie at energy E_r. We also assume the transmission probability $T(E)$ has a total area ΔE under its resonance peak, that is $\int T(E-)dE = \Delta E$, where the integral is over the resonance peak. If the emitter to collector bias is small, then the device current is simply proportional to the applied voltage V. We can then replace the convolution with $W(E,V)$ by multiplication with eV and can approximate $T(E,V) \simeq T(E,V=0) \equiv T(E)$. Thus, we assume that the energy E_r of the resonant level does not change significantly under applied bias if the voltage V is small. The energy scale ΔE defines small voltages as $eV \ll \Delta E$.

The tunneling probability $T(E)$ near a resonance has a Lorentzian shape[16,17,23] which we denote $L(E - E_r)$:

$$L(E - E_r) = \frac{\hbar^2 \Gamma_L \Gamma_R}{(E - E_r)^2 + (\hbar/2)^2(\Gamma_L + \Gamma_R)^2}. \qquad (14)$$

Here Γ_L and Γ_R are the partial leakage rates for the electron out of the quasi-bound state to the left and the right. For the transmission probability $L(E)$ we find $\Delta E = 2\pi\hbar\Gamma_L\Gamma_R/(\Gamma_L + \Gamma_R)$, while the full width at half maximum (FWHM) is just the sum of the partial leakage rates $\text{FWHM} = \hbar(\Gamma_L + \Gamma_R)$. For our purposes, we shall not be overly concerned with the exact shape of the resonance peak and can approximate

$$L(E - E_r) \simeq \theta(E - E_r + \Delta E/2) - \theta(E - E_r - \Delta E/2). \qquad (15)$$

This approximation assumes we will not try to resolve the exact shape of the resonance peak. In addition, if we do not attempt to resolve features on an energy scale of ΔE, we can approximate the resonant peak in $L(E)$ as a delta function whenever it occurs underneath an integral sign as

$$L(E - E_r) \simeq (\Delta E)\delta(E - E_r). \qquad (16)$$

The tunneling probability $T(E)$ through the potential $U(z,V=0)$ of Fig. 2(a) can now be well approximated as

$$T(E) \simeq L(E - E_r)\theta(E)\theta(E_T - E) + \theta(E - E_T), \qquad (17)$$

where E_T is the top of the potential barrier as in Fig. 2. We have sketched Eq. (17) for the transmission probability in the first column on the right-hand side of Fig. 4. This second piece of the transmission probability arises from nonresonant transmission above the double barrier. The approxima-

tion in Eq. (17) ignores the slight increase in transmission as the energy approaches the top of the barriers. Equation (17) also ignores any oscillations in $T(E)$ resembling the Ramsauer resonances, which occur when the Fermi energy is larger than the barrier height E_T. If we calculate $T(E)$ numerically and numerically apply our convolution method, we can remove these restrictions. However, the numerically calculated transmission probability will still qualitatively resemble Eq. (17).

The 1DE device current from Eq. (1) is proportional to the transmission coefficient

$$I_{1D}(E) = (e^2V/\pi\hbar)\,[L(E-E_r)\theta(E)\theta(E_T-E) + \theta(E-E_T)]. \tag{18}$$

We display this result graphically in Fig. 6(a). The quantum contact resistance of $h/2e^2$ implied[13] by this formula has recently been observed in a striking series of experiments.[26]

To calculate the 2DE device current as a function of emitter Fermi energy, we must convolve Eq. (18) with $N_{1D}(E)$ as in Eq. (2). We first make an intuitive argument for the resonant piece of the current following the reasoning leading to Eq. (12). Since only electrons with $E_z \simeq E_r$ contribute to the device current, we only have to evaluate the convolution at one value of E_z. The current at the resonant peak for the 1DE device from Eq. (18) is

$$I_{1D}(E_r) = e^2V/\pi\hbar. \tag{19}$$

This current can be pulled outside the integral in Eq. (12). The number of parallel one dimensional channels having energy near $E_z = E_r$ is given by taking $dE_z \simeq \Delta E$ in Eq. (12). The resonant piece of the current for the 2DE device is then found by multiplying the current carried by each one dimensional channel by the number of channels as

$$I_{2D}(E) = \frac{e^2V}{\pi\hbar}\frac{L_x}{2}(\Delta E)N_{1D}(E-E_r). \tag{20}$$

By similar arguments to the ones above combined with those leading to Eq. (13), the resonant piece of the current for the 3DE device is

$$I_{3D}(E) = \frac{e^2V}{\pi\hbar}\frac{L_xL_y}{2}(\Delta E)N_{2D}(E-E_r). \tag{21}$$

Again, this result is simply the current carried in each one-dimensional channel multiplied by the number of channels in parallel. The amount of current carried in each channel is limited by the quantum contact resistance in Eq. (19), so that the current in a large resonant tunneling device may be thought of as adding a large number of quantum contact resistances in parallel.

We now obtain the 2DE and 3DE device currents. Using the delta function approximation from Eq. (16) for the resonance peak, we carry out the convolution of Eq. (18) with $N_{1D}(E)$ to obtain the 2DE device current from Eq. (2) as

$$I_{2D}(E) = \frac{e^2V}{\pi\hbar}\frac{L_x}{2}\left(\frac{1}{\pi}\sqrt{\frac{2m}{\hbar^2}}\right) \times\left((\Delta E)\frac{\theta(E-E_r)}{\sqrt{E-E_r}}\theta(E_r) + 2\sqrt{E-E_T}\,\theta(E-E_T)\right). \tag{22}$$

The resonant piece of Eq. (22) agrees with Eq. (20). We display this result graphically in Fig. 6(b). To obtain the result for the nonresonant piece of the current in Eq. (22) we used Eq. (9).

We carry out the convolution of Eq. (18) with $N_{2D}(E)$ in Eq. (11) to obtain the 3DE device current. We again use the delta function approximation from Eq. (16) for $L(E-E_r)$ so that Eq. (11) gives

$$I_{3D}(E) = \frac{e^2V}{\pi\hbar}\frac{L_xL_y}{2}\left(\frac{m}{\pi\hbar^2}\right) \times[(\Delta E)\theta(E-E_r)\theta(E_r) + (E-E_T)\theta(E-E_T)]. \tag{23}$$

The resonant piece of Eq. (23) agrees with Eq. (21). We display this result graphically in Fig. 6(c). To obtain the result for the nonresonant piece of the current in Eq. (23) we used Eq. (10). We have included the step function $\theta(E_r)$ in the resonant terms of Eqs. (22) and (23) because the resonant piece of the transmission coefficient from Eqs. (16) and (17) is zero if $E_r < 0$. While it is not possible to have $E_r < 0$ for the potential we have drawn in Fig. 2(a), it is possible to imagine a potential in which the bottom of the well is below the zero of energy shown in the emitter. For that case we could have $E_r < 0$ and the resonant piece of the transmission probability would be zero.

We can sketch the energy dependencies of the three currents in Fig. 6 from physical arguments alone. One can argue physically for the energy dependence of the transmission coefficient needed to obtain the current in one dimension in Fig. 6(a). Furthermore, we can draw the energy dependence of the convolutions in Eqs. (8) and (9), which are also known from physical arguments,[20] to graph the energy dependencies for all three cases in Fig. 6. And as we show in the

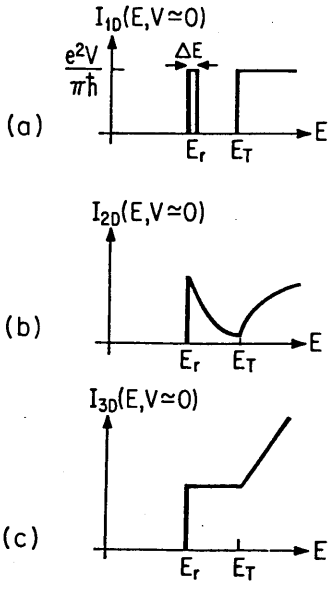


FIG. 6. Device current vs Fermi energy in the emitter for (a) the 1DE device, (b) the 2DE device, and (c) the 3DE device. These curves correspond to transistor operation of the devices, since E_F could be controlled by a gate voltage or by light. The curves in (b) and (c) follow directly from the curve in (a) by convolution with the one- and two-dimensional free-electron density of states, respectively.

next section, the energy dependencies from Fig. 6 determine the shape of the resonant peak when the devices are operated as diodes.

It is possible to incorporate a substrate gate into standard HEMT field-effect transistors which controls the Fermi energy. We can then interpret the energy axis of the graphs in Fig. 6 as being controlled by a transistor gate voltage. We can also interpret it as being controlled by LED light as argued at the beginning of this section. Figure 6 shows that the 1DE device has a strong negative transconductance, the 2DE device a weaker negative transconductance, while the 3DE device can at best show a zero transconductance. The 3DE and 2DE devices show positive transconductance when the emitter Fermi level is above the top of the potential barrier, but the 1DE device shows zero transconductance. This has recently been observed in GaAs HEMTs.[26] Only in the 1DE case does the current fall to zero when the device is operated as a transistor. This has been observed in Si inversion layers.[27]

Other modes of field-effect transistor operation are possible, in which the gate voltage controls the barrier height and width instead of the emitter Fermi energy. This has been demonstrated in a 2DE resonant tunneling transistor.[5] In this mode of transistor operation it may be possible to make the current fall to zero by raising the lowest resonant level either above the emitter Fermi energy or below the bottom of the emitter distribution. Negative transconductances were also observed in a 3DE device by controlling the well with a base contact.[8,9] We discuss this mode of transistor operation in the next section.

The temperature dependence of the current versus Fermi energy in Fig. 6 is calculated by convolving each of the three curves with the thermal smearing function $-df/dE$ as in Eq. (1). This is most easily done numerically. For small emitter-collector voltages, a convolution with the voltage broadening function $W(E,V)$ has an effect similar to the convolution with $-df/dE$ on the curves in Fig. 6. Both the thermal and voltage convolutions act to broaden the curves, leading to the criterion that any feature in the transmission coefficient at zero temperature and small voltage must have an energy spacing larger than about $3.5kT$ or eV to be observable at that temperature or voltage.[28] Note that the convolution with $-df/dE$ includes both "thermal broadening" of the resonance peak, as well as "thermionic emission" over the top of the barrier entering through the nonresonant piece of the current. The voltage convolution can never have an effect analogous to thermionic emission, since the function $W(E,V)$ has a sharp cutoff instead of the long Fermi tail present in $-df/dE$.

IV. DIODE OPERATION

In this section we consider the devices of Fig. 1 operated as diodes. That is, we calculate the current as a function of source to drain bias V. We obtain the diode curves in each dimension by averaging the transistor curves from Sec. III over the energy distribution in the emitter, as outlined in Sec. II.

The energy diagram for the resonant tunneling device operated as a diode is shown in Fig. 2(b). We again assume that only a single resonant level at energy E_r having area ΔE exists in the well under zero bias. We assume that the resonant level for our symmetric tunnel barriers moves down in energy an amount $eV/2$ under applied bias,[11] based on a simple geometrical argument using the device structure in Fig. 2(b). This conclusion seems to hold even if the position of the resonant level is calculated quantum mechanically including electrostatic self-consistency with the Poisson equation.[29] We also assume the resonance area ΔE does not significantly change under an applied bias.[30] These assumptions are valid as long as the structure remains approximately symmetrical under an applied bias, that is, for energy differences $E_T - E_r$ which are large compared to the applied voltage V. We also neglect the small Stark shift in the energy of the resonant level, as well as the accumulation in the emitter and depletion in the collector which would be present in a molecular-beam-epitaxy-grown device.[31] Thus, the following calculations are valid for large barrier heights, low-lying energy levels in the well, and moderate voltages.

Our simplified tunneling probability as a function of biasing voltage under these assumptions becomes

$$T(E,V) \simeq L(E - E_r^*)\theta(E)\theta(E_T - E) + \theta(E - E_T), \tag{24}$$

where $E_r^* = E_r - eV/2$. The transmission coefficient has the same form as before, but the applied voltage distorts the tunnel barriers so that the position of the resonant level depends on the applied voltage. The resonant term in the transmission probability is nearly zero when $eV/2 > E_r$, leading to a negative differential resistance in all three dimensions. As is well known, the negative differential resistance arises because the resonant level in the well falls below the electron distribution in the emitter, leaving no electrons which can tunnel resonantly.

The Fermi level E_F in the emitter is constant independent of V.[31] Therefore, it is not necessary to evaluate the convolution with $W(E,V)$ for all energies. We only evaluate this convolution at the Fermi energy. We also assume $E_F < E_T$. Using only the resonant portion of the current in Eqs. (18), (22), and (23), we first divide these currents by eV, replace E_r by E_r^*, then convolve with $W(E,V)$ at the Fermi energy to obtain

$$I_{1D}(E_F,V) = \frac{e}{\pi\hbar}(\Delta E)\int_{E_F - eV}^{E_F} \delta(E - E_r^*)\theta(E)dE \tag{25}$$

for the 1DE device,

$$I_{2D}(E_F,V) = \frac{e}{\pi\hbar}(\Delta E)\frac{L_x}{2}\left(\frac{1}{\pi}\sqrt{\frac{2m}{\hbar^2}}\right) \times \int_{E_F - eV}^{E_F} \frac{\theta(E - E_r^*)}{\sqrt{E - E_r^*}}\theta(E_r^*)dE \tag{26}$$

for the 2DE device, and

$$I_{3D}(E_F,V) = \frac{e}{\pi\hbar}(\Delta E)\frac{L_x L_y}{2}\left(\frac{m}{\pi\hbar^2}\right) \times \int_{E_F - eV}^{E_F} \theta(E - E_r^*)\theta(E_r^*)dE \tag{27}$$

for the 3DE device. In Eqs. (25), (26), and (27) we have

finally carried out the procedure outlined in Fig. 4. Again we have approximated the resonance peak as a delta function inside the integrals, which is valid assuming $eV > \Delta E$. We need the step function $\theta(E_r^*)$ in Eqs. (26) and (27) to enforce the condition from Eq. (25) that the current I_{1D} becomes zero when the resonant level E_r^* falls below the electron distribution in the emitter. The result when convolving zero current with the free-electron density of states must be zero. From Eqs. (25), (26), and (27) we can consider either the case where the resonant level is initially above or below the emitter Fermi level, that is $E_r > E_F$ or $E_r < E_F$.

A. Case of $E_r > E_F$

The simplest case to analyze is when the resonant level E_r lies above the Fermi energy E_F at $V = 0$, that is $E_r > E_F$. This is the standard case for resonant tunneling diodes grown using molecular-beam epitaxy. The resonant level initially lies above the Fermi level and moves down at a rate $eV/2$, while $W(E,V)$ initially opens at the Fermi level and grows wider at a rate eV. Yet if the initial width of the emitter distribution is only eV, how we are able to represent the device current as being averaged over the full width E_F of electron energies in the emitter as in Fig. 4?

To answer the question in the preceding paragraph graphically, we display the convolution of Eqs. (18), (22), and (23) with $W(E,V)$, as in Eqs. (25), (26), and (27), respectively, on the left-hand side of Fig. 7. We again ignore any structure on the scale ΔE. The lower limit of $W(E,V)$ at $E_F - eV$ is always below the energy threshold E_r^* for resonant current to flow, as seen in Fig. 7. Therefore, we can always extend the lower limit of $W(E,V)$ from $E_F - eV$ to 0 and obtain the same device currents. We conclude that the device currents for $E_r > E_F$ will be the same if we average the transmission probability over the full width E_F of electron energies in the emitter instead of the true width eV, even for $eV < E_F$ as in Fig. 4. The importance of this argument is that only in this special case of $E_r > E_F$ can we picture resonant tunneling currents in the standard way as on the left-hand side of Fig. 4.

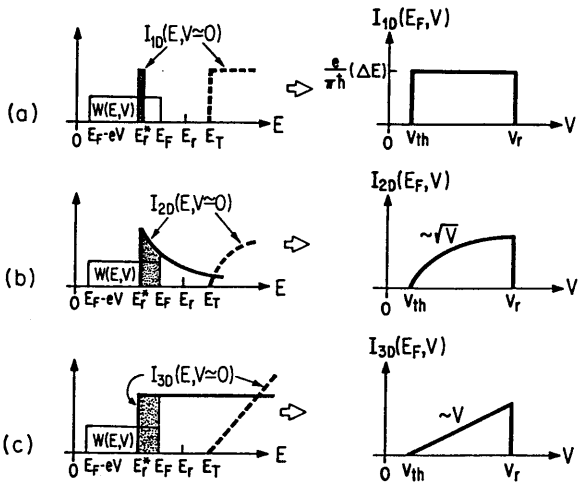


FIG. 7. Device current vs emitter to collector voltage assuming $E_r > E_F$ for (a) the 1DE device, (b) the 2DE device, and (c) the 3DE device. These curves correspond to operating the devices as a diode. The shaded areas on the left-hand side give the total diode current at each voltage V through a convolution of the transistor currents from Fig. 6 with $W(E,V)$ shown on the figure. The integrated area as a function of voltage follows as shown on the right-hand side, illustrating the connection between transistor and diode operation of the devices. The dotted lines on the left-hand side indicate the nonresonant piece of the transmission coefficient, which does not contribute to the device current on the right-hand side.

The shaded area underneath the small voltage current, $I(E,V \simeq 0)$ on the left-hand side of Fig. 7, corresponds to the total current of the resonant electrons after averaging $I(E,V \simeq 0)$ over the emitter distribution $W(E,V)$ for each value of V. The total current when the voltage is large is then simply the addition of the small voltage currents over an energy range set by $W(E,V)$. The solid lines on the left-hand side of Fig. 7 represent the resonant portion of the current which shifts down with the applied voltage. The dotted lines represent the nonresonant above barrier contribution to the current, which under our set of approximations does not shift in energy with the applied voltage and does not contribute to the diode current at $T = 0$. This is why we have neglected the nonresonant current in Eqs. (25), (26), and (27).

Integrating the resonant portion of the current from Eqs. (25), (26), and (27) we obtain the current magnitudes

$$I_{1D}(E_F,V) = (e/\pi\hbar)(\Delta E)\theta(|V| - V_{th})\theta(V_r - |V|) \quad (28)$$

for the 1DE device,

$$I_{2D}(E_F,V) = \frac{e}{\pi\hbar}(\Delta E)\frac{L_x}{2}\left(\frac{1}{\pi}\sqrt{\frac{2m}{\hbar^2}}\right)\left[2\sqrt{e(|V| - V_{th})/2}\right] \times \theta(|V| - V_{th})\theta(V_r - |V|) \quad (29)$$

for the 2DE device, and

$$I_{3D}(E_F,V) = \frac{e}{\pi\hbar}(\Delta E)\frac{L_x L_y}{2}\left(\frac{m}{\pi\hbar^2}\right)[e(|V| - V_{th})/2] \times \theta(|V| - V_{th})\theta(V_r - |V|) \quad (30)$$

for the 3DE device. In Eqs. (28), (29), and (30), the threshold voltage is $V_{th} = 2(E_r - E_F)/e$ and the resonance voltage is $V_r = 2E_r/e$. The current is symmetric about $V = 0$. We graph Eqs. (28), (29), and (30) on the right-hand side of Fig. 7. Below the threshold voltage V_{th} no current flows through the diode. The negative differential resistance occurs at the resonance voltage V_r.

The 1DE diode current in Fig. 7(a) is especially interesting because the current value is the product of fundamental constants $e/\pi\hbar$ times the area ΔE under the resonance peak. If such a diode could be fabricated, it would provide a macroscopic measurement of the quantum mechanical resonance width ΔE. The form of each diode current in two and three dimensions from Eqs. (29) and (30) is also easy to understand, because they again simply multiply the diode current in one dimension by the number of channels in the transverse dimension in an energy range $e(|V| - V_{th})/2$ available to carry current. Equations (29) and (30) therefore describe a continuum of one-dimensional diodes in parallel.

We now wish to argue for the shapes of the diode current in Fig. 7, knowing only that we must average the transistor characteristics from Fig. 6 over the energy distribution in the emitter. Looking at the left-hand side of Fig. 7, we see that the convolution with $W(E,V)$ requires us to add up all the resonant current from E_r^* to E_F. Integrating the delta-function-like shape of the 1DE device from Fig. 6(a) gives the constant voltage dependence for the 1DE diode current in Fig. 7(a). Integrating the inverse square root of energy dependence of the 2DE device from Fig. 6(b) gives the square-root dependence of the 2DE diode current on voltage in Fig. 7(b). And integrating the constant energy dependence of the 3DE device from Fig. 6(c) gives the linear voltage dependence of the 3DE diode current in the voltage in Fig. 7(c). We can therefore obtain the shape of the diode characteristics on the right-hand side of Fig. 7 by graphically averaging the transistor characteristics on the left-hand side of Fig. 7 over the emitter energy distribution. Figure 7 therefore shows how transistor operation (sweeping the Fermi energy) and diode operation (sweeping the source-drain voltage) are connected via our convolution method.

Although we have already physically argued for the shapes of the diode current in Fig. 7 using the transistor currents in Fig. 6 (which we physically argued for in the previous section), we wish to give yet another way to understand the shapes of the diode current in different spatial dimensions. Examining the diode current for each of the three emitter dimensions will show how our convolutions with the free-electron density of states are again merely counting the number of resonant electrons in each dimension.

The current for a 1DE device is given by counting the charge, times the velocity of the resonant electrons v_r, times the density of resonant electrons n_r.[12] Since the resonance width is narrow compared to the applied voltage, the number of resonant electrons for the one dimensional device can be obtained by multiplying the density of states at the resonance energy by the area ΔE under the resonance peak in $T(E,V)$,

$$I_{1D} = ev_r n_r^+ = ev_r(E_r^*)\left[N_{1D}^+(E_r^*)\Delta E\right]. \tag{31}$$

In one dimension $v_r(E)N_{1D}^+(E) = 1/\pi\hbar$ for all E, where $N_{1D}^+(E) = N_{1D}(E)/2$ is the density of states for electrons moving only along the positive z direction, so we recover the diode current[32]

$$I_{1D} = (e/\pi\hbar)(\Delta E) \tag{32}$$

as in Eq. (28). Because of the separable potential $U(z,V)$, we can count the density of resonant electrons in two and three dimensions as in Secs. II and III. The only difference here is that, instead of finding the number of resonant electrons at an energy E, we wish to find the density of resonant electrons in a range of energies between E_r^* and E_F. Therefore, the density of electrons in two dimensions by analogy with our arguments in Secs. II and III is

$$n_r^+ = N_{1D}^+(E_r^*)(\Delta E)\int_{E_r^*}^{E_F}\frac{L_x}{2}N_{1D}(E-E_r^*)dE. \tag{33}$$

We obtain the diode current in two dimensions as

$$I_{2D} = \frac{e}{\pi\hbar}(\Delta E)\frac{L_x}{2}\left(\frac{1}{\pi}\sqrt{\frac{2m}{\hbar^2}}\right)2\sqrt{E_F-E_r^*}, \tag{34}$$

agreeing with Eq. (29). The number of resonant electrons in three dimensions we find from

$$n_r^+ = N_{1D}^+(E_r^*)(\Delta E)\int_{E_r^*}^{E_F}\frac{L_xL_y}{2}N_{2D}(E-E_r^*)dE, \tag{35}$$

yielding the current

$$I_{3D} = \frac{e}{\pi\hbar}(\Delta E)\frac{L_xL_y}{2}\left(\frac{m}{\pi\hbar^2}\right)(E_F-E_r^*), \tag{36}$$

consistent with Eq. (30).

We wish to return here to the origin of the negative differential resistance in Fig. 7 when $V = V_r$. From the left-hand side of Fig. 7 it looks as though in two and three dimensions there will still be resonant electrons even when $V > V_r$. Mathematically, it is clear from Fig. 7(a) that the diode current in one dimension will be zero when $V > V_r$. Thus, convolving the one-dimensional device current with the free-electron density of states must give zero when $V > V_r$. Physically, in two and three dimensions, even though there are still electron states available in the well for the range of emitter energies $0 < E < E_F$ when $V > V_r$, the emitter cannot put electrons into these states and simultaneously satisfy all the required conservation laws. In two dimensions for example, the total energy E must be conserved and $E = E_z + E_x$. Because of translational invariance along the x direction, the momentum k_x must be conserved. Therefore, E_x must be conserved. But if E and E_x must be conserved during tunneling, then the energy E_z along the tunneling direction must also be conserved. The only range of electron energies E_z available in the emitter are between zero and the Fermi energy. Therefore, if the resonant level moves outside this energy range, the diode current must fall to zero. If there are defects in the diode such that momentum perpendicular to the tunneling direction no longer must be conserved, there can be a finite diode current even when the resonant level falls below the electron energy distribution in the emitter in two and three dimensins.

B. Case of $E_r < E_F$

We now consider the case $E_r < E_F$, where the resonant level is initially below the Fermi level. Beginning with one dimension, no current can flow until the applied voltage is at least $eV = (E_F - E_r)$. But by the time we have applied this much voltage, the resonant level has moved down to a position $E_r^* = E_r - (E_F - E_r)/2$ and so on. The diode threshold voltage V_{th} in one dimension for $E_r < E_F$ is therefore found from the infinite series

$$eV_{th} = (E_F - E_r)(1 + \tfrac{1}{2} + \tfrac{1}{4} + \tfrac{1}{8} + \cdots) = 2(E_F - E_r). \tag{37}$$

Also, the resonant level must be at least $E_F - E_r$ above the bottom of the conduction band when $V = 0$, or else the level will fall below the bottom of the conduction band before the applied voltage reaches the threshold voltage. In other words, if the resonant level is lower than half the Fermi level there will be no resonant diode current in one dimension. This naturally divides the case $E_F > E_r$ into two subcases:

$E_F > E_r > E_F/2$ and $E_r < E_F/2$.

We display the integrands of Eqs. (25), (26), and (27) graphically on the left-hand side of Fig. 8 for the case $E_F > E_r > E_F/2$ and the bias condition $V < V_{th}$. As in Fig. 7, the shaded area in Fig. 8 corresponds to the total current of the resonant electrons. Note there are no resonant electrons in one dimension for $V < V_{th}$ in the figure. The threshold voltage corresponds to the emitter energy distribution overtaking the resonant level as it moves down in energy. When the applied voltage is greater than the threshold voltage, these energy diagrams on the left-hand side of Fig. 8 will look almost exactly the same as the ones on the left-hand side of Fig. 7. In the case where $E_r > E_F$ described in Fig. 7, the emitter energy distribution always extended below the resonant level. Thus we expect the same dependence of the current on the applied voltage for both the case $E_r > E_F$, and for the case $E_r < E_F$ when the applied voltage is greater than the threshold voltage.

Integrating Eqs. (25), (26), and (27) directly for $E_F > E_r$ we find

$$I_{1D}(E_F,V) = (e/\pi\hbar)(\Delta E)\theta(|V| - V_{th})\theta(V_r - |V|) \quad (38)$$

for the 1DE device,

$$I_{2D}(E_F,V) = \frac{e}{\pi\hbar}(\Delta E)\frac{L_x}{2}\left(\frac{1}{\pi}\sqrt{\frac{2m}{\hbar^2}}\right) \times\{[2\sqrt{e(V_{th}+V)/2}]\theta(V_{th}+V) - [2\sqrt{e(V_{th}-V)/2}]\theta(V_{th}-V)\} \times\theta(V_r - |V|) \quad (39)$$

for the 2DE device, and

$$I_{3D}(E_F,V) = \frac{e}{\pi\hbar}(\Delta E)\frac{L_x L_y}{2}\left(\frac{m}{\pi\hbar^2}\right) \times\{[e(V_{th}+V)/2]\theta(V_{th}+V) - [e(V_{th}-V)/2]\theta(V_{th}-V)\} \times\theta(V_r - |V|) \quad (40)$$

for the 3DE device. Equation (38) is a current magnitude while Eqs. (39) and (40) are signed currents.

We display the three diode currents from Eqs. (38), (39), and (40) graphically for $E_F > E_r > E_F/2$ on the right-hand side of Fig. 8. In contrast to the 1DE diode, a nonzero current flows for voltages below the threshold voltage for the 2DE and 3DE diodes when $E_r < E_F$. This can be inferred from the left-hand side of Fig. 8, where there will always be some resonant electrons in the 2DE and 3DE diodes for small V. The 1DE device current in Eq. (38) is always zero if the resonant level is below half the emitter Fermi level.

The form of Eqs. (39) and (40) motivate the following useful construction of the diode currents for $E_F > E_r$: The two terms in Eqs. (39) and (40), shown as dotted lines on the right-hand side of Fig. 8, have exactly the same functional form as Eqs. (29) and (30) from our previous case of $E_r > E_F$. The difference of these two dotted lines gives the diode current in the case of $|V| < V_{th}$. These two dotted lines may be thought of as the "forward" and "reverse" currents with their difference giving the net current. The forward current is the current from the emitter assuming there are no filled electron states in the collector and vice versa. For our previous case of $E_r > E_F$ the reverse current was zero and did not contribute.

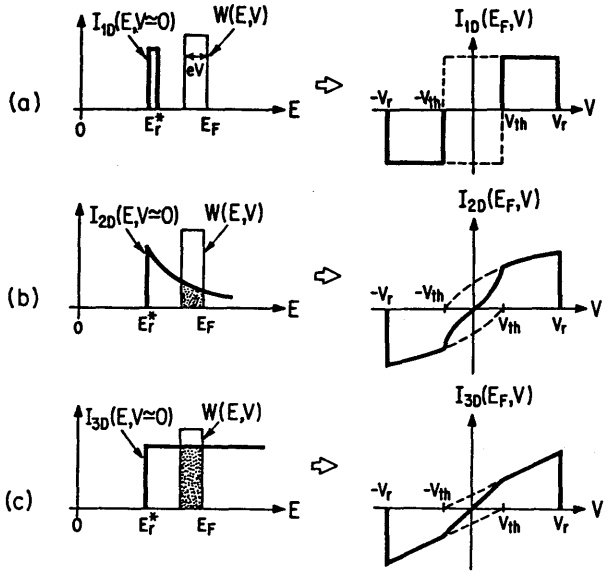


FIG. 8. Device current vs emitter to collector voltage assuming $E_F/2 < E_r < E_F$ for (a) the 1DE device, (b) the 2DE device, and (c) the 3DE device. The left-hand side of this figure, similar to Fig. 6, is shown for the bias condition $V < V_{th}$. The left- and right-hand sides of the figure again illustrate the connection between transistor and diode operation of the devices. The shaded areas on the left-hand side again add up the total diode current for each value of V. The diode current when $V < V_{th}$ may be pictured by subtracting the dotted curves shown on the right-hand side, corresponding to subtracting the "forward" and "reverse" diode currents. In (c) describing the 3DE device, the slope of the I-V curve changes by a factor of 2 at V_{th}.

The diode current in Fig. 8 increases faster with voltage when $|V| < V_{th}$ than it does above the threshold voltage, there being a factor-of-2 difference in the slopes in three dimensions. The graphical construction in Fig. 8 includes both subcases $E_F > E_r > E_F/2$ and $E_r < E_F/2$, only in the case $E_r < E_F/2$ the negative differential resistance voltage occurs before the threshold voltage. This construction can also be done in one dimension as shown in Fig. 8(a). And similar to the discussion following Eqs. (28), (29), and (30) the diode currents in two and three dimensions can be thought of as arising from adding many one-dimensional diodes in parallel. The difference for the case of $E_r < E_F$ is that we must subtract the reverse diode current from the forward current in each one-dimensional diode.

C. Base contact to the quantum well

In this section we explore a simple model for operating the devices in Fig. 1 as transistors by adding an additional base contact to the quantum well. We assume the base current is zero, and that the base acts only to alter the shape of the tunneling potential. We show that a strong negative transconductance is possible for all three devices if one can contact the quantum well. In contrast, for the mode of transistor operation in which we sweep the Fermi energy as in

Sec. III, only the 1DE device showed a strong negative transconductance.

Under the same caveats as at the beginning of our discussion of the diode operation, we can model the effect of the additional base contact as simply shifting the position of the resonant level linearly with the base voltage as

$$E_r^* = E_r - eV_B, \tag{41}$$

where V_B is the emitter to base voltage. Note that an applied base voltage can shift the position of the resonant level to either a higher or lower value in energy. The emitter to collector bias V no longer has any effect on the energy of the resonant level, and simply determines the width of $W(E,V)$ as before. Equations (25), (26), and (27) still describe the effect of the base contact if we use Eq. (41) to determine the position of the resonant level E_r^*.

For the case of large bias across the device so that $eV > E_F$, the full electron distribution in the emitter participates in conduction as in Fig. 7. Therefore, from Eq. (41), the diode curves on the right-hand side of Fig. 7 can be reinterpreted as current-versus-base voltage if we rescale the voltage axis by a factor of 2 and similarly change the threshold and resonance voltages. We graph the current-versus-base voltage when $eV > E_F$ on the left-hand side of Fig. 9. In this figure the threshold voltage $V_B = V_{th}$ occurs when $E_r^* = E_F$ so that $eV_{th} = E_r - E_F$. The resonance voltage $V_B = V_r$ occurs when $E_r^* = 0$ and is given by $eV_r = E_r$. Note that the threshold voltage for the base can be either positive or negative. The number of resonant electrons can be pictured in the same way as on the left-hand side of Fig. 7.

When the applied bias is smaller than the Fermi energy, $eV < E_F$, we can obtain the current-versus-base voltage by a construction similar to the one we used in the diode in Fig. 8 when $E_r < E_F$. We shown this construction on the right-hand side of Fig. 9, subtracting the dotted lines to obtain the total current. The threshold and resonance voltages are the same as the case of large bias, but in addition there is a "break" voltage $V_B = V_{br}$ defined when $E_r^* = E_F - e|V|$ giving $V_{br} = V_{th} + |V|$. This break voltage in the transistor currents only occurs when the applied voltage is less than the Fermi energy. We do not write down any equations for the currents shown in Fig. 9, as they are easily obtained from the ones found in our treatment of the diode earlier in this section.

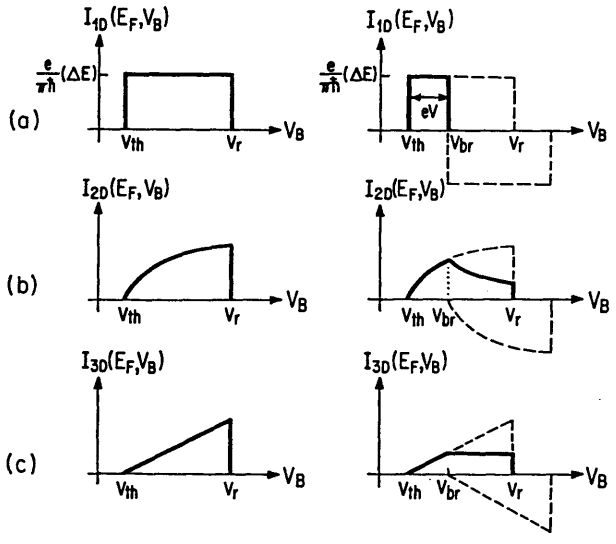


FIG. 9. Transistor characteristics for the three resonant tunneling devices having a base contact to the quantum well. We show the device current vs base voltage V_B for (a) the 1DE device, (b) the 2DE device, and (c) the 3DE device. Our calculation assumes the resonant level moves linearly with base voltage and that the base current is zero. The left-hand side shows the transistor characteristics when the emitter-to-collector bias is greater than the Fermi energy, $eV > E_F$. The curves on the right-hand side are for $eV < E_F$.

V. OTHER RESONANT TUNNELING DEVICES

In this section we examine other possible resonant tunneling devices in addition to the ones in Fig. 1. Our purpose in this section is to show how currents in these other possible resonant tunneling devices are related through a convolution with the free-electron density of states. Both the emitter spatial dimension and the degree of confinement in the quantum well can vary. This gives a total of nine possibilities as shown in Fig. 10. The columns of Fig. 10 hold the emitter dimension fixed while the rows hold the degree of confinement in the quantum well constant.

In this paper we have shown how the convolution method connects the conductance of devices along the main diagonal of Fig. 10, as indicated by the arrows in the figure. The top left-hand corner of Fig. 10 is the usual resonant tunneling diode grown by molecular-beam epitaxy. We label this device the 3DE/1DC for its three-dimensional emitter (3DE) and one-dimensionally confined (1DC) quantum well. We proceeded in this paper from the 1DE/3DC device by convolution to the 2DE/2DC device, and similarly from the 2DE/2DE device by convolution to the standard 3DE/1DC device.

The conductance of the devices in the left-hand column of Fig. 10, the 3DE/1DC, 3DE/2DC, and 3DE/3DC de-

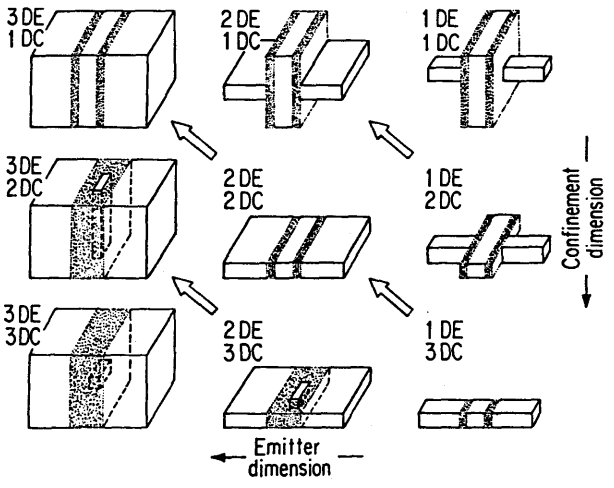


FIG. 10. Nine possible combinations of resonant tunneling devices having different emitter dimensions (1DE, 2DE, and 3DE) and different confinement dimensions in the quantum well (1DC, 2DC, and 3DC). The convolution method relates the conductance of the devices along the diagonals, as indicated by arrows in the figure.

vices, was calculated in Ref. 11. This was called "tunneling into a quantum well, wire, and dot," respectively in Ref. 11. The transmission coefficient for the 3DE/3DC device in the lower left-hand corner was also calculated in Ref. 33 and found to have a Lorentzian shape. It is unclear whether these two calculations[11,33] are in agreement. The case of resonant tunneling in two dimensions, the 2DE/3DC device, was studied in Ref. 34. The transmission coefficient was also found there to have a Lorentzian shape. It should be possible to relate this calculation[34] for two-dimensional tunneling into a quantum dot (2DE/3DC) to the case of three-dimensional tunneling into a quantum wire (3DE/2DC) in Ref. 11 using our convolution method. This is shown by an arrow along a diagonal in Fig. 10. Büttiker[16] has also argued that the Lorentzian shape for the transmission probability of the 3DE/3DC device in Ref. 33 and the 2DE/3DC device in Ref. 34 must follow from the Breit–Wigner formula. And the transmission coefficient of the 1DE/3DC device is already known to have a Lorentzian shape[23] near the resonance peak. Thus, all the devices along the bottom row have Lorentzian-shaped transmission coefficients.

The devices in the upper right-hand corner of Fig. 10 have not been fabricated and only recently studied,[35] but the 1DE/2DC device is connected to the 2DE/1DC device through convolution as shown by an arrow along a diagonal in the figure. Thus, for the nine possible devices of Fig. 10, convolution with the one-dimensional free-electron density of states relates their conductance along the diagonals.

Including quantization in the device emitter produces many more possibilities for resonant tunneling devices. Some recent studies in this area are summarized in Ref. 36. The 2DE devices could be regarded as having a quasi-one-dimensional emitter (Q1DE), while the 3DE devices can have either a quasi-two-dimensional emitter (Q2DE) or a Q1DE. Obtaining the conductance for these devices is quite difficult in general. But for the devices along the main diagonal of Fig. 10, which are described by a separable potential, including the effect of the emitter quantization is much easier. Because the potential depends separately on each coordinate z, x, and y for the 1DE/3DC, 2DE/2DC, and 3DE/1DC devices, an electron cannot scatter from one subband to another as it traverses these devices. The separate subbands then act like independent devices in parallel.

Suppose that confinement in the emitter gives rise to a number of subbands at energies E_1, E_2, E_3, etc. To obtain the conductance of the devices along the main diagonal when the emitter energies are quantized into confinement subbands, we must convolve the device currents obtained in the previous sections with the function

$$\delta(E-E_1)+\delta(E-E_2)+\delta(E-E_3)+\cdots. \quad (42)$$

Again, this corresponds to adding classical resistors in parallel since there is no scattering between the different confinement subbands. One can sketch all the diode and transistor currents very easily for this case. It has also been speculated that accumulation in the emitter of a standard MBE-grown tunneling diode would give rise to a bound level in the emitter. In that case, we would have to add the current of a 2DE in parallel with the 3DE to obtain the total device current.

VI. CONCLUSIONS

We have studied the transistor- and diodelike operation of resonant tunneling devices having a one-, two-, and three-dimensional electron emitter. Our method emphasizes the use of convolutions, allowing us to separate the effects of finite voltage, finite temperature, and increasing spatial dimension of the emitter on the device current. Using this convolution method, we obtained an intuitive physical picture of the tunneling current. Although we did not emphasize the temperature dependence of the current here, it can also be included as a simple convolution.

We considered transistor operation by sweeping the Fermi energy in the resonant tunneling device. This is possible in a field-effect transistor using either a substrate contact or light. For this case, the 1DE device shows a strong negative transconductance, the 2DE device a weaker negative transconductance, while the 3DE device can at best show a zero transconductance. Only in the 1DE case does the current fall to zero when the device is operated as a transistor.

For diodelike operation of the resonant tunneling devices, we recover the standard triangular-shaped resonant peak in the 3DE device where the resonant device current is proportional to the emitter to collector voltage V. For the 2DE device the resonant current is proportional to $\sqrt{V}$. In the 1DE device the resonant current is independent of source-to-drain voltage. The resonant current for the 1DE diode is equal to the fundamental constants $e/\pi\hbar$ times the area ΔE under the resonance peak in $T(E,V)$. We examined the case where the bound level is initially below the Fermi energy in the emitter and the different resulting voltage dependencies of the diode current.

We also examined transistor operation of the resonant tunneling devices in each dimension by sweeping the energy of the quantum well. We modeled this energy as being directly proportional to a base voltage V_B and assumed the base current was zero. A strong negative transconductance is then possible in each spatial dimension sweeping the base voltage, because the base contact can pull the resonant level either above or below the electron distribution in the emitter. The dependence of the transistor currents on the base voltage for large bias across the device is the same as for the diode current, that is, proportional to V_B for the 3DE, rising like $\sqrt{V_B}$ for the 2DE, and independent of V_B in one dimension. We also examined the transistor currents-versus-base voltage at small emitter to collector voltages.

The unifying theme of our paper is that the effects from various physical process on the tunneling current can be more easily understood when expressed as convolutions. We have used convolutions to treat dimensional effects as well as broadening effects. As a final example, we have shown how our convolution method can give insight into resonant tunneling devices having different degrees of confinement in the quantum well, in addition to having a different spatial dimension in the device emitter.

ACKNOWLEDGMENTS

We acknowledge useful discussion with K. Ismail, J. del Alamo, R. A. Ghanbari, D. A. Antoniadis, and Henry I.

Smith. P. F. B. gratefully acknowledges financial support under U. S. Air Force Office of Scientific Research contract No. AFOSR-88-0304.

[1]L. L. Chang, L. Esaki, and R. Tsu, Appl. Phys. Lett. **24**, 593 (1974); R. Tsu and L. Esaki, *ibid.* **49**, 562 (1973). In this second paper, the ratio of squared wave-function magnitudes tt^* should be replaced by the current transmission probability $(k_r/k_l)tt^*$ as pointed out by D. D. Coon and H. C. Liu [Appl. Phys. Lett. **47**, 172 (1985)].

[2]T. C. L. G. Sollner, W. D. Goodhue, P. E. Tannenwald, C. D. Parker, and D. D. Peck, Appl. Phys. Lett. **43**, 588 (1983); T. C. L.G. Sollner, E. R. Brown, and H. Q. Le, in *Physics of Quantum Electron Devices*, edited by F. Capasso (Springer, Berlin, 1990).

[3]T. P. E. Broekaert, W. Lee, and C. G. Fonstad, Appl. Phys. Lett. **53**, 1545 (1988).

[4]M. A. Reed, J. N. Randall, R. J. Aggarwal, R. J. Matyi, T. M. Moore, and A. E. Wetsel, Phys. Rev. Lett. **60**, 535 (1988).

[5]K. Ismail, D. A. Antoniadis, and H. I. Smith, Appl. Phys. Lett. **55**, 589 (1989).

[6]A. Palevski, M. Heiblum, C. P. Umbach, C. M. Knoedler, A. N. Broers, and R. H. Koch, Phys. Rev. Lett. **62**, 1776 (1989).

[7]S. Y. Chou, J. S. Harris, Jr., and R. F. W. Pease, Appl. Phys. Lett. **52**, 1982 (1988); S. Y. Chou, E. Wolak, and J. S. Harris, Jr., *ibid.* **52**, 657 (1988); S. Y. Chou, D. R. Allee, R. F. W. Pease, and J. S. Harris, Jr., *ibid.* **55**, 176 (1989); D. R. Albee, S. Y. Chou, J. S. Harris, Jr., and R. F. W. Pease, Superlattices and Microstructures, **7**, 131 (1990).

[8]M. A. Reed, W. R. Frensley, J. Matyi, J. N. Randall, and A. C. Seabaugh, Appl. Phys. Lett. **54**, 1034 (1989).

[9]F. Beltram, F. Capasso, S. Luryi, S. G. Chu, A. Y. Cho, and D. L. Sivco, Appl. Phys. Lett. **53**, 219 (1988); F. Capasso, S. Sen, F. Beltram, L. M. Lunardi, A. S. Vengurlekar, P. R. Smith, N. J. Shah, R. J. Malik, and A. Y. Cho, IEEE Trans. Electron Devices **36**, 2065 (1989).

[10]S. Luryi, Appl. Phys. Lett. **47**, 490 (1985). A similar picture was developed by L. V. Iogansen, Sov. Phys. Usp. **8**, 413 (1965).

[11]H. C. Liu and G. C. Aers, Solid State Commun. **67**, 1131 (1988); J. Appl. Phys. **65**, 4908 (1989).

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

[13]R. Landauer, J. Phys. Condensed Matter **1**, 8099 (1989); R. Landauer, in *Analogies in Optics and Micro-Electronics*, edited by W. van Haeringen and D. Lenstra (Academic, New York) (in press). Equation (1) is valid only through a "constriction" geometry as emphasized in this reference.

[14]The convolution symbol $\otimes$ has its usual meaning: $A(E)\otimes B(E)=\int_{-\infty}^{\infty}A(E')B(E-E')dE'=\int_{-\infty}^{\infty}B(E')A(E-E')dE'$.

[15]In Ref. 12 for the case of $eV>E_F$ we must set $E=eV$ rather than $E=\mu$ after doing the convolutions. We avoid that problem here by referencing all electron energies to the emitter instead of to the collector and allowing only positive voltages.

[16]M. Büttiker, Phys. Rev. B **33**, 3020 (1986); IBM J. Res. Develop. **32**, 63 (1988).

[17]A. D. Stone and P. A. Lee, Phys. Rev. Lett. **54**, 1196 (1985).

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

[19]To prove that Eq. (3) arises from the "two-terminal" multichannel Landauer formula, one imposes periodic boundary conditions on the sum over transverse "channels" to show that (see Ref. 12) $\Sigma_{ij}T_{ij}(E)\rightarrow T(E)\otimes\frac{1}{2}N_{2D}(E)$. Imposing periodic boundary conditions implies that the electron has no knowledge of the lateral boundaries of the conductor. This assumes both that electron's phase coherence length is much less than the device width and that no significant reflection is introduced when the phase is randomized. The phase coherence length is still assumed much longer than the distance between the tunnel barriers.

[20]P. F. Bagwell and T. P. Orlando, Phys. Rev. B **40**, 3757 (1989).

[21]Although this convolution is easy to perform, we could also obtain the result by calculating the density of states for an electron with four spatial degrees of freedom.

[22]B. Ricco and M. Ya. Azbel, Phys. Rev. B **29**, 1970 (1984).

[23]P. J. Price, Superlattices and Microstructures **2**, 593 (1986); Phys. Rev. B **38**, 1994 (1988).

[24]S. Datta, *Quantum Phenomena* (Addison-Wesley, New York, 1989), pp. 32–35.

[25]J. N. Schulman and M. Waldner, J. Appl. Phys. **63**, 2859 (1988).

[26]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); 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).

[27]T. E. Kopley, P. L. McEuen, and R. G. Wheeler, Phys. Rev. Lett. **61**, 1654 (1988); Bull. Am. Phys. Soc. **34**, 778 (1989).

[28]Also interesting is that there is no observability criterion for increasing the emitter dimensionality, similar to that for voltage or temperature, because there is no characteristic width in energy for either $N_{1D}(E)$ or $N_{2D}(E)$.

[29]M. Cahay, M. McLennan, S. Datta, and M. S. Lundstrom, Appl. Phys. Lett. **50**, 612 (1987); see alo M. Cahay, Ph.D. thesis, Purdue University, 1987.

[30]T. P. E. Broekaert, S. M. thesis, M.I.T., 1989. The dependence of the resonance width ΔE on applied bias is calculated here.

[31]Accumulation in the emitter and depletion in the collector is a significant effect in tunneling diodes grown using molecular-beam epitaxy. This effect can be handled either by (1) using a new emitter distribution set by the size of the accumulation or (2) keeping the same emitter distribution and including the accumulation and depletion effects in the transmission coefficient. These two different methods seem to correspond to different assumptions concerning the effects of inelastic scattering in the emitter and is an unresolved question.

[32]Equation (32) has an interesting form: $I=2e/\Delta\tau$, where $\Delta E\Delta\tau=h$ and ΔE is the area under the transmission resonance. This resistance is independent of the amount of energy the electron must dissipate to come into equilibrium in the right contact. For a completely ballistic conductor $\Delta E=eV$ until $eV>E_F$, at which point $\Delta E=E_F$ so that $I=(e/\pi\hbar)E_F$ for a completely ballistic conductor in one dimension. This is also independent of the disequlibrium energy of the electron in the right contact. Even though one might like to interpret these resistances as arising from inelastic scattering in the contacts, the independence of these two results from the energy to equilibrate in the contact argues against such an interpretation of Eq. (1).

