AD-A237 948

DTIC
ELECTE
JUL 0 9 1991
S C D

AFOSR-TR- 91 0567

Report on Grant AFOSR-88-0096

MODELING, SIMULATION, AND ANALYSIS OF QUANTUM TRANSPORT

Supriyo Bandyopadhyay, Craig S. Lent,
and Wolfgang Porod
Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN 46556

Approved for [illegible]
distribution [illegible]

AIR FORCE OF SC[illegible] RESEAR[illegible]
NOTICE OF [illegible]
This tech[illegible]
appr[illegible]
Dis[illegible]
G[illegible]
STI[illegible]

Department of
ELECTRICAL ENGINEERING

UNIVERSITY OF NOTRE DAME, NOTRE DAME, INDIANA

91 7 08 003

91-04295

Report on Grant AFOSR-88-0096

MODELING, SIMULATION, AND ANALYSIS OF QUANTUM TRANSPORT

Supriyo Bandyopadhyay, Craig S. Lent,
and Wolfgang Porod
Department of Electrical Engineering
University of Notre Dame
Notre Dame, IN 46556

Accession For
NTIS GRA&I ☑
DTIC TAB ☐
Unannounced ☐
Justification
By
Distribution/
Availability Codes
Dist | Avail and/or Special
A-1

15 March 1991

Final Report for Period 1 January 1988 – 31 December 1990

Prepared for
U.S. AIR FORCE OFFICE OF SCIENTIFIC RESEARCH
Bolling Air Force Base
Washington, DC 20332-6448

SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION/AVAILABILITY OF REPORT unlimited
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION University of Notre Dame Electrical Engineering Dept. 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Same as 8a
6c. ADDRESS (City, State and ZIP Code) Notre Dame, IN 46556	7b. ADDRESS (City, State and ZIP Code) Same as 8c
8a. NAME OF FUNDING/SPONSORING ORGANIZATION U.S. Air Force Office of Scientific Research 8b. OFFICE SYMBOL (If applicable) NE	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-88-0096
8c. ADDRESS (City, State and ZIP Code) Bolling Air Force Base Washington, D.C. 20332-6448	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 61102F PROJECT NO.: 2306 TASK NO.: B1 WORK UNIT NO.
11. TITLE (Include Security Classification) Modeling, Simulation, and Analysis of Quantum Transport	
12. PERSONAL AUTHOR(S) Bandyopadhyay, Supriyo; Lent, Craig S.; Porod, Wolfgang	
13a. TYPE OF REPORT Final	13b. TIME COVERED FROM 88-1-1 TO 90-12-31
14. DATE OF REPORT (Yr., Mo., Day) 91-3-15	15. PAGE COUNT
16. SUPPLEMENTARY NOTATION	

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Research carried out by the Nanostructures Group in the Department of Electrical Engineering at Notre Dame was concerned with a variety of quantum transport and optical phenomena in mesoscopic structures. The major issues examined included analysis of the scope and performance of velocity modulation transistors for ultrafast switching applications, novel methods of fabricating quantum wires that can provide extremely high electron mobility and high optical absorption, the Aharonov-Bohm effect along with its basic physics and possible applications in quantum interference transistors with ultra-low power-delay product, a variety of interference phenomena - both optical and electronic, study of quantum transport in ballistic electron waveguides and bends with special attention given to space charge effects, quantum transport in heavily doped structures with strong elastic scattering as well as magnetotransport theory, accurate treatment of real space transfer in quantum wells, a critical examination of the scope of quantum devices as both analog and digital elements either as a discrete device or in integrated circuits, and finally, interesting properties of periodic ballistic structures.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☐ SAME AS RPT ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION Unclass
22a. NAME OF RESPONSIBLE INDIVIDUAL Pomerene	22b. TELEPHONE NUMBER (Include Area Code) 202-767-4931 22c. OFFICE SYMBOL NE

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

SECURITY CLASSIFICATION OF THIS PAGE

TABLE OF CONTENTS

1. SUMMARY OF RESEARCH 2

2. RESEARCH DESCRIPTION 3
   2.1: Velocity Modulation Transistors 3
   2.2: Quantum Wires and Their Importance 4
   2.3: The Aharonov-Bohm Effect and Device Applications 5
      2.3.1: Aharonov-Bohm Interferometers 5
      2.3.2: Twin Minima in Electrostatic Aharonov-Bohm Conductance Oscillations 6
      2.3.3: Mode Quenching 6
   2.4: Coherence and Interference Phenomena 7
      2.4.1: Optical Fluctuations Due to Coherence 7
      2.4.2: Quantum Interference Effects in Transient Transport 7
   2.5: Quantum Effects in Real-Space Transfer 8
      2.5.1: Phonon Assisted Transition From Bound to Unbound States 8
      2.5.2: Anisotropy in Real Space Transfer 8
   2.6: The Quantum Transmitting Boundary Method 9
   2.7: Electron Transport in Coherent Ballistic Channels 9
      2.7.1: Nonlinear Conductance of Ballistic Constrictions 9
      2.7.2: Space-Charge Effects in Ballistic Constrictions 10
      2.7.3: Transmission Through a Bend in an Electron Waveguide 10
      2.7.4: Electron Waveguide Cavities 10
      2.7.5: Current Vortex Formation 10
   2.8: Diffusive Quantum Transport 12
   2.9: Magnetic Edge States 12
   2.10: Analysis of Quantum Devices - Merits and Drawbacks 13
   2.11: Landauer Resistances of Finite Repeated Structures 13

3. PUBLICATIONS 14

4. CONFERENCE PRESENTATIONS 17

5. PERSONNEL 20

6. APPENDIX: REPRINTS OF PUBLICATIONS 21

1. SUMMARY OF RESEARCH

During the period January 1, 1988 to December 31, 1990, research carried out by the Nanostructures Group in the Department of Electrical Engineering at Notre Dame was concerned with a variety of quantum transport and optical phenomena in mesoscopic structures. This research was funded by the Air Force Office of Scientific Research under Grant No. AFOSR-88-0096. The major issues examined included analysis of the scope and performance of velocity modulation transistors for ultrafast switching applications, novel methods of fabricating quantum wires that can provide extremely high electron mobility and high optical absorption, the Aharonov-Bohm effect along with its basic physics and possible applications in quantum interference transistors with ultra-low power-delay product, a variety of interference phenomena - both optical and electronic, study of quantum transport in ballistic constrictions and bends with special attention given to space charge effects and bend resistance, quantum transport in heavily doped structures with strong elastic scattering as well as magnetotransport theory, accurate treatment of real space transfer in quantum wells, a critical examination of the scope of quantum devices as both analog and digital elements either as a discrete device or in integrated circuits, and finally interesting properties of and phenomena in periodic structures.

The research supported by this grant resulted in twenty one journal papers, ten articles in books, and thirty four conference presentations. The research contributed to the granting of five M.S. degrees and have supported two continuing M.S. research and three continuing Ph.D. research.

Individuals who participated in the grant are Profs. S. Bandyopadhyay, C. S. Lent, W. Porod and graduate students S. Bhobe, M. Leng, L. Liang, S. Sivaprakasam, H. Harbury, S. Chaudhuri, V. Deshpande, and S. Subramaniam.

2. RESEARCH DESCRIPTION

2.1 Velocity Modulation Transistors

Velocity Modulation Transistors (VMT) were originally proposed as field effect switching transistors in which the channel conductance is modulated by modulating the mobility rather than the concentration of the carriers in the channel by an applied gate voltage. The advantage is of this is that since the device is not switched by physically moving carriers out of the channel, the switching speed is not limited by the transit time of electrons; instead, it is limited by the momentum relaxation time which is typically much shorter than the transit time. This leads to an ultrafast electronic switch.

We investigated this device in a normally-off (enhancement mode) configuration. Enhancement-mode operation is important to prevent standby power dissipation in circuits. The relevant structure consists of a quantum well one half of which is intentionally doped while the other half is left undoped. In the absence of any external electric field, electrons mostly reside in the doped half due to the electrostatic attraction it feels from the dopants (ionized donors). In this region, the electron mobility is low due to strong impurity scattering, so that when the electrons are mostly in the doped region, the overall mobility of electrons in the well (channel) is low. As a result, the channel conductance is low and the device is off. Now if an electric field is applied perpendicular to the well-barrier interface, it skews the electron wavefunction to the undoped half in which the mobility is much higher because of the absence of in-situ impurity scattering. Consequently the channel mobility and conductance goes up dramatically. This switches the device on.

We investigated this switching mechanism using a combination of quantum-mechanical and semiclassical analysis. Our quantum-mechanical analysis was fully self-consistent in the sense that we solved the Schrödinger and Poisson Equations to find the potential and wavefunction in the well both in the presence and absence of the electric field. From the wavefunction, the impurity scattering rates were calculated using Fermi's Golden Rule and used in a Monte Carlo simulation to evaluate the momentum relaxation times and hence the mobility. Our analysis revealed that wells with high carrier concentrations are unsuitable for this device since screening increases the momentum relaxation times dramatically and reduces the switching speed. However, regardless of the carrier concentration, the field (voltage) required to change the conductance by 90 % at a temperature of 4.2 K was very small ($\sim$ hundreds of millivolts).

Therefore, we found that this device can have extremely large transconductance at liquid helium temperatures resulting in small RC time constants for switching and large unity gain frequency.

We also investigated the performance of the device when instead of doping one-half of the well uniformly, it is delta-doped with a series of planar doped sheets. Delta doping has been shown to result in improved performance for MESFETs by several groups, but primarily because of increased carrier concentration. We found that delta doping in a VMT does improve the switching speed by increasing the momentum relaxation rate dramatically. The physics underlying this effect is extremely interesting. If the scattering from the various delta doped layers

are *correlated*, that is the scattering amplitudes add in phase, then the strength of the scattering interaction is increased many-fold. On the other hand, if the scatterings are uncorrelated, then scattering is not enhanced significantly. We expect strong correlations between the delta-doped layers at low temperatures so that delta-doping may indeed result in better device performance. Such realizations are extremely useful for the design of VMTs.

Relevant Publications:

- Modulated Interface Roughness Scattering in Quantum Wells and its Device Applications, Suyog Bhobe, Wolfgang Porod and Supriyo Bandyopadhyay, Surface and Interface Analysis, **14**, 590 (1989).
- Modulation of Impurity Scattering Rates by Wavefunction Engineering and Its Device Applications, Suyog Bhobe, Wolfgang Porod and Supriyo Bandyopadhyay, Solid State Electronics, **32**, 1651 (1989)
- Tailoring Transport Properties By Wavefunction Engineering in Quantum Wells and its Device Applications, Suyog Bhobe, Wolfgang Porod and Supriyo Bandyopadhyay, Nanostructure Physics and Fabrication, eds. M. A. Reed and W. P. Kirk, (Academic Press, Boston, 1989), p. 201.

2.2: Quantum Wires and Their Importance

Quantum wire structures are expected to provide extremely high electron mobility at low enough temperatures due to the drastic suppression of elastic scatterings that dominate at cryogenic temperatures. Additionally, they show strong optical non-linearities and high optical absorption due to extremely large density of states. Moreover, quantum wire structures have recently found a large number of applications in "waveguide electronics" (or mesoscopic phenomena) where electron transport is analogous to microwave propagation in a waveguide. It has also been claimed by us that the performance of certain quantum devices improve dramatically if they are realized from quantum wires rather than quantum wells.

The obvious method for realizing quantum wells is by etching mesas or V-grooves into quantum wells. This is a challenging process since precise control of etching is required so as not to deplete the wires of mobile carriers due to Fermi level pinning. We have proposed an alternate technique where quantum confinement into one-dimensionality is achieved through space-charge effect rather than using quantum wells. The main feature of our design is that the confinement of electrons in the quasi 1-d wire is exclusively produced by the internal electric fields resulting from the conduction band discontinuities at crossing heterointerfaces. The idea is to use 1-d confinement created at the intersection of two 2-d structures. The quasi 1-d quantum wire can, in principle, be fabricated by crossing two accumulation layers formed along separate heterointerfaces, thereby realizing an 'accumulation wire.' Electron confinement is thus solely provided by the inherent electric fields associated with heterointerfaces, without the need for confinement in artificial quantum wells. This feature distinguishes our design of a quantum wire

from earlier proposals where confinement in a quantum well was necessary. Our design of the quantum wire compares to previous designs in the same manner in which a quasi 2-d system in an accumulation layer compares to a quasi 2-d system in a quantum well.

In particular, we investigate accumulation wires which are produced at the intersection of two perpendicular heterointerfaces. We study the electronic states at the intersection by solving Schrödinger's and Poisson's equations self-consistently, taking into account the potential discontinuities at the interfaces. An isolated quantum wire will only be produced if the Fermi energy is such that the electrons accumulate only at the crossing point, but not at the heterointerfaces themselves. We study under which doping conditions and for which material combinations such an isolated accumulation wire can be achieved.

Relevant Publications:

- A Proposed Quantum Wire Structure: An 'Accumulation Wire' at Crossing Heterointerfaces, Henry K. Harbury and Wolfgang Porod, J. Vac. Sci. Technol. B 8, 923 (1990).
- Numerical Study of Electronic States in a Quantum Wire at Crossing Heterointerfaces, *Computational Electronics, Semiconductor Transport and Device Simulation*, ed. by K. Hess, J. P. Leburton, and U. Ravaioli (Kluwer Academic Publishers, 1990), p. 243.

2.3: The Aharonov-Bohm Effect

2.3.1: Aharonov-Bohm Interferometers

The Aharonov-Bohm effect has long been proposed as a quantum interference effect for realizing novel electronic devices with very low power-delay product, multifunctionality and low threshold voltage. The low threshold voltage is especially attractive since that makes it possible to switch these devices optically rather than electronically resulting in ultrafast opto-electronic switches whose switching speed is not limited by any RC time constant.[1] The major drawback of these devices however is that switching is accomplished through quantum-mechanical interference which requires tight control over an electron's phase throughout the structure. In the case of ballistic transport, when no elastic or inelastic scattering takes place, there are two effects that introduce an uncertainty in the phase and thereby dilutes the interference effect - (a). thermal spread in the longitudinal (along the direction of current flow) momenta of electrons which increases with increasing temperature, and (b). spread in the transverse momenta in polydimensional structures. In one-dimensional interferometers, the latter cause can be eliminated so that these structures can operate at higher temperatures. The advantage of one-dimensional structures is even more pronounced when transport in diffusive rather than ballistic. In that case, one dimensional confinement can drastically reduce elastic scattering which helps to retain control over the electron's transit time which determines the electrostatic Aharonov-Bohm phase shift.

[1]Optical switching using virtual charge polarization induced by sub-bandgap optical excitation results in ultrafst switching since the switching time is neither limited by transit time, nor by RC time constant, nor by radiative recombination lifetimes.

Based on the above realizations, we have designed new electrostatic Aharonov-Bohm interferometers that could operate at 77 K which is far above the temperature of operation of conventional interferometers (typically less than 4.2 K). These structures, albeit difficult to fabricate, are realizable with present day technology. They utilize double quantum wires fabricated by etching a very narrow V-groove (presumably by focussed ion beam etching) in a quantum well. Our analysis reveals that these structures can provide $\sim$ 90 % modulation of the conductance at 77 K. We later extended our analysis to examine the performance of both 2-d and 1-d interferometers in the diffusive regime. This resulted in the determination of temperature-carrier concentration maps for obtaining various levels of performance for both 1-d and 2-d interferometers. Such results are extremely useful to device engineers in designing Aharonov-Bohm devices.

2.3.2: Twin Minima in Electrostatic Aharonov-Bohm Conductance Oscillations

Our research with the electrostatic Aharonov-Bohm effect also revealed a novel fundamental feature associated with this effect in 1-d rings that was hitherto unsuspected. We predicted the existence of *two* different types of conductance minima in the conductance oscillations arising from different interference conditions. This doubles the frequency of oscillations making it *twice* that predicted by the Aharonov-Bohm effect. This novel feature, which gives rise to secondary minima, is not inhibited strongly by elastic scattering but is less robust than the primary Aharonov-Bohm effect so that it can be observed only at low temperatures. The opposite voltage dependences of the primary and secondary minima helps to distinguish between the two types and may serve as a fundamental test for the observation of the electrostatic Aharonov-Bohm effect.

2.3.3: Mode Quenching

We also found that interface roughness scattering can cause a peculiar mode quenching effect in quasi one dimensional Aharonov-Bohm interferometers whereby a propagating mode can suddenly become evanescent in a region containing an island disorder. This effect leaves its fingerprints on the oscillation characteristics in the form of well resolved kinks.

Relevant Publications:

- Double Quantum Wire Aharonov-Bohm Interferometers for Possible LN_2 Temperature Operation, S. Bandyopadhyay and W. Porod, Superlattices and Microstructures, **5**, 239 (1989).

- Performance of Electrostatic Aharonov-Bohm Interferometers in the Diffusive Regime, Appl. Phys. Lett., **53**, 2323 (1989).

- Doubled Frequency of the Conductance Minima in Electrostatic Aharonov-Bohm Oscillations in One-dimensional Rings, M. Cahay, S. Bandyopadhyay and H. L. Grubin, Nanostructure Physics and Fabrication, eds. M. A. Reed and W. P. Kirk, (Academic Press, Boston, 1989), p. 407.

- Two Types of Conductance Minima in Electrostatic Aharonov-Bohm Conductance Oscillations, M. Cahay, S. Bandyopadhyay and H. L. Grubin, Phys. Rev. B., **39**, 12989 (1989).

2.4: Coherence and Interference Phenomena in Mesoscopic Nanostructures

2.4.1: Optical Fluctuations due to Coherence

Fluctuation effects in the conductance of nanostructures arising from quantum interference of electrons have been studied widely in the context of universal conductance fluctuations. Recently, we predicted a novel quantum interference phenomenon that gives rise to fluctuations in the optical spectra of disordered nanostructures when the inelastic scattering time in the structures exceeds the radiative recombination lifetime. This is the first prediction of quantum fluctuation effects in the optical (rather than transport) properties of nanostructures. This effect arises as follows. If the dominant process for the relaxation of the optical dipole moment in the structure is elastic scattering which preserves phase relationships between the electron and hole states, then the optical spectrum of the structure will depend on these phase relationships and change if these phase relationships are altered by an external electric or magnetic field. In other words, the spectrum will fluctuate randomly but reproducibly in an external field. These fluctuations will be sample-specific since the phase-relationships in the absence of any field depends on the exact "configuration" of the elastic scatterers within the sample. Consequently, the fluctuation pattern will be a fingerprint of the exact locations of the defects (elastic scattering centers) within a sample. Since the spectrum depends on the configuration of the defects (elastic scattering centers) within a sample, it will have a unique consequence on the inhomogeneous broadening of lineshapes in multi quantum well structures or superlattices. In a lateral surface superlattice, each quantum dot will have a slightly different optical spectrum if the dots merely have different impurity configurations but are otherwise identical. This phenomenon therefore sets a fundamental lower limit to the inhomogeneous broadening in a superlattice structure.

Because of the essential similarity between the physics of this effect and that of "universal conductance fluctuations", we consider this effect to be an optical analog of the conductance fluctuations.

This effect will have fundamental implications for optical switching devices in integrated arrays.

2.4.2: Quantum Interference Effects in Transient Transport.

We also extended the formalism that we used to study the optical fluctuation effect to the study of transient transport effects in semiconductor nanostructures. We found that if the momentum relaxation time in a structure is shorter than the transit time which in turn is shorter than the inelastic scattering time (i.e. the momentum relaxes entirely due to elastic scattering), then the mobility in a disordered structure (calculated quantum-mechanically) depends not only the degree of disorder, but also on the precise configuration of the disorder (i.e. the locations of the scattering centers). We compared the results of the quantum-mechanical calculation of mobility with results obtained from the Fermi's Golden Rule formalism which does not account for interference between the scatterers. It was found that the quantum-mechanical result depends on the configuration unlike the Fermi's Golden Rule result and also the quantum mechanical result is always smaller in magnitude probably because of the coherent back-scattering effect that gives rise to Anderson localization.

Relevant Publications:

- Fluctuations in the Optical Spectra of Disordered Microstructures Due to Quantum Interference Effects, S. Bandyopadhyay, Phys. Rev. B., 38, 7466 (1988).
- Quantum Interference Effects in Transient Electronic Transport, D. R. Poole and S. Bandyopadhyay, J. Appl. Phys., 66 5422 (1989).
- Quantum Phase Coherent Effects in the Photoluminescence Spectra of Disoredered Mesoscopic Structures, Supriyo Bandyopadhyay, Nanostructure Physics and Fabrication, eds. M. A. Reed and W. P. Kirk, (Academic Press, Boston, 1989), p. 201.

2.5: Quantum Effects in Real-Space Transfer

2.5.1: Phonon Assisted Transitions From Bound to Unbound States

It is well known that quantum wells, in addition to confined bound states, produce resonant continuum states. These resonant states are a consequence of the reflections, and the consecutive interference, of the electronic wavefunctions at the edges of the quantum well. While the existence of these so-called virtual resonant states has been recognized, their influence on electronic transport has received little attention. We investigated the influence of these continuum resonances on the transfer of electrons in and out of quantum wells.

We find that the matrix elements which determine scattering rates exhibit structure at the resonant energies. This leads to suppression of scattering by polar optical phonons relative to non-polar optical and acoustic phonon scattering. The suppression by the effective matrix element of small q_z scattering can be understood from a simple qualitative argument. Resonant continuum states correspond to states which are orthogonal to bound states inside the well. This condition of orthogonality implies that for $q_z = 0$, the matrix element connecting bound states with resonant states vanishes. Thus the total scattering rate is dominated by the contribution of those phonons with larger momentum components normal to the interface. Polar optical scattering is predominantly small q, forward scattering. We conclude, therefore that the polar optical scattering rate will be suppressed relative to processes not so strongly weighted toward forward scattering. This implies that it is non-polar optical and acoustic phonons that are primarily responsible for carriers scattering out of the well and also their capture.

2.5.2: Anisotropy in Real Space Transfer

We have also studied the dependence of the real-space transfer rates upon the shape of the confining quantum well. We discovered a rather surprising asymmetry in the escape rates from wells with non-symmetrical potential profiles. Consider, for example, a well has one abrupt edge (confining wall) and one edge which is smoothly graded. Electrons transferred out of the well by polar optical phonon scattering will preferentially scatter in the direction of the more abrupt interface. We explored this asymmetry in several potential profiles. The effect has its origin in the overlap between the initial-state wavefunction and the final scattered-state wavefunction. This phenomenon may prove very useful in the design of real-space transfer devices.

Relevant Publications:

- Escape from Quantum Wells by Polar Optical Phonon Scattering, Craig S. Lent, Lie Liang and Wolfgang Porod, Appl. Phys. Lett. **54**, 2315 (1989).
- Real Space Transfer Rates for Polar Optical Phonon Scattering from Asymmetric Quantum Wells, Craig S. Lent and Lie Liang, Solid State Electronics **32**, 1479 (1989).
- Escape from Quantum Wells via Polar Optical Phonon Scattering, Lie Liang and Craig S. Lent, J. Appl. Phys. **68** 1741 (1990).

2.6: The Quantum Transmitting Boundary Method (QTBM)

A long-standing difficulty in computing the solutions to the Schrödinger equation for current-carrying states is the difficulty of expressing the boundary conditions for the wavefunction without prior knowledge of the transmission and reflection coefficients. We developed a technique for handling these transmitting boundaries which enables us to solve directly for the wavefunction of the current-carrying states in two dimensions. This is a very important tool in investigating ballistic quantum transport. It allows us to efficiently find both the transmission coefficients necessary for calculating the total current, and the charge density inside the device region, necessary for including self-consistent effects. Because the complete wavefunction is calculated, *any* quantum mechanical observable can be computed. An early benefit of this capability was the identification of vortices in the current inside a cavity (see discussion in Section 2.5.5).

Relevant Publications:

- Numerical Simulation of Single-Electron Transmission Through Two Dimensional Quantum Device by the Finite Element Method, D.J. Kirkner, Craig S. Lent, and Srinivas Sivaprakasam, International Journal for Numerical Methods in Engineering **29**, 1527 (1990).
- The Quantum Transmitting Boundary Method, Craig S. Lent and D.J. Kirkner, J. Appl. Phys. **67**, 6353 (1990).

2.7: Electron Transport in Coherent Ballistic Channels

2.7.1: Nonlinear Conductance of Ballistic Constrictions

Much attention has focussed on interpreting the first experiments which showed that conductance through a very small ballistic channel is quantized. This has been confirmed experimentally and the theory is now well understood. The quantization hold only for the linear response regime, however, when the voltage drop from one end of the channel to the other is very small. We have employed the QTBM to calculate the I-V curve for a constriction in a quantum channel when

the voltage drop is not small – the non-linear response. We chose a simple linear model for the potential drop down the length of the constriction. The contributions from all the occupied states were calculated at non-zero temperature. Our results showed a saturation in the current through the constriction. The saturation current varied nearly linearly with the width of the constriction. We see a smooth transition to quantized conduction at the constriction vanishes and we recover the straight-channel results. Since the constriction width can be controlled electrostatically, this response is very analogous to that of a conventional JFET.

2.7.2: Space-Charge Effects in Ballistic Constrictions

The importance of a built-in potential caused by space charge effect in narrow ballistic constrictions was investigated using a scattering matrix method coupled with a boundary element technique. It was found that in the linear response regime, the built in potential does not cause qualitative changes in the transport characteristics although it certainly causes quantitative changes.

2.7.3: Transmission Through a Bend in an Electron Waveguide

If ultimately quantum waveguide devices are to be interconnected, it is important to establish whether bends in waveguides present significant sources of reflection and mode-mixing. We again employed the QTBM to calculate the transmission through a circular, right-angle bend in an electron waveguide. We examined an idealized case of infinitely hard walls, and a more realistic model with softer parabolic walls. The latter allow some leakage around the bend by tunneling. Our calculation established the the problem of reflection by such a bend can easily be minimized with modest design considerations. We found that mode-mixing is significant and may require single-moded operation of waveguide devices.

2.7.4: Electron Waveguide Cavities

The wave nature of electrons in a ballistic channel implies that localized regions of the channel which are wider can behave like resonant cavities in microwave theory. One therefore expects transmission resonances at energies which depend on the cavity geometry. We have explored geometrical effects in transmission through symmetric cavities. Our numerical approach, using the QTBM, includes automatically the effects of all evanescent channel modes. Further, we were able to study cavities with soft walls rather than restricting ourselves to the idealized hard-wall case. Several device designs have been proposed which exploit the dramatic sensitivity of the transmission to the cavity geometry. Small changes can be made (electrostatically) to the cavity dimensions, which result in large changes in the transmitted current. Because all of the theory of such devices relies on models with infinitely hard cavity walls, it is important to ascertain whether the rapid variation of the transmission is an artifact of this feature of the model, or is more robust. We calculated the transmission coefficient through cavities with soft potential walls and found qualitatively similar abruptness in turn-on and turn-off of the transmission, thus verifying the robust character of the effect.

2.7.5: Current Vortex Formation

We examined the flow of current in a ballistic waveguide cavity. This is directly accessible because the QTBM yields the full wavefunction everywhere. Evaluating the current density in

the cavity region, we found the very surprising result that the current formed vortices. The center of the vortex is a point where the complex phase of the wavefunction is singular and the probability density vanishes. At a given energy, many vortices may be formed. If the cavity is symmetric, the vorticity occurs in vortex-antivortex pairs so the wavefunction as a whole has no net angular momentum. If the cavity is asymmetric, however, vortices of one sign may dominate, yielding a net angular momentum. This is possible only because the time-reversal symmetry of the Hamiltonian is broken by the boundary conditions at the current-carrying leads. (Note that no magnetic field is imposed here.)

We studied a circular resonant cavity which is connected by a tunnel barrier to one side of a straight waveguide channel. This structure was designed to maximize vorticity of one sign. We found that, indeed, a single very strong current vortex is present at nearly all energies. Resonances associated with the bound states of the circular cavity produce huge increases in the charge stored in the cavity. Such resonant structures may be useful as quantum capacitors. Additionally, the ability to couple directly states with non-zero angular momentum to current carrying states may yield novel opto-electronic coupling.

Relevant Publications:

- A Two-dimensional Hot Carrier Injector for Electron Waveguide Structures, Craig S. Lent, Srinivas Sivaprakasam and D.J. Kirkner, Solid State Electronics **32**, 1137 (1989).
- Calculation of Ballistic Transport in Two-dimensional Quantum Structures using the Finite Element Method, Craig S. Lent, S. Sivaprakasam and D.J. Kirkner, in *Nanostructure Physics and Fabrication*, 279, edited by M. A. Reed and W. P. Kirk, (Academic Press, Boston, 1989).
- Scattering Matrix Analysis of Electron Transport in Disordered Aharonov-Bohm Interferometers and Ballistic Constrictions, M. Cahay, S. Bandyopadhyay and H. R. Frohne, J. Vac. Sci. Technol., **8**, 1399 (1990).
- Transmission Through a Bend in a Quantum Waveguide, Craig S. Lent, Appl. Phys. Lett. **56**, 2554 (1990).
- Quantum Electron Waveguides: Bends, Constrictions, and Cavities, Craig S. Lent and S. Sivaprakasam, in *Nanostructures and Microstructure Correlation with Physical Properties of Semiconductors*, SPIE Proceedings Vol. 1284, 31, edited by H.G. Craighead and J.M. Gibson, (SPIE, Bellingham, Washington, 1990).
- Calculation of Transport Through Ballistic Quantum Structures, Craig S. Lent, in *Computational Electronics*, edited by K. Hess, J.P. Leburton, and U. Ravaioli, 259, (Kluwer, Boston, 1990).
- Ballistic Current Vortex Excitations in Electron Waveguide Structures, Craig S. Lent, Appl. Phys. Lett. **57**, 1678 (1990).

2.8: Diffusive Quantum Transport

We developed a *scattering matrix method* to study diffusive quantum transport in collaboration with Prof. Marc Cahay of the Department of Electrical and Computer Engineering in the University of Cincinnati. The method is ideally suited to study quantum transport in relatively large and heavily doped structures where other methods such as the recursive Green's function formalism will be computationally inefficient. The scattering matrix method has been successfully applied to reproduce the results of Anderson localization and universal conductance fluctuations in disordered structures. In particular, we have used this method to study the role of evanescent states in quantum transport. Many surprising results were found including the fact that evanescent states have a strong influence on the localization length and on the probability distribution of conductances and their higher order cumulants.

Relevant Publications:

- Influence of Evanescent States on Quantum Transport Through an Array of Elastic Scatterers, M. Cahay, S. Bandyopadhyay, M. A. Osman, and H. L. Grubin, Surface Science **228**, 301 (1990).
- The Generalized Scattering Matrix Approach: An Efficient Technique for Modeling Quantum Transport in Relatively Large and Heavily Doped Structures, S. Bandyopadhyay and M. Cahay, Computational Electronics, (Kluwer Academic Press, Norwell, 1990), p. 223
- Numerical Study of the Higher Order Moments in Conductance Fluctuations in Mesoscopic Structures, P. Marzolf, M. Cahay and S. Bandyopadhyay, Computational Electronics, (Kluwer Academic Press, Norwell, 1990), p. 263.

2.9: Magnetic Edge States

In an applied magnetic field, the electronic states in a confined structure break into two classes: states localized by the magnetic field (Landau levels), and circulating states which traverse the edges of the structure (edge states). We studied the formation of these states in a quantum dot structure as the magnetic field increases. Since the high-field limit produces degenerate Landau levels, this process is sometimes know as Landau condensation. States initially split by the presence of the confining potential, become degenerate as they are localized in the interior of the dot. The edge states corresponding to orbits which skip around the periphery yield a paramagnetic correction to the free electron diamagnetism. Semiclassical orbit are frequently invoked to explain edge-state behavior. We compared current densities calculated from the wavefunction to the semiclassical orbits. We found that a naive picture predicted even the direction of circulation incorrectly. We formulated a connection between the semiclassical orbits and the quantum results. The effects of softer confining potentials was also examined.

Relevant Publication:

- Edge States in a Circular Quantum Dot, Craig S. Lent Phys. Rev. B **43**, 4179 (1991).

2.10: Analysis of Quantum Devices – Merits and Drawbacks

In collaboration with Prof. Gary Bernstein, we investigated the scope of lateral quantum devices utilizing the Aharonov-Bohm effect, quantum diffraction, and stub tuning in T-structure electron waveguides. With respect to switching speed, we found that contrary to popular belief, these devices are not necessarily slow. Even though their current carrying capability is small, the threshold voltages for switching are also small so that the overall RC time constant can be quite small.

The small threshold voltage also makes it possible to switch the devices optically rather than electronically thereby eliminating the RC time constant limitation altogether. However, the major drawback of these devices is the extreme sensitivity of the characteristics to slight structural variations which may inhibit their applications in integrated circuits. In addition, the fact that these devices typically operate in the linear response regime also preclude their use in many conventional applications such as amplification or logic nodes requiring signal restoration. Overall, we found that the most attractive feature of these devices is their multifunctionality which makes it possible to use a single device perform the task of many devices.

Relevant Publications:

- Analysis of the Device Performance of Quantum Interference Transistors Utilizing Ultra-small Semiconductor T-structures, S. Subramaniam, S. Bandyopadhyay and W. Porod, J. Appl. Phys. **68**, 4861 (1990).
- Quantum Devices Based on Phase Coherent Lateral Quantum Transport, S. Bandyopadhyay, G. H. Bernstein and W. Porod, Nanostructure Physics and Fabrication, eds. M. A. Reed and W. P. Kirk, (Academic Press, Boston, 1989), p. 183.

2.11: Landauer Resistance of Finite Repeated Structures

Periodic structures are of interest since they are useful for many device applications and also for the study of fundamental band structure effects such as Bloch oscillations. We have also derived many useful theorems pertaining to the Landauer resistance of finite periodic structures. The Landauer resistance is the quantum-mechanical resistance of a structure in the linear response regime and is related to the transmission properties of electrons through the structures. These theorems, along with a sum rule that we derived, can be applied to the calculation of bandstructure of superlattices and their transmission properties.

Relevant Publications:

- Properties of the Landauer Resistance of Finite Repeated Structures, M. Cahay and S. Bandyopadhyay, Phys. Rev. B., **42**, 5100 (1990).

3. PUBLICATIONS

1. S. Bhobe, W. Porod, S. Bandyopadhyay and D. J. Kirkner, "Modulated Interfacial Disorder Scattering and its Device Applications," *Surface and Interface Analysis* **14**, 590 (1989).

2. S. Bhobe, W. Porod and S. Bandyopadhyay, "Modulation of Impurity Scattering Rates by Wavefunction Engineering," *Solid State Electronics.* **32**, 1651 (1989).

3. S. Bhobe, W. Porod and S. Bandyopadhyay, "Tailoring Transport Properties by Wavefunction Engineering in Quantum Wells and its Device Applications," *Nanostructure Physics and Fabrication*, eds. M. A. Reed and W. P. Kirk (Academic Press, Boston, 1989), p. 201.

4. Henry K. Harbury and Wolfgang Porod, "A Proposed Quantum Wire Structure: An 'Accumulation Wire' at Crossing Heterointerfaces,"*J. Vac. Sci. Technol.* **B 8**, 923 (1990).

5. Henry K. Harbury and Wolfgang Porod, "Numerical Study of Electronic States in a Quantum Wire at Crossing Heterointerfaces," *Computational Electronics, Semiconductor Transport and Device Simulation*, ed. by K. Hess, J. P. Leburton, and U. Ravaioli (Kluwer Academic Publishers, 1990), p. 243.

6. S. Bandyopadhyay and W. Porod, "A Proposed Aharonov-Bohm Interferometer for Possible LN_2 Temperature Operation," *Superlattices and Microstructures* **5**, 239 (1989).

7. S. Bandyopadhyay and W. Porod, "Performance of Electrostatic Aharonov-Bohm Interferometers in the Diffusive Regime," *Appl. Phys. Lett.*, **53**, 2323 (1988).

8. M. Cahay, S. Bandyopadhyay and H. L. Grubin, "Doubled Frequency of Conductance Minima in Elecrostatic Aharonov-Bohm Oscillations in One-dimensional Rings," *Nanostructure Physics and Fabrication*, eds. M. A. Reed and W. P. Kirk (Academic Press, Boston, 1989) p. 407.

9. M. Cahay, S. Bandyopadhyay and H. L. Grubin, "Twin Conductance Minima in Electrostatic Aharonov-Bohm Oscillations," *Phys. Rev. B.* **39** 12989 (1989).

10. S. Bandyopadhyay, "Fluctuations in the Optical Spectra of Disordered Microstructures due to Quantum Interference Effects," *Phys. Rev. B15* **38**, 7466 (1988).

11. D. R. Poole and S. Bandyopadhyay, "Quantum Interference Effects in Dissipationless and Diffusive Transient Transport," *J. Appl. Phys.* **66**, 5422 (1989).

12. S. Bandyopadhyay, "Phase-coherent Effects in the Photoluminescence Spectra of Disordered Mesoscopic Structures," *Nanostructure Physics and Fabrication*, eds. M. A. Reed and W. P. Kirk (Academic Press, Boston, 1989) p. 117.

13. Craig S. Lent, Lie Liang and Wolfgang Porod, "Escape from Quantum Wells by Polar Optical Phonon Scattering," *Appl. Phys. Lett.* **54**, 2315 (1989).

14. Craig S. Lent and Lie Liang, "Real Space Transfer Rates for Polar Optical Phonon Scattering from Asymmetric Quantum Wells," *Solid State Electronics* **32**, 1479 (1989).

15. Lie Liang and Craig S. Lent, "Escape from Quantum Wells via Polar Optical Phonon Scattering," *J. Appl. Phys.* **68** 1741 (1990).

16. D.J. Kirkner, Craig S. Lent, and Srinivas Sivaprakasam, "Numerical Simulation of Single-Electron Transmission Through Two Dimensional Quantum Device by the Finite Element Method," *International Journal for Numerical Methods in Engineering* **29** 1527 (1990).

17. Craig S. Lent and D. J. Kirkner, "The Quantum Transmitting Boundary Method," *J. Appl. Phys.* **67**, 6353 (1990).

18. Craig S. Lent, Srinivas Sivaprakasam and D. J. Kirkner, "A Two-dimensional Hot Carrier Injector for Electron Waveguide Structures," *Solid State Electronics* **32**, 1137 (1989).

19. Craig S. Lent, S. Sivaprakasam and D. J. Kirkner, "Calculation of Ballistic Transport in Two-dimensional Quantum Structures using the Finite Element Method," *Nanostructure Physics and Fabrication*, eds. M. A. Reed and W. P. Kirk, (Academic Press, Boston, 1989), p. 279.

20. M. Cahay and S. Bandyopadhyay, "Scattering Matrix Analysis of Electron Transport in Disordered Aharonov-Bohm Interferometers and Ballistic Constrictions," *J. Vac. Sci. Technol.* **8**, 1399 (1990)..

21. Craig S. Lent, "Transmission Through a Bend in a Quantum Waveguide," *Appl. Phys. Lett.* **56** 2554 (1990).

22. Craig S. Lent and S. Sivaprakasam, "Quantum Electron Waveguides: Bends, Constrictions, and Cavities," *Nanostructures and Microstructure Correlation with Physical Properties of Semiconductors*, SPIE Proceedings Vol. 1284, ed. by H.G. Craighead and J.M. Gibson, (SPIE, Bellingham, Washington, 1990), p. 31.

23. Craig S. Lent, "Calculation of Transport Through Ballistic Quantum Structures," *Computational Electronics*, eds. K. Hess, J. P. Leburton, and U. Ravaioli (Kluwer, Boston, 1990), p. 259.

24. Craig S. Lent, "Ballistic Current Vortex Excitations in Electron Waveguide Structures," *Appl. Phys. Lett.* **57**, 1678 (1990).

25. M. Cahay, S. Bandyopadhyay, M. A. Osman and H. L. Grubin, "Influence of Evanescent Modes on the Conductance of Quasi 1-D Mesoscopic Samples," *Surface Science* **228**, 301 (1990).

26. S. Bandyopadhyay and M. Cahay, "The Generalized Scattering Matrix Approach: An Efficient technique for Modeling Quantum Transport in Relatively Large and Heavily Doped Structures," *Computational Electronics: Semiconductor Transport and Device Simulation*, eds. K. Hess, J. P. Leburton and U. Ravaioli, (Kluwer Academic Publishers, Boston, Massachusetts, 1990), p. 223.

27. P. Marzolf, M. Cahay, and S. Bandyopadhyay, "Numerical Study of the Higher Order Moments of Conductance Fluctuations," *Computational Electronics: Semiconductor Transport and Device Simulution*, eds. K. Hess, J. P. Leburton and U. Ravaioli, (Kluwer Academic Publishers, Boston, Massachusetts, 1990), p. 263.

28. Craig S. Lent, "Edge States in a Circular Quantum Dot," *Phys. Rev. B* **43**, 4179 (1991).

29. S. Subramaniam, S. Bandyopadhyay and W. Porod, "Analysis of the device performance of Quantum Interference Transistors utilizing ultrasmall semiconductor T-structures," *J. Appl. Phys.* **68**, 4861 (1990)..

30. S. Bandyopadhyay, G. H. Bernstein and W. Porod, "Quantum Devices Based on Phase-coherent Lateral Quantum Transport," *Nanostructure Physics and Fabrication*, eds. M. A. Reed and W. P. Kirk (Academic Press, Boston, 1989) p. 183.

31. M. Cahay and S. Bandyopadhyay, "Properties of the Landauer Resistances of Finite Repeated Structures," *Phys. Rev. B* **42**, 5100 (1990).

4. CONFERENCE PRESENTATIONS

1. Craig S. Lent and Wolfgang Porod, "Polar Optical Phonon Scattering from Quantum Wells," *15th Annual Conference on the Physics and Chemistry of Semiconductor Interfaces*, Asilomar, California, January 1988.

2. S. Bandyopadhyay and W. Porod, "Double Quantum Wire Aharonov-Bohm Interferometers for LN_2 Temperature Operation," *Fourth International Conference on Superlattices, Microstructures and Microdevices*, Trieste, Italy, August, 1988.

3. S. Bhobe, W. Porod and S. Bandyopadhyay, "Velocity Modulation Transistor," *Midwest Solid-State Theory Conference*, Cincinnati, Ohio, October 1988.

4. S. Bandyopadhyay, "Fluctuations in the Photoluminescence Spectra of Semiconductor Mesoscopic Structures due to Quantum Phase Coherence," *International Symposium on Nanostructure Physics and Fabrication*, College Station, Texas, March 1989.

5. M. Cahay, S. Bandyopadhyay and H. L. Grubin, "Prediction of an Additional Conductance Minimum due to the Electrostatic Aharonov-Bohm Effect in One-dimensional Doubly-Connected Structures," *International Symposium on Nanostructure Physics and Fabrication*, College Station, Texas, March 1989.

6. Craig S. Lent, S. Sivaprakasam and D.J. Kirkner, "Calculation of Ballistic Transport in Two-dimensional Quantum Structures Using the Finite Element Method," *International Symposium on Nanostructure Physics and Fabrication*, College Station, Texas, March 1989.

7. S. Bandyopadhyay, G. H. Bernstein and W. Porod, "Phase-coherent Lateral Quantum Transport Phenomena in Semiconductor Nanostructures," *International Symposium on Nanostructure Physics and Fabrication*, College Station, Texas, March 1989.

8. S. Bhobe, W. Porod and S. Bandyopadhyay, "Selective Doping in Quantum Wells and its Device Applications," *International Symposium on Nanostructure Physics and Fabrication*, College Station, Texas, March 1989.

9. S. Bandyopadhyay, "Fluctuations in the Optical Spectra of Semiconductor Mesoscopic Structures," *Spring Meeting of the American Physical Society*, St. Louis, Missouri, March 1989.

10. M. Cahay and S. Bandyopadhyay, "Two Minima in the Conductance Modulation of Electrostatic Aharonov-Bohm Interferometers," *Spring Meeting of the American Physical Society*, St. Louis, Missouri, March 1989.

11. W. Porod and S. Bandyopadhyay, "Performance of 2-d and 1-d Electrostatic Aharonov-Bohm Interferometers in the Presence of Elastic Scattering," *Spring Meeting of the American Physical Society*, St. Louis, Missouri, March 1989.

12. D. R. Poole and S. Bandyopadhyay, "Phase-coherent Effects in Dissipationless and Diffusive Transient Transport," *Spring Meeting of the American Physical Society*, St. Louis, Missouri, March 1989.

13. S. Bhobe, W. Porod and S. Bandyopadhyay, "Device Applications of Modulated Interface Scattering," *Spring Meeting of the American Physical Society*, St. Louis, Missouri, March 1989.

14. Craig S. Lent,S. Sivaprakasam and D.J. Kirkner, "A Two-dimensional Hot Carrier Injector for Electron Waveguide Structures," *Sixth International Conference on Hot Carriers in Semiconductors*, Scottsdale, Arizona, July 1989.

15. S. Bhobe, W. Porod and S. Bandyopadhyay, "Modulation of Electronic Mobility in Quantum Wells by Wavefunction Engineering and its Transistor Applications," *Sixth International Conference on Hot Carriers*, Scottsdale, Arizona, July 1989.

16. Craig S. Lent and Lie Liang, "Real Space Transfer Rates for Polar Optical Phonon Scattering from Asymmetric Quantum Wells," *Sixth International Conference on Hot Carriers in Semiconductors*, Scottsdale, Arizona, July 1989.

17. M. Cahay, S. Bandyopadhyay, M. A. Osman and H. L. Grubin, "The Role of Evanescent Channels in Diffusive Quantum Transport," *Fourth International Conference on Modulated Semiconductor Structures*, Ann Arbor, Michigan, July 1989.

18. S. Bhobe, W. Porod and S. Bandyopadhyay, "Configuration-dependent Impurity Scattering Rates in Quasi 2-D Electronic Systems," *Fourth International Conference on Modulated Semiconductor Structures*, Ann Arbor, Michigan, July 1989.

19. Henry Harbury and Wolfgang Porod, "A Proposed Quantum Wire Structure: An 'Accumulation Wire' at Crossing Heterointerfaces," *17th International Conference on the Physics and Chemistry of Semiconductor Interfaces*, Clearwater Beach, Florida, January/February 1990.

20. Harold P. Hjalmarson and Craig S. Lent, "Qualitative Electronic Properties of Interface Roughness," *American Physical Society March Meeting*, Anaheim, California, March 1990.

21. Craig S. Lent, "Quantum Electron Waveguide: Bends, Constrictions, and Cavities", *SPIE Nanostructures Conference*, San Diego, California, March 1990.

22. M. Cahay and S. Bandyopadhyay, "Scattering matrix analysis of electron transmission through quantum point contacts," *34th. International Symposium on Electron, Ion and Photon Beams*, San Antonio, Texas, March 1990.

23. S. Bandyopadhyay and M. Cahay, "The Influence of Evanescent States on Transport in Disordered Mesoscopic Systems," *Spring Meeting of the American Physical Society*, Anaheim, California, March 1990.

24. S. Subramaniam, S. Bandyopadhyay and W. Porod, "Analysis of Quantum Transistors Utilizing Mesoscopic Electron Waveguides," *Spring Meeting of the American Physical Society*, Anaheim, California, March 1990.

25. S. Bandyopadhyay and D. R. Poole, "Fluctuations in the Electronic Mobility at Disordered Semiconductor-Insulator Interfaces due to Variations in Configuration of Defects at the Interface," *Spring Meeting of the American Physical Society*, Anaheim, California, March 1990.

26. M.-H. Tsai, J. D. Dow, S. Bandyopadhyay and R. V. Kasowski, "Conduction Band Mediated Superexchange in FeTe, CoTe and NiTe," *Spring Meeting of the American Physical Society*, Anaheim, California, March 1990.

27. H. K. Harbury and W. Porod, "Proposal of an 'Accumulation Wire' Formed by Crossing Heterointerfaces,"*Spring Meeting of the American Physical Society*, Anaheim, California, March 1990.

28. Craig S. Lent,"Calculation of Transport Through Ballistic Quantum Structures," *Workshop on Computational Electronics*, Urbana, Illinois, May 1990.

29. Henry Harbury and Wolfgang Porod, "Numerical Study of Electronic States in a Quantum Wire at Crossing Heterointerfaces," *Workshop on Computational Electronics*, Urbana, Illinois, May 1990.

30. S. Bandyopadhyay and M. Cahay, "The Generalized Scattering Matrix Method: An Efficient Technique for Modeling Quantum Transport in Large and Heavily Doped Structures," *Workshop on Computational Electronics*, Urbana, Illinois, May 1990.

31. P. Marzolf, M. Cahay, and S. Bandyopadhyay, "Higher-Order Cumulants of Conductance Fluctuations in Mesoscopic Structures," *Workshop on Computational Electronics*, Urbana, Illinois, May 1990.

32. S. Subramaniam, S. Bandyopadhyay and W. Porod, "Electron Waveguide Transistors," *Fifth International Conference on Superlattices and Microstructures*, Berlin, German Democratic Republic, August 1990.

33. S. Bandyopadhyay, M. Cahay, D. Berman and B. Nayfeh, "The Role of Unoccupied Subbands in Quantum Conduction Through Disordered Mesoscopic Structures," *Fifth International Conference on Superlattices and Microstructures*, Berlin, German Democratic Republic, August 1990.

34. Manhua Leng and Craig S. Lent, "Bloch electrons in a Periodically Modulated Electron Waveguide," *Midwest Solid State Symposium*, Evanston, Illinois, October 1990.

5. PERSONNEL

Principal Investigators: Supriyo Bandyopadhyay, Craig S. Lent, and Wolfgang Porod

Graduate Students:

- Suyog Bhobe, Research Assistant, M.S. (1989), Thesis: "Analysis and Simulation of Velocity Modulation."
- Lie Liang, Research Assistant, M.S. (1989), Thesis: "Polar Optical Scattering from Quantum Wells," currently in Ph.D. program.
- Srinivas Sivaprakasam, Research Assistant, M.S. (1990), Thesis: 'Finite Element Simulation of Quantum Devices."
- Suresh Subramaniam, Research Assistant, M.S. (1990), Thesis: "Low Temperature Measurements of Quantum Interference Effects in Nanostructures," currently in Ph.D. program.
- Henry Harbury, Research Assistant, M.S. (1990), Thesis: "Proposed Quantum Wire Structure at Crossing Heterointerfaces, " currently in Ph.D. program.
- Manhua Leng, Research Assistant.
- Santanu Chaudhuri, Research Assistant.

APPENDIX

Reprints

SURFACE AND INTERFACE ANALYSIS, VOL. 14, 590–594 (1989)

Modulated Interfacial Disorder Scattering in Quantum Wells and its Device Applications

Suyog Bhobe, Wolfgang Porod and Supriyo Bandyopadhyay
Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, IN 46556, USA

David J. Kirkner
Department of Civil Engineering, University of Notre Dame, Notre Dame, IN 46556, USA

We investigate the nature of the electronic states and transient transport in a single, selectively doped, GaAs quantum well where the dopants are confined near one of the interfaces. The electronic states are calculated from self-consistent solutions of the Schrödinger and Poisson equations and are used to compute the scattering rates for electrons interacting with the dopants. These scattering rates are then used in a Monte Carlo simulation to study the time-dependent decay of the momentum of an ensemble of electrons traveling parallel to the interface.

We also investigate the momentum relaxation of electrons injected into the well under different transverse electric fields that skew the wavefunction towards one of the interfaces. When the wavefunction is skewed away from the doped interface, the momentum relaxation time (and hence the mobility) increases dramatically by three orders of magnitude. A device application of this phenomenon, namely the operation of a velocity modulation transistor (VMT), has been investigated, with special focus on the transistor's switching speed.

INTRODUCTION

The progress of microelectronics and computer technology relies heavily on the continued development of faster and smaller electronic devices. Over the past decade, electronic devices have shrunk rapidly in size and have concomitantly become faster, but this trend seems to be saturating. It is now realized that classical devices that operate on the principles of classical physics, have either reached their limits or are about to reach them, whereupon entirely new concepts are required for the next generation of devices. Future devices are expected to rely on quantum-mechanical principles for their operation and utilize quantum-confined structures such as semiconductor quantum wells fabricated by molecular beam epitaxy or by metal–organic chemical vapor deposition.

One such device that has been proposed in the literature is the velocity modulation transistor (VMT),[1] which operates on the basis of 'wavefunction engineering' in a semiconductor quantum well. A velocity modulation transistor is an ultrafast switching device whose switching time can be smaller than 1 ps. The concept that undergirds the operation of this transistor is the following. Any field effect transistor (FET) is switched between the 'on' and 'off' states by modulating the conductance of the transistor. The conductance, in $(\Omega/\square)^{-1}$, is given by

$$G = en_s\mu \qquad (1)$$

where e is the electronic charge, n_s is the carrier concentration in the two-dimensional conducting channel (such as the inversion layer in a conventional MOSFET) and μ is the carrier mobility.

In a conventional FET, the conductance is modulated by modulating the carrier concentration n_s. A transverse electric field, applied perpendicular to the plane of the conducting channel, either depletes the channel (and the transistor is switched off) or accumulates the channel, in which case the transistor is turned on. The problem with this mode of operation is that the transistor cannot be switched any faster than the time it takes to move carriers in and out of the channel, which is typically the transit time of carriers from one contact (termed 'source') to the other (termed 'drain'). An alternative way to realize the switching is to modulate the channel conductance by modulating the mobility μ instead of the concentration n_s. The mobility can change on timescales of the order of the momentum relaxation time, which is typically much shorter than the transit time. Thus, the latter mode of switching is much faster than the former.

The VMT employs the latter mode of switching. The mobility is modulated by *engineering* the electronic wavefunction inside a quantum well which acts as the channel. A basic schematic of the VMT is shown in Fig. 1. The quantum well is selectively doped so that one-half of the well is interspersed with impurities while the other half is undoped.

In the absence of any external transverse electric field, the electronic wavefunction is skewed towards the doped region in the well where the potential is lower due to space–charge effects. The application of a transverse electric field (or gate voltage), applied through the gate terminal in Fig. 1, can lower the potential on the undoped side of the well so that the electrons move towards the undoped side where the mobility is higher due to the absence of *in situ* impurity scattering. The gate voltage therefore modulates the mobility of the electrons by wavefunction engineering. The transistor is in the low conductance state (or 'off' state) when the gate voltage is zero, and switches on when the gate voltage is turned on. It shall be emphasized that the 'on'

0142-2421/89/100590-05 $05.00
© 1989 by John Wiley & Sons, Ltd.

Received 22 November 1988
Accepted 26 January 1989

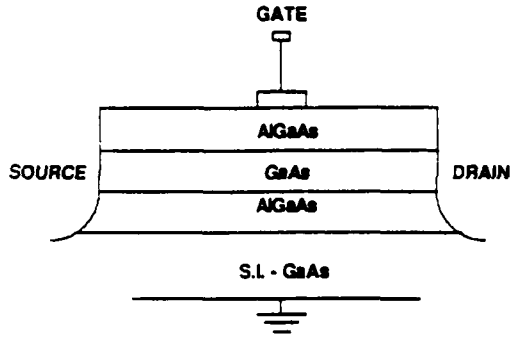


Figure 1. Schematic diagram of the velocity modulation transistor. Only one-half of the GaAs quantum well is selectively doped. Current transport takes place parallel to the heterointerfaces (between the source and drain contacts).

and 'off' states here refer to high and low conductance states, as opposed to the customary label for states with finite and zero current flows. (The wavefunctions inside the well in the 'off' and 'on' states are depicted in Figs 2 and 3, respectively).

The type of VMT just discussed was proposed by Hamaguchi and co-workers.[2,3] It is slightly different from the original proposal of Sakaki,[1] where the mobility was to be modulated by pushing the wavefunction closer to the well interface with a transverse electric field. This increases the interface roughness scattering and decreases the mobility. The problem with the original design of Ref. 1 is that the transistor is 'normally on' (at zero gate voltage), which causes stand-by power dissipation in a circuit, whereas in the present design the transistor is 'normally off' so that there is less stand-by power dissipation. The present design is therefore analogous to an enhancement mode FET, whereas the original design is analogous to a depletion mode FET. Note also that in the original design, the channel current decreases with increasing gate voltage so that the transconductance of the transistor is *negative* whereas in the present design the transconductance is *positive*. The opposite sign of the transconductance in the two cases presents the intriguing possibility of using the two designs to realize ultrafast complementary logic.

In the next section we outline our model for analysis of the VMT, and then we present the computational scheme for calculating the self-consistent electronic wavefunctions.

THEORY

The electronic states of a quasi two-dimensional system, such as a quantum well, have been studied widely in the literature.[4–7] Following the usual treatment, we separate the electronic motion parallel to the interface from the motion normal to the interface. As long as the effective mass of the electrons in the well is spatially invariant, the parallel motion is decoupled from the perpendicular motion. We label the direction perpendicular to the interfaces as the $\hat{z}$ direction and the plane parallel to the interfaces as the $\hat{x}$–$\hat{y}$ plane. Since the Hamiltonian is invariant in $\hat{x}$ and $\hat{y}$, the wavevectors k_x and k_y are 'good quantum numbers', so that the wavefunction is given by

$$\psi(x, y, z) = \exp(ik_x x)\exp(ik_y y)\psi_m(z) \tag{2}$$

The eigenstates in the $\hat{x}$ and $\hat{y}$ directions are plane wave states labeled by the wavevector quantum numbers k_x and k_y, while the eigenstates in the $\hat{z}$ direction are quantized sub-band states denoted by $\psi_m(z)$, where m is the sub-band index. The normalized envelope functions for the mth sub-band can be obtained from the Schrödinger equation

$$-\frac{\hbar^2}{2m^*}\frac{d^2\psi_m(z)}{dz^2} + [-e\phi(z) + V(z)]\psi_m(z) = E_m\psi_m(z) \tag{3}$$

The potential $V(z)$ accounts for all external fields, as well as any band offset. The potential $\phi(z)$ is the electrostatic potential given by the solution of the one-dimensional Poisson equation

$$\frac{d^2\phi(z)}{dz^2} = \frac{e}{\varepsilon_0\varepsilon}\left[\sum_{m=1}^{M} N_m|\psi_m(z)|^2 - N_D^+(z)\right] \tag{4}$$

The concentration $N_D^+(z)$ (in units of per unit volume) denotes the density of ionized background impurities. Each sub-band with energy E_m contributes to the total electron concentration with N_m electrons per unit area given by

$$N_m = \frac{m^* k_B T}{\pi\hbar^2}\ln\left[1 + \exp\left(\frac{E_F - E_m}{k_B T}\right)\right] \tag{5}$$

The Fermi energy is denoted by E_F and we determine it here such that the total charge inside the quantum well is zero (space charge neutrality).

NUMERICAL METHOD

The one-dimensional Schrödinger and Poisson equations were discretized using the finite element method (FEM).[8] This technique has recently been applied to the numerical solution of quantum mechanical problems.[9] Linear basis functions were used to interpolate the wavefunction and space charge potential between nodal values. For the purpose of analysis, a 500-Å quantum well was considered. The number of nodal points in the quantum well was taken to be 2000. Sufficiently converged solutions were obtained with this number of nodal points. It was assumed that the right half of the quantum well was doped by donor impurities, so that all the disorder was confined to near the right interface of the well rather than being distributed uniformly throughout the well. In all our analyses, the ambient temperature was assumed to be 4.2 K.

The eigenvalues E_m and the eigenfunctions $\psi_m(z)$ were first obtained by solving the Schrödinger equation. For this, the quantum well was assumed to be infinitely deep, and homogeneous boundary conditions were imposed on the eigenfunction. The assumption of an infinitely deep well greatly simplifies the numerical calculation, and results in accurate eigenvalues for energies which are not too close to the edge of the quantum well.[3] The eigenfunctions thus obtained were used in the Poisson equation to yield the space charge potential. To

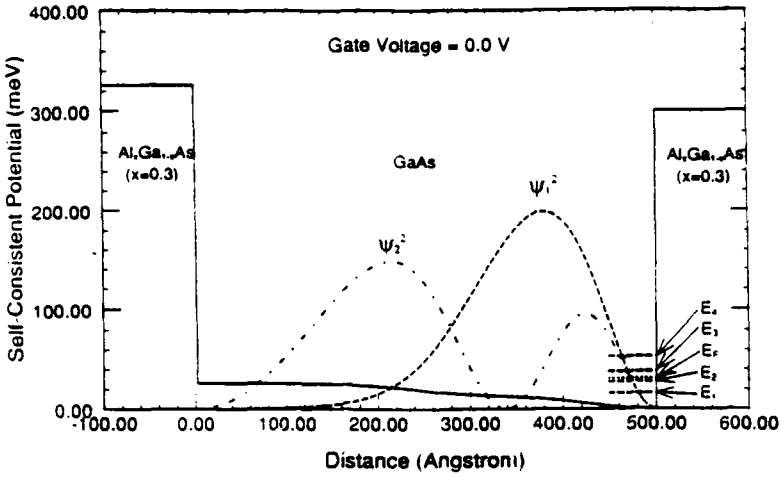


Figure 2. Self-consistent results of the quasi two-dimensional electron gas inside the quantum well for zero external bias. Shown are the potential profile, the sub-band energies and the square of the corresponding wavefunctions.

solve the Poisson equation, the space charge potential and its slope were assumed to be zero at the right end of the quantum well. This boundary condition reflects overall charge neutrality for the quantum well.

The following procedure was used to solve iteratively the coupled system of Schrödinger's and Poisson's equations. Let ϕ_S^n be the potential input to the Schrödinger equation in the nth iteration. The wavefunctions obtained by solving the Schrödinger equation were used to calculate the charge density on the right-hand side of the Poisson equation. The solution of the Poisson equation in the nth iteration is denoted by ϕ_P^n. The next update for the potential to be used in the $(n+1)$th iteration in the Schrödinger equation is then obtained from

$$\phi_S^{n+1} = \phi_S^n + (\phi_P^n - \phi_S^n) \cdot \alpha \tag{6}$$

where α is the so-called convergence factor, which is taken from the interval 0.0–1.0. A high α value can give rapid convergence, but may cause oscillations. A low α value, on the other hand, reduces the rate of convergence but eliminates such oscillations. The iterations were continued until convergence was achieved. Convergence was deemed satisfactory when $(\phi_S^{n+1} - \phi_S^n)/\phi_S^n < \varepsilon$, where ε was set to 0.0001. A lower value of ε did not result in any significant increase in accuracy.

The squares of the wavefunctions were integrated over the doped region in the quantum well to obtain the 'effective' number of impurities, N_{eff}, that interact *in situ* with the electrons

$$N_{eff} = \sum_m \int N_D(x) |\psi_m(x)|^2 \, dx \tag{7}$$

The above integration is performed over the selectively doped regions of the quantum well. This 'effective' number was calculated for various gate voltages that skew the wavefunctions by different amounts. From the 'effective' number, the scattering rates for impurity interaction,[10] $1/\tau(E)$, were calculated using Fermi's Golden Rule

$$\frac{1}{\tau(k)} = \frac{N_{eff} m^* e^4}{4\hbar^3 \varepsilon^2 \lambda \sqrt{\lambda^2 + 4k^2}} \tag{8}$$

where k is the Fermi wavevector given by $k = \sqrt{2\pi n_s}$, where n_s is the two-dimensional electron concentration, m^* is the effective mass of electrons in GaAs, ε is the dielectric constant and λ is the screening constant. Since the quantum well in our model is relatively wide, we used the bulk Debye screening model to calculate λ

$$\lambda = e\sqrt{\frac{N_D}{\varepsilon k T}} \tag{9}$$

The screening length $1/\lambda$ turns out to be *ca.* 11 Å in our case. Because of such strong screening (which is caused by the heavy doping), the scattering rates are rather small, typically between 2×10^7 and 8×10^9 s^{-1}. The strong screening, however, allows us to neglect the effects of remote impurity scattering when the wavefunction is skewed away from the doped region of the well.

The scattering rates obtained from Eqns (7), (8) and (9) were then used in a Monte Carlo simulation[11] to extract the momentum decay characteristics of an ensemble of electrons injected into the quantum well with the Fermi velocity. The Monte Carlo simulation was used to predict the momentum relaxation rate for various gate voltages. In the simulation, we have considered only *in situ* impurity scattering and neglected remote impurity scattering because of the strong screening. We have also neglected phonon scattering based on the premise that at 4.2 K, phonon interactions are significantly weaker than impurity interactions. In addition, since there is no electric field in the direction of electron motion, the electrons never gain enough energy to spontaneously emit polar optical phonons. Thus, there is no dissipation in our system. The simulations were run for sufficiently long times to ensure that the momentum relaxes to zero before the simulation is terminated. Each simulation was performed for an ensemble of 10000 electrons to obtain statistically reliable estimates.

RESULTS

We have performed calculations for a quantum well that is 500 Å wide. The right half of the well, from 250 to 500 Å, is selectively doped with donors. We chose a

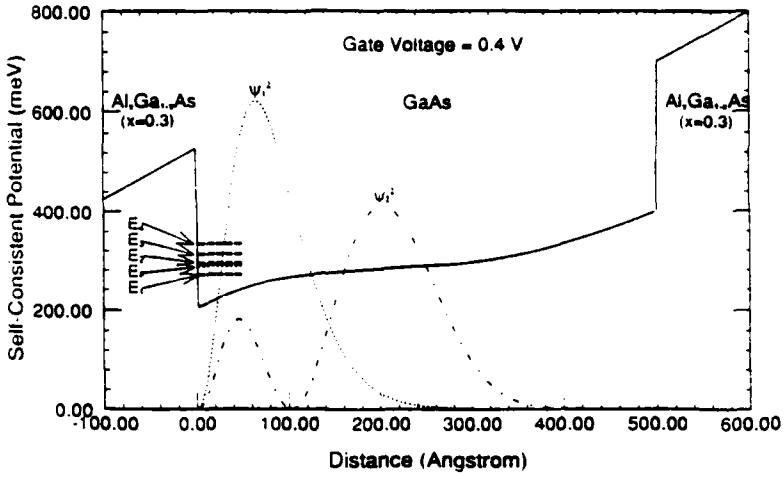


Figure 3. Self-consistent results of the quasi two-dimensional electron gas inside the quantum well for an external bias of 0.4 V. Shown are the potential profile, the sub-band energies and the square of the corresponding wavefunctions.

doping concentration of $N_D = 2 \times 10^{23}$ m^{-3}. This corresponds to a density of $n_s = 5 \times 10^{15}$ m^{-2} of the quasi two-dimensional electron gas.

In our calculations, we assume that all donors are ionized, i.e. that they contribute to the positive background charge in the right half of the quantum well. This is a reasonable approximation for very shallow-level donors at 4.2 K. All electrons are assumed to remain in the quantum well and not to spill over to the barriers, resulting in a constant density of the quasi two-dimensional electron gas.

We show typical results of our calculations in the following figures. Figure 2 shows the potential profile inside the quantum well, with the corresponding wavefunctions for zero gate bias. Figure 3 displays the same data for a gate bias of 0.4 V. All results are valid for a temperature of 4.2 K.

We see that by increasing the bias, the wavefunctions are indeed pushed to one side of the well—the undoped side in this case—which will increase the overall mobility. The variation of the Fermi level and sub-band energies with gate bias is also shown in Fig. 4. Note that all energies are measured with respect to the potential energy at the right edge of the quantum well, which is the energy associated with the unscreened externally applied gate bias (cf. also Figs 2 and 3).

Table 1 shows the rates for ionized impurity scattering for various gate biases. These rates were calculated from Eqns (7), (8) and (9). Note that the variation in the scattering rates spans almost three orders of magnitude. This shows the extreme sensitivity of the scattering rate to the gate bias.

The scattering rates given in Table 1 were used in a Monte Carlo simulation of electron transport to extract the momentum decay characteristics and the momentum relaxation time. Figure 5 shows the various momentum decay characteristics for various gate biases. The momentum relaxation time is defined as the time that elapses before the momentum decays to $1/e$ times its initial value. From the momentum relaxation time, we can calculate the effective 'mobility' in the quantum

Table 1. Scattering rates vs. gate voltage

Gate voltage (V)	Scattering rate (s^{-1})
0.0	8×10^9
0.1	6×10^9
0.2	1×10^9
0.3	2×10^8
0.4	2×10^7

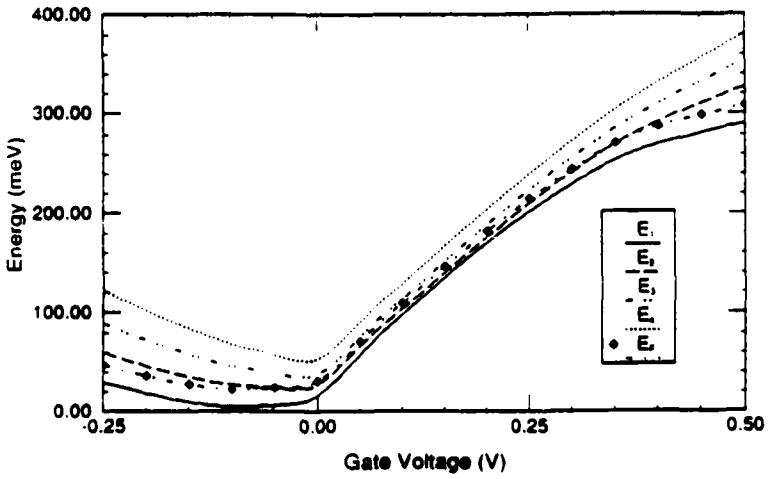


Figure 4. Variation of the sub-band energies with gate bias. The variation of the Fermi level, corresponding to charge neutrality, is also shown.

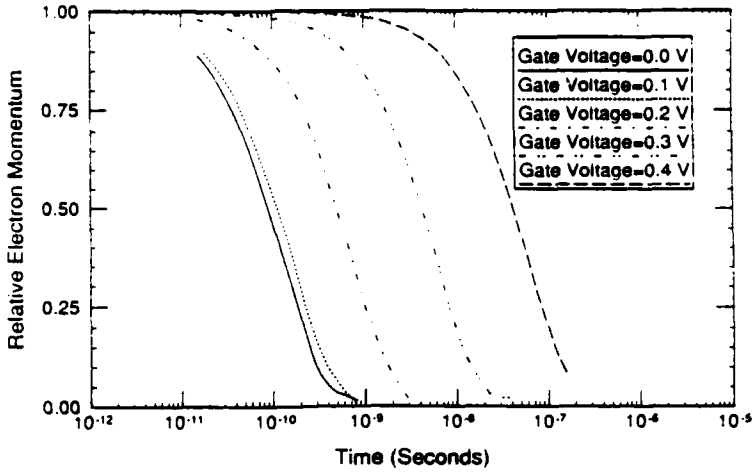


Figure 5. Momentum decay for an ensemble of electrons with initially aligned momenta. Note the large variation in the momentum decay times upon changes of the gate bias.

well by using the well-known formula

$$\mu = \frac{e\tau_m}{m^*} \tag{10}$$

Once the mobility is known, the conductance of the VMT can be calculated from Eqn (1). Table 2 lists the conductances for various gate biases. Note the conductance can be modulated over approximately three orders of magnitude (which gives a peak-to-valley ratio of 1000 for the current through the VMT) by varying the gate bias over a range of only 400 mV. This modulation can be achieved typically on timescales of the order of the momentum relaxation time. In our simulation, the momentum relaxation time is rather long, of the order of 10^{-10} s, which is caused by the fact that the impurity scattering is strongly screened by the rather large electron concentration. For ultrafast switching, the conductance modulation must be achieved in timescales of the order of a picosecond or less. Thus, our results indicate that, for optimum performance, the quantum well must be relatively lightly doped so that the carrier concentration (and hence the screening) is much lower.

Table 2. Conductance *vs.* gate voltage

Gate voltage (V)	Conductance $(\Omega/\square)^{-1}$
0.0	0.255
0.1	0.637
0.2	2.126
0.3	19.13
0.4	191.31

CONCLUSION

We have investigated electronic transport in a quantum well using the detailed nature of the electronic wavefunctions in the well. The results are self-consistent insofar as they account for space–charge effects. We have specifically investigated an intriguing device application of wavefunction engineering in a quantum well, namely the operation of a velocity modulation transistor. We found that such a transistor can be switched by a relatively small gate voltage of 400 mV to yield a peak-to-valley ratio of the 'on' and 'off' conductance approaching a factor of 1000. This gives rise to an extremely large transconductance, which translates into an extremely large frequency range of operation for the VMT that could be orders of magnitude larger than those achievable with conventional FETs.

Acknowledgement

This work was supported by the Air Force Office of Scientific Research, under Grant No. AFOSR-88-0096.

REFERENCES

1. H. Sakaki, *Jpn. J. Appl. Phys.* **21**, 381 (1982).
2. C. Hamaguchi, K. Miyatsuji and H. Hihara, *Jpn. J. Appl. Phys.* **23**, L132 (1984).
3. K. Miyatsuji, H. Hihara and C. Hamaguchi, *Superlattices Microstructures* **1**, 43 (1985).
4. F. Stern, *Phys. Rev. Sect. B* **5**, 4891 (1972).
5. F. Stern and S. Das Sarma, *Phys. Rev. Sect. B* **30**, 840 (1984).
6. K. Yokoyama and K. Hess, *Phys. Rev. Sect. B* **31**, 6872 (1985).
7. S. E. Laux, D. J. Frank and F. Stern, *Surf. Sci.* **196**, 101 (1988).
8. L. J. Segerling, *Applied Finite Element Analysis*. Wiley, New York (1984).
9. D. J. Searles and E. I. von Nagy-Felsobuki, *Am. J. Phys.* **56**, 444 (1988).
10. E. M. Conwell, *High Field Transport in Semiconductors*, Solid State Phys., Suppl. 9, Academic Press, New York (1967).
11. C. Jacoboni and L. Reggiani, *Rev. Mod. Phys.* **55**, 645 (1983).

Solid-State Electronics Vol. 32, No. 12, pp. 1651–1655, 1989
Printed in Great Britain. All rights reserved

0038-1101/89 $3.00 + 0.00
Copyright © 1989 Pergamon Press plc

MODULATION OF IMPURITY SCATTERING RATES BY WAVEFUNCTION ENGINEERING IN QUASI 2-D SYSTEMS AND ITS DEVICE APPLICATIONS

S. Bhobe, W. Porod, and S. Bandyopadhyay

Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, IN 46556

ABSTRACT

We investigate impurity - scattering - limited electronic transport in a quasi 2-d electronic system. For the calculation of the impurity scattering rates, we account for the self - consistent dependence of the electronic states on the precise configuration of the scattering centers. Based on this, we investigate the physical mechanism of mobility enhancement in delta-doped semiconductor structures.

KEYWORDS

Delta Doping; Impurity Scattering; Wavefunction Engineering; Monte Carlo

INTRODUCTION

Recent advances in epitaxial growth techniques have made it possible to dope semiconductor microstructures selectively. A major advantage of selective doping, such as delta-doping, is that the mobility of the resulting two-dimensional electron gas can be much higher than what can be attained with uniform doping (Gillman et al., 1988; Schubert et al., 1989; Hong et al., 1989). Recently, it was pointed out that spatial ordering of the dopant atoms in the plane of the delta dopants can also lead to a further enhancement of the mobility (Levi et al., 1989).

In this paper, we investigate the physical mechanism of mobility enhancement in delta - doped structures. Our formalism is based on a quantum-mechanical analysis in which we first evaluate the electronic states in a quantum well sel -consistently by solving the Schrödinger-Poisson equations. From the electronic states, we calculate the *exact* matrix element for impurity interaction and from this the scattering rate, or alternatively the impurity-scattering-limited mobility. We then use the scattering rate to perform a Monte Carlo simulation of the momentum decay of electrons in the quantum well. The momentum relaxation time is estimated from the Monte Carlo simulation from which the mobility and the conductance can be obtained.

The mobility depends on the precise configuration of the impurities (scattering centers) within the structure for two reasons. Firstly, the electronic wavefunctions depend on the exact *shape* of the self-consistent potential (produced by the space-charges) which in turn depends on the precise locations of the impurities within the structure. The wavefunctions determine the matrix element so that the scattering rate has an indirect dependence on the impurity coordinates. Secondly, the Coulombic interaction potential, that appears in the matrix element, is itself an explicit function of the impurity coordinates. This is a direct effect which makes the mobility strongly configuration-dependent. As a result, the mobility in a structure can be altered significantly by tailoring the positions of the dopants within the structure.

The so-called velocity modulation transistor (VMT) is a device designed to exploit the modulation of the mobility by wavefunction engineering (Sakaki, 1982; Hamaguchi et al., 1984). For the study of the switching behavior of the VMT at low temperatures, the impurity scattering rates have to be known as a function of the applied bias and the configuration of the impurities. We present here results for the momentum relaxation times for various gate voltages and various arrangements of delta - doped impurity sheets.

ELECTRONIC STATES

The electronic states of a quasi two-dimensional system have been studied widely in the literature (Stern, 1972; Yokoyama and Hess, 1985). Following the usual treatment, we write the electronic wavefunction, using $\vec{\rho} = (x, y)$,

$$\psi(x,y,z) = e^{i\vec{k}\cdot\vec{\rho}}\psi_m(z) \tag{1}$$

where the normalized envelope function $\psi_m(z)$ for the mth subband can be obtained from the 1-d Schrödinger equation, with z denoting the direction perpendicular to the interfaces.

$$-\frac{\hbar^2}{2m^*}\frac{d^2\psi_m(z)}{dz^2} + [-e\phi(z) + V(z)]\psi_m(z) = E_m\psi_m(z) \tag{2}$$

The potential $V(z)$ in the above equation accounts for all external fields and band offsets, while the potential $\phi(z)$ is the self-consistent space-charge potential obtained from the 1-d Poisson equation.

$$\frac{d^2\phi(z)}{dz^2} = \frac{e}{\epsilon_0\epsilon}[\sum_{m=1}^{M} N_m|\psi_m(z)|^2 - N_D^+(z)] \tag{3}$$

The concentration $N_D^+(z)$ denotes the density of ionized background impurities. The functional dependence of $N_D^+(z)$ on the z-coordinate is determined by the nature of the doping. In the case of delta-doping corresponding to a sheet of charges, $N_D^+(z)$ is a delta-function located at some coordinate z_0. For uniform doping, it is a Heaviside function, and so on. The quantity N_m is the contribution to the total electron concentration (per unit area) from the mth subband. It is given by

$$N_m = \frac{m^* k_B T}{\pi\hbar^2} ln[1 + exp(\frac{E_F - E_m}{k_B T})]. \tag{4}$$

where E_F is the Fermi energy which is determined in every case such that the total charge inside the quantum well is zero (space charge neutrality).

IMPURITY SCATTERING RATES

Once the wavefunctions are determined, we compute the impurity scattering rates as follows. The rate for an electron to be scattered from an initial state $\vec{k}$ to a final state $\vec{k}'$ by ionized impurities is given by Fermi's Golden Rule,

$$1/\tau(\vec{k},\vec{k}') = \frac{2\pi}{\hbar}\sum_m \left|\int e^{i(\vec{k}-\vec{k}')\vec{\rho}}d^2\rho \int |\psi_m(z)|^2 V_{imp}(\rho,z)dz\right|^2 \delta(E_k - E_{k'}) , \tag{5}$$

where $V_{imp}(\rho,z)$ is the impurity interaction potential which is screened Coulombic or Yukawa-type in nature. For a sheet of charge located at z_0, it is given by

$$V_{imp}(\rho,z) = V_{z_0}(\rho,z) = \frac{e^2}{4\pi\epsilon_0\epsilon}\frac{e^{-\lambda\sqrt{\rho^2+(z-z_0)^2}}}{\sqrt{\rho^2+(z-z_0)^2}} , \tag{6}$$

where λ is the screening constant.

It is important to note that in calculating the matrix element in Equation (5), we considered only intra-subband transitions. Impurity scattering can cause inter-subband transitions (without involving a change in the total energy), but the accompanying momentum change is so large that it can only be caused by the short-range (or large wavevector) components of the scattering potential. Since we consider weakly-screened impurities, this is not an important process. Finally, the total rate, $1/\tau(k)$, is obtained by summing $1/\tau(\vec{k},\vec{k}')$ over all final k's,

$$1/\tau(k) = \sum_{\vec{k}'} 1/\tau(\vec{k},\vec{k}'). \tag{7}$$

To study the configuration dependence of the impurity scattering rates, we arrange the constant number of dopants inside the well in several different ways. The case of uniform doping was treated by us in the past (Bhobe and co - workers, 1989) and will be omitted here. In the present study, we arrange the dopants in a number of delta - doped sheets whose total number is denoted by N_δ. The sheet density of impurities at location z_0 is denoted by n_{z_0} which is given by $n_{z_0} = n_S/N_\delta$, where n_S is the total sheet density of dopants in the quasi 2-d system. As one can see, the case of uniform doping is a limiting case corresponding to a large number of doped sheets ($N_\delta \rightarrow \infty$) with correspondingly low sheet doping density.

We now distinguish between two different cases for the calculation of the impurity scattering rates. In one case, which we term the "coherent" case, the scattering potential $V_{imp}(\rho,z)$ is the arithmetic sum of the potentials of the various delta-spikes

$$V_{imp}(\rho, z) = \sum_{i=1}^{N_\delta} V_{z_i}(\rho, z) \tag{8}$$

This corresponds to the situation where the electronic wavefunction is coherently scattered by all delta - doped sheets. This situation will be realized if ordering in the z-direction exists, yielding correlations of the dopant location in adjacent impurity sheets. In the second case, which we term the "incoherent" case, the scattering rate due to all the sheets is the sum of the scattering rates for each individual sheet. This case corresponds to the situation when there is no correlation or phase coherence between scattering from different sheets.

The scattering rates for the two cases are then

$$1/\tau(\vec{k}, \vec{k}')^{coherent} = \frac{2\pi}{\hbar} \sum_m \left| \int e^{i(\vec{k}-\vec{k}')\vec{\rho}} d^2\rho \int |\psi_m(z)|^2 \left[\sum_{i=1}^{N_\delta} V_{z_i}(\rho, z) \right] dz \right|^2 \delta(E_k - E_{k'}) \tag{9}$$

and

$$1/\tau(\vec{k}, \vec{k}')^{incoherent} = \frac{2\pi}{\hbar} \sum_m \sum_{i=1}^{N_\delta} \left| \int e^{i(\vec{k}-\vec{k}')\vec{\rho}} d^2\rho \int |\psi_m(z)|^2 V_{z_i}(\rho, z) dz \right|^2 \delta(E_k - E_{k'}) \tag{10}$$

It is easy to see that in all cases $1/\tau(\vec{k}, \vec{k}')^{coherent} > 1/\tau(\vec{k}, \vec{k}')^{incoherent}$.

Finally, these scaterring rates are used as input to a Monte Carlo computer simulation to study the momentum decay of an ensemble of electrons subjected to impurity scattering.

RESULTS

We investigate the configuration-dependence of the impurity scattering rates in a GaAs-AlGaAs quantum well which is 500 Å wide. The right half of the well, from 250 to 500 Å, is doped with a fixed concentration of donors of $5 \times 10^{15}\ m^{-2}$. For the case of uniform doping, this corresponds to a doping concentration of $N_D = 2 \times 10^{23}\ m^{-3}$. In particular, we consider the cases where this number of dopants are arranged in a discrete number of equally

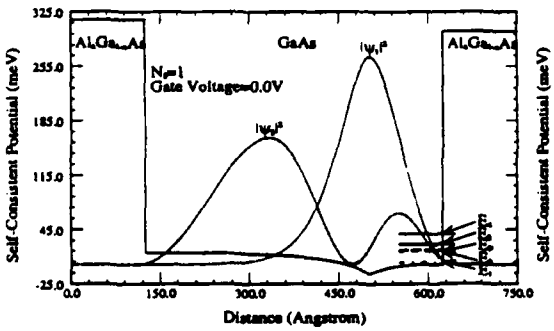


Fig. 1. Shown are the self-consistent potential, the lowest two eigenstates and the corresponding eigenvalues for a single delta-doped sheet, $N_\delta = 1$. No external bias is applied.

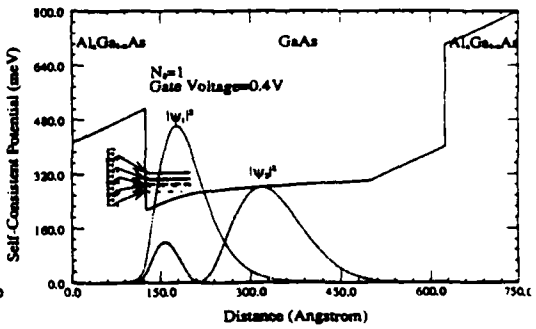


Fig. 2. Shown are the self-consistent potential, the lowest two eigenstates and the corresponding eigenvalues for a single delta-doped sheet, $N_\delta = 1$. An external bias of 0.4 V is applied.

spaced delta-doped sheets, N_δ. We report results for 12 different values of N_δ, namely $N_\delta = 1, \cdots 10, 15, 20$. All our calculations are performed for a temperature of 4.2 K.

We assume that all donors are ionized and that all carriers remain inside the quantum well, thus leaving the total system charge-neutral. This appears justified for shallow donors at 4.2 K. The quasi two-dimensional electron gas, therefore, has a constant density of $n_S = 5 \times 10^{15}\ m^{-2}$, independent of the spatial arrangement of the donors. In Fig. 1 and Fig. 2, we show for two different biases the self-consistent potential and the corresponding wavefunctions for the case of a single delta-doped sheet, $N_\delta = 1$. The location of the delta-doped sheet can be inferred from the cusp in the potential. For the case without external bias, which is depicted in Fig. 1, the electrons are concentrated at the location of the delta-doped sheet because of the Coulombic attraction of the ionized donors. The external bias, as can be seen in Fig. 2, has the effect of pushing the carriers into the left half of the quantum well, which is undoped.

Similar behavior is observed if several sheets of dopant atoms are placed with equal spacings inside the doped,

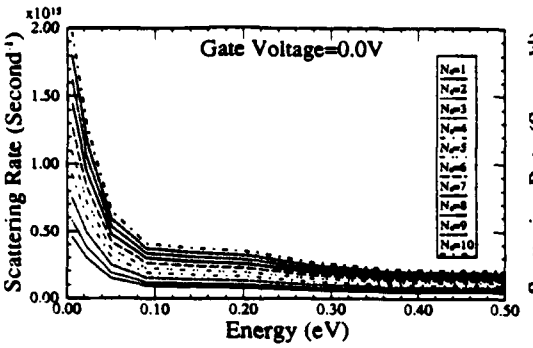


Fig. 3. Shown is the energy dependence of the coherent scattering rate for several numbers of dopant sheets. No external bias is applied.

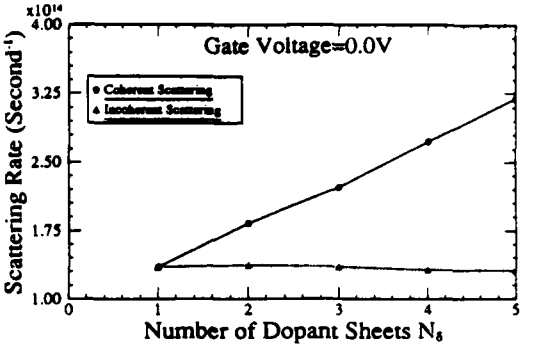


Fig. 4. Shown are the scattering rates for the coherent and for the incoherent cases as a function of the number of dopant sheets, N_δ.

right hand side of the quantum well. We have obtained the self - consistent potentials and wavefunctions for the 12 values of N_δ and for 5 values of external gate bias, namely $V_g = 0$, 0.1, 0.2, 0.3, and 0.4 V. Larger values of the bias would lead to a significant leakage of carriers out of the well because of the smallness of the potential barrier between GaAs and GaAlAs.

These wavefunctions are now used for the calculation of the coherent and incoherent scattering rates for various doping conditions. A screening length of 313 Å is used. In Fig. 3, we show the energy dependence of the coherent impurity scattering rate for an unbiased quantum well. We see that the scattering rate increases as the number of delta-doped sheets increases. A comparison between the coherent and the incoherent rates is shown in Fig. 4.

We see that the coherent rate increases almost linearly as the number of doping sheets N_δ increases. This is due to the fact that the square of the matrix element increases as $\sim N_\delta^2$ whereas the concentration in each sheet decreases as $\sim 1/N_\delta$. Consequently, the scattering rate increases as $\sim N_\delta$. In the incoherent case, the main difference is that the square of the matrix element increases as $\sim N_\delta$ instead of as $\sim N_\delta^2$ so that the rate remains relatively independent of N_δ. It therefore appears that in the presence of coherence, the highest mobility is achieved when $N_\delta = 1$ or when a single sheet, rather than many sheets, is used to dope the layers.