[33]V. Kalmeyer and R. B. Laughlin, Phys. Rev. B **35**, 9805 (1987).

[34]W. Xue and P. A. Lee, Phys. Rev. B **38**, 3913 (1988).

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

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

Appendix F

Philip F. Bagwell, 'Evanescent Modes and Scattering in Quasi-One-Dimensional Wires', Physical Review B, 41, 10354 (1990).

Evanescent modes and scattering in quasi-one-dimensional wires

Philip F. Bagwell
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139
(Received 12 January 1990)

We calculate the current transmission amplitudes and electrical conductance as a function of Fermi energy for electrons scattering from a single defect in a quasi-one-dimensional wire. In a confined geometry the scattering boundary conditions couple propagating modes in the wire to nonpropagating or evanescent modes. Therefore, the applied steady current causes localized or evanescent modes to build up around any defects in the wire. These extra stored electrons strongly affect the scattering boundary conditions for the propagating modes whenever the Fermi energy approaches either a new quasi-one-dimensional subband or a quasi-bound-state splitting off of the higher confinement subbands. We show that the presence of evanescent modes can lead to either perfect transparency or perfect opaqueness for the scattering modes, even in the presence of scattering defects. For the special case of a δ-function scatterer in the wire we analytically obtain the scattering amplitudes. We also numerically examine a finite-range scatterer.

I. INTRODUCTION

When an electron scatters elastically from an imperfection in an open geometry, such as the scattering from a potential-energy barrier or well in an infinite three-dimensional space, it scatters into a traveling wave which propagates away from the defect. In contrast, if the electron is restricted to a wire such that confinement subbands are formed, the incident electron can elastically scatter into evanescent modes available in the wire. Thus for a steady current flow incident on a defect in the wire, a localized mode will build up around the defect even if the scatterer is repulsive. Unlike the case in electron localization theory where the "localization" is either a consequence of the coherent scattering of electrons from *multiple* defects or of electrons trapped in potential-energy wells, the localized electron trapped around a *single* scattering defect in a narrow wire is maintained there by the applied incident current, regardless of the sign of the scattering potential.

For the special case of an attractive scatterer in a narrow wire, it is possible to have quasi-bound-states splitting off from one of the higher-lying confinement subbands. The bound state associated with the lowest subband is analogous to a donor level below the conduction-band minima of a semiconductor. An electron trapped in the lowest subband's "donor level" at zero temperature cannot escape the region of space near the donor; it is truly bound. But since the confinement potential of the wire gives rise to multiple subbands, there will be a new "quasi-donor-level" associated with each subband. These additional "bound-state remnants" are also spatially localized near the defect and composed of evanescent waves, but are not true bound states. Because they are degenerate in energy with propagating modes in the wire and are coupled to the propagating modes, they will decay with time. But in contrast to the unbound evanescent waves which can build up around either an attractive or a repulsive scattering defect, we believe the quasi-bound-states can persist for much longer times if the incident current is reduced to zero. Analyzing how these new quasi-bound-states will decay with time is an interesting problem in its own right, but for the purposes of this paper it is sufficient to realize that these states exist and are important for electron transport in confined geometries.

We show in this paper that the presence of evanescent modes strongly affects the scattering of propagating modes from a defect in a quasi-one-dimensional wire. For interacting electrons this is clearly true because an electron trapped near a scattering defect will alter the scattering potential near the defect. But even for the noninteracting electrons which we consider in this paper, the building up of evanescent modes near the scattering defect alters the boundary conditions for the scattering event. This leads to unusual scattering properties when the Fermi energy approaches either a subband minima or a quasi-bound-state splitting off of a higher-lying confinement subband, so that evanescent modes cannot be neglected when analyzing scattering in a confined geometry. We are careful in our analysis to consider the effects of both propagating and evanescent modes during the scattering event. It is well known from the analogous case of electromagnetic wave scattering in microwave waveguides that one cannot neglect the electromagnetic energy stored near any defects or sudden spatial variations in the guide when calculating its scattering properties. For optical waveguides one must also consider electromagnetic radiation generated at defects, an effect whose analog is absent in our problem.

Our main goal is to understand the simplest possible scattering problem in a quasi-one-dimensional wire: scattering from a single δ-function defect in an infinitely

tain an equation of motion for the Fourier coefficients $c_n(x)$ as[2]

$$\frac{d^2c_n(x)}{dx^2}+k_n^2c_n(x)=\sum_m \Gamma_{nm}(x)c_m(x) \tag{4}$$

where the wave vector k_n is

$$k_n^2=\frac{2m}{\hbar^2}(E-E_n) \tag{5}$$

and the Γ_{nm}'s given by

$$\Gamma_{nm}(x)=\frac{2m}{\hbar^2}\int dy\, \chi_n^*(y)V_d(x,y)\chi_m(y) \tag{6}$$

are the mode coupling constants. Equation (4) allows any normal mode to scatter elastically into any other normal mode through the defect potential $V_d(x,y)$. One must therefore solve a scattering problem with an infinite number of coupled modes to obtain the exact answer for the scattering properties of any one normal mode. The sum on the right side of Eq. (4) includes both the term $n=m$ and the coupling to any evanescent modes for which we set $k_n=i\kappa_n$.

The boundary conditions for solving this scattering problem are analogous to the ones from textbook quantum-mechanical scattering problems in one dimension. Continuity of the wave function $\psi(x,y)$ requires that each Fourier coefficient $c_n(x)$ be continuous. If the scattering potential $V_d(x,y)$ is nonsingular, then it follows by integrating Eq. (4) and using the continuity of the $c_n(x)$'s that the derivative $dc_n(x)/dx$ must be continuous for each n. We discuss the boundary conditions for singular potentials in the next section.

In the regions where the scattering potential is zero, namely, regions I and III as shown in Fig. 1, the solutions to Eq. (4) are

$$c_n(x)=\begin{cases} A_ne^{ik_nx}+B_ne^{-ik_nx}, & x<0 \quad (7)\\ C_ne^{ik_n(x-L)}+D_ne^{-ik_n(x-L)}, & x>L\ , \quad (8)\end{cases}$$

for the propagating modes and, by setting $k_n=i\kappa_n$, obtain

$$c_n(x)=\begin{cases} A_ne^{-\kappa_nx}+B_ne^{\kappa_nx}, & x<0 \quad (9)\\ C_ne^{-\kappa_n(x-L)}+D_ne^{\kappa_n(x-L)}, & x>L\ , \quad (10)\end{cases}$$

for the evanescent modes. Solving Eq. (4) with the appropriate boundary conditions will determine the coefficients A_n, B_n, C_n, and D_n in each region. We describe a method for doing this in Sec. III and V.

Suppose the coefficients A_n, B_n, C_n, and D_n are now known. Define the current transmission probability $T_{i\to j}=T_{ij}$ from normal mode or "channel" i on the left to channel j on the right following the diagram in Büttiker *et al.*[15] by

$$T_{ij}=\frac{J_{j,\text{right}}}{J_{i,\text{left}}}=\tilde{t}_{ij}\tilde{t}^*_{ij}=\frac{k_j}{k_i}t_{ij}t^*_{ij} \tag{11}$$

when all other incoming currents are zero except those in channel i. Here $\tilde{t}_{ij}$ is the current transmission amplitude as defined in the scattering matrix.[20] In terms of the coefficients A_n, B_n, C_n and D_n these are

$$\tilde{t}_{i\to j}=\frac{(k_j)^{1/2}}{(k_i)^{1/2}}t_{i\to j}=\frac{(k_j)^{1/2}}{(k_i)^{1/2}}\frac{C_j}{A_i} \tag{12}$$

where we have defined $t_{i\to j}=C_j/A_i$ as the ratio of the wave-function amplitudes. Similarly, the reflection coefficients R_{ij} are

$$R_{ij}=\frac{J_{j,\text{left}}}{J_{i,\text{left}}}=\tilde{r}_{ij}\tilde{r}^*_{ij}=\frac{k_j}{k_i}r_{ij}r^*_{ij} \tag{13}$$

where $r_{i\to j}=B_j/A_i$ is the ratio of the reflected wave-function amplitude to the incident amplitude. The two-probe current at small voltages we obtain from one of Landauer's formulas[15–22]

$$G_{\text{two-probe}}=\frac{I}{V}=\frac{e^2}{\pi\hbar}\sum_{i,j}T_{ij}=\frac{e^2}{\pi\hbar}\sum_{i,j}\tilde{t}_{ij}\tilde{t}^*_{ij}=\frac{e^2}{\pi\hbar}\mathrm{Tr}(\tilde{t}\,\tilde{t}^\dagger) \tag{14}$$

where the sum in Eq. (14) runs only over the propagating normal modes of the wire. Equation (14) and its generalizations to finite voltages are equivalent to standard tunneling conductance formulas.[24–26] The quantum contact resistance implied by Eq. (14), discussed in Refs. 17–22, has been seen in a striking and clear fashion.[27,28]

We shall also examine the four-probe Landauer conductance.[14,15,17,21,22] For the cases we consider in this paper, an electron incident on the scatterer from either end of the conductor will have the same transmission probability. Or in the notation of Ref. 15, we will have $T_{ij}=T'_{ij}$ and $R_{ij}=R'_{ij}$ in all our calculations. For this case the four-probe conductance can be written more simply as

$$G_{\text{four-probe}}=\frac{I}{V'}=\frac{e^2}{\pi\hbar}\left[\sum_{i,j}T_{ij}\right]\left[\sum_i N_i\right]\times\left[\sum_i N_i\left[\sum_j R_{ij}\right]\right]^{-1}. \tag{15}$$

Equation (15) is just the two-probe conductance from Eq. (14) multiplied by a voltage division factor. The currents I in Eqs. (14) and (15) are the same but the voltages V and V' are measured differently.[14] In Eq. (15) N_i is the one-dimensional free-electron density of propagating states for the ith occupied subband. Also following Ref. 15, we assume that the weakly coupled voltage probes are far enough away from the scatterer so that evanescent modes existing near the voltage probes can be neglected. This assumption is always violated near the subband minima. However, for our special case, the evanescent mode amplitude is the same on either side of the scatterer so the four-probe formula of Ref. 15 is likely valid for our problem despite its neglect of charge accumulation near the voltage probes in the evanescent modes.

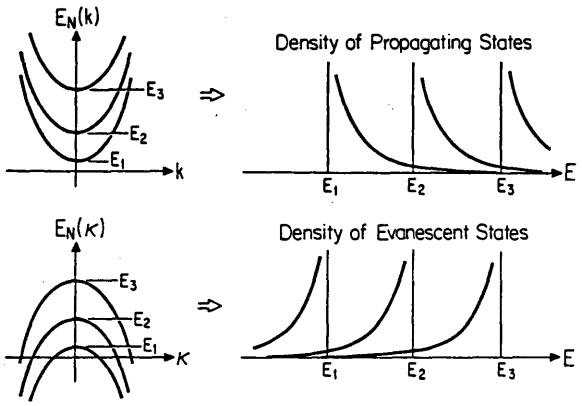


FIG. 2. Density of states in a quasi-one-dimensional wire. (a) shows the unusual dispersion relation and density of propagating states while (b) emphasizes that evanescent states also exist in the wire. Thus, in the presence of scattering defects, the states available in (a) are not the only relevant ones to consider. The evanescent density of states depends on position as explained in the text.

problem, as they must be for scattering in a confined geometry, then the position-dependent density of states

$$N(E;x,y)=\sum_{n,k}\delta(E-E_{nk})|\psi_{nk}(x,y)|^2 \tag{16}$$

will no longer be just the usual result for clean wire shown in Fig. 2(a). Instead, for positions near the defect, there will be an extra piece added to the position-dependent density of states which will resemble Fig. 2(b). Figure 2 again emphasizes that evanescent modes will have an effect on the scattering properties of the wire well before the subband minima is reached. If quasi-bound-states form in the wire for the case of an attractive scatter, then Fig. 2(b) will of course become more complicated.

The remaining sections will be concerned with solving the infinite set of coupled equations (4) for the wave functions $c_n(x)$, obtaining the various transmission and reflection coefficients, and studying the Landauer conductance of the wire.

III. CASE OF A δ-FUNCTION SCATTERER

In this section we consider electron scattering from a δ-function potential in a quasi-one-dimensional wire. We write down the detailed analytic solutions to the infinite set of coupled equations, Eq. (4), for the transmission coefficients T_{ij} and reflection coefficients R_{ij} in Appendix B. These solutions are valid for δ-function potentials of arbitrary strength and for any arbitrary number of propagating and evanescent modes. In this section we formulate the problem and discuss some of its most interesting features.

Let the δ-function scattering potential be

so that we take $L=0$ in Fig. 1. The weight γ can be either positive or negative. Integrating Eq. (4) across the δ function gives[2]

$$\left.\frac{dc_n(x)}{dx}\right|_{x=0^+}-\left.\frac{dc_n(x)}{dx}\right|_{x=0^-}=\sum_m \Gamma_{nm}c_m(0) \tag{18}$$

where

$$\Gamma_{nm}=\frac{2m\gamma}{\hbar^2}\chi_n^*(y_i)\chi_m(y_i) \tag{19}$$

are the mode coupling constants. The Γ_{nm} are proportional to the strength of the impurity and the size of the wave function at the impurity, and are zero if the wave function has a node at the impurity position. Equation (18) reduces to

$$ik_n(C_n-D_n)-ik_n(A_n-B_n)=\sum_m \Gamma_{nm}(A_m+B_m) \tag{20}$$

if n denotes a propagating mode and, setting $k_n=i\kappa_n$, becomes

$$-\kappa_n(C_n-D_n)+\kappa_n(A_n-B_n)=\sum_m \Gamma_{nm}(A_m+B_m) \tag{21}$$

if n denotes an evanescent mode. The sum over m again includes the term where $m=n$ as well as all the evanescent modes. In addition, the boundary condition that the wave function itself must be continuous, $c_n(0^-)=c_n(0^+)$, gives

$$A_n+B_n=C_n+D_n \tag{22}$$

for all normal mode indices n.

Imposing the remaining boundary conditions from our scattering problem, proceeding as in the normal textbook quantum-mechanical scattering problems in one dimension, allows further simplification of Eqs. (20), (21), and (22). We consider particles incident only from the left. Then $D_n=0$ for all propagating n. Because we take the wire to be infinitely long, we can eliminate all growing evanescent waves and set $D_n=0$ and $A_n=0$ for all evanescent n. This leaves the condition $B_n=C_n$ for all evanescent n and $A_n+B_n=C_n$ for all propagating n. Furthermore, for all the propagating A_n's, we allow only one normal mode incident at a time so that we have one more unknown than equations (as is usual in scattering problems). We solve for all the transmitted and reflected wave amplitudes with respect to the single incident A_n.

We now find a procedure for solving the infinite set of coupled equations (20), (21), and (22). Truncating the equations to a finite size will allow us to find the right procedure. For definiteness, let us write down the matrix equation we must solve if we consider two propagating modes and two evanescent modes. Let mode two be incident on the scatterer so that $A_1=0$ and $B_1=C_1$. We write down the set of four coupled equations valid when

$$\begin{bmatrix} 0 \\ -2ik_2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \Gamma_{11}-2ik_1 & \Gamma_{12} & \Gamma_{13} & \Gamma_{14} \\ \Gamma_{21} & \Gamma_{22}-2ik_2 & \Gamma_{23} & \Gamma_{24} \\ \Gamma_{31} & \Gamma_{32} & \Gamma_{33}+2\kappa_3 & \Gamma_{34} \\ \Gamma_{41} & \Gamma_{42} & \Gamma_{43} & \Gamma_{44}+2\kappa_4 \end{bmatrix} \begin{bmatrix} t_{21} \\ t_{22} \\ t_{23} \\ t_{24} \end{bmatrix} . \tag{23}$$

Now let us eliminate the highest evanescent mode, in this case mode four, from the 4×4 matrix Eq. (23). We obtain the 3×3 matrix equation

$$\begin{bmatrix} 0 \\ -2ik_2 \\ 0 \end{bmatrix} = \begin{bmatrix} \Gamma_{11,4}-2ik_1 & \Gamma_{12,4} & \Gamma_{13,4} \\ \Gamma_{21,4} & \Gamma_{22,4}-2ik_2 & \Gamma_{23,4} \\ \Gamma_{31,4} & \Gamma_{32,4} & \Gamma_{33,4}+2\kappa_3 \end{bmatrix} \begin{bmatrix} t_{21} \\ t_{22} \\ t_{23} \end{bmatrix} \tag{24}$$

which has the same form as the matrix equation we would have started with if we had originally decided to allow two propagating modes and only one evanescent mode. Equation (23) looks formally like Eq. (24) if we simply truncate the highest evanescent mode in Eq. (23). The only difference between Eq. (24) and truncating Eq. (23) is that all the mode coupling constants in Eq. (24) have been rescaled by the evanescent mode four where

$$\Gamma_{ij,4}=\Gamma_{ij}\frac{2\kappa_4}{\Gamma_{44}+2\kappa_4}, \quad i,j=1,2,3 . \tag{25}$$

Now eliminate the next highest evanescent mode, mode three, from Eq. (24). We obtain

$$\begin{bmatrix} 0 \\ -2ik_2 \end{bmatrix} = \begin{bmatrix} \Gamma_{11,3-4}-2ik_1 & \Gamma_{12,3-4} \\ \Gamma_{21,3-4} & \Gamma_{22,3-4}-2ik_2 \end{bmatrix} \begin{bmatrix} t_{21} \\ t_{22} \end{bmatrix} . \tag{26}$$

Again, this is the same matrix equation we would have originally written down if we had allowed only the two propagating modes and completely neglected the evanescent modes in Eq. (23), except that the mode coupling constants [which were initially only rescaled by mode four in Eq. (24)] are now rescaled by both the evanescent mode three and mode four as

$$\Gamma_{ij,3-4}=\Gamma_{ij,4}\frac{2\kappa_3}{\Gamma_{33,4}+2\kappa_3}, \quad i,j=1,2 . \tag{27}$$

The total effect of the evanescent modes is to rescale the mode coupling constants of the propagating modes. The rescaling procedure makes the mode coupling constants energy dependent. And although we have written down this procedure only for two propagating modes and two evanescent modes, by induction it clearly holds for any number of propagating and evanescent modes. If there are p propagating modes and q evanescent modes, we can eliminate the q evanescent modes by the procedure outlined above and solve a $p\times p$ matrix equation to obtain the transmission and reflection amplitudes for the propagating modes, and hence determine the conductance. We solve this matrix equation explicitly for the lowest three subbands in Appendix B and assert that, because of the symmetry of the matrix equation analogous to Eq. (23), we can solve the $p\times p$ matrix equation for any integer p. Hence we find a complete solution to our scattering problem. We emphasize at this point that our scheme for including the total effect of the evanescent modes as a simple rescaling of the mode coupling constants for the propagating modes is only valid for the special case of a single δ-function scatterer in a narrow wire.

Clearly we must truncate this infinite set of coupled equations in order to solve them, but how many equations must we include in order to obtain a physically correct answer? The above rescaling procedure tells us that it depends on the strength of the scatterer. From Eq. (19) all the mode coupling constants are of order $\Gamma_{nm}\sim 2m\gamma/W\hbar^2$ so that for strong scatterers the larger will be Γ_{nn} and hence it will be necessary to include more evanescent modes in order to obtain the correct transmission probabilities. For evanescent mode n, the rescaling of the mode couplings constants will have negligible effect when the evanescent density of states times the strength of the scatterer is small, $(m/\hbar^2\kappa_n)(\gamma/W)\ll 1$, and so it is necessary to include enough evanescent modes in the calculations such that this inequality is satisfied. In real wires, which have a potential well of finite depth, evanescent states from the continuum may play some role in determining their scattering properties if the scattering potential is strong enough.

Our mode rescaling procedure tells us another important fact: The modes completely "decouple" when the particle's energy aligns with the bottom of a subband, and perfect transmission results for each of the separate modes. Consider the case of two propagating modes described by Eq. (26). From Eq. (27), as we approach the bottom of subband 3, we must rescale all the mode coupling constants by the factor $2\kappa_3/(\Gamma_{33}+2\kappa_3)$. This factor approaches zero since $\kappa_3\to 0$ at the third subband minima. Hence all the mode coupling constants become zero, making it appear as though the scatterer were absent from the wire as we approach each new quasi-one-dimensional subband. This "perfect transparency" effect can be seen here for $T_{22}=1$ and $T_{21}=0$ by substituting the mode coupling constants near the third subband minima into Eq. (26). A similar equation written down in

Appendix B shows $T_{11}=1$ and $T_{12}=0$ near the third subband minima. By our mode rescaling arguments, all the mode coupling constants will become zero at each new Q1D subband minimum for the δ-function scatterer.

One can see this perfect transmission effect even more clearly by writing down the set of equations valid for one propagating mode and one evanescent mode for $E_1<E<E_2$ as

$$\begin{bmatrix} -2ik_1 \\ 0 \end{bmatrix} = \begin{bmatrix} \Gamma_{11}-2ik_1 & \Gamma_{22} \\ \Gamma_{21} & \Gamma_{12}+2\kappa_2 \end{bmatrix} \begin{bmatrix} t_{11} \\ t_{12} \end{bmatrix}, \tag{28}$$

or, writing this slightly differently,

$$\begin{aligned} -2ik_1(A_1-C_1)&=\Gamma_{11}C_1+\Gamma_{12}C_2 , \\ -2\kappa_2 C_2&=\Gamma_{21}C_1+\Gamma_{22}C_2 . \end{aligned} \tag{29}$$

Including more evanescent modes in Eq. (28) will not affect the structure of the equation, since their total effect is just to alter the mode coupling constants in a qualitatively unimportant way. The first row of Eq. (28) or Eq. (29) asserts that the discontinuity in the derivative of mode one at the defect is proportional to the probability amplitudes of each of the modes. The derivative jump of the evanescent mode, from the second row of Eq. (28), must also be supported by the buildup of the wave function at the defect. The second row of Eq. (28) emphasizes that the modes are coupled, since if the amplitude of the evanescent mode is taken to be zero, the propagating mode must also have zero amplitude.

Setting $\kappa_2=0$, meaning that there is no change in the derivative of mode two at the defect (mode two stays a constant value for the entire length of the wire), gives perfect transmission of mode one, $T_{11}=1$ when $E=E_2$. To obtain this result from Eq. (28) we use the identity $\Gamma_{21}\Gamma_{12}=\Gamma_{11}\Gamma_{22}$. The analogous result holds for higher subbands; all the modes have perfect transmission whenever the electron energy aligns with the bottom of a subband. Equation (28) asserts that, at the bottom of the second subband, the evanescent mode occupation will build up to exactly the right amount to compensate for the derivative jump required of mode one at the scatterer.

There is another interesting feature of Eq. (28). If the δ-function potential is attractive, that is if $\gamma<0$, then it is possible to have $\Gamma_{22}+2\kappa_2=0$. Or more generally, if there are m total modes, then $\Gamma_{22,3-m}+2\kappa_2=0$. We show in Appendix A that the energy satisfying this equation corresponds to a quasi-bound-state in the δ function splitting off from the second subband. In that case Eq. (28) predicts zero transmission, that is $T_{11}=0$. A qualitatively similar thing happens in higher subbands as we show in the next section, although the transmission does not fall to zero for the higher subbands. So the presence of evanescent modes in the wire leads to perfect transparency when $\kappa_n=0$ for any n, and we can also lead to perfect opaqueness in the lowest subband if the δ-function potential is negative but not too strong. If the δ potential is so strong and negative that $\Gamma_{22,3-m}+2\kappa_2=0$ has no solu-

longer be a point in energy of perfect reflection in that region because the bound state (which originally split off of the second subband) has dropped below the first subband and therefore out of the scattering problem. We discuss the bound states of the δ-function potential in a quasi-one-dimensional wire in detail in Appendix A.

We are now interested in showing, from Eq. (28), how the occupation probability of the evanescent mode accumulates as the Fermi energy approaches the second subband. The solutions to Eq. (28), valid in the energy range from $E_1<E<E_2$, are

$$t_{11}=\left[1-\left[\frac{\Gamma_{11}}{2ik_1}\frac{2\kappa_2}{\Gamma_{22}+2\kappa_2}\right]\right]^{-1}, \tag{30}$$

leading to the transmission coefficient T_{11}:

$$T_{11}=\left[1+\left[\frac{\Gamma_{11}}{2k_1}\right]^2\left[\frac{2\kappa_2}{\Gamma_{22}+2\kappa_2}\right]^2\right]^{-1} . \tag{31}$$

Since there is only one propagating mode we have $1=T_{11}+R_{11}$. The wave-function amplitude in the evanescent mode is

$$\begin{aligned} t_{12}=\frac{B_2}{A_1}&=\frac{\Gamma_{21}}{\Gamma_{22}+2\kappa_2}t_{11} \\ &=\Gamma_{21}\left[(\Gamma_{22}+2\kappa_2)-\left[\frac{\Gamma_{11}}{2ik_1}2\kappa_2\right]\right]^{-1} , \end{aligned} \tag{32}$$

leading to a building up of probability density

$$\begin{aligned} |t_{12}|^2&=\left|\frac{B_2}{A_1}\right|^2 \\ &=\Gamma_{11}\Gamma_{22}\left[(\Gamma_{22}+2\kappa_2)^2+\left[\frac{\Gamma_{11}}{2k_1}\right]^2(2\kappa_2)^2\right]^{-1} . \end{aligned} \tag{33}$$

The transmission coefficient T_{11} from Eq. (31), as well as the size of the evanescent mode $|t_{12}|^2=|B_2|^2/|A_1|^2$ from Eq. (33), are shown in Fig. 3. The strengths of the δ functions are given in the figure caption and other relevant parameters at the beginning of the next section. Figure 3(a) shows the case of a repulsive δ-function scatterer. The two solid curves in Fig. 3(a) are for a weak repulsive scatterer, while the two dotted curves describe a slightly stronger repulsive scatterer. The transmission coefficient T_{11} is shown near the top of Fig. 3(a). Perfect transmission, $T_{11}=1$, occurs when the incident electron energy aligns with the second subband minima at $E=E_2=25$ meV. The transmission coefficient T_{11} decreases as the scatterer is made stronger. The electron probability to occupy the evanescent mode, $|t_{12}|^2$ shown near the bottom of Fig. 3(a), grows steadily as the energy approaches the second subband. $|t_{12}|^2$ qualitatively follows the density of evanescent states. As the strength of the repulsive scatterer increases and the transmission probability T_{11} correspondingly decreases, the probability of occupying the evanescent mode increases. This is a reasonable result since (1) a more opaque barrier reduces transmission and (2) the evanescent modes must be popu-

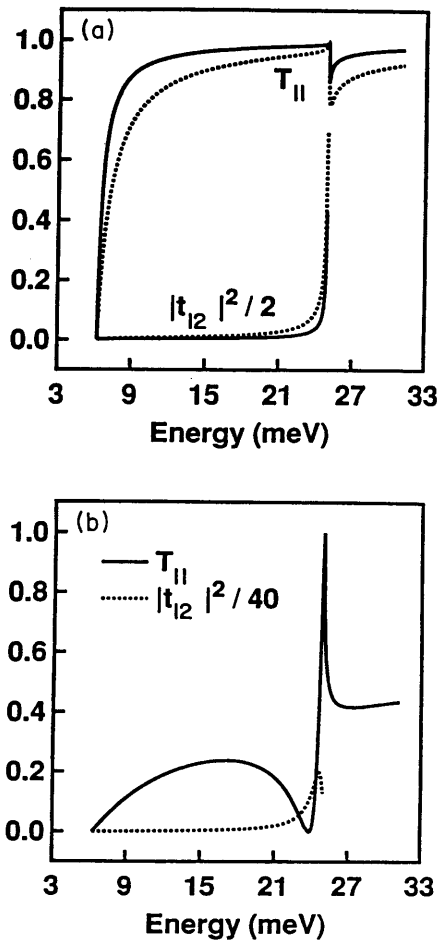


FIG. 3. Transmission coefficient T_{11} through δ-function scattering potential $V(x,y)=\gamma\delta(x)\delta(y-y_i)$ shown next to the strength $|t_{12}|^2$ of evanescent mode two. Perfect transmission $T_{11}=1$ results when $E=E_2=25$ meV. (a) shows the case of a repulsive scatterer having strength $\gamma=10$ feV cm^2 (two solid lines) and $\gamma=80$ feV cm^2 (two dotted lines). As the scatterer is made stronger, transmission decreases and more electrons are stored in evanescent mode two. (b) shows an attractive scatterer of strength $\gamma=-70$ feV cm^2. The transmission coefficient T_{11} (solid line) becomes zero at the quasi-bound-state energy, while the number of electrons $|t_{12}|^2$ (dotted line) stored in evanescent mode two reaches a maximum near the quasi-bound-state and then decreases as the incident electron energy approaches the second subband minima.

scattering is enhanced if the scattering strength increases.

Figure 3(b) shows T_{11} (solid line) and $|t_{12}|^2$ (dotted line) for an attractive scatterer which is slightly weaker than the repulsive ones of Fig. 3(a). The point where $T_{11}=0$ is the quasi-bound-state energy of an electron in the attractive scatterer which has split off from the second subband. Again $T_{11}=1$ at the second subband minima. The building up of the evanescent mode $|t_{12}|^2$ is largest near the quasi-bound-state energy, but its maximum does not occur at the bound-state energy. Equation (33) for $|t_{12}|^2$ has a Lorentzian-like shape whose peak is above the bound-state energy. Note also that, for the attractive scatter, the occupation of the evanescent have more than ten times the number of electrons stored in evanescent modes near the defect in Fig. 3(b) for the attractive scatterer as compared with the repulsive scatterer of Fig. 3(a). This is due to the quasi-bound-state nearby in energy. We shall see in the next section, however, that if the scatterer is made so strong that the quasi-bound-state energy moves below even the first subband, the transmission coefficient T_{11} in Fig. 3(b) for the attractive scatterer will evolve to qualitatively resemble the T_{11} found in Fig. 3(a) for the repulsive scatterer. For the moderately strong attractive scatterer, evanescent modes will then become less important. Another important difference between the attractive and repulsive scatterers is that, for the attractive scatterer of Fig. 3(b), the occupation of evanescent mode two first rises, reaches a maximum near the quasi-bound-state energy, and then falls as the energy approaches the second subband minima. This is in contrast to the repulsive scatterer of Fig. 3(a) where the occupation of evanescent mode two increases continuously as the energy approaches the second subband minima.

IV. RESULTS AND DISCUSSION OF δ-FUNCTION SCATTERER

In this section we graph the detailed solutions to the infinite set of coupled Eqs. (4) for the δ-function scatterer which are written down in Appendix B. We display all the intersubband and intrasubband transmission and reflection coefficients for the single δ-function scatterer as follows: As in Sec. III, we consider both a repulsive and an attractive scatterer. In Fig. 4 we consider a weak repulsive δ-function scatterer, while Fig. 5 shows a weak attractive scatterer of approximately the same strength as the repulsive scatterer. As the repulsive scatterer is made stronger, the energy dependence of the transmission or reflection coefficients does not qualitatively change. For an attractive scatterer the energy dependence of the reflection and transmission coefficients does qualitatively change as the scatterer is made stronger due to the movement of the quasi-bound-state. We do not show the transmission coefficients for the case of a strong attractive scatterer because they qualitatively resemble the ones for the repulsive scatterer in Fig. 4. As the attractive scatterer is made stronger all the graphs of Fig. 5 will evolve into ones qualitatively resembling Fig. 4. We then display the two-probe and four-probe Landauer conductance for these sets of scatterers in Figs. 6, 7, and 8.

For definiteness in all the calculations, we consider an infinite square well along the y direction which we take to be $W=300$ Å wide. We take the mass of the electron to be the effective mass for GaAs, that is 0.067 of the free-electron mass. For this wire, the subband energies are $E_1=6.236$ meV, $E_2=24.94$ meV, $E_3=56.12$ meV, and $E_4=99.78$ meV. We place the impurity $\frac{5}{12}$ of the distance across the wire. We include 100 total modes in our calculations. These are the same parameters used in Fig. 3. It is easy to include a greater or fewer number of modes, but including more modes does not have a great qualitative influence on the results. As we argued in Sec.

mode n we satisfy $2\kappa_n \gg 2m\gamma/W\hbar^2$, including more evanescent modes makes only a small quantitative difference in the results. Using Eq. (25) we see that, for a repulsive scatterer, including more evanescent modes increases the transmission because the mode coupling constants are made smaller. In our rescaling procedure each additional evanescent mode then multiplies the coupling constants by a number less than one. For an attractive scatterer where Γ_{nn} is negative, including more evanescent modes decreases the transmission because each additional evanescent mode then multiplies all the coupling constants by a number greater than one.

Figure 4 shows the transmission probabilities for a repulsive δ-function scatterer having $\gamma = 10$ feV cm^2. Or expressed more usefully $\gamma/W^2 = 1.111$ meV, so that the potential is relatively weak. Consider first the intrasubband transmission and reflection coefficients. Figure 4(a) shows that the transmission probability T_{11} is unity whenever the electron energy aligns with the bottom of a new subband. Figure 4(b) shows a similar result for T_{22} and T_{33}. The reflection coefficient R_{11} is shown in Fig. 4(c). Note that R_{11} is zero at the minima of the second and third subbands and undergoes no discontinuous change as does T_{11}. Reflection into the lowest normal mode is suppressed at the bottom of the second and third subband. A similar behavior for the intrasubband reflection coefficient R_{22} is shown in Fig. 4(d). If only the lowest mode can propagate, the intrasubband transmission and reflection coefficients T_{11} and R_{11} have the same functional form as those through a δ-function potential in one dimension given in Appendix A.

The intersubband transmission coefficients T_{12}, T_{13}, and T_{23} are shown on Figs. 4(a) and 4(b) multiplied by a factor of 10 so they can be clearly seen. The intersubband transmission and reflection coefficients are equal for the δ-function scatterer, that is $R_{12} = T_{12}$, etc. If there were no scattering the intersubband transmission would be zero, so the intersubband transmission should increase

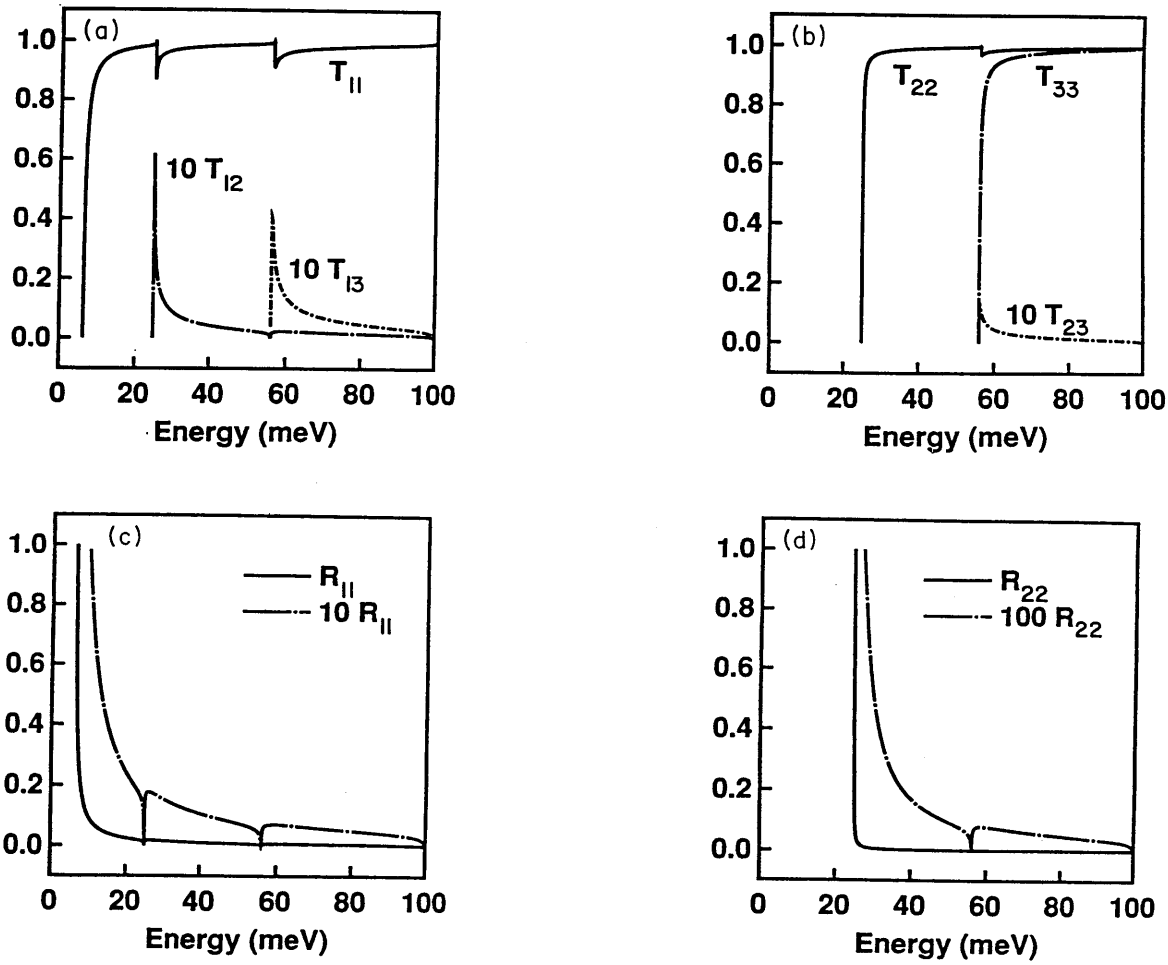


FIG. 4. Transmission coefficients for scattering from a weak repulsive δ-function potential in a quasi-one-dimensional wire having strength $\gamma = 10$ feV cm^2. We assume infinite square well confinement of width $W = 300$ Å. (a) shows T_{11}, $T_{12} = T_{21} = R_{12} = R_{21}$, and $T_{13} = T_{31} = R_{13} = R_{31}$. (b) shows T_{22}, T_{33}, and $T_{23} = T_{32} = R_{23} = R_{32}$. (c) shows R_{11} and (d) shows R_{22} all as a function of the incident electron energy. The overall shapes of the reflection coefficients R_{ij} can be understood from the golden-rule scattering rate as described in Appendix B. In addition, perfect transmission with no mode conversion occurs whenever the incident energy aligns with a new Q1D subband. The functional form of each curve is given in Appendix B.

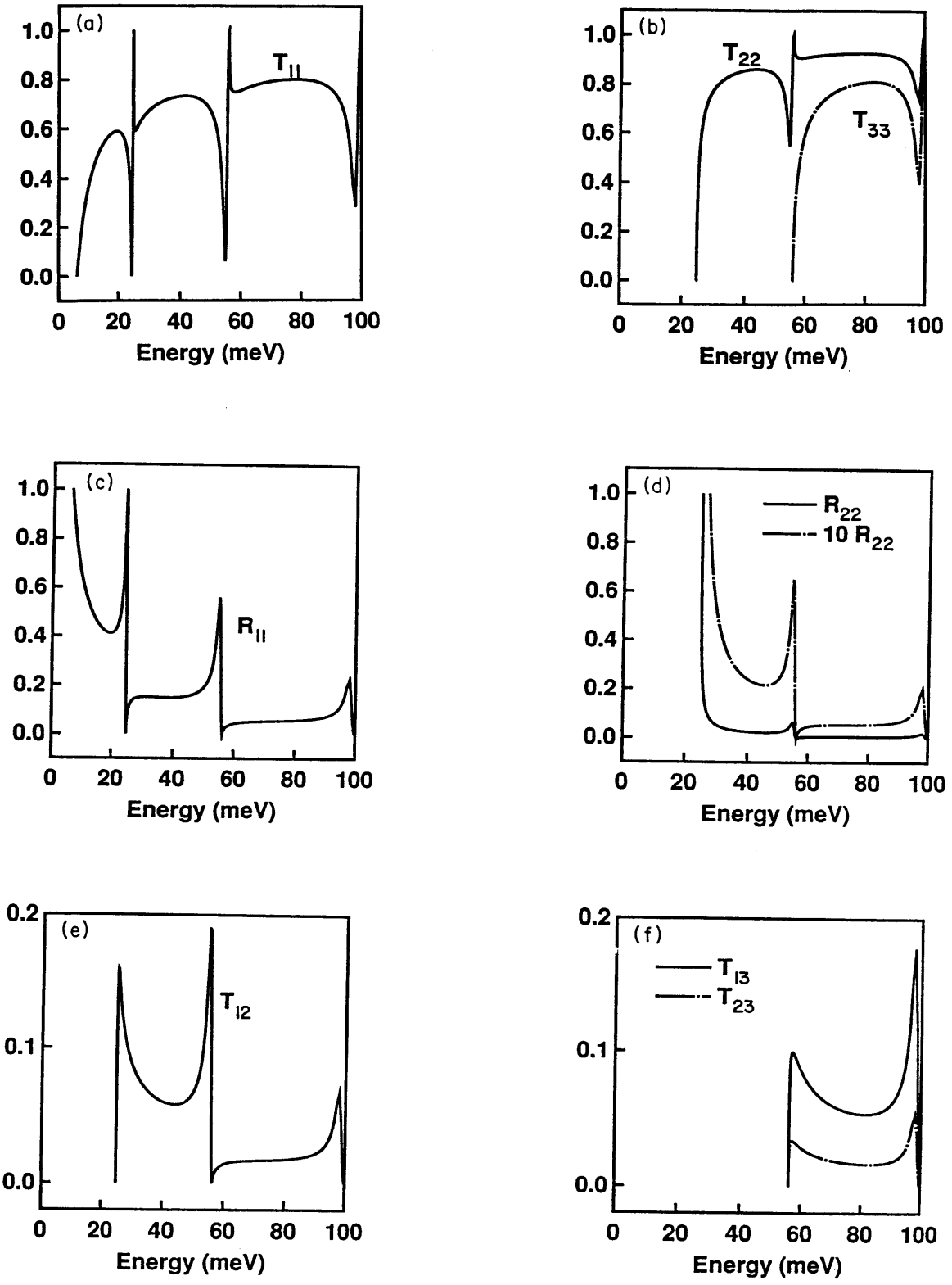


FIG. 5. Scattering from a weak attractive δ function having strength $\gamma = -6$ feV cm^2. (a) shows T_{11}, (b) gives T_{22} and T_{33}, (c) shows R_{11}, (d) shows R_{22}, (e) gives T_{22}, while (f) shows T_{13} and T_{23}. The qualitatively different behavior of these transmission coefficients from the ones in Fig. 4 arises from the presence of new quasi-bound-states forming in the attractive δ-function potential. For example, the new dips in the intrasubband transmission T_{11}, T_{22}, and T_{33} occur *well before* reaching subbands 2, 3, and 4, the distance between these dips and the subband minima being simply the bound-state energy as we show in Appendix B. The intrasubband reflection increases strongly at the quasi-bound-state energy as shown in (c) and (d). The intersubband transmission and reflection,

with increasing strength of the scatterer. This is indeed true, although not shown in the figures. In Fig. 4 the intersubband transmission is small because the potential is relatively weak. At the onset of the second subband in Fig. 4(a) only about 6% of the incident carriers are converted into the second normal mode through T_{12}, and 4–5 % are converted into the third normal mode via T_{13} at the bottom of the third subband. Figure 4(b) gives only between 1% and 2% conversion from the second to the third mode at the bottom of the third subband via T_{23}.

We can understand some features of Fig. 4 by arguing from the Fermi "golden-rule" scattering rate. To do this we do not consider the intrasubband transmission T_{11}, T_{22}, or T_{33}, as they are simply the result of leftover particles which did not scatter and can be obtained from the requirements of current conservation. Consider first the intersubband transmission T_{12}, T_{13}, and T_{23}. The intersubband transmission has a maximum near the onset of a subband and decays like the inverse square root of energy away from the maximum. This can be understood from a Fermi's "golden-rule" viewpoint, where the probability of scattering is proportional to the final density of states in the subband which decays like $1/\sqrt{E}$. In Appendix B we show that the dominant term in the intersubband scattering probability is indeed given by an expression similar to the golden rule. The intersubband transmission and reflection coefficient $T_{12}=R_{12}$ in Fig. 4(a) also shows interesting behavior around the bottom of the third subband, staying zero on both sides of the subband minima. There is no scattering out of mode one into mode two at the bottom of the third subband. We have yet to find a good explanation for this lack of mode conversion or reflection at the subband minima. However, the overall shapes of the transmission and reflection coefficients are still well understood by golden-rule arguments.

Given the golden-rule-like shapes of the intersubband transmission and reflection coefficients and the intrasubband reflection, we can argue for the shape of the intrasubband transmission. Let us do so for T_{11}. Because particles must be conserved so that $1=T_{11}+T_{12}+R_{12}+R_{11}$, and since $R_{11}=0$ on both sides of the subband minima, the drop in T_{11} after reaching perfect transmission at the second subband must be equal to $T_{12}+R_{12}=2T_{12}$, or just twice the intersubband transmission coefficient. This is shown in Fig. 4(a). Similarly, the discontinuity in T_{11} in Fig. 4(a) at the minima of the third subband is just twice T_{13}.

Next, let us examine the scattering coefficients for an attractive potential. Figure 5 shows a δ-function scatterer of comparable strength to the one in Fig. 4, but when

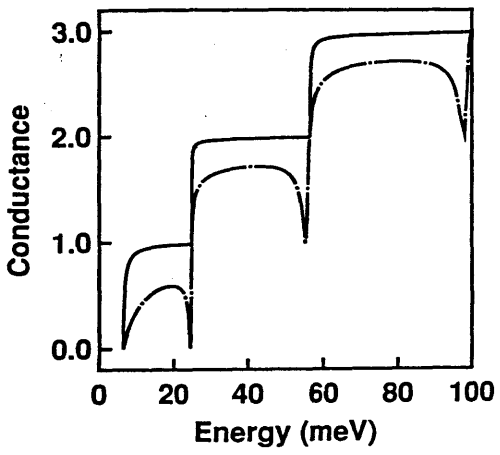


FIG. 6. Two-probe conductance through a δ-function defect in the quasi-one-dimensional wire in units of $2e^2/h$. The solid line corresponds to the repulsive scatterer from Fig. 4, while the dashed line gives the conductance of the attractive scatterer from Fig. 5. When the electron energy aligns with a subband minimum, the conductance through the defect is equal to the ballistic conductance. At these special energies the wire is perfectly transparent as if no scatterer were present. There is only a small difference between the conductance for the weak repulsive scatterer and the ideal ballistic conductance throughout the entire range of electron energies. For the attractive scatterer, the new dips in conductance correspond to quasi-bound-states developing in the wire. The distance in energy from these dips to the subband minimum is the quasi-bound-state energy. Note also that, even though the repulsive scatterer is stronger, the conductance of the attractive scatterer is much smaller due to the presence of the quasi-bound-state

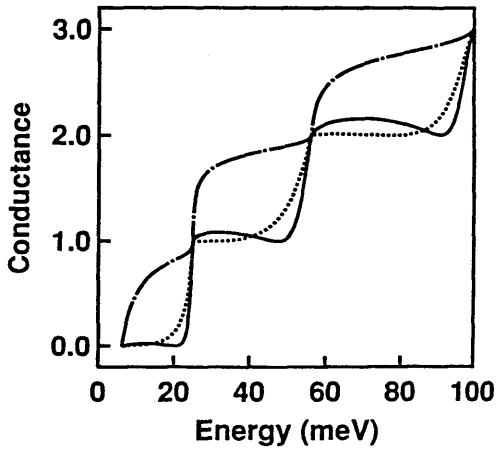


FIG. 7. Two-probe conductance in units of $2e^2/h$ for an attractive scatterer having $\gamma=-8$ feV cm^2 (solid line), $\gamma=-9$ feV cm^2 (dotted line),and $\gamma=-20$ feV cm^2 (dashed line). Beginning with the dotted line from Fig. 6 showing the weakest attractive scatterer having $\gamma=-6$ feV cm^2, the overall conductance level decreases and the new dips corresponding to the quasi-bound-states move lower in energy as the scatterer is made more attractive. As the scatterer becomes so attractive that the quasi-bound-states move below the bottom of the next lowest subband, the new dips first disappear and the conductance then *increases* as the scatterer is made stronger. This unusual effect occurs because the bound states have now moved below the energy range in which they can block conduction

the scatter is made attractive. The strength of the scatterer in Fig. 5 is $\gamma = -6$ feV cm^2, or equivalently $\gamma/W^2 = -0.666$ meV, 60% of the strength of Fig. 4. The transmission coefficient T_{11} through this potential is shown in Fig. 5(a). The new dips in T_{11} immediately before each subband minimum correspond to a quasi-bound-state forming in the wire. The distance in energy between each local minimum in T_{11} and the following subband is simply the quasi-bound-state energy. Although these new minima appear to be immediately below the subband minima, they are in fact at a lower energy. We can see this clearly from Fig. 3(b) where the attractive scatterer is stronger. Since the scatterer in Fig. 5 is very weak, the quasi-bound-states lie very near in energy to their respective subbands. Figure 5(b) shows T_{22} and T_{33} for this potential. We see again that intrasubband transmission decreases at the quasi-bound-state energy and is perfect at the subband minima.

The intrasubband reflection R_{11} and R_{22} increase at the quasi-bound-state energies and are zero at each successive subband minimum as shown in Figs. 5(c) and 5(d), respectively. Intersubband transmission also increases at the quasi-bound-state energies as shown for T_{12} in Fig. 5(e) and again for T_{13} and T_{23} in Fig. 5(f). Note that both the intrasubband reflection and intersubband transmission are much larger for this attractive scatterer than for the repulsive scatterer we considered in Fig. 4, even though the scattering strength is weaker than in Fig. 4. This is also due to the quasi-bound-state nearby in energy.