In comparing our results for the coherent and incoherent cases, we see that coherence, which arises from ordering in the z-direction, *increases* the scattering rate. This should be contrasted with the usually found *decrease* of the scattering rate for ordering within the delta-doped plane (ρ plane). We believe that this behavior arises because of the different character of the electronic states in the directions perpendicular and parallel to the dopant sheet. In the z-direction the wavefunctions are standing waves which leads to constructive coherent superposition in the matrix element. Along the sheet, however, the wavefunctions are propagating plane waves and ordering has the effect of sharpening the Bragg reflection peaks at regions of high momentum in the Brillouin zone which are no longer occupied by electrons.

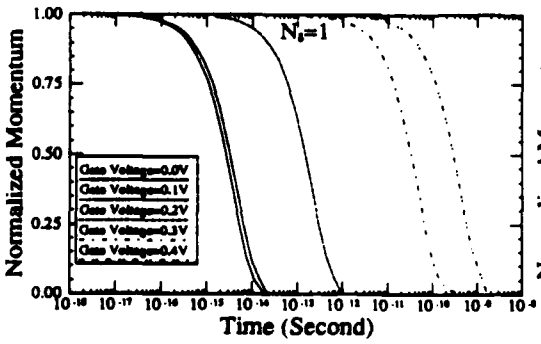


Fig. 5. Shown is the momentum decay for several bias voltages and for the case of one dopant sheet.

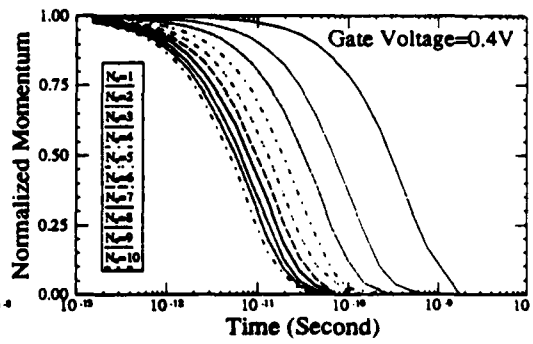


Fig. 6. Shown is the momentum decay for a bias voltage of 0.4 V and for various values of N_δ.

We also investigate the effect of an externally applied electric field on the scattering rates in the quantum well. Previously, we have seen that by applying an appropriate gate voltage the electronic wavefunction can be shifted to a region of the quantum well which is free of dopants. As a result, the impurity scattering rate decreases because the impurity potential is removed from the peak in the electronic wavefunction. We have calculated the coherent scattering rates for the full sequence of N_δ and bias voltages.

These scattering rates then are used to investigate the momentum decay of an ensemble of electrons moving in the plane of the quantum well. The Monte Carlo technique is used to simulate the behavior of carriers injected

with energies at the Fermi energy. The momentum decay for several combinations of bias voltages and numbers of dopant sheets is shown in Figs. 5 and 6.

We see that momentum relaxation is a strong function of the applied bias. We define a momentum relaxation time, τ_m which is the time after which the momentum of the ensemble of electrons decays to $1/e$ of its initial value. In Table 1, we shown our results for the momentum relaxation time τ_m as a function of the applied bias and the number of delta - doped sheets.

Table 1 Momentum Relaxation Times

	$V_g = 0.0$	$V_g = 0.1$	$V_g = 0.2$	$V_g = 0.3$	$V_g = 0.4$
$N_\delta = 1$	3.92×10^{-15}	4.80×10^{-15}	2.20×10^{-13}	4.95×10^{-11}	4.25×10^{-10}
$N_\delta = 2$	2.99×10^{-15}	4.13×10^{-15}	1.34×10^{-13}	1.22×10^{-11}	9.93×10^{-11}
$N_\delta = 3$	2.29×10^{-15}	3.18×10^{-15}	9.45×10^{-14}	6.00×10^{-12}	4.77×10^{-11}
$N_\delta = 4$	1.93×10^{-15}	2.62×10^{-15}	7.18×10^{-14}	3.64×10^{-12}	2.76×10^{-11}
$N_\delta = 5$	1.64×10^{-15}	2.30×10^{-15}	5.91×10^{-14}	2.54×10^{-12}	1.88×10^{-11}
$N_\delta = 6$	1.39×10^{-15}	1.97×10^{-15}	4.94×10^{-14}	1.91×10^{-12}	1.46×10^{-11}
$N_\delta = 7$	1.23×10^{-15}	1.75×10^{-15}	4.19×10^{-14}	1.59×10^{-12}	1.19×10^{-11}
$N_\delta = 8$	1.10×10^{-15}	1.60×10^{-15}	3.68×10^{-14}	1.31×10^{-12}	9.58×10^{-12}
$N_\delta = 9$	9.85×10^{-16}	1.45×10^{-15}	3.27×10^{-14}	1.14×10^{-12}	8.36×10^{-12}
$N_\delta = 10$	8.88×10^{-16}	1.28×10^{-15}	2.96×10^{-14}	9.90×10^{-13}	7.26×10^{-12}
$N_\delta = 15$	6.21×10^{-16}	9.04×10^{-16}	1.95×10^{-14}	6.28×10^{-13}	4.25×10^{-12}
$N_\delta = 20$	4.77×10^{-16}	7.00×10^{-16}	1.46×10^{-14}	4.34×10^{-13}	2.99×10^{-12}

Our calculations show that application of a gate voltage of 400 mV increases the relaxation time by approximately 5 orders of magnitude when the quantum well is doped with one impurity sheet. However, when the quantum well is doped with twenty impurity sheets, application of a 400 mV gate voltage increases the momentum relaxation time (assuming "coherent" scattering) by only 4 orders of magnitude (see Table 1). This modulation of the impurity scattering rates translates into a significant modulation of the channel conductance for a VMT.

CONCLUSION

We investigate the dependence of impurity scattering rates in a quantum well on the configuration of the impurities and on the external bias. Application of an external bias reduces the impurity scattering rate by several orders of magnitude with be a corresponding increase in electron mobility and channel conductance. One can thus realize a switching transistor by skewing the wavefunction with an appropriate external electric field. Obviously, the performance of such a transistor will be influenced by dopant ordering.

We find that selective doping, in which the dopants are confined to a single plane rather than being distributed over several planes, has the effect of decreasing the ionized impurity scattering rates for coherent scattering. In the case of incoherent scattering, i. e. in the absence of correlations between impurity centers, the mobility exhibits a weak dependence on the impurity configuration. This immediately suggests that a given gate voltage is more effective in modulating the conductance when the quantum well is doped with a smaller number of dopant sheets. Thus, the switching performance of the VMT is improved when doping in the quantum well is concentrated into fewer dopant sheets.

ACKNOWLEDGEMENTS

This work was supported by the Air Force Office of Scientific Research under Contract No. AFOSR 88-0096 and by an IBM Development Award.

REFERENCES

Bhobe, S., W. Porod, S. Bandyopadhyay, and D. J. Kirkner, (1989). *Nanostructure Physics and Fabrication* (Academic, Boston, in press)

Gillman, G., B. Vinter, E. Barbier, and A. Tardella, (1988). *Appl. Phys. Lett.* **52**, 972 .

Hamaguchi, C., K. Miyatsuji, and H. Hihara, (1984). *Jpn. J. Appl. Phys.* **23**, L132.

Hong, W.-P., F. DeRosa, R. Bhat, S.J. Allen, and J.R. Hayes, (1989). *Appl. Phys. Lett.* **54**, 457.

Levi, A.F J., S.L. McCall, and P M. Platzman, (1989). *Appl. Phys. Lett.* **54**, 940.

Sakaki, H., (1982). *Jpn. J. Appl. Phys.* **21**, 381.

Schubert, E. F., L. Pfeiffer, K.W. West, and A. Izabelle, (1989). *Appl. Phys. Lett.* **54**, 1350.

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

Yokoyama, K., and K. Hess, (1985). *Phys. Rev.* **B 31**, 6872.

TAILORING TRANSPORT PROPERTIES BY WAVEFUNCTION ENGINEERING IN QUANTUM WELLS AND ITS DEVICE APPLICATIONS[1]

Suyog Bhobe, Wolfgang Porod, and Supriyo Bandyopadhyay
Department of Electrical and Computer Engineering
University of Notre Dame

David J. Kirkner
Department of Civil Engineering
University of Notre Dame

We investigate a semi-classical mesoscopic phenomenon in which the dependence of a system's macroscopic transport properties on the microscopic details of the electronic wavefunction is exploited to realize an ultrafast switching transistor. The conductance of a quantum well with a selectively-doped region depends on the precise nature of the wavefunction in the well which can be altered by an external field that "pushes" the wavefunction in and out of the doped region. This modulates the conductance of the well (by few orders of magnitude at liquid helium temperature) on timescales of the order of 100 femtoseconds. We have investigated this phenomenon using a combination of self-consistent Schrödinger Equation solution and ensemble Monte Carlo simulation to model transient electronic transport in the well.

I. INTRODUCTION

It is generally believed that mesoscopic "quantum devices", whose operations rely on quantum mechanical phenomena, will be much faster than classical devices such as an ordinary field-effect-transistor. The reason for this is that classical devices are switched by moving carriers in and out of the device so that the switching time is limited by the transit time of carriers. Quantum devices, on the other hand, do not usually require infusion and extraction of carriers. Typically, they are switched by inducing constructive or destructive interference of electrons

[1]This work was supported by the Air Force Office of Scientific Research.

Copyright © 1989 by Academic Press, Inc.
All rights of reproduction in any form reserved
ISBN 0-12-585000-X

which does not require physical movement of charges. Therefore, the switching speed of quantum devices is not likely to be limited by the transit time of carriers[2].

Although quantum devices have this inherent advantage, they also have certain disadvantages. Devices that are based on phase-coherent phenomena (such as interference) must have dimensions smaller than the phase-breaking length of electrons which is given by

$$L_\phi = \sqrt{D\tau_{in}}\ , \tag{1}$$

where τ_{in} is the inelastic scattering time and D is the diffusion coefficient of electrons which depends on the elastic scattering time or "mobility".

It is evident that quantum devices will have the following drawbacks. Firstly, they must be operated well below room temperature so that τ_{in} is sufficiently large. Room temperature operation would require such small feature sizes that the lithographic capabilities necessary for delineating them are presently unavailable. Secondly, the material for the devices must be sufficiently "clean" since D depends on the elastic scattering time and hence the mobility μ. This is quite critical in two- or three-dimensional structures in which elastic scattering is far more frequent than inelastic scattering at cryogenic temperatures. Only in one-dimensional structures, the cleanness of the material is not that critical because of the drastic suppression of elastic scattering by one-dimensional confinement[3]. But one-dimensional structures (quantum wires) are not easy to fabricate and their current carrying capability is inherently low which makes them inappropriate for many applications.

There is however at least one semi-classical device that combines the advantages of both quantum devices (fast switching speed not limited by the transit time) and classical devices (no requirement of phase coherence and associated complications). The principle behind the operation of this device is very simple. The conductance of a two-dimensional structure such as a quantum well is given by

$$G\left(in\ (\Omega/\square)^{-1}\right) = qn_s\mu \tag{2}$$

Instead of modulating G by modulating n_s (as is done traditionally), one can modulate it by changing μ. The advantage is that μ can be changed on timescales of the order of the momentum relaxation time so that the switching speed of such a device is not limited by the transit time.

[2]There are exceptions however. An example is the electrostatic Aharonov-Bohm interferometer in which the switching speed is in fact limited by the transit time of carriers.

[3]This does not mean that one-dimensional structures can be arbitrarily "dirty" since many quantum interference effects may not survive in the strong localization regime.

The important question now is how to modulate μ. The mobility depends on the scattering rates of electrons which (even in the semiclassical formalism of Fermi's Golden Rule) depends explicitly on the electronic states (wavefunctions) in the system. By applying an external field, the wavefunction in a quantum well can be altered - it can be skewed and pushed in and out of a selectively doped region within the well - which modulates the scattering rates and the mobility. Such wavefunction engineering is essentially similar to the quantum-confined Franz-Keldysh effect. When the wavefunction resides mostly the doped region, the mobility is low, otherwise it is high. If we neglect the time required for skewing the wavefunction (which is very small) then the switching time of such a device is essentially the momentum relaxation time which can be less than a picosecond.

Such a device, termed a velocity modulation transistor (VMT), has been proposed by Hamaguchi and his co-workers [1]. In this paper, we analyze this device and evaluate the magnitude of the conductance modulation as well as the switching time.

II. THEORY AND COMPUTATION

The electronic states of a quasi two-dimensional system, such as a quantum well, have been studied widely in the literature [2]. We have calculated the wavefunctions inside a 500 Å GaAs-AlGaAs quantum well, whose right-half [250 - 500 Å] is doped with impurities. The wavefunctions are obtained by solving self-consistently the Schrödinger and Poisson equations using the Finite Element Method. The details of the calculation are presented in Ref. 3. We assume a carrier concentration of 5 x 10^{11} cm^{-2} and an impurity concentration of 2 x 10^{17} cm^{-3} in the selectively doped region. The wavefunctions and the energy levels are shown in Fig. 1.

The scattering rate of two-dimensionally confined electrons in the well is obtained from Fermi's Golden Rule and is given by

$$\frac{1}{\tau(k)} = \frac{N_{eff} m^* e^4}{4\hbar^3 \epsilon^2 \lambda \sqrt{\lambda^2 + 4k^2}} , \tag{3}$$

where k is the Fermi wavevector $\left(k = \sqrt{2\pi n_s}\right)$, n_s is the two-dimensional electron concentration, m^* is the effective mass of electrons, ϵ is the dielectric constant and λ is the screening constant which was taken as 300 Å for GaAs.

The term in the above equation that depends explicitly on the precise details of the wavefunction is the so-called effective impurity density N_{eff} which is related to the "effective" number of impurities that interact *in situ* with the electrons. This quantity is obtained as

$$N_{eff} = \sum_m \int_{doped\ region} N_D(x)\ |\psi_m(x)|^2 dx , \tag{4}$$

where $\psi_m(x)$ is the wavefunction of the mth. subband in the well. The integration is performed over the selectively doped regions of the quantum well. This "effective" density was calculated for various electric fields applied transverse to the well interfaces that skew the wavefunctions away from the doped region by different amounts. It is easy to see now that the scattering rate can be modified by altering $\psi_m(x)$ by a transverse field (or equivalently a "gate voltage") which alters N_{eff}.

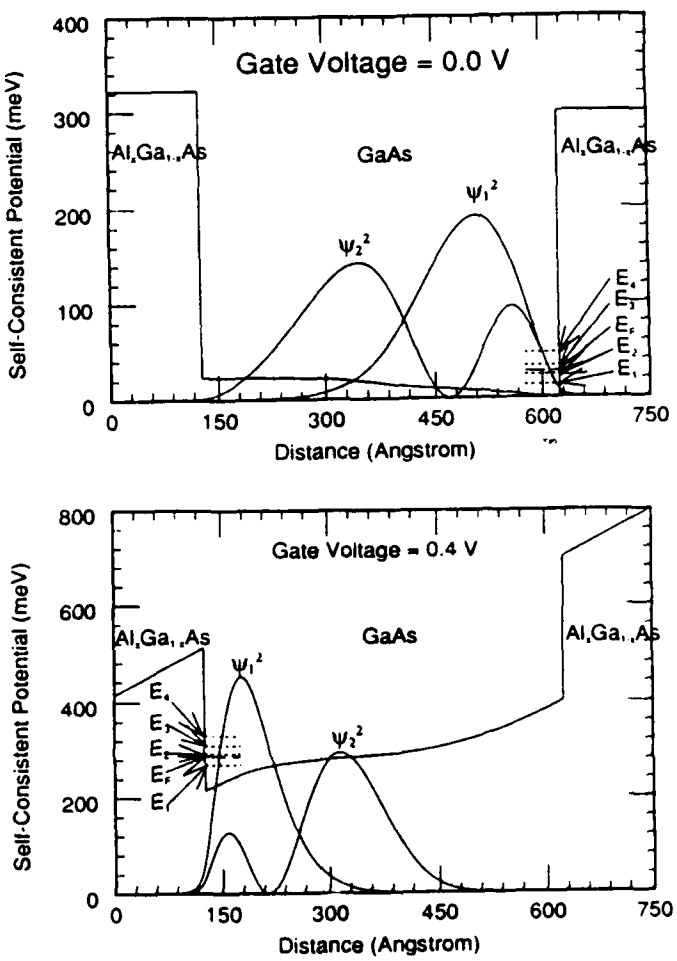


Fig. 1. The self-consistent potential and the wavefunctions in the two lowest occupied subbands in the selectively-doped GaAs-AlGaAs well in the absence (top figure) and presence (bottom figure) of a gate voltage.

The scattering rates obtained from Equations (3) and (4) for various gate-voltage-dependent $\psi_m(x)$ were used in a two-dimensional Monte Carlo simulation to model transient electronic transport at 4.2 K. From the simulation results, we extracted the momentum decay characteristics of an ensemble of electrons injected parallel to the interfaces of the well with the Fermi velocity. We included only the impurity scattering mechanism in the simulation and neglected all other kinds of scattering. The impurity scattering rate was found to be sufficiently high to be dominant at 4.2 K. Fig. 2. shows the decay of the initial momentum with time at various gate voltages. From the decay characteristics, we evaluated the momentum relaxation time τ_m by defining τ_m to be the time that elapses before the ensemble average momentum decays to $\frac{1}{e}$ times its initial value. From the momentum relaxation time, we calculate the effective "mobility" in the quantum well using $\mu = e\tau_m/m^*$, and from this we obtain the conductance G using Equation (2).

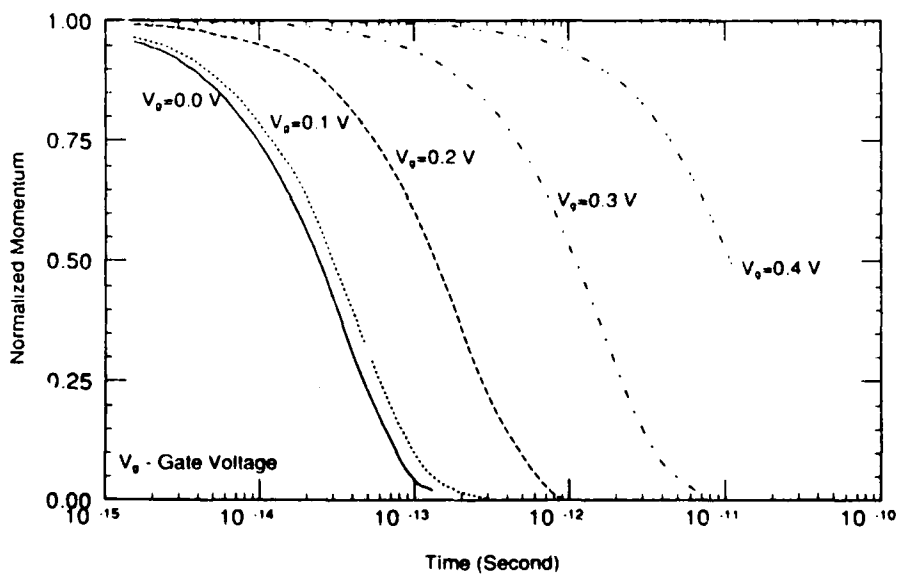


Fig. 2. The decay of the initial momentum of an electron injected parallel to the well interfaces at various gate voltages. The results are valid for a temperature of 4.2 K. These curves were obtained from ensemble Monte Carlo simulation.

Gate voltage (V)	Scattering rate (sec^{-1})	Conductance ($(\Omega/\square)^{-1}$)
0.0	3.8×10^{13}	5.5×10^{-5}
0.1	3.0×10^{13}	6.9×10^{-5}
0.2	6.5×10^{12}	3.2×10^{-4}
0.3	8.0×10^{11}	2.6×10^{-3}
0.4	8.3×10^{10}	2.5×10^{-2}

Table 1: Scattering rates and conductance vs. gate voltage

Table I above lists the conductances for various gate voltages. Note that the conductance can be modulated over more than three orders of magnitude by varying the gate bias over a range of only 400 millivolts at 4.2 K. The modulation will certainly decrease significantly at higher temperatures due to phonon scattering, but the rather small threshold voltage of 0.4 V still indicates a large transconductance and also a very low power-delay product. In addition, the switching speed, being of the order of the momentum relaxation time, is about 100 fs which is comparable to the switching speeds of quantum devices presently extant. In these respects, the performance of this device is quite comparable to the performance of ultra-high performance quantum devices.

III. CONCLUSION

In this paper, we have explored a semi-classical device whose performance is comparable to those of mesoscopic quantum devices but whose fabrication is much easier. It is an interesting example of a mesoscopic phenomenon where a macroscopic property, namely the conductance, depends on the microscopic details of the electronic wavefunction. The advantage here is that there is no requirement of phase coherence so that the device could operate at elevated temperatures and the demand on the material quality is not stringent.

V. REFERENCES

[1] C. Hamaguchi, K. Miyatsuji, and H. Hihara, *Jpn. J. Appl. Phys.* **23**, L132 (1984). A similar concept has been proposed by H. Sakaki, *Jpn. J. Appl. Phys.* **21**, 381 (1982).

[2] See, for example, F. Stern, *Physical Review* **B 5**, 4891 (1972).

[3] S. Bhobe, W. Porod and S. Bandyopadhyay, *Surface and Interface Analysis*, (1989) (in press).

A proposed quantum wire structure: An "accumulation wire" at crossing heterointerfaces

Henry Harbury and Wolfgang Porod
Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556

(Received 31 January 1990; accepted 12 March 1990)

We demonstrate that a quasi-one-dimensional quantum wire can, in principle, be produced at the crossing point of heterointerfaces. We show that the intersection of two accumulation layers, which are formed along separate heterointerfaces, realizes an "accumulation wire." Our theoretical analysis yields potential distributions which possess confined electronic states in the vicinity of the point of intersection. In our design, confinement in the quantum wire is solely provided by the inherent electric fields associated with heterointerfaces without the need for additional lithography. In other words, our proposal of the quantum wire compares to previous designs in the same manner in which a quasi-two-dimensional (2D) system in an accumulation (or inversion) layer compares to a quasi-2D system in a quantum well.

I. INTRODUCTION

Semiconductor structures with reduced dimensionality have been the subject of much recent research. These quantum confined systems are expected to exhibit superior properties as compared to more conventional semiconductor structures. Improved optical characteristics[1,2] include larger electroabsorption and electrorefraction, enhanced optical nonlinearities, and higher differential optical gain. These structures are also expected to show unique electrical transport properties and to have extremely high electron mobilities due to suppressed impurity scattering.[3]

Quasi-one-dimensional (1D) systems, so-called quantum wires, have been realized using a variety of techniques. Almost all of these techniques are based on the principle of providing additional, lateral confinement for a quasi-two-dimensional electron gas (2DEG) by lithographic means. The 2DEG is realized in Si systems at the Si–SiO_2 interface, and in the GaAs/GaAlAs material system either at a single heterointerface or inside a quantum well. Lateral confinement is achieved by additional processing steps, such as etching and regrowth,[4,5] deep mesa etching,[6,7] electric field confinement,[8,9] shallow etching,[10] epitaxial growth on a tilted substrate,[11] and focused ion beam implantation.[12]

In this paper, we explore the possibility of creating quantum wires by the exclusive use of the internal electric fields associated with heterointerfaces. Specifically, we investigate whether quasi-two-dimensional (2D) confinement of electrons can be achieved at crossing heterointerfaces. It is well known that a heterojunction may provide 1D confinement in the direction perpendicular to the interface, which is the basis for the device operation in a high electron mobility transistor (HEMT). It is then natural to ask whether and under what conditions a pair of heterointerfaces may provide 2D confinement. In that case, quasi-2D confinement is achieved solely by built-in electric fields which distinguishes our proposal of a quantum wire from the earlier designs where lithographic techniques are required. In other words, our proposal of the quantum wire compares to previous designs in the same manner in which a quasi-2D system in an accumulation (or inversion) layer compares to a quasi-2D system in a quantum well.

More specifically, we concentrate on a model system which consists of perpendicular accumulation layers, and we investigate if it is possible to produce an "accumulation wire" at their point of intersection. We theoretically study the potential distributions and electronic states at the intersection by solving Poisson's and Schrödinger's equation, taking into account the potential discontinuities at the interfaces. An isolated quantum wire will only be produced if the Fermi energy is such that the electrons accumulate only at the crossing point, and not at the heterointerfaces themselves. We study doping conditions and material combinations forming such an isolated accumulation wire.

The theory of potential distributions and electronic states in quantum wires has previously been studied. The potential distributions have been determined for narrow channel[13] geometries. The electronic states have been obtained using variational wave functions[14] and self-consistent solutions to Schrödinger and Poisson's equations.[15]

We proceed in Sec. II by outlining our model system to investigate crossing heterointerfaces. In Sec. III, we detail the numerical techniques used for this theoretical study. Results are presented in Sec. IV for the GaAs–AlGaAs material system. Finally, we conclude in Sec. V.

II. CROSSING HETEROINTERFACES

The model structure which we consider here for the study of crossing heterointerfaces is schematically depicted in Fig. 1. The two-dimensional spatial domain with coordinates (x,y) is subdivided into three different regions with interfaces between them. Each of these regions, denoted by I, II, and III, may be occupied by a different material. In general, we then have three different heterointerfaces, denoted by A, B, and C, which intersect in the center point, W. It is right at this center point, where the accumulation wire is expected to exist.

Such a structure could, in principle, be realized by growing material II epitaxially onto I, then polishing the sides, and epitaxially growing material III onto the side faces. Alternatively, region III might represent a filled-in V groove etched through the interface between materials I and II, although the heterointerfaces would not be perpendicular to

 © 1990 American Vacuum Society

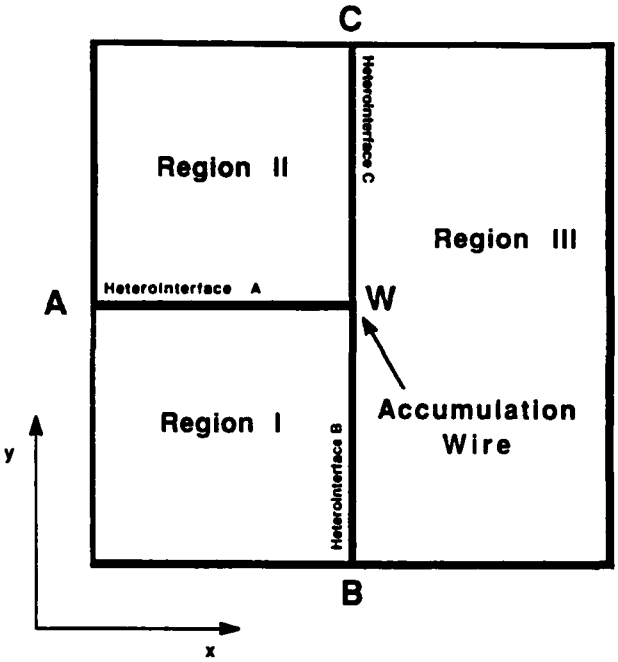


FIG. 1. A schematic diagram is shown for a geometry of crossing heterointerfaces. Special points along the heterointerfaces and the point of their intersection are indicated.

each other in this structure. Regions I and II could also be repeated periodically in the fashion of a superlattice which would result in an array of parallel wires. Another possibility for realizing this structure might be to utilize a V groove with its sidewalls at heterointerfaces A and B. In that case, regions II and III would be identical, and material I would represent the filled-in groove. A structure similar to the latter case was realized very recently[16] and it was shown that a quasi-2D electronic system can exist on the sidewalls of the V groove. Also very recently it was demonstrated that an optically active quantum wire[17] may be realized at the tip of a V groove which is overgrown with a quantum well structure.

We will concentrate on the heterointerfaces A and B, which are the ones along the intersection of regions I–II and I–III, respectively. We want to study conditions for which there is no significant accumulation along the heterointerfaces at locations A and B, i.e., far from the center point, W. At the same time, however, we would like to achieve quasi-one-dimensional accumulation in a region close to the point of intersection.

In the theoretical treatment of this problem, we obtain the potential distribution for the complete two-dimensional spatial domain. Within a Thomas–Fermi screening model, the electrostatic potential $\phi(\mathbf{r})$ determines the charge distribution, and therefore the band bending. We solve Poisson's equation for the geometry shown in Fig. 1. The effects of the different materials enter through the background doping and the band discontinuities at the heterointerfaces.

$$\nabla^2\phi(\mathbf{r}) = \frac{e}{\epsilon\epsilon_0}\left[n(\mathbf{r}) - N_D^+\right]. \quad (1)$$

The charge term contains the sum of the electron density and the background doping. We assume that all dopants are ionized, i.e., N_D^+ is constant throughout each region and has the value of the nominal background doping. The electronic charge density, for degenerate statistics, is given by the Fermi–Dirac integral of order 1/2, $n(\mathbf{r}) = N_c(\pi/2)F_{1/2}(\eta)$, where η is the energy separation between the local conduction band edge $E_c(\mathbf{r})$ from the Fermi level, E_F measured in units of k_BT;$\eta = -(E_c - E_F)/k_BT$.

At the heterointerfaces, the electronic potential has a discontinuity which is given by the conduction band offsets. These quantities have been the subject of numerous investigations and are available for the various heterojunctions. Here, we are faced with a more complicated situation in that the conduction band discontinuities along the interfaces may be a function of the distance to the center point. Lacking any model for this spatial dependence, we assume the conduction band offset to be constant along a particular heterointerface. Concerning the crossing point itself, we assume that the sum of the potential energies has to be zero along a path looping around the center point. This implies that the potential discontinuities for the three heterointerfaces have to add up to zero. This is consistent with a transitivity property of the conduction band offsets for any given combination of materials.

Of special interest are the electronic states in the vicinity of the crossing point. We are particularly interested in bound states which are confined both in the x and y directions. Quasi-one-dimensional motion is then only possible in the z direction. We solve Schrödinger's equation for the potential which we obtain from Poisson's equation.

$$-\frac{\hbar^2}{2m^*}\nabla^2\psi(\mathbf{r}) + [V(\mathbf{r}) - e\phi(\mathbf{r})]\psi(\mathbf{r}) = E\psi(\mathbf{r}). \quad (2)$$

Here, $V(\mathbf{r})$ includes external potentials and band offsets and $\phi(\mathbf{r})$ is the electrostatic potential obtained from Poisson's equation.

III. NUMERICAL METHOD

The determination of the potential distribution and electronic states at the crossing heterointerfaces represents a challenging numerical problem since the spatial scales for the two problems are very different. Poisson's equation has to be solved on a domain on the order of microns to obtain accurate potential distributions with the correct asymptotic behavior. Schrödinger's equation, however, should yield accurate bound electronic states in the vicinity of the center point, with spatial scales on the order of hundreds of Angstroms. The difference in the size of the domain for the two equations implies that the Schrödinger problem has to be solved on a subset of the domain used to solve Poisson's equation.

We employ the finite element method for the numerical treatment of both the Poisson and Schrödinger equations. Uniform rectangular elements are used with standard first order Chapeau basis functions. We typically use a mesh size of 100×100 nodal points for both equations. The finite element method results in a linear system of equations for the unknowns at each nodal point. For our mesh size, we have to solve problems of order 10 000! Fortunately, these matrices

are sparse, and efficient numerical techniques are available for their solution.

Poisson's equation is solved on a spatial domain of typically 6000×6000 Å which we found to be sufficiently large to satisfy the boundary condition that the potential asymptotically approaches its value in the bulk. The boundary conditions fix the potential at the four corners of the rectangular domain such that the corresponding electron density is equal to the background doping of that region, simulating charge neutrality in the bulk. A standard LU (lower and upper triangular) decomposition method is used to solve the linear system of equations after it has been reduced to banded symmetric form.

Schrödinger's equation is solved on a spatial domain of typically 1000×1000 Å surrounding the center point. This region is sufficiently large to obtain bound states by requiring the wavefunction to be equal to zero at the boundary of this domain. The potential used as input to Schrödinger's equation is a bilinearly interpolated detail around the center point of the potential obtained from the solution of Poisson's equation. The resulting eigenvalue problem can be reduced to banded symmetric form of typical dimension $10\,000 \times 100$. Since we are primarily interested in the bound electronic states, only the lowest eigenvalues and eigenvectors need to be determined. The Lanczos alogorithm[18] is used for the efficient numerical computation of the lowest eigenvalues and eigenvectors.

Most numerical computations are performed on a Convex C-2. The solution of Poisson's equation typically takes 600 s of CPU time. The solution of the eigenvalue problem for the first 10 states typically requires 200 s of CPU time. Computations were also performed on a Cray X-MP 48, with execution times about an order of magnitude faster.

IV. RESULTS

We concentrate on the GaAs–$Al_xGa_{1-x}As$ material system because of the good lattice match in these heterostructures. The various regions (I, II, and III) in Fig. 1 correspond to different Al mole fractions, x_I, x_{II}, and x_{III}, and different background doping concentrations, N_I, N_{II}, and N_{III}. We linearly interpolate the relevant material parameters, like effective masses, dielectric constants, etc. At the interfaces between regions of different Al content, the potential energy for electrons will exhibit a discontinuity. For the values of these conduction band offsets, we use the data given in Ref. 19.

A typical potential landscape is shown in Fig. 2 from two different viewing angles to display the location of the heterointerfaces and their point of intersection. The front edge is highlighted in both graphs to aid spatial perception. Using the same convention as adopted in Fig. 1, we label the heterointerfaces (A, B, and C) and the crossing point (W). The top portion of the figure shows the potential as seen from the top at a viewing angle of $+80°$. The heterointerfaces with their potential discontinuity are clearly discernible. The bottom portion of the figure shows the potential as seen from below at a viewing angle of $-10°$. Clearly visible now is the dip in the potential at the center point, which provides the quantizing potential "funnel" for the quantum wire. The sidefaces show the familiar potential variation across a heterointerface.

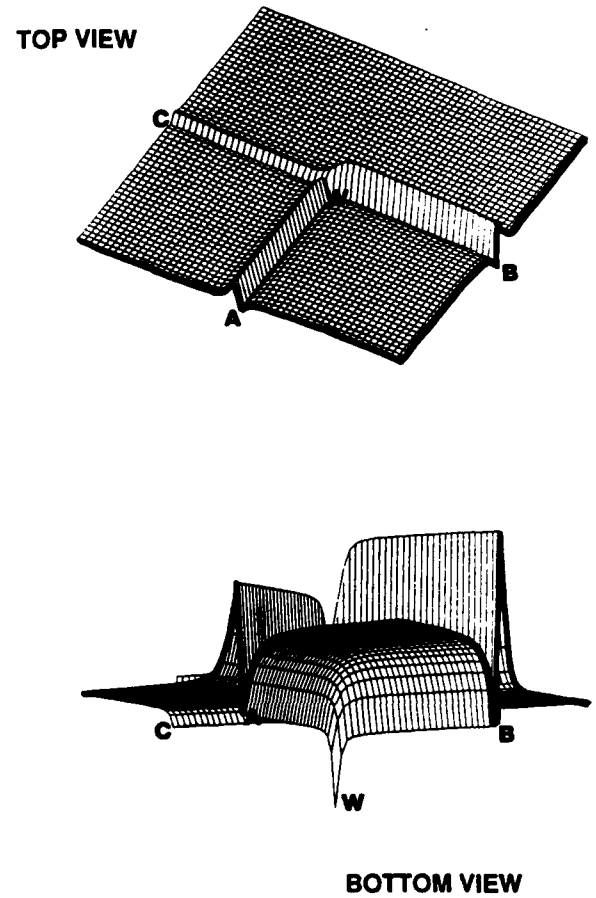


FIG. 2. A typical potential landscape is shown from two different viewing angles. The top part shows the top view from an angle of $+80°$ and the bottom part shows the bottom view from an angle of $-10°$. Special points are labeled using the same convention as in Fig. 1.

In order to study accumulation, we choose material I to be lightly n-type doped GaAs with a doping concentration of $N_I = 5 \times 10^{15}$ cm^{-3} throughout. We first study the case in which regions II and III are occupied by the same material

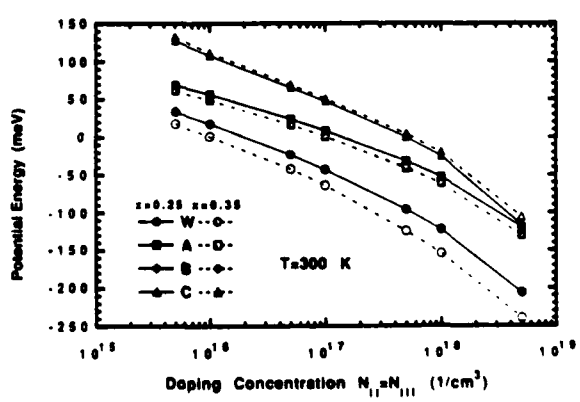


FIG. 3. Shown is the potential variation at selected locations as a function of the doping concentration in regions II and III. Note that the center point W dips below the Fermi level, which is the zero on the energy scale.

$x_{II} = x_{III}$ with the same background doping concentrations $N_{II} = N_{III}$. In Fig. 3, we show the potential energy for electrons, as a function of the doping concentration in regions II and III, at points A, B, and C at the heterointerfaces as well as at the center point, W. A temperature of 300 K is used in these calculations. The figure shows the value corresponding to the lower potential at the respective heterojunction. The zero of potential energy is taken to be the Fermi energy for convenience. Because of the assumed symmetry in regions II and III, the potentials at points A and B are the same. We show results for two different mole fractions $x_{II} = x_{III} = 0.25$ indicated by the solid symbols and solid lines, and $x_{II} = x_{III} = 0.35$ indicated by the open symbols and dotted lines. The larger band discontinuity for $x = 0.35$ lowers the potentials at points W, and A and B. For point C, the larger effective mass and smaller dielectric constant leads to the slight increase in the values for $x = 0.35$. Note that the center point, W, dips below the Fermi energy, indicating an accumulation of electrons. Also note that the energies at the heterojunctions far from the intersection do not dip down as far as the center point This means that accumulation in the 1D wire starts before 2D accumulation layers form. For doping larger than about 5×10^{17} cm^{-3} in the layers II and III, the heterointerfaces themselves begin to accumulate electrons.

To demonstrate the existence of a confined electronic state at the center point, we solve Schrödinger's equation for promising potential profiles, as just determined. As can be seen from Fig. 3, the case for a doping concentration $N = 1\times10^{17}$ cm^{-3} looks promising since point W lies below the Fermi energy while the other points have positive potential energies. The top portion of Fig. 4 shows the potential

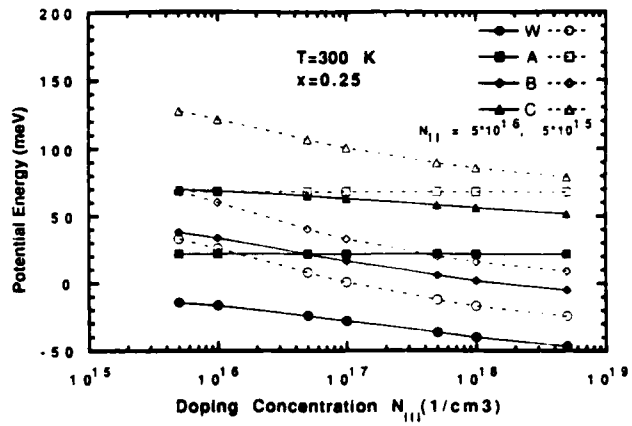


FIG. 5. Shown is the potential variation at selected locations as a function of the doping concentration in region III, with N_{II} as a parameter. Note that the center point W dips below the Fermi level, which is the zero on the energy scale.

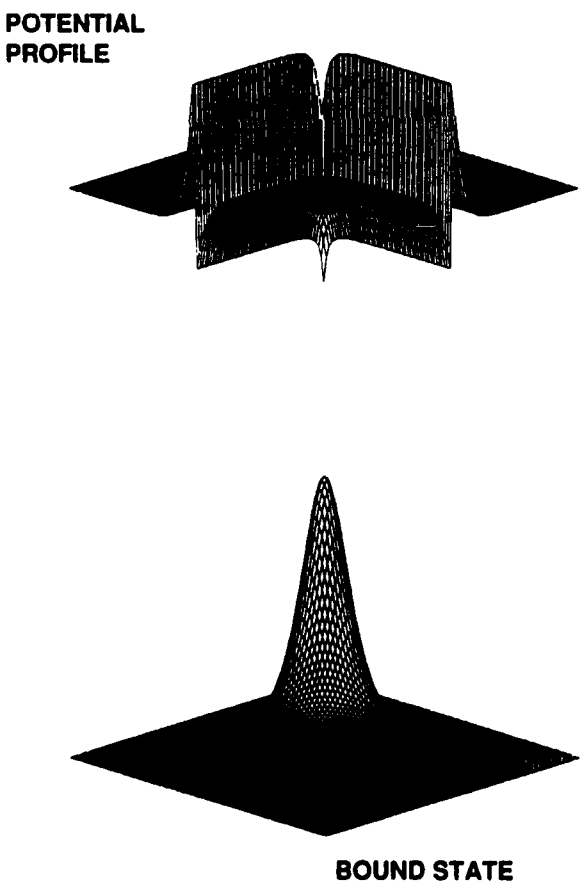


FIG. 4. The top part shows the potential landscape for the symmetric case in which regions II and III are assumed identical. The potential dip in the center produces a bound electronic state whose wave function is displayed in the lower portion of the figure. For more details refer to the text.

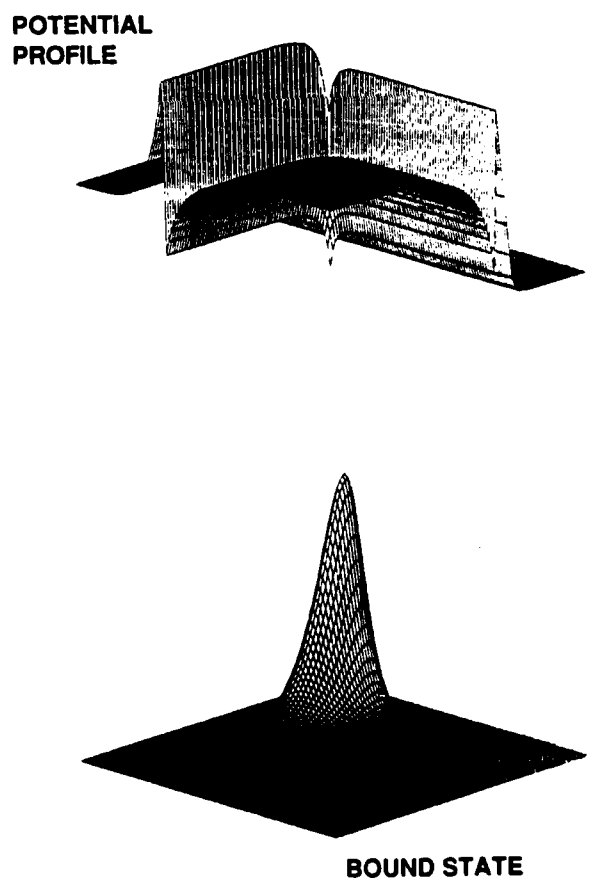


FIG. 6. The top part shows the potential landscape for the asymmetric case where region III contains a spacer layer. The potential dip in the center produces a bound electronic state whose wave function is displayed in the lower portion of the figure. For more details refer to the text.

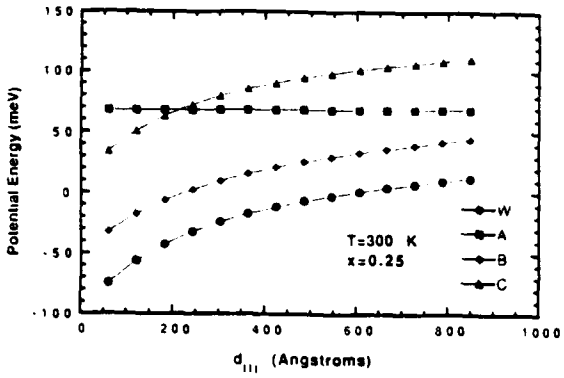


FIG. 7. Shown is the potential variation at selected locations as a function of the width of the spacer layer as described in the text. Note that the center point W dips below the Fermi level, which is the zero on the energy scale.

landscape for this case and $x = 0.35$. Schrödinger's equation is solved for this potential and a bound state is found with an energy of 14.3 meV. The corresponding bound state wave function is also shown in the bottom part of Fig. 4. The spatial extent of the bound electronic state is about 150 Å in the direction along the heterointerfaces and about 75 Å in the direction perpendicular to the heterointerfaces.