Now we turn to a study of the Landauer conductance through these δ-function scatterers. Figure 6 shows the two-probe conductance from Eq. (15) for both the repulsive δ-function defect of Fig. 4 (solid line) and the attractive δ-function defect of Fig. 5 (dashed line). The conductance for the weak repulsive scatterer is nearly the same as that of a perfect ballistic wire throughout the whole range of Fermi energy, and is exactly equal to the ballistic conductance when the Fermi energy aligns with each new subband minima. The shoulders of the quantized ballistic conductance steps are slightly rounded due to increased reflection immediately above each subband minima. The weak attractive scatterer, however, has a much smaller conductance reflecting the presence of the quasi-bound-state nearby in energy. The extra dips in the conductance through the attractive δ function again occur before reaching each subband, the difference being just the quasi-bound-state energy. The conductance of the attractive δ function is also exactly equal to the ballistic conductance when the Fermi energy aligns with new subband minima.

Figure 7 shows the two-probe conductance evolving as we increase the strength of the attractive scatterer from $\gamma = -6$ feV cm^2, shown as the dashed line in Fig. 6, to $\gamma = -8$ feV cm^2 (solid line), $\gamma = -9$ feV cm^2 (dotted line), and $\gamma = -20$ feV cm^2 (dashed line). Note that, as for the repulsive scatterer, perfect transmission still occurs at the subband minima. However, the additional quasi-bound-state in the attractive scatterer causes increased reflection well before the subband minima are reached, leading to an additional dip in the conductance (solid curve). As the strength of the attractive potential is increased, the conductance becomes smaller and the dip moves lower in energy until the quasi-bound-state drops below the next subband as shown by the dotted curve, where the conductance resembles a "cobweb" (dotted curve). As the potential is made even more attractive (dashed curve), the conductance increases and begins to resemble the conductance through a repulsive δ function. This is because each quasi-bound-state energy which originally split off from, say, subband i, has now moved below the bottom of the next lower subband $i-1$, and therefore out of the problem. It is paradoxical that, as the scatterer is made stronger over this small range of strengths, the conductance actually increases. The bound states have now moved out of the way and the electron can transmit

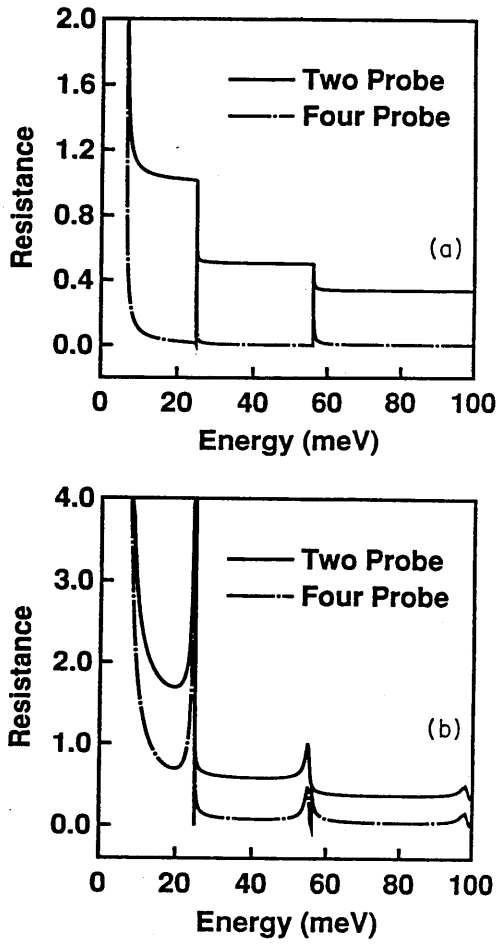


FIG. 8. Two-probe resistance (solid line) and four-probe resistance (dashed line) in units of $h/2e^2$ for the case of (a) the repulsive scatterer from Fig. 4 and (b) the attractive scatterer from Fig. 5. The overall resistance is much lower for the four-probe measurement due to the absence of contact resistance. The four-probe resistance $R_{\text{four-probe}}$ is zero at each subband minimum in both (a) and (b) due to the perfect transmission. $R_{\text{four-probe}}$ rises just above each subband minimum to equal the two-probe resistance as a result of the screening described in Ref. 15 for both the attractive and repulsive defects. In addition, for the attractive scatterer of (b), $R_{\text{four-probe}}$ also rises immediately below each subband minimum due to increased

through the defect. Usually stronger scatterers lead to a reduced conductance. But we see this is not necessarily the case when the electrons are confined to move in a wire.

Figure 8 compares the two- and four-probe Landauer conductance for the scatterers from Figs. 4 and 5. Since the four-probe conductance is so large, we plot its inverse. Figure 8(a) shows the resistance of the repulsive scatterer having the same strength as in Figs. 4 and 6. The solid line in Fig. 8(a) is the two-probe resistance, the inverse of the two-probe conductance from Fig. 6. The dashed line in Fig. 8(a) shows the four-probe resistance for the same repulsive scatterer. The four-probe resistance is of course much smaller than the corresponding two-probe resistance due to the absence of any contact resistance. The four-probe resistance is close to zero except when the electron energy is slightly above each subband minimum, where it suddenly rises to equal the two-probe resistance and then falls again. The four-probe resistance being zero slightly below each subband is a consequence of the scatterer being perfectly transmitting, while the rise in the four-probe resistance slightly above each subband minimum is a consequence of screening.[15] Near the bottom of each subband the voltage division factor in Eq. (15) is dominated by the transmitted or reflected waves for that subband. Since each intrasubband reflection coefficient is one at the bottom of its respective subband, while the intersubband reflection is initially zero, the voltage division factor is simply 1 and the two- and four-probe conductances are equal at these special points.

Figure 8(b) shows the resistance of the attractive scatterer having the same strength as in Figs. 5 and 6. The two-probe resistance (solid line) is the inverse of the dashed line from Fig. 6. The four-probe resistance (dashed line) for the same attractive scatterer is again lower than the two-probe resistance due to the absence of contact resistance. Both resistances increase sharply at the quasi-bound-state energies, and are both equal and infinite at the quasi-bound-state in the first subband due to perfect reflection at that point. At subsequent quasi-bound-states, the four-probe resistance is smaller because the reflection at the quasi-bound-states in higher subbands is not perfect. The four-probe resistance is again zero at each subband minimum, and again rises to equal the two-probe resistance just above each subband. Near the subband minima the resistance for the attractive scatterer first rises due to the presence of the quasi-bound-state, then falls to zero due to the unexplained mode "decoupling," then rises again to equal the two-probe resistance as a consequence of screening.

We have also numerically examined the transmission through two δ-function scatterers in a narrow wire,[29] and find unusual modifications to resonant tunneling whenever the Fermi energy is near the second subband minima. In particular resonant transmission is suppressed when distance between the two barriers is equal to an odd number of half wavelengths of the lowest mode, but unaffected when the distance between the barriers equals an even number of half wavelengths. It is also possible to exactly solve the Dyson equation for scattering from a δ-function defect in a Q1D wire, including mode conversion and coupling to the evanescent modes. The same transmission coefficients are obtained as in Appendix B.

V. FINITE-RANGE SCATTERER

In Sec. III and IV we considered the transmission and reflection from a single δ-function scatterer inside a narrow wire. In this section we examine the next simplest case where the scatterer inside the Q1D wire is a rectangle with a finite width and length. Recall the geometry of Fig. 1, shown there for an arbitrarily shaped defect. Analyzing the transmission through a rectangular defect will allow us to understand how to calculate transmission through an arbitrarily shaped defect.

As in Sec. II, the wave functions $c_n(x)$ in region I have coefficients A_n and B_n given by Eqs. (7) and (9), wave vectors k_n given by Eq. (5), and are described by the normal modes $\chi_n(y)$. In region II we will also have coefficients parametrizing the exponential solutions for x motion C_n and D_n. But now, due to the presence of the extra defect, we will have a new set of subbands $\tilde{E}_n$ and a new set of normal modes $Y_n(y)$ which must be calculated numerically by diagonalizing a matrix expanded in the old basis set $\chi_n(y)$. In addition, we have a new set of wave vectors $\tilde{k}_n$ defined by

$$\tilde{k}_n^2=\frac{2m}{\hbar^2}(E-\tilde{E}_n) \tag{34}$$

such that the $c_n(x)$'s in the region $0<x<L$ can be expressed

$$c_n(x)=C_n e^{i\tilde{k}_n x}+D_n e^{-i\tilde{k}_n x} \tag{35}$$

if n denotes a propagating mode in region II and

$$c_n(x)=C_n e^{-\tilde{\kappa}_n x}+D_n e^{\tilde{\kappa}_n x} \tag{36}$$

if n denotes an evanescent mode in region II. Since region III has the same confinement potential as region I, the normal modes $\chi_n(y)$ and wave vectors k in region III will be the same as for region I. We take the coefficients for x motion to be E_n and F_n in region III where

$$c_n(x)=E_n e^{ik_n(x-L)}+F_n e^{-ik_n(x-L)} \tag{37}$$

describes the propagating modes for $x>L$.

Following our discussion of the scattering boundary conditions in Sec. II, enforcing continuity of the wave function and its derivative at each boundary leads to the transfer matrix

$$C_\beta=\tfrac{1}{2}\sum_\alpha\left[\int dy\, Y_\beta^*(y)\chi_\alpha(y)\right]\left[1+\frac{k_\alpha}{\tilde{k}_\beta}\right]A_\alpha + \tfrac{1}{2}\sum_\alpha\left[\int dy\, Y_\beta^*(y)\chi_\alpha(y)\right]\left[1-\frac{k_\alpha}{\tilde{k}_\beta}\right]B_\alpha , \tag{38}$$

$$D_\beta=\tfrac{1}{2}\sum_\alpha\left[\int dy\, Y_\beta^*(y)\chi_\alpha(y)\right]\left[1-\frac{k_\alpha}{\tilde{k}_\beta}\right]A_\alpha + \tfrac{1}{2}\sum_\alpha\left[\int dy\, Y_\beta^*(y)\chi_\alpha(y)\right]\left[1+\frac{k_\alpha}{\tilde{k}_\beta}\right]B_\alpha \tag{39}$$

describing the mode coupling at $x=0$, the transfer matrix describing the propagation on top of the rectangular barrier as

$$C'_\beta = C_\beta e^{i\tilde{k}_\beta L}, \quad D'_\beta = D_\beta e^{-i\tilde{k}_\beta L}, \tag{40}$$

and which does not couple the modes, and finally the transfer matrix

$$E_\gamma = \tfrac{1}{2}\sum_\beta \left[\int dy\, \chi^*_\gamma(y)Y_\beta(y)\right]\left[1+\frac{\tilde{k}_\beta}{k_\gamma}\right]C'_\beta + \tfrac{1}{2}\sum_\beta \left[\int dy\, \chi^*_\gamma(y)Y_\beta(y)\right]\left[1-\frac{\tilde{k}_\beta}{k_\gamma}\right]D'_\beta, \tag{41}$$

$$F_\gamma = \tfrac{1}{2}\sum_\beta \left[\int dy\, \chi^*_\gamma(y)Y_\beta(y)\right]\left[1-\frac{\tilde{k}_\beta}{k_\gamma}\right]C'_\beta + \tfrac{1}{2}\sum_\beta \left[\int dy\, \chi^*_\gamma(y)Y_\beta(y)\right]\left[1+\frac{\tilde{k}_\beta}{k_\gamma}\right]D'_\beta \tag{42}$$

incorporating mode mixing at the $x=L$ potential step.

Multiplying these three transfer matrices numerically will give all the transmission coefficients through the rectangular defect. However, there is a numerical problem with the growing real exponential for D_n in Eq. (40) if we include too many evanescent modes, or if the length of the scatterer L is made too long. The highest evanescent modes, which should become unimportant to the physics of the problem, wind up dominating the transmission properties as well as violating particle number conservation. In place of multiplying the transfer matrices, if we instead cascade scatter matrices as described in Ref. 2 we avoid these numerical problems. An arbitrarily shaped defect can be broken up into many rectangular shaped defects of short lengths, so that equations similar to Eqs. (38), (39), and (40) can be used to calculate the transmission and reflection coefficients through an arbitrarily shaped defect in a narrow wire.

We display the two-probe Landauer conductance for a rectangular shaped scatterer in Fig. 9. Figure 9(a) shows a repulsive scatterer that is 50 nm long (along the x direction), 8.7 nm wide, offset 1.2 nm from the side of the wire, and 50 meV tall. We include six total modes in our calculations, and again assume the original confinement potential to be an infinite square well of 30 nm width. The shape of this defect potential forms an extra constriction inside the Q1D wire. Extra oscillations analogous to the Ramsauer resonances appear in the conductance of Fig. 9(a), known from many previous calculations including Refs. 9, 11, and 12. These Ramsauer-like resonances appear in all the inter- and intrasubband transmission and reflection coefficients, though we do not show these here. We emphasize, however, that the rises in conductance corresponding to the "quantized steps" do not occur at the original subband energies E_n of 6.23, 24.9, 56.1, and 99.7 meV, but rather at the new subband energies $\tilde{E}_n$ of 10.3, 39.4, and 78.3 meV. Due to the presence of evanescent modes below the $\tilde{E}_n$, some carriers can leak through below the new subband minima as shown in the figure.

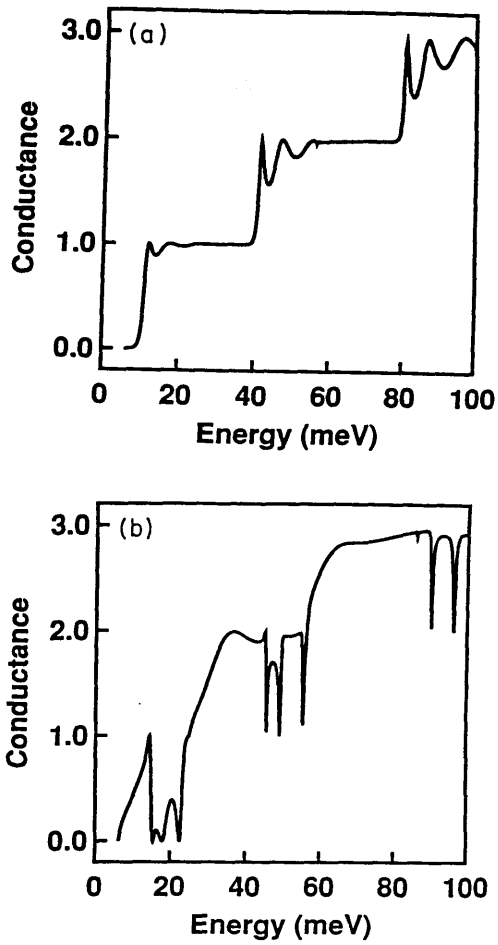


FIG. 9. Conductance in units of $2e^2/h$ through a finite size rectangular barrier whose dimensions are given in the text. (a) shows a repulsive barrier while (b) considers the same size barrier when the sign of the potential is reversed. Ramsauer-like resonances are observed in (a) due to the finite barrier size. The "quantized steps" in (a) occur at the new subband minima inside the barrier, not those of the clean wire. Multiple quasi-bound-states where the conductance falls sharply occur in (b) again due to the finite size of the defect. Transmission is no longer perfect at each minimum as was the case for the δ-function scattering defect.

rises above the 50-meV barrier height as shown in Fig. 9(a).

When we reverse the sign of the scatterer, making it attractive but leaving all other parameters unchanged, we obtain the two-probe conductance shown in Fig. 9(b). The new subband energies $\tilde{E}_n$ for those potential are at -20.6, 13.3, 44.0, and 84.3 meV. Three bound states can be seen in the conductance of Fig. 9(b) immediately before the original second and third subband minima. Our point in Figs. 9(a) and 9(b) is that, as we give the scatterer a finite size, (1) we obtain geometrical resonances with the defect analogous to the Ramsauer resonances, (2) there can be multiple bound states if the scatterer is attractive, (3) there is more than one set of subbands in the problem which are important, and (4) the transmission is no

shown in the figures. For example, T_{11} has quite a complex behavior for both potentials considered in this section. Also we have verified that, as we shrink the size of the scatterer while maintaining the total area underneath the potential, the results for transmission coefficients evolve into those for the δ-function scatterer studied in the preceding two sections.

VI. CONCLUSIONS

In this paper we have considered the scattering from defects in a confined geometry, in our case for electrons confined to a quasi-one-dimensional wire. The scattering properties of electrons in these confined geometries are qualitatively different from the usual case of scattering in open geometries due to the building up and storage of electrons in evanescent waves near the scattering defect. This phenomenon is well known for the analogous case of electromagnetic wave scattering in microwave waveguides. We obtained both the intersubband and intrasubband transmission and reflection coefficients by hand for the special case of a δ-function scatterer, and numerically found these same coefficients for the case when the scatterer has a finite size. We then used these coefficients to study the two- and four-probe Landauer conductance through the defect.

For the case of the δ-function scatterer we found that, for electron energies equal to the subband minima, all the normal modes completely decouple and that conductance through the defect is equivalent to the ballistic conductance in agreement with Ref. 23. This unusual result is independent of where the scatterer is located in the wire, the shape of the confining potential, and the subband separation (i.e., independent of the incident electron wavelength). This "perfect transmission" effect therefore seems to be independent of any exact geometrical feature in the problem except for the existence of confinement. These facts support our conclusion that the existence of evanescent modes is the qualitatively new and important feature responsible for the unusual scattering properties in the confined geometry, rather than a geometrical resonance between the scatterer and the walls of the wire. In addition, if the δ-function scatterer is made attractive, a quasi-bound-state splitting off of each Q1D subband forms in the potential and dominates the scattering properties when the electron energy is near the quasi-bound-state energy.

For the finite size scatter which we explored briefly in Sec. V, the overall features in the transmission amplitudes and the conductance qualitatively resemble those of the δ-function scatterer. However, the transmission is no longer perfect at each subband minimum. We therefore attribute the perfect transmission seen for the δ-function potential to the special shape of the scatterer, and conclude that it is not a general feature for any possible shape of the scattering potential. Reference 23 concludes that any *s*-wave scatterer will exhibit the perfect transmission property. The δ function is an *s*-wave scatterer while the square barrier is not. We can conclude in general that the transmission is enhanced near a subband minimum and reflection increased near any exhibits geometrical resonances (analogous to the Ramsauer resonances) in all the transmission coefficients, both intersubband and intrasubband. A finite size scatterer which is attractive can have multiple quasi-bound-states splitting off of the higher confinement subbands, and the effects of these can be seen in the transmission probability through the defect.

Note added in proof. We have learned after submitting this manuscript of a complementary study of impurities in a constriction by Tekman and Ciraci,[31] of a quantum cavity resonator by Peeters,[32] and of the influence of evanescent waves on electron localization in the limit of a large number of impurities by Cahay *et al.*[33]

ACKNOWLEDGMENTS

We thank Terry P. Orlando, Arvind Kumar, Marc A. Kastner, Charles Kane, Dimitri A. Antoniadis, and Henry I. Smith for useful discussions. This work was sponsored by the U.S. Air Force Office of Scientific Research under Grant No. AFOSR-88-0304 and the U.S. Joint Services Electronics Program Contract No. DAAL03-89-C-0001.

APPENDIX A: QUASI-BOUND-STATES FOR AN ATTRACTIVE δ-FUNCTION POTENTIAL IN A Q1D WIRE

The bound-state energy and transmission coefficient through a δ-function potential $V(x)=V_0\delta(x)$ in one dimension can easily be evaluated[30] to find

$$E=-\frac{\hbar^2}{2m}\left[\frac{\Gamma}{2}\right]^2 \tag{A1}$$

$$T(E)=\left[1+\frac{\Gamma^2}{4k^2}\right]^{-1} \tag{A2}$$

where $\Gamma=2mV_0/\hbar^2$ and the bound-state energy is a solution of the equation $2\kappa+\Gamma=0$ where κ is the inverse decay length.

In this appendix we wish to discuss the quasi-bound-states for a δ function in a Q1D wire in order to understand the similarities and differences between the Q1D case and the 1D case worked out in Ref. 30. The potential $V(x,y)=\gamma\delta(x)\delta(y-y_i)$ is attractive so that $\gamma<0$. We consider a wave function which consists of a sum of exponentially decaying modes as

$$\psi(x,y)=\sum_n B_n e^{-\kappa_n|x|}\chi_n(y)\ , \tag{A3}$$

so that the matrix equation for the B_n's, taking four evanescent modes for definiteness, is

$$\begin{bmatrix}0\\0\\0\\0\end{bmatrix}=\begin{bmatrix}\Gamma_{11}+2\kappa_1 & \Gamma_{12} & \Gamma_{13} & \Gamma_{14}\\ \Gamma_{21} & \Gamma_{22}+2\kappa_2 & \Gamma_{23} & \Gamma_{24}\\ \Gamma_{31} & \Gamma_{32} & \Gamma_{33}+2\kappa_3 & \Gamma_{34}\\ \Gamma_{41} & \Gamma_{42} & \Gamma_{43} & \Gamma_{44}+2\kappa_4\end{bmatrix}\begin{bmatrix}B_1\\B_2\\B_2\\B_4\end{bmatrix}. \tag{A4}$$

The same method of rescaling the mode coupling constants applies for this case as in the previous discussion of Sec. III, so that for m total modes the bound-states ener-

$$\Gamma_{11,2-m}+2\kappa_1=0\ . \tag{A5}$$

Call the energy which solves this equation E_{1b}. This bound-state energy will satisfy $E_{1b}<E_1$, so that it is below the first subband energy. The energy E_{1b} is a true bound state which is a stationary solution of the Schrödinger equation. However, if mode one is a propagating mode and we set $k_1=i\kappa_1$ in Eq. (A4), we cannot obtain a solution to the equation. This implies there are no "truly bound" states in the wire which are degenerate in energy with any of the propagating modes. But if we throw away all of the mode coupling to the propagating modes, in this case by neglecting the first row and first column of the matrix in Eq. (A4), we obtain a quasi-bound-state energy E_{2b} below the second subband given by

$$\Gamma_{22,3-m}+2\kappa_2=0\ , \tag{A6}$$

where the quasi-bound-state solution E_{2b} to this equation satisfies $E_{2b}<E_2$. If the δ function is made attractive enough we can even have $E_{2b}<E_1$. The energies E_{2b}, E_{3b}, etc., are important in understanding the scattering properties of electrons from the δ-function defect. Since the quasi-bound-states at energies E_{b2}, E_{b3}, etc., actually do couple to the propagating modes of the wire, they will decay with time. Only the bound state splitting off of the lowest subband at energy E_{b1} is a truly bound state.

Equation (A5) can be rewritten as

$$1+\sum_{i=1}^{m}\frac{\Gamma_{ii}}{2\kappa_i}=0\ . \tag{A7}$$

The bare mode coupling constants appear in Eq. (A7), not rescaled by any of the evanescent modes. Equation (A6) can be rewritten similarly as

$$1+\sum_{i=2}^{m}\frac{\Gamma_{ii}}{2\kappa_i}=0\ , \tag{A8}$$

and so on for all the possible quasi-bound-states. Clearly, the terms in the sum become smaller and smaller quite rapidly as the κ_i's become larger and larger.

Sketching the terms in the sums of Eqs. (A7) and (A8) shows that there can only be one bound-state or quasi-bound-state solution to each equation. The sketch also shows that the presence of all the other evanescent modes acts to lower the bound-state energy below its value in the absence of all the other evanescent modes. Therefore an upper bound on the solution of Eq. (A7), for example, is $\Gamma_{11}+2\kappa_1=0$. Since the solution to this equation is below E_1, and the solution to Eq. (A7) will have an even lower energy as we can see graphically, we must have the bound-state solution to Eq. (A5) satisfy $E_{1b}<E_1$.

Next consider Eq. (A8), which describes only modes two and higher allowed as evanescent waves. The quasi-bound-state energy solution to Eq. (A8) does not have to be below the lowest subband, in contrast to the case of the δ function in 1D. Even if the first mode is propagating, there can still be a quasi-bound-state split of from the second subband given by the solution to Eq. (A8) at any energy below E_2. An upper limit to the energy E_{b2}

APPENDIX B: SOLUTIONS TO COUPLED SCATTERING EQUATIONS

Here we wish to write down the general solution to the infinite set of coupled scattering equations, Eq. (4), for the case of a single δ-function scatterer. For this case Eq. (4) can be expressed in a form similar to Eq. (23) for any number of propagating incident modes. We proceed as follows: We write down the solutions first for the case of a single propagating mode, then for two propagating modes, then for three, etc. It is these solutions which we have plotted in Sec. IV. In writing down these solutions we will recognize a pattern that, when generalized, leads to the transmission coefficients for any arbitrary number of propagating modes. We have already considered the effect of the evanescent modes in Sec. III, so our solutions are valid for any arbitrary number of evanescent modes by simply rescaling the mode coupling constants. This rescaling procedure is only valid for the case of a single δ-function scatterer in a wire.

Our main result in this appendix is to show that the reflection amplitudes $\tilde{r}_{ab}$ can be written as

$$\tilde{r}_{ab}(E)=\frac{-i\dfrac{\Gamma_{ab}}{2\sqrt{k_a k_b}}}{1+\sum_{n}^{e}(\Gamma_{nn}/2\kappa_n)+i\sum_{n}^{p}(\Gamma_{nn}/2k_n)}\ . \tag{B1}$$

In Eq. (B1) $\sum^e$ denotes a sum over all the evanescent modes, $\sum^p$ denotes a sum over the propagating modes, and modes a and b are assumed propagating. Equation (B1) also holds for the intersubband transmission amplitudes since $\tilde{r}_{ab}(E)=\tilde{t}_{ab}(E)$ for $a\neq b$. The intrasubband transmission amplitude $\tilde{t}_{aa}(E)$ can then be obtained from the constraint of current conservation. The simple analytic result of Eq. (B1) can be interpreted in terms of the Fermi golden-rule scattering rate as we argue below, and has been obtained more transparently from the solution of Dyson's equation for this same scattering problem. Equation (B1) is a compact way of writing down the results listed in this appendix, and can be obtained from them by the same type of algebra used to go, for example, from Eq. (A5) to Eq. (A7).

We now write down the solutions for the transmission amplitudes and transmission probabilities for m total modes as the number of propagating modes ranges between one and three. When transport is only in the lowest subband, namely, $E_1<E<E_2$, we must solve the matrix Eq. (28) or Eq. (29) given in Sec. III to obtain

$$t_{11}=1/d_1=1+r_{11}\ , \tag{B2}$$

$$T_{11}=\frac{k_1}{k_1}t_{11}t_{11}^{*}=\frac{1}{D_1}\ , \tag{B3}$$

where the denominator amplitude d_1 is

$$d_1=1-\frac{\Gamma_{11,2-m}}{2ik_1} \tag{B4}$$

yielding the denominator

$$D_1=d_1d_1^{*}=1+\left[\frac{\Gamma_{11,2-m}}{2k_1}\right]^2\ . \tag{B5}$$

Note that Eq. (B3) has the same form as the transmission coefficient through a δ function in one dimension given in Appendix A, except that we make the substitution $\Gamma \rightarrow \Gamma_{11,2-m}$.

The reflection coefficients are

$$r_{11}=\left[\frac{\Gamma_{11,2-m}}{2ik_1}\right]\frac{1}{d_1}\ , \tag{B6}$$

$$R_{11}=\frac{k_1}{k_1}r_{11}r_{11}^{*}=\left[\frac{\Gamma_{11,2-m}}{2k_1}\right]^2\frac{1}{D_1} \tag{B7}$$

from which the current conservation identity

$$R_{11}+T_{11}=1 \tag{B8}$$

is satisfied. Note also that $0\le T_{11}\le 1$ since $1\le D_1\le\infty$, so that the normalized two-probe conductance in units of $2e^2/h$ is between zero and one when only the lowest subband is occupied.

We would like to give Eq. (B3) for T_{11} and Eq. (B7) for R_{11} an interpretation in terms of the Fermi golden rule for scattering rates. Our viewpoint is that T_{11} does not describe any scattering processes (for this problem), and is just the remaining particle flux left over by current continuity after any scattering processes (reflection or intersubband transmission) subtract particles from the incident beam. The only scattering process when one subband is occupied is intrasubband reflection. The numerator for R_{11} can be approximated as

$$\left[\frac{\Gamma_{11}}{2k_1}\right]^2=\frac{2}{\pi}\frac{m}{\hbar^2k_1}\left[\frac{\pi}{2}\gamma|\chi_1(y_i)|^2\right]^2\frac{2}{\pi}\frac{m}{\hbar^2k_1}$$
$$=N_1(E)\left|\frac{\pi}{2}\langle 1|\gamma\delta(y-y_i)|1\rangle\right|^2N_1(E) \tag{B9}$$

which is reminiscent of the Fermi golden rule for transition rates. The numerator of R_{11} is simply the density of initial states in subband one, the square of a matrix element coupling subband one to itself through the scattering potential, and the density of final states in subband one. Each term in Eq. (B1) can be similarly interpreted.

To lowest order in the scattering potential we would just have $T_{11}=1$ (i.e., the particle would pass through the obstacle undisturbed) and R_{11} just equal to its numerator from Eq. (B9), which is the Fermi golden-rule transition rate. But then particle continuity would be violated. Hence the denominator D_1 from Eq. (B5) is determined by the sum of the scattering probabilities for each process to the lowest order in the scattering potential. This denominator D_1 preserves particle continuity.

Now we consider transport for energies in which both the first and second subbands are occupied. For $E_2<E<E_s$ we have two propagating modes and four transmission coefficients. Consider first the case where only mode one is incident on the scatterer. We obtain a matrix analogous to Eq. (23)

$$\begin{bmatrix}-2ik_1\\0\\0\\0\end{bmatrix}=\begin{bmatrix}\Gamma_{11}-2ik_1 & \Gamma_{12} & \Gamma_{13} & \Gamma_{14}\\ \Gamma_{21} & \Gamma_{22}-2ik_2 & \Gamma_{23} & \Gamma_{24}\\ \Gamma_{31} & \Gamma_{32} & \Gamma_{33}+2\kappa_3 & \Gamma_{34}\\ \Gamma_{41} & \Gamma_{42} & \Gamma_{43} & \Gamma_{44}+2\kappa_4\end{bmatrix}\begin{bmatrix}t_{11}\\t_{12}\\t_{13}\\t_{14}\end{bmatrix}. \tag{B10}$$

Solving the matrix Eq. (B10) we obtain the transmission amplitude t_{11} as

$$t_{11}=\left[1-\frac{\Gamma_{22,3-m}}{2ik_2}\right]\frac{1}{d_2}=1+r_{11} \tag{B11}$$

yielding the transmission coefficient T_{11}

$$T_{11}=\frac{k_1}{k_1}t_{11}t_{11}^{*}=\left[1+\left[\frac{\Gamma_{22,3-m}}{2k_2}\right]^2\right]\frac{1}{D_2} \tag{B12}$$

where the denominator amplitude d_2 is

$$d_2=1-\frac{\Gamma_{11,3-m}}{2ik_1}-\frac{\Gamma_{22,3-m}}{2ik_2} \tag{B13}$$

and the denominator D_2 is

$$D_2=d_2d_2^{*}=1+\left[\frac{\Gamma_{11,3-m}}{2k_1}+\frac{\Gamma_{22,3-m}}{2k_2}\right]^2. \tag{B14}$$

Note that from Eq. (B12) we can recover Eq. (B3) by setting $k_2\rightarrow i\kappa_2$. The reflection coefficient R_{11} is obtained from

$$r_{11}=\frac{\Gamma_{11,3-m}}{2ik_1}\frac{1}{d_2}\ , \tag{B15}$$

$$R_{11}=\frac{k_1}{k_1}r_{11}r_{11}^{*}=\left[\frac{\Gamma_{11,3-m}}{2k_1}\right]^2\frac{1}{D_2}\ . \tag{B16}$$

The remaining intersubband transmission amplitudes and transmission coefficients from Eq. (B10) are

$$t_{12}=\frac{\Gamma_{21,3-m}}{2ik_2}\frac{1}{d_2}=r_{12}\ , \tag{B17}$$

$$T_{12}=\frac{k_2}{k_1}t_{12}t_{12}^{*}=\frac{\Gamma_{11,3-m}}{2k_1}\frac{\Gamma_{22,3-m}}{2k_2}\frac{1}{D_2}=R_{12}\ . \tag{B18}$$

Equation (23) describes the transmission properties for two occupied subbands when mode two is incident on the defect. We obtain the solutions to Eq. (23) as

$$t_{21}=\frac{\Gamma_{12,3-m}}{2ik_1}\frac{1}{d_2}=r_{21}\ , \tag{B19}$$

$$T_{21}=\frac{k_1}{k_2}t_{21}t_{21}^{*}=\frac{\Gamma_{11,3-m}}{2k_1}\frac{\Gamma_{22,3-m}}{2k_2}\frac{1}{D_2}=R_{21} \tag{B20}$$

so that $T_{21}=T_{12}$ and

$$t_{22}=\left[1-\frac{\Gamma_{11,3-m}}{2ik_1}\right]\frac{1}{d_2}=1+r_{22}\ , \tag{B21}$$

$$T_{22}=\frac{k_2}{k_2}t_{22}t_{22}^*=\left[1+\left[\frac{\Gamma_{11,3-m}}{2k_1}\right]^2\right]\frac{1}{D_2}\,, \tag{B22}$$

$$r_{22}=\frac{\Gamma_{22,3-m}}{2ik_2}\frac{1}{d_2}\,, \tag{B23}$$

$$R_{22}=\frac{k_2}{k_2}r_{22}r_{22}^*=\left[\frac{\Gamma_{22,3-m}}{2k_2}\right]^2\frac{1}{D_2}\,. \tag{B24}$$

The current conservation identities

$$T_{11}+T_{12}+R_{11}+R_{12}=1\,, \tag{B25}$$

$$T_{21}+T_{22}+R_{21}+R_{22}=1 \tag{B26}$$

are also satisfied. The identity

$$T_{11}+T_{12}+T_{21}+T_{22}=1+\frac{1}{D_2} \tag{B27}$$

limits the normalized two-probe conductance to be between one and two for two occupied subbands since $1\le D_2\le\infty$.

We would like to consider a Fermi golden-rule-like interpretation of the transmission and reflection coefficients for two occupied subbands. Analogously to our interpretation for R_{11} when one subband is occupied, given by Eq. (B9), the numerators of R_{11} from Eq. (B16), T_{12} from Eq. (B18), R_{12} from Eq. (B18), T_{21} from Eq. (B20), R_{21} from Eq. (B20), and R_{22} from Eq. (B24) can be regarded as the golden-rule scattering rates for their respective scattering processes. The denominator D_2 from Eq. (B14) can similarly be regarded as the sum of all the scattering processes.

For transport for three occupied subbands in the energy range $E_3<E<E_4$, the pattern from our calculation of the transmission coefficients for one and two occupied subbands continues as before. The matrix equations we solve to find all the scattering coefficients are very similar to Eq. (23) from the text and Eq. (B10) from this appendix, so we do not write them down. Instead, we merely catalogue the solutions for three propagating subbands. The transmission amplitude t_{11} is

$$t_{11}=\left[1-\frac{\Gamma_{22,4-m}}{2ik_2}-\frac{\Gamma_{33,4-m}}{2ik_3}\right]\frac{1}{d_3}=1+r_{11}\,, \tag{B28}$$

$$T_{11}=\frac{k_1}{k_1}t_{11}t_{11}^*=\left[1+\left[\frac{\Gamma_{22,4-m}}{2k_2}+\frac{\Gamma_{33,4-m}}{2k_3}\right]^2\right]\frac{1}{D_3} \tag{B29}$$

where the denominator amplitude d_3 is

$$d_3=1-\frac{\Gamma_{11,4-m}}{2ik_1}-\frac{\Gamma_{22,4-m}}{2ik_2}-\frac{\Gamma_{33,4-m}}{2ik_3} \tag{B30}$$

and

$$D_3=d_3d_3^*=1+\left[\frac{\Gamma_{11,4-m}}{2k_1}+\frac{\Gamma_{22,4-m}}{2k_2}+\frac{\Gamma_{33,4-m}}{2k_3}\right]^2. \tag{B31}$$

The reflection coefficient is obtained from

$$r_{11}=\frac{\Gamma_{11,4-m}}{2ik_1}\frac{1}{d_3}\,, \tag{B32}$$

$$R_{11}=\frac{k_1}{k_1}r_{11}r_{11}^*=\left[\frac{\Gamma_{11,4-m}}{2k_1}\right]^2\frac{1}{D_3}\,. \tag{B33}$$

The remaining coefficients are

$$t_{12}=\frac{\Gamma_{21,4-m}}{2ik_2}\frac{1}{d_3}=r_{12}\,, \tag{B34}$$

$$T_{12}=\frac{k_2}{k_1}t_{12}t_{12}^*=\frac{\Gamma_{11,4-m}}{2k_1}\frac{\Gamma_{22,4-m}}{2k_2}\frac{1}{D_3}=R_{12}\,, \tag{B35}$$

$$t_{13}=\frac{\Gamma_{31,4-m}}{2ik_3}\frac{1}{d_3}=r_{13}\,, \tag{B36}$$

$$T_{13}=\frac{k_3}{k_1}t_{13}t_{13}^*=\frac{\Gamma_{11,4-m}}{2k_1}\frac{\Gamma_{33,4-m}}{2k_3}\frac{1}{D_3}=R_{13}\,, \tag{B37}$$

$$t_{21}=\frac{\Gamma_{12,4-m}}{2ik_1}\frac{1}{d_3}=r_{21}\,, \tag{B38}$$

$$T_{21}=\frac{k_1}{k_2}t_{21}t_{21}^*=\frac{\Gamma_{11,4-m}}{2k_1}\frac{\Gamma_{22,4-m}}{2k_2}\frac{1}{D_3}=R_{21}\,, \tag{B39}$$

giving $T_{21}=T_{12}$. We write

$$t_{22}=\left[1-\frac{\Gamma_{11,4-m}}{2ik_1}-\frac{\Gamma_{33,4-m}}{2ik_3}\right]\frac{1}{d_3}=1+r_{22}\,, \tag{B40}$$

$$T_{22}=\frac{k_2}{k_2}t_{22}t_{22}^*=\left[1+\left[\frac{\Gamma_{11,4-m}}{2k_1}+\frac{\Gamma_{33,4-m}}{2k_3}\right]^2\right]\frac{1}{D_3}\,, \tag{B41}$$

$$r_{22}=\frac{\Gamma_{22,4-m}}{2ik_2}\frac{1}{d_3}\,, \tag{B42}$$

$$R_{22}=\frac{k_2}{k_2}r_{22}r_{22}^*=\left[\frac{\Gamma_{22,4-m}}{2k_2}\right]^2\frac{1}{D_3}\,, \tag{B43}$$

$$t_{23}=\frac{\Gamma_{32,4-m}}{2ik_3}\frac{1}{d_3}=r_{23}\,, \tag{B44}$$

$$T_{23}=\frac{k_3}{k_2}t_{23}t_{23}^*=\frac{\Gamma_{33,4-m}}{2k_3}\frac{\Gamma_{22,4-m}}{2k_2}\frac{1}{D_3}=R_{23}\,, \tag{B45}$$

$$t_{31}=\frac{\Gamma_{13,4-m}}{2ik_1}\frac{1}{d_3}=r_{31}\,, \tag{B46}$$

$$T_{31}=\frac{k_1}{k_3}t_{31}t_{31}^*=\frac{\Gamma_{33,4-m}}{2k_3}\frac{\Gamma_{11,4-m}}{2k_1}\frac{1}{D_3}=R_{31}\,, \tag{B47}$$

so that $T_{13}=T_{31}$,

$$t_{32}=\frac{\Gamma_{23,4-m}}{2ik_2}\frac{1}{d_3}=r_{32}\,, \tag{B48}$$

$$T_{32}=\frac{k_2}{k_3}t_{32}t_{32}^*=\frac{\Gamma_{33,4-m}}{2k_3}\frac{\Gamma_{22,4-m}}{2k_2}\frac{1}{D_3}=R_{32}\,, \tag{B49}$$

leaving $T_{23}=T_{32}$, and finally

$$t_{33}=\left[1-\frac{\Gamma_{22,4-m}}{2ik_2}-\frac{\Gamma_{33,4-m}}{2ik_3}\right]\frac{1}{d_3}=1+r_{33}\ , \tag{B50}$$

$$T_{33}=\frac{k_3}{k_3}t_{33}t_{33}^*=\left[1+\left[\frac{\Gamma_{11,4-m}}{2k_1}+\frac{\Gamma_{22,4-m}}{2k_2}\right]^2\right]\frac{1}{D_3}\ , \tag{B51}$$

$$r_{33}=\frac{\Gamma_{33,4-m}}{2ik_3}\frac{1}{d_3}\ , \tag{B52}$$

$$R_{33}=\frac{k_3}{k_3}r_{33}r_{33}^*=\left[\frac{\Gamma_{33,4-m}}{2k_3}\right]^2\frac{1}{D_3}\ . \tag{B53}$$

The current conservation identities

$$T_{11}+T_{12}+T_{13}+R_{11}+R_{12}+R_{13}=1\ , \tag{B54}$$

$$T_{21}+T_{22}+T_{23}+R_{21}+R_{22}+R_{23}=1\ , \tag{B55}$$

$$T_{31}+T_{32}+T_{33}+R_{31}+R_{32}+R_{33}=1 \tag{B56}$$

are also satisfied. The identity

$$T_{11}+T_{12}+T_{13}+T_{21}+T_{22}+T_{23}+T_{31}+T_{32}+T_{33}=2+\frac{1}{D_3} \tag{B57}$$

limits the two-probe conductance to be between two and three when three subbands are occupied since $1\le D_3\le\infty$.

The equations in this appendix are plotted in Sec. IV. By continuing the pattern from the above equations, we can obtain a general solution for any number of propagating modes in the waveguide scattering from the δ-function defect given compactly in Eq. (B1).

[1]S. Datta, M. Cahay, and M. McLennan, Phys. Rev. B **36**, 5655 (1987).

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

[3]Patrick A. Lee and Daniel S. Fisher, Phys. Rev. Lett. **47**, 882 (1981); D. J. Thouless and S. Kirkpatrick, J. Phys. C **14**, 235 (1981); Alfredo Levy Yeyati, Mariana Weissmann, and Enrique Anda, J. Phys. Condens. Matter **1**, 5429 (1989).

[4]Harold U. Baranger and A. Douglas Stone, Phys. Rev. Lett. **63**, 414 (1989).

[5]D. G. Ravenhall, H. W. Wyld, and R. L. Schult, Phys. Rev. Lett. **62**, 1780 (1989).

[6]George Kirczenow, Phys. Rev. Lett. **62**, 2993 (1989).

[7]Yshai Avishai and Yehuda B. Band, Phys. Rev. Lett. **62**, 2527 (1989); Phys. Rev. B **40**, 3429 (1989).

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

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

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

[11]Aaron Szafer and A. Douglas Stone, Phys. Rev. Lett. **62**, 300 (1989).

[12]George Kirczenow, J. Phys. Condens. Matter **1**, 305 (1989); Solid State Commun. **68**, 715 (1988); Phys. Rev. B **39**, 10 452 (1989); see also Eleuterio Castaño and George Kirczenow (unpublished).

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

[14]H. L. Engquist and P. W. Anderson, Phys. Rev. B **24**, 1151 (1981).

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

[16]Daniel S. Fischer and Patrick A. Lee, Phys. Rev. B **23**, 6851 (1981).

[17]M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986); see also IBM J. Res. Dev. **32**, 317 (1988).

[18]A. Douglas Stone and Aaron Szafer, IBM J. Res. Dev. **32**, 384 (1988).

[19]Rolf Landauer, Z. Phys. B **68**, 217 (1987).

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

[21]Rolf Landauer, J. Phys. Condens. Matter. **1**, 8099 (1989).

[22]Rolf Landauer, in *Analogies in Optics and Micro-Electronics*, edited by W. van Haeringen and D. Lenstra (Kluwer Academic, in press).

[23]C. S. Chu and R. S. Sorbello, Phys. Rev. B **40**, 5941 (1989).

[24]Philip F. Bagwell and Terry P. Orlando, Phys. Rev. B **40**, 1456 (1989).

[25]Philip F. Bagwell, Tom P. E. Broekaert, Terry P. Orlando, and Clifton G. Fonstad (unpublished).

[26]M. Büttiker, in *Electronic Properties of Multilayers and Low Dimensional Semiconductor Structures*, edited by J. M. Chamberlain, L. Eaves, and J. C. Portal (Plenum, New York, in press).

[27]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).

[28]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).

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

[30]Kurt Gottfried, *Quantum Mechanics* (Benjamin, Reading, Mass., 1974), Vol. 1, see p. 51.

[31]E. Tekman and S. Ciraci (unpublished).

[32]Francois M. Peeters, *Bound and Resonant States in Quantum Wire Structures*, in *Science and Engineering of 1- and 0-Dimensional Semiconductors*, edited by S. Beaumont and C. Sotomayor-Torres (Plenum, New York, in press).

[33]M. Cahay, S. Bandyopadhyay, M. A. Osman, and H. L. Grubin (unpublished).

Appendix G

Philip F. Bagwell, 'Solution of Dyson's Equation in a Quasi-1D Wire', Journal of Physics: Condensed Matter, 2, 6179 (1990).

J. Phys.: Condens. Matter 2 (1990) 6179–6188. Printed in the UK

Solution of Dyson's equation in a quasi-1D wire

Philip F Bagwell

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

Received 15 February 1990

Abstract. We obtain the current transmission amplitudes as a function of Fermi energy for electrons scattering from a defect in a quasi-one-dimensional wire by solving Dyson's equation for the single-electron Green function. Dyson's equation in a confined geometry includes both mode conversion and coupling to all the evanescent modes in the wire. After obtaining the Green functions, we use Fisher and Lee's relationship between the single-electron Green functions and the current transmission amplitudes through the defect to find all the intersubband and intrasubband transmission probabilities. In agreement with a previous calculation of the transmission amplitudes performed by simply matching wavefunctions at the defect boundary, evanescent modes are shown to dominate the scattering properties whenever the Fermi energy approaches either a new confinement subband or a quasi-bound state splitting off from the higher-lying confinement subbands.

1. Introduction

Electron scattering in a confined geometry is qualitatively different from scattering in an open geometry due to the existence of evanescent modes introduced by the confinement [1]. Figure 1 shows a case where only the lowest normal mode is incident on a defect in a wire. In figure 1 the second and higher normal modes are evanescent waves which decay along the x direction of propagation. The scattering defect couples propagating modes in the wire both to each other and to all the evanescent modes through the scattering boundary conditions. Therefore, for a steady current flow incident on a defect in the wire, a localised mode will build up around the defect even if the scatterer is repulsive. These extra stored electrons cannot collect around a single defect in an open geometry where the electrons must scatter into a travelling wave which propagates away from the defect.

In this paper we consider the scattering from a single delta function defect in a quasi-one-dimensional wire. Dyson's equation for the single-electron Green function is exactly soluble for this special potential. Mode conversion as well as scattering into all the evanescent modes from each higher-lying confinement subband are included in the Dyson equation describing scattering in a confined geometry. We have already examined this problem using a simpler method of matching wavefunctions and their derivatives at the defect [1]. Here we show that the same transmission and reflection coefficients result from the solution of Dyson's equation. Some additional insight can be gained into the resulting unusual scattering properties [1–4] by considering the class of scattering diagrams which dominate our solution of Dyson's equation.

0953-8984/90/286179+10$03.50

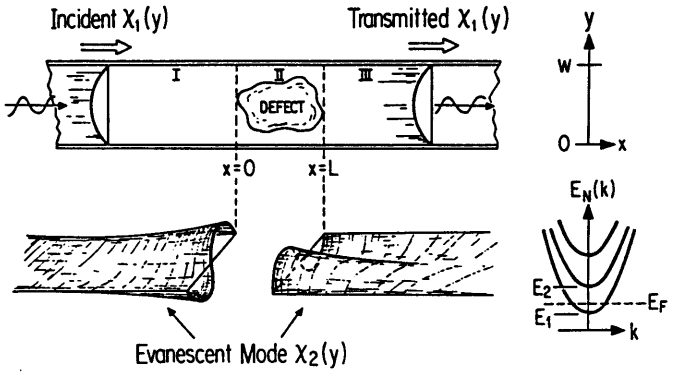


Figure 1. A single scattering defect in a quasi-one-dimensional wire. The wire is assumed to be infinitely long on either side of the defect. For carriers incident only in the lowest subband as shown, evanescent waves build up on either side of the defect in the second and higher normal modes. The building up and storage of electrons around a single scatterer is a unique feature of scattering in a confined geometry such as a wire.

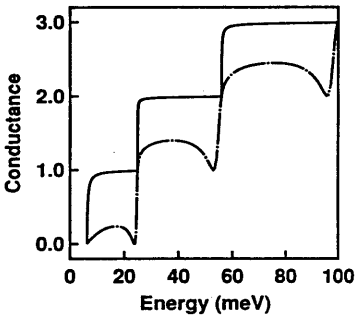


Figure 2. Two-probe Landauer conductance versus electron Fermi energy through a repulsive delta function defect (full curve) and an attractive delta function defect (chain curve) in a quasi-1D wire. A point of 'perfect transparency', where the conductance is equal to its ballistic value of $(n-1)$ times $2e^2/h$, appears immediately below the nth subband minima. The extra drops in the conductance through the attractive potential correspond to extra quasi-bound states in the wire which have split off from the confinement subbands.

The paradigm for calculating conductivity in phase coherent structures has shifted almost exclusively to the viewpoint of Landauer and the various Landauer conductance formulae [5]. Derivations of these conductance formulae from linear response theory have been criticised by Landauer [6] for failure to include the necessary adiabatic widening from a narrow to a wide region. These criticisms can be further appreciated by comparing the electrostatic and electrochemical potentials for a geometry with and without a constriction, given in figure 1 and figure 2 of [7], to the geometries without a constriction considered in [8–10]. For the geometries considered in [8–10] it is difficult to understand how the necessary electrostatic potential drop associated with the quantum contact resistance can develop, for at what point in space can the electrostatic potential $V(x)$ in the perfect leads correspond to that in figure 1 of [7]?

However, in these same studies of conductance using linear response theory [8–10], a relation has been derived between the one-particle Green functions and the current transmission amplitudes through a disordered region beginning with the work of Fisher and Lee [8] and also discussed in [9–11]. Here we confirm that the Fisher–Lee relationship is satisfied for the special case of a delta function potential in a quasi-1D wire by explicit calculation of both the single-particle Green functions and the transmission amplitudes. 'Two-probe' Landauer formulae can therefore be used to calculate conductance when the transmission probabilities are obtained either by straightforwardly matching wavefunctions and their derivatives at the disordered regions or by the more complicated recursive evaluation of Dyson's equation for the single-particle Green functions [12]. Detailed background and bibliography of previous work on the conductance of phase-coherent electron devices can be obtained from the citations in [1–11].

2. Dyson's equation in a quasi-1D wire

Consider again the quasi-one-dimensional wire having electrons confined along the y direction but free to move along the x direction shown in figure 1. The two-dimensional 'wire' of this paper is a reasonable approximation to real physical systems where the confinement along z, usually normal to a semiconductor heterojunction interface, is much stronger than the 'lateral' confinement along y. Furthermore, if the additional confinement along z is taken into account the results for the transmission coefficients presented in section 4 do not change substantially, so it is adequate to work with the simpler two-dimensional 'wire'.

The equation of motion for the Green function G in the quasi-1D wire of figure 1 is

$$\left\{E - \left[-\frac{\hbar^2}{2m}\left(\frac{\mathrm{d}^2}{\mathrm{d}x^2} + \frac{\mathrm{d}^2}{\mathrm{d}y^2}\right) + V_c(y) + V_d(x,y)\right]\right\} G(xy;x'y') = \delta(x-x')\delta(y-y') \tag{1}$$

where the confinement potential $V_c(y)$ depends only on the transverse direction y and $V_d(x,y)$ is the potential of any defects or impurities in the wire. Throughout our discussion we assume propagation at a constant energy E and do not write the energy argument in the Green functions. $G(xy;x'y')$ from (1) has the standard interpretation as being the 'transmission amplitude' that a unit impulse of probability amplitude originally deposited at position (x',y') will propagate to position (x,y). The one-dimensional Schrödinger equation along y including only the confinement potential $V_c(y)$

$$\left(-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}y^2} + V_c(y)\right)\chi_n(y) = E_n\chi_n(y) \tag{2}$$

gives rise to a set of normal modes $\chi_n(y)$ and subband energies E_n where n is the subband index. In addition to their standard completeness property, the $\chi_n(y)$ can be chosen real and obey the useful relation

$$\sum_n \chi_n(y)\chi_n(y') = \delta(y-y'). \tag{3}$$

Multiplying (1) on the left by $\chi_a(y)$, on the right by $\chi_c(y)$, and applying the useful equation (3), we obtain the equation of motion for the Green function in a Q1D wire as

$$\sum_b \left\{ \left[E - E_b - \left(-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} \right) \right] \delta_{ab} - V_{ab}(x) \right\} G(xb;x'c) = \delta(x-x')\delta_{ac}. \tag{4}$$

The matrix elements of the defect potential in (4) are

$$V_{ab}(x) = \int \mathrm{d}y\, \chi_a(y) V_{\mathrm{d}}(x,y) \chi_b(y) \tag{5}$$

and the matrix elements of the Green function are

$$G(xb;x'c) = \int \mathrm{d}y \int \mathrm{d}y'\, \chi_b(y) G(xy;x'y') \chi_c(y'). \tag{6}$$

$G(xb;x'c)$ from (6) has the interpretation as being the 'transmission amplitude' that a unit impulse of probability amplitude originally deposited at position x' in normal mode c will propagate to position x in normal mode b. Also, the 'free' Green function in the absence of a defect potential obeys an equation of motion

$$\left\{ E - E_a - \left(-\frac{\hbar^2}{2m}\frac{\mathrm{d}^2}{\mathrm{d}x^2} \right) \right\} G^0(xa;x'c) = \delta(x-x')\delta_{ac} \tag{7}$$

so $G^0(xa;x'c)$ is diagonal in the mode indices a and c as $G^0(xa;x'c) = G^0_a(x;x')\delta_{ac}$.

By standard manipulations of (4) and (7), repeatedly applying (3), we obtain the Dyson equation for a Q1D wire as

$$G(xa;x'c) = G^0(xa;x'c) + \sum_{bd} \int \mathrm{d}x''\, G^0(xa;x''b) V_{bd}(x'') G(x''d;x'c). \tag{8}$$

This Dyson equation can be given its usual interpretation of summing over the probability amplitudes of all the possible scattering processes for a particle starting at position x' in normal mode c arriving at position x in normal mode a.

3. Delta function scatterer

We choose the scattering potential $V_{\mathrm{d}}(x,y)$ to be a delta function

$$V_{\mathrm{d}}(x,y) = \gamma\delta(x)\delta(y-y_i) \tag{9}$$

so that its matrix elements $V_{ab}(x)$ from (5) are

$$V_{ab}(x) = \delta(x)\gamma\chi_a(y_i)\chi_b(y_i). \tag{10}$$

The weight γ can be either positive or negative.

Now let us iterate the Dyson equation (8) assuming initially for simplicity that only the lowest two normal modes are present. The infinite series for $G(x1;x'1)$ is

$$\begin{aligned}G(x1;x'1) = {} & G_1^0(x;x') + G_1^0(x;0)V_{11}G_1^0(0;x') \\ & + G_1^0(x;0)V_{11}G_1^0(0;0)V_{11}G_1^0(0;x') \\ & + G_1^0(x;0)V_{12}G_2^0(0;0)V_{21}G_1^0(0;x') \\ & + G_1^0(x;0)V_{11}G_1^0(0;0)V_{11}G_1^0(0;0)V_{11}G_1^0(0;x') \\ & + G_1^0(x;0)V_{11}G_1^0(0;0)V_{12}G_2^0(0;0)V_{21}G_1^0(0;x') \\ & + G_1^0(x;0)V_{12}G_2^0(0;0)V_{21}G_1^0(0;0)V_{11}G_1^0(0;x') \\ & + G_1^0(x;0)V_{12}G_2^0(0;0)V_{22}G_2^0(0;0)V_{21}G_1^0(0;x') + \cdots \end{aligned} \tag{11}$$

where $V_{ab} = \gamma\chi_a(y_i)\chi_b(y_i)$. Equation (11) can be regrouped as a power series

$$\begin{aligned}G(x1;x'1) = {} & G_1^0(x;x') + G_1^0(x;0)V_{11}G_1^0(0;x')[1 + \left(V_{11}G_1^0(0;0) + V_{22}G_2^0(0;0)\right)^1 \\ & + \left(V_{11}G_1^0(0;0) + V_{22}G_2^0(0;0)\right)^2 \\ & + \left(V_{11}G_1^0(0;0) + V_{22}G_2^0(0;0)\right)^3 + \cdots] \end{aligned} \tag{12}$$

and summed as

$$G(x1;x'1) = G_1^0(x;x') + \frac{G_1^0(x;0)V_{11}G_1^0(0;x')}{1 - \left(V_{11}G_1^0(0;0) + V_{22}G_2^0(0;0)\right)}. \tag{13}$$

Equation (13) is valid for any energy of the incident electron. That is, equation (13) is valid when modes one and two are either propagating modes or evanescent modes. We later discuss the convergence of the power series (12) leading to (13), and show that the resulting equation (13) is better than the method used to obtain it.

By analogy with our calculation of G_{11} from (13), the result for an arbitrary intrasubband Green function $G(xa;x'a)$ including all the normal confinement modes is

$$G(xa;x'a) = G_a^0(x;x') + G_a^0(x;0)V_{aa}G_a^0(0;x')\left[1 - \left(\sum_n V_{nn}G_n^0(0;0)\right)\right]^{-1}. \tag{14}$$

Detailed discussion of one-dimensional results similar to (14) are given in [13–14]. The expression for the intersubband Green function $G(xa;x'b)$ again including all the normal confinement modes is

$$G(xa;x'b) = +G_a^0(x;0)V_{ab}G_b^0(0;x')\left[1 - \left(\sum_n V_{nn}G_n^0(0;0)\right)\right]^{-1}. \tag{15}$$

Equations (14) and (15) are the exact solutions of the Dyson equation (8) with the scattering potential (9), and can be obtained by straightforward algebra after inserting (9) into (8). The iteration proceedure in (11), which fails if the Fermi energy is equal to a subband minimum but is useful for visualizing the possible scattering processes, is not necessary to solve (8).

Equations (14) and (15) can be evaluated by noting that the 'free' Green function for normal mode a is [13]

$$G_a^0(x;x') = -\frac{m}{\hbar^2\kappa_a}\exp\left(-\kappa_a|x-x'|\right) \qquad E < E_a \tag{16}$$

if mode a is an evanescent mode where

$$\kappa_a = +\sqrt{\frac{2m(E_a - E)}{\hbar^2}} \tag{17}$$

and

$$G_a^{0\pm}(x;x') = \mp i\frac{m}{\hbar^2 k_a}\exp\left(\pm ik_a(x-x')\right) \qquad E > E_a \tag{18}$$

if mode a is a propagating mode where

$$k_a = +\sqrt{\frac{2m(E - E_a)}{\hbar^2}}. \tag{19}$$

From (16) and (18) we see that $G_a^0(0,0)$, which occurs repeatedly in the power series (12) and corresponds diagrammatically to the particle repeatedly looping around the delta function [14], is simply proportional to the density of propagating or evanescent states for mode a. Thus, the series in the denominator of (14) and (15) of the form $\sum_n V_{nn}G_n^0(0;0)$ should be interpreted as a Golden Rule amplitude involving the square root of the initial density of states, a matrix element connecting initial to final state, and the square root of the density of final states. Taking the square magnitude of this denominator will yield an infinite series of Golden-Rule-type scattering terms between all possible normal modes. The numerators of (14) and (15) can also be interpreted in this way.

4. Transmission coefficients

Let us evaluate the Green functions for the special case where x and x' are on opposite sides of the scatterer and hence $x' < 0 < x$. Since we are interested in transmission through the scatterer from left to right we consider only G^+. Because we work with the time-independent form of the Green functions, implicitly assumed in all our calculations is that a constant applied incident current is imposed on the scatterer from the left. If this were not true the scattering problem could not reach a time-independent solution. We must leave on the applied current long enough that evanescent modes can build up around the scatterer until a steady state is reached as described in the introduction and in [1].

To evaluate the intrasubband Green function from (14), we use the identity

$$G_a^{0+}(x;0)G_a^{0+}(0;x') = G_a^{0+}(0;0)G_a^{0+}(x;x') \tag{20}$$

to rewrite (14) as

$$G^+(xa;x'a) = t_{aa}(E)\, G_a^{0\,+}(x;x'). \tag{21}$$

Here $t_{aa}(E)$ is the current transmission amplitude through the defect

$$t_{aa}(E) = \left(1 + \sum_n^{\mathrm{e}} V_{nn}\frac{m}{\hbar^2\kappa_n} + \mathrm{i}\sum_{n\neq a}^{\mathrm{P}} V_{nn}\frac{m}{\hbar^2 k_n}\right) \times \left(1 + \sum_n^{\mathrm{e}} V_{nn}\frac{m}{\hbar^2\kappa_n} + \mathrm{i}\sum_n^{\mathrm{P}} V_{nn}\frac{m}{\hbar^2 k_n}\right)^{-1} \tag{22}$$

given in [1]. In (22) $\sum^{\mathrm{e}}$ denotes a sum over all the evanescent modes, $\sum^{\mathrm{P}}$ denotes a sum over the propagating modes, and mode a is assumed propagating. The ability to factor the Green function into a product of the free Green function multiplied by the transmission amplitude depends critically on the shape of the scatterer. Only for delta function scatterers is it possible to make the simple factorization in (21). In (21), the 'free' Green function keeps track of the particle's phase while $t_{aa}(E)$ gives the current transmission amplitude including any possible phase shifts. Factorization of the Green function as in (21) for the case of two delta function scatterers in one dimension (modelling a resonant tunneling problem) has been noted by Garcia-Calderón [15].

The new physics of scattering in a confined geometry, discussed in detail for all the transmission coefficients in [1], can be briefly illustrated by considering a simple case of (22) where mode one is propagating and mode two is evanescent:

$$t_{11}(E) = \frac{1 + V_{22}m/\hbar^2\kappa_2}{1 + V_{22}m/\hbar^2\kappa_2 + \mathrm{i}V_{11}m/\hbar^2 k_1}. \tag{23}$$

At the minima of the second subband we have $\kappa_2 = 0$ resulting in perfect transmission of the incident mode $t_{11} = 1$. This 'perfect transparency' effect, first pointed out by Chu and Sorbello [2], is a consequence of evanescent modes building up near the scattering defect [1]. In addition, the numerator of (23) is zero when the incident electron energy lines up with the quasi-bound state which has split off from the second subband [1] resulting in perfect reflection, $t_{11} = 0$. Setting the real part of the Green function's denominator in (21) to zero we recover the quasi-bound-state energy.

Evaluating the intersubband Green function from (15) yields

$$G^+(xa; x'b) = t_{ab}(E)\left(-\mathrm{i}\frac{m}{\hbar^2\sqrt{k_a k_b}}\right)\exp(+\mathrm{i}k_a x - \mathrm{i}k_b x'). \tag{24}$$

Here $t_{ab}(E)$ is the current transmission amplitude through the defect from the incident normal mode b on the left to the transmitted normal mode a on the right

$$t_{ab}(E) = -\mathrm{i}V_{ab}\frac{m}{\hbar^2\sqrt{k_a k_b}}\left(1 + \sum_n^{\mathrm{e}} V_{nn}\frac{m}{\hbar^2\kappa_n} + \mathrm{i}\sum_n^{\mathrm{P}} V_{nn}\frac{m}{\hbar^2 k_n}\right)^{-1} \tag{25}$$

given in [1].

Equation (25) gives the transmission amplitudes $t_{ab} = t_{ba}$ for $a \neq b$. But for the delta function scattering potential of (9), $t_{ab} = r_{ab}$ for $a \neq b$ simply by wavefunction continuity at the scatterer [1]. Furthermore $1 + r_{aa} = t_{aa}$, so (25) gives the reflection

amplitudes r_{aa} if $a = b$. These results can also be shown using the Green function approach of this paper. Therefore

$$r_{ab}(E) = -\mathrm{i}V_{ab}\frac{m}{\hbar^2\sqrt{k_a k_b}}\left(1+\sum_n^{\mathrm{e}} V_{nn}\frac{m}{\hbar^2\kappa_n}+\mathrm{i}\sum_n^{\mathrm{p}} V_{nn}\frac{m}{\hbar^2 k_n}\right)^{-1} \tag{26}$$

which holds for any two propagating normal modes a and b, should be considered the fundamental result of this paper. The two factors of (26) have a simple interpretation in terms of the Fermi Golden Rule as described in the previous section. Conversely, the intrasubband transmission from (22) appears to be a result of 'leftover' particles not deflected by the delta function, and can be obtained by applying (26) together with wavefunction continuity at the scatterer.

The relationship obtained by Fisher and Lee [8] between the current transmission amplitudes and the Green functions for an arbitrary defect potential is

$$t_{ab}(E) = -\mathrm{i}\hbar\sqrt{v_a v_b}G^+(xa;x'b)\exp(-\mathrm{i}k_a x+\mathrm{i}k_b x') \tag{27}$$

where we again require that $x' < 0 < x$. We have inserted an extra $\hbar$ in their relationship which must be there on purely dimensional grounds. The current transmission amplitudes t_{ab} through the delta function potential were explicitly calculated in [1]. In this paper we have calculated all the Green functions $G(xa;x'b)$ through the delta function defect. Both the intrasubband Green functions from (21) and the intersubband Green functions from (24) clearly obey the relation in (27) (up to an unimportant phase factor of -1). Alternately, had we not previously calculated the current transmission amplitudes using another method, we could have used the calculation in section 3 and (27) to obtain them. Therefore, we can also regard the transmission coefficients t_{aa} from (22) and t_{ab} from (25) as being obtained by solving the Dyson equation (8) and applying the Fisher–Lee relation (27).

The current transmission coefficients T_{ab} necessary to calculate the conductance through a defect we may now obtain as

$$T_{ab} = t_{ab}t_{ab}^* = v_a v_b|\hbar G^+(xa;x'b)|^2. \tag{28}$$

The-two probe Landauer conductance can then be written as

$$G = \frac{e^2}{\pi\hbar}\sum_{ab}T_{ab} = \frac{e^2}{\pi\hbar}\sum_{ab}v_a v_b|\hbar G^+(xa;x'b)|^2 \tag{29}$$

which is similar to the expression in [8]. Relation (27) of Fisher and Lee for the transmission coefficients, inserted into the two-probe Landauer conductance formula, gives the expression for the conductance in terms of Green functions as in (29).

The two-probe conductance versus Fermi energy from (29) for both a repulsive delta potential (full curve) having $\gamma = 7$ feV cm^2 and an attractive delta potential (chain curve) with $\gamma = -7$ feV cm^2 are shown in figure 2. Figure 2 assumes an infinite square well confinement of 30 nm width and an electron mass of 0.067 times the free electron mass appropriate for GaAs heterojunctions. Conductance through the repulsive potential (full curve) is lower than its ballistic value of n times $2e^2/h$ (where n is the subband index) due to increased reflection immediately above the bottom of the nth subband. This increased reflection rounds the shoulders of the

quantised conductance steps. Immediately below the minima of the nth subband the conductance rises to its ballistic value of $(n-1)$ times $2e^2/h$ as a result of the 'perfect transparency' effect described in this section. For the two-probe conductance through an attractive defect (chain curve), a single quasi-bound state splits off from each quasi-one-dimensional subband and is visible as the extra pronounced dips in the conductance. The quasi-bound state associated with the nth subband appears when the Fermi energy lies in the $(n-1)$th subband, resulting in increased reflection and a correspondingly lower conductance near the quasi-bound-state energy. The 'perfect transparency' effect is also present when the scatterer is attractive. We have investigated the conductance through a delta function scatterer in a wire in detail in [1].

5. Conclusions

The scattering properties of electrons in a confined geometry are qualitatively different from the usual case of scattering in open geometries due to the building up and storage of electrons in evanescent waves near the scattering defect. To illustrate these properties, we solved Dyson's equation for the single-electron Green function describing electrons scattering from a delta function defect in a quasi-one-dimensional wire. We then used Fisher and Lee's relationship between the single-electron Green functions and the current transmission amplitudes to obtain the current transmission coefficients of electrons through a delta function defect in the Q1D wire. The transmission coefficients so obtained agree with those found by simply matching wavefunctions and their derivatives at the defect [1]. For the delta function scatterer, all normal modes completely 'decouple' at a new subband minimum resulting in perfect transmission despite the presence of a scatterer. If the delta function scatterer is attractive, a single quasi-bound state splits off from each confinement subband and causes increased reflection if the Fermi energy is near the quasi-bound state.

Acknowledgments

We thank Terry P Orlando, Arvind Kumar, Marc A Kastner, Rolf Landauer, Charles Kane, Dimitri A Antoniadis, and Henry I Smith for useful discussions. This work was sponsored by the US Air Force Office of Scientific Research under grant AFOSR-88-0304.

References

[1] Bagwell P F 1990 Evanescent modes and scattering in quasi-one-dimensional wires *Phys. Rev.* B **41** at press

[2] Chu C S and Sorbello R S 1989 Effect of impurities on the quantized conductance of narrow channels *Phys. Rev.* B **40** 5941

[3] Masek J, Lipavsky P and Kramer B 1989 Coherent potential approach for the zero temperature DC conductance of weakly disordered narrow systems *J. Phys.: Condens. Matter* **1** 6395

[4] Tekman E and Ciraci S Ballistic transport through quantum point contacts: elastic scattering by impurities unpublished

[5] Landauer R 1989 Conductance determined by transmission: probes and quantized constriction resistance *J. Phys.: Condens. Matter* **1** 8099

[6] Landauer R 1990 'Electrons as guided waves in laboratory structures: strengths and problems' *Analogies in Optics and Micro-Electronics* ed W van Haeringen and D Lenstra (Deventer: Kluwer Academic) at press

[7] Payne M C 1989 Electrostatic and electrochemical potentials in quantum transport *J. Phys.: Condens. Matter* **1** 4931

[8] Fisher D S and Lee P A 1981 Relation between conductivity and transmission matrix *Phys. Rev.* B **23** 6851

[9] Stone A D and Szafer A 1988 What is measured when you measure a resistance?—The Landauer Formula revisited *IBM J. Res. Dev.* **32** 384

[10] Baranger H U and Stone A D 1989 Electrical linear-response theory in arbitrary magnetic field: a new Fermi surface formulation *Phys. Rev.* B **40** 8169

[11] Wang Lihong and Feng Shenchao 1989 Correlations and fluctuations in reflection coefficients for coherent wave propagation in disordered scattering media *Phys. Rev.* B **40** 8284

[12] Lee P A and Fisher D S 1981 Anderson localization in two dimensions *Phys. Rev. Lett.* **47** 882

[13] Economou E N 1983 *Green Functions in Quantum Physics* 2nd edn (New York: Springer) pp 14, 66, 98

[14] Inkson J C 1984 *Many-Body Theory of Solids* (New York: Plenum) pp 36–38

[15] Garcia-Calderón G 1987 The effect of asymmetry on resonant tunneling in one dimension *Solid State Commun.* **62** 441

Garcia-Calderón G and Rubio A 1987 Characteristic times for resonant tunneling in one dimension *Solid State Commun.* **62** 441

Appendix H

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

Solid State Communications, Vol. 75, No. 12, pp. 949-953, 1990.
Printed in Great Britain.

0038-1098/90$3.00+.00
Pergamon Press plc

RESONANT TUNNELING IN A MULTI-CHANNEL WIRE

Arvind Kumar and Philip F. Bagwell
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

(received May 7, 1990 by G. Burns; in revised form June 22, 1990)

We calculate analytically the transmission coefficient through two delta function barriers in a quasi-one-dimensional wire as the Fermi energy and distance between scatterers are varied. In a purely one-dimensional or single-mode system, the transmission coefficient is periodic as a function of distance between scatterers with successive resonances separated by half the wavelength of the incident electron. In a two-mode system we find unusual modifications to this result when the Fermi energy is near the second subband minimum. First, the presence of the evanescent mode causes successive resonances in the transmission coefficient to be separated by a full wavelength rather than by a half wavelength. Furthermore, when the Fermi energy coincides with the second subband minimum, the transmission approaches unity for scatterer separations larger than a few wavelengths and is therefore no longer a periodic function of distance between scatterers.

1. Introduction

Electron scattering in a confined geometry exhibits unusual properties because of the existence of evanescent modes induced by the confinement. In a confined geometry, such as the quasi-one-dimensional wire shown in Fig. 1, scattering into the evanescent modes causes a local buildup of particle density around the scatterers. The novel scattering properties of a single obstacle in a multi-channel wire have been studied in Refs. [1]-[4]. In particular, it was found that the transmission through a single delta function scatterer becomes perfect whenever the Fermi energy coincides with the bottom of a subband regardless of the sign of the scatterer. Additional novel transmission properties were found for an attractive scatterer when the Fermi energy is near a quasi-bound state of the wire.

In this communication we calculate the transmission through two delta function barriers in a quasi-one-dimensional wire. In the presence of only a single mode, one obtains the usual one-dimensional resonant tunneling problem, which is discussed in detail in Refs. [5]-[7]. When the Fermi energy is near the bottom of the second subband, we find that the two-mode transmission coefficient is not periodic as a function of distance between scatterers and has resonances separated by a full wavelength of the incident electron. These properties suggest that the two-mode transmission resonances are caused by a mechanism different from the geometrical interference effect found in the one-dimensional problem. When the Fermi energy is near the bottom of a subband, we find that coupling to the evanescent mode dominates the scattering in the wire.

Our purpose in this communication is to present only the basic and most dominant features of this multi-mode scattering problem. Therefore, we choose to focus our attention here on the case where the two tunnel barriers have equal strengths and are aligned, though we comment on the more complicated results obtained for other scattering potentials. We also concentrate for simplicity only on the transmission of the lowest mode for

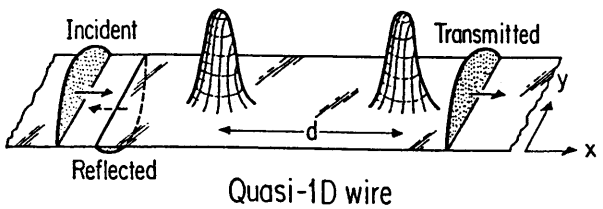


Fig. 1. An electron, incident from the left in the lowest normal mode, scattering from two barriers in a quasi-one-dimensional wire.

energies near the second subband minimum. If one or both of the scatterers are made attractive, there are also complicated effects arising from quasi-bound states in the wire which we do not investigate here. The method and some results we present encompass these more complicated cases, as we note at the appropriate points.

2. Scattering From Two Barriers in a Wire

The system we consider is a quasi-one-dimensional wire with a confining potential $V_c(y)$ and a scattering potential $V_S(x,y)$. The time-independent Schrödinger equation (in the effective mass approximation) is

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

For our two-barrier problem, we take the scattering potential to consist of two delta functions:

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

This potential is shown schematically in Fig. 1 where the barriers are separated along the x direction by a distance d.

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

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dy^2}+V_c(y)\right]\chi_n(y) = E_n\chi_n(y), \quad (3)$$

where $n = 1,2,3,....$ In the quasi-one-dimensional wire each transverse eigenstate n is associated with a mode which is propagating if the particle energy $E > E_n$, and a mode which is evanescent if $E < E_n$. The wavefunction of the clean wire is separable so that the x dependent part of the wavefunction is a linear combination of two plane waves for the propagating modes, and a linear combination of growing and decaying exponentials for the evanescent modes.

Away from the scattering potential we can therefore expand the wavefunction as follows [8]:

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

where

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

Here the wavevector 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. The boundary conditions require $A_n = D_n = 0$ for evanescent modes so that there are no growing exponentials away from the scatterers.

The wavefunction 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 [8] so that $\alpha_n(0^+) = \alpha_n(0^-)$. Therefore, at $x = 0$,

$$A_n + B_n = F_n + G_n \quad (7)$$

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

$$\left(\frac{d^2}{dx^2}+k_n^2\right)\alpha_n(x) = \sum_m \left(\Gamma_{nm}^{(1)}\delta(x) + \Gamma_{nm}^{(2)}\delta(x-d)\right)\alpha_m(x). \quad (8)$$

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

$$\Gamma_{nm}^{(i)} = \frac{2m\gamma^{(i)}}{\hbar^2}\chi_n(y_S^{(i)})\chi_m(y_S^{(i)}), \qquad i = 1,2. \quad (9)$$

At $x = 0$ the discontinuity in the first derivative of each mode, by integrating Eq. (8), must satisfy

$$\left.\frac{d\alpha_n(x)}{dx}\right|_{0+} - \left.\frac{d\alpha_n(x)}{dx}\right|_{0-} = \sum_m \Gamma_{nm}^{(1)}\alpha_m(0)\,. \quad (10)$$

Inserting Eq. (5) in Eq. (10) yields

$$ik_n(F_n - G_n) - ik_n(A_n - B_n) = \sum_m \Gamma_{nm}^{(1)}(A_m + B_m). \quad (11)$$

Equation (11) couples mode n to all other modes in the problem. For simplicity, we cut off the sum in Eq. (11) at a finite number of modes. Likewise, equations similar to Eq. (7) and Eq. (11) enforcing the boundary conditions at $x = d$ can also be written down. We note that the form of Eq. (11) does not depend on the y dependence of the scattering potential in Eq. (2). A different choice for the y dependence would only change the values of the $\Gamma_{nm}^{(i)}$.

The previous equations also do not depend on the choice of the confinement potential, so that none of the results of this paper depend fundamentally on $V_c(y)$. For definiteness we take $V_c(y)$ to be zero for $0 \le y \le W$ and infinite elsewhere, yielding

$$\chi_n(y) = \sqrt{\frac{2}{W}}\sin\left(\frac{n\pi y}{W}\right), \qquad E_n = \frac{\pi^2\hbar^2}{2mW^2}n^2. \quad (12)$$

3. Transmission Coefficients

We take the particles to be incident from the left, so that $D_n = 0$ in Eq. (5), and define wavefunction transmission and reflection amplitudes normalized to the amplitude of the incident mode:

$$t_{nm} = \frac{C_n}{A_m}, \qquad r_{nm} = \frac{B_n}{A_m}. \quad (13)$$

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

$$T_{nm} = \frac{k_n}{k_m} t^*_{nm} t_{nm}, \qquad R_{nm} = \frac{k_n}{k_m} r^*_{nm} r_{nm}. \tag{14}$$

For one mode, by solving Eqs. (7) and (11) and the analogous equations at $x = d$, we straightforwardly obtain

$$t_{11} = e^{ik_1 d}/\mathcal{D}_1 \tag{15}$$

where the denominator $\mathcal{D}_1$ is

$$\mathcal{D}_1 = \left(1 + i\frac{\Gamma^{(1)}_{11}}{2k_1}\right)\left(1 + i\frac{\Gamma^{(2)}_{11}}{2k_1}\right) + \frac{\Gamma^{(1)}_{11}\Gamma^{(2)}_{11}}{4k_1^2} e^{2ik_1 d} \quad . \tag{16}$$

The denominator of the transmission amplitude $\mathcal{D}_1$ is a periodic function of distance d with period $\lambda_1/2$, where $\lambda_1 = 2\pi/k_1$. The numerator of Eq. (15) is a simple phase factor having unit modulus. Therefore, the transmission coefficient T_{11} is periodic as a function of distance with period $\lambda_1/2$.

We now consider the case of two modes. For simplicity we take the scatterers to be identical and aligned with $\Gamma^{(1)}_{nm} = \Gamma^{(2)}_{nm} = \Gamma_{nm}$, although the two scatterers can be either attractive or repulsive. Solving Eqs. (7) and (11), along with their analogues at $x = d$, and applying Eq. (13) yields

$$t_{11} = \frac{1}{\mathcal{D}_2}\left[\frac{i}{2k_2}\sum_{n=1}^{2}\frac{\Gamma^2_{n2}}{k_n}\sin k_n d + \left(1 + i\frac{\Gamma_{22}}{k_2}\right)e^{-ik_2 d}\right] \tag{17}$$

where

$$\begin{aligned}\mathcal{D}_2 = {} & \frac{i}{2}\sum_{n=1}^{2}\left[\frac{\Gamma^2_{n1}}{k_1 k_n}e^{-ik_2 d} + \frac{\Gamma^2_{n2}}{k_2 k_n}e^{-ik_1 d}\right]\sin k_n d \\ & + \left[1 + i\frac{\Gamma_{11}}{k_1} + i\frac{\Gamma_{22}}{k_2} - \frac{\Gamma^2_{12}}{2k_1 k_2}\right]e^{-i(k_1+k_2)d} \\ & + \frac{\Gamma^2_{12}}{4k_1 k_2}\left(e^{-2ik_1 d} + e^{-2ik_2 d}\right) \quad . \end{aligned} \tag{18}$$

To derive this expression, we have used the identity $\Gamma^2_{nm} = \Gamma_{nn}\Gamma_{mm}$. Of course, Eqs. (17)-(18) reduce to Eqs. (15)-(16) if $\Gamma_{n2} = \Gamma_{2n} = 0$ so that there is no coupling to the second mode. If we set $d = 0$ in Eqs. (17)-(18) we recover the results of Refs. [1]-[2] for a single scatterer.

In the limits $k_2 \ll \Gamma_{22}$ and $k_2 \ll d^{-1}$, we obtain

$$t_{11} = \frac{\left[\frac{i}{2}\sum_{n=1}^{2}\frac{\Gamma^2_{n2}}{k_n}\sin k_n d + i\Gamma_{22}\right]e^{ik_1 d}}{\left[\frac{i}{2}\sum_{n=1}^{2}\frac{\Gamma^2_{n2}}{k_n}\sin k_n d + i\Gamma_{22}\right] + \frac{\Gamma^2_{12}}{2k_1}(\cos k_1 d - 1)}. \tag{19}$$

Since the bracketed terms in the numerator and denominator of t_{11} are equal, we obtain unity transmission when $(\cos k_1 d - 1)$ vanishes, *i.e.*, when $d = j\lambda_1, j = 0, 1, 2, \ldots$. We additionally note that if $\lambda_1 \ll d \ll \lambda_2$, the $n = 2$ term in the above summations will dominate both the numerator and the denominator, driving the transmission to unity.

Plotting the transmission coefficient as a function of energy gives no special insight in this problem, since the strongest modifications to the one-dimensional scattering properties occur mainly for a range of electron energies near $E = E_2$. Therefore, we choose to plot T_{11} as a function of scatterer separation d. We now take the scatterers to be repulsive, with $\gamma^{(1)} = \gamma^{(2)} = 10^{-14}$ eV cm^2 and $y_s^{(1)} = y_s^{(2)} = 13.3$ nm. The variation of the transmission coefficient T_{11} with scatterer separation d is illustrated in Fig. 2 for energies $E = 0.9E_2$ (bottom), $E = E_2$ (middle), and $E = 1.1E_2$ (top). We have chosen parameters $m = 0.067m_0$ and $W = 30$ nm appropriate for GaAs heterojunctions.

For $E = 0.9E_2$, when the Fermi energy is below the second subband minimum, the evanescent modes have little qualitative effect on the transmission properties of the system so that we recover the standard one-dimensional result: perfect transmission resonances separated by $\lambda_1/2$ with the first resonance offset from $d = 0$, consistent with Eqs. (15)-(16). However, the transmission is enhanced when the evanescent modes are included. When $E = E_2$, we note that successive resonances in T_{11} become separated by a full wavelength. The first maximum occurs at $d = 0$, so that we recover the result of perfect transmission at $E = E_2$ for a single scatterer that was proved in Refs. [1]-[3]. Furthermore, the transmission asymptotically approaches unity when the scatterer separation becomes larger than a few wavelengths of the incident electron. Finally, at an energy $E = 1.1E_2$ such that the second mode is propagating, we observe a beat-like modulation in T_{11}. This is not surprising, since there are now two allowed propagating wavelengths in the transmission problem. This beating pattern would not occur if the coupling between the two modes were zero, and is therefore a qualitatively new feature of scattering in the multi-mode wire.

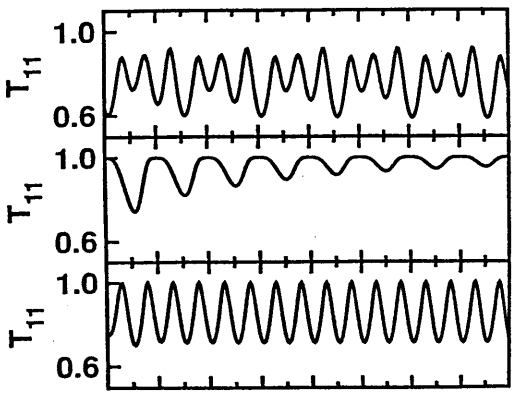


Fig. 2. Transmission coefficient T_{11} versus scatterer separation d for two aligned delta function barriers in a quasi-one-dimensional wire. The scatterer separation d is normalized to the incident electron wavelength λ_1, which is different for each energy E. The graphs shown are for constant energies $E = 0.9E_2$ (bottom), $E = E_2$ (middle), and $E = 1.1E_2$ (top).

We have also considered disaligned scatterers, unequal scatterers, and attractive scatterers [9]. Importantly, we still find successive resonances in T_{11} separated by a full wavelength and T_{11} approaching unity in the limit of large scatterer separations as before. These properties fail only when one of the barriers is exactly on the node in $\chi_2(y)$, eliminating the coupling to the second mode at that scatterer. Even if the scatterers have unequal strengths, the transmission maxima are still perfect provided the scatterers are aligned, in sharp contrast to the one-dimensional case of Eqs. (15)-(16). Disaligning the scatterers does slightly degrade the transmission maxima, but this effect is appreciable only when either of the scatterers is near the node in $\chi_2(y)$. The first transmission maximum is also offset from $d = 0$ if the scatterers are disaligned. If the scatterers are attractive we additionally find $T_{11}=0$ near the quasi-bound state energies splitting off from the confinement subbands.

4. Discussion

Near the subband bottom, it is evident from Eqs. (17)-(19) that the dominant terms, which have k_2^{-1} dependences, all involve coupling to mode 2. These terms are nonexistent in the single-mode case, Eqs. (15)-(16). In particular, the terms in Eqs. (17)-(18) corresponding to the geometrical interference effect in the one-dimensional problem are completely overshadowed by terms involving coupling to mode 2. Hence, mode conversion dominates the scattering properties near the subband minimum. This is reasonable since the one-dimensional density of states for mode 2 diverges as k_2^{-1} on either side of the subband minimum.

The relative phase of the incident mode at each of the barriers plays an important role in determining the spatial variation of the evanescent modes, which can only be populated by scattering from the incident applied current. Since the relative phase of the incident mode is clearly periodic with a period λ_1, we have some basis for understanding the change in the spacing between successive resonances in the transmission coefficient from $\lambda_1/2$ to λ_1 when coupling to the second mode becomes dominant. To understand this in more detail we examine the spatial variation of the evanescent mode $\alpha_2(x)$.

For the simple case of identical, aligned scatterers, so that the mode coupling constants of the two barriers are equal, we have found [9] that the evanescent mode amplitude $\alpha_2(x)$ can be chosen real and has the form shown in Fig. 3. If the scatterers are separated by an integer multiple of λ_1, the evanescent mode amplitude $\alpha_2(x)$ is symmetric about $x = d/2$, the midpoint between the scatterers. By analogy to molecular physics, we have labelled this a 'bonding' state. Transmission maxima occur when the evanescent mode forms a 'bonding' state. In contrast, if the scatterer separation is a half integer multiple of λ_1, the evanescent mode amplitude $\alpha_2(x)$ has opposite signs at each scatterer, resulting in a node

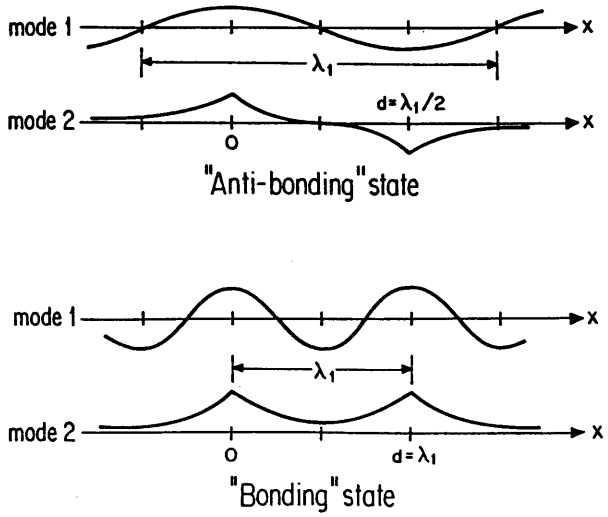


Fig. 3. The evanescent mode is populated by scattering from the incident applied current, so that electrons accumulate locally in evanescent modes around each tunnel barrier. For the case of identical, aligned scatterers, the evanescent mode forms an 'anti-bonding' state when the barriers are separated by half the wavelength of the incident electron, and a 'bonding' state when the separation is a full wavelength.

in the evanescent mode halfway between the scatterers. We have labelled this an 'anti-bonding' state. We emphasize that the form of the evanescent mode is more complicated if the coupling at each scatterer is different.

5. Conclusion

We have compared the transmission of an electron through two delta function barriers in a one-mode wire to that in a two-mode wire. The substantial differences we find between the one-mode and the two-mode case suggest that resonances in the transmission coefficient arise from different mechanisms when the Fermi energy is near the bottom of a subband in the multi-mode wire. Geometrical resonances analogous to those in a Fabry-Perot optical cavity dominate the transmission properties in the single-mode case, whereas the intermode coupling and the accumulation of evanescent modes around the scatterers dominate the transmission properties of the two-mode wire.

Acknowledgement-The authors have enjoyed stimulating discussions with Terry P. Orlando, Ray Ghanbari, Dimitri A. Antoniadis, and Henry I. Smith. This work was supported by the Air Force Office of Scientific Research under grant AFOSR-88-0304.

References

[1] P.F. Bagwell, *Phys. Rev. B* **41**, 10354 (1990).

[2] P.F. Bagwell, to appear in *J. Phys.: Cond. Matter.*

[3] C.S. Chu and R.S. Sorbello, *Phys. Rev. B* **40**, 5941 (1989).

[4] E. Tekman and S. Ciraci (unpublished).

[5] G.A. Toombs and F.W. Sheard, to appear in *Electronic Properties of Multilayers and Low Dimensional Semiconductor Structures,* J.M. Chamberlin, L. Eaves, and J.C. Portal, eds. (Plenum, New York, in press).

[6] P.J. Price, *Phys. Rev. B* **38**, 1994 (1988).

[7] P.F. Bagwell, T.P.E. Broekaert, T.P. Orlando, and C.G. Fonstad, to appear in *J. Appl. Phys.*

[8] M. Cahay, M. McLennan, and S. Datta, *Phys. Rev. B* **37**, 10125 (1988).

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

Appendix I

Arvind Kumar and Philip F. Bagwell, 'Resonant Tunneling in a Quasi-One-Dimensional Wire: Influence of Evanescent Modes', submitted for publication.

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

November 27, 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 new pattern of conductance oscillations emerges. If either of the scatterers is attractive, the transmission is suppressed abruptly near the 'quasi-donor' levels formed in the scatterer below each new subband minimum.

PACS 73.20.Dx, 71.55.Jv, 72.20.Dp, 73.40.Gk
submitted to Physical Review B-15 (received July 30, 1990)

1 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 make the first 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 microns) compared to 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], so we choose to examine a highly *non-adiabatic* 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 multi-mode 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 the decay length of the evanescent mode, so that the tail of the evanescent wavefunction accumulating around one of the scatterers will overlap with the other scatterer. If the Fermi energy is away from a subband minimum or quasi-bound 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]-[6].

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 novel transmission properties, as investigated in Sec. IV. The appendix describes how the transmission coefficients are calculated. [17]

2 Scattering in a multi-mode 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

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

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

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

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 $E \leq E_n$, 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

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

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

$$\psi(x,y)=\sum_n \alpha_n(x)\chi_n(y), \tag{4}$$

where

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

Here the wavevector k_n is

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

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 wavefunction transmission and reflection amplitudes normalized to the amplitude of the incident mode A_m:

$$t_{nm} = \frac{C_n}{A_m}, \qquad r_{nm} = \frac{B_n}{A_m}. \tag{7}$$

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. (18)-(19) in the appendix.

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

$$T_{nm} = \frac{k_n}{k_m} t^*_{nm} t_{nm}, \qquad R_{nm} = \frac{k_n}{k_m} r^*_{nm} r_{nm}. \tag{8}$$

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

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

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 $W = 30\,\text{nm}$ and an electron mass $m = 0.067 m_e$ appropriate for GaAs heterojunctions.

3 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. 3-4. 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 Fig. 4(a)-(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.

The variation of the transmission coefficient in Fig. 4(b)-(c) is markedly different from the standard one-dimensional resonant tunneling result. In case (b), when the Fermi energy aligns 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 Fig. 4(a) and (b) demonstrate that *the Fabry-Perot wave-interference between propagating modes is not the mechanism giving rise to conductance oscillations when the Fermi energy aligns with a subband minimum.* In case (c) T_{11} has a beat-like 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 wavefunction $\alpha_2(x)$ when the scatterers are separated by a half wavelength $d = \lambda_1/2$, while Fig. 5(b) shows the lowest evanescent wavefunction 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 results for the incident propagating wavefunction. 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 wavefunction, 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. The evanescent wavefunction $\alpha_2(x)$ for the unequal barrier strengths studied in Figs. 3-4 is slightly more complex.

We can now give an intuitive argument to explain the electron transmission in Fig. 4(b). Because the evanescent density of states diverges like $\kappa_2{}^{-1}$ 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 wavefunction 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 wavefunction 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]

$$\alpha_2(x) = C_2 = t_{21} e^{ik_1 x_s} A_1, \tag{10}$$

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 Eq. (10) and Eq. (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

$$e^{ik_1 x_s} \sin(\pi y_s/W)/\sin(2\pi y_s/W) = \text{const} \ . \tag{12}$$

Note that Eqs. (11)-(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 wavefunction at either scatterer only when $d = j\lambda_1$, where $j = 0, 1, 2, ...$ 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 wavefunction 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 $E = E_2$. [31] 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

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. 3-4.

In Fig. 6(c) the transmission resonances are not exactly perfect, but the deviation from unity is not noticable 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 Fig. 6(a-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 wavefunction, 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 noticably 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 wavefunction $e^{ik_1 x}\sin(\pi y/W)$ 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 delta 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 delta function scatterers yielding

perfect transmission $T_{11} = 1$ when $E = E_2$.

4 Attractive scatterers

If attractive scatterers are present in a multi-mode wire, quasi-bound or bound states made up of evanescent waves form below each subband minimum, as shown in Fig. 2. The quasi-bound 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 zeroes in the transmission coefficient in the vicinity of the quasi-bound 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 quasi-bound state.

In Fig. 8 we plot the zeroes 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. 8-9. 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 $y_s^{(1)} = y_s^{(2)} = 13.3$ nm and have equal strength $|\gamma^{(1)}| = |\gamma^{(2)}| = 25$ feV cm^2. The energy near the top of the figure where the zeroes 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 'anti-bonding' state is forced into the continuum above the second subband.

The zeroes of T_{11} in Fig. 8(a)-(b) qualitatively follow the bound state energies of the single-mode problem shown in the inset. These zeroes 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 quasi-bound states, so that the zeroes 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 quasi-bound states of a multi-channel system [32] and zeroes 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 zeroes.

We plot the conductance 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 zeroes in Fig. 9(c), and no transmission zeroes in Fig. 9(a), as expected from Fig. 8(a)-(b). [33]

5 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 multi-mode transmission through two point barriers is perfect only for certain resonant values of the scatterer separation. Even though each single barrier in the multi-mode 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 [34] 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.

6 Acknowledgments

We thank Terry P. Orlando, both for contributing his ideas and for his improvements to 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 AFOSR-88-0304. A.K. gratefully acknowledges support from a Semiconductor Research Corporation fellowship during part of this work.

7 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 wavefunction 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 $\alpha_n(d^-) =$

$\alpha_n(d^+)$. [17] Therefore, at $x = 0$,

$$A_n + B_n = F_n + G_n \tag{13}$$

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

$$\left(\frac{d^2}{dx^2} + k_n^2\right) \alpha_n(x) = \sum_m \left(\Gamma_{nm}^{(1)}\delta(x) + \Gamma_{nm}^{(2)}\delta(x-d)\right) \alpha_m(x). \tag{14}$$

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

$$\Gamma_{nm}^{(i)} = \frac{2m\gamma^{(i)}}{\hbar^2}\chi_n(y_s^{(i)})\chi_m(y_s^{(i)}), \qquad i = 1, 2. \tag{15}$$

At $x = 0$ the discontinuity in the first derivative of each mode, by integrating Eq. (14), must satisfy

$$\left.\frac{d\alpha_n(x)}{dx}\right|_{0+} - \left.\frac{d\alpha_n(x)}{dx}\right|_{0-} = \sum_m \Gamma_{nm}^{(1)}\alpha_m(0)$$
$$= \chi_n(y_s^{(1)}) \left(\frac{2m\gamma^{(1)}}{\hbar^2}\sum_m \chi_m(y_s^{(1)})\alpha_m(0)\right) . \tag{16}$$

Since Γ_{nm} from Eq. (15) factors into a product form, the right side of Eq. (16) is simply $\chi_n(y_s^{(1)})$ times a constant independent of n. If the right hand side of Eq. (16) is not zero so that mode n has a derivative jump, Eq. (16) implies that every other mode with a nonzero wavefunction 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 delta function scatterer, wavefunction continuity from Eq. (13) requires that the lowest evanescent mode wavefunction $\alpha(x)$ must be constant everywhere in space. The term in large brackets in Eq. (16) 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. (16) yields

$$ik_n(F_n - G_n) - ik_n(A_n - B_n) = \sum_m \Gamma^{(1)}_{nm}(A_m + B_m). \tag{17}$$

Equation (17) couples mode n to all other modes in the problem. For simplicity, we cut off the sum in Eq. (17) at a finite number of modes. Likewise, equations analogous to Eqs. (13) and (17) 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 $\mathbf{t}'$ and $\mathbf{r}'$ directly. [17] We then interchange $\Gamma^{(1)}$ and $\Gamma^{(2)}$ to obtain t_{11} as:

$$t_{11} = \frac{1}{\mathcal{D}_2}\left\{\frac{i}{2k_2}\sum_{n=1}^{2}\frac{\Gamma^{(1)}_{n2}\Gamma^{(2)}_{2n}}{k_n}\sin k_n d + \left[1 + \frac{i(\Gamma^{(1)}_{22} + \Gamma^{(2)}_{22})}{2k_2}\right]e^{-ik_2 d}\right\} \quad , \tag{18}$$

$$\begin{aligned}\mathcal{D}_2 = {} & \frac{ie^{-ik_2 d}}{2k_1}\sum_{n=1}^{2}\frac{\Gamma^{(1)}_{n1}\Gamma^{(2)}_{1n}}{k_n}\sin k_n d + \frac{ie^{-ik_1 d}}{2k_2}\sum_{n=1}^{2}\frac{\Gamma^{(1)}_{n2}\Gamma^{(2)}_{2n}}{k_n}\sin k_n d + \\ & \left[1 + \frac{i(\Gamma^{(1)}_{11} + \Gamma^{(2)}_{11})}{2k_1} + \frac{i(\Gamma^{(1)}_{22} + \Gamma^{(2)}_{22})}{2k_2} - \frac{\Gamma^{(1)}_{11}\Gamma^{(2)}_{22} + \Gamma^{(1)}_{22}\Gamma^{(2)}_{11}}{4k_1 k_2}\right]e^{-i(k_1+k_2)d} \\ & + \frac{\Gamma^{(1)}_{12}\Gamma^{(2)}_{21}}{4k_1 k_2}e^{-i2k_1 d} + \frac{\Gamma^{(1)}_{21}\Gamma^{(2)}_{12}}{4k_1 k_2}e^{-i2k_2 d}\end{aligned} \tag{19}$$

Most of the results of this paper can be obtained by taking appropriate limits of Eqs. (18)-(19). To obtain the curves in Fig. 8 we have set $t_{11} = 0$ from Eq. (18). The resulting expression is

$$\frac{1}{2}\left(\frac{\Gamma^{(1)}_{12}\Gamma^{(2)}_{21}}{k_1\kappa_2}\sin k_1 d + \frac{\Gamma^{(1)}_{22}\Gamma^{(2)}_{22}}{\kappa_2^2}\sinh \kappa_2 d\right) + \left(1 + \frac{\Gamma^{(1)}_{22} + \Gamma^{(2)}_{22}}{2\kappa_2}\right)e^{\kappa_2 d} = 0, \tag{20}$$

where we have written $\kappa_2 = -ik_2$ since the second mode is evanescent in the energy range over which quasi-bound 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. (20).