For realizations of this structure in which region III is grown epitaxially onto I and II, modulation doping of region III might be desirable. In order to study this case, we now subdivide region III into two parts. Next to the heterointerfaces we place a lightly doped "spacer layer" of width d_{III}, and the remainder of region III is doped with N_{III} as before. Figure 5 shows the potentials at the center point and heterointerfaces as a function of the doping concentration N_{III}. The spacer layer is assumed to have a width of $d_{III} = 333$ Å, and a doping density equal to N_{II}, which takes on two different values, as shown in the figure. In this case, there is no symmetry between points A and B and the corresponding potential values are no longer the same. Note that the center point W again dips below the Fermi energy while the potentials along the heterointerfaces remain above. A confined electronic state exists at the center, which is demonstrated in Fig. 6. The top part of this figure shows the potential landscape for the current asymmetric case for $N_{III} = 5\times10^{16}$ cm^{-3} and $x = 0.25$. The bottom part of the figure displays the electronic wave function of the bound state which has an energy of 26 meV. In addition, we also show the dependence of the electronic potentials as a function of the width of the spacer layer, d_{III}. Figure 7 shows the results for $N_{III} = 5\times10^{15}$ cm^{-3} and $x = 0.25$. Note again that a potential funnel is produced at the intersection of the heterointerfaces which extends below the Fermi energy indicating the presence of an accumulation wire.

V. CONCLUSIONS

We have demonstrated that a quasi-1D quantum wire can, in principle, be produced at the point of intersection of heterointerfaces. By crossing two accumulation layers formed along separate heterointerfaces, one may thus realize an accumulation wire. This demonstrates that electron confinement solely provided by the inherent electric fields associated with heterointerfaces is sufficiently strong to produce a bound electronic state in two dimensions.

In our theoretical treatment of the problem we use a semiclassical approach to calculate the amount of charge in the structure and we then calculate the quantum confined states from the resulting potential. A shortcoming of the present approach is that it is not self-consistent. In the future, we intend to extend our numerical treatment to calculate the potential and the charge distributions self-consistently. Due to the lack of self-consistency, we have also refrained in the present study from reporting the amount of charge residing in the structure which depends rather sensitively on the energy levels.

Our proposed structure does not require lithography to provide lateral confinement, however, it requires the formation of a lateral interface for regrowth. This interface can be realized either by polishing or by etching. It is hard to assess the relative merits of these two processing steps for the quality of the resulting interface. Likely, both techniques are effectively equivalent which makes our proposed structure similar in feasibility to the ones proposed in Refs. 4 and 5.

The density of the 1D gas in our structure is determined by the material composition and the doping levels in the different regions. We have given results that provide a sample of the rather large parameter space for the various combinations of material parameters. In its present form, our structure is without some sort of gate to control the amount of charge in the wire. Future studies will address the issue of gaining control over the Fermi level such that it can be moved through the quasi-1D levels which have been formed. A related issue is the possible transfer between quasi-two-dimensional electrons in the accumulation layers and quasi-one-dimensional electrons in the accumulation wire. Self-consistent solutions to Schrödinger's and Poisson's equations will then be required to investigate these questions.

ACKNOWLEDGMENTS

The authors would like to thank Dr. S. Bandyopadhyay, Dr. D. J. Kirkner, Dr. C. S. Lent, and S. Srinivasan for helpful discussions. This work was supported, in part, by the Air Force Office of Scientific Research, the Office of Naval Research, and an IBM Faculty Development Award. Computer time was obtained through the National Center of Supercomputing Applications at UIUC and the National Center for Computational Electronics.

[1] D. A. B. Miller, D. S. Chemla, and S. Schmitt-Rink, Appl. Phys. Lett. 52, 2154 (1988).
[2] Y. Arakawa and A. Yariv, IEEE J. Quantum Electron. 22, 1887 (1986).
[3] H. Sakaki, Jpn. J. Appl. Phys. 19, L735 (1980); J. Vac. Sci. Technol. 19, 148 (1981).
[4] P. M. Petroff, A. C. Gossard, R. A. Logan, and W. Wiegman, Appl. Phys. Lett. 41, 635 (1982).
[5] Y.-C. Chang, L. L. Chang, and L. Esaki, Appl. Phys. Lett. 47, 1324 (1985).
[6] W. J. Skocpol, L. D. Jackel, E. L. Hu, R. E. Howard, and L. A. Fetter, Phys. Rev. Lett. 49, 951 (1982).
[7] K. K. Choi, D. C. Tsui, and S. C. Palmateer, Phys. Rev. B 32, 5540 (1985).

[8] T. J. Thornton, M. Pepper, H. Ahmed, D. Andrews, and G. J. Davies, Phys. Rev. Lett. **56**, 1198 (1986).
[9] A. C. Warren, D. A. Antoniadis, and H. I. Smith, Phys. Rev. Lett. **56**, 1858 (1986).
[10] S. Thoms, I. McIntyre, S. P. Beaumont, M. Al-Mudares, R. Cheung, and C. W. Wilkinson, J. Vac. Sci. Technol. B **6**, 127 (1988).
[11] J. M. Gaines, P. M. Petroff, H. Kroemer, R. J. Simes, R. S. Geels, and J. H. English, J. Vac. Sci. Technol. B **6**, 1378 (1988).
[12] T. Hiramoto, K. Hirakawa, Y. Iye, and T. Ikoma, Appl. Phys. Lett. **54**, 2103 (1989).
[13] A. Ya. Shik, Sov. Phys. Semicond. **19**, 915 (1985).
[14] W. Y. Lai and S. Das Sarma, Phys. Rev. B **33**, 8874 (1986).
[15] S. E. Laux and F. Stern, Appl. Phys. Lett. **49**, 91 (1986).
[16] M. R. Frei and D. C. Tsui, Appl. Phys. Lett. **55**, 2432 (1989).
[17] E. Kapon, D. M. Hwang, and R. Bhat, Phys. Rev. Lett. **63**, 430 (1989); E. Kapon, S. Simhony, R. Bhat, and D. M. Hwang, Appl. Phys. Lett. **55**, 2715 (1989).
[18] J. K. Cullum and R. A. Willoughby, *Lanczos Algorithms for Large Symmetric Eigenvalue Computations*, (Birkhäuser, Boston, 1985), Vol. I: Theory, Vol. II: Programs.
[19] J. Batey and S. L. Wright, J. Appl. Phys. **59**, 200 (1986); J. Batey, S. L. Wright, and D. J. DiMaria, *ibid.* **57**, 484 (1985).

Reprinted from COMPUTATIONAL ELECTRONICS, Eds. K. Hess, J. P. Leburton, and U. Ravaioli (Kluwer Academic Press) 1990.

NUMERICAL STUDY OF ELECTRONIC STATES IN A QUANTUM WIRE AT CROSSING HETEROINTERFACES

Henry K. Harbury and Wolfgang Porod
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, IN 46556

Abstract

We demonstrate that the confinement produced by the internal electric fields at crossing heterointerfaces is sufficient to produce a bound electronic state in two dimensions. The potential profiles and electronic states are obtained by numerically solving the Poisson and Schrödinger equations using the Finite Element Method. We also investigate the feasibility of parallelizing the calculations on the local element level by using a cluster of networked workstations as a distributed computational resource.

Introduction

Semiconductor structures with reduced dimensionality have been the subject of much recent research. These quantum confined systems are expected to exhibit superior electrical [1] and optical [2] properties as compared to more conventional semiconductor structures. Various design ideas for systems with confinement in 2 dimensions and free propagation in the third, so-called quantum wires, have been proposed [4,5]. Several evaluations of the electronic states in quantum wires have been reported in the literature, e.g. [5,6].

Here, we present calculations of the electronic states in a novel quantum wire structure, which we recently proposed [7]. This so-called 'accumulation wire' is formed at the intersection of two perpendicular accumulation layers. Electronic confinement is solely provided by the electric fields associated with the heterointerfaces. We study the potential profiles and electronic wavefunctions by solving Poisson's and Schrödinger's equation, taking into account the potential discontinuities at the interfaces. The Finite Element Method (FEM) is employed for the numerical solution of this set of equations [8].

Our calculations are performed in an environment of networked SUN workstations in which network daemons can utilize idle CPU's on the network. These daemons can also assign computational tasks to available CPU's. We investigate the possibility of speeding up the code by parallelizing the FEM calculations at the local element level, using the network as a distributed computational resource.

Crossing Heterointerfaces

The model geometry of crossing heterointerfaces is schematically depicted in Fig. 1. The two - dimensional spatial domain with coordinates (x,y) is subdivided into three different regions, denoted by I, II, and III. In general,

we then have three different heterointerfaces, denoted by A, B, and C, which intersect in the center point, W, where the quantum wire is expected to exist. The realization of such a structre appears to be within the realm of possibility, considering recent progress in processing for etching and regrowth [9], and cleaving and regrowth [10].

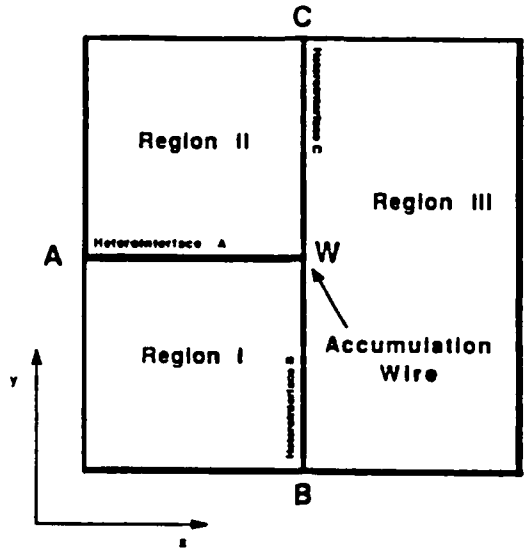


Figure 1: A schematic diagram is shown for a geometry of crossing heterointerfaces. Special points along the heterointerfaces and their point of intersection are indicated.

In the theoretical treatment of this problem, we obtain the potential distribution for the complete two - dimensional spatial domain. Within a Thomas - Fermi screening model, the electrostatic potential determines the charge distribution, and consequently the band bending. We solve Poisson's equation taking into account the respective background doping and the band discontinuities at the heterointerfaces. Of special interest are the electronic states in the vicinity of the crossing point which we obtain from solving Schrödinger's equation. For bound states which are confined both in the x- and y- directions, quasi one - dimensional motion is possible in the z- direction.

Numerical Method

We employ the Finite Element Method for the numerical treatment of both the Poisson and Schrödinger equations. Uniform rectangular elements are used with standard first order Chapeau basis functions. We typically use a mesh size of 100 x 100 nodal points for both equations. The Finite Element Method results in a linear system of equations for the unknowns at each nodal point.

Poisson's equation is solved on a spatial domain of typically 6000 x 6000 Å which we found to be sufficient large to satisfy the Dirichlet boundary condition

that the potential asymptotically approaches its value in the bulk. The boundary conditions fix the potential at the four corners of the rectangular domain such that the corresponding electron density is equal to the background doping of that region, forcing charge neutrality in the bulk. A standard LU decomposition method is used to solve the linear system of equations after it has been reduced to banded symmetric form, and a Newton-Raphson iteration scheme implements the nonlinearity due to the Thomas - Fermi screening model.

Schrödinger's equation is solved on a spatial domain of typically 1000 × 1000 Å surrounding the center point. This region is sufficiently large to obtain bound states by requiring the wavefunction to be equal to zero at the boundary of this domain. The potential used as input to Schrödinger's equation is a bilinearly interpolated detail around the center point of the potential obtained from the solution of Poisson's equation. The resulting eigenvalue problem can be reduced to banded symmetric form of typical dimension 10,000 × 100. Since we are primarily interested in the bound electronic states, only the lowest eigenvalues and eigenvectors need to be determined. The subspace iteration technique is used for the efficient numerical computation of the lowest eigenvalues and eigenvectors.

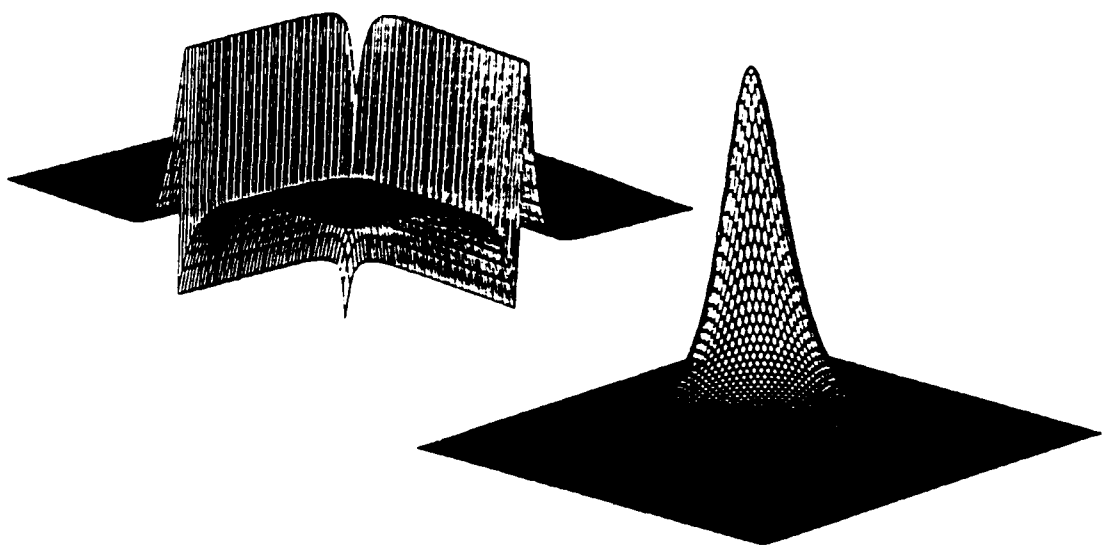


Figure 2: The left part shows the potential landscape for the case in which regions II and III are assumed identical. The potential dip in the center produces a bound electronic state whose wavefunction is displayed in the right half of the figure.

In Fig. 2, we show as a result the potential landscape for a GaAs/$Al_xGa_{1-x}As$ system with mole fraction $x_{II} = x_{III} = 0.35$. Region I is taken to be GaAs with a doping of $N_I = 1 \times 10^{15}$ cm^{-3}, and regions II and III are assumed to be identical with doping $N_{II} = N_{III} = 1 \times 10^{17}$ cm^{-3}. Schrödinger's equation is solved for this potential and a bound state is found with an energy of 14.3

meV relative to the Fermi energy. The corresponding bound state wavefunction is also shown in Fig. 2. The spatial extent of the bound electronic state is about 150 Å in the direction along the heterointerfaces and about 75 Å in the direction perpendicular to the heterointerfaces. More details of our calulations and results can be found in a forthcoming paper [7].

Most numerical computations are performed on a Convex C-2. The solution of Poisson's equation typically takes 600 seconds of CPU time. The solution of the eigenvalue problem for the first 10 states typically requires 200 seconds of CPU time. Computations were also performed on a Cray X-MP/48, with execution times about an order of magnitude faster.

Distributed Computing

We also explored the possibility of utilizing a cluster of networked workstations for distributed computing. The main idea is to take advantage of the fact that in the finite element method all calculation on the local element level are independent and can be done in parallel. Different machines on the computer network can work on different parts of the problem at the same time.

We have implemented such an algorithm on a cluster of 30 SUN-4 Sparc workstations. Remote procedure calls are used to execute the local element calculations on remote machines. So-called network *daemons* handle the transfer of information between the individual computers, dispatch assignments, and manage dynamic network CPU loading. The overall performance depends upon the number of remote machines utilized, and on the number of elements sent to a particular remote machine at a given time. Detailed performance measures will be given in a separate paper.

Acknowledgements: The authors would like to thank Drs. S. Bandyopadhyay, M. A. Herro, C. S. Lent, W. Pötz, and U. Ravaioli for helpful discussions. This work was supported, in part, by the Air Force Office of Scientific Research, the Office of Naval Research, and an IBM Faculty Development Award. Computer time was obtained at NCSA through NCCE.

References

[1] H. Sakaki, *Jap. J. Appl. Phys.* 19, L735 (1980).
[2] Y. Arakawa and A. Yariv, *IEEE J. Quantum Electron.* 22, 1887 (1986).
[3] P. M. Petroff, A. C. Gossard, R. A. Logan, W. Wiegman, *Appl. Phys. Lett.* 41, 635 (1982).
[4] Y.-C. Chang, L. L. Chang, and L. Esaki, *Appl. Phys. Lett.* 47, 1324 (1985).
[5] S. E. Laux and F. Stern, *Appl. Phys. Lett.* 49, 91 (1986).
[6] T. Kerkhoven, A. T. Galick, J. H. Arends, U. Ravaioli, and Y. Saad, *J. Appl. Phys.*, in press.
[7] H. K. Harbury and W. Porod, *J. Vac. Sci. Technol.* B 8, in press.
[8] K.-J. Bathe, *Finite Element Procesure in Engineering Analysis* (Prentice, 1982).
[9] M. R. Frei and D. C. Tsui, *Appl. Phys. Lett.* 55, 2432 (1989).
[10] L. Pfeiffer, K. W. West, H. L. Störmer, J. P. Eisenstein, K. W. Baldwin, D. Gershoni, and J. Spector, *Appl. Phys. Lett.* 56, 1697 (1990).

DOUBLE QUANTUM WIRE AHARONOV-BOHM INTERFEROMETERS FOR POSSIBLE LN_2 TEMPERATURE OPERATION

S. Bandyopadhyay and W. Porod
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

(Received: August 7, 1988)

In this paper, we discuss the design of semiconductor electrostatic and magnetostatic Aharonov-Bohm interferometers that could operate at liquid nitrogen temperature. We find that for elevated temperature operation, one dimensional structures constructed from quantum wires are invariably the only choice, especially when transport is diffusive instead of ballistic. We have proposed such a structure which can be fabricated by present day technology. It may exhibit large conductance modulation in an electric field at 77 K and is an ideal configuration for "Quantum Interference Transistors" (QUITS) based on the electrostatic Aharonov-Bohm effect.

Introduction

Semiconductor interferometers based on the magnetostatic and electrostatic Aharonov-Bohm effect have received widespread attention in recent years because of their potential application in novel transistors with excellent power-delay product[1]. Recently, the electrostatic Aharonov-Bohm effect has been demonstrated in metallic rings[2] leading to an upsurge of interest in this area. In this paper, we discuss various issues pertaining to the device applications of Aharonov-Bohm interferometers at elevated temperatures. We discuss both magnetostatic and electrostatic interferometers but with greater emphasis on the latter since the electrostatic effect is more pertinent to transistor applications.

The basic structure for an Aharonov-Bohm quantum interference transistor consists of two conduction paths connected at both ends. Electrons enter at one end and exit at the other. The quantum-mechanical phase shifts suffered by an electron in traversing the two paths can be controlled by an external electric or magnetic field (the Aharonov-Bohm effect) which controls the interference between the paths and therefore the total conductance. This realizes the transistor operation.

For switching transistor applications, it is important to design the interferometer in such a way as to obtain the largest possible conductance modulation at the highest possible temperature. The size of the conductance modulation is critical. It determines the ratio of the maximum ('ON') conductance to the minimum ('OFF') conductance and therefore the seperation between the logic levels in digital circuits. This seperation must be sufficiently large so that the bit error rate (e.g. in digital communications) is tolerable[3].

For interferometers that rely on the *magnetostatic* Aharonov-Bohm effect, it is possible, in principle, to make the conductance modulation approach 100 % (or the ratio of maximum to minimum conductance approach infinity) by making the minimum (OFF) conductance of the structure almost zero. This can happen under two different circumstances[4]: (1) when transport is perfectly ballistic, and (2) when both elastic and inelastic scattering are present, but the scattering potentials in the two interfering paths are perfectly "correlated". Perfectly correlated scattering gives rise to identical phase shifts in the two paths, so that the *relative* phase shift between the paths, which determines the interference, is not affected at all by such scattering. An example of correlated elastic scattering is the case of impurity scattering when the impurity concen-

0749-6036/89/020239+07 $02.00/0 © 1989 Academic Press Limited

tration *and* the impurity configuration (i.e. the locations of the impurities) in the two paths are identical[5]. An example of partially correlated inelastic scattering is scattering due to polar optical phonons which is a dominant inelastic scattering mechanism in most compound semiconductors at moderate field strengths. For this type of scattering, the interaction of electrons is much stronger with long wavelength phonons than with short wavelength phonons[6]. If the phonon wavelength happens to be much longer than the center-to-center seperation between the two paths, both paths are perturbed almost identically by the scattering event and the relative phase shift between them is not affected at all. Perfect correlation between scattering events is of course an unlikely situation, but in general, if the two paths are physically very close, there can be significant correlation between their scattering potentials, especially when the potentials have long range such as those due to weakly screened impurities or long-wavelength phonons. In any case, it is possible, although not very probable, to observe $\sim$ 100 % conductance modulation in magnetostatic Aharonov-Bohm interferometers under favorable conditions.

The case of *electrostatic* Aharonov-Bohm interferometers however is not so fortuitous. It is not necessarily possible to observe a $\sim$ 100 % conductance modulation in electrostatic interferometers even if transport is ballistic or all scattering events in the two paths are perfectly correlated. The basic reason for this is that unlike the magnetostatic Aharonov-Bohm phase shift which depends only on the magnetic flux enclosed by the interfering paths, the electrostatic phase-shift depends not only on the electrostatic potential difference between the paths but also on the transit time of electrons through the paths. If there is a "spread" in the transit time arising from the fact that different electrons traversing the structure can have different transit times, then there will be a corresponding spread in the phase-shift even when transport is ballistic or scattering events are perfectly correlated. This will dilute the interference effect because of ensemble averaging and reduce the conductance modulation. This deleterious effect of ensemble averaging can be eliminated in magnetostatic inteferometers since the magnetostatic phase-shift can be made unique for all electrons by making the aspect ratio of the structure (the ratio of the distance between the paths to the width of the paths) large. For the electrostatic effect however, the only way to eliminate this deleterious effect is to reduce the spread in the transit time to zero. As we shall see shortly, this can be achieved only if the interfering paths are "strictly single-moded electron waveguides", i.e. one-dimensional structures or quantum wires.

In the regime of ballistic transport, the spread in the transit time of electrons in two or three dimensional structures can arise from two sources: (1) non-zero temperature giving rise to a thermal spread in the electron velocity, and (2) the large spread in the transverse momenta which is present even at 0 K. In one dimensional structures, such as "quantum wires", the latter source is absent. In these structures, there is only one transverse electronic mode (in either transverse direction) and hence the transverse momentum is unique. It is therefore possible to make the spread in the transit time approach zero in 1-d structures by lowering the temperature sufficiently[7]. Consequently, a one dimensional structure can, in principle, exhibit $\sim$ 100 % conductance modulation at low enough temperatures which a two or three dimensional structure can never do even under the most ideal conditions (ballistic transport and zero temperature)[8].

In the case of diffusive transport, the difference between one and poly-dimensional structures is even more pronounced. This is because in diffusive transport, carriers execute a "random walk" motion due to elastic scattering and consequently the spread in the transit time in poly-dimensional structures can be very large. But in 1-d structures, the spread is still very small since the random walk motion is severely restricted. The only permitted random walk in 1-d structures is "backwards and forwards" motion (but no "sideways" motion) since all elastic scattering events involve a 180° deflection of the electron which corresponds to a reflection. Even this reflection is a highly unlikely occurence, especially for high velocity electrons, since the accompanying momentum change is so large that it can only be caused by the short-range (i.e. large wavevector) components of the scattering potential. As long as the scattering potential varies smoothly in space (compared to the scale of a DeBroglie wavelength), such scatterings are practically absent and random walk is essentially prohibited.

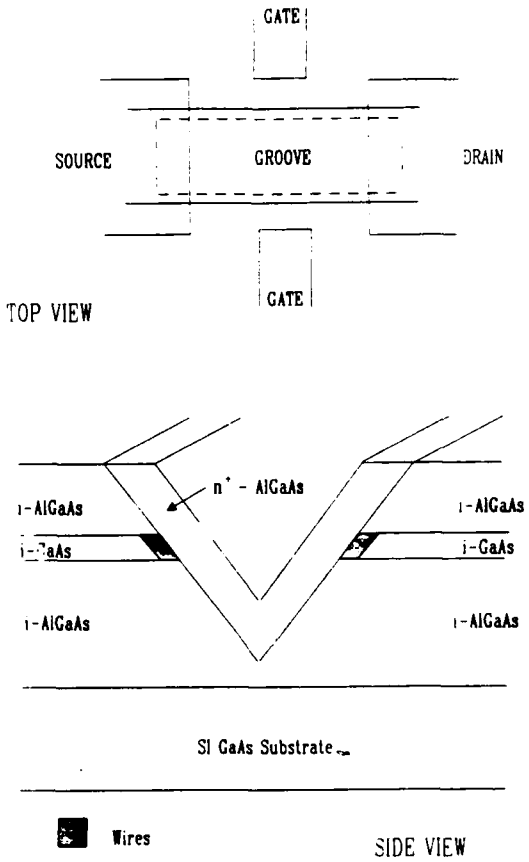
The suppression of elastic scattering in 1-d structures also implies that as long as the length of the structure is shorter than the inelastic mean-free-path, the structure behaves essentially as a ballistic structure (no elastic or inelastic scattering). Consequently, just like in the case of ballistic transport, the spread in the transit time and phase-shift can be reduced almost to zero by reducing the temperature sufficiently. Based on this, we have recently shown[9] that as electrostatic Aharonov-Bohm interferometers, one dimensional structures are vastly superior to two dimensional structures in the diffusive regime. This is especially true at elevated temperatures. Two dimensional structures cannot exhibit sufficiently good performance for device applications at 77 K but one dimensional structures can. Our analysis showed that the maximum temperature of operation for GaAs two-dimensional interferometers (for reasonably fair performance) is $\sim$ 26 K and the maximum allowed carrier concentration is $\sim 6.3 \times 10^{10}$ cm^{-2}. On the other hand, one dimensional GaAs structures can, in principle, exhibit excellent performance even at 77 K if the carrier concentration exceeds 2.5×10^{6} cm^{-1}.

There is another advantage with one-dimensional structures. In order to be able to observe quantum interference effects, the length of the structure must be shorter than the "phase-breaking length". In poly-dimensional structures, the phase-breaking length is the inelastic diffusion length[10] which depends on the diffusion coefficient and hence on the amount of elastic scattering in the sample. But in one-dimensional structures since diffusive motion or random walk is essentially forbidden, the phase-breaking length is actually the inelastic mean-free-path rather than the inelastic diffusion length. This means that the phase-breaking length is longer in 1-d structures (which makes the lithography easier) and it is also completely independent of the degree of disorder or elastic scattering in the sample.

In view of the above, it seems that one-dimensional structures are the best choice for electrostatic Aharonov-Bohm interferometers, especially in the diffusive regime. In the next section we present such a one-dimensional structure. It consists of two closely spaced parallel quantum wires and is derived from a configuration proposed by Sakaki[11]. Such a structure can be fabricated by present-day technology. Our analysis shows that this structure can exhibit very large (approaching 100% in principle) conductance modulation in a magnetic field and can also exhibit quite large conductance modulation in an electric field at temperatures as high as 77 K.

Proposed Structure

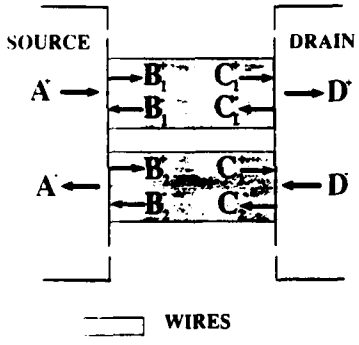
The proposed structure is shown in Fig. 1. It consists of a single undoped GaAs quantum well sandwiched between intrinsic AlGaAs layers. After etching a *narrow* V-groove through the quantum well by focussed ion beam milling or electron beam exposure[12], a n^+ AlGaAs layer is regrown on the etched surface. These steps can all be performed in ultrahigh vacuum without ever breaking the vacuum. The process of "etching and regrowth" is certainly a difficult step, but it has been demonstrated recently[13]. Following successful regrowth, two parallel closely-spaced "quantum wires" will form as accumulation layers in the GaAs quantum well if spatial transfer of charges from the n^+ AlGaAs layer to the GaAs layer takes place. Even if the spatial transfer does not occur, there may still be enough carriers in the channel generated by positively charged interface states. The mobility of these carriers will be poor, but the mobility is not important in this case. The only major problem may arise due to Fermi level pinning. If the Fermi level gets pinned inside the bandgap, the wires will be depleted of carriers. This problem does not arise in InAs systems. There have been reports of inversion layers forming under natively-grown oxides on InAs[14]. It may be advantageous to replace the GaAs quantum well with an InAs quantum well since then one merely has to grow a native oxide on the etched surface of the V-groove to generate the carriers. This is much easier than effecting spatial transfer of charges across the V-groove surface.



• Fig. 1. Proposed double quantum wire structure for electrostatic Aharonov-Bohm interferometer. The figure shows both the top view and the side view.

Another problem may arise due to electron localization effects. It is important to ensure that the device does not operate in the regime of strong localization. This can be ensured by making the length of the structure shorter than the localization length.

The quantum wires formed at the surface of the V-groove can be contacted by either Au-Ge alloying or by Si implantation[15]. The latter is preferable since it creates rather lightly doped contacts and this has an advantage. This issue is discussed later in the paper. The electrostatic



• Fig. 2. Schematic representation of the structure in Fig. 1 showing the incident and reflected electron amplitudes.

potential inducing the Aharonov-Bohm effect is applied between two gate pads (see the top view in Fig. 1).

In order for the structure to exhibit a strong interference effect, its length must be shorter than the "phase-breaking length" at the operating temperature. As discussed before, in quantum wires, this length is the inelastic mean-free-path which depends on the temperature and carrier concentration. The inelastic scattering time in heavily doped GaAs wires at 4.2 K has been reported to exceed 5 ps[16] so that assuming a $T^{-\frac{1}{2}}$ dependence of the inelastic scattering time on temperature[17], we find that the inelastic mean free path at 77 K is larger than 0.3 μm if the carrier concentration is 10^6 cm^{-1}. The length of the structure can be easily made to be 0.25 μm or shorter with electron beam lithography. Our analysis shows that, in principle, such a structure can exhibit $\sim$ 100 % modulation of the conductance in a magnetic field and $\sim$ 90 % modulation of the conductance in an electric field even at liquid nitrogen temperature.

Analysis

For purposes of analysis, the proposed structure is schematically represented as shown in Fig. 2. The current I through the structure is given by the Tsu-Esaki formula[18]

$$I = \frac{2q}{h} \int dE \, |T_{total}(E)|^2 \, [f(E,T) - f(E+qV,T)] \quad (1)$$

For a small applied bias voltage V, this reduces to

$$\begin{aligned} I &= \tfrac{2q^2V}{h} \int dE \, |T_{total}(E)|^2 \, [-\tfrac{\partial f(E,T)}{\partial E}] \\ &= \tfrac{q^2V}{2hKT} \int dE \, |T_{total}(E)|^2 \, sech^2(\tfrac{E - E_F}{2KT}) \end{aligned} \quad (2)$$

which gives the conductance as

$$G = \frac{I}{V} = \frac{q^2}{2hKT} \int dE \, |T_{total}(E)|^2 \, sech^2(\frac{E - E_F}{2KT}) \quad (3)$$

Here T_{total} is the amplitude of the total transmission through the structure from contact to contact, E is the electron's energy, T is the temperature and f is the Fermi-Dirac distribution function in the contacts. The problem is now to calculate T_{total}.

As discussed before, a one-dimensional structure essentially acts as a ballistic structure if its length is shorter than the inelastic mean-free-path. Elastic scattering is suppressed by one-dimensional confinement and inelastic scattering is absent since the transit time through the structure is smaller than the inverse of the inelastic scattering rate. In that case, the transmission amplitude T_{total} can be found in a straightforward manner by using the formalism of ballistic transport. This is done by cascading the three scattering matrices[19] representing propagation from the left contact region to the interfering paths, propagation along the paths, and propagation from the paths to the right contact. The first and the last of these scattering matrices (for junctions A-B and C-D; see Fig. 2) can, in principle, be found exactly by matching the wavefunctions and their derivatives along the junction between the contacts and paths[20]. However, for simplicity, we will represent these scattering matrices by the Shapiro matrix[21]

$$\begin{pmatrix} A^- \\ B_1^+ \\ B_2^+ \end{pmatrix} = \begin{pmatrix} -(a+b) & \sqrt{\epsilon} & \sqrt{\epsilon} \\ \sqrt{\epsilon} & a & b \\ \sqrt{\epsilon} & b & a \end{pmatrix} \cdot \begin{pmatrix} A^+ \\ B_1^- \\ B_2^- \end{pmatrix} \quad (4)$$

where

$$a = \frac{1}{2}(\sqrt{1 - 2\epsilon} - 1) \quad (5)$$

$$b = \frac{1}{2}(\sqrt{1 - 2\epsilon} + 1) \quad (6)$$

The amplitudes A and B are defined in Fig. 2. The superscript '+' refers to waves traveling from left to right and '-' refers to waves traveling in the opposite direction. In the Shapiro matrix, ϵ represents the probability of transmission from the contact into any one path; ϵ = 0.5 corresponds to perfect transmission (no end-reflection). The Shapiro matrix implicitly assumes that transmission into the two paths from the contact are equal in *both* magnitude and phase. The latter condition is more difficult to meet in practice, but if the carrier concentration in the contacts is not too large and the channels are physically very close so that the seperation between them is comparable to the DeBroglie wavelength in the contacts, then some degree of phase coherence in the injection and detection process can be expected[22]. This consideration makes it necessary to make the V-groove in the proposed structure very narrow so that the wires are closely spaced and also have the contacts defined by Si implantation rather than by Au-Ge alloying. The transmission amplitudes for propagation from junction B to C are represented by t^+_{12} for

left to right propagation along paths 1 and 2 and $t^-_{1,2}$ for reverse propagation along these paths. For single moded structures, the task of cascading these scattering matrices is relatively simple and can be performed analytically to yield analytical expressions for T_{total}[23] in the presence of an external magnetic or electric field.

Magnetostatic Effect

Assuming that the two paths are identical so that in the absence of any external field $t^+_1 = t^+_2 = t^+$ and $t^-_1 = t^-_2 = t^-$, T_{total} is given by[23]

$$T_{total} = \frac{\epsilon t^+(1+e^{-i\theta})(1-t^+t^-)}{[1-t^+t^-(a^2+b^2e^{i\theta})][1-t^+t^-(a^2e^{i\theta}+b^2)]-a^2b^2t^{+2}t^{-2}(e^{i\frac{\theta}{2}}+e^{-i\frac{\theta}{2}})^2} = (1+e^{-i\theta})T'_{total}(\theta, E) \quad (7)$$

where θ is the magnetostatic Aharonov-Bohm phase shift given by

$$\theta = \frac{qBA}{\hbar} \quad (8)$$

It is to be noted that θ depends only on the magnetic flux density B and the area A enclosed by the two paths and *does not depend on the electron's energy*. Consequently, substitution of Equation (7) in Equation (3) gives

$$G = |1+e^{-i\theta}|^2 \frac{q^2}{2hKT} \int dE\, |T'_{total}(\theta, E)|^2\, sech^2(\frac{E-E_F}{2KT}) = cos^2(\frac{\theta}{2})G_0(\theta) \quad (9)$$

where

$$G_0(\theta) = \frac{4q^2}{hKT} \int dE\, |T'_{total}(\theta, E)|^2\, sech^2(\frac{E-E_F}{2KT}) \quad (10)$$

Equation (9) predicts a 100% modulation of the conductance in a magnetic field. The conductance oscillation however is not necessarily sinusoidal since $G_0(\theta)$ depends on the magnetic field through θ. The non-sinusoidal shape is caused by higher harmonics generated from multiple reflections of the electron between the contacts. A 100 % conductance modulation is also possible for a 2-d structure if transport is perfectly ballistic[4] or if the scattering events in the two interfering paths are perfectly correlated. Correlated scattering is difficult to realize in practice and for ballistic transport, the length of the structure has to be shorter than *both* the elastic and inelastic mean free path. The elastic mean-free-path in two-dimensional (double quantum well) structures can be quite small, even when the wells are modulation-doped, since one of the interfaces of each well will always be an inverted interface. Fabrication of ballistic structures therefore places extreme demands on the lithography. In contrast, for a 1-d structure, transport does not have to be ballistic; it merely has to be non-dissipative, i.e. there should not be any inelastic scattering. Consequently the length has to be merely shorter than the inelastic mean-free-path which can be quite long at low temperatures. This makes the lithographic delineation of these structures much easier.

Electrostatic Effect

For the electrostatic effect, Equation (7) is modified to[23]

$$T_{total} = \frac{\epsilon t^+(1+e^{i\phi})(1-t^+t^-e^{i\phi})}{[1-t^+t^-(a^2+b^2e^{i\phi})][1-t^+t^-(a^2e^{i\phi}+b^2e^{2i\phi})]-a^2b^2t^{+2}t^{-2}e^{i\phi}(1+e^{i\phi})^2} \quad (11)$$

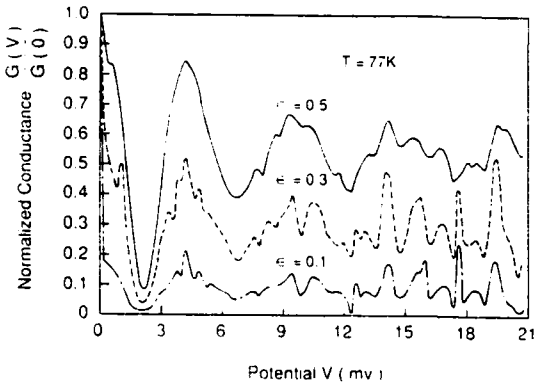
where ϕ is the electrostatic Aharonov-Bohm phase shift given by

$$\phi = \frac{e}{\hbar}V\tau_t = \frac{\sqrt{2m^*E}}{\hbar}[\sqrt{1+\frac{eV}{E}} - 1]L \quad (12)$$

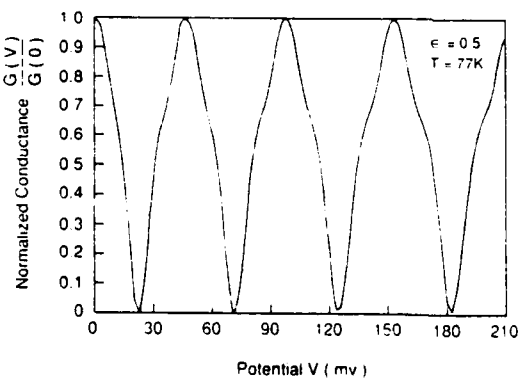
Here τ_t is the average (harmonic mean) of the transit times through the two paths, L is the length of each path and V is the electrostatic potential difference between the two paths.

Since ϕ depends explicitly on the transit time and hence on the electrons energy (unlike in the case of the magnetostatic effect), different electrons having different energy (at elevated temperatures) suffer different phase-shifts so that the electrostic effect does not escape the deleterious effects of ensemble averaging unlike its magnetostatic counterpart. In order to prevent this from happening, the temperature must be lower than the so-called Thouless temperature T_{Th} where kT_{Th} is equal to the energy seperation between the electron modes near the Fermi energy. For a 1-d structure, $kT_{Th} = \frac{\hbar v_F}{L}$ where v_F is the Fermi velocity and L is the length of the structure.

To examine the conductance modulation in a electric field, the integral in Equation (3) was performed numerically and the results are shown in Fig. 3 (at T = 77 K) for various values of ϵ. Note that $\epsilon = 0.5$ corresponds to perfect transmission from the contact into the paths and $\epsilon = 0.1$ corresponds to 10 % transmission probabilty. In the calculation, we neglected any dependence of ϵ on the electron's wave-vector. The carrier concentration was assumed to be 10^6 cm^{-1}, the length of the structure was 0.25 μm and the material was GaAs. For this structure, the Thouless temperature is 8 K. For a 1000 Angstroms long structure



• Fig. 3. Conductance vs. potential characteristics for the structure in Fig. 1. The carrier concentration is 10^6 cm^{-1}, the length is 0.25 μm, the material is GaAs and the temperature is 77 K. For this structure, the Thouless temperature is 8 K. The characteristics are shown for various values of ϵ.



• Fig. 4. Conductance vs. potential characteristic for the same structure at 77 K when the carrier concentration is 4 x 10^6 cm^{-1} and the length is 1000 Angstroms. In this case the Thouless temperature is $\approx$ 77 K which is the ambient temperature.

with a carrier concentration of 4 x 10^6 cm^{-1}, the Thouless temperature is about 77 K. For the sake of comparison, we have also shown the conductance modulation of such a structure in Fig. 4 for $\epsilon = 0.5$.

Several interesting features are found in Fig. 3. The conductance modulation (for the first half-cycle) is larger than 90 % at 77K for all values of ϵ. This is promising for switching transistor applications. For such applications, it is only the first half-cycle of the oscillation which is important since all that is required is to switch the conductance between the ON and OFF states. Consequently a 90 % conductance modulation over the first half-cycle is encouraging. Another interesting feature is the decay of the oscillations with increasing values of the electrostatic potential. This happens because the uncertainty in the phase-shift for a fixed potential V is given by $\Delta\phi = V\Delta\tau_t$ and this is proportional to the potential itself. At low values of the potential, the uncertainty $\Delta\phi$ is small and the conductance modulation is large. At higher values of the potential, the uncertainty increases thereby decreasing the conductance modulation.

Perhaps the most interesting feature in Fig. 3 is the effect of ϵ or the role of multiple reflections. If ϵ is small (large end-reflections), only the first half-cycle of the oscillation is discernible and the later cycles are not[24]. This will make it impossible to detect the presence of the Aharonov-Bohm oscillations in a direct experiment if the test structure is not cleverly designed to eliminate such reflections. An explanation for this role of multiple reflections is the following. If an electron suffers many reflections back and forth between the contacts before it exits the structure, its effective transit time increases[25]. This, in turn, increases the spread in the transit time thereby increasing the uncertainty in the phase-shift for a given potential. As a result, the interference effect "dies off" much more rapidly in the presence of multiple reflections. There is another harmful effect of multiple reflections. An increase in the effective transit time may cause it to exceed the inelastic scattering time and the interference effect may be summarily destroyed by inelastic scattering. For experiments designed to demonstrate the electrostatic Aharonov-Bohm oscillations, it is important to minimize multiple reflections. This can be achieved by designing the geometry of the structure in such a way that the contacts are "transparent" to the electrons. Conventional semiconductor "ring" structures that have been used for experiments[26] are not ideal in this respect. In such structures, the radius of curvature of the ring may be comparable to the DeBroglie wavelength of electrons which will inevitably lead to severe reflections between the leads[27]. The present structure is much better designed in this respect since there are no sharp curvatures in the geometry to induce reflections.

In conclusion, we have discussed the relative merits of one dimensional Aharonov-Bohm interferometers over two and three dimensional interferometers in the regime of both ballistic and diffusive transport. We have shown that cleverly designed double quantum wire structures with suitable geometries to minimize multiple reflection effects can exhibit large Aharonov-Bohm interference. We have proposed such a structure that can be fabricated by present-day technology. This structure is especially suited for device applications at elevated temperatures (77K).

Acknowledgement: This work was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-88-0096 and, in part, by the Office of Naval Research under Grant No. N00014-86-K-0506.

REFERENCES

1. S. Datta, M. R. Melloch, S. Bandyopadhyay and M. S. Lundstrom, *Applied Physics Letters*, **48**, 487 (1986); S. Bandyopadhyay, S. Datta and M. R. Melloch, *Superlattices and Microstructures*, **2**, 539 (1986); S. Bandyopadhyay et. al., Technical Digest of the IEDM, 76 (1986). For a review, see, S. Datta, "Quantum Interference Devices" Ch. 8, in *Physics of Quantum Electron Devices*, ed. F. Capasso, (Springer-Verlag, New York, to appear) and references therein.