References

[1] B.J. van Wees, H. van Houten, C.W.J. Beenakker, J.G. Williamson, L.P. Kouwenhoven, D. van der Marel, 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, G.A.C. Jones, J. Phys. C: Solid State Phys., **21**, L209, (1988).

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

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

[5] P.F. Bagwell, Phys. Rev. B, **41**, 10354 (1990). P.F. Bagwell, J. Phys.: Conden. Matt., **2**, 6179 (1990).

[6] C.S. Chu and R.S. Sorbello, Phys. Rev. B, **40**, 5941 (1989). C.S. Chu and R.S. Sorbello, Phys. Rev. B, **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,* S. Beaumont and C. Sotomayor-Torres eds., (Plenum, New York, in press).

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

[10] R.L. Schult, H.W. Wyld, and D.G. Ravenhall, Phys. Rev. B, **41**, 12760 (1990).

[11] H.U. Baranger, 'Multiprobe Electron Waveguides: Filtering and Bend Resistance', (preprint).

[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 Comm., **75**, 949 (1990).

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

[17] M. Cahay, M. McLennan, and S. Datta, Phys. Rev. B **37**, 10125 (1988).

[18] M. Cahay, S. Bandyopadhyay, M.A. Osman, and H.L. Grubin, 'Influence of Evanescent States on Quantum Transport Through an Array of Elastic Scatterers,' to appear in Surface Science. See also S. Bandyopadhyay and M. Cahay, 'The Generalized Scattering Matrix Approach', to appear in the Proceedings of the Workshop on Computational Electronics, Urbana-Champaign, May 21-22, 1990.

[19] J. Masek, P. Lipavsky, and B. Kramer, J. Phys.: Conden. Matt., **1**, 6395 (1989).

[20] I. Kander, Y. Imry, and U. Sivan, Phys. Rev. B, **41**, 12941 (1990).

[21] P.F. Bagwell and A. Kumar, 'Comment on "Effects of Channel Opening and Disorder on the Conductance of Narrow Wires"', (preprint).

[22] L.I. Glazman, G.B. Lesovik, D.E. Khmel'nitskii, and R.I. Shekhter, Pisma 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, 'The Background to Resonant Tunneling Theory,' in *Electronic Properties of Multilayers and Low-Dimensional Structures*, J.M. Chamberlain, L. Eaves, and J.C. Portal, eds., (Plenum, New York, in press).

[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.: Conden. Matt., **1**, 8099 (1989). R. Landauer, in *Analogies in Optics and Micro-Electronics*, W. van Haerington and D. Lenstra eds., (Kluwer, Academic Press, in press).

[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 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).

[32] A.I.Baz', Y.B. Zel'dovich, and A.M. Perelomov, *Scattering, Reactions, and Decay in Nonrelativistic Quantum Mechanics*, (Israel Program for Scientific Translations Ltd. Press, Jerusalem, 1969).

[33] 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 (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.

[34] U. Meirav, M.A. Kastner, and S.J. Wind, Phys. Rev. Lett., **65**, 771 (1990).

Figure 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.

Figure 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 ($E < E_n$), where $k_n = i\kappa_n$, having a dispersion relation given by the dashed lines. In addition, attractive scatterers give rise to bound or quasi-bound states associated with each subband formed from the evanescent wavefunctions.

Figure 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.

Figure 4: Transmission coefficient T_{11} versus 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.

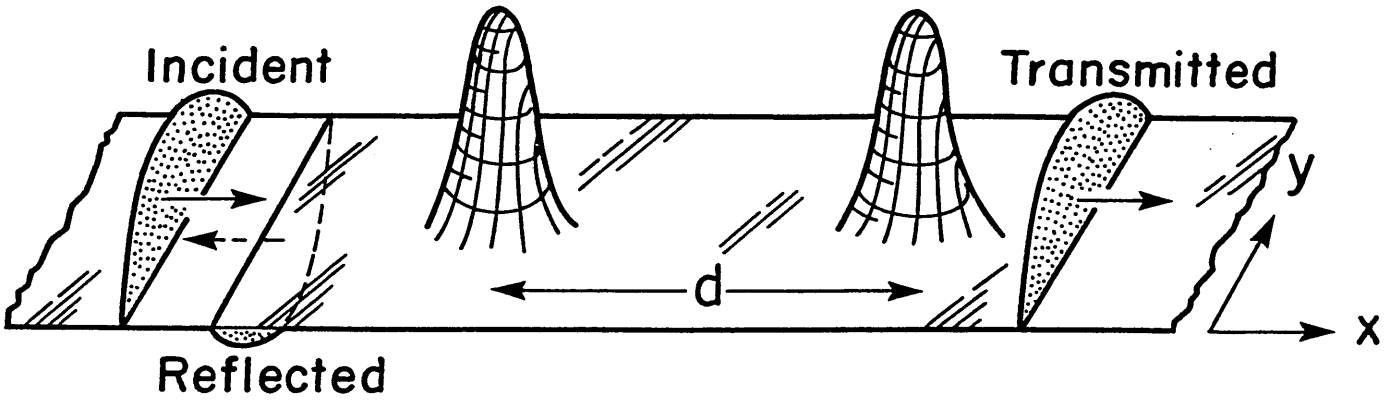
Figure 5: Wavefunction amplitudes of the second mode ($\alpha_2(x)$) when $E = 0.9E_2$ (dashed) and $E = E_2$ (solid) for scatterer separations (a) $d = \lambda_1/2$ and (b) $d = \lambda_1$. When the scatterers are separated by a half wavelength in (a), an 'anti-bonding' state is formed. As the Fermi energy aligns with the second subband at $E = E_2$ the 'anti-bonding' wavefunction 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' wavefunction is uniform along the length of the wire, forcing perfect transmission of the incident wavefunction at $E = E_2$.

Figure 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 wavefunction $\chi_2(y)$, 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.

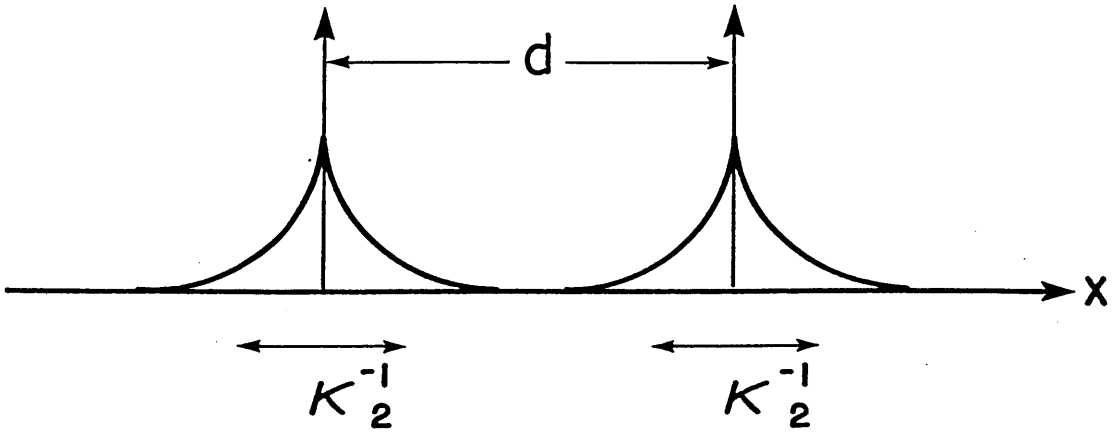
Figure 7: Point scatterers of arbitrary signs and strengths can be placed at any combination of these positions in the wire, and perfect transmission $T_{11} = 1$ still results when the Fermi energy is aligned with the second subband minimum at $E = E_2$. Any of the scatterers can be taken to have zero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

Figure 8: Zeroes in the transmission coefficient $T_{11} = 0$ in d-E plane for (a) one attractive and one repulsive scatterer and (b) two attractive scatterererers. Zeroes 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 quasi-bound states.

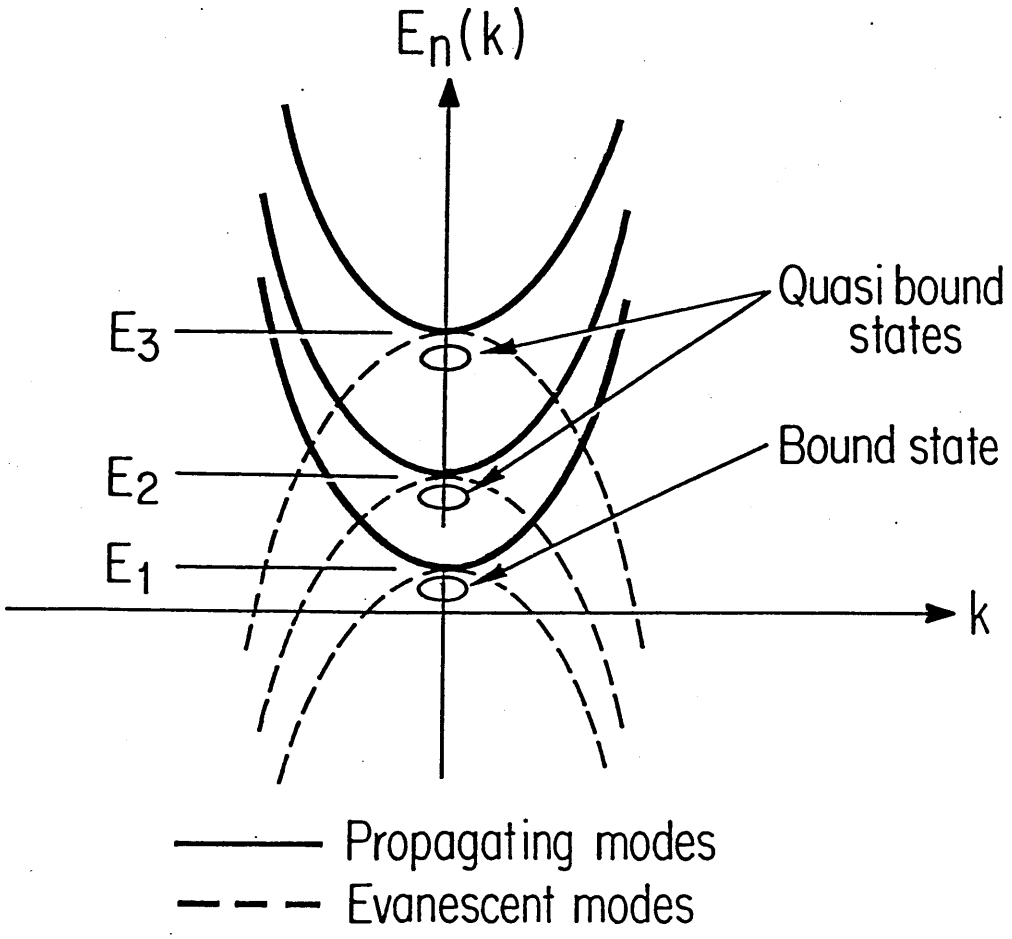
Figure 9: Normalized conductance g versus Fermi energy for (a) two repulsive scatterers, (b) one attractive and one repulsive scatterer, and (c) two attractive scatterers. The insets show more detailed behavior near the second subband minimum. Conductance 'dips' below the second subband minimum in (b)-(c)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



Quasi-1D wire



Accumulation of
Evanescent Mode



Dispersion Relation for a Wire

Figure 2

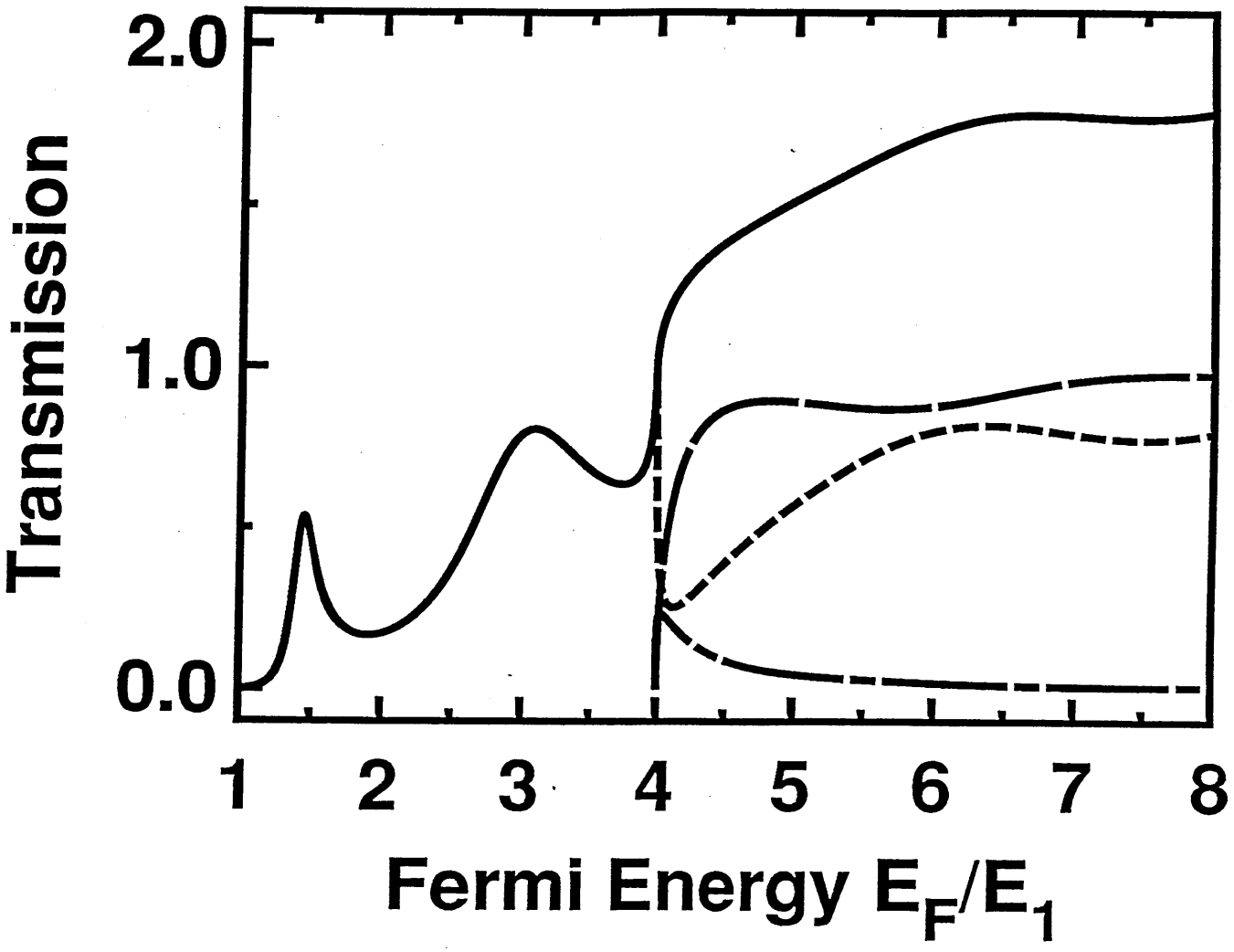


Figure 3

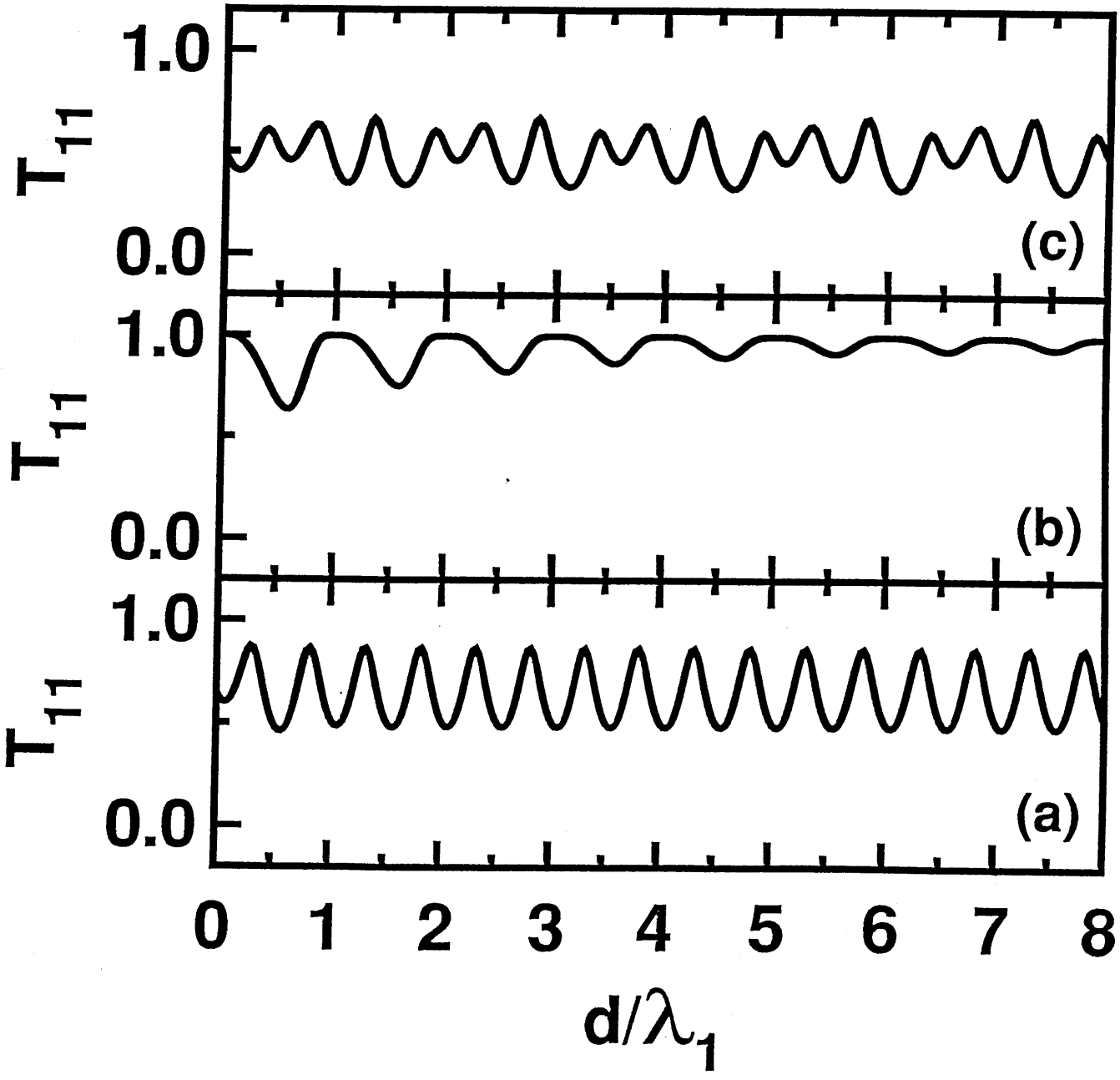


Figure 4

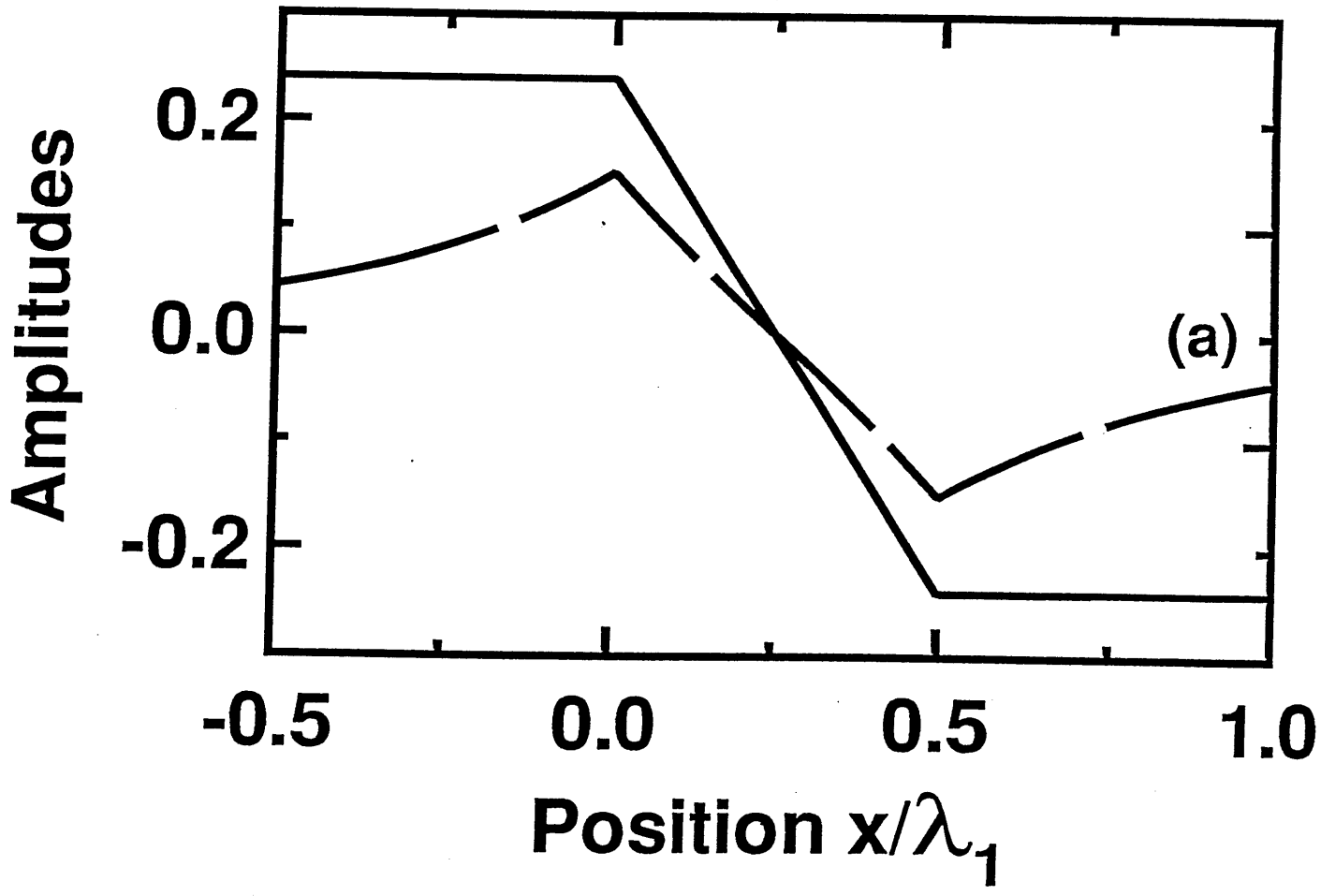


Figure 5(a)

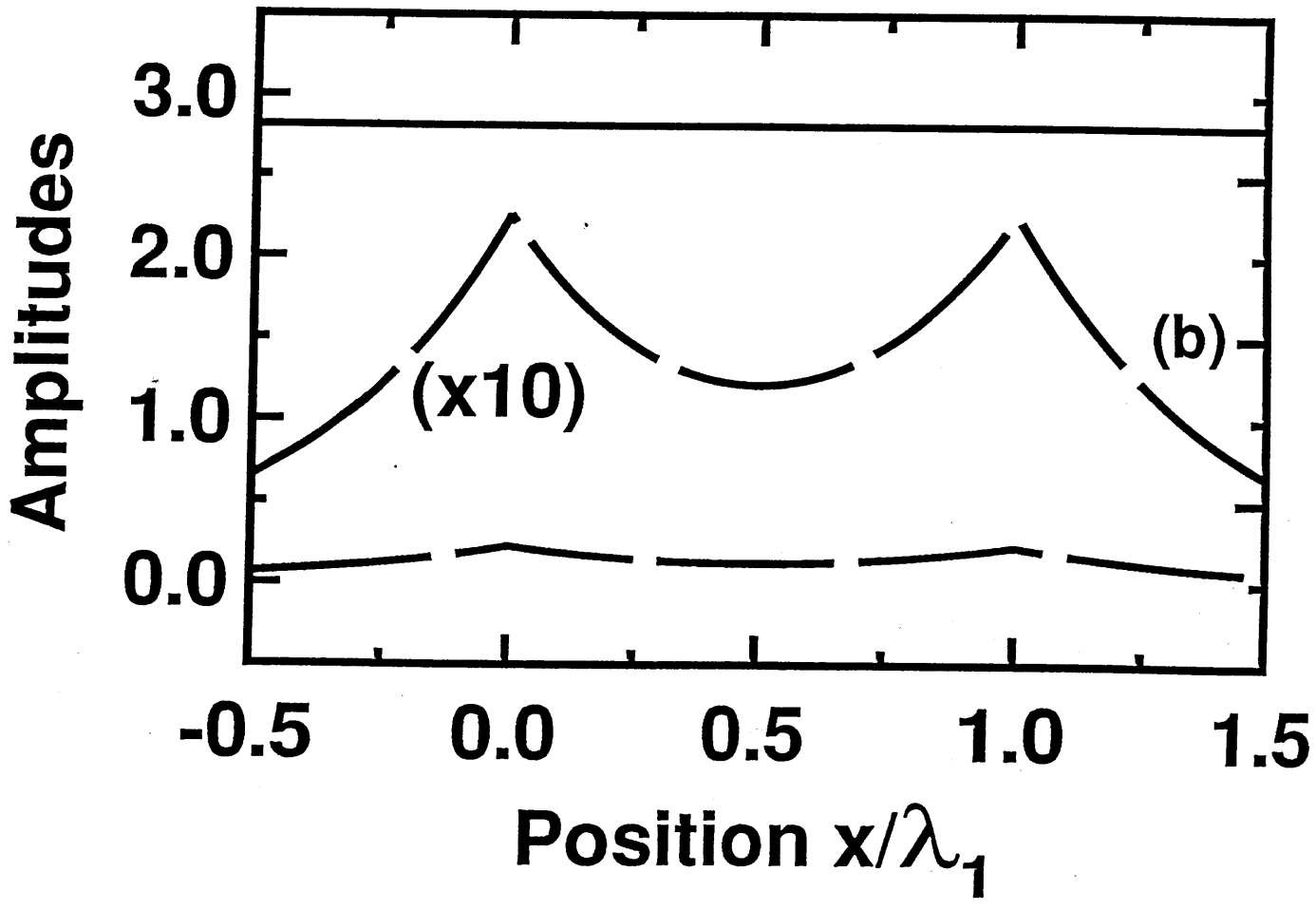


Figure 5(b)

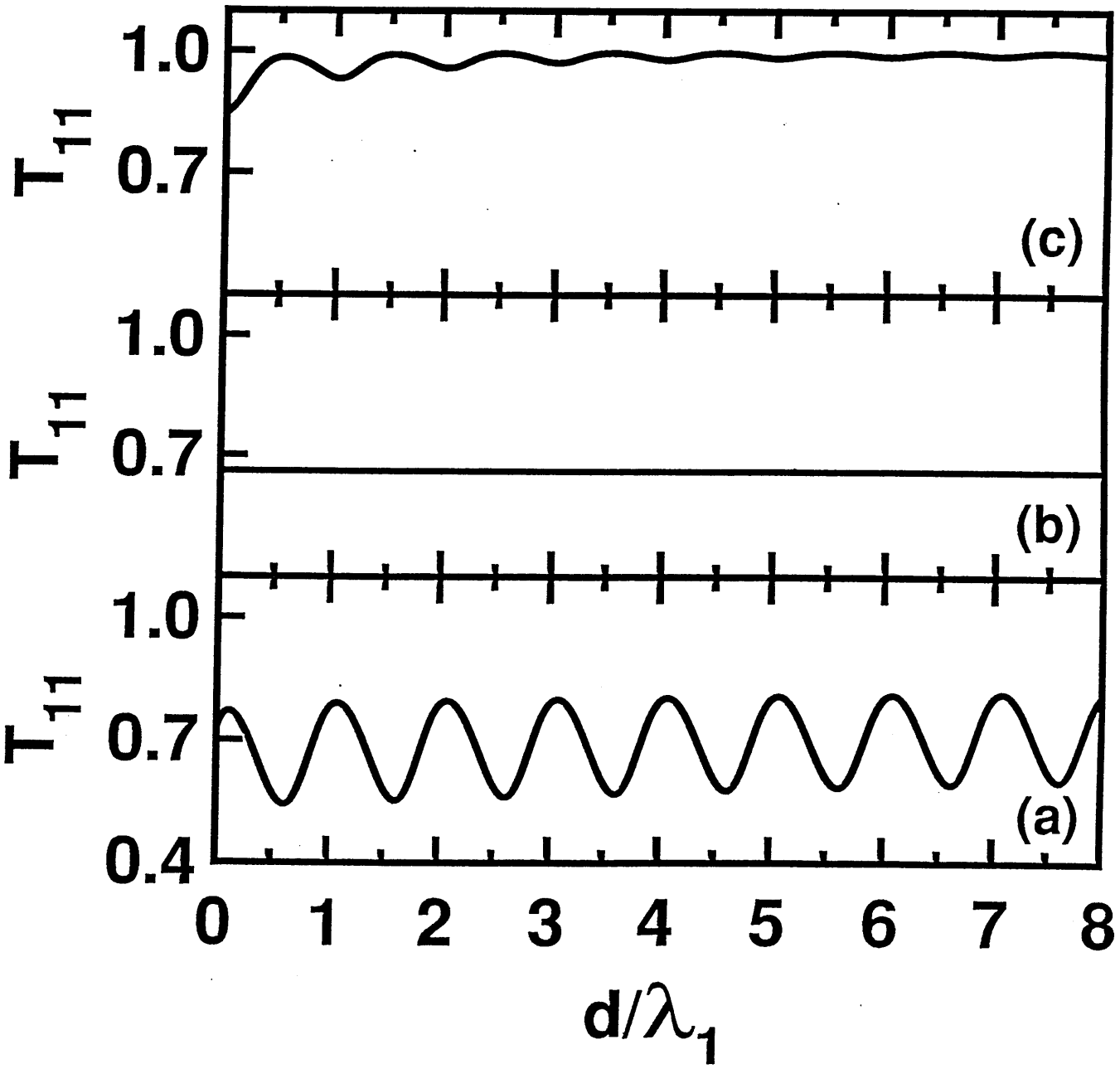
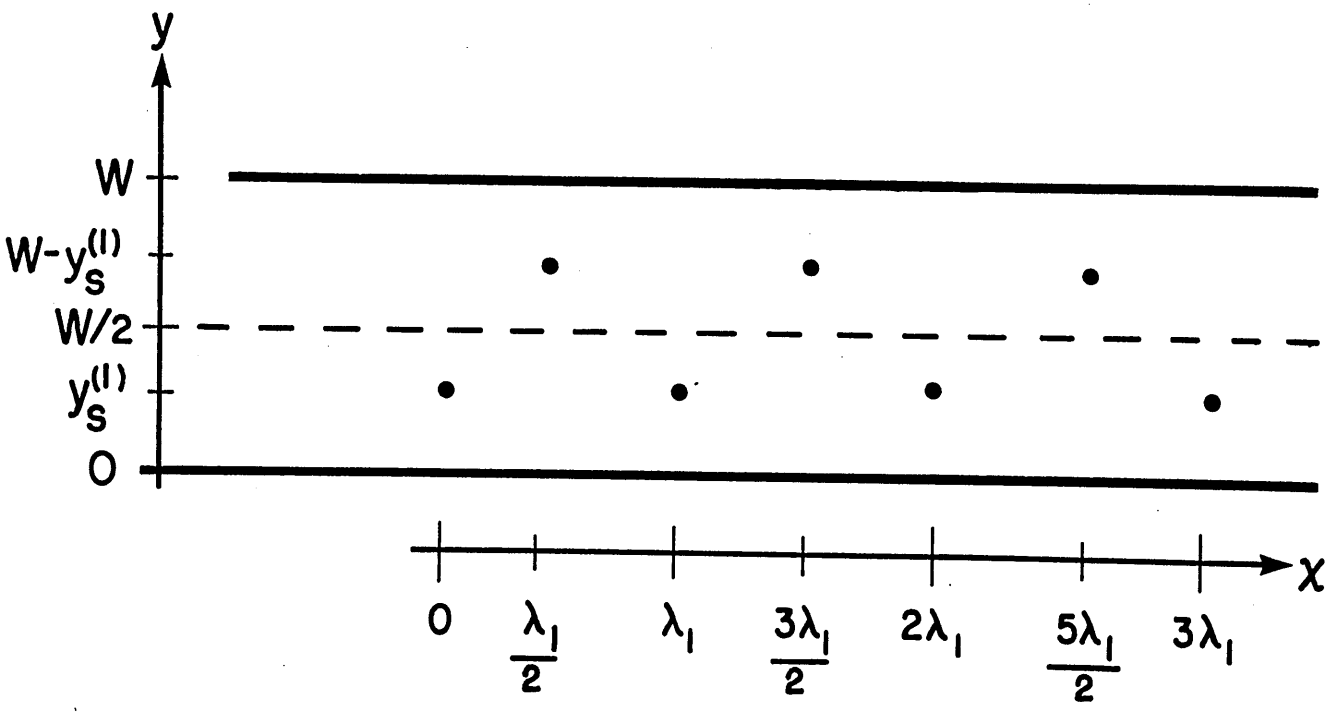


Figure 6

Scatterer Locations for
Perfect Transmission at $E = E_2$



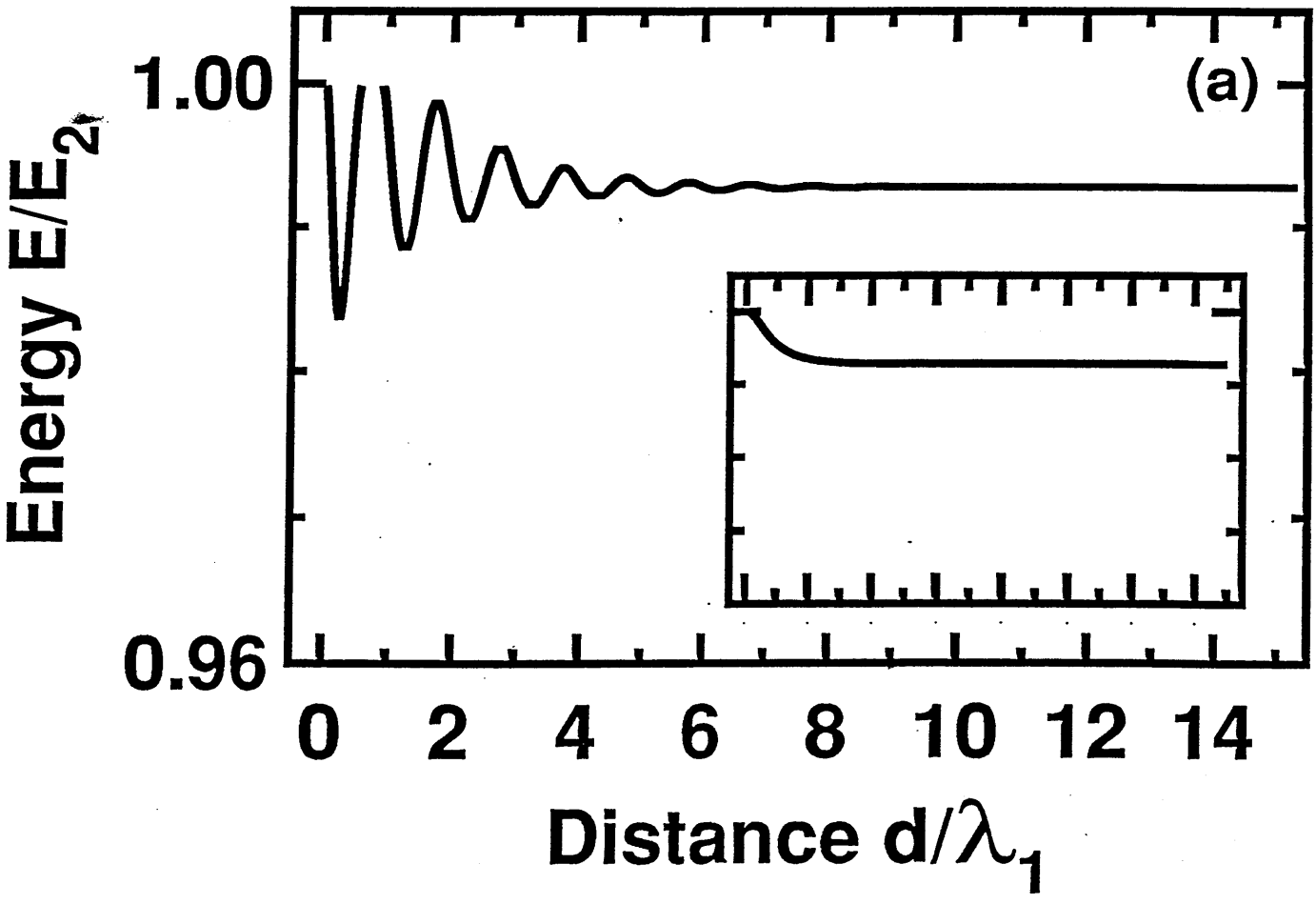


Figure 8(a)

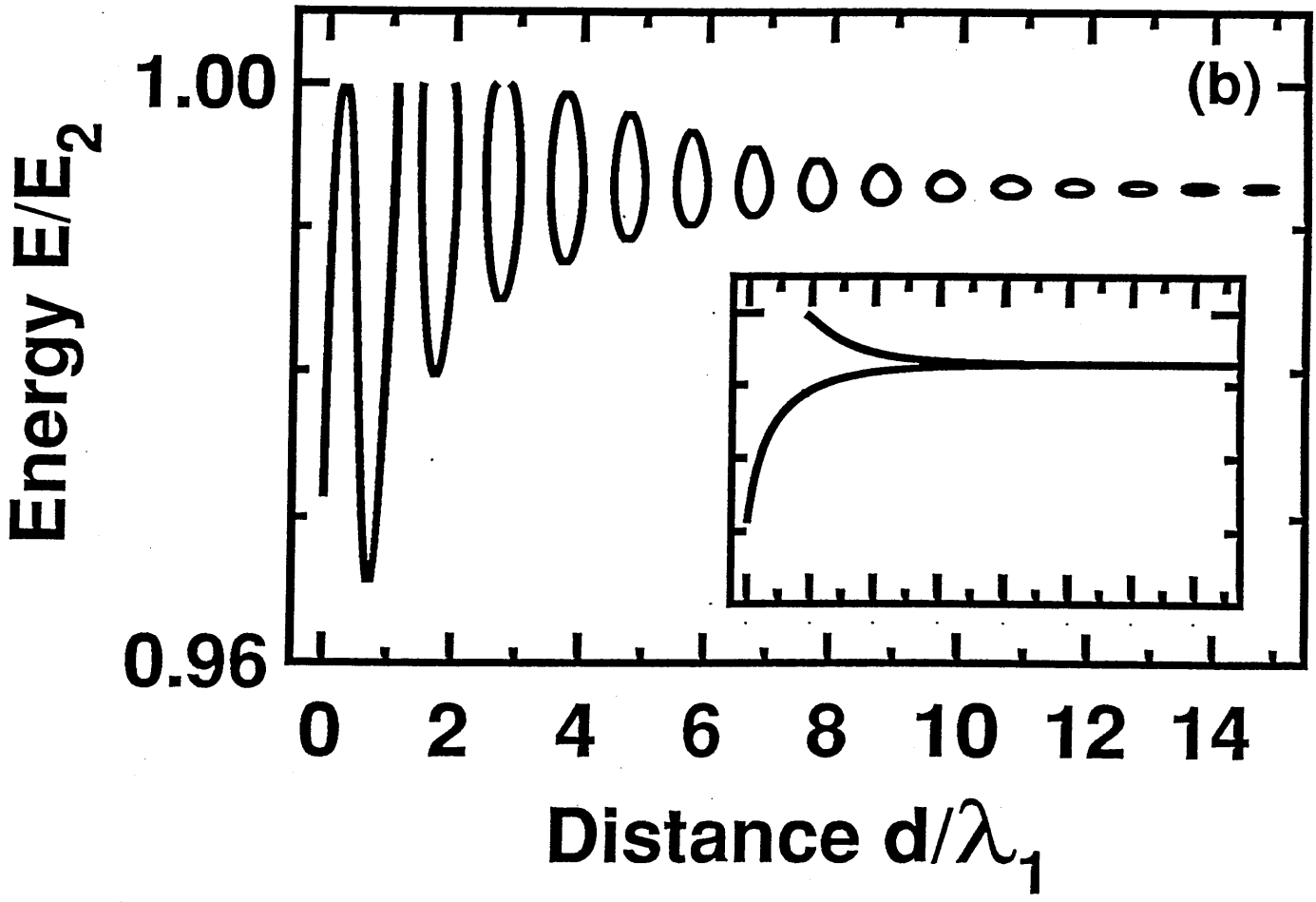
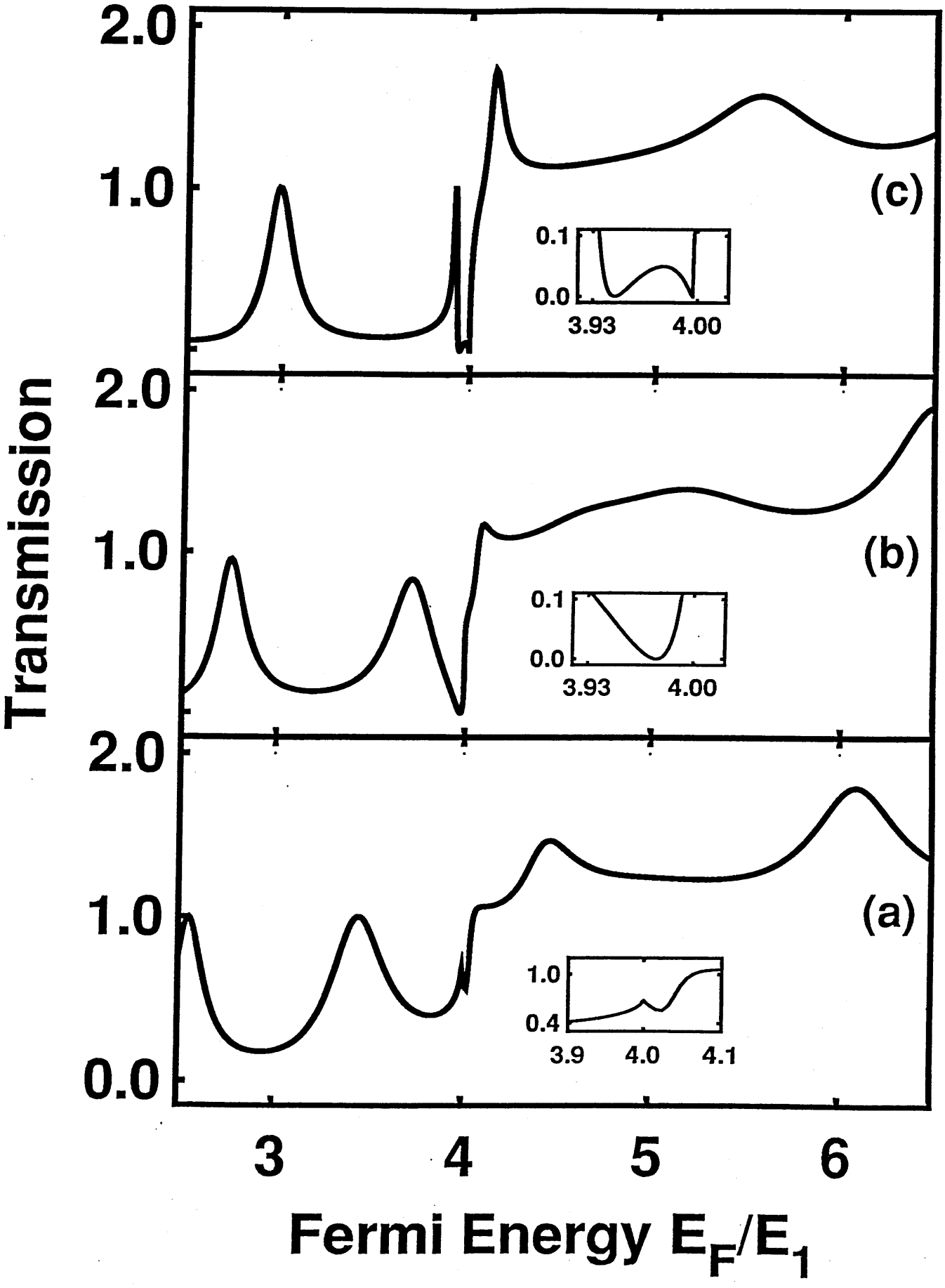


Figure 8(b)

Figure 9


2.0
1.0
(c)
0.1
0.0
3.93
4.00
2.0
1.0
(b)
0.1
0.0
3.93
4.00
2.0
1.0
0.0
(a)
1.0
0.4
3.9
4.0
4.1
Transmission
3
4
5
6
Fermi Energy E_F/E_1

Appendix J

Philip F. Bagwell, Terry P. Orlando, and Arvind Kumar, 'Low-Dimensional Resonant Tunneling', in *Resonant Tunneling in Semiconductors: Physics and Applications*, L.L. Chang and E.E. Mendez editors. (Plenum, London, in press).

To appear in: 'Resonant Tunneling
in Semiconductors: Physics and
Applications', L.L. Chang and
E.E. Mendez, editors, (Plenum
Press, London, 1991).

LOW-DIMENSIONAL RESONANT TUNNELING

Philip F. Bagwell, Terry P. Orlando, and Arvind Kumar

Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139, U.S.A.

ABSTRACT.- Resonant tunneling occurs for a wide variety of possible device geometries in one, two, and three dimensional conductors. The simplest example is the one and two dimensional analogue of the standard three dimensional resonant tunneling diode. We consider these three resonant tunneling devices operated as either a diode or a transistor by developing a 'convolution method' to calculate the device currents. Next, we consider resonant tunneling in a quasi-one-dimensional wire where the normal modes transverse to the tunneling direction cannot be neglected. Many unusual scattering properties, due to the existence of evanescent modes induced by the confinement, are found near quasi-one-dimensional subband minima or quasi-bound states in the wire.

INTRODUCTION

Many possible geometries are available for resonant tunneling in reduced dimensions, some of which are shown in Fig. 1. The overwhelming majority of studies so far have focused on resonant tunneling diodes with a three dimensional electron emitter and one direction of confinement in the quantum well, the 3DE/1DC geometry shown in Fig. 1. Space does not permit us to list the numerous experimental and theoretical contributions to resonant tunneling in reduced dimensional geometries which motivated our own work, but that is done in detail elsewhere[1]. Our purpose in this article is twofold: (1) to illustrate in section 2 an alternative 'convolution picture' for resonant tunneling currents in diodes and transistors for the simplest one, two, and three dimensional resonant tunneling geometries, and (2) to show in section 3 that if quantum confinement is strong enough that one cannot neglect the existence of normal modes induced by the confinement, there are unusual modifications to electron scattering, and resonant tunneling in particular, when the Fermi energy is near any subband minimum or quasi-bound state in a quantum wire.

The currents for all of the resonant tunneling devices in Fig. 1 can be calculated

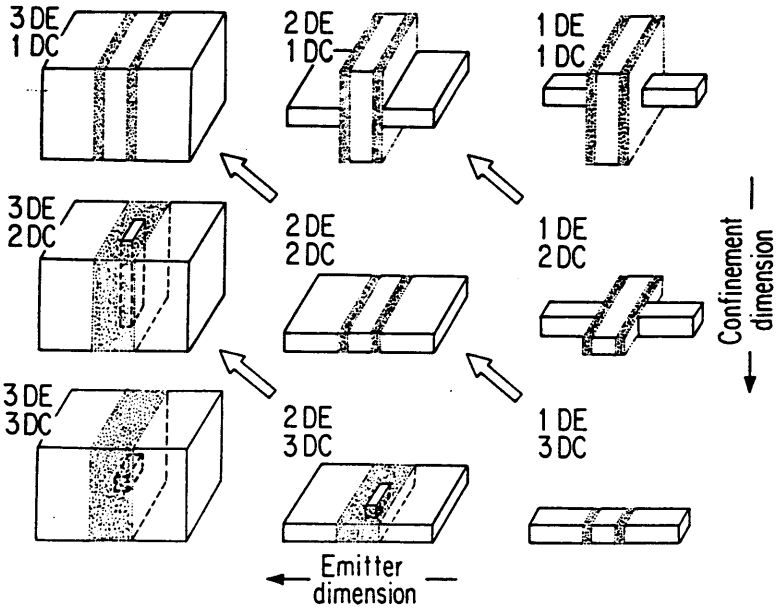


Fig. 1. Possible resonant tunneling devices having different emitter dimension (1DE, 2DE, and 3DE) and different confinement dimension in the quantum well (1DC, 2DC, and 3DC). The 'convolution method' relates the conductance of the devices along the diagonals as indicated by arrows. (From Ref. 1.)

using the two terminal Landauer formula[2 – 3] as

$$I(\mu, V, T) = \frac{e}{\pi\hbar} \sum_{ij} \int dE\, T_{ij}(E, V)\, [f(E-\mu) - f(E-\mu+eV)] \quad . \qquad (1)$$

Equation. (1) includes a factor of 2 to describe spin degenerate electrons. The indices i and j refer to the quasi-one-dimensional normal modes or scattering 'channels' induced by confinement on opposite sides of the conductor[3]. Equation. (1) can also describe interacting electrons and inelastic scattering[2,4]. For the standard one-dimensional resonant tunneling geometry $T_{11}(E) = T(E)$ is the usual transmission coefficient[5 – 6]. Section 2 shows how to separate the effects of temperature, voltage across the device, and the dimensionality of the device in Eq. (1), while section 3 presents analytical results for the transmission coefficient T_{ij} through one and two point scatterers in a quantum wire.