2. S. Washburn, H. Schmid, D. Kern and R. A. Webb, *Physical Review Letters*, **59**, 1791 (1987)).

3. This is especially true if the devices have a soft threshold. Quantum Interference Transistors based on the electrostatic Aharonov-Bohm effect typically have a soft threshold.

4. S. Datta and S. Bandyopadhyay, *Physical Review Letters*, **58**, 717 (1987).

5. M. Cahay, PhD thesis, Purdue University, 1987.

6. This is due to the anisotropic nature of polar optical phonon scatterring. Smaller angle scatterings are preferred over larger angle scatterings so that electrons preferentially interact with longer wavelength phonons.

7. The temperature should be lower than the Thouless temperature which depends on the length of the structure and the carrier concentration.

8. The maximum conductance modulation for a two-dimensional structure under the most ideal conditions is ~ 75 %.

9. S. Bandyopadhyay and W. Porod (to be published).

10. For a discussion of the "phase breaking length" determined by inelastic scattering, see, for example, S. Washburn and R. A. Webb, *Adv. Phys.*, **35**, 375 (1986).

11. H. Sakaki, *Journal of Vacuum Science and Technology*, **19**, 148 (1981).

12. Linewidths of ~ 80 Angstroms are achievable with electron beam exposure. See, for example, F. Emoto, K. Gamo, S Namba, N. Samoto and R. Shimizu, *Japanese Journal of Applied Physics*, **24**, L809 (1985).

13. S. M. Goodnick, private communication.

14. D. A. Baglee, D. K. Ferry, C. W. Wilmsen and H. H. Wieder, *Journal of Vacuum Science and Technology*, **17**, 1032 (1980).

15. R. E. Williams in *GaAs Processing Techniques*, Ch. 11, (Artech House, 1987).

16. K. Ishibashi et. al., *Solid State Communication*, **61**, 385 (1987).

17. P. Santhanam, S. Wind and D. E. Prober, *Physical Review Letters*, **53**, 1179 (1984).

18. R. Tsu and L. Esaki, *Applied Physics Letters*, **22**, 562 (1973). This is also the Landauer formula for a two-terminal measurement.

19. Rules for cascading scattering matrices are discussed in: P. W. Anderson, *Physical Review*, **B23**, 4828 (1981) and B. Shapiro, *Physical Review*, **B35**, 8256 (1987).

20. A. M. Kriman and P. Ruden, *Physical Review*, **B32**, 8013 (1985). A specific application of such a technique to Aharonov-Bohm devices has been reported by R. Frohne and S. Datta (to appear in *Journal of Applied Physics*).

21. B. Shapiro, *Physical Review Letters*, **50**, 747 (1983). The Shapiro matrix has been widely used in the analysis of doubly-connected geometries; see, for example, M. Büttiker, *Physical Review*, **B32**, 1846 (1985).

22. R. Frohne and S. Datta (to appear in *Journal of Applied Physics*).

23. M. Cahay (preprint).

24. For switching transistor applications, this may actually be a desirable feature since it allows for some tolerance in the threshold voltage and may in fact mitigate the problem of threshold variability. However for experiments intended to demonstrate the electrostatic Aharonov-Bohm effect, this is disastrous outcome since it can mask the effect totally.

25. As a result of the increase in the effective transit time, the threshold voltage for turning off the device also decreases. The threshold voltage is given by $V_{th} = \frac{\epsilon\hbar}{e<\tau_t>}$. If $<\tau_t>$ increases, V_{th} decreases and an indication of this is clearly seen in Fig. 3 where the conductance drops off much more rapidly for lower values of ϵ.

26. G. L. Timp et. al., *Physical Review Letters*, **58**, 2814 (1987).

27. Alternate ring structures that alleviate this problem have been proposed in Ref. 1 and Ref. 4.

Performance of electrostatic Aharonov–Bohm interferometers in the diffusive regime

S. Bandyopadhyay and W. Porod
Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556

(Received 20 July 1988; accepted for publication 26 September 1988)

In this letter we discuss the performance of semiconductor electrostatic Aharonov–Bohm interferometers that operate in the *diffusive* regime. We find that the performance is primarily determined by temperature and carrier concentration, and we have identified the conditions for "fair," "good," and "excellent" performance. Our analysis shows that two-dimensional interferometers cannot operate at elevated temperatures but one-dimensional interferometers could operate at 77 K if the carrier concentration is sufficiently high and the structure is cleverly designed to minimize end reflections.

Semiconductor electrostatic Aharonov–Bohm interferometers have received serious attention in recent years because of their potential application in novel transistors with excellent power-delay product.[1] In this letter we examine the performance of these interferometers in the *diffusive* regime, i.e., when elastic scattering is present, but phase-randomizing inelastic scattering is absent. In the past, most of the research dealing with such interferometers has focused on the *ballistic* regime where both elastic and inelastic scatterings are absent. Fabrication of ballistic structures, with dimensions shorter than the inelastic *and* elastic mean free path, places extreme demands on semiconductor technology. The demand is twofold. Firstly, modulation doping is required to eliminate *in situ* impurity scattering. Secondly, the lithographic demands can be imposing since the elastic mean free path can be quite small, even in modulation-doped quantum well structures, since one of the interfaces of the well will always be an inverted interface. Structures meant for the diffusive regime, on the other hand, are much easier to fabricate. Modulation doping is not required and the lithographic demands are significantly relaxed since the length of the structure has to be merely shorter than the phase-breaking length which is typically much longer than the elastic mean free path at low temperatures.

Recently, the electrostatic Aharonov–Bohm effect has been demonstrated in metallic rings[2] in which transport is diffusive. The observed effect was small and indirect since metallic rings are not ideal for this purpose. They are not strictly one-dimensional structures (the diameter of the wires is much larger than the DeBroglie wavelength of carriers) and this has a deleterious effect. For electrostatic Aharonov–Bohm interferometers, one-dimensional structures are best. They are inherently superior to two- or three-dimensional structures, especially when transport is diffusive rather than ballistic. This is elucidated below.

The electrostatic Aharonov–Bohm phase shift depends on the transit time of electrons through a structure. If there is a "spread" in the transit time, there will be a corresponding spread in the phase shift and this will dilute the interference effect as a result of ensemble averaging.

In the *ballistic* regime, the spread in the transit time in 2-d structures can arise from two sources: (1) nonzero temperature giving rise to a nonzero spread in the velocity of electrons, and (2) the large spread in the transverse momenta which is present even at 0 K. In 1-d structures, the latter source is absent (since there is only one transverse mode), and consequently the spread in the transit time can approach zero at low enough temperatures. As a result, a 1-d electrostatic Aharonov–Bohm interferometer can exhibit almost a 100% conductance modulation at low enough temperatures which a 2-d interferometer can never do even at zero temperature.[3]

In the *diffusive* regime, the difference between 1-d and 2-d structures is even more pronounced. The spread in the transit time in 2-d structures can be very large since carriers execute a "random walk" motion due to elastic scattering. But in 1-d structures, the spread is still very small since the "random walk" motion is severely restricted. The only permitted "random walk" in 1-d structures is "backwards and forwards" motion (but no "sideways" motion) since elastic scattering events involve a 180° deflection of the electron. Even this is highly unlikely, especially for high-velocity electrons, since the accompanying momentum change is so large that it can only be caused by the short-range (i.e., large wave vector) components of the scattering potential. As long as the scattering potential in a 1-d structure varies smoothly in space (compared to the scale of a DeBroglie wavelength) such scatterings are practically absent.

The suppression of elastic scattering and random walk in 1-d structures makes the spread in the transit time almost zero at low enough temperatures. In the next paragraphs, we focus on the specific issue of phase randomization due to the spread in the transit time and the resultant performance degradation. We then establish a performance criterion based on this consideration to evaluate various 2-d and 1-d interferometers in the diffusive regime under different conditions of temperature and carrier concentration.

The electrostatic Aharonov–Bohm phase shift between two interfering paths in an interferometer is given by

$$\phi = (e/\hbar) V \tau_t, \tag{1}$$

where V is the potential difference between the paths and τ_t is the average (harmonic mean) of the transit times through the two paths. Any spread in the transit time will give rise to a spread in the phase shift and dilute the interference effect thereby causing the conductance modulation to decrease.

 © 1988 American Institute of Physics

For switching transistor application, it is necessary to ensure that the conductance of the interferometer is close to zero at the trough of the first half-cycle of the Aharonov–Bohm oscillations. This in turn requires that when one electron is interfering destructively corresponding to an Aharonov–Bohm phase shift $\phi = \pi$, every other electron in the ensemble also suffers a phase shift close to π. If the spread in the phase shift $\Delta\phi$ is also close to π, then some electrons will interfere constructively while others are interfering destructively and the effect is washed out. In other words, it is necessary to have $\Delta\phi < \pi$. For switching transistor applications, it is only the first half-cycle of the conductance modulation that is important. Therefore, we need to have $\Delta\phi < \pi$ when $\phi = \pi$. This is equivalent to the condition that the ratio $\Delta\phi/\phi < 1$. We now adopt this ratio as a "performance index" (η) and require that for a sufficiently large conductance modulation (peak-to-valley ratio), η be less than unity. We then calculate this dimensionless quantity η for 2-d and 1-d interferometers.

$$\eta = \Delta\phi/\phi = \langle \Delta\tau_t \rangle / \langle \tau_t \rangle, \tag{2}$$

where $\langle \Delta\tau_t \rangle$ is the spread in the transit time and $\langle \tau_t \rangle$ is the average transit time for the ensemble.

For 2-d structures and in the case of diffusive transport, $\langle \Delta\tau_t \rangle$ can be calculated the same way as is done for the Shockley–Haynes experiment.[4]

$$\langle \Delta\tau_t \rangle = 4/v_d \sqrt{DL/v_d}, \tag{3}$$

where D is the diffusion coefficient, v_d is the drift velocity, and L is the length of the structure.

$$\langle \tau_t \rangle = L/v_d. \tag{4}$$

Therefore,

$$\eta = 4/L \sqrt{DL/v_d}. \tag{5}$$

For a low electric field ϵ, $v_d = \mu\epsilon$ where μ is the mobility. Also $\epsilon \approx V_{\text{bias}}/L$, where V_{bias} is the bias voltage over the structure. This gives

$$\eta = \sqrt{(16/V_{\text{bias}})(D/\mu)}. \tag{6}$$

For 2-d interferometers, η depends on the bias. The minimum value of η corresponds to the maximum allowed value for V_{bias}. The upper limit on V_{bias} is the voltage at which an electron, arriving at one contact from the other, just reaches sufficient energy to cross the threshold for polar optical phonon emission (strong inelastic scattering). Hence from Eq. (6), invoking the generalized Einstein relation for a carrier concentration n_s, we get that for a 2-d interferometer

$$\eta_{\min} = \sqrt{(16kT/\epsilon_{\text{POP}}) \ln\{1 + \exp[(E_F - E_0)/kT]\}\{1 + \exp[(E_0 - E_F)/kT]\}}, \tag{7}$$

where E_0 is the energy of the lowest electronic subband (the only one presumed to be occupied), E_F is the Fermi energy, and ϵ_{POP} is the polar optical phonon energy (= 36 meV for GaAs). Equation (7) gives us the limiting values of temperature and carrier concentration for which $\eta_{\min} < 1$. In Fig. 1 we show the performance of 2-d GaAs interferometers as a function of temperature and carrier concentration. The performance is considered "fair" if $\eta_{\min} < 1$, "good" if $\eta_{\min} < 0.5$, and "excellent" if $\eta_{\min} < 0.1$. For 2-d interferometers, there is no visible region of "excellent" performance. We also find that 2-d interferometers cannot operate at 77 K which is far outside the range of "fair" performance. In addition, we find that even for "fair" performance, the maximum temperature of operation is ~ 26 K (for the lowest carrier concentration) and the maximum allowed carrier concentration is $\sim 6.3 \times 10^{10}$ cm^{-2} (at the lowest temperature). Two-dimensional interferometers are therefore not a judicious choice for device application in the diffusive regime.

We now discuss 1-d interferometers. In such systems, the only source of a spread in the transit time is the thermal smearing of the electron distribution. Therefore,

$$\eta = \langle \Delta\tau_t \rangle / \langle \tau_t \rangle \approx \langle \Delta v_t \rangle / \langle v_t \rangle, \tag{8}$$

where $\langle v_t \rangle$ is the average transit velocity and $\langle \Delta v_t \rangle$ is the spread in the electron velocity arising from the thermal spread in energy.

For degenerate carrier concentrations, $\langle v_t \rangle = v_F$ (the Fermi velocity) and $\langle \Delta v_t \rangle = (1/2)\sqrt{kT/m^*}$. Hence,

$$\eta = \sqrt{kT/8E_F}. \tag{9}$$

In contrast to the case of 2-d interferometers, η for 1-d interferometers does not depend on the bias.

In Fig. 2 we show the performance of 1-d interferome-

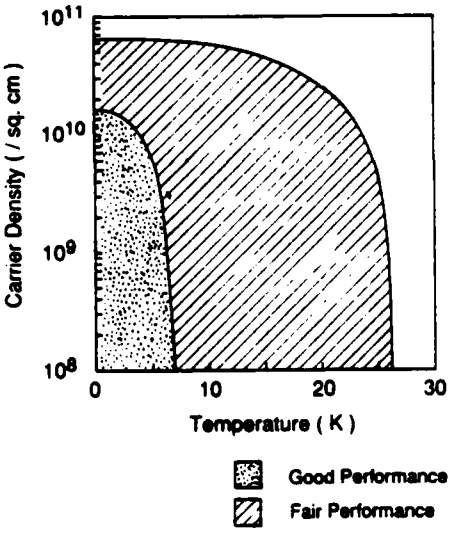


FIG. 1. Performance diagram for two-dimensional GaAs electrostatic Aharonov–Bohm interferometers operating in the diffusive regime. Regions of "fair" and "good" performance are shown in the diagram. There is no region of excellent performance.

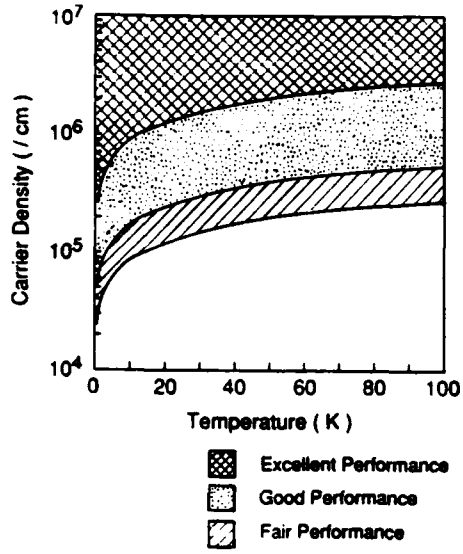


FIG. 2. Performance diagram for one-dimensional GaAs electrostatic Aharonov–Bohm interferometers operating in the diffusive regime. There are regions of "fair," "good," and "excellent" performance. "Good" to "excellent" performance can be expected at 77 K for practical carrier concentrations of $\sim 10^6$ cm^{-1}.

ters for various temperatures and (degenerate) carrier concentrations. *Unlike in the case of 2-d interferometers, there is a region of "excellent" performance.* We find that one can expect "fair" performance at liquid-nitrogen temperature if the carrier concentration exceeds 2.5×10^5 cm^{-1}, "good" performance if it exceeds 5×10^5 cm^{-1}, and "excellent" performance if it exceeds 2.5×10^6 cm^{-1}. The performance improves with increasing carrier concentration. However, as the carrier concentration is increased, electron-electron scattering (which is an inelastic mechanism) also becomes more frequent and the inelastic mean free path (the phase-breaking length) becomes shorter which increases the demands on lithography. Nevertheless, we have found that a 0.25-μm-long structure with a carrier concentration of $\sim 10^6$ cm^{-1} can exhibit $>90\%$ conductance modulation at 77 K.[5] Realization of such structures is well within the capability of present day technology.

Although 1-d interferometers have the *potential* for excellent performance at 77 K, there are other critical design issues that must be addressed before such performance can be expected. Perhaps the most critical issue is the role of the contacts. The contacts must infuse and extract carriers from the two interfering paths phase coherently and this is achievable to some extent if the interfering paths are closely spaced and the carrier concentration in the contacts is not too large.[6] Lightly doped contacts can be realized by Si implantation (for GaAs-AlGaAs structures).[7] In addition, the contact geometry must also be such that the contacts are "transparent" to the electrons. Otherwise, an electron will suffer many reflections back and forth between the contacts before it finally exits the structure. Multiple reflections have two deleterious effects. Firstly, since the transit time through the structure increases proportionately with the number of reflections, the spread in the transit time also increases, which in turn reduces the conductance modulation. Secondly, the dwell time of an electron within the structure increases and this enhances its chances of encountering a phase-randomizing inelastic collision. The geometry of the structure is therefore a critical consideration in the design. Semiconductor ring structures, which are conventionally used for experiments,[8] are a poor design in this respect since the radius of curvature of the ring is usually comparable to the DeBroglie wavelength of carriers so that multiple reflections between the leads (contacts) can be severe.[9] Alternate structures that do not have sharp bends and curvatures are possibly a better choice. One such structure has been proposed by us in Ref. 5.

To summarize, we have shown that well-designed one-dimensional interferometers have the potential to operate at liquid-nitrogen temperature. We have also identified the temperatures and carrier concentrations required for various levels of performance.

This work was supported by the Air Force Office of Scientific Research under grant No. AFOSR 88-0096.

[1] S. Datta, M. R. Melloch, S. Bandyopadhyay, and M. S. Lundstrom, Appl. Phys. Lett. **48**, 487 (1986); S. Bandyopadhyay, S. Datta, and M. R. Melloch, Superlatt. Microstruct. **2**, 539 (1986); S. Bandyopadhyay, M. R. Melloch, S. Datta, B. Das, J. A. Cooper, Jr., and M. S. Lundstrom, Technical Digest of the IEDM, Cat. No. 86CH2381-2, 76 (1986).
[2] S. Washburn, H. Schmid, D. Kern, and R. A. Webb, Phys. Rev. Lett. **59**, 1791 (1987).
[3] The maximum conduction modulation for a two-dimensional structure under ideal conditions is $\sim 75\%$ (Ref. 1).
[4] B. G. Streetman, *Solid State Electronic Devices* (Prentice–Hall, Englewood Cliffs, NJ, 1972).
[5] S. Bandyopadhyay and W. Porod, presented at the *Fourth International Conference on Superlattices, Microstructures and Microdevices*, Miramare, Trieste, Italy, August 1988 (unpublished).
[6] R. Frohne and S. Datta (unpublished).
[7] R. E. Williams, in *GaAs Processing Techniques* (Artech House, Dedham, MA, 1987), Chap. 11.
[8] G. Timp, A. M. Chang, J. E. Cunningham, T. Y. Chang, P. Mankiewich, R. Beringer, and R. E. Howard, Phys. Rev. Lett. **58**, 2814 (1987).
[9] This issue has been discussed by S. Datta and S. Bandyopadhyay, Phys. Rev. Lett. **58**, 717 (1987).

DOUBLED FREQUENCY OF THE CONDUCTANCE MINIMA IN ELECTROSTATIC AHARONOV-BOHM OSCILLATIONS IN ONE-DIMENSIONAL RINGS[1]

M. Cahay[(a)], S. Bandyopadhyay[(b)] and H. L. Grubin[(a)]

[(a)]Scientific Research Associates, Inc.
Glastonbury, Connecticut 06033

[(b)]Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

We predict the existence of *two* different sets of conductance minima in the conductance oscillation of a one-dimensional ring due to the electrostatic Aharonov-Bohm effect. The two sets of minima arise from two different conditions and effectively *double* the frequency of the conductance troughs in the oscillations. This makes the frequency of the troughs *twice* that predicted by the Aharonov-Bohm effect. We discuss the origin of this feature along with the effects of temperature and elastic scattering. We also compare it with the magnetostatic Aronov-Al'tshuler-Spivak effect and point out the similarities and differences.

I. INTRODUCTION

Oscillatory conductance due to the electrostatic Aharonov-Bohm effect has been predicted for a variety of ring structures along with potential device applications of that effect. In this paper, we point out an intriguing feature in the conductance oscillation of a one-dimensional ring due to the electrostatic Aharonov-Bohm effect. Unlike in the magnetostatic effect, the conductance in the electrostatic effect reaches its minimum under two *different* conditions which gives rise to *two*

[1]The work at SRA was supported by the Air Force Office of Scientific Research under contract no. F49620-87-C-0055. The work at Notre Dame was supported by the same agency under grant no. AFOSR-88-0096 and by an IBM Faculty Development Award.

Copyright © 1989 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-585000-X

distinct and independent sets of conductance minima in the oscillations. One set of minima arises from the usual destructive interference of transmitted electrons and the other arises from constructive interference of reflected electrons. The minima in each individual set recur in the oscillations with the periodicity predicted by the Aharonov-Bohm effect, but the separation between two adjacent minima (belonging to the two different sets) is smaller than and unrelated to the Aharonov-Bohm periodicity. In the following Sections, we establish this feature and discuss various issues related to it.

II. THEORY

The conductance G of a one-dimensional structure in the linear response regime is given by the two-probe Landauer or Tsu-Esaki formula [1]

$$G = \frac{e^2}{2hkT} \int dE \, |T_{total}(E)|^2 \, sech^2(\frac{E - E_F}{2kT}) \, , \qquad (1)$$

where $T_{total}(E)$ is the transmission coefficient of an electron with incident energy E through the entire structure (i.e. from one contact to the other), T is the temperature and E_F is the Fermi level.

The problem of calculating the conductance G is essentially the problem of calculating T_{total}. The quantity T_{total} can be found from the overall scattering matrix for the structure. For a ring structure, the overall scattering matrix is determined by cascading three scattering matrices [2] representing propagation from the left lead of the ring to the two interfering paths, propagation along the paths, and propagation from the paths to the right lead. For simplicity, we will represent the first and the last of these scattering matrices by the so-called Shapiro matrix which is defined in Ref. 3.

A. Ballistic Transport

In the case of ballistic transport, cascading the aforementioned three scattering matrices (according to the prescription of Ref. 2) yields the overall scattering matrix and the transmission T_{total} [1,4] as

$$T_{total} = \frac{\epsilon[(t_1 + t_2) - (b - a)^2 t_1 t_2 (t_1' + t_2')]}{[1 - t_1(a^2 t_1' + b^2 t_2')][1 - t_2(a^2 t_2' + b^2 t_1'] - a^2 b^2 t_1 t_2 (t_1' + t_2')^2} \qquad (2)$$

where ϵ, a and b are the elements of the Shapiro matrix[2], and t and r stand for transmission and reflection amplitudes within the two interfering paths. The subscripts '1' and '2' identify the corresponding path and the unprimed and primed quantities are associated with forward and reverse propagation of the electron.

[2] For a definition of these elements, see Ref. 1, 3 or 4.

In the presence of an external potential V inducing the electrostatic Aharonov-Bohm effect, t_1, t_2, $t_1{}'$ and $t_2{}'$ transform according to the following rule [4]:

$$\begin{pmatrix} t_1 \rightarrow \hat{t}_1 & t_1{}' \rightarrow \hat{t}'_1 \\ t_2 \rightarrow \hat{t}_1 e^{i\phi} & t_2{}' \rightarrow \hat{t}'_1 e^{i\phi} \end{pmatrix} , \tag{3}$$

where the quantities with the "hats" represent the transmission amplitudes in the absence of the external potential V, and ϕ is the electrostatic Aharonov-Bohm phase-shift between the two paths induced by V and given by

$$\phi = \frac{e}{\hbar} V < \tau_t >= \frac{\sqrt{2m^* E}}{\hbar} [\sqrt{1 + \frac{eV}{E}} - 1] L \tag{4}$$

Here $< \tau_t >$ is the harmonic mean of the transit times through the two paths which depends on V and the kinetic energy E of the electrons, m^* is the electron's effective mass and L is the length of each path.

Using the transformations given by Equation (3) in Equation (2) and assuming that in the absence of the external potential V the two paths are identical in all respects (i.e. $\hat{t}_1 = \hat{t}_2$ and $\hat{t}_1{}' = \hat{t}_2{}'$), we obtain

$$T_{total}(\phi) = \frac{\epsilon \hat{t}_1 (1 + e^{i\phi})(1 - (b-a)^2 \hat{t}_1 \hat{t}'_1 e^{i\phi})}{D(\hat{t}_1, a, b, \phi)} . \tag{5}$$

where the denominator D is a function of $\hat{t}_1$, a, b and ϕ.

We find from the above equation that $T_{total}(\phi)$ vanishes and hence the conductance (see Equation (1)) reaches a minimum whenever

$$\phi = (2n+1)\pi , \quad i.e. \, when \quad \frac{\sqrt{2m^* E}}{\hbar} [\sqrt{1 + \frac{eV}{E}} - 1] L = (2n+1)\pi \tag{6}$$

This gives the usual conductance minima (which we call the *primary* minima) associated with destructive interference of transmitted electrons.

However, we find from Equation (5) that $T_{total}(\phi)$ also vanishes whenever

$$(b-a)^2 \hat{t}_1 \hat{t}'_1 e^{i\phi} = 1 \tag{7}$$

From the unitarity of the Shapiro matrix (see Ref. 4) it can be shown that $b - a$ differs from unity only by a constant phase factor, i.e

$$b - a = e^{i\nu} \tag{8}$$

Now since in ballistic transport $\hat{t}_1 = \hat{t}'_1 = e^{ikL}$ (where k is the electron's wavevector in either path in the absence of the external potential V), Equation (7) really corresponds to the condition

$$2k_1L + \phi + 2\nu = \frac{\sqrt{2m^*E}}{\hbar}[\sqrt{1+\frac{eV}{E}} + 1]L + 2\nu = 2m\pi \qquad (9)$$

Whenever Equation (9) is satisfied, another set of conductance minima should appear in the oscillations since the numerator of $T_{total}(\phi)$ goes to zero and the conductance should fall to a minimum unless the denominator of $T_{total}(\phi)$ also happens to go to zero at the same time. It is easy to see that the denominator of $T_{total}(\phi)$ vanishes whenever ϕ is an even multiple of π. Hence, unless Equation (9) is satisfied only by those values of ϕ that are even multiples of π (which requires $2(k_1L + \nu)$ to be also an even multiple of π), the conductance of the structure should reach a minimum whenever ϕ satisfies Equation (9). This gives rise to an additional set of minima which we call the *secondary minima*. Actually, the secondary minima always occur unless $2(k_1L + \nu)$ is an even or an odd multiple of π. The latter case is not proved here for the sake of brevity, but is proved in Ref. 4.

B. Diffusive Transport

In the case of diffusive transport, $T_{total}(\phi)$ can again be found from the prescription of Ref. 2, except that now we have to evaluate it numerically. We have calculated the conductance G vs. the electrostatic potential V for both ballistic and diffusive transport. The results are displayed in Fig. 1. The secondary minima are not washed out by elastic scattering in the weak localization regime. However, they begin to wash out with the onset of strong localization and with increasing temperature. The effect of temperature has been discussed in Ref. 4. Note also the interesting feature exhibited by the secondary minima; they become more and more pronounced in the higher cycles of oscillations (increasing V) unlike the primary minima. This implies that in an experimental situation, even if the secondary minima cannot be observed in the first few cycles, they could show up in the later cycles.

III. DISCUSSION

Before concluding this paper, we briefly discuss the origin of the secondary minima. Equation (9), which predicts the existence of the secondary minima in the ballistic case, physically represents the condition that an electron reflected *around* the ring interferes constructively with itself at its point of entry into the ring. This minimizes the conductance by maximizing the reflection. Such a phenomenon can be viewed as some kind of "coherent backscattering", but it is not exactly similar to the magnetostatic Aronov-Al'tshuler-Spivak (AAS) effect which also involves backscattering, but specifically involves interference of two backscattered *time-reversed paths*. Conductance modulation due to the interference of time-reversed

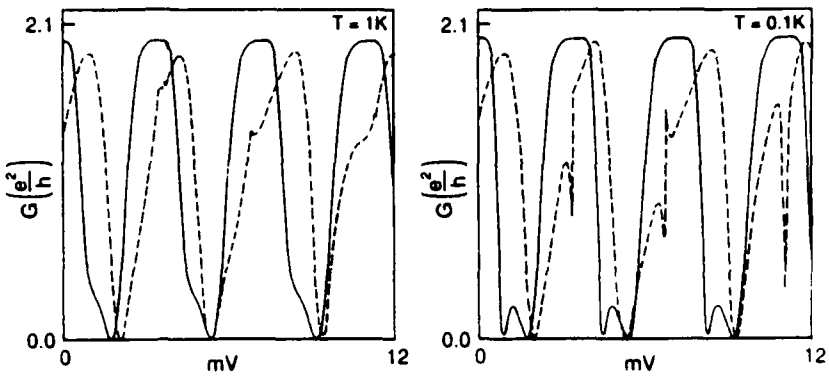


Fig. 1. Electrostatic A-B oscillations in a 1-d ring. The length of each path is 5000 Å. The carrier concentration is 1.55 x 10^6 cm^{-1} and the parameter ϵ = 0.5. The solid curve is for ballistic transport and the broken curve is for diffusive transport. In the latter case, there are 10 elastic scatterers in each path arbitrarily located. Strong localization would have set in if there were 33 scatterers in either path. In both ballistic and diffusive transport, the secondary minima are bleached out much more rapidly than the primary minima as the temperature is increased.

paths cannot occur in the electrostatic case since the time reversed paths always interfere constructively and an external electrostatic potential cannot change that[3]. However, in spite of this basic difference, there is undeniably the superficial similarity between the two effects in that they both double the frequency of the conductance troughs in the oscillations.

REFERENCES

1. Bandyopadhyay, S., and Porod, W. (1988a). Appl. Phys. Lett. 53, 2323. ; Bandyopadhyay, S., and Porod, W. (1988b). Superlattices and Microstructures, to appear.

2. Cahay, M., McLennan, M., and Datta, S. (1988). Phys. Rev. B15., 37, 10125.

3. Shapiro, B. (1983). Phys. Rev. Lett. 50, 747.

4. Cahay, M., Bandyopadhyay, S., and Grubin, H. L. (1989). to appear Phys. Rev. B15 (Rapid Communication).

[3]This happens because the phase-shifts suffered by an electron in traveling along opposite directions have the *same* sign in the electrostatic case, but *opposite* signs in the magnetostatic case.

Two types of conductance minima in electrostatic Aharonov-Bohm conductance oscillations

M. Cahay
Scientific Research Associates, Inc., Glastonbury, Connecticut 06033

S. Bandyopadhyay
Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556

H. L. Grubin
Scientific Research Associates, Inc., Glastonbury, Connecticut 06033
(Received 17 January 1989)

We predict the existence of *two different* types of conductance minima, arising from *different* interference conditions, in the conductance oscillation of a one-dimensional ring due to the electrostatic Aharonov-Bohm effect. The occurrence of *two* types of minima doubles the frequency of the conductance troughs in the oscillations, making it *twice* that predicted by the Aharonov-Bohm effect. This feature, which is not inhibited by elastic scattering in the weak-localization regime, can, however, be observed only at sufficiently low temperatures. At elevated temperatures, one of the two types of minima is bleached out and the normal Aharonov-Bohm frequency is restored.

In this Rapid Communication, we point out an intriguing feature in the conductance oscillation of a one-dimensional ring due to the electrostatic Aharonov-Bohm effect.[1] Unlike in the magnetostatic effect, the conductance oscillation of a ring due to the electrostatic effect exhibits *two distinct* sets of minima arising from two *different* interference conditions. One set of minima is caused by the usual destructive interference of transmitted electrons, and the other arises due to the constructive interference of an electron traveling completely around the ring and interfering with itself at its point of entry into the ring. In the next paragraphs we establish the existence of this feature and discuss various issues related to it.

For purposes of analysis, we represent a one-dimensional ring structure as shown in Fig. 1. We assume that phase randomization in the two contacts (termed "source" and "drain") occur sufficiently far away from the junctions between the contacts and the paths.[2] The (two-terminal) conductance of the structure, in the linear-response regime, is given by[3]

$$G=\frac{e^2}{2hkT}\int dE\,|T_{\text{total}}(E)|^2\operatorname{sech}^2\left(\frac{E-E_F}{2kT}\right), \quad (1)$$

where T_{total} is the transmission coefficient of an electron through the entire structure, E is the kinetic energy of the electron, and E_F is the Fermi level.

The conductance G depends on the transmission T_{total}. The transmission T_{total} can be found from the overall scattering matrix for the ring structure determined by cascading three scattering matrices.[4] They represent propagation from the source to the two paths, propagation along the paths and propagation from the paths to the drain, respectively. For simplicity, we represent the first and the last of these scattering matrices (for junctions A-B and C-D in Fig. 1) by the so-called Shapiro matrix[5] which relates the incident, reflected, and transmitted amplitudes according to (see Fig. 1)

$$\begin{pmatrix} A^- \\ B_1^+ \\ B_2^+ \end{pmatrix} = \begin{pmatrix} -(a+b) & (\sqrt{\epsilon})^* & (\sqrt{\epsilon})^* \\ \sqrt{\epsilon} & a & b \\ \sqrt{\epsilon} & b^* & a \end{pmatrix} \begin{pmatrix} A^+ \\ B_1^- \\ B_2^- \end{pmatrix}, \quad (2)$$

where the asterisk denotes complex conjugate.

The scattering matrix representing propagation along the two paths (i.e., across the junction B-C in Fig. 1) is given by[4]

$$\begin{pmatrix} B_1^- \\ B_2^- \\ C_1^+ \\ C_2^+ \end{pmatrix} = \begin{pmatrix} r_1 & 0 & t_1' & 0 \\ 0 & r_2 & 0 & t_2' \\ t_1 & 0 & r_1' & 0 \\ 0 & t_2 & 0 & r_2' \end{pmatrix} \begin{pmatrix} B_1^+ \\ B_2^+ \\ C_1^- \\ C_2^- \end{pmatrix}, \quad (3)$$

where t and r stand for the transmission and reflection coefficients within the paths. The subscripts 1 and 2 identify the corresponding path and the unprimed and primed quantities are associated with forward and reverse propagation of an electron from the source to the drain.

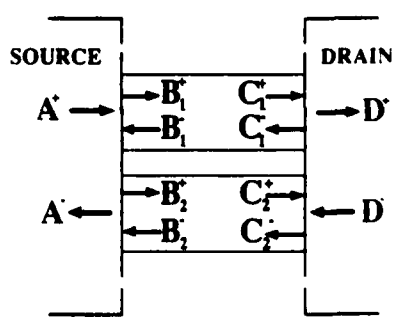


FIG. 1. Schematic representation of a one-dimensional ring-like structure showing the incident, reflected, and transmitted electron amplitudes.

 © 1989 The American Physical Society

If we assume ballistic transport in the two paths, in which case $r_1 = r_2 = r_1' = r_2' = 0$, then cascading the three scattering matrices for the three junctions A-B, B-C, and C-D yields the overall scattering matrix and thence the transmission T_{total} $(= D^+/A^+)$ as

$$T_{\text{total}} = \frac{\epsilon[(t_1+t_2)-(b-a)^2 t_1 t_2 (t_1'+t_2')]}{[1-t_1(a^2 t_1'+b^2 t_2')][1-t_2(a^2 t_2'+b^2 t_1')]-a^2 b^2 t_1 t_2 (t_1'+t_2')^2} . \tag{4}$$

The above equation is a perfectly general expression for the transmission through a *ballistic* one-dimensional ring. It may be pointed out that ballistic transport, although difficult to achieve, is not totally unexpected in strictly one-dimensional semiconductor microstructures at low enough temperatures since elastic scattering events are highly unlikely in one-dimensional structures.[3] The case of diffusive transport, when elastic scattering is present, is discussed later.

Let us now consider the conductance of the ballistic ring in a magnetic field. In the presence of a magnetic flux inducing the *magnetostatic* Aharonov-Bohm effect, t_1, t_2, t_1', and t_2' in Eq. (4) transform according to the following rule:[6]

$$\begin{pmatrix} t_1 \to \hat{t}_1 e^{-i\theta/2} & t_1' \to \hat{t}_1' e^{i\theta/2} \\ t_2 \to \hat{t}_2 e^{i\theta/2} & t_2' \to \hat{t}_2' e^{-i\theta/2} \end{pmatrix} , \tag{5}$$

where the quantities with carets represent transmission amplitudes in the absence of any magnetic flux and θ is the magnetostatic Aharonov-Bohm phase shift given by $\theta = (e/\hbar)\Phi$, Φ being the flux threaded by the ring. Using the transformations given by Eq. (5) in Eq. (4) and assuming that the two arms of the ring are identical in all respects so that (in the absence of any flux) $\hat{t}_1 = \hat{t}_2$ and $\hat{t}_1' = \hat{t}_2'$, we get

$$T_{\text{total}}(\theta) = \frac{\epsilon \hat{t}_1 e^{i\theta/2}(1+e^{-i\theta})[1-(b-a)^2 \hat{t}_1 \hat{t}_1']}{[1-\hat{t}_1\hat{t}_1'(a^2+b^2 e^{-i\theta})][1-\hat{t}_1\hat{t}_1'(a^2+b^2 e^{i\theta})]-a^2 b^2 \hat{t}_1^2 \hat{t}_1'^2 (e^{i\theta/2}+e^{-i\theta/2})^2} . \tag{6}$$

The numerator in $T_{\text{total}}(\theta)$ goes to zero and hence the conductance of the ring [see Eq. (1)] reaches a minimum[7] whenever

$$\theta = \frac{e}{\hbar}\Phi = (2n+1)\pi . \tag{7}$$

This gives the usual conductance minima in the magnetostatic Aharonov-Bohm oscillations associated with destructive interference of transmitted electrons. Note, however, that the numerator in $T_{\text{total}}(\theta)$ also becomes identically zero *(independent of the magnetic flux)* if the following condition is satisfied,

$$(b-a)^2 \hat{t}_1 \hat{t}_1' = 1 . \tag{8}$$

It can be shown from the required unitarity of the Shapiro matrix that the quantity $b-a$ differs from unity by a constant phase factor, i.e., $b-a = e^{i\nu}$. Also, in ballistic transport, $\hat{t}_1 = \hat{t}_1' = e^{ikL}$ (where L is the length of each path and k is the electron's wave vector in either path at zero magnetic flux). Therefore, Eq. (8) really corresponds to the condition,

$$2kL + 2\nu = 2n\pi . \tag{9}$$

It appears that if condition (9) is satisfied [in which case the numerator in $T_{\text{total}}(\theta)$ remains identically zero independent of θ], the conductance of the ring should always remain at its minimum, regardless of the magnetic flux. However, that is not quite true since the denominator in $T_{\text{total}}(\theta)$ could also become zero at some values of the magnetic flux. It is easy to see that the denominator does vanish whenever $\theta = 2n\pi$ or $\Phi = nh/e$ ($n = 0$ or an integer) in which case, application of L'Hospital's rule shows that $|T_{\text{total}}(\Phi = nh/e)| = 1$. It is interesting to note that if the ring's parameters (wave vector and length) are such that Eq. (9) is satisfied (which actually implies that the ring is "Fabry-Perot resonant" at zero magnetic flux), then $|T_{\text{total}}(\Phi)| = \delta_{\Phi, nh/e}$, where the δ is a Krönicker δ. In that case, at a temperature of 0 K, the magnetoconductance $G(\Phi)$ of the ring will appear as a series of "spikes" occurring at $\Phi = nh/e$; the spikes, however, will broaden with increasing temperature.

In the case of the *electrostatic* effect, the transformations in Eq. (5) are replaced by

$$\begin{pmatrix} t_1 \to \hat{t}_1 & t_1' \to \hat{t}_1' \\ t_2 \to \hat{t}_1 e^{i\phi} & t_2' \to \hat{t}_1' e^{i\phi} \end{pmatrix} , \tag{10}$$

where ϕ is the electrostatic Aharonov-Bohm phase shift between the two paths given by

$$\phi = \frac{e}{\hbar} V \langle \tau_t \rangle = \frac{\sqrt{2m^* E}}{\hbar} \left[\left(1 + \frac{eV}{E} \right)^{1/2} - 1 \right] L . \tag{11}$$

Here $\langle \tau_t \rangle$ is the harmonic mean of the transit times through the two paths which depends on the incident energy E of the electrons and also the potential difference V between the paths.[8]

The difference between the transformations in Eqs. (5) and (10) accrue from the fact that the magnetostatic Aharonov-Bohm phase shifts suffered by an electron in traveling along opposite directions (time-reversed paths) have *opposite* signs, whereas the electrostatic phase shifts will have the *same* sign. This is an important distinction which ultimately causes two different sets of minima to appear in the electrostatic effect but not in the magnetostatic effect. Is is also this difference that precludes the existence of an electrostatic analog of the magnetostatic

Aronov-Al'tshuler-Spivak effect.

Using the transformations given by Eq. (10) in Eq. (4), we obtain (for the electrostatic case),

$$T_{\text{total}}(\phi)=\frac{\epsilon\hat{t}_1(1+e^{i\phi})[1-(b-a)^2\hat{t}_1\hat{t}_1'e^{i\phi}]}{[1-\hat{t}_1\hat{t}_1'(a^2+b^2e^{i\phi})][1-\hat{t}_1\hat{t}_1'(a^2e^{2i\phi}+b^2e^{i\phi})]-a^2b^2\hat{t}_1^2\hat{t}_1'^2e^{i\phi}(1+e^{i\phi})^2}\,. \tag{12}$$

The numerator of $T_{\text{total}}(\phi)$ vanishes and the conductance reaches a minimum[8] whenever

$$\phi=(2n+1)\pi$$

or (13)

$$\frac{\sqrt{2m^*E}}{\hbar}\left[\left(1+\frac{eV}{E}\right)^{1/2}-1\right]L=(2n+1)\pi\,.$$

This gives the usual conductance minima (which we call the *primary* minima) associated with destructive interference of transmitted electrons. Note, however, that the numerator of $T_{\text{total}}(\phi)$ also vanishes if

$$(b-a)^2\hat{t}_1\hat{t}_1'e^{i\phi}=1\,. \tag{14}$$

In ballistic transport, this corresponds to the condition

$$2kL+\phi+2\nu=\frac{\sqrt{2m^*E}}{\hbar}\left[\left(1+\frac{eV}{E}\right)^{1/2}+1\right]L+2\nu=2m\pi\,. \tag{15}$$

It is obvious that whenever condition (15) is satisfied, the numerator of $T_{\text{total}}(\phi)$ goes to zero and the conductance should fall to a minimum unless the denominator of $T_{\text{total}}(\phi)$ also happens to go to zero at the same time. The denominator vanishes whenever $\phi=2n\pi$. Hence, unless Eq. (15) is satisfied only by those values of ϕ that are even multiples of π (which requires $2kL+\nu$ to be an even multiple of π or the ring to be Fabry-Pérot resonant at $V=0$), the conductance of the ring should reach a minimum whenever ϕ satisfies Eq. (15). This gives rise to an additional set of minima which we call the *secondary minima.* The physical origin of the secondary minima is the following: Eq. (15) represents the condition that *an electron, entering one of the paths from the left contact, gets reflected into the other path at the right contact, travels full circle around the ring and interferes constructively with itself at its point of entry at the left contact. This maximizes the reflection and hence minimizes the transmission and conductance.* This phenomenon could also cause a secondary set of minima to appear in the magnetostatic oscillations, but there the conditions for the occurrence of the primary and secondary minima are exactly identical (they occur at exactly the same value of the magnetic flux), so that they are always indistinguishable. But in the electrostatic case, the two conditions are different so that the two minima are *distinguishable.*

Let us now establish the requirements for the distinguishability. For this, we first find the difference between the phase shifts that give rise to the primary and secondary minima. From Eqs. (13) and (15),

$$\phi_{\text{primary}}-\phi_{\text{secondary}}=(2n+1)\pi-(2m\pi-2kL-2\nu)\,. \tag{16}$$

This difference becomes an even multiple of π (or, equivalently zero) if $2(kL+\nu)$ is an odd multiple of π, i.e., if the ring happens to be Fabry-Pérot antiresonant at $V=0$. In that case, the primary and secondary minima will overlap and remain indistinguishable. Barring this case, and the case of the ring being Fabry-Pérot resonant at $V=0$, both types of minima will not only occur in the oscillations, but also remain distinguishable.

It is clear that the appearance of *two different* sets of minima doubles the frequency of the conductance troughs in the oscillations. It is interesting to examine whether this can ever give rise to *exactly* half-periodic $(h/2e)$ oscillations. For this to happen, the secondary minima must occur when $\phi=2m\pi$ since the primary minima always occur when $\phi=(2n+1)\pi$. But the secondary minima cannot occur when $\phi=2m\pi$ since [from Eq. (15)] that would require $2(kL+\nu)$ to be an even multiple of π in which case the secondary minima do not even appear. Hence, exactly half-periodic oscillations can never arise from this effect in ballistic transport.

We now examine the effect of nonzero temperature on the two types of minima. Nonzero temperature gives rise to a thermal spread in the electron's energy which results in a bleaching out of the conductance minima due to ensemble averaging over the electron's energy. The primary minima are bleached out when the spread in the quantity on the left-hand side of Eq. (11) (due to a spread in the electron's energy) exceeds π and the secondary minima are bleached out when the spread in the quantity on the left-hand side of Eq. (15) exceeds π. These two spreads are

$$\Delta_{\text{primary}}=\left(\frac{m^*}{2\hbar^2}\right)^{1/2}L\left(\frac{1}{\sqrt{E+eV}}-\frac{1}{\sqrt{E}}\right)kT\,,$$

$$\Delta_{\text{secondary}}=\left(\frac{m^*}{2\hbar^2}\right)^{1/2}L\left(\frac{1}{\sqrt{E+eV}}+\frac{1}{\sqrt{E}}\right)kT\,, \tag{17}$$

where kT is the thermal spread in the energy.

From Eq. (17), we can find two critical temperatures T^c_{primary} and $T^c_{\text{secondary}}$ above which the primary minima and the secondary minima, respectively, are bleached out. These two temperatures are estimated by equating Δ_{primary} and $\Delta_{\text{secondary}}$ to π which gives (assuming the electron energy E to be the Fermi energy E_F)

$$kT^c_{\text{primary}}\sim\frac{\hbar}{L}\frac{1}{\sqrt{2m^*}}\left(\frac{1}{\sqrt{E_F}}-\frac{1}{\sqrt{E_F+eV_p}}\right)^{-1}, \tag{18}$$

and

$$kT^c_{\text{secondary}}\sim\frac{\hbar}{L}\frac{1}{\sqrt{2m^*}}\left(\frac{1}{\sqrt{E_F}}+\frac{1}{\sqrt{E_F+eV_s}}\right)^{-1}, \tag{19}$$

where V_p is the potential at which the first primary minimum occurs and V_s is the potential at which the first secondary minimum occurs in the oscillations.

Note that both T^c_{primary} and $T^c_{\text{secondary}}$ increase with increasing E_F or carrier concentration and decrease with increasing length of the structure. It is therefore necessary to have short structures with high-carrier concentration in order to observe both minima at sufficiently high temperatures.

In Fig. 2, we show the effect of temperature on both types of minima in the oscillations. While the primary minima can persist up to rather high temperatures, the secondary minima bleach out at much lower temperatures since [as we can see from Eqs. (18) and (19)] $T^c_{\text{secondary}} < T^c_{\text{primary}}$. This means that in an experimental situation, raising the temperature will gradually wash out the secondary minima and the oscillations will gradually revert to the normal Aharonov-Bohm oscillations with only the primary minima visible at higher temperatures.

Finally, another interesting feature, which is clearly visible in the oscillations, is that the primary minima tend to bleach out more and more in the higher cycles of the oscillations whereas the secondary minima exhibit the *opposite* behavior. This allows one to distinguish between the two types of minima in experimental data. It is a very interesting behavior and is easily understood from Eq. (17) which shows that at a given temperature, Δ_{primary} *increases* with increasing V while $\Delta_{\text{secondary}}$ actually *decreases* with increasing V. The significance of this is that at elevated temperatures, even if the secondary minima are not visible in the first few cycles of the oscillations, *they could eventually show up in the later cycles.*

Before concluding this Rapid Communication, we briefly discuss the effect of elastic scattering. We have carried out an analysis in the presence of elastic scattering following Ref. 4 and found that elastic scattering does not inhibit the twin-minima feature in the weak localization regime as long as the temperature is well below $T^c_{\text{secondary}}$.[9] We have also found that the feature is not completely inhibited in multichanneled transport as long as the number of channels (propagating modes) is not too large.[9] It therefore appears that the feature predicted in this Rapid Communication is quite robust and should be observable in realistic semiconductor structures at sufficiently low temperatures.

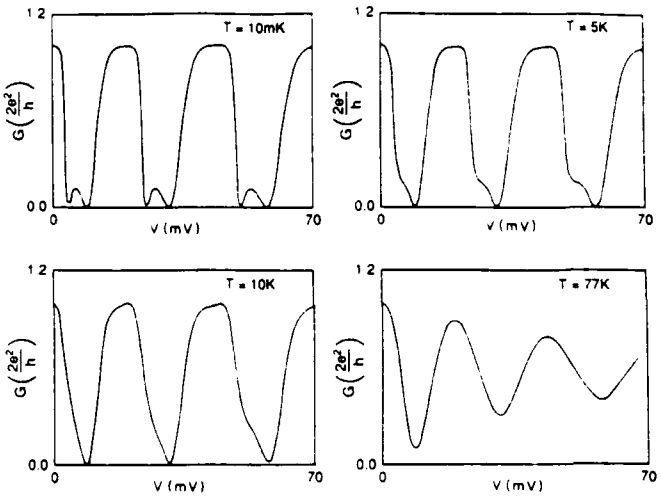


FIG. 2. The electrostatic Aharonov-Bohm conductance oscillations in a "ballistic ring" made of GaAs showing both types of minima. Each type of minima recur with the usual Aharonov-Bohm periodicity, but the separation between two adjacent minima (belonging to the two different types) is smaller than and unrelated to the Aharonov-Bohm periodicity. Note that the secondary minima are bleached out at much lower temperatures than the primary. These curves were obtained by performing the integral in Eq. (1) numerically. The parameters for the ring were carrier concentration equals 1.55×10^6 cm^{-1}, path length equals 1000 Å, $\epsilon=0.5$, and $\nu=0$.

In conclusion, we have established the existence of a hitherto unsuspected feature in the conductance oscillation of a one-dimensional ring due to the electrostatic Aharonov-Bohm effect. We have identified the origin of this feature and discussed the conditions for its observability.

The work at Scientific Research Associates was supported by the Air Force Office of Scientific Research under Contract No. F49620-87-C-005. The work at Notre Dame was supported by the same agency under Grant No. AFOSR 88-0096 and by IBM.

[1]For a discussion of the electrostatic Aharonov-Bohm effect which is a type-II nonlocal effect, see, for example, G. Matteucci and G. Pozzi, Phys. Rev. Lett. **54**, 2469 (1985).

[2]We thank S. Washburn for pointing out the importance of this assumption.

[3]S. Bandyopadhyay and W. Porod, Appl. Phys. Lett. **53**, 2323 (1988); Superlattices Microstruc. **5**, 239 (1989).

[4]P. W. Anderson, Phys. Rev. B **23**, 4828 (1981); B. Shapiro, *ibid.* **35**, 8256 (1987); M. Cahay, M. McLennan, and S. Datta, *ibid.* **37**, 10 125 (1988).

[5]B. Shapiro, Phys. Rev. Lett. **50**, 747 (1983).

[6]Y. Gefen, Y. Imry, and M. Ya Azbel, Phys. Rev. Lett. **52**, 139 (1984); Surf. Sci. **142**, 203 (1984).

[7]The minimum value of the conductance in the magnetostatic effect is exactly zero in ballistic transport even at nonzero temperatures. However, in the electrostatic case, the minimum value is zero only at 0 K.

[8]Note that since the electrostatic Aharonov-Bohm phase shift ϕ is not a linear function of the potential V, the conductance oscillations are not exactly periodic in V. Actually, the period increases continuously with increasing V.

[9]M. Cahay, S. Bandyopadhyay, and H. L. Grubin (unpublished); *Nanostructure Physics and Fabrication*, edited by M. A. Reed and W. P. Kirk (Academic, Boston, in press).

Fluctuations in the optical spectra of disordered microstructures due to quantum-interference effects

Supriyo Bandyopadhyay
Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556
(Received 11 February 1988)

We describe a novel quantum-interference phenomenon that gives rise to fluctuations in the optical spectra of disordered microstructures when the inelastic scattering time in the structures exceeds the radiative recombination lifetime. The origin of this phenomenon lies in the fact that an electron or hole, forming an optical dipole, does not lose its phase memory in the absence of *inelastic* scattering. Consequently, when the dominant relaxation process for the optical dipole moment is elastic impurity scattering, the optical spectra of disordered samples depend sensitively on the phase relationships between the various electron (hole) states in the system due to quantum interference. Since these phase relationships themselves depend on the exact locations of the impurities (scattering centers) within the structure, the optical spectra will also depend on the precise details of the impurity "configuration" inside the structure. In addition, if the phase relationships are altered with an external field which perturbs the states, the optical spectrum will exhibit sample-specific fluctuations. In many ways, this phenomenon is an optical analog of "universal conductance fluctuations" and indeed has the same physical origin. An important consequence of this phenomenon is that in a superlattice structure, each quantum well will have a slightly different optical spectrum if they merely have different impurity configurations but are otherwise identical. Consequently, this phenomenon will induce a unique type of *inhomogeneous broadening* in such a structure. This inhomogeneous broadening can be quite large and at low enough temperatures can even be the dominant cause of linewidth broadening.

I. INTRODUCTION

It is now well established that elastic scattering processes, such as impurity scattering, do not destroy the phase memory of electrons, unlike inelastic scattering.[1] Recent observations of the Aronov-Al'tshuler-Spivak effect,[2] the Aharonov-Bohm effect,[3] and universal conductance fluctuations[4] in disordered mesoscopic structures (with sample dimensions shorter than the inelastic diffusion length) bear ample testimony to this fact. A striking consequence of this phase-memory conservation in impurity scattering is that certain macroscopic properties of a disordered sample, such as the conductance, can depend sensitively on certain microscopic features, such as the exact locations of the impurities within the sample.[5] If the impurity configuration inside a sample is altered while keeping the total number of impurities the same, the conductance of the sample also changes. This is purely a quantum-mechanical effect and arises from the fact that the conductance depends on the superposition of the transmission amplitudes of various Feynman paths inside the sample. Since elastic scattering does not destroy the phase memory of electrons, the interference terms in the superposition do not ensemble average to zero. Consequently, changing the impurity configuration (which alters the phase relationships between the various Feynman paths) changes the conductance. A convenient way of demonstrating this phenomenon is to use an external magnetic field to introduce an additional Aharonov-Bohm phase shift between the various electron trajectories. If the field is scanned, the interference conditions between the trajectories change causing the conductance to fluctuate. These aperiodic fluctuations are "sample-specific" since the impurity configuration is sample specific, but the rms value of the fluctuations has a universal magnitude. This is the essence of universal conductance fluctuations.

In this paper, the primary objective is to study an optical analog of the above transport phenomenon. Elastic intraband relaxation processes do not destroy the phase memory of an electron or a hole forming an optical dipole. Consequently the decay of the electron and hole states in the presence of impurity scattering is influenced by quantum-interference effects. In particular, since the impurity configuration determines the phase relationships between the various states that the electron and hole are coupled to by the impurity interaction, the configuration plays an important role in determining the exact nature of the decay processes for the electronic and hole states and therefore the optical dipole. The precise details of how the optical dipole moment decays with time determines several features of the optical spectrum. Consequently, any influence of the impurity configuration on this decay process is manifested in the optical spectra. In other words, the optical spectrum is influenced by the impurity configuration.

In order to study this phenomenon, we have developed a simple quantum-mechanical model to calculate the

 ©1988 The American Physical Society

damping of an optical dipole moment with time in the presence of impurity scattering. The formalism is based on *coupled-mode theory* and is derived directly from the Schrödinger equation. We first compute the time evolution of an electronic state and a hole state that form an optical dipole by solving a set of coupled-mode equations. From this, we obtain the damping characteristics of the dipole moment. The impurity coordinates appear *explicitly* in the Hamiltonian for the system and affect the decay characteristics directly. Inelastic processes are excluded from the model based on the premise that at low enough temperatures, the inelastic scattering time (for both electrons and holes) can exceed the radiative recombination lifetime of a photogenerated electron-hole pair. Hence a typical pair may not experience a single inelastic collision during its lifetime. For instance, the inelastic scattering time for electrons in GaAs may exceed 1 nsec at temperatures lower than 100 mK (Ref. 6) whereas the radiative recombination lifetime for typical carrier concentrations in GaAs is ~ 1 nsec. From the decay characteristics of the optical dipole moment, the photoluminescence line shape is calculated and compared for varying impurity configurations.

In Sec. II, we describe the theoretical model. In Sec. III we show how the amplitudes of an electronic state and a light- or heavy-hole state in a disordered two-dimensional quantum box decay with time due to impurity scattering. We also show how the decay characteristics are influenced by the precise details of the impurity configuration. We then show the influence of the impurity configuration on the decay characteristics of the optical dipole moments themselves. In Sec. IV we show the photoluminescence spectra for both electron–light-hole transitions and electron–heavy-hole transitions and how each is affected by the details of the impurity configuration. Finally, in Sec. V, we discuss the analogy between the optical fluctuations and the universal conductance fluctuations, and present the conclusions.

II. THEORY

We start from the single-particle effective-mass Schrödinger equation for a two-dimensional quantum box,

$$i\hbar\frac{\partial\psi}{\partial t}=-\frac{\hbar^2}{2m^*}\nabla_\rho^2\psi+H'\psi\ , \tag{1}$$

where ∇_ρ^2 is the two-dimensional Laplacian and H' is the impurity interaction Hamiltonian which is time independent (elastic scattering).

The wave function ψ is expanded in a complete orthonormal set

$$\psi=\sum_p c_p(t)\phi_p\ , \tag{2}$$

where the ϕ's are the so-called "normal modes" of the system which are the solutions of the Schrödinger equation in the absence of impurity interaction. These are therefore the "particle-in-a-box" states given by

$$\phi_p=\frac{2}{W}\sin\left[\frac{m\pi x}{W}\right]\sin\left[\frac{n\pi y}{W}\right]=|\,m,n\,\rangle\ , \tag{3}$$

where W is the lateral dimension of the quantizing (square) box. Each basis state is labeled by two indices m and n corresponding to the two-dimensional subbands in the two-dimensional quantum box.

Substitution of Eq. (2) into (1) yields the matrix equation which gives the time evolution of the amplitude $c_p(t)$ of any state ϕ_p,

$$i\hbar\frac{\partial[\mathbf{c}(t)]}{\partial t}=\mathcal{H}_0[\mathbf{c}(t)]+\mathcal{H}'[\mathbf{c}(t)]\ , \tag{4}$$

where $[\mathbf{c}(t)]$ is a column vector whose elements are the amplitudes of the various states in the quantum box at time t, $\mathcal{H}_0$ is the unperturbed Hamiltonian (a diagonal matrix whose elements are the kinetic energies of the basis states), and $\mathcal{H}'$ is the impurity interaction Hamiltonian whose matrix elements are given by

$$H'_{pq}=-\frac{\Gamma q^2}{4\pi\epsilon}\sum_i\langle\phi_p\,|\,\delta(\mathbf{r}-\mathbf{r}_i)\,|\,\phi_q\rangle\ . \tag{5}$$

Γ is the strength of the interaction, and $\mathbf{r}_i$ is the two-dimensional coordinate of the ith impurity. The summation is carried out over the coordinates of all the impurities in the quantum box.

For the scattering potential, we use a "delta potential" rather than the screened Coulomb potential so as to be able to obtain an analytical expression for the matrix elements of $\mathcal{H}'$. The choice of δ scatterers makes impurity scattering isotropic, but does not change the essential physics. The interference between the scattered amplitudes enters through the summation over the impurity coordinates. This underscores the importance of the impurity "configuration," since the exact locations of the impurities, appearing explicitly in the Hamiltonian, determine the phase relationships between the various ϕ's. The present formalism, which deals directly with the scattered amplitudes rather than with the scattering probabilities, is different from the semiclassical formalism based on "Fermi's golden rule," which deals only with the probability of scattering so that all interference effects between the scattered amplitudes are inevitably masked. Consequently, in the semiclassical formalism, it is only the net impurity concentration or the total number of impurities inside a sample that is important in determining any physical parameter; the precise details of the configuration are irrelevant. However, in a quantum-mechanical treatment, one must take the impurity coordinates explicitly into account since they affect the interference between the scattered amplitudes which in turn can affect certain macroscopic properties of a sample.

Equation (4) has the solution

$$[\mathbf{c}(t)]=\exp\left[-\frac{i\mathcal{H}t}{\hbar}\right][\mathbf{c}(0)]\ , \tag{6}$$

where $\mathcal{H}=\mathcal{H}_0+\mathcal{H}'$. It may be noted that since $\mathcal{H}$ is Hermitian, $\exp(-i\mathcal{H}t/\hbar)$ is always unitary, as it must be, in

order to conserve probability. Equation (4) is a set of "coupled-mode equations" which are inherently reversible. Since no inelastic processes are operative, there is no dissipation and consequently the system described by Eq. (4) is reversible. As a result of this reversibility, the complex amplitude of any state (both the magnitude and phase) is completely deterministic at any time and can be found from Eq. (6). Specifically, since the phase is deterministic at all times, phase memory is never lost. Moreover, since $\mathcal{H}'$ depends explicitly on the impurity coordinates $\mathbf{r}_i$ [see Eq. (5)], the impurity configuration will have a direct effect on both the magnitude and phase of any state at any given time.

Equations (1)–(6) are valid for both electrons and holes. The amplitude of any conduction-band state or valence-band state in the presence of impurity scattering can be found from Eq. (6). The time-dependent optical dipole moment $\mathbf{P}(t)$ associated with transitions between two quantum states in the valence and conduction bands of a semiconductor is given by

$$\mathbf{P}(t)=\tfrac{1}{2}\mathbf{P}(0)\sum_{p} c_{e_p}^{*}(t)c_{h_p}(t)+\text{c.c.}\ , \tag{7}$$

where c_e and c_h are the complex amplitudes of an electron and a hole state. Therefore, the damping of the optical dipole moment $\mathbf{P}(t)$ can also be calculated from Eq. (6). The optical line shape $\mathcal{F}(\omega)$, representing the *homogeneous broadening,* is obtained in a straightforward manner by Fourier transforming the time-dependent dipole moment $\mathbf{P}(t)$ into the frequency domain ω of the incident photons.[7] $\mathcal{F}(\omega)$ is thus also affected by the impurity configuration, insofar as the damping of $\mathbf{P}(t)$ is influenced by the configuration.

III. DECAY OF THE OPTICAL DIPOLE MOMENT WITH TIME

In this section we first show how an electronic state and a hole state decay with time in the presence of elastic impurity scattering. We choose a two-dimensional GaAs quantum box of area 1000×1000 Å^2 which can be fabricated by electron-beam lithography.[8] The impurity concentration is chosen to be 5×10^{11} cm^{-2} so that there are 50 impurities in the box. The impurity configuration is generated by two independent uniform random-number generators that determine the x and y coordinates. A typical impurity configuration is shown in Fig. 1. The various parameters are chosen as $\Gamma=333$ Å (for both electrons and holes), $m_e^*=0.067m_0$, $m_{\rm HH}^*=0.45m_0$ (heavy holes), $m_{\rm LH}^*=0.082m_0$ (light holes), and $\epsilon=12.9\epsilon_0$.

The initial photoexcited state for the electron is taken to be the state $|m,n\rangle\equiv|3,3\rangle$. This state has an energy 10 meV above the conduction-band edge so that both spontaneous and stimulated polar-optical-phonon emission from this state are inhibited. Other inelastic mechanisms (such as carrier-carrier scattering, optical-phonon absorption, and acoustic-phonon interactions) can be suppressed by lowering the temperature sufficiently. In a quantum dot, the phonon interactions are weak since quantum confinement gives rise to only discrete states

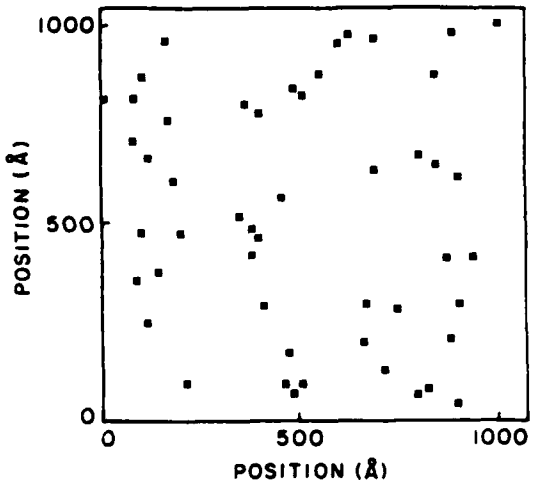


FIG. 1. A 1000×1000-Å^2 quantum box showing a typical random configuration of impurities.

and it is very difficult to simultaneously conserve both momentum and energy in even multiple phonon processes. In addition, radiative transitions between the subband states within the conduction band or between the subband states within the valence band are also forbidden because of the orthogonality of the envelope wave functions of the subband states. In other words, all intraband inelastic processes that could cause the initial electron or hole state to decay are weak and negligible compared to elastic interactions at low enough temperatures.

The initial heavy-hole state is also chosen to be the state $|3,3\rangle$ since the only radiative transitions between conduction-band states and valence-band states that are allowed are those between subbands with identical indices. The initial heavy-hole state thus has an energy 1.5 meV below the valence-band edge. Similarly, the light-hole state is also chosen to be the state $|3,3\rangle$ which has an energy 8.2 meV below the valence-band edge. It may be noted that for a given resonant photon energy, either only a heavy-hole transition or a light-hole transition is involved (but not both) since quantum confinement lifts the degeneracy between the heavy- and light-hole bands.

In an actual experimental situation, however, the choice of the initial electron and hole states will be governed by the energy of the incident photons. Since there are only discrete levels (and no continuum states) in a true "quantum dot" with infinite barriers, photoexcitation or absorption can occur only at discrete frequencies corresponding to discrete energy separations between the electron and hole subbands. In fact, by tuning the excitation frequency, one can select either a specific heavy-hole transition or a specific light-hole transition. Assuming the band gap of GaAs to be 1.42 eV, the present choice of initial states for the heavy-hole transition corresponds to a photon energy of about 1.4315 eV, and for the light-hole transition it corresponds to a photon energy of 1.4382 eV. We have neglected any strain-induced effects.

A very intriguing question at this point is whether the initial state $|3,3\rangle$, which is *not* degenerate in energy with any other state in the system, can decay at all in the ab-

sence of inelastic scattering. The only way this state can decay is if the electron gets coupled to other electronic states in the quantum box (not at the same energy) through *elastic* coupling. Alternately, one may ask if there can be any homogeneous broadening of the optical spectrum in the absence of inelastic scattering. Any homogeneous broadening or nonzero electronic linewidth indicates that the electron has either gained or lost energy after it was excited to the conduction band, which it apparently cannot do without inelastic scattering events. The same is true of holes. Indeed in the semiclassical picture, elastic scattering cannot couple an electron between states that are not isoenergetic so that an electron cannot gain or lose energy through elastic scattering. In the Fermi's-golden-rule formalism, this is ensured by the energy-conserving δ function which, however, is inexact for short time scales because of the uncertainty principle. In a strictly quantum-mechanical formalism, coupling between nondegenerate states is allowed for sufficiently strong elastic interaction, although it is always much weaker than coupling between degenerate states. Indeed in the coupled-mode formalism that we have adopted, this is a well-known fact in the context of integrated optics.[9] For instance, it is possible to couple light from one optical waveguide to another even if the two waveguides have slightly different characteristic frequencies. By the same token, an electron can be coupled to a higher or lower energy state for short times by sufficiently strong elastic coupling, thereby causing homogeneous broadening. The efficiency of this coupling, however, decreases rapidly with increasing energy separation between the states.

To illustrate this particular point, we have purposely chosen the initial state as the state $|3,3\rangle$. We have included 64 states in the calculation ($m=1,2,3,\ldots,8$ and $n=1,2,3,\ldots,8$) for both electrons and holes. In Fig. 2 we show how the initial electron and hole states decay with time. The two curves (solid and dashed lines) are for two different impurity configurations. These curves were obtained directly from Eq. (6). The initial states decay as the electron or hole is coupled away to other electron or hole states in the system. To illustrate this further, we show in Figs. 3(a) and 3(b) the real and imaginary parts of the electron "density matrix" at time $t=275$ fs. The electron density matrix in reciprocal space (momentum representation) is defined as

$$\rho_{pq}(t)=c_p^*(t)c_q(t)\ , \tag{8}$$

where c_p and c_q are the amplitudes of any two electronic states.

In the labeling scheme that we have adopted, the index p used to label a state is chosen such that the state $|1,1\rangle$ ($m=1$, $n=1$) has the index $p=1$, the state $|1,2\rangle$ has the index $p=2,\ldots$, the state $|2,1\rangle$ has the index $p=9,\ldots$. The initial state $|3,3\rangle$ therefore corresponds to $p=19$. In Fig. 3(a), the dominant peak appears at $p=19$, $q=19$ (i.e., $\rho_{19,19}$ is the largest element in the density matrix) which shows that the initial state is still the dominant state. Nevertheless states that are close to the initial state in energy have developed quite large amplitudes which indicates that the electron has been significantly coupled to

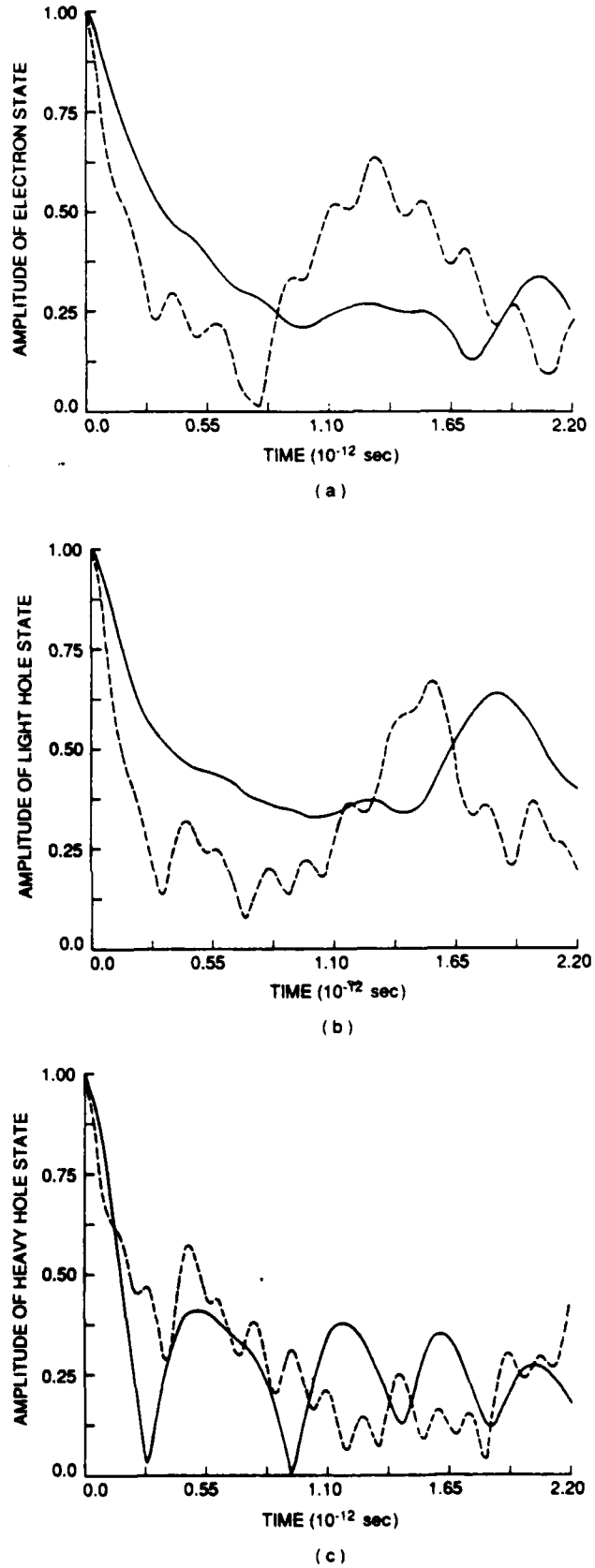


FIG. 2. Decay of the amplitudes of the initial (a) electron state, (b) the light-hole state, and (c) the heavy-hole state with time due to impurity scattering. The impurity concentration is 5×10^{11} cm^{-2}. The solid and dashed curves are for two different impurity configurations.

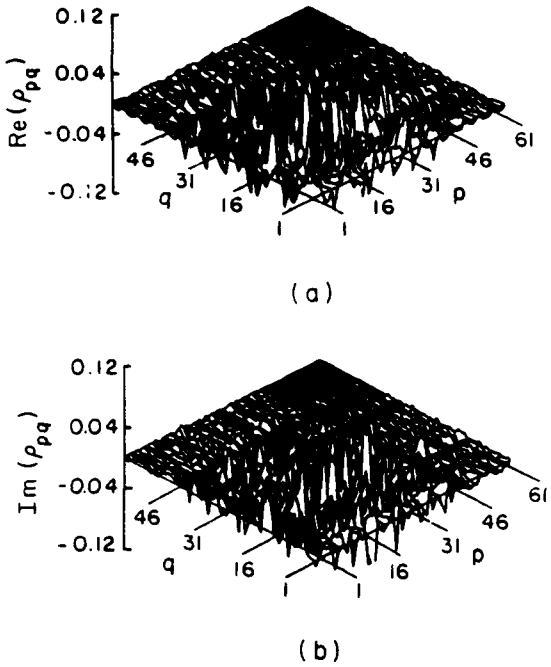


FIG. 3. The (a) real and (b) imaginary part of the electron density matrix ρ_{pq} in momentum space at time $t=275$ fs.

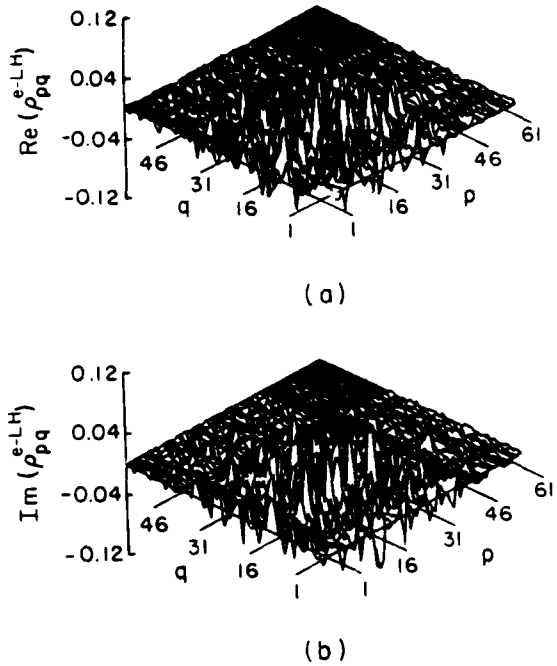


FIG. 4. The (a) real and (b) imaginary part of the electron–light-hole joint density matrix $\rho_{pq}^{e\text{-}LH}$ at time $t=275$ fs.

these states. This demonstrates the fact that in the presence of sufficiently strong elastic interaction, significant coupling may occur even between states that are nondegenerate in energy. However, states that are farther from the initial state in energy have correspondingly smaller amplitudes since coupling to these states is much weaker. For instance, the state $|8,8\rangle$ ($p=64$, $q=64$) is the state with the highest energy which has an energy 61 meV above the energy of the initial state. Consequently the element $\rho_{64,64}$ is almost zero since very little coupling can take place between this state and the initial state. (This also shows that 64 was a sufficient number of states to include in the calculation since states lying above the 64th state in energy are barely coupled to the initial state for this interaction strength.)

In Figs. 4 and 5 we also show the real and imaginary parts of the electron-hole *joint* density matrices for both light and heavy holes. The joint density matrix is defined as

$$\rho_{pq}^{e\text{-}h}(t)=c_{h_p}^{*}(t)c_{e_q}(t)\ , \tag{9}$$

where c_h is a light- or heavy-hole state and c_e is an electronic state.

There is an obvious difference between Fig. 4(a) and Fig. 5(a). Comparing the two, it is found that a much larger fraction of the elements are significantly large for the electron–heavy-hole joint density matrix than for the electron–light-hole joint density matrix. This is due to the fact that since the heavy holes have a larger effective mass, the heavy-hole subbands are spaced much closer in energy than the light-hole subbands so that elastic coupling between the heavy-hole states is much more efficient. Consequently, a much larger fraction of the heavy-hole states will acquire significant amplitudes after a given time. This also implies that a heavy-hole state will damp much more rapidly than a light-hole state in the presence of elastic scattering which is clearly seen in Figs. 2(b) and 2(c). Hence, an optical transition involving a heavy hole will have a shorter associated T_2 time in this case than a transition involving a light hole.

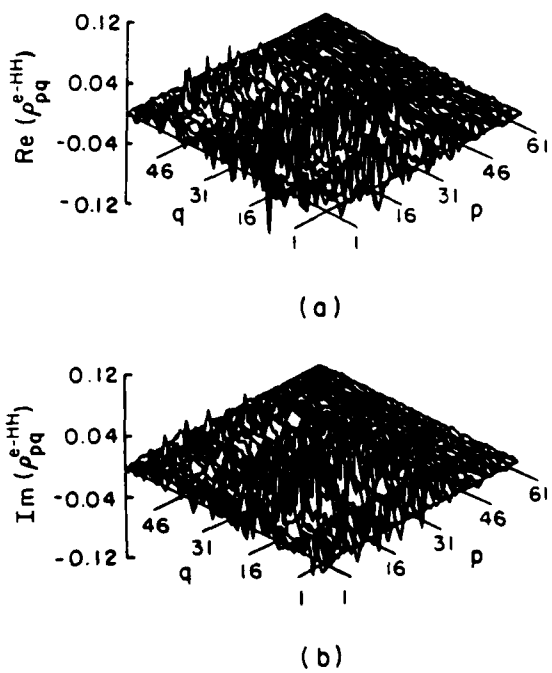


FIG. 5. The (a) real and (b) imaginary part of the electron–heavy-hole joint density matrix $\rho_{pq}^{e\text{-}HH}$ at time $t=275$ fs.

The decay characteristics in Figs. 2 have several interesting features. Each characteristic is *nonmonotonic*. This is related to the fact that in a dissipationless and reversible system, the initial state is never quenched permanently and should recur after sufficiently long time intervals. More importantly, the decay characteristic is distinctly different for two different impurity configurations (the solid and dashed curves). This is fully expected in view of the fact that the characteristics in Fig. 2 can be interpreted as interference patterns. They are influenced by the interference of the waves scattered from various impurities and also the interference of the waves reflected from the walls of the quantum box. Changing the impurity configuration will alter the phase relationships between these interfering waves and therefore the interference pattern. Consequently, any decay characteristic has a direct dependence on the impurity coordinates.

The monotonic parts of the decay characteristics merit close scrutiny. It is found that over a short interval of time ($t \leq 30$ fsec) the amplitude decays with time as $1-\alpha t^2$. This parabolic dependence has also been observed in Ref. 7. Over longer time intervals, the decay is almost linear with time. These features play an important role in determining the line shape of the photoluminescence spectra.

Finally, in Figs. 6(a) and 6(b) we show the decay characteristics of the optical dipole moments associated with an electron–light-hole transition and an electron–heavy-hole transition. As expected, the dipole moment involving a heavy-hole transition decays more rapidly. The so-called T_2 times (defined as the time that elapses before the dipole moment decays to 37% of its initial value) are approximately 150 and 250 fs for the heavy- and light-hole transitions, respectively, if one considers the solid curves. This implies that the homogeneous broadening of the optical spectrum will be larger for heavy-hole transitions than for light-hole transitions. However, it is interesting to note that the T_2 times are extremely sensitive to the impurity configuration. For the light-hole transition, the difference between the T_2 times for the two different impurity configurations is about 135 fs or the difference between the corresponding linewidths is about 5 meV which is a significant fraction ($\sim 50\%$) of the individual linewidths (full width at half maximum) themselves.

IV. FLUCTUATIONS IN THE LINEWIDTH, LINE SHAPE, AND "RED SHIFT" OF OPTICAL SPECTRA

In Figs. 7(a) and 7(b) we show the photoluminescence spectra obtained by Fourier transforming the decay of the dipole moment $\mathbf{P}(t)$ into the frequency domain ω of the incident photons for light- and heavy-hole transitions. Again, the two curves (solid and dashed) are for two different impurity configurations. There are three salient features associated with the line shapes that merit discussion. First, the line shape is *asymmetric*. It is neither Gaussian, nor Lorentzian as predicted by semiclassical theory. Second, the spectrum peaks at a frequency different from the frequency of the incident photons. In other words, there is a clear red shift. Third, the linewidth, line shape, and the red shift are distinctly different for the two different impurity configurations. In the following paragraphs, we discuss these features.

For a simple two-level system, semiclassical theory (which invokes only Markovian relaxation processes) predicts that the optical dipole moment decays with time exponentially,

$$\mathbf{P}(t)=\tfrac{1}{2}\mathbf{P}(0)\exp(i\omega_0 t)\exp\left[-\frac{t}{\langle T_2\rangle}\right]+\text{c.c.}\ , \qquad (10)$$

where $\langle T_2\rangle$ is an average "relaxation time" and ω_0 is the resonant photon frequency. Semiclassical theory therefore predicts a Lorentzian line shape with no red shift.

In our case, $\mathbf{P}(t)$ can be expressed as [see Eq. (7)]

$$\mathbf{P}(t)=\tfrac{1}{2}\mathbf{P}(0)\exp(i\omega_0 t)\sum_p b_{e_p}^*(t)b_{h_p}(t)+\text{c.c.}\ , \qquad (11)$$

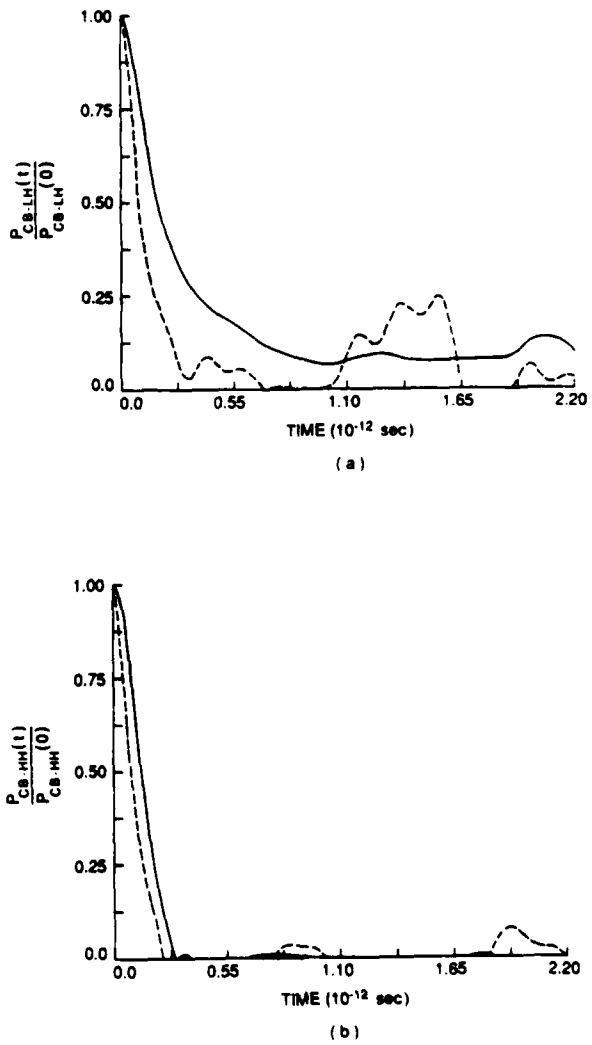


FIG. 6. Time evolution of (a) the electron–light-hole dipole moment $P_{CB\text{-}LH}(t)$, and (b) the electron–heavy-hole dipole moment $P_{CB\text{-}HH}(t)$ in the presence of elastic impurity scattering. The solid and dashed curves are for two different impurity configurations. In both cases only the envelope is plotted.

where

$$b_e(t)=\exp\left[i\frac{E_e t}{\hbar}\right]c_e(t) \tag{12}$$

and

$$b_h(t)=\exp\left[i\frac{E_h t}{\hbar}\right]c_h(t)\ . \tag{13}$$

E_e and E_h are the subband energies for the initial electron and hole states and

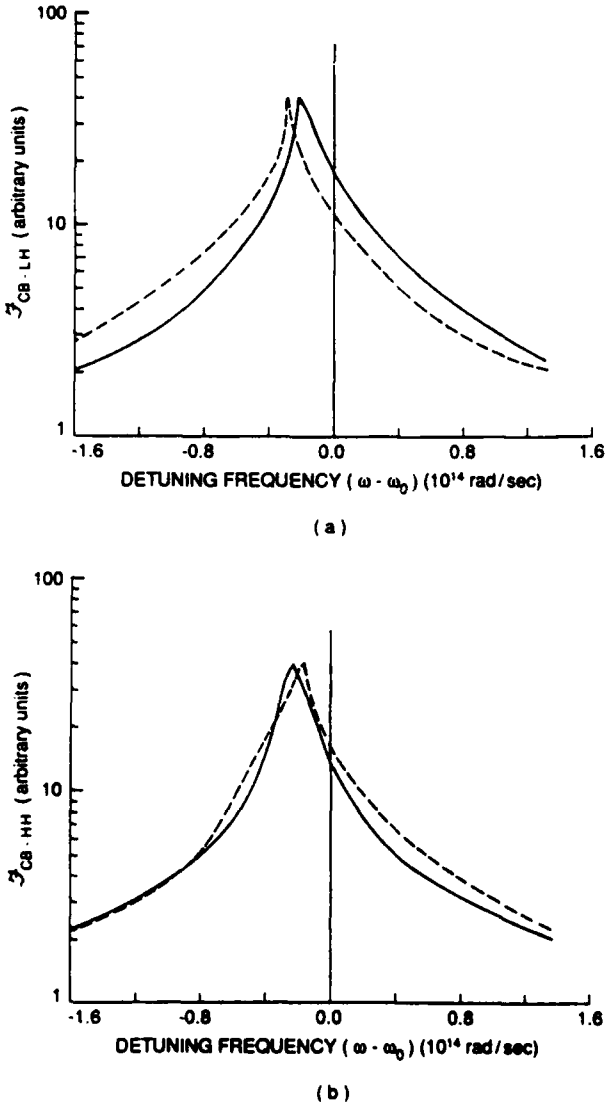


FIG. 7. (a) The photoluminescence spectrum for a photon energy corresponding to 1.4315 eV in a 1000×1000 Å^2 GaAs quantum box. This corresponds to a conduction-band–light-hole band transition. The solid and dashed curves are for two different impurity configurations. (b) The photoluminescence spectrum for a photon energy corresponding to 1.4382 eV. This corresponds to a conduction-band–heavy hole band transition. The solid and dashed curves are for two different impurity configurations.

$$E_e-E_h=\hbar\omega_0\ . \tag{14}$$

In other words, ω_0 is the resonant angular frequency of the incident photons corresponding to the effective "band gap" between the chosen electron and hole subbands.

Equation (11) can now be recast as

$$\begin{aligned}\mathbf{P}(t)&=\tfrac{1}{2}\mathbf{P}(0)\exp(i\omega_0 t)A(t)e^{-i\theta(t)}+\text{c.c.}\\&=\mathbf{P}(0)A(t)\cos[\omega_0 t-\theta(t)]\ ,\end{aligned} \tag{15}$$

where $A(t)$ is the magnitude and $\theta(t)$ the phase of the complex product $\sum b_{e_p}^*(t)b_{h_p}(t)$ representing the decay.

Comparing Eqs. (10) and (15), one finds two essential differences. First, the decay $A(t)$ is not exponential. In fact, over short time intervals it is parabolic and over longer time intervals it is linear. Consequently, the optical line shape is neither Lorentzian nor Gaussian. Secondly, there is a phase shift $\theta(t)$ associated with the decay which causes the line shape to be asymmetric about the center frequency while at the same time giving rise to a red shift. Physically, the red shift is associated with the real part of the self-energy correction (for both electrons and holes) due to the impurity interaction which renormalizes the effective band gap. Similar features were observed in Ref. 7 which included non-Markovian relaxation processes in calculating the optical line shape.