CONVOLUTION METHOD

Equation. (1) implies that the currents in the devices of Fig. 1 are not independent but, for the devices along the diagonal arrows shown in Fig. 1, are related by a 'convolution method' [1,7 – 8]. This 'convolution method' separates the effects of finite voltage, finite temperature, and the addition of free electron motion along one or more directions in the device, and embodies Eq. (1) in a very intuitive way for a certain class of devices. Here we apply this method to the devices along the main diagonal in Fig. 1 with the purpose of understanding conceptually how the ideal resonant tunneling transistor and diode I-V characteristics change when the electron emitter becomes three dimensional, two dimensional, or one dimensional. This presents a physical picture of resonant tunneling currents complementary to that in Refs. 9 and 10. The

final mathematical expressions for the tunneling current are the same as if calculated using the method in Refs. 9 and 10.

Consider a strictly one dimensional device where one completely ignores the other two directions perpendicular to the transport direction. Transport only in the lowest subband of a quasi-one-dimensional wire, or the 1DE/3DC device in Fig. 1, would correspond to this situation. In that case Eq. (1) reduces to[1]

$$I_{1D}(E,V,T) = \frac{e}{\pi\hbar} T(E,V) \otimes W(E,V) \otimes \left(-\frac{df}{dE}(E,T)\right) \quad . \tag{2}$$

One merely convolves[11] ($\otimes$) the transmission coefficient $T(E,V)$ with the voltage broadening function $W(E,V)$

$$W(E,V) = \begin{cases} \theta(E) - \theta(E-eV) & eV < E_F \\ \theta(E) - \theta(E-E_F) & eV > E_F \end{cases} \quad , \tag{3}$$

where $\theta(E)$ in Eq. (3) is the unit step function, and with the derivative of the Fermi distribution function

$$-\frac{df}{dE}(E,T) = \frac{1}{4k_BT}\mathrm{sech}^2\left(\frac{E}{2k_BT}\right) \quad , \tag{4}$$

to obtain the total device current. The surprising feature of Eq. (2) is that one can separate the effects of thermal and voltage broadening in Eq. (1). The convolution with $W(E,V)$ describes the emitter launching electrons over a range of speeds near the Fermi energy, while the convolution with $-df/dE$ gives the thermal variance in those speeds. These are shown as the first two convolutions in Fig. 2.

Next, we wish to show how the currents for devices along the main diagonal in Fig. 1 are related. The essential physical difference between the 1D, 2D, and 3D electron emitter is that the 1D emitter can launch electrons only forwards, the 2D emitter launches electrons at all angles in a half plane with equal probability, and the 3D emitter launches electrons at all solid angles in a half-sphere with equal probability. This is illustrated in momentum space on the left side of Fig. 3. To make any further progress, one must specify the shape of the tunneling potential. For the simplest case along the main diagonal in Fig. 1, the potential is free electron like perpendicular to

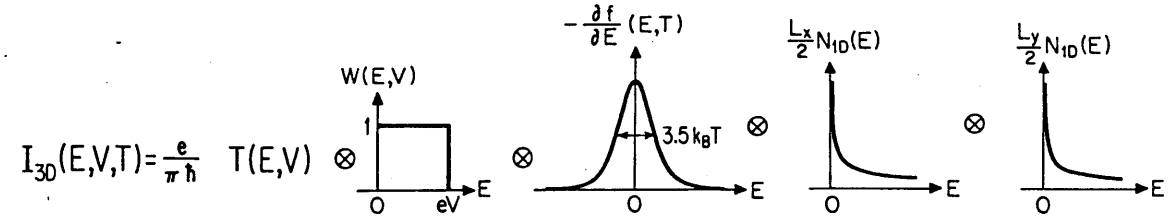


Fig. 2. The convolution method for calculating the 3DE device current corresponding to Eq. (6). The transmission coefficient $T(E,V)$ is successively convolved ($\otimes$) with the range of emitter velocities $W(E,V)$, with the thermal variance in those velocities $-df(E,T)/dE$, and with the free electron density of states $N_{1D}(E)$ corresponding to the spatial distribution of emitted electrons. (From Ref. 1.)

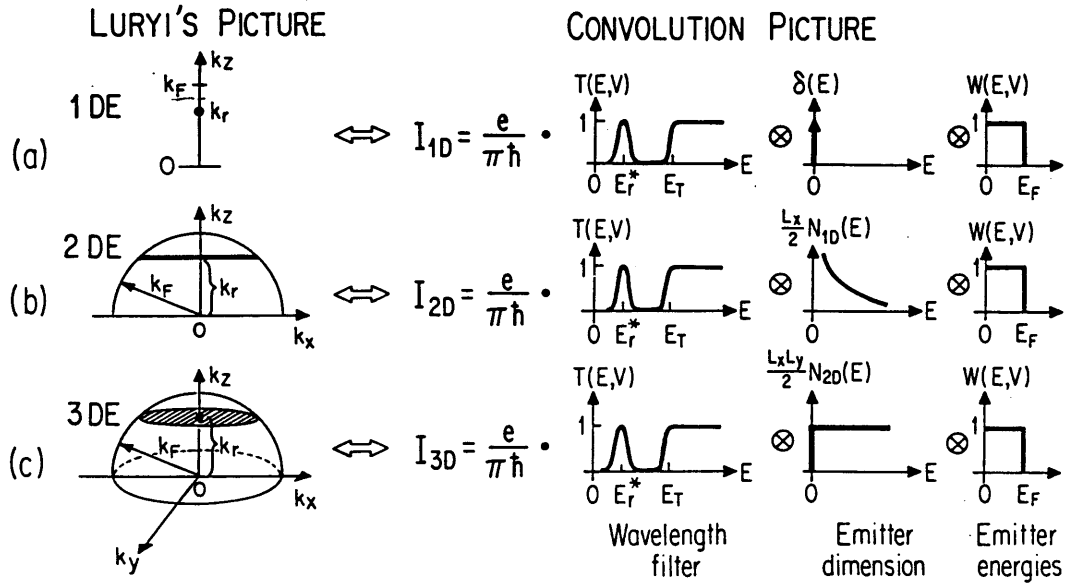


Fig. 3. Luryi's picture (left side) compared with the 'convolution picture' (right side) for calculating tunneling currents. In the convolution method, the electron transmission coefficient is convolved both with the emitter distribution $W(E,V)$ and the number density of available transverse states $N(E)$ to give the total tunneling current. (From Ref. 1.)

the tunneling direction such that transverse momentum is conserved. Therefore, along the main diagonal in Fig. 1 where $V(x,y,z) = V(z)$, so that $V(z)$ is the tunneling potential for a double barrier structure, the sum of the transverse 'channels' in Eq. (1) can be written for the 2DE/2DC device as[1]

$$I_{2D}(E,V,T) = I_{1D}(E,V,T) \otimes \frac{L_x}{2} N_{1D}(E) \quad , \tag{5}$$

and similarly for the three dimensional emitter 3DE/1DC device as

$$I_{3D}(E,V,T) = I_{2D}(E,V,T) \otimes \frac{L_y}{2} N_{1D}(E) \quad . \tag{6}$$

L_x and L_y are the size of the device along the x and y directions and $N_{1D}(E)$ is the one-dimensional free electron density of states including spin degeneracy

$$N_{1D}(E) = \frac{1}{\pi}\sqrt{\frac{2m}{\hbar^2}}\sqrt{\frac{1}{E}}\theta(E) \quad . \tag{7}$$

Equation. (5) assumes the 2DE/2DC device in Fig. 1 is strictly two dimensional, and completely ignores the vertical confinement. This is a good assumption if transport occurs only in the lowest subband normal to the confinement surface of an inversion layer or thin film. The 'convolution' formula for the 3DE/1DC resonant tunneling current from Eq. (6) reduces to the standard formula[10] when one inserts the definition of the convolution[11] into Eq. (6).

Convolution with the one-dimensional free electron density of states needed to go from the one-dimensional to the three-dimensional resonant tunneling device is shown

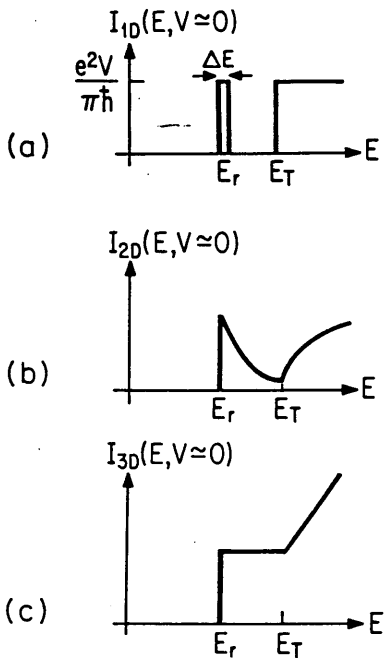


Fig. 4. Device current versus Fermi energy in the emitter for (a) the 1DE device, (b) the 2DE device, and (c) the 3DE device. These curves can correspond to transistor operation of the devices by varying E_F either by a gate voltage or by light. (From Ref. 1.)

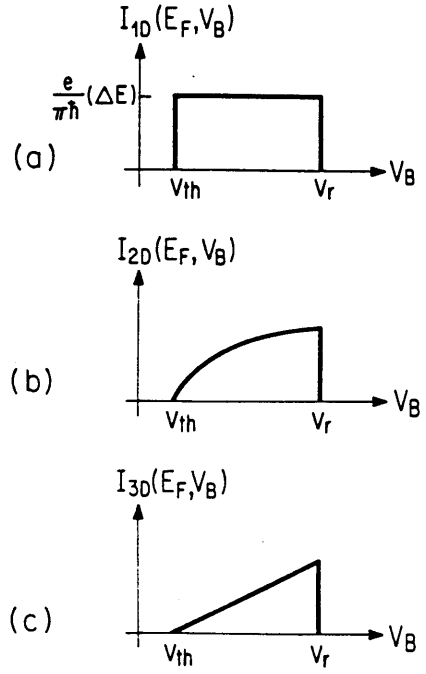


Fig. 5. Transistor current versus base voltage V_B in the quantum well for (a) the 1DE device, (b) the 2DE device, and (c) the 3DE device. We assume $2eV > E_F$, that the resonant level moves linearly with base voltage, and that the base current is zero. (From Ref. 1.)

graphically in Fig. 2 and Fig. 3. The convolution with the one dimensional free electron density of states in Eq. (5) and Eq. (6) results from simply adding classical resistors in parallel. This is due to the special shape of the tunneling potential $V(x,y,z)$. Since $V(x,y,z) = V(z)$ there is no scattering potential to couple the different transverse channels, so the scattering channels labeled by the different values of the transverse momentum are non-communicating. The transverse channels therefore add 'in parallel', by summing the current in each channel. The convolutions in Eq. (5) and Eq. (6) may also be though of as embodying the difference in the spatial distribution of the emitted electrons from a 1D emitter to a 2D and 3D emitter. A convolution method similar to that in Eq. (5) above can be developed to connect the calculated conductance for devices along any of the diagonal arrows in Fig. 1, not just those devices along the main diagonal[1]. The standard method, sometimes called 'Luryi's method', has recently been applied to tunneling from a 3D emitter through quantum wires and dots (the left hand column of Fig. 1) by Liu and Aers[12]. Luryi's method has also been applied by Ohno, Mendez, and Wang[13] to study tunneling currents when the effective mass differences in the semiconductor well and barrier material are taken into account.

A plot of current versus Fermi energy for the devices along the main diagonal of Fig. 1, assuming the emitter to collector voltage V is small, is shown in Fig. 4. In Fig. 4 the energy E_T signifies the top of the double barrier potential under zero bias, E_r is the position of the resonant level, and ΔE the area under the transmission resonance

peak so that $\int_p T(E)dE = \triangle E$. The dependence of the current on Fermi energy is qualitatively different for each different spatial dimension of the emitter. Fig. 4 corresponds to a type of transistor operation of resonant tunneling devices which can be realized either by incorporating a substrate gate into HEMT devices where the energy axis can be thought of as a gate voltage, or by varying the Fermi energy with LED light[14] where the energy axis may be thought of as the light intensity. Current in the 1DE device of Fig. 4(a) is proportional to the transmission coefficient. The current traces the resonant peak when E_F is below the top of the double barriers and yields the quantum contact resistance of $h/2e^2$ when the Fermi energy is above the barriers. Current in the 2DE device of Fig. 4(b) follows by convolving $T(E)$ with the 1D free electron density of states. Current in the 3DE device of Fig. 4(c) follows by convolving $T(E)$ with the 2D free electron density of states.

Next, we consider adding a contact to the quantum well base. If we assume the base contact moves the resonant level in the base linearly with the applied base voltage, and that the emitter to collector voltage is large, one obtains[1] the curves shown in Fig. 5. Fig. 5 results from simply convolving the curves in Fig. 4 with the electron energy distribution $W(E,V)$ in the emitter. Negative transconductance is obtained for the 1DE, 2DE, and 3DE devices. Note that in one dimension the current saturates at $I = e(\triangle E)/\pi\hbar$. Therefore, if one could fabricate a truly one dimensional resonant tunneling device, it would be possible to measure the resonance width of the transmission coefficient on a current meter. The shapes of tunneling diode curves in each spatial dimension are qualitatively similar[1] to the transistor curves of Fig. 5. The peak current in Fig. 5(b) in two dimensions is

$$I_{2D} = \frac{e}{\pi\hbar}(\triangle E)\frac{1}{\pi}\sqrt{\frac{2mE_F}{\hbar^2}}\,L_z \quad , \tag{8}$$

while for Fig. 5(c) the three dimensional peak current is

$$I_{3D} = \frac{e}{\pi\hbar}(\triangle E)\frac{mE_F}{2\pi\hbar^2}\,L_xL_y \quad , \tag{9}$$

the same as in the standard method[10]. We also comment that, for a ballistic conductor in one dimension, the current saturates at a value $I = eE_F/\pi\hbar$ proportional to the Fermi energy when $eV > E_F$.

We discuss one problem with this approach to calculating tunneling currents (which is also a problem in the standard approach): In the limit of a standard wide tunneling device, where the conductor opens up into a continuum of plane wave scattering states and one imposes periodic boundary conditions on the problem, Eq. (1) evolves into Eq. (6). While this limit of Eq. (1) is mathematically appealing, it involves a subtle physical assumption. Neglecting the quantum confinement in Eq. (1) implies that the electrons do not feel the confining walls of the conductor so that one can impose periodic boundary conditions. This can only be true if weak 'phase randomizing' scattering is invoked inside the conductor. The phase breaking scattering must be weak enough so that the electrons are not significantly reflected by the scattering, but strong enough so that the information about the boundaries of the conductor is not retained by the tunneling electron. A rigorous treatment of the conditions necessary for Eq. (1) to cross over into Eq. (6) has yet to be supplied.

TUNNELING IN A QUANTUM WIRE

We wish to obtain the conductance of a quantum wire when one or two potential barriers are present using Eq. (1), the Landauer formula. Transport in the quantum wire of Fig. 6 is described by the Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\left(\frac{\partial^2}{\partial x^2}+\frac{\partial^2}{\partial y^2}\right)+V_c(y)+V_d(x,y)\right]\psi(x,y)=E\psi(x,y) \quad , \tag{10}$$

where $V_d(x,y)$ is the potential of any impurities or defects in the wire. The confinement potential $V_c(y)$ defines a set of normal modes $\chi_n(y)$ satisfying

$$\left[-\frac{\hbar^2}{2m}\frac{d^2}{dy^2}+V_c(y)\right]\chi_n(y)=E_n\chi_n(y) \quad , \tag{11}$$

where E_n is the subband energy for mode n. The electron dispersion relation $E = E_n + \hbar^2 k_n^2/2m$ for each subband n is also shown in Fig. 6.

Single subband transport never strictly exists in a confined geometry such as the quantum wire shown in Fig. 6. If the Fermi energy lies in the lowest subband there are propagating states for which the wavefunction has the form

$$\psi(x,y) \sim e^{ik_1 x}\chi_1(y) \quad , \tag{12}$$

and evanescent states for which the wavefunction has the form

$$\psi(x,y) \sim e^{-\kappa_2 x}\chi_2(y) \quad . \tag{13}$$

Here κ_2, defined by

$$\kappa_2 = ik_2 = i\sqrt{\frac{2m(E_2-E)}{\hbar^2}} \quad , \tag{14}$$

is the imaginary wavevector for the second normal mode. Note that the dispersion

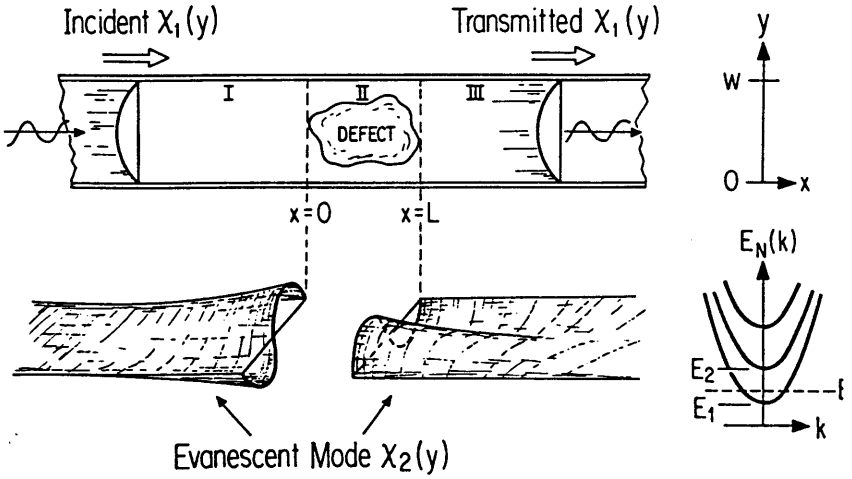


Fig. 6. Scattering from a defect in a quasi-one-dimensional wire where a steady current incident in the lowest mode is applied from the left. The scattering potential couples the incident propagating mode to the evanescent modes, causing probability density to accumulate locally around the defect. (From Ref. 18.)

relation for the evanescent modes is $E = E_n - \hbar^2\kappa_n^2/2m$ for each subband n. The wavefunction on the left side of the scattering defect where $x < 0$ can be expanded as a linear combination of both the propagating and evanescent modes of the clean wire as

$$\psi(x,y) = \sum_n \left(A_n e^{ik_n x} + B_n e^{-ik_n x}\right) \chi_n(y) \quad . \tag{15}$$

On the right of the defect in Fig. 6 where $x > L$, the wavefunction can be expanded as

$$\psi(x,y) = \sum_n \left(C_n e^{ik_n x} + D_n e^{-ik_n x}\right) \chi_n(y) \quad . \tag{16}$$

If a defect is present in the wire, propagating modes such as Eq. (12) are coupled through the scattering potential to the evanescent modes such as Eq. (13). Consider a steady current flow incident from the left on a defect in a quantum wire shown in Fig. 6. The scattering potential will cause some of the electrons to be scattered out of the propagating incident wave into an evanescent wave, and a localized mode will therefore build up around the defect even if the scatterer is repulsive. This phenomenon is similar to the electromagnetic energy storage near any defects or sudden discontinuities in a microwave waveguide. Since the scattering boundary conditions require the wavefunction and its derivative be continuous everywhere, the presence of the localized mode built up around the scatter will change the scattering of the incident waves. This is true even though the evanescent waves carry no current. References. 15-22 discuss some of the unusual scattering properties of electrons in a confined geometry.

In this section we first consider the case of a wire having uniform width subject to a delta function scattering potential[17 – 19]

$$V_d(x,y) = \gamma\delta(x)\delta(y - y_i) \quad . \tag{17}$$

Requiring the wavefunction be continuous at $x = 0$ forces

$$A_n + B_n = C_n + D_n \tag{18}$$

for all propagating and evanescent modes n. Integrating Eq. (10) across the defect gives the derivative jump condition

$$ik_n(C_n - D_n) - ik_n(A_n - B_n) = \sum_m \Gamma_{nm}(A_m + B_m) \quad , \tag{19}$$

where $\hbar^2\Gamma_{ab}/2m = \gamma\chi_a(y_i)\chi_b(y_i) \equiv V_{ab}$. Forcing particles incident only from the left requires $D_n = 0$ for all propagating modes. We take only a single incident propagating mode A_n to be nonzero. Additionally we require $A_n = D_n = 0$ for all evanescent modes so that the wavefunction is normalizable. The transmission coefficients T_{mn} for propagating modes m and n required to evaluate Eq. (1) can then be obtained as

$$T_{mn} = \frac{k_m}{k_n}\frac{C_m}{A_n} \qquad \text{and} \qquad R_{mn} = \frac{k_m}{k_n}\frac{B_m}{A_n} \quad , \tag{20}$$

where R_{mn} are the reflection coefficients.

This scattering problem through a single point defect in Eq. (17) is easy to state, and the transmission and reflection coefficients easily obtained by solving Eqs. (18) and (19), but the results are not intuitive. We have solved this problem both by simply matching wavefunctions and their derivatives at the scatterer[18], and by a sum over scattering histories approach solving the Dyson Equation in a quasi-1D wire[19]. Both approaches give the same answer for the current transmission coefficients T_{mn}. One essential insight obtained is that, just as the propagating electron density of states diverges as $1/\sqrt{E-E_n}$ when the Fermi energy approaches each subband minimum from above, the evanescent electron density of states diverges as $1/\sqrt{E_n-E}$ when the Fermi energy approaches each subband minimum from below[18]. Thus, the evanescent modes determine how the electrical conductance behaves as the Fermi energy crosses each new quasi-one-dimensional subband.

The reflection amplitudes r_{ab} through the point scatterer in Eq. (17) have the simplest form and are

$$r_{ab}(E) = \frac{-iV_{ab}\dfrac{m}{\hbar^2\sqrt{k_a k_b}}}{1+\sum_n^e V_{nn}\dfrac{m}{\hbar^2\kappa_n}+i\sum_n^p V_{nn}\dfrac{m}{\hbar^2 k_n}} \quad , \tag{21}$$

where the reflection coefficients are found from $R_{ab}=|r_{ab}|^2$. In Eq. (21) the $\sum_n^p$ denotes a sum over the propagating modes of the wire, $\sum_n^e$ denotes a sum over the evanescent modes, and modes a and b are assumed propagating. The current transmission amplitudes are determined as $t_{ab}=r_{ab}$ when $a\neq b$ and $t_{aa}=1+r_{aa}$ when $a=b$. This follows from enforcing wavefunction continuity at $x=0$. The transmission coefficients in the conductance formula Eq. (1) are then found from $T_{ab}=|t_{ab}|^2$. Note that $m/\hbar^2 k_n$ is simply the density of propagating or evanescent states for mode n, while V_{nn} is a scattering matrix element. Thus, each term in Eq. (21) can be interpreted in terms of a Fermi Golden Rule scattering rate[18].

The conductance determined from Eq. (1) with the transmission coefficients determined from Eq. (21) is shown in Fig. 7. Figure. 7(a) shows a weak repulsive scatterer (γ =10 feV cm^2, solid line) and a weak attractive scatterer (γ =-6 feV cm^2, dashed line). At each subband minimum perfect transmission is obtained for all the modes in the wire, a result first pointed out in Ref. 17. The shoulders of the quantized conductance steps are rounded for the case of the repulsive scatterer due to increased reflection immediately above each new subband minimum. The extra dips in the conductance through the attractive scatterer correspond to new quasi-bound states forming in the attractive potential. Outside the attractive potential the evanescent mode decays on either side, but locally inside of the potential well the evanescent mode can propagate. The condition that an integral number of electron wavelengths of the evanescent mode can fit inside the attractive defect roughly defines the condition for a quasi-bound-state to form. These quasi-bound states resemble a donor level below the conduction band minima of a semiconductor, but now there is one new 'donor level' for each subband. These new minima in the conductance through a single attractive barrier in a wire are therefore a form of 'resonant tunneling' occurring through the evanescent mode, and lead to minima rather than maxima in transmission.

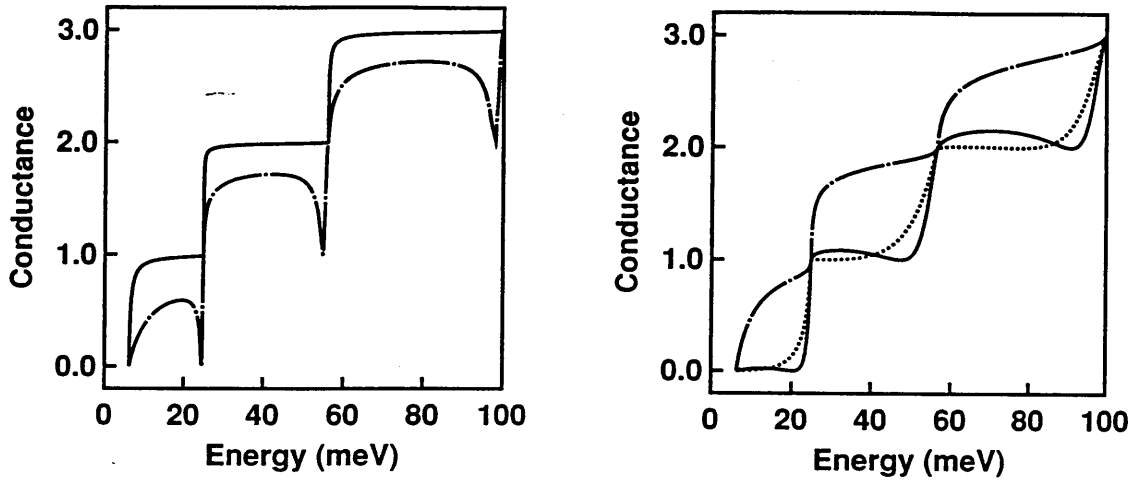


Fig. 7. Conductance through a single scatterer in a quasi-one-dimensional wire as a function of Fermi energy. (a) compares the conductance through a repulsive defect (solid) and an attractive defect (dashed). (b) studies the motion of the quasi-bound-state in the attractive defect. Paradoxically in (b), the strongest scatterer (dashed) gives the largest conductance. (From Ref. 18.)

Figure. 7(b) shows the movement of the quasi bound state as the attractive scatterer is made stronger. For an attractive scatterer slightly stronger than Fig. 7(a) (γ =-8 feV cm^2, solid line), the quasi bound state moves lower in energy and the conductance decreases. For an even stronger attractive scatterer (γ =-9 feV cm^2, dotted line) the minima in transmission due to the quasi-bound state move so close to the next lowest subband minimum that they are is no longer visible, while for the strongest attractive scatterer (γ =-20 feV cm^2, dashed line) each quasi-bound state has moved completely out of the subband in which it originated and into the next lowest subband. Note that the electrical conductance actually *increases* for all values of the Fermi energy in Fig. 7(b) from the medium strength scatterer (dotted line) to the strongest scatterer (dashed line). This paradoxical result in Fig. 7(b), that conductance actually increases as the scatterer is made stronger, occurs because the quasi-bound states have now moved out of the energy range in which they can block transmission of the incident electron.

We next consider electron transmission through two delta function barriers in a quasi-one-dimensional wire where the scattering potential is

$$V_d(x,y) = \gamma_1 \delta(x)\delta(y-y_1) + \gamma_2 \delta(x-d)\delta(y-y_2) \quad , \tag{22}$$

shown schematically in Fig. 8. Evanescent modes, which have a spatial extent of $1/\kappa_2$ if the electron Fermi energy is in the lowest subband, now accumulate around both barriers in the wire as shown in Fig. 8. When the separation between the barriers $d < 1/\kappa_2$, the evanescent modes building up around one barrier will extend to the other barrier and alter the scattering properties of the electrons.

The transmission coefficient T_{11} through the potential of Eq. (22) for the case of two allowed modes is calculated analytically in Refs. 21-22, and analyzed numerically when many modes are present in Ref. 22. The equations which must be solved to obtain the transmission coefficients are very similar to Eqs. (18) and (19) in this

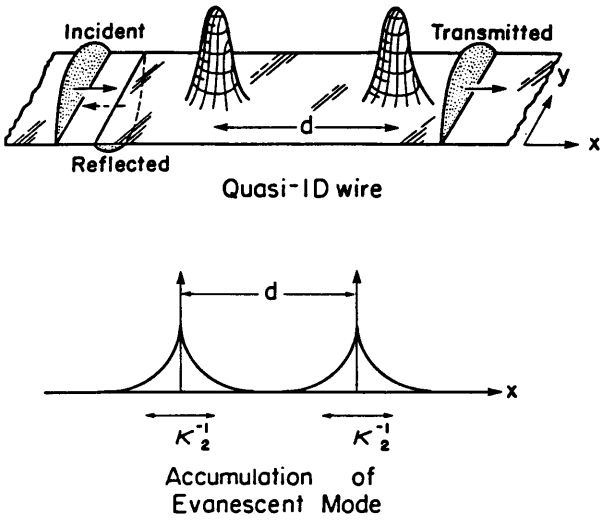


Fig. 8. Scattering from two defects (resonant tunneling) in a quasi-one-dimensional wire. Evanescent modes accumulate around both scattering defects. When the evanescent mode building up around one barrier overlaps the second barrier, deviations from standard one-dimensional resonant tunneling theory occur. (From Ref. 22.)

section. We display T_{11} in Fig. 9, where we have chosen two repulsive barriers having $\gamma_1 = \gamma_2 = 10$ feV cm^2 and $y_1 = y_2$. When $E = 0.9E_2$ in Fig. 9(bottom), we recover almost the same qualitative behavior for T_{11} as a function of distance d between the scatterers as in the usual 1D transmission coefficient[5 – 6]. In the 1D case the transmission coefficient is periodic with period $\lambda_1/2$ where λ_1 the incident electron wavelength. However, the first two transmission minima in Fig. 9(bottom) are slightly different than all the others, indicating that the evanescent modes are overlapping at small d. Thus, when the Fermi energy is below the second subband minimum, we find only small modifications to the standard one-dimensional resonant tunneling theory. This is because the density of evanescent states far below the second subband minimum is small.

At the second subband minimum, $E = E_2$ in Fig. 9(middle), the evanescent modes completely dominate over the Fabry-Perot wave interference effect. Firstly, the transmission coefficient becomes highly aperiodic. The transmission minima rise as the scatterer separation increases and the transmission actually approaches unity for large d. Secondly, the perfect transmission resonances are separated by a wavelength of the incident electron, rather than by a half wavelength. Furthermore, perfect transmission occurs for $d = 0$ in agreement with the case of a single scatterer. The scattering phase shift present in the 1D case and in Fig. 9(bottom) is completely absent at $E = E_2$. The change in oscillation period from half the incident electron wavelength to a full wavelength is robust and does not depend critically on the shape of the scatterers. The 'beating' pattern in T_{11} when both modes are propagating, shown for $E = 1.1E_2$ in Fig. 9(top), is also a new feature of the multimode wire.

At the second subband minimum, $E = E_2$, the spatial shape of the lowest evanescent mode in the wire determines the scattering properties of the propagating mode.

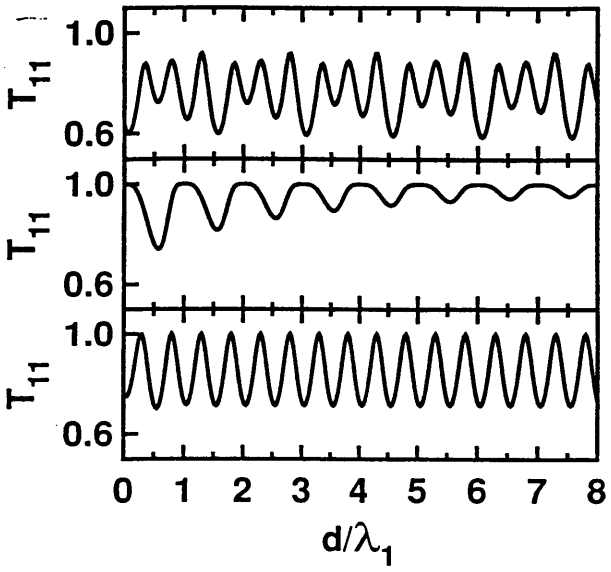


Fig. 9. Transmission coefficient T_{11} through two point defects in a wire as a function of distance d between the scatterers for energies $E = 0.9E_2$ (bottom), $E = E_2$ (middle), and $E = 1.1E_2$ (top). The Fabry-Perot effect present when $E = 0.9E_2$ is completely extinguished when $E = E_2$. (From Ref. 21.)

Since the evanescent modes are populated by scattering from the applied incident current, and the coupling to the evanescent mode is sensitive to the phase of the incident electron, the general shape of the evanescent mode is repeated when the barrier separation is increased by a full wavelength[22]. If the barriers are equal and aligned, a shape resembling a 'bonding' orbital in molecular physics recurs when $d = j\lambda_1$, $j = 0, 1, 2...$, and perfect transmission results. The x-dependence of the lowest evanescent mode resembles an 'antibonding' state when the barriers are separated by a half integer multiple of the electron wavelength as $d = (j + 1/2)\lambda_1$, where some fraction of the incident electrons are reflected. If the evanescent mode must change its value from one scatterer to the next, as it does when an 'anti-bonding' state is formed, it can do so more gradually as d becomes large. This suppresses reflection of the incident mode at each barrier so that the transmission becomes perfect when the barriers have infinite separation.

CONCLUSIONS

We have presented an intuitive 'convolution method' for the transistor and diode currents in a resonant tunneling device. The method is conceptually important because it separates the effects of finite temperature, finite voltage, and free electron motion perpendicular to the tunneling direction in the formula for tunneling current. The voltage, thermal, and dimensional convolutions can be easily evaluated graphically to calculate the tunneling current. Each current has a characteristic dependence on Fermi energy, emitter to collector voltage, and potential energy in the quantum well base which is different for 1D, 2D, and 3D electron emitters.

We then examined the transmission coefficients through one and two point scatterers in a quantum wire. Evanescent modes are shown to dominate the transmission properties and electrical conductance when the Fermi energy is near a subband minimum or quasi-bound-state in the wire. When a single attractive scatterer is present in the wire, a type of 'resonant tunneling' occurs via the evanescent mode and leads to minima in the conductance as a function of Fermi energy. When two repulsive barriers are placed in a quantum wire, and when $E \simeq E_2$, the transmission coefficient T_{11} as a function of distance between the scatterers has maxima separated by the full wavelength of the incident electron, rather than the usual 1D result in which successive transmission maxima are separated by a half wavelength. The spatial shape of the lowest evanescent mode in 'bonding' and 'anti-bonding' orbitals determines the behavior of the conductance at a subband minimum, rather than wave interference between propagating modes in the wire, so that the Fabry-Perot interference effect between propagating modes is completely suppressed.

ACKNOWLEDGEMENTS.- We thank T.P.E. Broekaert, K. Ismail, C.G. Fonstad, J. del Alamo, R.A. Ghanbari, P.J. Price, M.A. Kastner, D.A. Antoniadis, and Henry I. Smith for useful discussions. This work was sponsored by the U.S. Air Force Office of Scientific Research under grant AFOSR-88-0304 and the Joint Services Electronics Program contract DAAL03-89-C-0001.

REFERENCES

1. P.F. Bagwell, T.P.E. Broekaert, T.P. Orlando, and C.G. Fonstad, 'Resonant Tunneling Diodes and Transistors with a One, Two, or Three Dimensional Electron Emitter', to appear in J. Appl. Phys., October 15, 1990.

2. R. Landauer, 'Electrons as Guided Waves in Laboratory Structures: Strengths and Problems', *Analogies in Optics and Micro-Electronics*, W. van Haeringen and D. Lenstra eds., (Kluwer, Academic Press, in press).

3. M. Büttiker, 'Symmetry of Electrical Conduction', IBM J. Res. Dev., 32:317 (1988).

4. M. Büttiker, 'Coherent and Sequential Tunneling in Series Barriers', IBM J. Res. Dev., 32:63 (1988).

5. P.J. Price, 'Theory of Resonant Tunneling in Heterostructures', Phys. Rev. B, 38:1994 (1988).

6. G.A. Toombs and F.W. Sheard, 'The Background to Resonant Tunneling Theory', *Electronic Properties of Multilayers and Low-Dimensional Structures*, J.M. Chamberlain, L. Eaves, and J.C. Portal, eds., (Plenum, New York, in press).

7. P.F. Bagwell and T.P. Orlando, 'Landauer's Conductance Formula and its Generalization to Finite Voltages', Phys. Rev. B, 40:1456 (1989).

8. P.F. Bagwell and T.P. Orlando, 'Broadened Conductivity Tensor and Density of States for a Superlattice Potential in One, Two, and Three Dimensions', Phys. Rev. B, 40:3757 (1989).

9. S. Luryi, 'Frequency Limit of Double-Barrier Resonant Tunneling Oscillators', Appl. Phys. Lett., 47:490 (1985). See also L.V. Iogansen, 'Thin-Film Electron Interferometers', Sov. Phys. Usp., 8:413 (1965).

10. R. Tsu and L. Esaki, 'Tunneling in a Finite Superlattice', Appl. Phys. Lett., 49:562 (1973).

11. The convolution symbol $\otimes$ has its usual meaning

$$A(E) \otimes B(E) = \int_{-\infty}^{\infty} A(E')B(E-E')dE' = \int_{-\infty}^{\infty} A(E-E')B(E')dE' \qquad (23)$$

12. H.C. Liu and G.C. Aers, 'Resonant Tunneling Through One-, Two-, and Three-Dimensionally Confined Quantum Wells', J. Appl. Phys., 65:4908 (1989).

13. H. Ohno, E.E. Mendez, and W.I. Wang, 'Effects of Carrier Mass Differences on the Current Voltage Characteristics of Resonant Tunneling Structures', Appl. Phys. Lett., 56:1793 (1990).

14. K. Ismail, D.A. Antoniadis, and H.I. Smith, 'Lateral Resonant Tunneling in a Double-Barrier Field-Effect Transistor', Appl. Phys. Lett., 55:589 (1989).

15. F.M. Peeters, 'Bound and Resonant States in Quantum Wire Structures', *Science and Engineering of 1- and 0- Dimensional Semiconductors*, S. Beaumont and C. Sotomayor-Torres eds., (Plenum, New York, in press).

16. J. Masek, P. Lipavsky, and B. Kramer, 'Coherent Potential Approach for the Zero Temperature DC Conductance of Weakly Disordered Narrow Systems', J. Phys.: Cond. Matt., 1:6395 (1989).

17. C.S Chu and R.S. Sorbello, 'Effect of Impurities on the Quantized Conductance of Narrow Channels', Phys. Rev. B, 40:5941 (1989).

18. P.F. Bagwell, 'Evanescent Modes and Scattering in Quasi-One-Dimensional Wires', Phys. Rev. B, 41:10354 (1990).

19. P.F. Bagwell, 'Solution of Dyson's Equation in a Quasi-1D Wire', J. Phys.: Cond. Matt., 2:6179 (1990).

20. E. Tekman and S. Ciraci, 'Ballistic Transport Through Quantum Point Contacts: Elastic Scattering by Impurities', to be published.

21. A. Kumar and P.F. Bagwell, 'Resonant Tunneling in a Multi-Channel Wire', to appear in Solid State Comm., 75:949 (1990).

22. A. Kumar and P.F. Bagwell, 'Resonant Tunneling in a Quasi-One-Dimensional Wire: Influence of Evanescent Modes', to be published.

Appendix K

Philip F. Bagwell and Arvind Kumar, 'Comment on "Effects of Channel Opening and Disorder on the Conductance of Narrow Wires" ', submitted for publication.

Comment on 'Effects of Channel Opening and Disorder on the Conductance of Narrow Wires'

Philip F. Bagwell and Arvind Kumar

Department of Electrical Engineering and Computer Science,

Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

October 10, 1990

Kander, Imry, and Sivan [Phys. Rev. B **41**, 12941 (1990)] have argued that electron scattering in a disordered quasi-one-dimensional wire increases whenever the Fermi energy is near a subband minimum. We show that the conductance drops in the calculation of Kander *et al.* arise when quasi-bound states form in a locally attractive minimum in the scattering potential. Adding even a single attractive impurity to a disordered wire produces these conductance drops, so that their occurrence does not depend on wave interference between multiple impurity scattering events.

PACS 73.20.Dx, 73.20.Hb, 73.50.Bk

submitted to Physical Review B-15 (Received July 23, 1990)

Kander, Imry, and Sivan [1] calculate the electron transmission through an array of elastic scatterers placed on a square grid, where half the scatterers are attractive and half are repulsive. A similar calculation was also done in Ref. [2]. In interpreting the conductance 'dips' below each new subband in their calculation Kander *et al.* state: 'Since the longitudinal energy of a barely open channel is small, this channel is localized and reflects many of the incoming electrons to backwards propagating states. Such a mechanism will indeed result in conductance dips each time a new channel is opened. The effect also exists as the evanescent channel is about to be opened.'

If the interpretation of Kander *et al.* is correct, it implies that the conductance should drop on *both* sides of a subband minimum, not just below the subband. Furthermore it is known, from studies of electron transmission through a point barrier in a narrow wire [3]-[4], that in some models the electron transmission always increases as the Fermi energy passes through a subband minimum from below. So the argument given in Ref. [1], if correct, cannot be a general one.

The conductance 'dips' in the calculation of Ref. [1] actually arise from quasi-bound states forming in a local attractive minimum in the disordered

potential [3]-[5]. This 'quasi-bound state' is analogous to a donor level below the conduction band minimum of a semiconductor, where the confinement introduces a new 'quasi-donor level' below each subband in the narrow wire. The formation of quasi-bound states explains why the conductance drops occur only below the subband, so that the energy separation from the 'dip' to the next subband is related to the binding energy of the state. Furthermore, as the attractive 'hole' is made larger to admit more bound states, the 'dips' move lower in energy and more such 'dips' appear [4]. 'Level repulsion between Lyapunov exponents of the transfer matrix', emphasized in Ref. [1], seems unnecessary to understand this suppressed transmission.

In Fig. 1 we show the two terminal Landauer conductance [6] versus Fermi energy through twenty repulsive delta-function scatterers (solid line), randomly positioned in a 30 nm wide wire. The conductance was calculated by cascading scattering matrices [7] including the evanescent wire modes. There is no evidence of suppressed transmission near a subband minimum. We then made one of the twenty scatterers attractive by reversing its sign. The resulting conductance (dashed line) shows two pronounced 'dips' *below* the second and third subband minima, corresponding to the quasi-bound

states forming in the attractive scatterer.

The essential aspect for the manifestation of the conductance drops studied Refs. [1]-[5] is the formation of quasi-bound states in the disordered potential, not the global sign of the scattering potential. In our point scatterer model a locally attractive potential is necessarily also globally attractive, but this need not be the case. One can easily imagine a slow (adiabatic) rise and fall of the band edge in the conduction channel to form a large repulsive barrier. A small locally attractive hole can be placed on top of this potential barrier to admit one or more quasi-bound states. The same arguments used above would then apply to the *local* subbands formed around the locally attractive hole. For this case, the position of the 'dips' in conductance with respect to the *global* subbands depends on the height of the large repulsive barrier.

This work was supported by the U.S. Air Force Office of Scientific Research under Grant No. AFOSR-88-0304.

References

[1] I. Kander, Y. Imry, and U. Sivan, Phys. Rev. B **41**, 12941 (1990).

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

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

[4] P.F. Bagwell, Phys. Rev. B **41**, 10354 (1990). P.F. Bagwell, J. Phys: Condens. Matter., **2**, 6179 (1990).

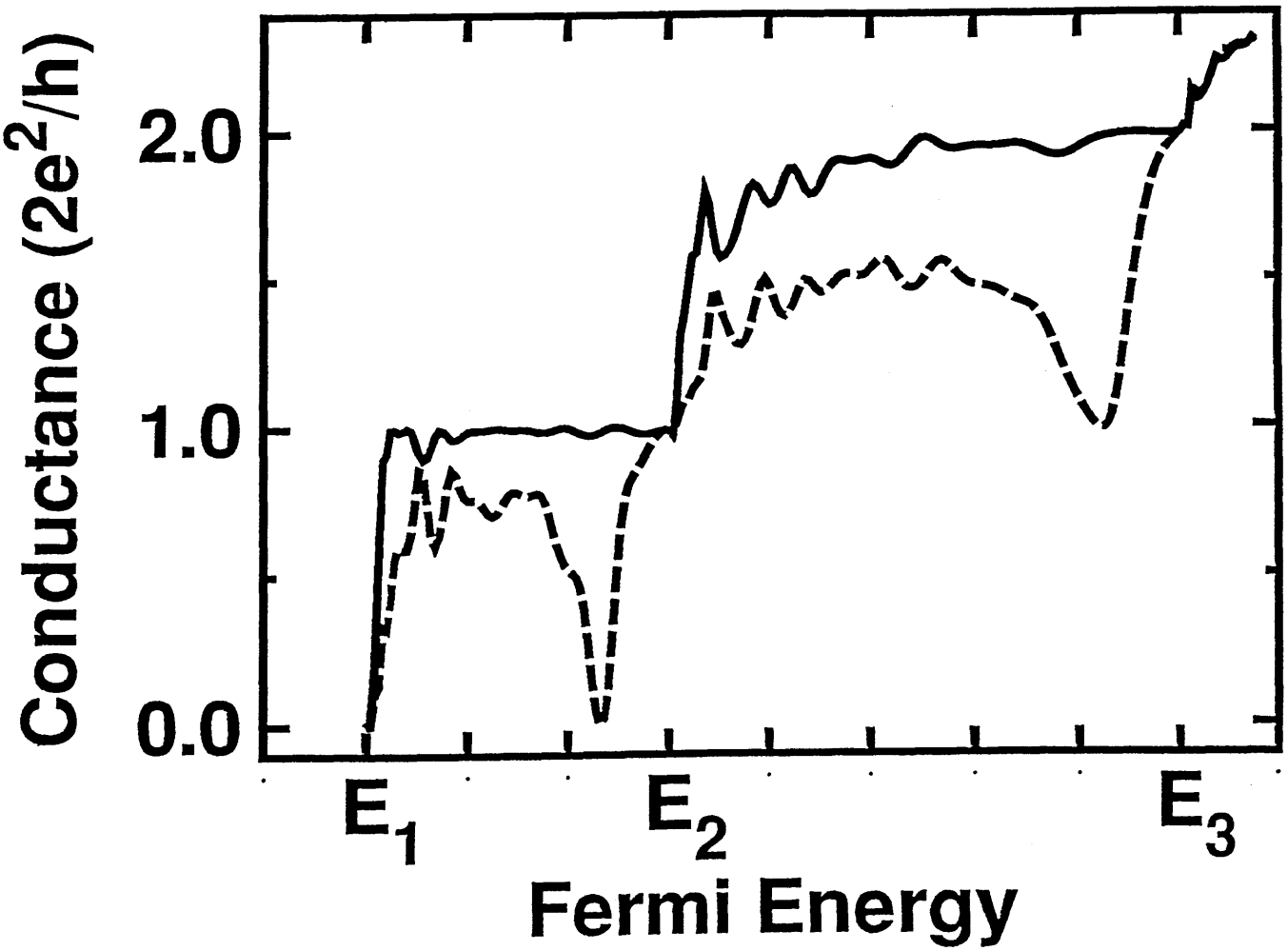
[5] E. Tekman and S. Ciraci, (unpublished).

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

[7] M. Cahay, M. McLennan, and S. Datta, Phys. Rev. B **37**, 10125 (1988).

Figure Caption

Fig. 1: Two terminal Landauer conductance versus Fermi energy through twenty delta function scatterers in a 30 nm wide wire. When all the scatterers are repulsive (solid line) the conductance rises at each new subband minimum. When one of the scatterers is made attractive (dashed line) by reversing its sign a pronounced 'dip' appears *before* the opening of each new quasi-one-dimensional channel.



Appendix L

Arvind Kumar and Philip F. Bagwell, 'Evolution of the Quantized Ballistic Conductance with Increasing Disorder in Narrow Wire Arrays', submitted for publication.

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

October 29, 1990

We study the Landauer conductance averaged over a parallel array of disordered narrow wires as the Fermi energy and length of the disordered region are varied. When quantum diffusion is the dominant electron transport mechanism, we find numerically that the ballistic conductance steps evolve into conductance drops after a 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 'quasi-donor levels' forming in the impurities.

PACS 73.20.Dx, 71.55.Jv, 72.20.Dp, 73.40.Gk

submitted to Physical Review Letters

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]-[16]. The essential conclusion of these studies is that the average conductance rises after the opening of each new subband channel, although structure in the conductance of a single wire may be obscured by wave interference fluctuations. If some of the scatterers are attractive, pronounced conductance drops before the opening of each new subband were also found to occur, due to the formation of 'quasi-donor levels' in the impurities [3]-[9]. 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 [17]-[20] in narrow quantum wires. Indeed, based on a modified Drude model, Refs. [21]-[28] 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 [17]-[19], but it is has been unclear what physics this structure might represent or where it occurs in relation to subband minima in the wires.

In this letter we calculate the Landauer conductance [29]-[33] of a parallel array of quantum wires [17]-[19] to obtain the 'ensemble-averaged' conductance. We show that there is a clear transition 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. Finally, we find conductance drops before the opening of a new subband channel only when attractive scatterers are present in the wires. We also show that the standard semiclassical 'Golden-Rule' or 'Born approximation' scattering theory is invalid near a subband minimum, so that the correct dependence of the conductance versus Fermi energy is not obtained in Refs. [21]-[28] even when the carrier diffusion is classical.