Figures 7(a) and 7(b) also show an appreciable difference between the two line shapes that correspond to two different impurity configurations. Both the red shift and the linewidth vary significantly with the impurity configuration. For this example, the rms values of the fluctuations in the red shift and linewidth (as the impurity configuration is randomly charged) are a few meV. Such fluctuations will give rise to a unique type of *inhomogeneous broadening* of the line shape for a lateral surface superlattice that consists of multiple two-dimensional quantum boxes with varying impurity configuration. From this example it appears that this homogeneous broadening can be an appreciable fraction of the total (inhomogeneous plus homogeneous) broadening. Moreover, in this case, the inhomogeneous broadening is larger than the energy spacing between the lower subbands for the 1000×1000 Å^2 quantum dot. Consequently, such inhomogeneous broadening can sometimes make it impossible to resolve the discrete optical spectra characteristic of quasi zero-dimensional structures.

V. CONCLUSION

In this paper we have discussed an optical analog of the "universal conductance fluctuations." We have shown that the line shape, linewidth, and the red shift in the photoluminescence spectra of a disordered sample depend sensitively on the impurity configuration inside the sample and fluctuate if the configuration is randomly altered. The practical importance of this phenomenon is that it can cause significant inhomogeneous broadening of the line shape for an otherwise ideal superlattice or multiple-quantum-well structure.

Physically, both the conductance fluctuations and their optical analog, the line-shape fluctuations, have the same origin. They both require coherence of the electronic (and hole) wave function which is preserved in the presence of *elastic* impurity scattering. As a result, there can be pronounced interference effects between waves scattered from various impurities within a sample which have a marked influence on macroscopic sample properties. In the case of conductance fluctuations, this can be demonstrated indirectly by using an external magnetic field. The field will introduce an additional Aharonov-Bohm phase shift between the various Feynman trajectories inside the sample so that changing the field will change the phase relationships between the trajectories and give rise to aperiodic fluctuations in the magnetoconductance. The nature of these fluctuations are sample specific since they depend on the precise details of the impurity configuration within the sample. By the same token, it is possible in principle to use a magnetic field to induce the optical fluctuations. The magnetic field will perturb the various modes in the system thereby changing the interference conditions between them. The exact nature of the change would depend on the specific impurity configuration so that one also expects to find sample-specific fluctuations in the optical spectra as the field is scanned. These fluctuations are optical *magnetofingerprints* since they are signatures of the exact impurity configuration within the sample. However, the field must be low enough so that level splitting and other magnetic-field-induced effects are negligible.

The primary requirement to observe these quantum-interference-induced fluctuations is to inhibit phase-breaking inelastic scattering events. In the case of conductance fluctuations, this is ensured by having the sample dimension smaller than the inelastic diffusion length so that an electron typically does not experience an inelastic interaction during its journey from one contact to the other. In the case of optical fluctuations, the requirement is to have the inelastic scattering time (for both electrons and holes) exceed the radiative recombination lifetime, so that an electron-hole dipole does not experience any inelastic scattering event during its lifetime.

There is another important consideration that distinguishes the two fluctuation phenomena. For typical semiconductor microstructures with dimensions approaching 1 μm, the inelastic diffusion length may exceed the sample dimensions at temperatures of a few kelvins so that the conductance fluctuations can become observable at temperatures of a few kelvins. On the other hand, the inelastic scattering time may exceed the radiative recombination lifetime only at temperatures of a few tens of millikelvins. Hence the optical fluctuations will usually be more difficult to observe than the conductance fluctuations. However, this problem is somewhat less serious in quantum-dot structures where spatial confinement of the electron and hole can significantly shorten the radiative recombination lifetime. A radiative lifetime of 20 ps has been calculated for CdS quantum dots with dimensions of 100 Å.[10] Hence the optical fluctuations may become observable in quasi zero-dimensional structures at temperatures much higher than a few tens of millikelvins.

Finally, in drawing an analogy between the conductance and optical fluctuations, an important point is whether there is any "universality" associated with the optical fluctuations. The conductance fluctuations are universal in the sense that the rms value of the fluctuations is $\approx e^2/h$ which depends only on universal constants. In the case of the optical fluctuations, no such universality is evident as yet, but more theoretical as well as experimental work is necessary to answer this question satisfactorily.

ACKNOWLEDGMENTS

The author is indebted to M. Cahay, W. Porod, C. S. Lent, S. Datta, M. Yamanishi, and Y. Lee for many insightful discussions. This work was supported by the U.S. Air Force Office of Scientific Research under Grant No. AFOSR-88-0096.

[1]For a review, see M. Büttiker, *SQUID 85: Superconducting Quantum Interference Devices and Their Applications*, edited by D. A. Halbohm and H. Lübbig (de Gruyter, New York, 1985), p. 529; S. Washburn and R. A. Webb, Adv. Phys. **35**, 375 (1986).

[2]D. Yu. Sharvin and Yu. V. Sharvin, Pis'ma Zh. Eksp. Teor. Fiz. **34**, 285 (1981) [JETP Lett. **34**, 272 (1981)]; M. Gijs, C. Van Haesendonck, and Y. Bruynseraede, Phys. Rev. Lett. **52**, 2069 (1984); B. Pannetier, J. Chaussy, R. Rammal, and P. Gandit, *ibid.* **53**, 718 (1984); D. J. Bishop, J. C. Licini, and G. J. Dolan, Appl. Phys. Lett. **46**, 1000 (1985).

[3]R. A. Webb, S. Washburn, C. P. Umbach, and R. B. Laibowitz, Phys. Rev. Lett. **54**, 2696 (1985); V. Chandrasekhar, M. J. Rooks, S. Wind, and D. E. Prober, *ibid.* **55**, 1610 (1985); S. Datta, M. R. Melloch, S. Bandyopadhyay, R. Noren, M. Vaziri, M. Miller, and R. Reifenberger, *ibid.* **55**, 2344 (1985); G. Timp, A. M. Chang, J. E. Cunningham, T. Y. Chang, P. Mankiewich, R. Behringer, and R. E. Howard, *ibid.* **58**, 2814 (1987).

[4]J. C. Licini, D. J. Bishop, M. A. Kastner, and J. Melngailis, Phys. Rev. Lett. **55**, 2987 (1985); W. J. Skocpol, P. M. Mankiewich, R. E. Howard, L. D. Jackel, D. M. Tennant, and A. D. Stone, *ibid.* **56**, 2865 (1986).

[5]B. L. Al'tshuler, Pis'ma Zh. Eksp. Teor. Fiz. **41**, 530 (1985) [JETP Lett. **41**, 648 (1985)]; B. L. Al'tshuler and B. Z. Spivak, *ibid.* **42**, 363 (1985) [**42**, 447 (1985)]; A. D. Stone, Phys. Rev. Lett. **54**, 2692 (1985); P. A. Lee and A. D. Stone, *ibid.* **55**, 1622 (1985).

[6]B. J. F. Lin, M. A. Paalanen, A. C. Gossard, and D. C. Tsui, Phys. Rev. B **29**, 927 (1984).

[7]M. Yamanishi and Y. Lee, IEEE J. Quantum Electron. **QE-23**, 367 (1987).

[8]G. Bernstein, Ph.D. thesis, Arizona State University, 1987; 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).

[9]A. Yariv, IEEE J. Quantum Electron **QE-9**, 919 (1973).

[10]T. Takagahara, Phys. Rev. B **36**, 9293 (1987).

Quantum interference effects in transient electronic transport

D. R. Poole[a)] and S. Bandyopadhyay
Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556

(Received 8 June 1989; accepted for publication 18 August 1989)

A simple quantum-mechanical model is presented for simulating transient electronic transport in disordered mesoscopic structures in the absence of phase-randomizing inelastic collisions. We have applied this model to study the time-dependent decay of an electron's momentum in ultrasmall GaAs structures with various impurity concentrations. As expected, we find that the momentum decay rate (and hence effectively the mobility) depends sensitively on the exact locations of the impurities within the structure because of long-range phase coherence. We also find that the momentum relaxation rate, calculated quantum mechanically, is larger than the "semiclassical" rate calculated from Fermi's Golden Rule possibly because of coherent backscattering that gives rise to the Anderson localization effect.

I. INTRODUCTION

Recent advances in nanolithography have made it possible to realize "mesoscopic structures" in which an electron can traverse the entire structure without encountering phase-randomizing inelastic collisions. In these structures, many new quantum-interference effects that arise from long-range phase coherence have been observed, with some of them promising novel device applications. Examples of such effects are Anderson localization,[1] the Aronov–Al'tshuler–Spivak effect,[2] the Aharonov–Bohm effect,[3] nonlocal voltage/current modulation,[4] and universal conductance fluctuations.[5] Recently, quantum-interference effects have also been predicted to influence not just the transport properties, but even the optical properties of mesoscopic samples, such as the photoluminescence spectrum.[6]

In order to study quantum-interference effects in disordered solids, it is necessary to develop practical, implementable models that can simulate quantum transport in the presence of elastic scattering. A number of such models have appeared in the literature. They are mostly based on either Green's function techniques[7] or scattering matrix formalisms.[8] However, almost all of these techniques have treated exclusively steady-state transport through spatially varying media. While steady-state transport is important in many cases, there are situations when transient transport is of interest. For example, transient phenomena (such as velocity overshoot) determines the operational limit of many modern ultrasmall devices. Consequently, it is transient transport, rather than steady-state transport, that is quite often of importance in the study of ultrasmall quantum devices.

In this paper, we have developed a simple but fully quantum-mechanical technique that can model time varying or transient transport phenomena in disordered structures. Using this technique one can extract the time evolution of an electron's wave function in a disordered medium (i.e., in the presence of elastic scattering) and hence calculate any time-dependent transport property of interest. In Sec. II, we describe the theoretical framework for this technique which is derived directly from the time-dependent Schrödinger Equation. The attractive feature of the technique is its utter simplicity and computational ease. In addition, it is extremely illustrative of the role of elastic scattering in quantum transport. In Sec. III, we apply this technique to compute the momentum (or velocity) relaxation of a single electron injected into a field-free region of a disordered two-dimensional semiconductor nanostructure. In Sec. IV we show how the momentum decay characteristics depend sensitively on the precise details of the impurity "configuration" because of quantum interference between the scatterers. We also compare the momentum decay rates calculated with our technique to the semiclassical results obtained from Fermi's Golden Rule. In Sec. V we discuss these results and finally, in Sec. VI, we present our conclusions.

II. THEORY

In order to model transient electronic transport through a two-dimensional array of elastic scatterers, we start from the time-dependent effective mass Schrodinger Equation

$$i\hbar\frac{\partial\psi}{\partial t} = -\frac{\hbar^2}{2m^*}\nabla_\rho^2\psi + H'\psi, \tag{1}$$

where $\psi \equiv \psi(\rho,t)$ is the electron's wave function, ∇_ρ^2 is the two-dimensional Laplacian, and H' is the impurity interaction Hamiltonian which is itself time independent since the scattering mechanism is elastic.

We can in most cases find any time-varying transport property of interest by simply solving the above equation to evaluate the time- and space-dependent wave function $\psi(\rho,t)$. While this can be done by straightforward brute-force techniques such as finite difference or finite element methods, there are other techniques that are simpler and at the same time more illustrative. We describe one such technique that we adopted. Since the Hamiltonian is time invariant, the wave function ψ can be expanded in a complete orthonormal set as follows:

$$\psi(\rho,t) = \sum_p c_p(t)\phi_p(\rho), \tag{2}$$

where the ϕ's are the so-called "normal modes" of the system

[a)] Present address: Ametek/Houston Instrument, Austin, TX 78757.

 0021-8979/89/235422-06$02.40 © 1989 American Institute of Physics

that are the solutions of the time-independent Schrödinger equation in the absence of impurity interaction.

The choice of ϕ_p depends on the situation. If we are interested in the electron's momentum, we must use basis states ϕ_p that are eigenstates of the momentum operator. In that case, the simplest choice will be to use plane-wave basis states which are momentum eigenstates

$$\phi_p = (1/\sqrt{A})\, e^{i\mathbf{k}\cdot\mathbf{p}} \equiv |\mathbf{k}\rangle \tag{3}$$

where A is the area of the sample.

Equation (2) now becomes

$$\psi = \sum_{\mathbf{k}} c_{\mathbf{k}}(t)|\mathbf{k}\rangle. \tag{4}$$

Substituting Eq. (4) into Eq. (1) and using the orthonormality of the basis states $|\mathbf{k}\rangle$, we get a differential equation for the time evolution of the amplitude $c_{\mathbf{k}}(t)$ of any state $|\mathbf{k}\rangle$;

$$i\hbar \frac{\partial c_{\mathbf{k}}(t)}{\partial t} = \frac{\hbar^2|k|^2}{2m^*} c_{\mathbf{k}}(t) + \sum_{\mathbf{k}'} H'_{\mathbf{k}'\mathbf{k}} c_{\mathbf{k}'}(t), \tag{5}$$

where

$$H'_{\mathbf{k}'\mathbf{k}} = \sum_{n=1}^{N} \langle \mathbf{k}'|V_n|\mathbf{k}\rangle. \tag{6}$$

Here V_n is the scattering potential due to the nth impurity and the sum accounts for the contribution of all the N impurities in the system.

If we now assume the scattering potential V_n to be a screened Coulomb potential, we get

$$H'_{\mathbf{k}'\mathbf{k}} = -\frac{q^2}{4\pi\epsilon} \sum_{n=1}^{N} \langle \mathbf{k}'|\frac{e^{-\lambda|\boldsymbol{\rho}-\boldsymbol{\rho}_n|}}{|\boldsymbol{\rho}-\boldsymbol{\rho}_n|}|\mathbf{k}\rangle$$

$$= \frac{1}{A}\sum_{n=1}^{N} \frac{-q^2}{2\epsilon\sqrt{|\mathbf{k}-\mathbf{k}'|^2+\lambda^2}} e^{i(\mathbf{k}-\mathbf{k}')\cdot\boldsymbol{\rho}_n}. \tag{7}$$

The quantity ϵ is the dielectric constant, λ is the screening constant, q is the electronic charge, and $\boldsymbol{\rho}_n$ is the two-dimensional coordinate of the nth impurity (scattering center).

The terms in the summation in Eq. (7) are complex quantities, i.e., they have phases associated with them which immediately indicate that the net impurity interaction (experienced by an electron) depends on interference between the impurities. The nature of this interference is determined by the precise locations of the impurities in the sample (or the impurity "configuration") since the impurity coordinate $\boldsymbol{\rho}_n$ appears explicitly in the phase factor $e^{i(\mathbf{k}-\mathbf{k}')\cdot\boldsymbol{\rho}_n}$. Note that the sum in Eq. (7), or the net impurity interaction, does not necessarily increase with the number of terms in the summation (i.e., the number of impurities in the sample) since the sum is a phasor sum, not an arithmetic sum. There can be phase cancellations between the terms so that the sum may either increase or decrease with an increasing number of impurities. Consequently, it is quite possible that an electron could sometimes experience *weaker* impurity interaction in a "dirtier" sample (interspersed with more impurities) than in a "cleaner" sample (with fewer impurities) if the impurity configuration is favorable. This means that a dirtier sample could sometimes exhibit a *higher* "mobility" than a cleaner sample purely as a result of quantum interference! Recently, this fact has received some attention in a slightly different context[9] leading to an investigation into the possibility of exploiting this feature to realize high mobility in selectively doped samples.

Returning to Eq. (5), we find that this equation clearly exposes the role of impurity scattering in electron transport. Impurity scattering (or for that matter any type of elastic scattering) does nothing more than simply couple an electron from an initial state $|\mathbf{k}\rangle$ to all the other states $|\mathbf{k}'\rangle$ in the system over time. Equation (5) is valid for any wave-vector state $|\mathbf{k}\rangle$ so that for a finite system with M number of states at the Fermi level, we can write M coupled differential equations of the type of equation (5) which together describe the overall behavior of the system in the linear response regime. This set of coupled differential equations can be combined into a matrix form,

$$i\hbar \frac{\partial[\mathbf{c}(t)]}{\partial t} = \mathscr{H}_0[\mathbf{c}(t)] + \mathscr{H}'[\mathbf{c}(t)], \tag{8}$$

where $[\mathbf{c}(t)]$ is a column vector whose elements are the amplitudes $c_{\mathbf{k}}(t)$ of the various wave-vector states $|\mathbf{k}\rangle$ at time t, $\mathscr{H}_0$ is the unperturbed Hamiltonian (a diagonal matrix whose elements are the kinetic energies $\hbar|k|^2/2m^*$ of the basis states $|\mathbf{k}\rangle$) and $\mathscr{H}'$ is the impurity interaction Hamiltonian whose matrix elements are given by Eq. (7).

Equation (8) is the governing equation for electron transport in the presence of elastic scattering. Mathematically, Eq. (8) belongs to the genre of *coupled mode equations* that are widely used in the analysis of microwave (or optical) waveguides and directional couplers.[10] It is not surprising that the governing equation of dissipationless electron transport should be the same as the governing equations of microwaves and optics since it is now widely recognized that in the phase-coherent regime, and in the absence of many-body effects, the physics of electron transport through a disordered solid is no different from the physics of microwave or light propagation through a disordered medium.[4] In fact, the one-dimensional (1D) Schrödinger Equation describing dissipationless "propagation" of electrons through a semiconductor structure (with spatially varying potential) is identical in mathematical form to the 1D Maxwell's Equation that describes propagation of monochromatic light or microwave through an inhomogeneous medium with a spatially varying refractive index. The potential (including the elastic scattering potential) plays the role of the refractive index.

Equation (8) has the general solution

$$[\mathbf{c}(t)] = \exp\{-[i(\mathscr{H}' + \mathscr{H}_0)t/\hbar]\}[\mathbf{c}(0)]. \tag{9}$$

Given the initial condition $[\mathbf{c}(0)]$, the above equation provides the amplitudes $\mathbf{c}_{\mathbf{k}}(t)$ of every state $|\mathbf{k}\rangle$ at any arbitrary instant of time t. Once these amplitudes are determined from Eq. (9), the time-dependent wave function $\psi(\boldsymbol{\rho},t)$ or $\psi(x,y,t)$ can be found readily from Eq. (4). From the wavefunction, one can calculate any time-dependent transport variable by simply calculating the expected value of the corresponding operator.

III. MOMENTUM RELAXATION OF AN ELECTRON IN A DISORDERED NANOSTRUCTURE

In this section, we have used the above formalism to study the momentum relaxation of a single electron injected into a field-free region of a disordered semiconductor nanostructure. The time-dependent momentum along any chosen direction in the structure (say, the **x** direction) can be found from the usual prescription,

$$\mathbf{p}(t) = \left\langle \psi(x,y,t) \left| -i\hbar \frac{\partial}{\partial x} \right| \psi(x,y,t) \right\rangle, \tag{10}$$

where $\psi(x,y,t)$ is obtained from Eqs. (4) and (9).

If we are interested in the time evolution of the momentum along the direction of the initial momentum $\hbar\mathbf{k}$, we can use Eqs. (4) and (10) to obtain

$$\mathbf{p}_{\mathbf{k}}(t) = \hbar\mathbf{k} \sum_{\mathbf{k}'} |c_{\mathbf{k}'}(t)|^2 \frac{\mathbf{k}\cdot\mathbf{k}'}{|\mathbf{k}|^2}. \tag{11}$$

We now define the momentum relaxation rate $1/\tau_m$ by calling τ_m the time it takes for the momentum to decay to $1/e$ times its initial value, i.e.,

$$\mathbf{p}_{\mathbf{k}}(\tau_m^{\mathrm{QM}}) = \frac{1}{e}\mathbf{p}_{\mathbf{k}}(0), \tag{12}$$

or equivalently

$$\sum_{\mathbf{k}'} |c_{\mathbf{k}'}(\tau_m^{\mathrm{QM}})|^2 \frac{\mathbf{k}\cdot\mathbf{k}'}{|\mathbf{k}|^2} = \frac{1}{e}. \tag{13}$$

The superscript QM stands for the quantum-mechanical result. Evidently, τ_m^{QM} depends on the impurity coordinates since the amplitudes $c_{\mathbf{k}'}$ in Eq. (13) depend on the coordinates. We have calculated the quantum-mechanical rates $1/\tau_m^{\mathrm{QM}}$ (using Eq. (12)] in two-dimensional GaAs structures for various impurity concentrations and configurations. We have then compared these results with the "semi-classical" results obtained from the usual Fermi's Golden Rule. The latter result is

$$\frac{1}{\tau_m^{\mathrm{FGR}}} = \frac{N_s m^* q^4}{\pi\hbar^3\epsilon^2}\int_0^{\pi/2} \frac{1}{\lambda^2 \operatorname{cosec}^2\theta + 4|k|^2}\, d\theta, \tag{14}$$

where N_s is the impurity concentration per unit area.

Unlike the quantum-mechanical result, the result of Eq. (14) does not depend on the impurity configuration. This obviously happens because Fermi's Golden Rule deals only with the probability of scattering (and not with the complex amplitude of scattering) so that it contains no information about the phase shifts associated with scattering. It is the phase shifts that depend on the impurity coordinates. In other words, the Golden Rule does not account for interference between the scattering events.[11] Alternatively, we can say that using the Golden Rule is equivalent to assuming that *all the impurities in a sample are located at exactly the same point within the sample* (point doping) in which case there can be no interference between the scatterers. In contrast, the quantum-mechanically calculated rate does take into account the precise configuration of the scatterers and exhibits strong dependence on the configuration because of quantum interference between the scatterers. In the next section we show that the dependence on the impurity configuration can be so strong that the standard deviation in the momentum relaxation rate (arising from a variation in the impurity configuration alone) can become comparable to the mean value of the rate itself.

IV. RESULTS

For purposes of analysis, we consider two-dimensional GaAs structures with dimensions of 1000×1000 A.[2] These structures can be fabricated individually by electron-beam-lithography or can be viewed as the constituents of a lateral surface superlattice.[12] For the impurity and carrier concentrations that we have considered in our simulations, the average transit time, i.e., the time it takes for an electron to diffuse across the sample, is always much shorter than the mean time between inelastic collisions. This condition can be expressed as

$$L < \sqrt{D\tau_{in}}, \tag{15}$$

where L is the lateral dimension of the structure, D is the diffusion coefficient, and τ_{in} is the mean time between inelastic collisions. The right-hand side is the inelastic diffusion length L_ϕ (also called the "phase-breaking length"). At temperatures around 4.2 K, L_ϕ in heavily doped GaAs samples typically exceeds 1 μm,[13] so that the condition of Eq. (16) is overly satisfied in all our simulations up to a temperature of 4.2 K.

The impurity configuration in our simulations is generated by two independent uniform random number generators that provide the x and y coordinates. A typical configuration is shown in Fig. 1. The initial momentum of the electron $\mathbf{p}_{\mathbf{k}}(0)$ is always assumed to be $\hbar\mathbf{k}_F$ where $\mathbf{k}_F$ is the Fermi wave vector whose magnitude is given by $|k_F| = \sqrt{2\pi n_s}$ with n_s being the two-dimensional carrier concentration. In the calculations, we have used 64 wave-vector states so that the matrix $\mathscr{H}'$ has a size 64×64. Inclusion of 64 wave-vector

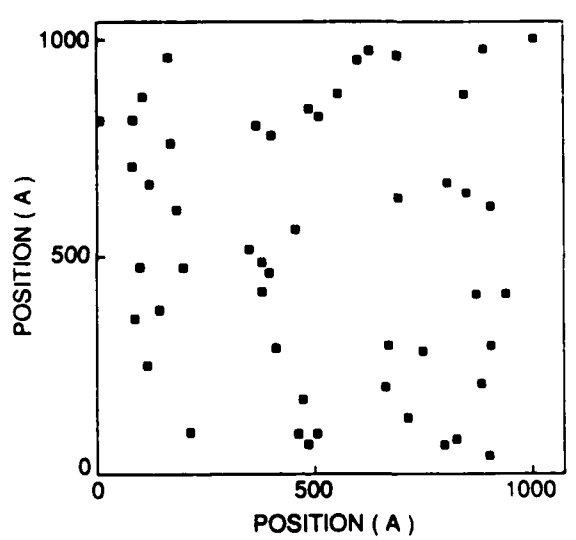


FIG. 1. A typical random configuration of impurities in a two-dimensional disordered structure. The configuration was generated by two independent uniform random number generators.

states was always sufficient to obtain convergence of the results. The 64 states that we have included in the calculation all have identical magnitudes of the wave vector (and are hence degenerate in energy), but the directions of the wave vectors are different. The inclusion of only degenerate states in the calculation is actually not quite rigorous since strong elastic scattering can couple an electron between nondegenerate states over short periods of time. In a more rigorous calculation, and also in order to obtain a reliable estimate of collisional broadening, one should include states with different magnitudes of the wave vector. Unfortunately, this taxes our present computational resources and is therefore left for future work.

In Fig. 2 we show a typical set of relaxation characteristics, i.e., how the normalized momentum $\mathbf{p}(t)/\mathbf{p}(0)$ decays with time for various impurity configurations and for a fixed impurity concentration. In this figure, the results are plotted over a time interval of 1 ps which is shorter than the time it takes for an electron to diffuse across the sample. The latter time is given by $t_{\rm diff} = t_{\rm transit} = A/D$, where A is the sample area and D is the semiclassical diffusion constant calculated from $D = (1/2)v_F^2\tau_m^{SC}$. In these examples, the value of $t_{\rm diff}$ is between 1.04 and 10.4 ps.

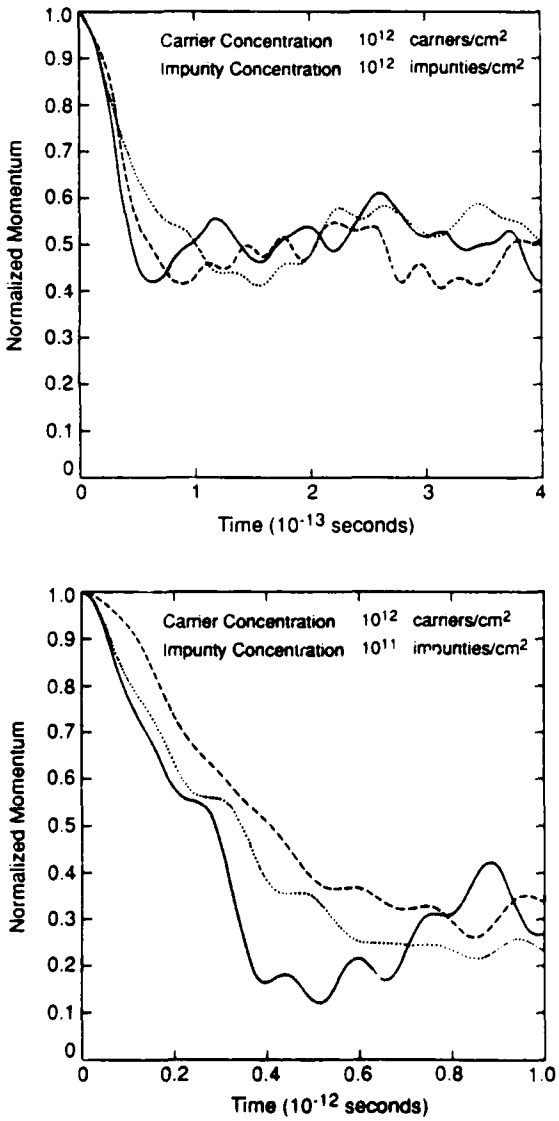


FIG. 2. The momentum relaxation characteristics for an electron in a disordered two-dimensional GaAs quantum box. The three different curves are for three different impurity configurations.

It appears from Figs. 2(a) and 2(b) that the momentum at first decays almost monotonically and then fluctuates around a steady-state value. Actually, this is somewhat deceptive since no real steady-state condition can ever be achieved in this system. We have a dissipationless (and hence reversible) finite system with a finite number of states. Such a system must obey Poincaré recurrence or the so-called wiederkehr effect.[14] That is, after a sufficiently long-time interval, the system must return to a state arbitrarily close to the initial state and the initial momentum will be restored. The time that elapses before such a return occurs is the so-called Poincaré cycle which depends on the number of states in the system. The larger the number of states, the larger is the period. In Figs. 2(a) and 2(b) we simply have not carried out the simulations long enough to observe the Poincaré recurrence. Nevertheless, the wiederkehr phenomenon is an essential feature of dissipationless transport and is actually implicit in Eqs. (8) and (9) that describe our system. Equation (8) is a set of coupled mode equations that are inherently reversible, and has a periodic solution given by Eq. (9). Consequently, the decay characteristics evaluated with Eqs. (8) and (9) will also exhibit periodic behavior and this period is the Poincaré cycle.

Since the decay characteristics are not monotonic over the entire simulation interval, we had to evaluate the momentum relaxation time τ_m^{QM} from the monotonic parts of the characteristics (sometimes employing extrapolation). The monotonic parts were found to converge rapidly with an increasing number of states. In every case, we obtained sufficient convergence after including 64 states in the calculation.

In Figs. 3(a) and 3(b) we show histograms of the momentum relaxation time τ_m^{QM} calculated for various "samples" with a given impurity concentration. Each "sample" is characterized by a particular impurity configuration. Each of our histograms is generated from 100 "samples" and in each histogram the carrier concentration and the impurity concentration are kept fixed.

In Fig. 4, we plot the semiclassical results for the momentum relaxation time along with the "average" quantum-mechanical results for various impurity concentrations. The "average" quantum-mechanical results are obtained by averaging over 100 different impurity configurations, which means that they are the mean values of the histograms. Finally, in Fig. 4, we also show the ratio of the standard deviation in τ_m^{QM} to the mean value of τ_m^{QM} for various impurity concentrations. Both the standard deviations and the mean values are calculated directly from the histograms.

V. DISCUSSION

The histograms in Figs. 3(a) and 3(b) exemplify the strong dependence of the momentum relaxation time τ_m^{QM} on the impurity configuration. As can be seen from the histograms, the relaxation time can vary over almost an order of

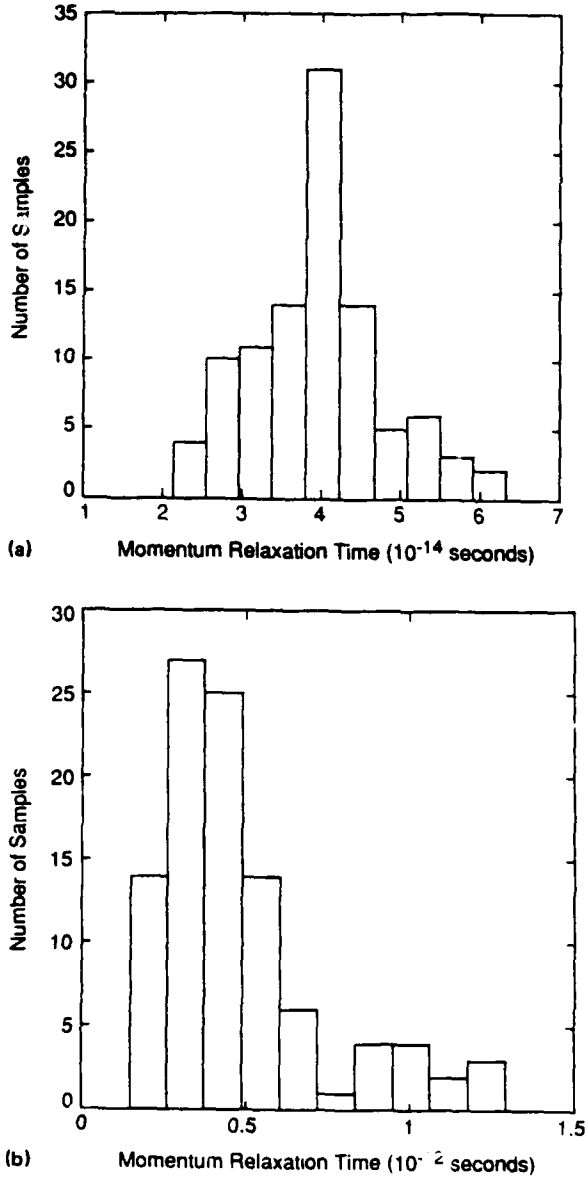


FIG. 3. The histograms showing the dependence of the momentum relaxation time on impurity configurations. (a) Carrier concentration = 10^{12} cm^{-2}, impurity concentration = 10^{12} cm^{-2}, average relaxation time = 39.6 fs, standard deviation = 8.06 fs. (b) Carrier concentration = 10^{12} cm^{-2}, impurity concentration = 10^{10} cm^{-2}, average relaxation time = 458 fs, standard deviation = 251 fs.

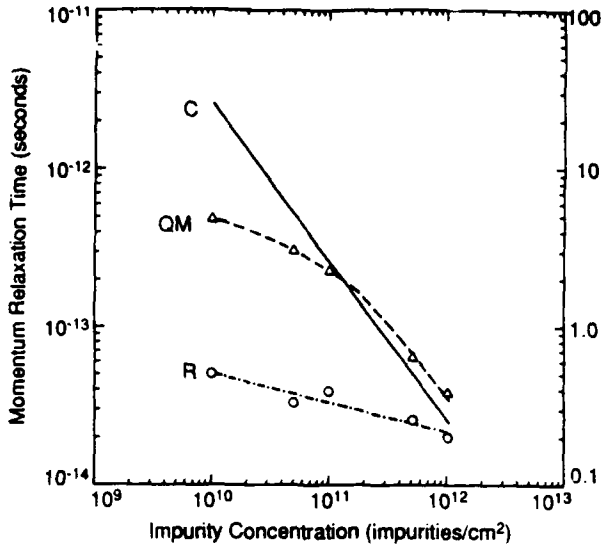


FIG. 4. Plots of the semiclassical and (average) quantum-mechanical momentum relaxation times vs impurity concentration. The curve labeled C is the semiclassical result and the curve labeled QM is the quantum-mechanical result. The average quantum-mechanical results are obtained by averaging over 100 different impurity configurations. The line through the quantum-mechanical results is a guide to the eye. The figure also shows the ratio of the standard deviation to the mean of the relaxation times (the curve labeled R).

magnitude depending on the configuration. Evidently the momentum relaxation time becomes large when the impurity configuration is such that the interference between the impurities [see Eq. (7)] becomes "destructive," i.e., there are phase cancellations. In that case, the net impurity interaction experienced by an electron is reduced which decreases the frequency of momentum-randomizing scattering events and thus increases the momentum-relaxation time. On the other hand, when the impurity configuration is "unfavorable," the interference between the impurities becomes "constructive." In that case, the net impurity interaction is enhanced which in turn reduces the momentum relaxation time. As an extreme case, we have found that the momentum relaxation times in a few samples with an impurity concentration of 10^{11} cm^{2} are actually *smaller* than those in some samples with a ten times higher impurity concentration of 10^{12} cm^{-2}. This is a strong manifestation of quantum-interference effects and demonstrates that dirtier samples can sometimes have higher mobilities than cleaner samples depending on the interference between the impurities!

It is interesting to note that the histogram in Fig. 3(a) is approximately Gaussian in shape. It has been observed before by several authors[15] that in the case of multichanneled diffusive transport, the conductances of samples (with varying impurity concentrations) exhibit a Gaussian distribution in the weak localization regime. In our case, we have two-dimensional samples of square geometry for which the conductance G is related to the momentum relaxation time as $G = e^2\tau_m n_s/m^*$ where n_s is the two-dimensional carrier concentration. For fixed n_s, the probability distribution of τ_m^{QM} will be the same as the probability distribution of the conductance G so that we also expect to observe a Gaussian profile. The observance of a Gaussian distribution in Fig. 3(a) is therefore in agreement with the observations of Ref. 15.

The histogram in Fig. 3(b) however is not Gaussian; it is significantly skewed to the right. We believe that this is due to the fact that this case corresponds to a very weakly disordered sample ($N_s = 10^{10}$cm^{-2}) in which transport is quasiballistic rather than diffusive. The skewing of the distribution to the right of the mean value is caused by the presence of ballistic electrons that hardly relax their momenta. It therefore appears that the deviation from the Gaussian profile is related to quasiballistic transport.

From Fig. 4, we find that the quantum-mechanically

calculated relaxation times are significantly different from the semiclassical results, especially for low impurity concentrations. The difference is as large as a factor of six for an impurity concentration of 10^{10} cm^{-2} which corresponds to quasiballistic transport. We also find that the quantum-mechanical result is always *less* than the semiclassical result for impurity concentrations less than $\sim 3\times 10^{11}$ cm^{-2}. This is probably caused by the well-known coherent backscattering effect[16] which is responsible for Anderson localization. Because of constructive interference between time-reversed Feynman paths (Cooperon channels) which gives rise to the coherent backscattering effect, there is an increased tendency for an electron to turn around inside a sample and travel backwards. This is obviously a drastic momentum relaxation process which will significantly decrease the momentum relaxation time. The phenomenon of backscattering is purely quantum mechanical in origin and is absent in the semiclassical picture. That is why the quantum-mechanically calculated relaxation time turns out to be smaller than the semiclassically calculated value in our simulations.

There is however an apparent reversal of this trend at higher impurity concentrations ($>3\times 10^{11}$ cm^{-2}) when the quantum-mechanical result exceeds the semiclassical result. The reason for this is the following. In our model, the randomness associated with the disorder is introduced through the random locations of the impurities (spatial disorder).[17] For $N_s > 3\times 10^{11}$ cm^{-2}, the average separation between impurities is $d = (N_s L)^{-1} < 34$ Å. Consequently, $k_F d < 0.83 < 1$. This means that the phase of an electron is not fully randomized over 2π between two neighboring impurities.[8] In that case, since there is no other source of randomization, the effect of disorder is not felt to the fullest extent and the quantum-mechanically calculated relaxation time is overestimated. In addition, if $k_F d$ is not much larger than unity, then evanescent states may also have a strong influence on transport.[17] Randomization of the phase over 2π is an essential ingredient of localization theory[18] and also a necessary condition in the model of Ref. 8. Failure to enforce this condition may lead to spurious results in quantum-mechanical models. The effect of having $k_F d < 2\pi$ has been discussed in Ref. 8.

Finally, the importance of quantum-interference effects in dissipationless transport becomes most evident from the ratio of the standard deviation to the mean of the relaxation time as plotted in Fig. 4. The ratio decreases with increasing impurity concentration which is expected since quantum-interference effects are more pronounced when transport is quasiballistic rather than diffusive. But more importantly, the ratio is quite large—it is larger than 0.5 for the lowest impurity concentrations—which shows that quantum-interference effects can introduce a significant spread in the apparent mobility (or diffusion coefficient) of an electron in the linear response regime. It is therefore important to include quantum-interference effects in calculating transport parameters such as the linear response mobility or diffusion constant when the ambient temperature is low enough that the phase coherence of the electron is preserved across the entire sample

VI. CONCLUSION

In this paper, we have introduced a simple quantum-mechanical formalism to study transient electronic transport in disordered mesoscopic structures in the absence of inelastic scattering. Application of this formalism to study the momentum relaxation of an electron in a disordered structure has revealed strong influences of quantum-interference effects in both diffusive and quasiballistic transport. This indicates that interference effects play an important role in transient transport and must be accounted for in modeling transient phenomena such as velocity overshoot that play a crucial role in the operation of many modern ultrafast devices.

ACKNOWLEDGMENT

The authors are indebted to Dr. Marc Cahay for many fruitful discussions. This work was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-88-0096 and by IBM.

[1] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).

[2] B. L. Al'tshuler, A. G. Aronov, and B. Z. Spivak, Pis'ma Zh. Eksp. Teor. Fiz. **33**, 101 (1981); JETP Lett. **33**, 94 (1981).

[3] Y. Aharonov and D. Bohm, Phys. Rev. **115**, 485 (1959).

[4] S. Datta, Superlatt. Microstruc. **6**, 83 (1989); D. C. Miller, R. K. Lake, S. Datta, M. R. Melloch, and R. Reifenberger, *Nanostructure Physics and Fabrication* (Academic, Boston, 1989), p. 165; F. Sols, M. Macucci, U. Ravaioli, and K. Hess, Appl. Phys. Lett. **54**, 350 (1989); F. Sols, M. Macucci, U. Ravaioli, and K. Hess, *Nanostructure Physics and Fabrication* (Academic, Boston, 1989), p. 157.

[5] A. D. Stone, Phys. Rev. Lett. **54**, 2692 (1985); P. A. Lee, Physica **140A**, 169 (1986).

[6] S. Bandyopadhyay, Phys. Rev. B **38**, 7466 (1988).

[7] P. A. Lee and D. S. Fisher, Phys. Rev. Lett. **47**, 882 (1981); D. J. Thouless and S. Kirkpatrick, J. Phys. **14**, 235 (1981); A. D. Stone, Phys. Rev. Lett. **54**, 2692 (1985).

[8] S. Datta, M. Cahay, and M. McLennan, Phys. Rev. B **36**, 5655 (1987); M. Cahay, M. McLennan, and S. Datta, Phys. Rev. B **37**, 10125 (1987).

[9] A. F. J. Levi, S. L. McCall, and P. M. Platzman, Appl. Phys. Lett. **54**, 940 (1989).

[10] A. Yariv, IEEE J. Quantum Electron. **QE-9**, 919 (1973).

[11] Fermi's Golden Rule can be modified to introduce some interference effects between scatterers by modifying the matrix element; see, for instance, S. Bhobe, W. Porod, and S. Bandyopadhyay, Proceedings of the 4th International Conference on Modulated Semiconductor Structures, Ann Arbor, MI, July 1989 (to be published).

[12] G. H. Bernstein and D. K. Ferry, J. Vac. Sci. Technol. B **5**, 964 (1987).

[13] B. J. F. Lin, M. A. Paalanen, A. C. Gossard, and D. C. Tsui, Phys. Rev. B **29**, 927 (1984).

[14] For a discussion of Poincaré recurrence or the wiederkehr effect see, for instance, G. H. Wannier, *Statistical Physics* (Dover, New York, 1966).

[15] P. W. Anderson, D. J. Thouless, E. Abrahams, and D. S. Fisher, Phys. Rev. B **22**, 3519 (1980); P. W. Anderson, Phys. Rev. B **23**, 4282 (1981); N. Giordano, Phys. Rev. B **38**, 4746 (1988).

[16] G. Bergmann, Phys. Rev. B **28**, 2914 (1983).

[17] M. Cahay, S. Bandyopadhyay, M. A. Osman, and H. L. Grubin, Proceedings of the 4th International Conference on Modulated Semiconductor Structures, Ann Arbor, MI, July 1989 (to be published).

[18] E. Abrahams, P. W. Anderson, D. C. Licciardello, and T. V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).

QUANTUM PHASE COHERENT EFFECTS IN THE PHOTOLUMINESCENCE SPECTRA OF DISORDERED MESOSCOPIC STRUCTURES [1]

S. Bandyopadhyay
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

At low enough temperatures, when the *inelastic* scattering time in a quantum dot exceeds the radiative recombination lifetime of photoexcited electrons and holes, the photoluminescence spectrum of the dot becomes sensitive to the exact *locations* of the elastic scatterers within the dot. This is a result of quantum interference whose nature is determined by the precise configuration of the elastic scatterers inside the dot. Several features of the photoluminescence spectra are influenced by the configuration, the most remarkable of which is the fact that the usual red-shift of the peak frequency, associated with bandgap renormalization, can change into a blue-shift depending on the configuration. The dependence of the optical spectra on the internal configuration is basically the same effect that makes (universal) conductance fluctuations sample-specific. An important consequence of this effect is that different quantum dots in a lateral surface superlattice will exhibit slightly different spectra if they merely have different impurity configurations, but are otherwise identical. The resulting inhomogeneous broadening can be comparable to the energy spacing between the subbands, so that it can sometimes mask the discreteness of the optical spectra expected of quasi-zero dimensional structures.

I. INTRODUCTION

It is well-known that elastic scattering does not destroy an electron's phase-memory so that quantum interference effects are not inhibited by impurity scattering at low enough temperatures. In a disordered semiconductor nanostructure,

[1]Supported by the Air Force Office of Scientific Research under grant no. AFOSR-88-0096 and by an IBM Faculty Development Award.