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

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

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

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

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

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

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

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

where T_{mn} denotes the transmission coefficient from mode n to mode m. The transmission coefficients 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.

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 5 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/AlGaAs heterostructures.

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 $f = 0.5$. 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 quasi-donor levels splitting off from the confinement subbands. The spacing $\triangle E$ from these quasi-donor levels to the next subband is of order $\triangle E \sim (m/\hbar^2)(\gamma/W)^2$, 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 which obscure any underlying subband structure.

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 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 quasi-bound state energies vary from wire to wire [6], 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 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), the quasi-donor 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 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 [34]-[35]. 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 $< G > \sim \exp(-L/\eta)$, 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 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 Fig. 2(a-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 wire [7], 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 [36].

A drop in conductance due to enhanced scattering after the opening of a subband channel has been argued previously [21]-[28], based on a Drude model in which the collision time is modified to account for scattering between quantum channels. Because wave interference between different scattering events is neglected in the Drude approximation, such a drop 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]-[5], the electron transmission through a single impurity increases when a new subband channel is opened. Furthermore, Refs. [21]-[28] use essentially a Golden-Rule (or first Born approximation) approach to calculating scattering rates, 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 in these calculations. Finally, the Born approximation of Refs. [21]-[28] breaks down at a subband minimum [4]. Even

though the neglect of wave interference between scattering events may be valid in some circumstances, the scattering at each impurity must be calculated properly, and this has not been done in Refs. [21]-[28]. In Refs. [3]-[9] pronounced conductance dips are shown below each subband minimum, as in Fig. 2(a), but we choose here to attribute these to formation of quasi-bound states in a local impurity potential energy minimum [4]-[7].

In conclusion, we find that the electrical conductance versus Fermi energy in a quasi-one-dimensional wire does 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. [21]-[28]. In the calculations of Refs. [21]-[28] the conductance decreases linearly with the length of the wire so the electrons are delocalized. In our calculation, we find that the electrical conductance in a quasi-one-dimensional wire decreases roughly exponentially with the length of the disordered region. However, the localization length of the decay is reduced when the Fermi energy crosses a confinement subband, leading to a fundamentally different subband structure in arrays of long quantum wires. Therefore, the new subband structure we find depends on quantum diffusion, not classical diffusion as in the Drude model of Refs. [21]-[28]. Similar results should follow from Anderson model calculations [8]-[9] if care is taken to exclude formation of quasi-donor levels described in Refs. [4]-[7]. If the potential energy is locally attractive, so that quasi-donor levels can form in an impurity, the new subband structure arising from localization phenomena studied in this letter is modified by the depressed transmission near the quasi-bound states.

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 AFOSR-88-0304. A.K. gratefully acknowledges support from a Semiconductor Research Corporation Fellowship.

References

[1] B.J. van Wees, H. van Houten, C.W.J. Beenakker, J.G. Williamson, L.P. Kouwenhoven, D. van der Marel, 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, G.A.C. Jones, J. Phys. C: Solid State, **21**, L209, (1988).

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

[4] P.F. Bagwell, Phys. Rev. B, **41**, 10354 (1990). P.F. Bagwell, J. Phys: Condens. Matter, **2**, 6179 (1990).

[5] E. Tekman and S. Ciraci, 'Ballistic Transport Through Quantum Point Contacts: Elastic Scattering by Impurities', (unpublished).

[6] A. Kumar and P.F. Bagwell, 'Resonant Tunneling in a Quasi-One-Dimensional Wire: Influence of Evanescent Modes', (unpublished). A. Kumar and P.F. Bagwell, Solid State Comm., **75**, 949 (1990).

[7] P.F. Bagwell and A. Kumar, 'Comment on "Effects of Channel Opening and Disorder on the Conductance of Narrow Wires" ', (unpublished).

[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**, 12941 (1990).

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

[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 Comm., **70**, 801 (1989).

[16] M. Cahay, M. McLennan, and S. Datta, Phys. Rev. B, **37**, 10125 (1988). S. Datta, M. Cahay, and M. McLennan, Phys. Rev. B, **36**, 5655 (1987).

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

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

[19] 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**, 12315 (1990).

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

[21] B.E. Sernelius, K.F. Berggren, M. Tomak, and C. McFadden, J. Phys. C: Solid State, **18**, 225 (1985).

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

[23] M.J. Kearney and P.N. Butcher, J. Phys. C: Solid State **20**, 47 (1987).

[24] N. Trivedi and N.W. Ashcroft, Phys. Rev. B, **38**, 12298 (1989).

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

[26] P. Vasilopoulos and F.M. Peeters, Phys. Rev. B **40**, 10079 (1989).

[27] T.P. Orlando, P.F. Bagwell, R.A. Ghanbari, and K. Ismail, in *Electronic Properties of Multilayers and Low-Dimensional Semiconductor Structures*, J.M. Chamberlain, L. Eaves, and J.C. Portal, eds., (Plenum, London, 1990).

[28] S.Y. Qiu and M.V. Jaric, Bull. APS, **35**, 492 (1990). W. Bao and Z. Tesanovic, Bull. APS, **35**, 492 (1990).

[29] Y. Imry, in *Directions in Condensed Matter Physics*, G. Grinstein and G. Mazenko, Editors, (World Scientific Press, Singapore, 1986).

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

[31] M. Büttiker, IBM J. Res. Develop., **32**, 317, (1988).

[32] R. Landauer, J. Phys.: Condens. Matter, **1**, 8099, (1989).

[33] R. Landauer, in *Analogies in Optics and Micro-Electronics*, W. van Haeringen and D. Lenstra eds., (Kluwer Academic Press, 1990).

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

[35] The exponential decrease of conductance with L at $E = E_2$ and $E = E_3$ we find surprising, since Ref. [6] 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.

[36] 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.

Figure 1: Landauer conductance versus Fermi energy for a single quasi-one-dimensional wire having (a) $L = 50$ nm and (b)-(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 'quasi-donor 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.

Figure 2: Landauer conducance 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. 'Quasi-donor' states are still observed for either short (a) or long (b, bottom curve) quantum wires when $f = 0.5$.

Figure 3: Average conductance versus 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), E_2 and E_3 (solid), and $1.1E_2$ and $1.1E_3$ (dot-dashed). The 'crossing over' of the conductance curves in (a) indicates a much shorter localization length when the Fermi energy moves into a new subband. 'Quasi-donor states' present in (b) depress the conductance below each subband, so there is no 'crossing over' of the conductance curves.

Figure 1(a)

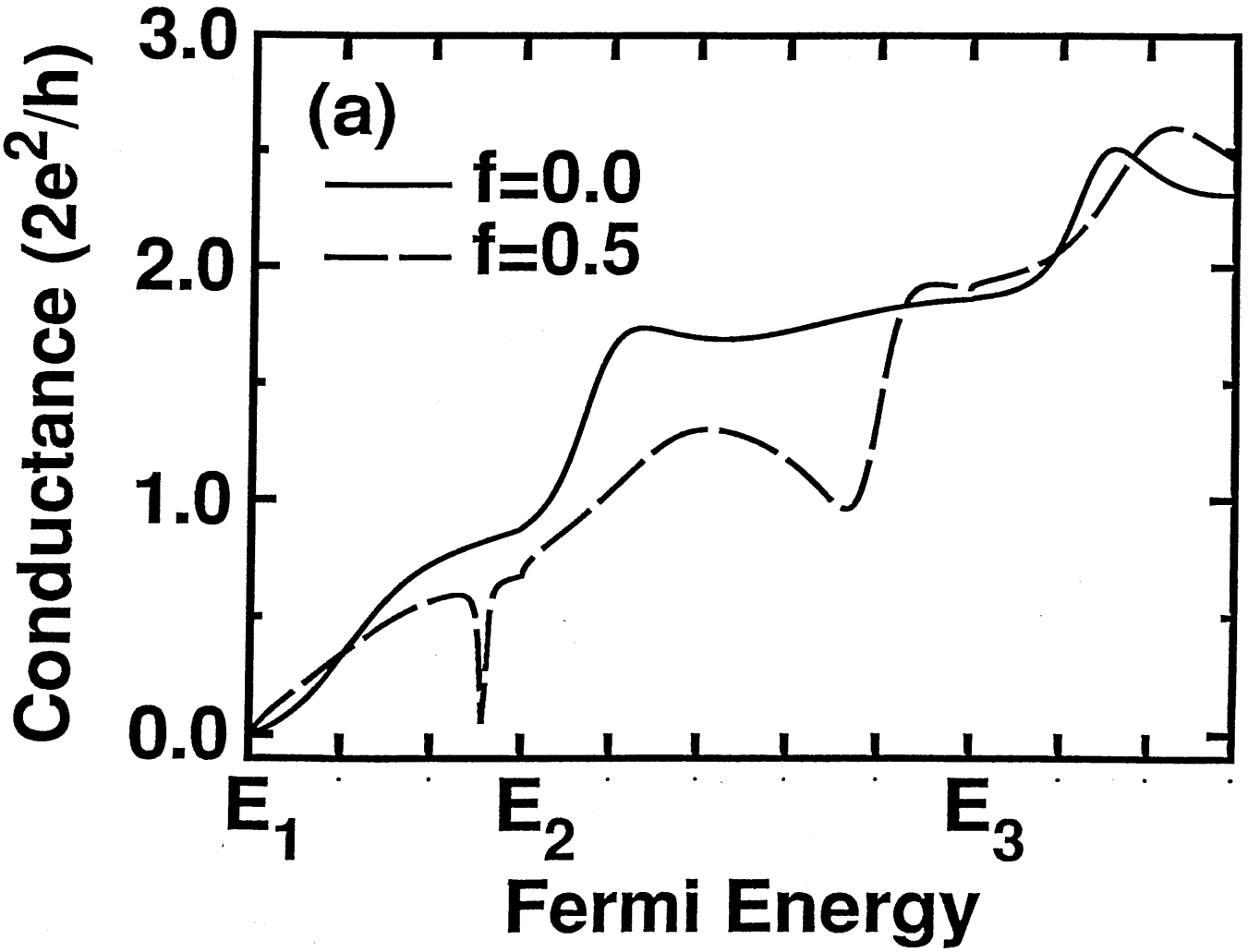


Figure 1(b)

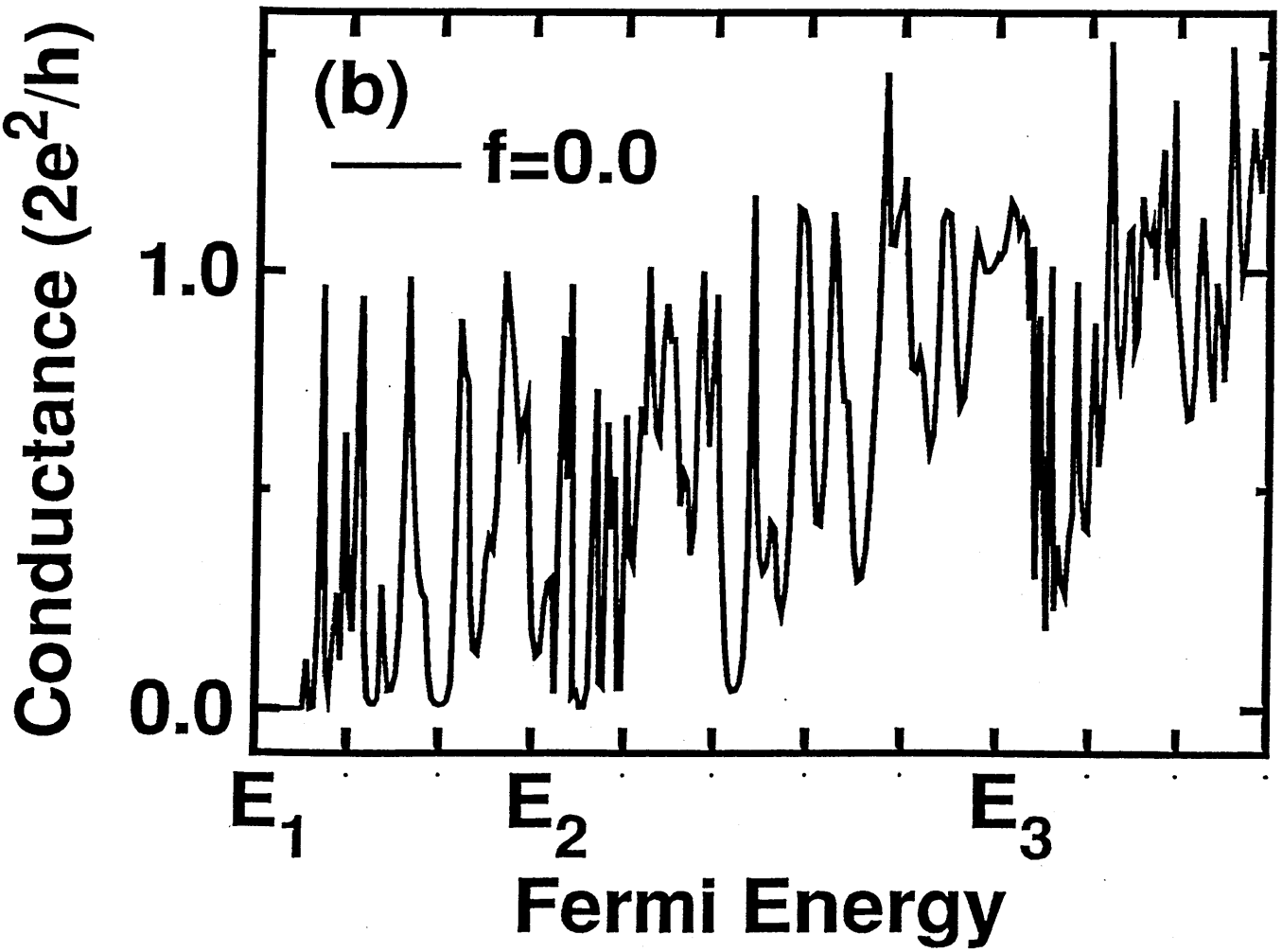


Figure 1(c)

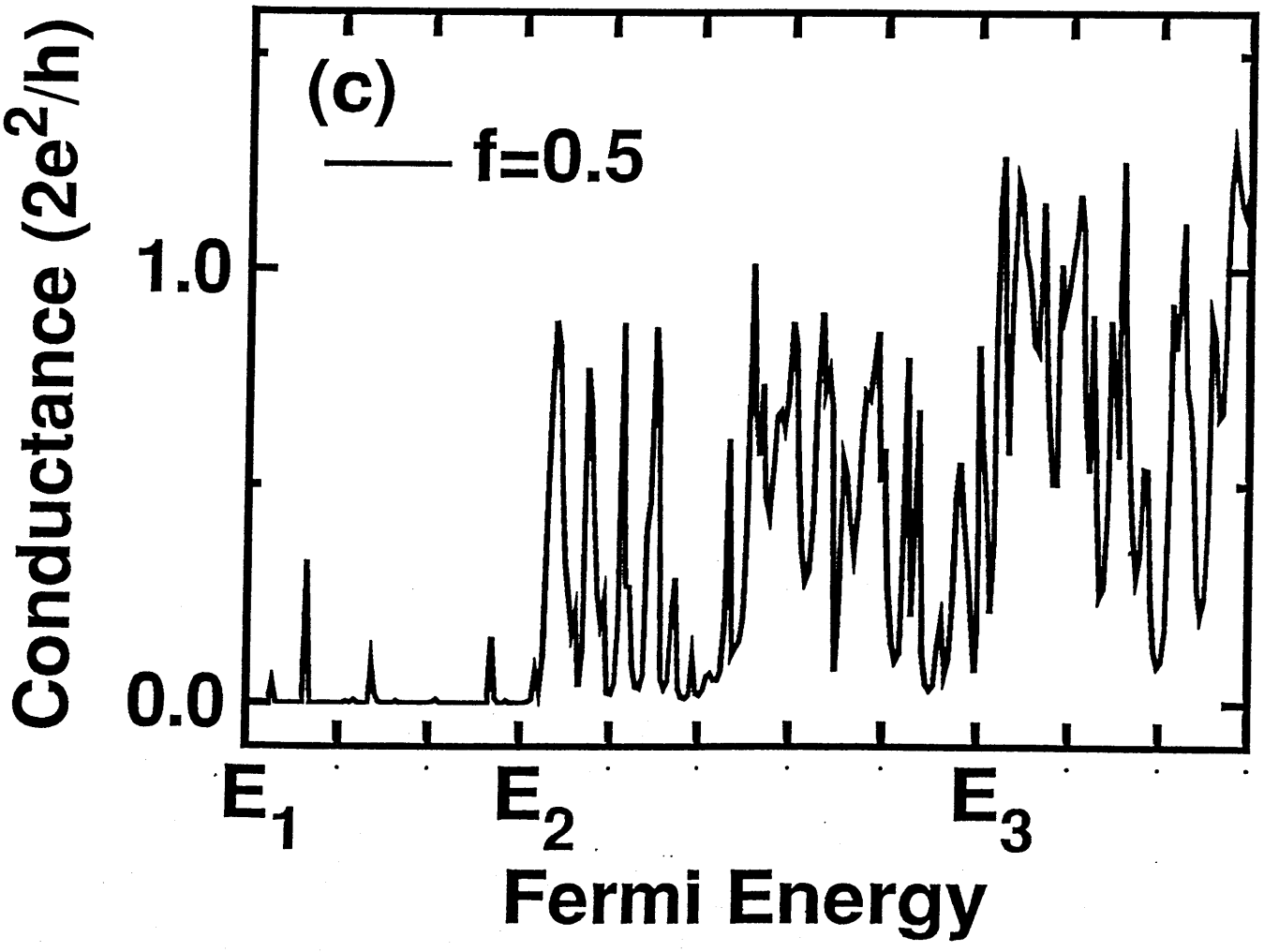


Figure 2(a)

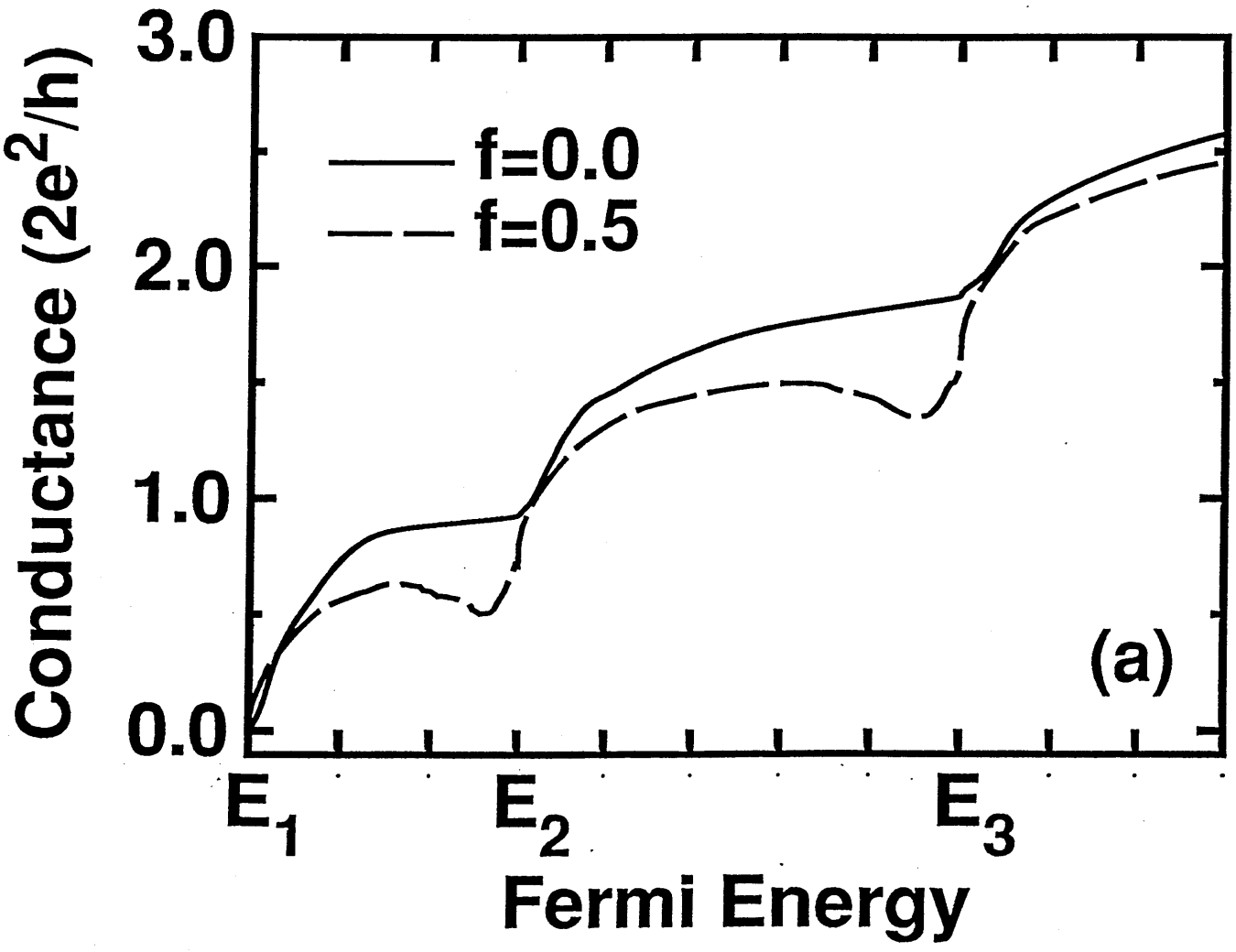


Figure 2(b)

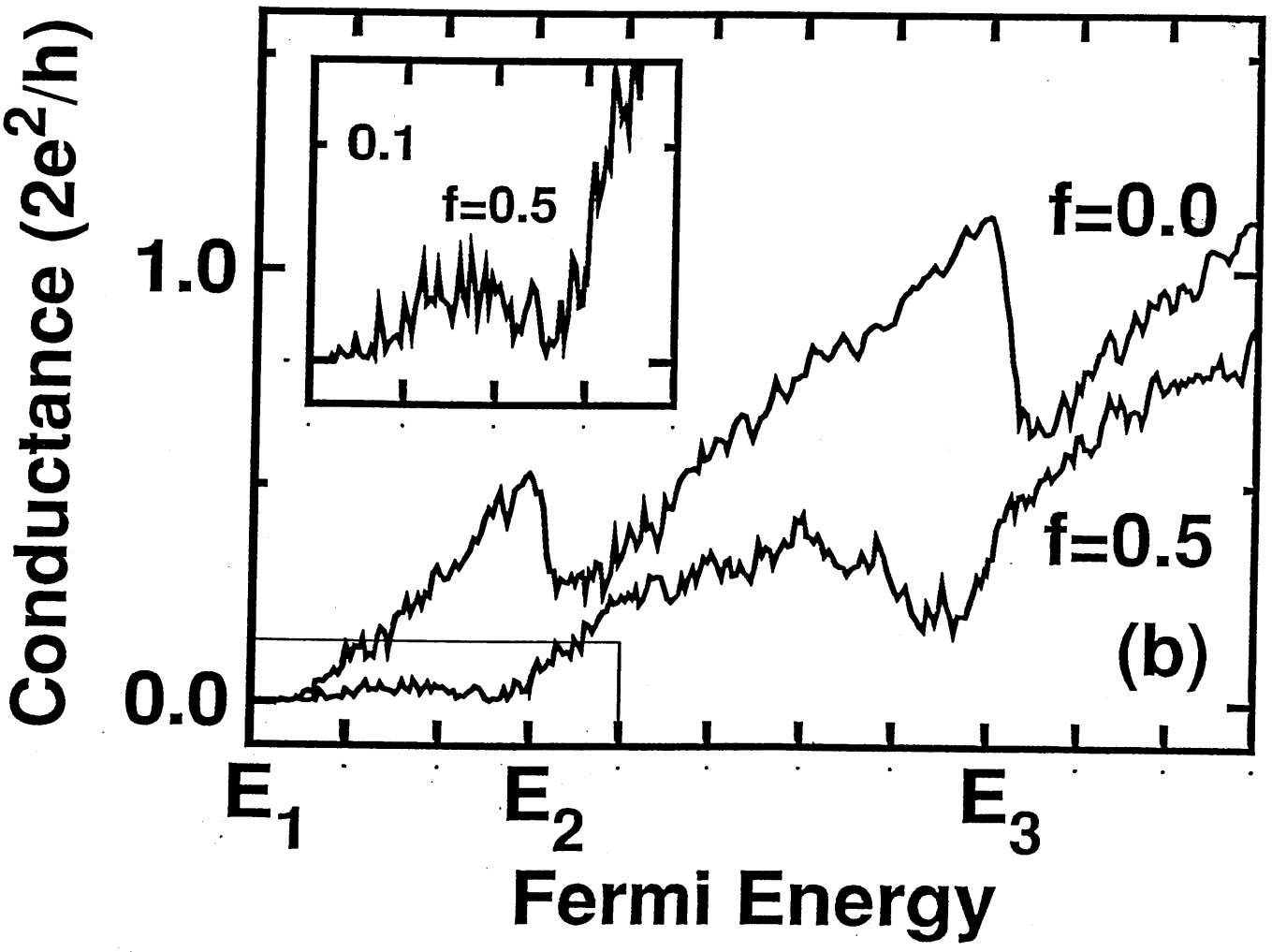


Figure 3(a)

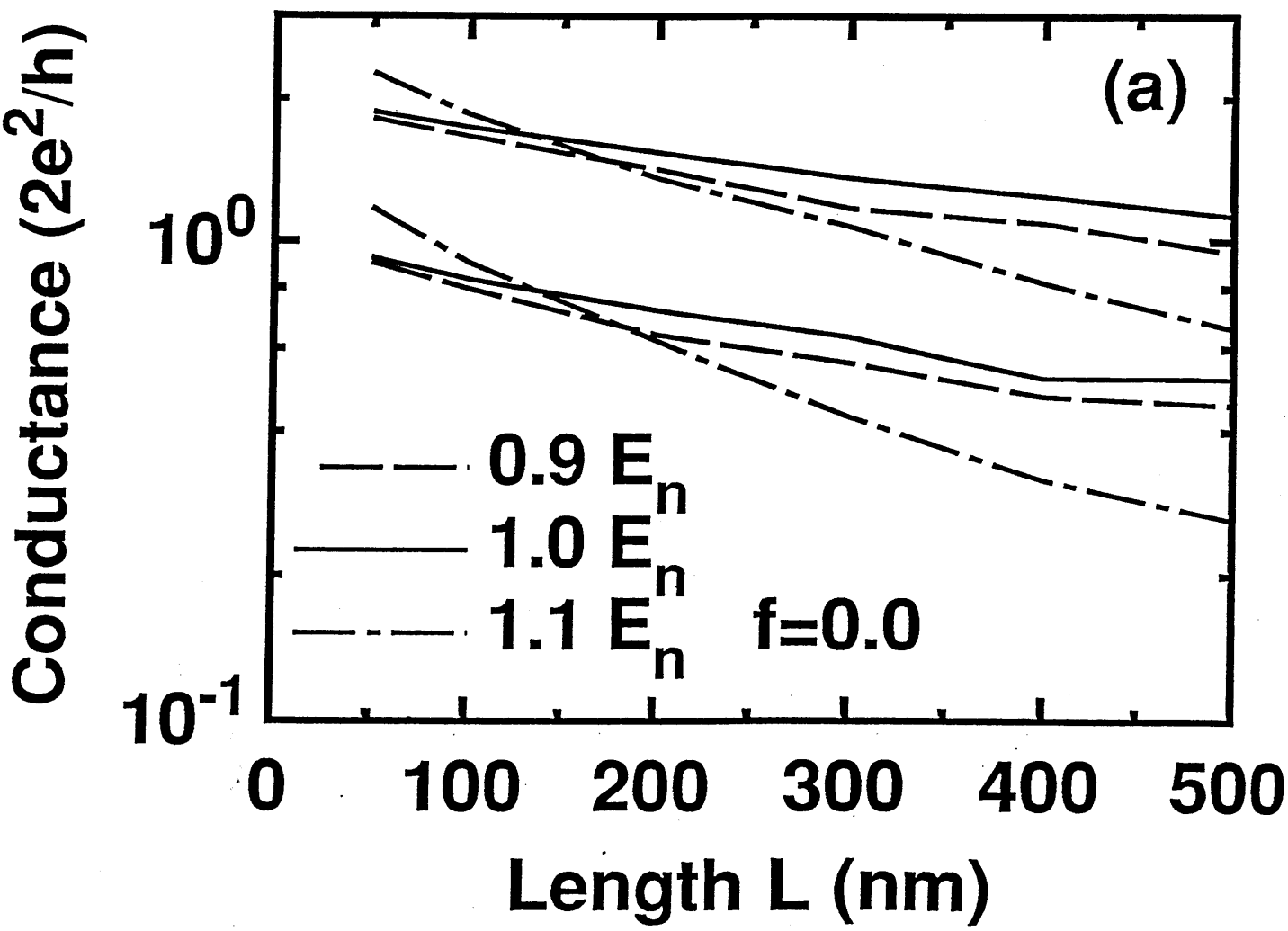
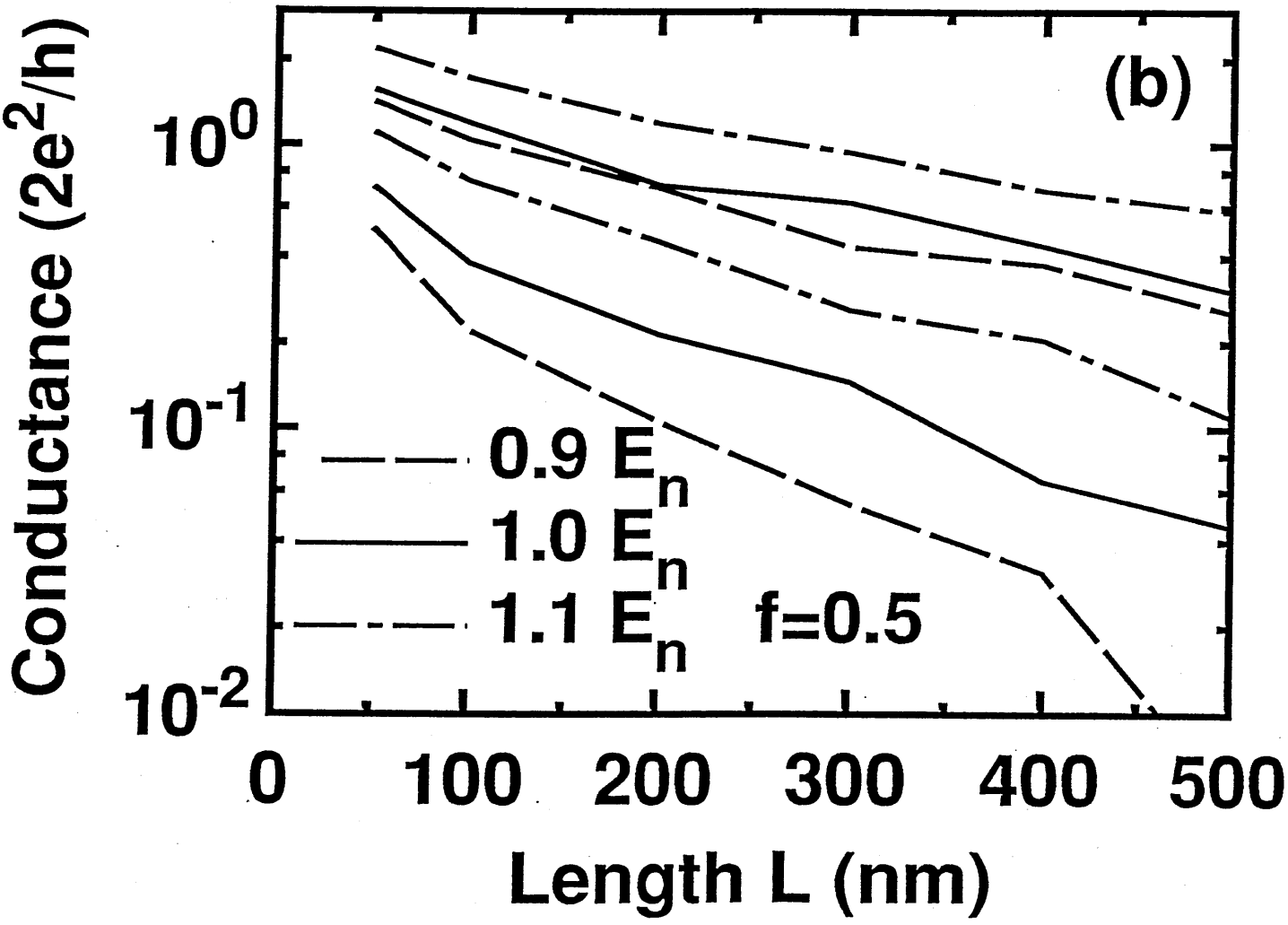

Conductance ($2e^2/h$)
10^0
10^{-1}
(a)
0.9 E_n
1.0 E_n
1.1 E_n
f=0.0
0
100
200
300
400
500
Length L (nm)

Figure 3(b)


(b)
Conductance ($2e^2/h$)
10^0
10^{-1}
10^{-2}
0.9 E_n
1.0 E_n
1.1 E_n
f=0.5
0
100
200
300
400
500
Length L (nm)

Appendix M

Philip F. Bagwell, Samuel L. Park, Anthony Yen, Dimitri A. Antoniadis, Henry I. Smith, Terry P. Orlando, and Marc A. Kastner, 'Magnetotransport in Multiple Narrow Silicon Inversion Channels Opened Electrostatically Into a Two-Dimensional Electron Gas', to be published.

Magnetotransport in Multiple Narrow Silicon Inversion Channels Opened Electrostatically Into a Two-Dimensional Electron Gas

Philip F. Bagwell, Samuel L. Park*, Anthony Yen,
Dimitri A. Antoniadis, Henry I. Smith, Terry P. Orlando, and Marc A. Kastner*
Department of Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

August 12, 1990

We study electrical transport in a dual gate Si metal-oxide-semiconductor field-effect transistor. The bottom gate is a grating which allows the inversion layer geometry to be controlled electrostatically. We compare the magnetoconductance of many parallel narrow inversion channels, a modulated potential, and a uniform two dimensional electron gas formed from the same background Si crystal. Electron weak localization becomes much more pronounced as the device is electrostatically pinched from a two dimensional inversion layer into many narrow wires in parallel, proving that the wire width can be reduced below the electron phase coherence length. For magnetic fields greater than 1 Tesla normal to the inversion layer there is a large drop in the two-terminal conductance of 90% or more, and which persists to room temperature, as electrons are added to the device so that it opens electrostatically from many narrow inversion layers in parallel into a two dimensional electron gas. This large negative transconductance results from electrostatically changing the dominant boundary condition on the classical Drude magnetoconductance tensor from that of a long and narrow to a short and wide conductor. Quantum edge states form at high magnetic fields, giving very different magnetoconductance for the parallel wires and wide electron gas. At a magnetic field of 30 Tesla the two-terminal conductance versus gate voltage of the narrow wires evolves into quantum Hall steps having a height of $2e^2/h$ multiplied by the number of wires in parallel. In contrast to a wide device the conduction band valley degeneracy is not resolved, giving rise to Hall steps of twice the expected size. The evolution from Shubnikov-de Haas oscillations to the quantum Hall effect qualitatively reproduces the 'anomalous magnetoresistance' of Kastner *et al.* [Phys. Rev. Lett., 60, 2535 (1988)].

* Department of Physics

to be published

1 Introduction

The confinement of electrons into narrow Si inversion channels [1]-[5] can be achieved by means of dual gates incorporated into a metal-oxide-semiconductor field-effect transistor (MOSFET). Varying the two gate voltages controls both the electron density and the width of the narrow channel. A variety of conductance fluctuation phenomena [6]-[9], weak localization [10], Shubnikov-de Haas oscillations and their evolution into the quantum Hall effect [11]-[13], and possible electron charging effects [4] have been demonstrated in narrow Si inversion layers. Although the mean free path in Si inversion layers is much shorter than in an electron gas formed at a GaAs/AlGaAs heterojunction, electrostatic control of the Si inversion layer geometry is much easier using a MOS field-effect capacitor. We exploit field-effect control of the Si inversion layer geometry to study in a single device the electrical conductance of an array of narrow wires, a modulated potential, and a uniform two dimensional electron gas as a function of electron density, magnetic field, and temperature.

Our device geometry is the dual gate MOSFET shown in Fig. 1. The bottom gate is a W grating gate having a 200nm period, and is separated from the inversion channel by a 20nm thick SiO_2 insulating layer. All the grating lines are electrically connected at the bottom gate contact pad. A futher 500nm of SiO_X separates the grating gate from a second continuous Al gate electrode. Standard MOSFET devices made with a continuous W gate on the same wafer had a mobility of 5000 cm^2/V-s at 4.2 K. The novel device fabrication technology is described elsewhere [3]-[5].

2 Device Characteristics

We measure the rms ac-drain current I_{DS} of the MOSFET device of Fig. 1 at low rms ac-drain voltage, $V_{DS} = 25\ \mu V$, using standard lock-in amplifier techniques. When the device is cooled to low temperature, T=50 mK, in a dilution refrigerator, the conductance $G=I_{DS}/V_{DS}$ versus gate voltage in Fig. 2 is obtained. In Fig. 2 the top gate voltage V_{TG} is swept from -16 V to 20 V holding the bottom gate (grating gate) voltage V_{BG} fixed at different values.

Fig. 2 shows that, for the family of curves where $V_{BG} = 2.5$ V, 3 V, 4 V, and 5 V there is a clear 'break' or 'kink' in the device I-V around $V_{TG} = 0$ V. For $V_{TG} \leq 0$ V on this family of curves the electron gas is confined only beneath the grating wires, while for $V_{TG} \geq 0$ V the electron gas exists everywhere in the channel. The 'break' in the device I-V curves as the entire channel becomes inverted can be understood in a simple model where conduction underneath the grating wires and in the gap between grating wires are viewed as two different MOSFETs in parallel having different threshold voltages. When the grating gate V_{BG} is held well above its threshold voltage of approximately 1 V, biasing the top gate negative can shut off the transistor current on this family of curves. This can only happen if electrostatic fringing fields around the bottom gate are significant, so that the field lines from the top gate can reach around the bottom gate fingers and turn off the electron gas.

For the family of bottom gate voltages $V_{BG} = 1$ V, 0 V, and -1 V in Fig. 2 there is

no 'kink' or 'break' in the I-V characteristics. For this family of curves the electron gas is confined only in the gap between grating lines. The entire MOSFET channel does not become inverted, as we subsequently prove by examining the device magnetoconductance. The small current which is independent of the top gate voltage on these three curves is a parasitic parallel current path around the side of the device which does not significantly affect the data we subsequently present. The current saturates in Fig. 2 due to the 330 Ω resistance of the wires running into the dilution refrigerator in series with the 50 Ω device contact resistance. This same series resistance is present in Figs. 2-5.

We now fix the top gate voltage above its threshold so that $V_{TG} = 13$ V, and gradually pinch the electron gas down into narrow inversion strips using $V_{BG} = 3$ V, 1.5 V, 0 V, and -1.25 V in Fig. 3. The magnetic field B is swept from -1 T to 8 T. As the electron gas is pinched into narrow channels the conductance decreases as expected. However, as the bottom gate voltage passes below its threshold from $V_{BG} = 1.5$ V to $V_{BG} = 0$ V, a large weak localization [14] magnetoresistance feature develops around B = 0. The development of this large weak localization feature proves that the wire width W is being made smaller than the electron's phase coherence length L_ϕ, and is further indication that narrow inversion channels are present. If the electron motion is diffusive, then in two dimensions an applied magnetic field extinguishes the weak localization when a magnetic flux quantum $\phi_0 = h/2e$ fits inside the average diffusion path to return to the origin. If this semiclassical diffusion path is constrained by a boundary, such as the side of a narrow wire, then the area of the semiclassical diffusion path is much smaller so that a higher magnetic field is needed to turn off the weak localization. The critical magnetic field B_c is approximately given by $B_c \simeq \phi_0/(L_\phi)^2$ in two dimensions, and increases to $B_c \simeq \phi_0/(W L_\phi)$ when the wire width is much smaller than the coherence length. The size of the conductance correction is also much larger in one dimension than in two dimensions. Using the semiclassical one-dimensional weak localization formula of Al'tshuler [15], after correcting for series resistance, we obtain an electron phase coherence length $L_\phi \simeq 0.6$ μm and a wire 'width' between 350-450 Å when $V_{BG} = 0$ V.

For the two dimensional electron gas in Fig. 3, formed when both gate voltages are large so that $V_{TG} = 13$ V and $V_{BG} = 3$ V, the two terminal device conductance decreases with magnetic field. This decreasing two-terminal current with increasing magnetic field is the same magnetoconductance as in a standard wide MOSFET device. As our grating gate device is pinched into long and narrow wires, such as when $V_{TG} = 13$ V and $V_{BG} = 0$ V in Fig. 3, there is almost no dependence of the background conductance on the magnetic field. All the curves display Shubnikov-de Haas oscillations, but the background magnetoconductance changes completely from a negative magnetoconductance in a wide device to almost no magnetoconductance in the narrow device. This is a third indication that narrow inversion channels are actually being formed. The device curve $V_{TG} = -7$ V and $V_{BG} = 3$ V in Fig. 3, when the electrons are now confined underneath the grating lines, shows a qualitatively similar magnetoconductance as the two previous curves where the electrons are confined in the gap between the grating lines, $V_{TG} = 13$ V and $V_{BG} = 0$ V and -1.25 V in Fig. 3.

The different magnetoconductance of a wide versus a narrow MOSFET is manifest in a quite dramatic way when the magnetic field is held fixed and the inversion layer geometry is varied with the two gates as in Fig. 4. We fix the bottom gate voltage at $V_{BG} = 3$ V

and vary the inversion layer geometry with the top gate. The 'break' or 'kink' in the static I-V curve around $V_{TG} \simeq 0$ V becomes quite prominent as the magnetic field is increased. For $V_{TG} \leq 0$ V the current is almost independent of magnetic field, consistent with the magnetoconductance of the array of narrow wires in Fig. 3. However, when $V_{TG} \geq 0$ V, the current drops very strongly with increases in the magnetic field, the drop becoming more pronounced as the magnetic field is further increased. This behavior is consistent with the magnetoconductance of a standard wide MOSFET device. The counterintuitive result in Fig. 4 is that, as more electrons are added to the device, the current actually decreases when a magnetic field is present. The decrease in the current is not small, a drop of more than two-thirds of the original current is observed at a magnetic field of 8 T in Fig. 4.

We argue that, in Fig. 4, the region of the curve where $V_{TG} \leq 0$ is where the electron gas is confined into narrow inversion channels. The decreasing conductance with increasing gate voltage occurs when the electron gas first becomes continuous across the channel and forms a modulated potential, and continues until a minimum conductance occurs at $V_{TG} \simeq 8$ V when B = 8 T. The conductance minimum occurs when the gate voltages permit an unmodulated two dimensional potential. The rising conductance when $V_{TG} \geq 8$ V occurs partially because more carriers are being added to the inversion layer, but mainly because the combination of gate voltages now again produces a modulated potential. [16]. On the rising part of the curve where $V_{TG} \geq 8$ V, the electrostatic potential minimum has shifted over by half a grating period to lie in the gap between grating wires. When $0 \text{ V} \leq V_{TG} \leq 8$ V on the falling part of the I-V curve, the electrostatic potential minimum was underneath the grating gate wires. A similar large negative transconductance can be observed by fixing the top gate voltage above its threshold and sweeping the bottom gate in an applied perpendicular magnetic field.

If our interpretation of Fig. 2 is correct, that the 'breaks' in the device I-V curve correspond to an opening of the narrow inversion channels into a wide two-dimensional electron gas, then there should be very little magnetoconductance of the family of bottom gate voltages below threshold, V_{BG} = 1 V, 0 V, and -1 V, where there is no 'break' in the static I-V. This is indeed the case as shown in Fig. 5, confirming that for this family of I-V curves the electron gas exists only in narrow strips between the grating gate wires.

3 Temperature Dependence

To understand if the large negative transconductance in Fig. 4 is a classical or quantum mechanical effect, we study its temperature dependence. Fig. 6 shows the current on a different device when V_{BG} = 3.5 V. We sweep the top gate voltage for temperatures T = 4.2 K, 20 K, 50 K, 100 K, 200 K, and 280 K. The total series resistance is 120 Ω in Fig. 6, since different measurement equipment is now being used. Although fringing fields are able to shut off the device when T = 4.2 K, we see in Fig. 6(a) that by the time the device is at room temperature the fringing fields are much less effective. We do not understand this different in detail. The decrease in conductance with increasing temperature is presumably due to more electron-phonon scattering at higher temperatures.

We apply a 15 Tesla magnetic field normal to the Si-SiO_2 interface in Fig. 6(b). The

T = 4.2 K curve when B = 0 is also shown for reference. At T = 4.2 K and B = 15 T, a single Shubnikov-de Haas oscillation is observable over the range of gate voltages shown. By T = 20 K the Shubnikov-de Haas oscillation is no longer present, indicating that the Landau level structure has completely detereorated. Yet the large negative transconductance seen in Fig. 6(b) persists even to room temperature. We conclude that the effect is completely classical.

Could this classical magnetoconductance effect arise from electrons being classically confined to a wire width smaller than the cyclotron orbit or a mean free path? In this sample a generous estimate of the mean free path which enters the Drude conductance is roughly 600 Å when T = 4.2 K. Thus, the mean free path can be smaller than the wire width at low temperature, but this does not seem possible at room temperature, when the mean free path is 5-10 times smaller. Also, at a magnetic field of B = 15 T, the cyclotron radius in (100) Si is less than 100 Å, yet the negative transconductance becomes larger, not smaller, as the magnetic field is increased.

The correct explanation for the behavior of the magnetoconductance as the device geometry changes was provided by Park [12]. Assume the magnetoconductance is described by the Drude conductance tensor. This assumption is appropriate if the device is not in the quantum Hall limit so that local resistances can still be defined. One must obtain the two terminal device conductance using the Drude tensor, by calculating the current consistent with the boundary conditions on the current and voltage. The boundary conditions are (1) that no current can flow through the side of the device, and (2) that the voltage at the source terminal be zero and at the drain terminal be V_{DS}, since the device contacts are heavily doped and can thus be approximated as an equipotential surface.

In a long and narrow geometry where L $\gg$ W, and if one is well away from the device ends, the current distribution from one differential slice of the device to the next must be the same as shown in Fig. 7(a). Therefore, the current mostly flows parallel to the edge of the device and the electric field is pointed away from the current path by the Hall angle. Since the current distribution is known, one can integrate

$$E_x = \rho_{xx} J_x + \rho_{xy} J_y \tag{1}$$

from source to drain. The term involving ρ_{xy} is completely negligible when L $\gg$ W and $\rho_{xx} \neq 0$, *i.e.* the device is not in the quantum Hall limit, so one obtains the two terminal resistance of the device as

$$R_{2P} \simeq L\rho_{xx}/W \quad . \tag{2}$$

This result makes sense, as it is just the standard argument for measuring ρ_{xx} in a Hall bar geometry. To make an accurate measurement of ρ_{xx} in a Hall bar geometry one wishes to make a sample very long and narrow, and keep the Hall probes well away from the ends of the device. Equation (2) for the two terminal resistance has been previously suggested by Syphers and Stiles [17].

Conversely, in the short and wide geometry of Fig. 7(b) where W $\gg$ L, the edges of the device are far away so that the current will have to adjust to the boundary condition on the voltage. For the short and wide geometry, the equipotential surfaces must be the same as one moves laterally across the channel as in Fig. 7(b). Therefore, the electric field points directly from source to drain and the current flows at the Hall angle across the channel.

There are minor corrections to this picture at the edges of the device, which will be small if W $\gg$ L. Consequently, since the electric field is known, the two terminal current can be obtained by integrating

$$J_x = \sigma_{xx} E_x + \sigma_{xy} E_y \tag{3}$$

from source to drain. One obtains in the limit W $\gg$ L and $\sigma_{xx} \neq 0$, where the second term in Eq. (3) is negligible, that

$$G_{2P} \simeq W\sigma_{xx}/L \quad . \tag{4}$$

Ref. [18] has suggested a similar expression for the two terminal magnetoconductance of a short and wide conductor.

The classical Eq. (2) and Eq. (4) qualitatively explain the data in Fig. 4. Since ρ_{xx} is independent of magnetic field, Eq. (2) predicts that in a long and narrow device where L $\gg$ W there is no magnetoconductance. Conversely, since σ_{xx} depends on magnetic field like $1/[1 + (\mu B)^2]$, where μ is the electron mobility and B the magnetic field, the two terminal conductance of a short and wide device where W $\gg$ L decreases as the magnetic field increases. To completely explain the data in Fig. 4, such as the minima occurring at different values of V_{TG} at different values of the magnetic field, one needs a theory of classical magnetoconductance in a modulated potential which is not available at this time. However Eq. (2) and (4) qualitatively explain most of the data.

How Eq. (2) and (4) relate to the more complicated expressions [19]-[20] of two terminal magnetoconductance in terms of the conductance tensor elements is not clear. One would expect to obtain Eq. (2) and (4) from Refs. [19] and [20] in appropriate limits. Park [12] has constructed a proof using the conformal mapping technique of Ref. [19] that there is a symmetry between G_{2P} in the W $\gg$ L limit and R_{2P} in the L $\gg$ W limit. As a special case, this symmetry implies that, if Eq. (2) holds in the limit of L $\gg$ W, then Eq. (4) holds when W $\gg$ L.

4 High Magnetic Fields

We turn to the behavior of the conductance at large magnetic fields. Fig. 8 shows the conductance (corrected for series resistance) when V_{BG} = 3.5 V and the top gate is swept for a family of magnetic fields B = 0 T, 5 T, 10 T, 15 T, 20 T, and 23 T. When the device aspect ratio is short and wide, the conductance continues to decrease with the large applied magnetic field. When the electrons are confined to a long and narrow conductor, the current remains roughly constant up to a magnetic field of 15 T. However, for the magnetic field values of 20 T and 23 T, there is a large increase in the current as B increases in the high aspect ratio device. Thus, the magnetoconductance is moving in different directions on different parts of the same device curve in Fig. 8.

We claim that this rise in conduction is due to edge states forming in the wire array as outlined by Büttiker [22]. Our device is 60 μm wide and the grating period is 0.2 μm, so that roughly 300 wires are present in the array. In a high magnetic field each wire must have a conductance of $2e^2/h$, the same conductance as that in a quantum ballistic conductor [21]. Since the wires have many scattering centers, their initial conductance is much less than the ballistic value. Therefore the conductance must rise with magnetic

field when edge states begin to be formed in the wire, giving a conductance of $2e^2/h$ times the number of wires. The conductance of the wide MOSFET will also approach $2e^2/h$ at large magnetic fields. Since many conduction channels are initially occupied in the wide MOSFET, its conductance must continue to fall at large magnetic fields to approach the quantum Hall conductance limit of $2e^2/h$. Therefore, the magnetoconductance must move in opposite directions at high magnetic fields in the wire array, depending on whether conduction occurs in many narrow wires or in a two-dimensional MOSFET.

To prove that edge state conduction is occurring, we measured the conductance of electrons confined in the gap between grating lines where $V_{BG} = 0$ V and B = 0 T, 10 T, 20 T, and 30 T in Fig. 9 (also corrected for series resistance). For B = 0 T and 10 T, the current versus top gate voltage qualitatively resembles Fig. 5, where Shubnikov-de Haas oscillations appear in the conductance and the background magnetoconductance is flat. When B = 30 T a well defined quantum Hall plateau develops at $1200e^2/h$, with weaker precursors to quantum Hall plateaus appearing at $600e^2/h$ and $1800e^2/h$. Quantum Hall plateau's were first reported in the two terminal conductance by Fang and Stiles [23], and fit nicely with the theory developed by Büttiker [22]. Since the device has 300 parallel wires, we interpret the Hall plateau at $1200e^2/h$ as the first filled Landau level with spin and valley degeneracy, due to the two conduction band valleys at a (100) Si surface. Note, however, that only one of the degeneracies is resolved in the conductance. Following Refs. [11], [12], [24], and [25], we interpret the valley degeneracy to be restored in the wire.

The magnetic field value B = 20 T is an intermediate case between the Shubnikov-de Haas oscillations and the quantum Hall effect. This curve reproduces roughly and qualitatively the basic phenomenology observed by Kastner *et al.* [11] in narrow MOSFET's, the 'anomalous magnetoresistance'. The anomalous filling factors in Ref. [11] have since been explained [12] in a model which takes into account how edge states form in the various thick and thin oxide regions underneath the gate of the samples in Ref. [11]. In our present device the narrow wire runs straight between the three dimensional degenerately doped source and drain, and not through any two dimensional regions or under different gate oxide thicknesses.

5 Conclusion

We originally undertook this study in an effort to see some manifestation of quasi-one dimensional subbands in the device conductance [1]-[2], [26] in a diffusive sample. This was not observed in these particular devices. However, a variety of magnetoresistance effects were studied. Weak localization was found to become a much stronger effect in the conductance, and persist to a much higher value of magnetic field, when the electron gas is pinched into a wire narrower than the phase coherence length. The grating gate enabled us to compare the magnetoresistance of a two dimensional sample with a one dimensional conductor, since the conductance of both systems are the same order of magnitude and can be observed on the same graph.

We also observed that the two terminal magnetoconductance of an inversion layer is strongly determined by the device aspect ratio (when a local magnetoconductance tensor

description of conduction is appropriate). This is in contrast to the two terminal resistance of a quantum Hall conductor, in which the resistance is independent of the size and shape of the conductor. A long and narrow conductor was found to have very little dependence of the resistance on magnetic field, while a short and wide conductor has a two terminal resistance which increases with magnetic field. This is manifested as a large negative transconductance of our grating gate MOSFET when the magnetic field is held fixed and the device geometry is varied.

Finally, for conduction in very high magnetic fields, the magnetoconductance was shown to move in opposite directions for the device with many parallel wires and a two dimensional MOSFET. We interpret this result in terms of magnetic quantum edge state formation in the wire array, as confirmed by the appearance of quantum Hall plateaus at the highest magnetic fields. The quantum Hall steps have size $2e^2/h$ since the conduction band valley degeneracy is not resolved.

6 Acknowledgements

We thank John-Scott Thomas, Udi Meirav, Norris Preyer, Seiji Horiguchi, and Akira Toriumi for experimental help. The high magnetic field measurements were carried out at the M.I.T. Francis Bitter National Magnet Laboratory, which is supported by the National Science Foundation. We thank Larry Rubin, Bruce Brandt, and Jing Luo for assistance with the high magnetic field measurements. James Carter, Mark Schattenburg, the staff at the M.I.T. Submicron Structures Laboratories and Microsystems Technology Laboratories greatly assisted in device fabrication. This work was sponsored by the U.S. Air Force Office of Scientific Research under Grant No. AFOSR-88-0304 and the U.S. Joint Services Electronics Program Contract No. DAAL03-89-C-0001.

References

[1] A.C. Warren, D.A. Antoniadis, and H.I. Smith, 'Quasi-One-Dimensional Conduction in Multiple, Parallel, Inversion Lines', Phys. Rev. Lett., **56**, 1858 (1986).

[2] 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, 'One-Dimensional Subband Effects in the Conductance of Multiple Quantum Wires in Si Metal-Oxide-Semiconductor Field-Effect Transistors', Phys. Rev. B., **37**, 6181 (1990).

[3] J.H.F. Scott-Thomas, M.A. Kastner, D.A. Antoniadis, H.I. Smith, and S.B. Field, 'Si Metal-Oxide Semiconductor Field-Effect Transistor with 70nm Slotted Gates for Study of Quasi-One-Dimensional Quantum Transport', J. Vac. Sci. Technol. B, **6**, 1841 (1988).

[4] 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).

[5] A. Yen, 'Grating Gate Si MOSFET for Study of Quantum Transport Effects', S.M. Thesis, Massachusetts Institute of Technology, 1987.

[6] A.B. Fowler, A. Hartstein, and R.A. Webb, 'Conductance in Resitricted Dimensionality Accumulation Layers', Phys. Rev. Lett., **48**, 196 (1982).

[7] M. Pepper and M.J. Uren, 'The Wigner Glass and Conductance Oscillations in Silicon Inversion Layers', J. Phys. C., **15**, L617 (1982).

[8] W.J. Skocpol, P.M. Mankiewich, R.E. Howard, L.D. Jackel, and D.M. Tennant, 'Universal Conductance Fluctuations in Silicon Inversion Layer Nanostructures', Phys. Rev. Lett., **56**, 2865 (1986).

[9] M.A. Kastner, R.F. Kwasnick, J.C. Licini, and D.J. Bishop, 'Conductance Fluctuations Near The Localized to Extended Transition in Narrow Si Metal-Oxide-Semiconductor Field-Effect Transistors', Phys. Rev. B, **36**, 8015 (1987).

[10] C.C. Dean and M. Pepper, 'One-Dimensional Localization Beyond First Order in Narrow Silicon MOSFETs', in *Proceedings of the 17th International Conference on the Physics of Semiconductors*, D.J. Chadhi and W.A. Harrison, eds., (Springer-Verlag, Berlin, 1984), p.425.

[11] M.A. Kastner, S.B. Field, J.C. Licini, and S.L. Park, 'Anomalous Magnetoresistance of the Electron Gas in a Restricted Geometry', Phys. Rev. Lett., **60**, 2535 (1988).

[12] S.L. Park, 'The Anomalous Magnetoresistance of the Electron Gas in a Restricted Geometry', Ph.D. Thesis, Massachusetts Institute of Technology, 1990.

[13] C.C. Dean and M. Pepper, 'Quasi-One-Dimensional Transport in Narrow Silicon Accumulation Layers', in *Localization, Interaction, and Transport Phenomena*, eds. B. Kramer, G. Bergmann, and Y. Bruynseraede, (Springer-Verlag, Berlin, 1985).

[14] G. Bergman, "Weak Localization in Thin Films", Physics Reports, **107**, 1 (1984)

[15] B.L. Al'tshuler and A.G. Aronov, 'Magnetoresistance of Thin Films and Wires in a Longitudinal Magnetic Field', Sov. Phys. JETP Lett., **33**, 499 (1981).

[16] This point must be considered tentative, as an explicit solution for classical transport in a modulated potential which properly incorporates the device boundaries is not available.

[17] D.A. Syphers and P.J. Stiles, 'Contiguous Two-Dimensional Regions in the Quantized Hall Effect', Phys. Rev. B., **32**, 6620 (1985).

[18] J.P. Harrang, R.J. Higgins, R.K. Goodall, P.R. Jay, M. Laviron, and P. Delescluse, 'Quantum and Classical Mobility Determination of the Dominant Scattering Mechanism in the Two-Dimensional Electron Gas of an AlGaAs/GaAs Heterojunction', Phys. Rev. B., **32**, 8126 (1985).

[19] R.W. Rendell and S.M. Girvin, 'Hall Voltage Dependence on Inversion Layer Geometry in the Quantum Hall Regime', Phys. Rev. B., **23**, 6610 (1981).

[20] G.L.J.A. Rikken, J.A.M.M. van Haaren, W. van der Wel, A.P. van Gelder, H. van Kempen, P. Wyder, J.P. Andre, K. Ploog, and G. Weimann, 'Two-Terminal Resistance of Quantum Hall Devices', Phys. Rev. B., **37**, 6181 (1988).

[21] B.J. van Wees, H. van Houten, C.W.J. Beenakker, J.G. Williamson, L.P. Kouwenhoven, D. van der Marel, C.T. Foxon, 'Quantized Conductance of Point Contacts in a Two Dimensional Electron Gas', Physical Review Letters, **60**, 848, (1988). 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, G.A.C. Jones, 'One Dimensional Transport and the Quantization of the Ballistic Resistance', Journal of Physics C: Solid State Physics, **21**, L209, (1988).

[22] M. Büttiker, 'Absence of Backscattering in the Quantum Hall Effect in Multiprobe Conductors', Phys. Rev. B, **38**, 9375 (1988).

[23] F.F. Fang and P.J. Stiles, 'Quantized Magnetoresistance in Two-Dimensional Electron System', Phys. Rev. B., **27**, 6487 (1983).

[24] A.B. Fowler, A. Hartstein, and R.A. Webb, '', Physica **117B & 118B**, 661 (1983).

[25] J.M. Kinaret and P.A. Lee, 'Exchange Interaction in a Quantum Wire in a Strong Magnetic Field', preprint.

[26] K. Ismail, D.A. Antoniadis, and H.I. Smith, 'One-Dimensional Subbands and Mobility Modulation in GaAs/AlGaAs Quantum Wires', Appl. Phys. Lett., **54**, 1130 (1989).

Figure 1: 'Grating Gate' MOSFET geometry. The split gates permit control of the inversion layer geometry via the field effect. A narrow inversion layer can form underneath the grating lines, in the gap between grating lines, or a wide inversion layer having a weakly modulated electron density can also form depending on the applied gate voltages.

Figure 2: Static device I-V curve for zero applied magnetic field. The five curves on the left show a pronounced 'kink' near zero volts, indicating a transition where inversion layer opens from an array of narrow channels underneath the grating wires to a two-dimensional electron gas. The three rightmost curves, where no 'kink' is observed, indicates the electron gas exists only in the gap between grating lines over the range of voltages shown.

Figure 3: A pronounced weak localization minimum developes as the two-dimensional inversion layer is pinched into narrow wires, indicating that the electron phase coherence length has crossed over the wire width. A corresponding change in background magnetoconductance also occurs, so that the magnetoconductance of the wide electron gas is large but the magnetoconductance of the array of narrow wires is small.

Figure 4: A large negative transconductance appears in a perpendicular magnetic field as the inversion layer geometry opens from the narrow wire array into a two dimensional electron gas. The additional carriers added to the device cannot overcome the decrease in two-terminal current required by the new boundary conditions on the Drude conductance tensor.

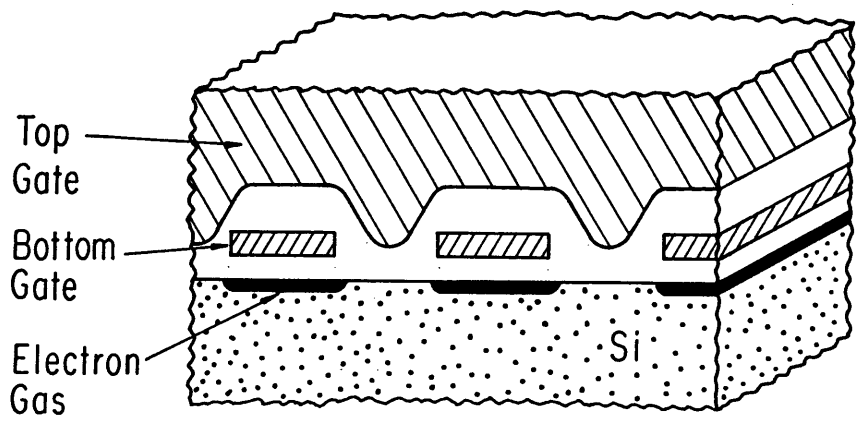
Figure 5: When there is no change in the inversion layer geometry with gate voltage, for the device curves without a 'kink', little magnetoconductance is observed.

Figure 6: Temperature dependence of the device current shown for T = 4.2 K, 20 K, 50 K, 100 K, 200 K, and 280 K. The magnetic field is (a) B = 0 T and (b) B = 15 T normal to the interface. The large negative transconductance persists to room temperature, confirming its classical origin.

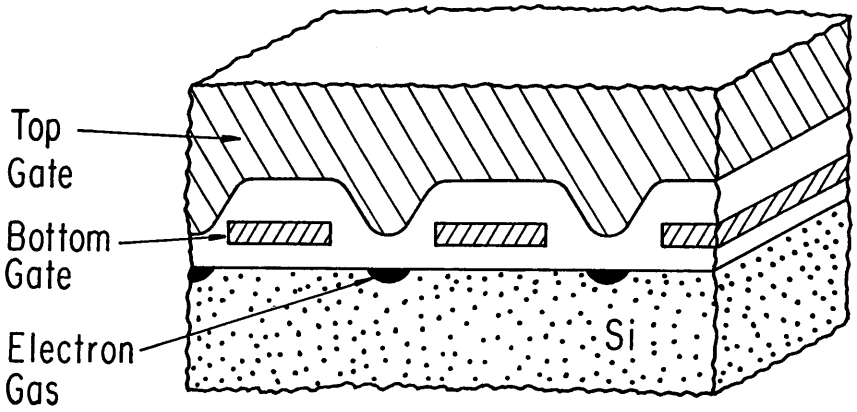
Figure 7: To obtain the two-terminal conductance from the Drude conductance tensor, we apply boundary conditions on the current and electric field appropriate to the device geometry. For (a) a long and narrow conductor, the current must continue to flow parallel to the side of the device even at high magnetic fields. If the conductor is (b) short and wide, current flows at the Hall angle across the device.

Figure 8: Conduction in a large magnetic field having values 0, 5, 10, 15, 20, and 23 Tesla. The bottom gate is held at 3.5 V. For the long and narrow conductors below 0 V in the figure, there is a large increase in the current at the two highest values of the magnetic field indicating the formation of quantum edge states and the approach to the quantum Hall effect.

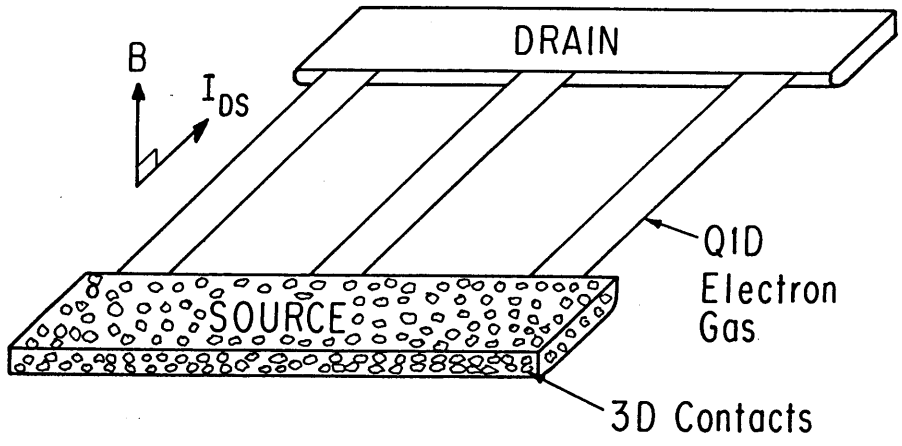
Figure 9: Two-terminal conductance of a narrow wire in applied magnetic fields of 0, 10, 20, and 30 Tesla. Shubnikov-de Haas oscillations are seen when B = 10 T, and quantum Hall plateaus appear when B = 30 T.


Top
Gate
Bottom
Gate
Electron
Gas
Si

(a)


Top
Gate
Bottom
Gate
Electron
Gas
Si

(b)


B
I_{DS}
DRAIN
Q1D
Electron
Gas
SOURCE
3D Contacts

(c)

Figure 1

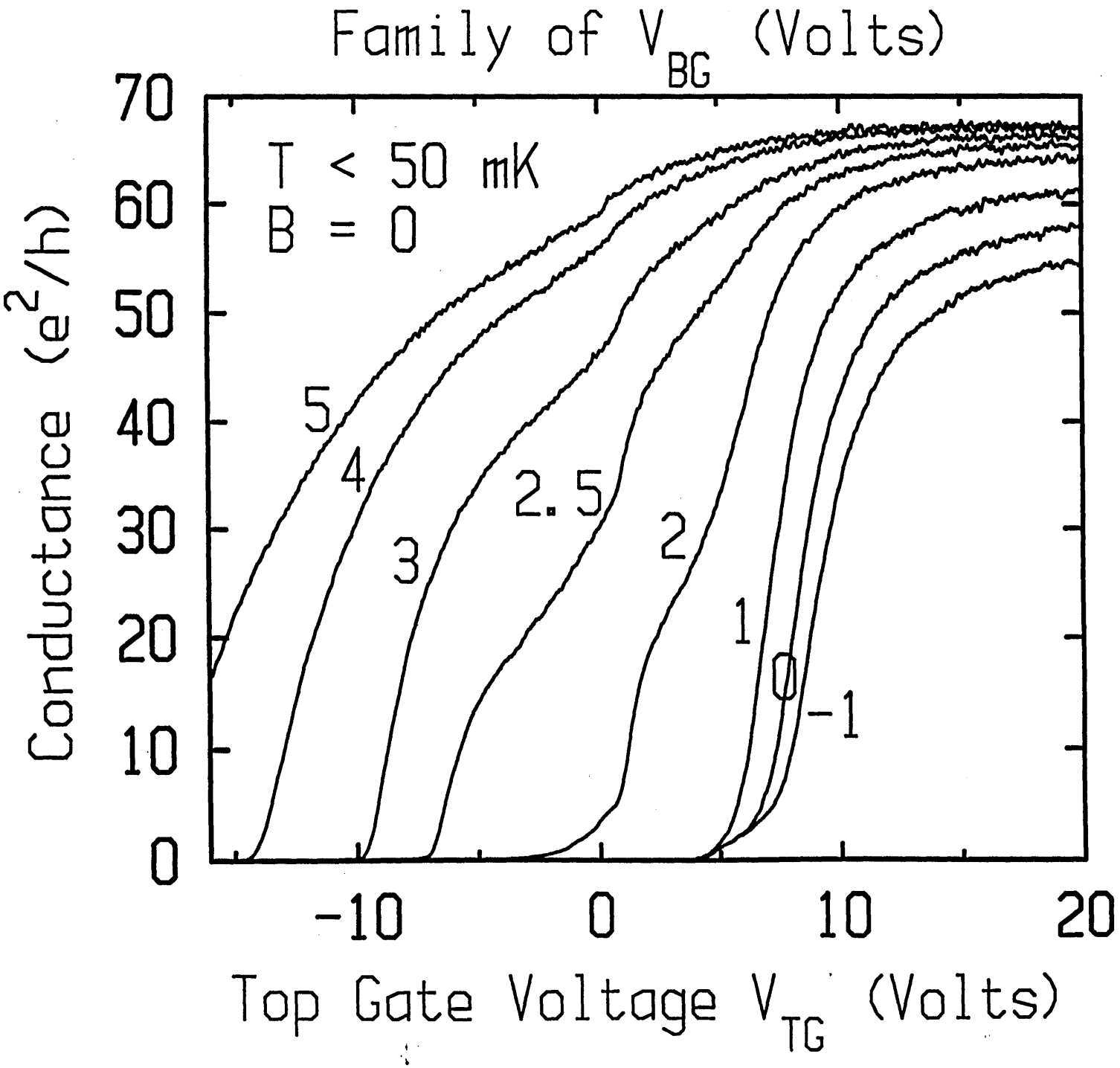

Family of V_{BG} (Volts)
T < 50 mK
B = 0
5
4
3
2.5
2
1
0
-1
Conductance (e^2/h)
70
60
50
40
30
20
10
0
-10
0
10
20
Top Gate Voltage V_{TG} (Volts)

Figure 2

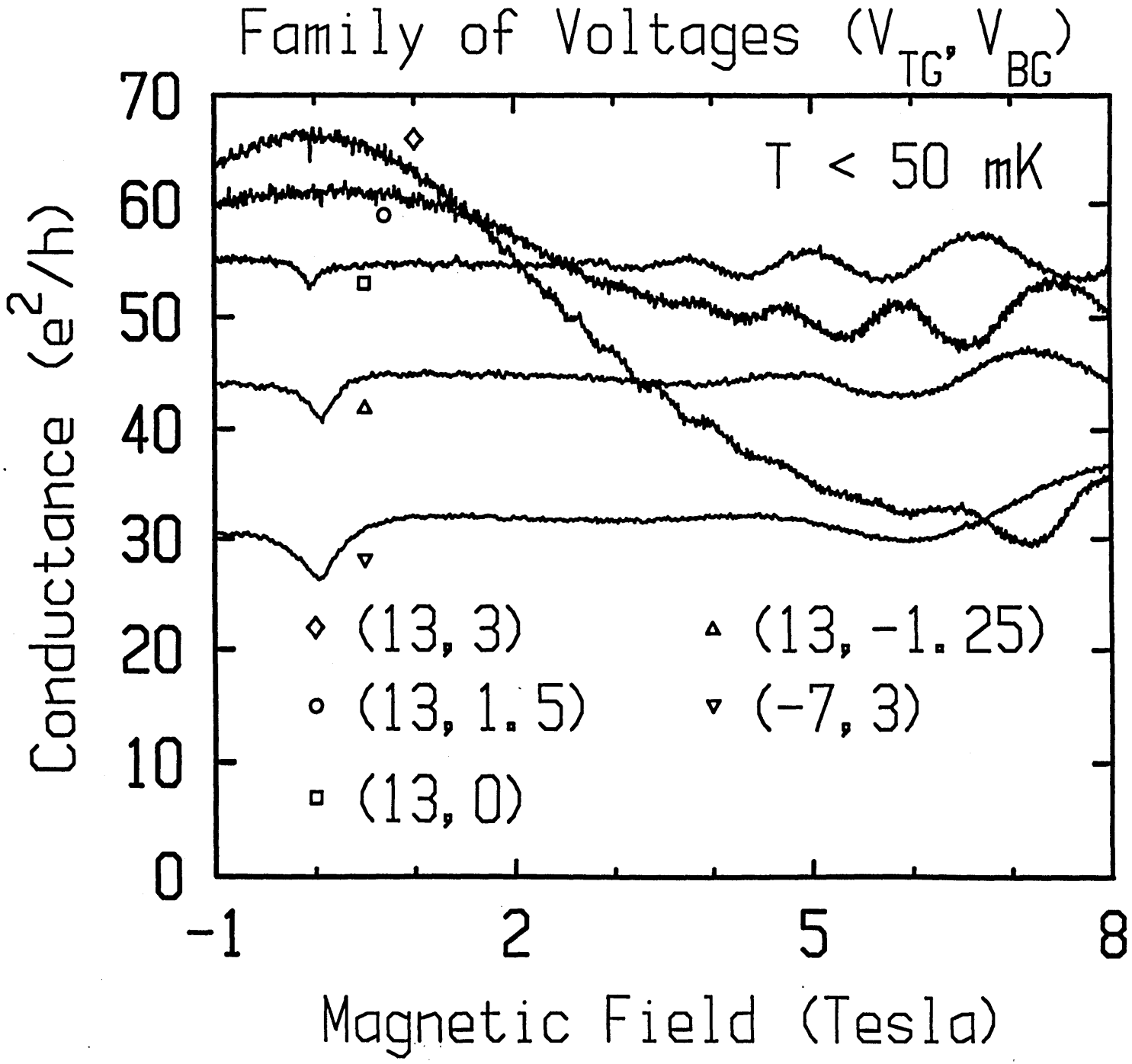

Family of Voltages (V_{TG}, V_{BG})
T < 50 mK
Conductance (e^2/h)
70
60
50
40
30
20
10
0
-1
2
5
8
Magnetic Field (Tesla)
(13, 3)
(13, 1.5)
(13, 0)
(13, -1.25)
(-7, 3)

Figure 3

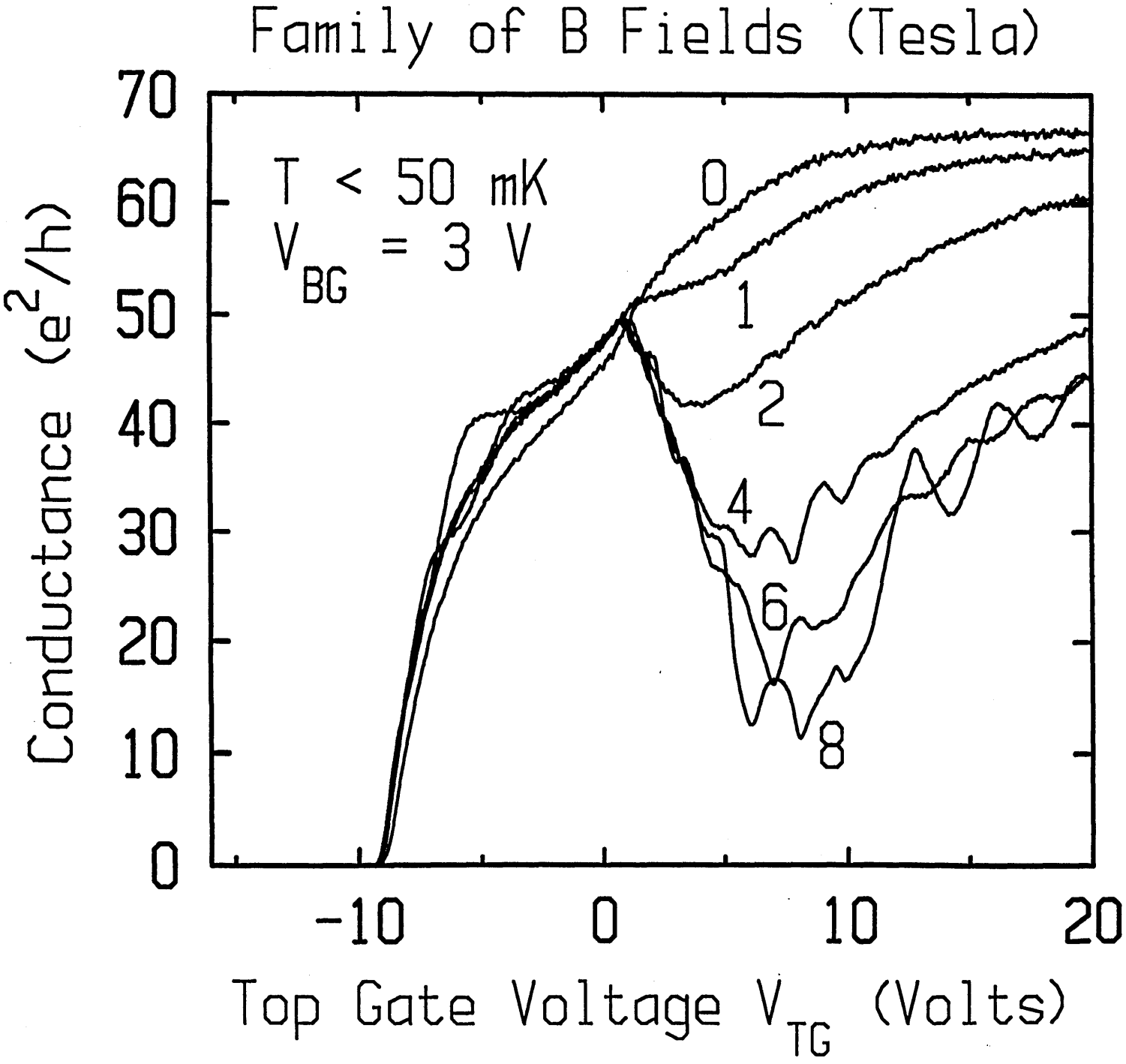

Family of B Fields (Tesla)
T < 50 mK
V_{BG} = 3 V
0
1
2
4
6
8
Conductance (e^2/h)
70
60
50
40
30
20
10
0
-10
0
10
20
Top Gate Voltage V_{TG} (Volts)

Figure 4

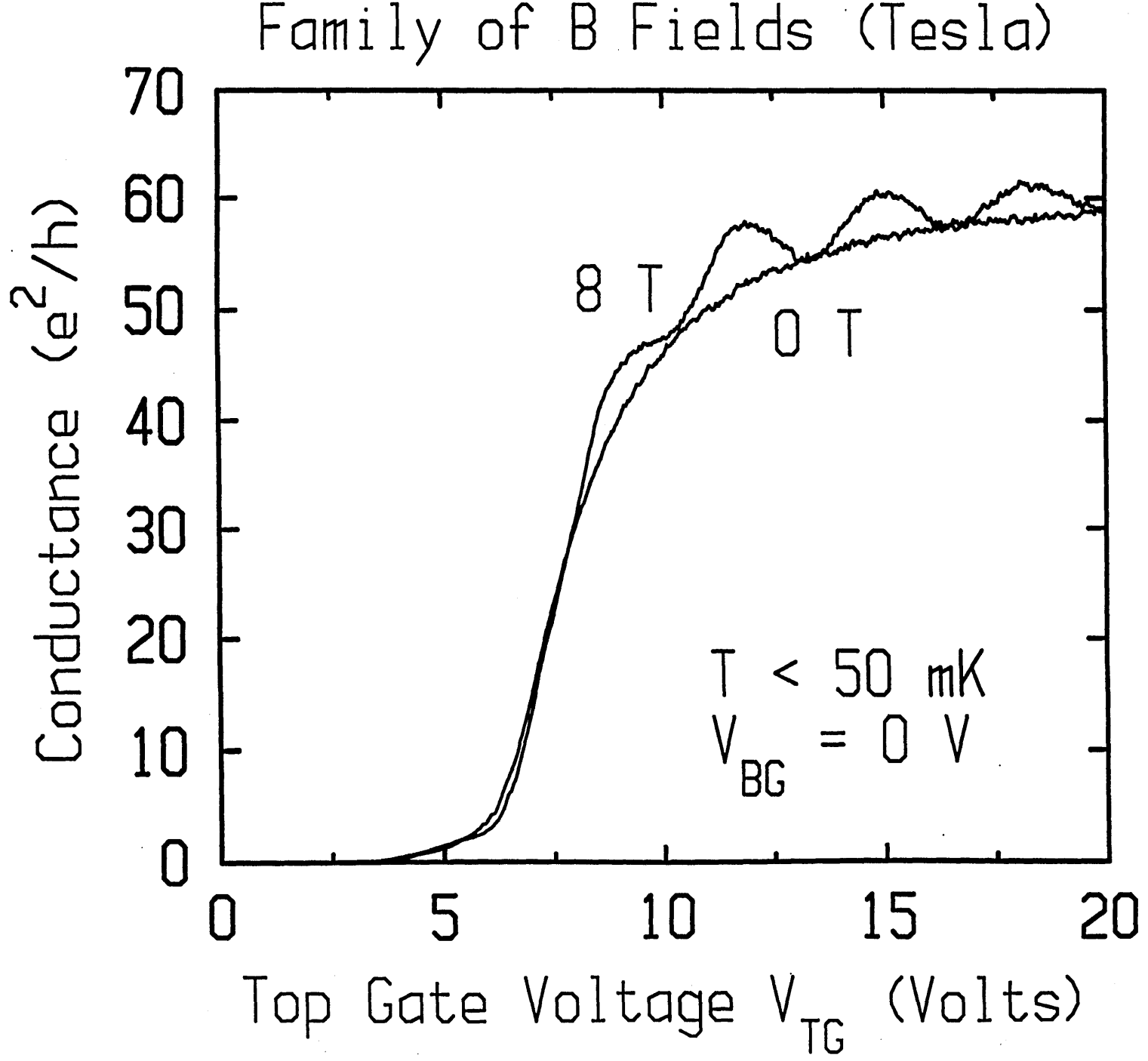

Family of B Fields (Tesla)
Conductance (e^2/h)
70
60
50
40
30
20
10
0
8 T
0 T
T < 50 mK
V_{BG} = 0 V
0
5
10
15
20
Top Gate Voltage V_{TG} (Volts)

Figure 5

Figure 6(a)

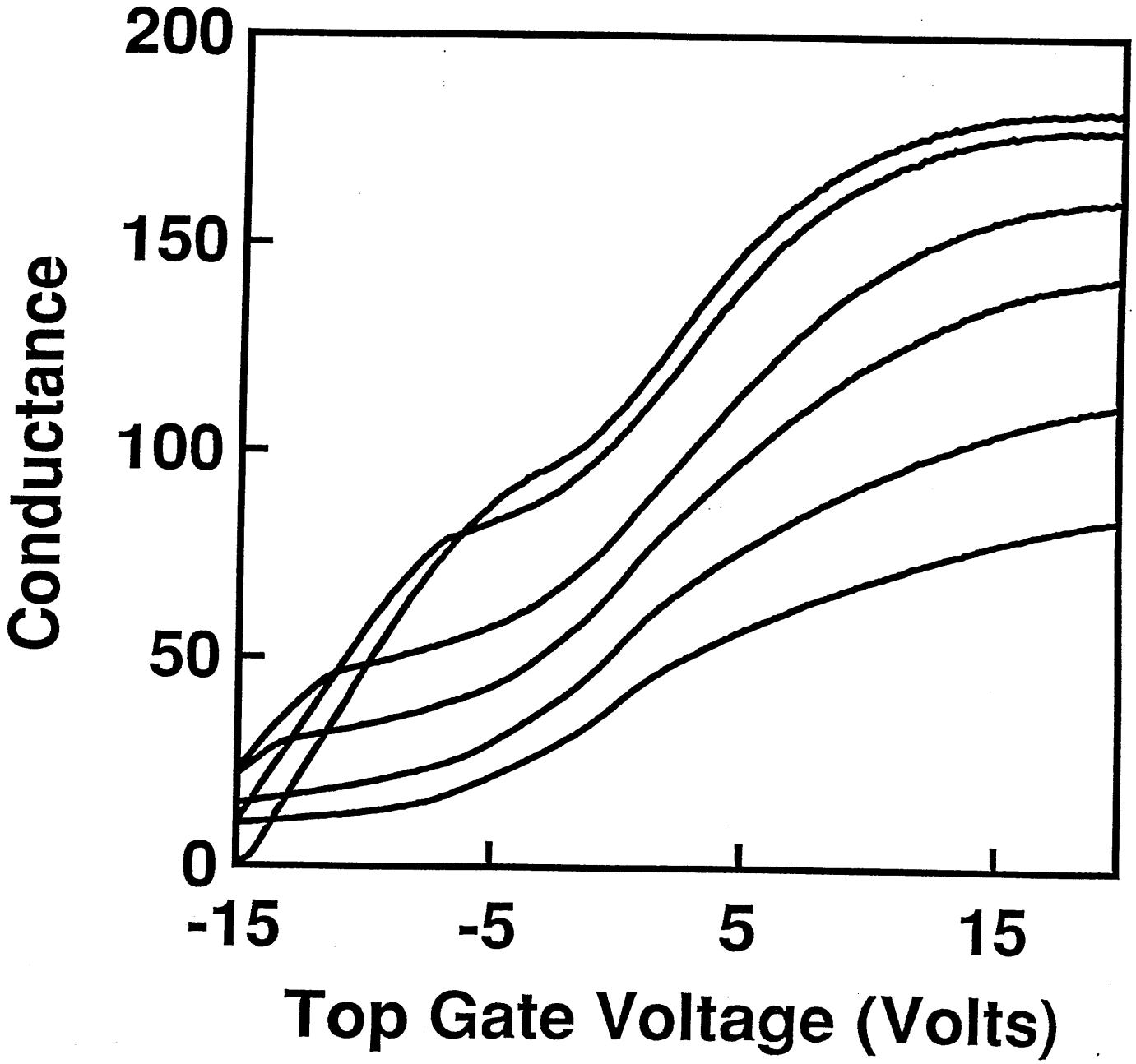
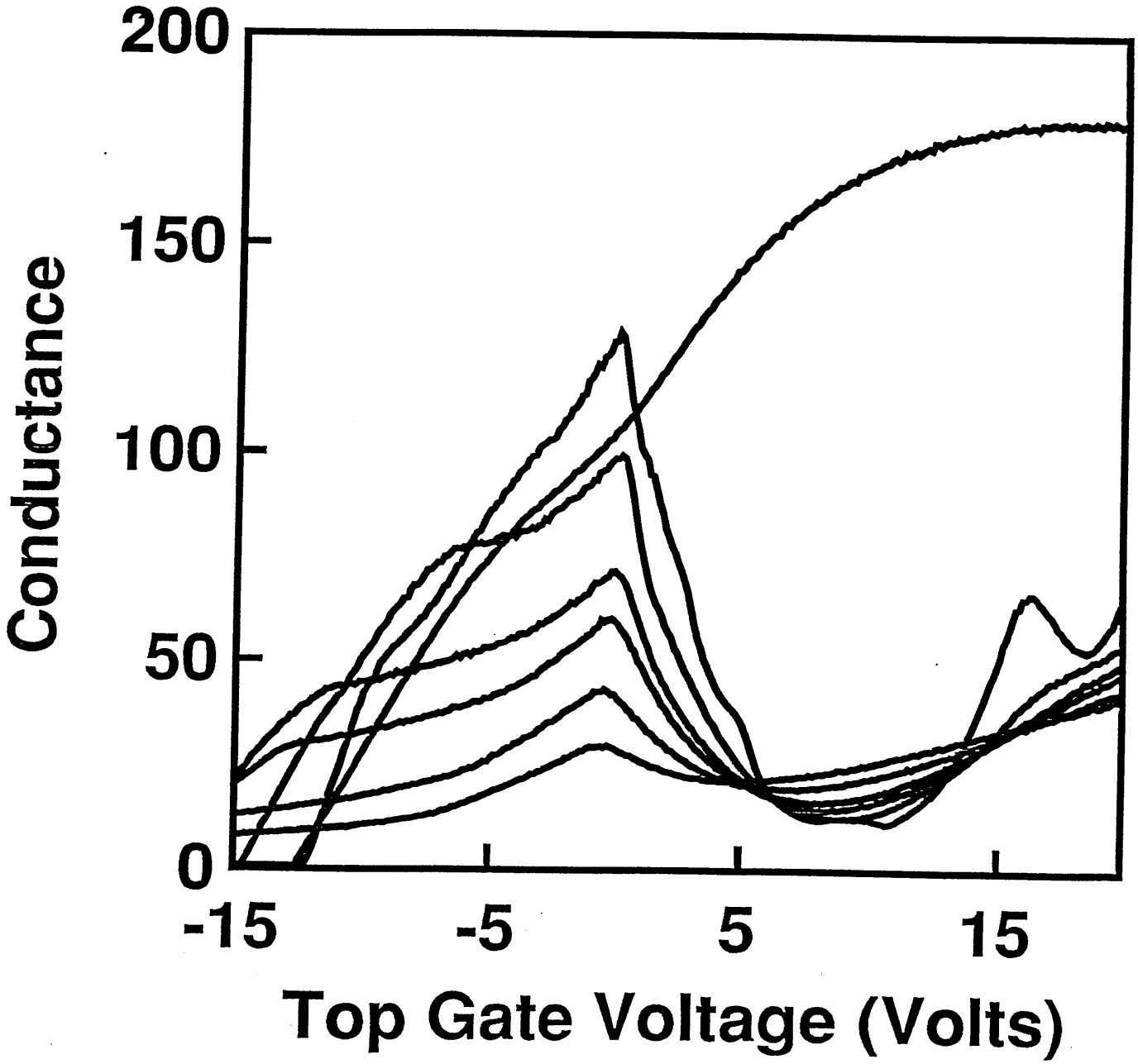


Figure 6(b)



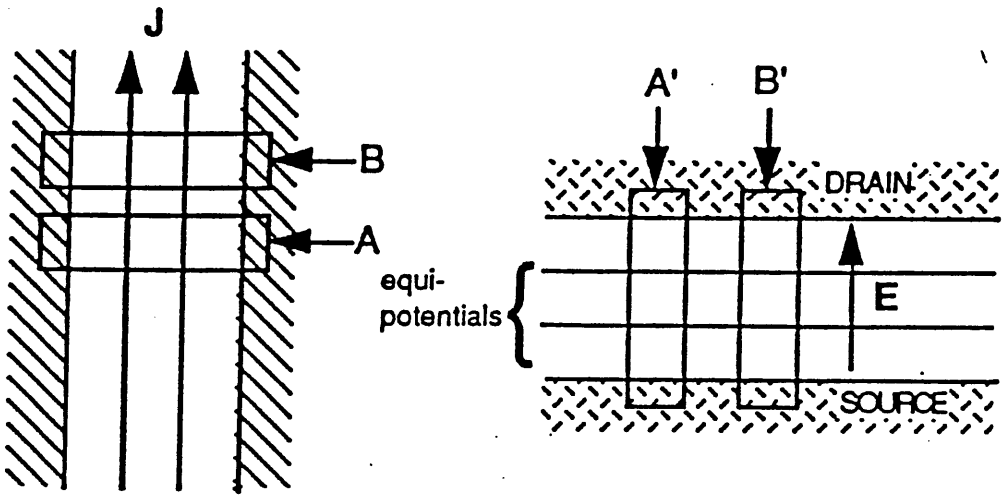


Figure 7

Figure 8

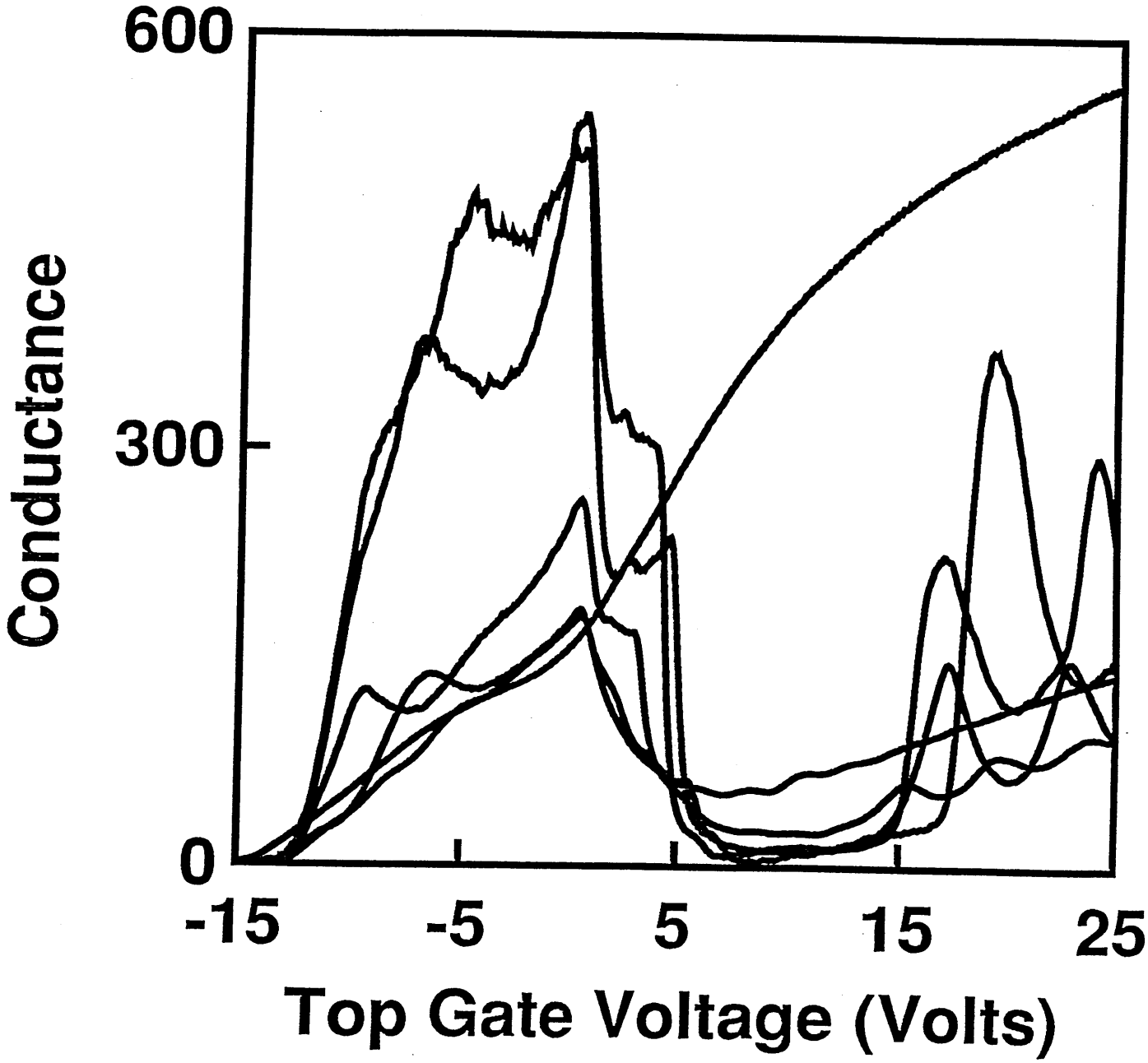
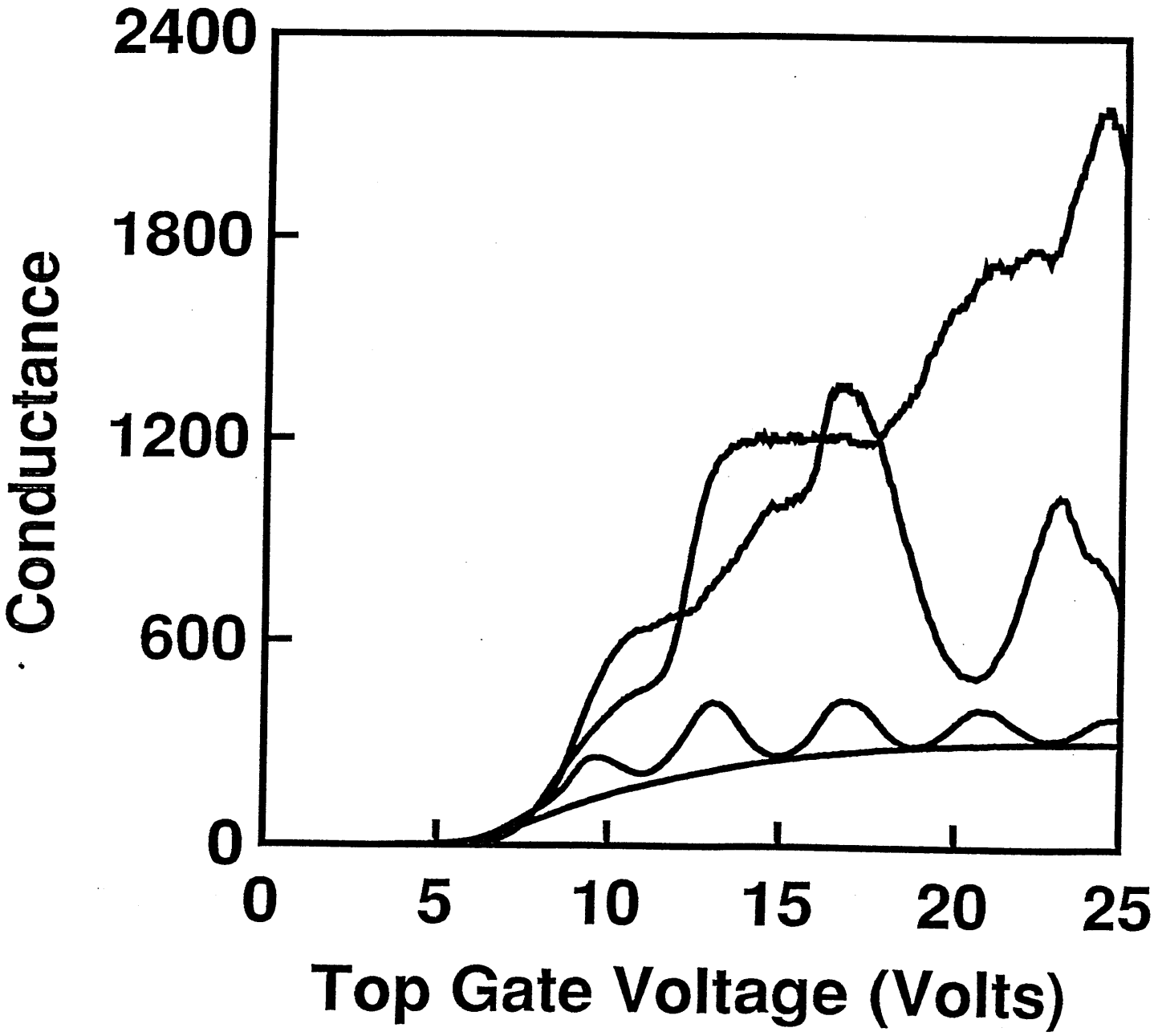


Figure 9


2400
1800
1200
600
0
Conductance
0
5
10
15
20
25
Top Gate Voltage (Volts)