Copyright © 1989 by Academic Press, Inc
All rights of reproduction in any form reserved
ISBN 0-12-583000-X

if the inelastic scattering times of both electrons and holes exceed the radiative recombination lifetime of a photoexcited electron-hole pair, then the optical dipole constituted by the pair never loses its phase-memory during its entire lifetime. In that case, the time-dependent decay of the optical dipole, due to impurity scattering, is apt to be influenced by quantum interferences between the electron and hole states in the system. The decay of the optical dipole moment determines the lineshape of the optical spectrum - the lineshape is in fact obtained by Fourier transforming the decay characteristic - and consequently the lineshape will also be influenced by quantum interference effects and depend on the precise configuration of the elastic scatterers that determines the nature of this interference [1]. The important implication here is that in the presence of phase-coherence, not only are macroscopic *transport* properties (such as the conductance of a sample) a function of such microscopic details as the impurity configuration [2], but so are *optical* properties like the photoluminescence spectra. In other words, *"mesoscopic physics" includes not only transport phenomena, but also optical phenomena.*

A practical consequence of the latter is that different quantum dots in a lateral surface superlattice will exhibit slightly different spectra if they merely have different impurity configurations but are otherwise identical. This phenomenon therefore induces a unique kind of inhomogeneous broadening in the photoluminescence linewidth of a superlattice structure whose origin is purely quantum-mechanical and specifically arises from phase-coherence. In some instances, this inhomogeneous broadening can be so large that it can even mask the discreteness of the optical spectra expected of quasi-zero dimensional structures.

II. THEORY

The time-dependent decay of the optical dipole moment P(t) associated with transitions between a conduction band state and a valence band state in a disordered quantum dot is given by [1]

$$P(t) = \frac{1}{2} P(0) \sum_p c_{ep}^*(t) c_{hp}(t) \; + \; c.c., \tag{1}$$

where $c_{ep}(t)$ and $c_{hp}(t)$ are the time-dependent complex amplitudes of the pth electronic state and the pth hole state that the photoexcited electron and hole couple to at time t via the impurity interaction, and the summation over p is carried out to include all such states in the system. The lineshape of the photoluminescence spectra $\mathcal{F}(\omega)$ is obtained by Fourier transforming the time-dependent decay of P(t) into the frequency domain ω of the incident photons.

The task here is to evaluate the amplitudes $c_{ep}(t)$ and $c_{hp}(t)$. For both electrons and holes, these amplitudes are found from [1]

$$[c(t)] = exp\left[-\frac{i\mathcal{H}t}{\hbar}\right][c(0)] \; , \tag{2}$$

where $\mathcal{H}$ is the Hamiltonian matrix for the disordered system and $[c(t)]$ is a column vector whose elements are the complex amplitudes $c_p(t)$ of the various electronic or hole states that the photoexcited electron or hole couples to. The Hamiltonian $\mathcal{H}$ can be expressed as

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}' , \tag{3}$$

where $\mathcal{H}_0$ is the unperturbed Hamiltonian (a diagonal matrix) whose elements are the kinetic energies of the various subband states that the photoexcited electron and hole couple to, and $\mathcal{H}'$ is the impurity interaction Hamiltonian whose elements are given by

$$H'_{pq} = -\frac{\Gamma q^2}{4\pi\epsilon}\sum_i < \phi_p(r)|\delta(r-r_i)|\phi_q(r) >= -\frac{\Gamma q^2}{4\pi\epsilon}\sum_i \phi_p^*(r_i)\phi_q(r_i) \tag{4}$$

In the above equation, $\phi_p(r)$ is the wavefunction of the pth subband state that the electron or hole couples to as a result of the impurity interaction, i.e. it is the pth eigenfunction of $\mathcal{H}_0$. The impurity potentials were assumed to be δ potentials located at coordinates at r_i. The summation is carried out over the coordinates of all the impurities in the system and the sum obviously *depends on the exact locations of the impurities*. The parameter Γ is a parameter representing the strength of the interaction. The choice of δ-scatterers (instead of screened Coulomb scatterers) in our model is merely a matter of convenience; it does not alter the essential physics.

Since the Hamitonian for the system is now clearly dependent on the coordinates of the impurities, it is obvious that the amplitudes $c_p(t)$ (see Equation (2)) and hence the time-dependent optical dipole moment **P**(t) (see Equation (1)) will also depend on the exact *locations* of the impurities within the system. Consequently, the optical spectrum of a sample will be a "fingerprint" of the internal configuration of the scatterers.

III. EXAMPLE

By way of an example, we have calculated the photoluminescence lineshape (corresponding to an electron-light hole transition) for a two-dimensional quantum dot with a parabolic confining potential. The material was assumed to be GaAs. The impurity coordinates were generated by random number generators and the concentration was 5 x 10^{11} cm^{-2}. In our calculation, we included 36 electronic states and 36 hole states[2]. The calculated lineshape is plotted in Fig. 1.

[2]These states need not be degenerate in energy, since strong impurity scattering, even though elastic, can couple an electron or hole between states that are non-degenerate in energy over short periods of time.

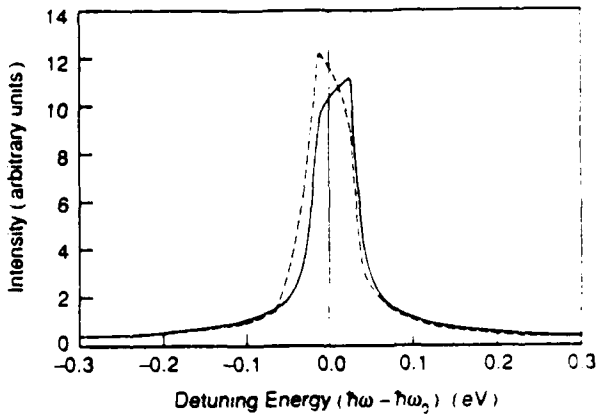


Fig.1 The photoluminescence spectrum (corresponding to an electron-light hole transition) for a quantum dot with parabolic confining potential. The incident photon energy $\hbar\omega_0$ is 1.4623 ev (the bandgap of GaAs is assumed to be 1.42 ev). The solid and dashed lines are for two different impurity configurations. Note that for one configuration, the peak frequency is red-shifted and for the other, it is blue-shifted. The difference corresponds to an energy of $\sim$ 4 mev which is comparable to the energy separation between the subband states ($\sim$ 6 mev). Consequently, the inhomogeneous broadening caused by varying impurity configuration, in different quantum dots in a lateral surface superlattice, may mask the discreteness of the optical spectra expected of quantum dots.

IV. DISCUSSION

To understand the nature of the photoluminescence lineshape in Fig. 1., we have to first recast Equation (1) in the form

$$P(t) = \frac{1}{2}P(0)exp(i\omega_0 t)\sum_p b_{ep}^*(t)b_{hp}(t) \; + \; c.c., \qquad (5)$$

where

$$b_{ep}(t) = exp\left[i\frac{E_e t}{\hbar}\right]c_{ep}(t) \quad ; \quad b_{hp}(t) = exp\left[i\frac{E_h t}{\hbar}\right]c_{hp}(t) \qquad (6)$$

In Equation (6), E_e and E_h are the energies of the states to which the electron and hole are photoexcited by the incident radiation and ω_0 is the resonant photon frequency corresponding to this transition, i.e.

$$\hbar\omega_0 = E_e - E_h \qquad (7)$$

We can now recast Equation (5) as

$$\begin{aligned} P(t) &= \frac{1}{2}P(0)exp(i\omega_0 t)A(t)e^{-i\theta(t)} + c.c. \\ &= P(0)A(t)cos\left[\omega_0 t - \theta(t)\right] , \end{aligned} \tag{8}$$

where $A(t)$ is the magnitude and $\theta(t)$ the phase of the complex product $\sum_p b_{e_p}^*(t)b_{h_p}(t)$ representing the decay of the dipole moment. Obviously, both $A(t)$ and $\theta(t)$ depend sensitively on the impurity configuration.

The photoluminescence lineshape $\mathcal{F}(\omega)$ is obtained by Fourier transforming $P(t)$

$$\mathcal{F}(\omega) = \frac{1}{2\pi}\int e^{i\omega t}P(t)dt \tag{9}$$

Hence we see that the phase $\theta(t)$ in Equation (8) has two effects. Firstly, it makes the lineshape asymmetric about the peak frequency, and secondly, it shifts the peak frequency away from the resonant frequency ω_0. This shift is associated with the real part of the self-energy correction (for both electrons and holes) due to impurity interaction which renormalizes the effective bandgap. Ordinarily, one would expect a shift to lower frequencies, i.e. a red-shift. However, we find from Equation (8) and (9) that depending on $\theta(t)$, or the precise details of the impurity configuration, the shift can be either a red-shift or a blue shift! That means that in the phase-coherent regime, quantum interference effects influence even the bandgap renormalization! This is truly a surprising result and is verifiable experimentally. A change in the sign of the shift is a remarkable effect of quantum interference and an intriguing case of microscopic features affecting macroscopic observables in a non-trivial way.

Finally, the only issue that remains to be discussed is the temperature at which such an effect could be observed. Inelastic scattering times of $\sim$ 10 ps have been measured at 4.2 K in GaAs samples [3] with a carrier concentration exceeding 7 x 10^{11} cm^{-2}, whereas a radiative recombination lifetime of $\sim$ 20 ps has been calculated for quantum dots [4]. Since the only requirement to observe the above effect is to ensure that the inelastic scattering time exceeds the radiative recombination lifetime, it is conceivable that this effect can be observed at temperatures not too far below liquid helium temperature. This makes it practical to verify this effect in semiconductor quantum dots.

REFERENCES

1. Bandyopadhyay, S. (1988). Phys. Rev. B15 38, 7466.
2. Al'tshuler, B. L. (1985). Pis'ma Zh. Eksp. Teor. Fiz. 41, 530 [JETP Lett. 41, 363].
3. Lin, B. J. F., Paalanen, M. A., Gossard, A. C., and Tsui, D.C. (1984). Phys. Rev. B. 29, 927.
4. Takagahara, T., (1987). Phys. Rev. B15. 36, 9293.

Real-space transfer from quantum wells by polar optical phonon scattering

Craig S. Lent, Lie Liang, and Wolfgang Porod
Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556

(Received 7 February 1989; accepted for publication 29 March 1989)

We calculate the rate at which electrons bound in a semiconductor quantum well are scattered out of the well by absorption or emission of polar optical phonons. This escape rate is important in real-space transfer devices and as a loss mechanism in high electron mobility transistors. Both quantum confinement effects and the two-dimensional (2D) to three-dimensional (3D) nature of the scattering are included. For all but very shallow wells, the real-space transfer rate is much smaller than either the bulk (3D) scattering rate or the confined (2D) scattering rate. Quantum resonances in the final state result in oscillations in the scattering rate as a function of electron energy.

Electrons bound in a one-dimensional quantum well can be scattered out of the well if they have sufficient energy. This real-space transfer effect has been exploited in storage and switching devices[1] and may limit performance of high electron mobility transistors.[2,3] A phonon scattering event can provide the necessary momentum to transfer the electron out of the well. This scattering has been treated thus far using either bulk scattering rates or the two-dimensional scattering rates appropriate to electrons confined in the well. As noted by Brennan and Park[3] in reporting recent Monte Carlo calculations of real-space transfer, one should really take into account the change in effective dimensionality between the initial and final states. Scattering from pseudo-two-dimensional (2D) states into three-dimensional (3D) states presents some subtleties, particularly in the normalization of the states. We present calculations for these escape rates for a model system, correctly accounting for both the quantum confinement and the change in dimensionality of the states. The results reveal that the escape scattering rates are much smaller than either the bulk or the 2D rates. The effects of continuum resonances are also apparent.

Our calculation extends the work of Müller *et al.*, who examined the effect of quantum reflections on optical scattering rates in quantum well structures.[4,5] They focused particularly on effective mass discontinuities, which we neglect here, and the conditions necessary to minimize quantum reflections by impedance matching. While noting that the real-space transfer rate appeared small, they did not actually calculate it. The 2D to 3D nature of the transition was simulated by placing the finite quantum well inside a much larger well with infinite barriers. Here we use free, current-carrying final states explicitly. They also examined scattering from states with initial momentum perpendicular to the heterointerfaces, whereas we consider only initial states bound in the well.

We calculate the escape rates for electrons scattered by emission or absorption of polar optical phonons using the Fermi golden rule. We use effective mass wave functions and a model semiconductor system with spherical energy bands and an effective mass of $m^* = 0.063 m_0$. Differences in the effective masses in the well and boundary material are neglected. For the polar optical phonon energy we use $\hbar\omega_0 = 0.036$ eV. The initial and final state wave functions are calculated by solving the one-dimensional Schrödinger equation for the effective heterostructure potential. The final state wave functions are chosen to be the scattering states with current incident from the left or right. Thus, our final states are truly free 3D states and carry current. We normalize the wave function by "box normalizing" the incoming plane wave. For the state with current to the right this means

$$\psi_{k_z}^{+}(z) = \begin{cases} L_z^{-1/2}(e^{ik_z z} + r^{+} e^{-ik_z z}) & \text{if } z<0 \\ u^{+}(z) & \text{if } 0<z<a, \\ L_z^{-1/2} t^{+} e^{ik_z z} & \text{if } a<z \end{cases}$$

where the well extends from 0 to a along the z axis and has a depth of V_0. L_z is the (arbitrary) normalization length for the incoming plane wave.

The escape rate for an electron in a bound state labeled by n, with total energy E is then

$$W_n(E) = \frac{L_z}{(2\pi)^3}\left(\frac{e^2\omega_0}{2\epsilon_p}\right) N^{\pm}(\omega_0) \times \int dk_z' \int d^3q \frac{1}{q^2} G^2_{n,k_z'}(q_z)\delta(E' - E \mp \hbar\omega_0). \tag{1}$$

Here ω_0 is the optical phonon frequency, $N^{\pm}(\omega_0)$ is the phonon occupation factor, and $1/\epsilon_p \equiv 1/\epsilon_\infty - 1/\epsilon_0$. The integrations are performed over all k_z''s, the z component of the final state wave vector, and over all $\mathbf{q}$, the wave vectors of the absorbed or emitted phonon. The square of the effective matrix element, $G^2_{nk_z'}(q_z)$, is given by

$$G^2_{n,k_z'}(q_z) = \left| \int_{-\infty}^{+\infty} \psi^*_{k_z'}(z) e^{iq_z z} \psi_n(z) dz \right|^2 . \tag{2}$$

Notice that because the final state wave function, $\psi'_{k_z}(z)$ includes a factor of $L_z^{-1/2}$, the factor of L_z in Eq. (1), which comes from the final density of states, is exactly canceled. As a result the rate is manifestly independent of this normalizing factor, as it should be.

For a very shallow well the phonon scattering rate should approach the bulk scattering rate. The rate for scattering from a well with 300 Å width and 0.005 eV depth is shown in Fig. 1 for $T = 300$ K. The rates are shown normal-

 0003-6951/89/232315-03$01.00 © 1989 American Institute of Physics

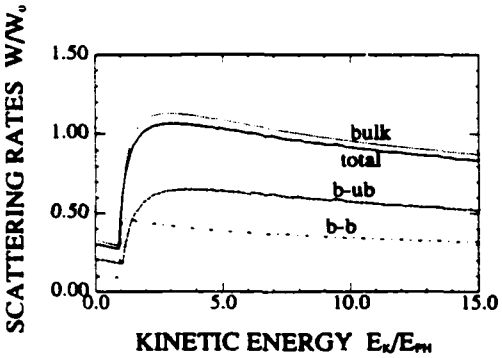


FIG. 1. Optical phonon scattering rates for a very shallow well. The well has a width of 300 Å and a depth of 0.005 eV. Shown are the bound-to-bound rate (dot-dashed curve) and the bound-to-unbound (i.e., escape) rate (dashed curve). The total rate, bound-to-bound plus bound-to-unbound, is the solid curve. For comparison the bulk scattering rate is shown with a dotted curve.

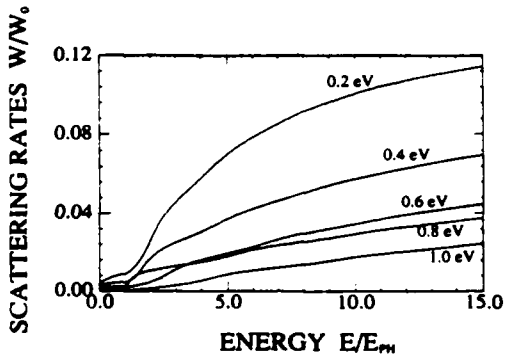


FIG. 3. Escape rates for scattering out of a well by polar optical phonon scattering. The initial state is the lowest bound state of the well. The physical parameters are the same as in Fig. 2 with the well depth increasing as indicated. Some continuum resonance effects are apparent at low energies.

ized to $W_0 \equiv 2\alpha_{ep}\omega_0$, where α_{ep} is the dimensionless electron-phonon coupling constant given by $\alpha_{ep} = (e^2/4\pi\hbar\epsilon_p)(m^*/2\hbar\omega_0)^{1/2}$. The bulk scattering rate in these units is of order unity. When the total escape rate is added to the bound-bound rate, the rate for scattering and staying in the well, the result is very close to the bulk scattering rate.[6,7] Notice that even for this very shallow well scattering in the well is almost as likely as scattering out of the well.

As the well depth increases, the escape rate decreases, as is shown in Figs. 2 and 3. These figures show the scattering rates for escape from the lowest bound state of the well. The total scattering rate is still comparable to the bulk rate but most of the contribution comes from the bound-bound rate. Wells of 0.2 eV depth (300 Å width) result in an escape rate which is already an order of magnitude smaller than the bulk rate. By the time the well depth reaches 1.0 eV, the escape rates have dropped by nearly a factor of 100 (see Fig. 3).

In addition to the large effect in the magnitude of the rates there is a change in the energy dependence of the scattering rate as the well deepens. For free carriers in the bulk the scattering rate rises rapidly to its maximum (roughly W_0) when the initial energy is one or two optical phonon energies. It then decreases slowly (roughly as $E^{-1/2}$) at

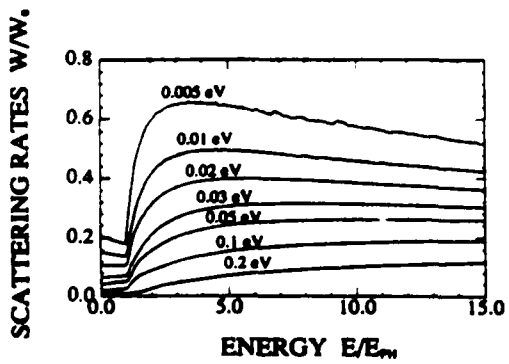


FIG. 2. Escape rates for scattering out of a well by polar optical phonon scattering. The well has a width of 300 Å and the various depths shown. The energy scale is the initial electron energy in units of the optical phonon energy (in this case, 0.036 eV). Initially the electron is in the lowest bound state of the well.

higher energies. This bulk-like behavior is preserved for very shallow wells. In contrast, for deeper wells the rate increases monotonically, although slowly, with energy. The transition from bulk-like behavior to the characteristic bound-unbound behavior takes place when the well depth is comparable to the optical phonon energy, as can be seen in Fig. 2. This change in the character of the energy dependence has its origin in the loss of strict momentum conservation in the z direction, which results in a larger number of allowed electron and phonon final states.

The dominant feature of the escape rates is their small magnitude relative to bulk rates. Notice that these greatly suppressed scattering rates apply even for electrons with energies significantly larger than the well depth. These electrons have large kinetic energy, with all of their momentum directed down the well, parallel to the well walls. It would seem, at least at first, that all they need to escape is to shift their momentum slightly so that they have a nonzero z component. Why is it so difficult to leave the well?

The full analysis of this behavior requires a careful examination of the effective matrix element and the range of final state integration permitted in Eq. (1). This will be presented in a longer and more detailed treatment than space limitations here permit.[8] A simple qualitative argument will suffice to make the major points. Consider the matrix element defined in Eq. (2). The initial state is the lowest bound state in the well, roughly a sine wave with wavelength $2a$, and varies slowly across the well width. Consider now a final state which is just barely free, i.e., k_z' is small and $E_z' \equiv (\hbar^2 k_z'^2/2m^*)$ is small. Its wavelength outside the well is very long, but it oscillates more rapidly in the well. The wave vector associated with this oscillation is

$$\kappa_z' = \sqrt{2m^*(E_z' + V_0)}/\hbar.$$

Since the initial state is nearly zero outside the well, the integral in Eq. (2) is dominated by the well region. If the final state wave function oscillates rapidly in this region, then the integral will be nearly zero unless the $e^{iq_z z}$ term from the phonon cancels out this oscillation. Thus, the dominant contribution to the matrix element will be from phonons with $q_z = \kappa_z'$. Because the electron-phonon coupling has a factor of

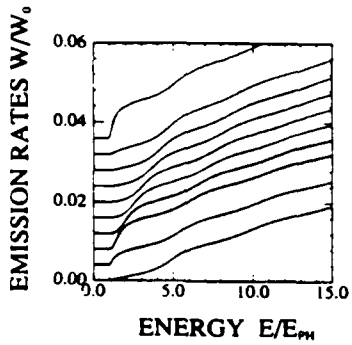


FIG. 4. Escape rates for scattering out of a well by polar optical phonon emission for several deep wells. The initial state is the lowest bound state of the well. The well width is 300 Å and the well depth is 0.80 eV (top curve), 0.82, 0.84, 0.86, 0.88, 0.90, 0.92, 0.94, 0.96, and 0.98 eV (bottom curve). The curves have the zero offset on the vertical scale for clarity. The enhancement in scattering rates near continuum resonances is clearly evident.

$1/q^2$ in it, the rate is suppressed accordingly. As the well becomes deeper, κ_z' increases, requiring larger q phonons to scatter out of the well, and lowering the scattering rate further. Thus, although only a small perpendicular momentum is needed to transfer an energetic, bound electron out of the well, only phonons with a sufficiently large wave vector can contribute to escape.

The quantum mechanical argument has a simple classical analogue. A classical particle traveling in a square well trough may have enough kinetic energy due to its high speed that its total energy is greater than the potential barriers of the well. Consider a particle with just barely enough total energy to escape, i.e., one which has a total energy of nearly zero (the zero of potential energy being at the top of the well walls). If it is elastically scattered, say by a bump in the well floor, it may escape. Energy conservation requires, however, that its velocity after it escapes is small. Thus, the change in its momentum must be large—almost all of its previous momentum down the trough is now gone. Thus, only scattering with a large momentum transfer can free the particle from the well. The $1/q^2$ factor associated with optical phonon scattering suppresses large momentum transfer events.

Finally, we consider briefly the effect of continuum resonances on the scattering rate. For shallow wells these are not important, but as the well strength increases, the effect of quantum reflections at the well walls becomes significant. The scattering rates exhibit oscillations at energies where the final state amplitude is enhanced by being resonant with the well. Figure 4 shows the rate for escape by phonon emission from the lowest bound state for a sequence of wells with increasing depths. The enhancements can be seen to occur near the resonance conditions. These resonances move down in energy as the well is deepened. Were the system completely one dimensional these oscillations would have genuine maxima and minima.[9,10] The ability to spread the final state energy over parallel momentum components smears out these oscillations. This results in the step-like structure observed.

The work was supported by the Air Force Office of Scientific Research under grant No. AFOSR-88-0096, and the National Science Foundation under contract No. EET87-07628.

[1]K. Hess and G. J. Iafrate, in *Hot-Electron Transport in Semiconductors*, edited by L. Reggiani (Springer, Berlin, 1985), p. 209.
[2]K. Hess, H. Morkoç, H. Shichijo, and B. G. Streetman, Appl. Phys. Lett. 35, 469 (1979).
[3]Kevin F. Brennan and Duke H. Park, J. Appl. Phys. 65, 1156 (1989).
[4]J. F. Müller, S. Schmitt-Rink, and A. F. J. Levi, Appl. Phys. Lett. 52, 236 (1988).
[5]J. F. Müller, A. F. J. Levi, and S. Schmitt-Rink, Phys. Rev. B 38, 9843 (1988).
[6]F. A. Riddoch and B. K. Ridley, J. Phys. C 16, 6971 (1983).
[7]B. K. Ridley, *Quantum Processes in Semiconductors* (Oxford, New York, 1982).
[8]Craig S. Lent, Lie Liang, Wolfgang Porod (unpublished).
[9]Wolfgang Porod and Craig S. Lent, Solid-State Electron. 31, 359 (1988).
[10]Craig S. Lent and Wolfgang Porod, Superlatt. Microstruct. 4, 77 (1988).

Solid-State Electronics Vol. 32, No. 12, pp. 1479–1483, 1989
Printed in Great Britain. All rights reserved

0038-1101 89 $3.00 + 0.00
Copyright © 1989 Pergamon Press plc

REAL SPACE TRANSFER RATES FOR POLAR OPTICAL PHONON SCATTERING FROM ASYMMETRIC QUANTUM WELLS

Craig Lent and Lie Liang

Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

ABSTRACT

We calculate the rate at which electrons bound in asymmetric semiconductor quantum wells are scattered out of the well by absorption or emission of polar optical phonons. The 2-D to 3-D nature of the scattering is included. The final states after scattering are states which carry current either to the right or left. We find that rates for scattering out of the well can be significantly smaller than bulk scattering rates. We also show that asymmetries in the well shape result in a directional dependence for the final state current. That is, electrons scatter out preferentially to the left or to the right depending on the details of the well potential.

KEYWORDS

Quantum wells; real-space transfer; phonon scattering; hot electron transport.

INTRODUCTION

Electrons in a quantum well can be heated by applying a field parallel to the well walls. Such electrons can gain energy sufficient to escape the well. This real space transfer phenomena has been studied extensively by Hess and others (Hess, 1981), and is exploited in the CHINT and NERFET devices (Kastalsky, 1984; Luryi, 1984). Extensive analysis of these structures has been done using semiclassical Monte Carlo techniques (Brennan and Park, 1989). A necessary input to such calculations is the scattering rate from the confined electron states to the free states out of the well. Currently, approximations are used which neglect the 2-D to 3-D nature of the transition. We calculate the rates for scattering from the bound 2-D states of a quantum well into free 3-D states by emission or absorption of polar optical phonons. We focus here on the effects of well shape on the scattering-out process. In particular, we investigate the effect of asymmetrically shaped wells on the symmetry of the scattering.

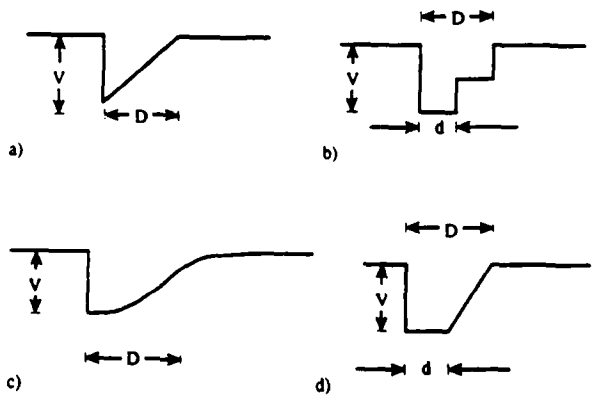


Fig. 1. Potential energy profiles for some asymmetric quantum wells

THEORY

We calculate polar optical phonon scattering rates using the Fermi Golden rule,

$$W = \frac{2\pi}{\hbar}\int |<\psi^f|\mathbf{H}_{e-ph}|\psi^i>|^2\delta(E_i - E_f)\ dS_f, \tag{1}$$

where the integral is over all final states. Initial and final wavefunctions are solutions to the single-band effective mass Schrödinger equation. We assume a parabolic band with effective mass $m^* = 0.063m_0$. Differences in the effective mass in the well and barrier materials are neglected. For the optical phonon energy we use $\hbar\omega = 0.036$ eV. The quantum well is in the z-direction, extending from $z = 0$ to $z = a$. The initial bound-state wavefunction consists of a product of plane waves in the x and y directions and a function $\psi_n(z)$ which is calculated numerically for each potential. Our approach is similar to Müller and coworkers (1988).

The final state wavefunctions are chosen to be eigenfunctions of the current operator which carry current either to the right $(+z)$ or to the left $(-z)$.

$$\psi^R_{k_z}(z) = \begin{cases} L_z^{-1/2}\left(e^{ik_z z} + r_R e^{-ik_z z}\right) & \text{if } z < 0 \\ u_R(z) & \text{if } 0 < z < a \\ L_z^{-1/2}\ \ t_R e^{ik_z z} & \text{if } a < z \end{cases}, \tag{2}$$

$$\psi^L_{k_z}(z) = \begin{cases} L_z^{-1/2}\ \ t_L e^{-ik_z z} & \text{if } z < 0 \\ u_L(z) & \text{if } 0 < z < a \\ L_z^{-1/2}\left(e^{-ik_z z} + r_L e^{ik_z z}\right) & \text{if } a < z \end{cases}. \tag{3}$$

We normalize the wavefunction by "box-normalizing" the incoming plane wave with L_z being the (arbitrary) normalization length. The final state wavevector, k_z, is always taken to be positive. The functions $u_R(z)$ and $u_L(z)$ are found by numerically solving the Schrödinger equation in the well region and matching to the form of the solution outside the well. The complete final-state wavefunctions can be written

$$\psi_{k'_x,k'_y,k'_z}(x,y,z) = Ae^{ik_x x + ik_y y}\psi_{k_z}(z)$$

where the final state energy is

$$E' = \frac{\hbar^2}{2m^*}(k'^2_x + k'^2_y + k'^2_z).$$

The rate for an electron to escape from a bound state labeled by n, with total energy E, into a state with current to the right (R) or left (L) is given by (Ridley,1982; Lent, Liang and Porod, 1989)

$$W_n^{(R/L)}(E) = \frac{L_z}{(2\pi)^3}\left(\frac{e^2\omega_0}{2\epsilon_p}\right)N^{\pm}(\omega_0)\int_0^\infty dk'_z\int d^3q\,\frac{1}{q^2}|G^{(R/L)}_{n,k'_z}(q_z)|^2\delta(E' - E \mp \hbar\omega_0). \tag{4}$$

The upper and lower sign denote phonon emission absorption respectively. The phonon occupation factor is given by

$$N^{\pm}(\omega_0) = \frac{1}{exp(\hbar\omega_0/k_BT) - 1} + \frac{1}{2} \pm \frac{1}{2},$$

and $1/\epsilon_p \equiv 1/\epsilon_\infty - 1/\epsilon_0$. The integrations are performed over all positive k'_z and over all q, the wavevector of the absorbed or emitted phonon. The square of the effective matrix element is given by

$$|G^{(L/R)}_{n,k'_z}(q_z)|^2 = \left|\int_{-\infty}^{+\infty}[\psi^{(L/R)}_{k'_z}(z)]^* e^{iq_z z}\psi_n(z)dz\right|^2. \tag{5}$$

Because the final state wavefunction includes the normalizing factor of $L_z^{-1/2}$, the factor of L_z in Equation (4), which comes from the final density of states, is exactly canceled. As a result, the rate is manifestly independent of this normalizing factor, as it must be.

RESULTS

We examine optical phonon scattering out of several asymmetric quantum well structures. Figure 1a shows the profile of a well with one abrupt interface and one linearly graded interface. The length of the well is 300 Å and the maximum depth is 0.2 eV. We assume a temperature $T = 300$ K. Figure 2 shows the total rate (phonon emission plus absorption) for escape from the well by polar optical phonon scattering. The initial state is taken to be the lowest bound state of the well. The rate is plotted as a function of the initial kinetic energy, in units of the optical phonon energy. Note that this energy can be much larger than the well depth because the bound electron can have a large momentum parallel

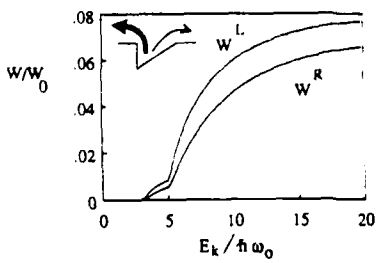


Fig. 2. The normalized escape rates for the potential profile shown in Figure 1a. The rates are shown for scattering into final state with current going to the right (W^R) and to the left (W^L).

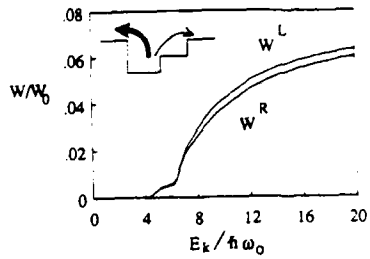


Fig. 3 The normalized escape rates for the potential profile shown in Figure 1b.

to the well walls. The rate is shown for scattering out of the well into states with current in the positive z direction (W^R) and in the negative z direction (W^L).

An important feature of this result is the magnitude of the rates. The rates are shown normalized to $W_0 \equiv 2\alpha_{ep}\omega_0$, where α_{ep} is the dimensionless electron-phonon coupling constant given by $\alpha_{ep} = (e^2/4\pi\hbar\epsilon_p)(m^*/2\hbar\omega_0)^{1/2}$. The bulk scattering rate in these units is of order unity, as is the bound-bound scattering rate. The rate for transfer out of the well is an order of magnitude smaller. We have discussed this suppression of the real-space transfer rates elsewhere (Lent and Porod, 1988; Lent, Liang, and Porod, 1989) and shown that it grows to a factor of 100 for a 1.0 eV well. The reason for this can be seen by considering the quantity $E_z \equiv E - \hbar^2/2m^*(k_x^2 + k_y^2)$. In the initial state E_z is negative, roughly -0.2 eV. The final state has an E_z which is positive, although it may be very small if the final-state electron is just barely free. Since the total energy, E can change only by $\hbar\omega_0$, most of the additional energy comes at the expense of momentum in the x and y directions. Since momentum parallel to the well walls is conserved, this implies a phonon scattering event that has a large $q_{\parallel} \equiv \sqrt{q_x^2 + q_y^2}$. The $1/q^2$ factor in the scattering rate, which is characteristic of polar optical scattering, suppresses large momentum transfer events. Thus the rate of transfer out of the well is small compared to bulk rates or bound-bound scattering rates.

Here we focus on the other obvious feature of Figure 2, the difference in the scattering rate for scattering to the right and to the left (*i.e.*, into states with current in the $+z$ and $-z$ directions). The scattering out of the well is preferentially to states with current going in the direction from the graded interface to the abrupt interface. We can define the relative difference of the two rates,

$$\Delta W(E)/W = \frac{W^L(E) - W^R(E)}{W^R(E)}$$

$\Delta W/W$ varies with energy between 20% and 40% . Since the ramp in the potential energy corresponds to a classical force in the negative direction it is tempting (but incorrect) to conclude that this is the source of the preference for scattering into that direction. In an effort to understand the physical origin of this asymmetry in escape rates we have calculated the the phonon scattering escape rates for a number of other well shapes.

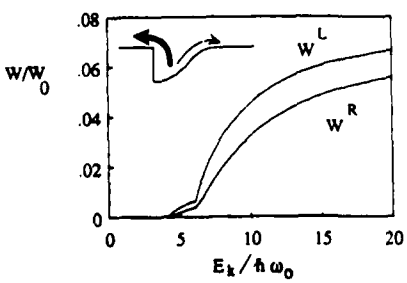


Fig. 4. The normalized escape rates for the potential profile shown in Figure 1c.

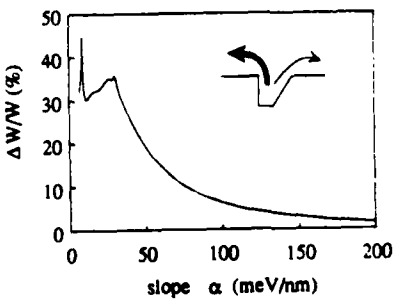


Fig. 5. The percent difference between the escape rates for scattering to the right and left for wells with profiles as shown in Figure 1d. The difference is plotted as a function of the slope of the graded interface.

Figure 1b shows a well with one large potential step on the left and two smaller potential steps on the right, a very crude approximation to a graded junction. We take $V = 0.2$ eV, $d = 150$Å, and $D = 300$ Å, so that a reasonable comparison can be make to the previous linearly graded potential. The potential is reduced by one half on the shallow end of the well. The corresponding scattering rates are shown in Figure 3. Again, a clear preference for final states with current moving from the "graded" interface to the abrupt interface is evident. $\Delta W(E)/W$ is about 5%, somewhat less than the previous potential. That the effect persists at all is evidence that an explanation based on the classical force in the well is inadequate.

Figure 1c illustrates a well with a very soft grading of the potential on one side. The maximum well depth remains 0.2 eV and the width (measured to the point of inflection) is again 300 Å. The two escape rates are shown in Figure 4. The difference in escape rates $\Delta W(E)/W$ is very comparable to the linearly graded well shown in Figure 2. Evidently, the precise form of the grading is not crucially important.

To examine the effect of the steepness of the potential grading on the rate asymmetry we calculate the escape rate for a family of curves shown in Figure 1d. The left wall of the potential well is always abrupt, with a depth of 0.2 eV. The total width of the well is held constant at 300 A. The length of the flat portion of the well d is varied and correspondingly, the slope α of the right side of the well. We want to examine the effect of the changing slope on the difference in the scattering rates. Figure 5 shows $\Delta W/W$, evaluated at $E_k = 10\hbar\omega_0$, as a function of α for this family of potential wells. Note that the asymmetry in the escape rates vanishes smoothly as the slope of the right wall becomes large. In the limit $\alpha \to \infty$, the well becomes a symmetric square well and any asymmetry must vanish. For small values of the slope there are two sharp peaks in $\Delta W/W$. These correspond to resonance conditions which occur when the number of bound states in the well changes from 4 to 5 and from 5 to 6. They are caused by the presence of a continuum resonance, which will become the new bound state, just above the well. A full treatment of these resonance effects will not be presented here.

DISCUSSION

We turn to a discussion of the origin of the asymmetry in the escape rates. Four basic ingredients go into the calculation of the the rates via Equation (1): the initial state wavefunctions, the final state wavefunctions, the perturbing potential, and the density of final states. For the two processes we are considering, scattering into positive and negative current states, the initial state wavefunctions are identical. The perturbing potential, the electron-phonon interaction, is also clearly the same. Both positive and negative current states have the same density of states. The asymmetries in the escape rates are due to differences in the final-state wavefunctions.

Figure 6 shows the absolute square of the final state wavefunctions in the well region for a well as shown in Figure 1d with $d = 60$ Å. The final states have $E_z' = .01$ eV. Such "barely free" states have a dominant contribution to the scattering rates because of the $1/q^2$ suppression of high momentum transfers. The state ψ^L has a considerably larger amplitude in the well region than the state ψ^R. Since the initial bound-state wavefunction is only large appreciable in the well region, it is this region that contributes most to the integral in Equation (5). We define F_L and F_R to be the integral over the well region of the final state probability density.

$$F_L = \int_0^a |\psi^L_{k_z}(z)dz|^2$$

$$F_R = \int_0^a |\psi^R_{k_z}(z)dz|^2$$

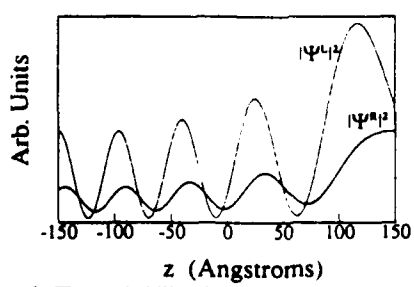


Fig. 6. The probability density for two possible unbound final states. Both states have the same energy but one carries current to the right and the other carries current to the left.

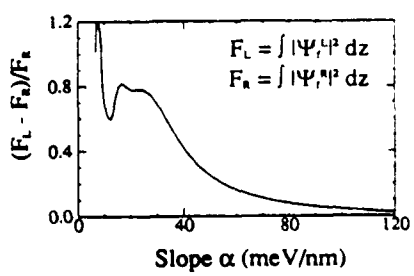


Fig. 7. The difference between the integrated probability densities for final states with current in opposite directions. The difference is plotted as a function of the slope of the graded side of the well (shown in Figure 1d).

Note that these are equal to the effective square matrix elements $|G^{(L/R)}_{n,k'_z}(q_z)|^2$ if we make the (very rough) approximation that $q_z = 0$ and $\psi_n(z)$ =constant. Figure 7 showns the percent difference between F_L and F_R as a function of the slope α for the family of linearly graded potential wells (Figure 1d). These are all evaluated at an E_z of 0.01 eV. Clearly the basic structure of the rate asymmetry shown in Figure 5 has as its origin this difference in final state amplitudes.

The importance of the potential grading on one side of the well now becomes clear and we can offer the following qualitative explanation for the asymmetry in the escape rates. Consider again the potential shown in Figure 1a. A plane wave incident from the left, as in ψ^R, reflects first off of the abrupt interface. That part of the wave which is reflected, contributes nothing to the effective matrix element in Equation (5). There is some additional reflection at the second interface, but it is weak because that interface is graded. By contrast, a plane wave incident from the right, as in ψ^L, is reflected minimally by the first (graded) interface it encounters. Most of the reflection occurs at the second, abrupt interface. Reflection there increases the amplitude of the wavefunction in the well region, and thus its contribution to the matrix element. The total reflection coefficient is, of course, equal for ψ^R and ψ^L. Yet because one is reflected *before* it gets to the well region and one is reflected *over* the well, a difference in amplitudes in the well region results. The matrix element for scattering to the left (into ψ^L) is therefore larger than the matrix element for scattering to the right (into ψ^R). It is this difference which is at the root of the asymmetries in the scattering rates.

CONCLUSIONS

We have reported calculations of the real-space transfer rates for electrons being scattered out of asymmetric quantum wells by polar optical phonon emission and absorption. Including the correct form for the final state wavefunctions (calculated numerically) and the explicitly 2-D to 3-D nature of the scattering yielded two new results. The magnitude of the real-space transfer rate is significantly smaller than the bulk or bound-bound scattering rates. Additionally, we find that asymmetric wells result in carriers scattering preferentially into states which carry current a particular direction which depends on shape of the well. The magnitude of the asymmetry is in the 20 – 40% range for wells with one linearly graded interface and one abrupt interface. We have shown that the asymmetry in the scattering rates is due to an asymmetry in the amplitude of the final state wavefunctions over the well region.

ACKNOWLEDGEMENT

The authors wish to Wolfgang Porod and Alfred Kriman for helpful discussions.

The work was supported by the Air Force Office of Scientific Research under grant number AFOSR-88-0096, and the National Science Foundation under grant number EET87-07628 This work was partially supported by National Science Foundation under grant number ECS890025 through the National Center for Computational Electronics, and utilized the Cray-2 at the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign

REFERENCES

Brennan, Kevin F., and Duke H. Park, J. Appl. Phys. **65**, 1156 (1989).
Hess, K., J. de Physique C7, **42**, 3 (1981).
Kastalsky, A., S. Luryi, A.C. Gossard, and R.H. Hendel, IEEE Electron Device Lett. **EDL-5**, 57 (1984).
Lent, Craig S., Wolfgang Porod, Superlattices and Microstructures **4**, 77 (1988).
Lent, Craig S., Lie Liang and Wolfgang Porod, Appl. Phys. Lett. **54**, 2315 (1989).
Luryi, S., A. Kastalsky, A.C. Gossard, and R.H. Hendel, IEEE Trans. Electron Devices **ED-31**, 832 (1984).
Müller, J. F., S. Schmitt-Rink, and A. F. J. Levi, Appl. Phys. Lett. **52**, 236 (1988).
Müller, J. F., A. F. J. Levi, and S. Schmitt-Rink, Phys. Rev. B **38**, 9843 (1988).
Ridley, B. K., *Quantum Processes in Semiconductors*, (Oxford, New York, 1982).
Porod, Wolfgang, and Craig S. Lent, Solid State Electronics **31**, 359 (1988).

Escape from quantum wells via polar optical phonon scattering

Lie Liang and Craig S. Lent
Department of Electrical and Computer Engineering University of Notre Dame Notre Dame, Indiana 46556

(Received 11 January 1990; accepted for publication 27 April 1990)

We calculate the rate for electrons confined in a quantum well to escape to unbound, current-carrying states by polar optical phonon scattering. We explicitly include the two-dimensional to three-dimensional nature of the transition and choose final state wave functions which are genuinely unbound. We find that the escape rate decreases dramatically as the well depth increases, even for electrons with more than enough energy to surmount the well walls. The real-space transfer rate is much smaller than either the bulk scattering rate or the confined bound-to-bound scattering rate. We explore this phenomenon and give a detailed explanation of the rate depression. We also show that the electron escape rates have a strong directional dependence in asymmetric wells. Electrons scatter out preferentially in the direction of the more abrupt interface.

I. INTRODUCTION

Electrons confined in a quantum well which are heated by an applied field can gain enough energy to escape from the well. The momentum necessary for getting out of the well may be provided by electron-phonon scattering. This real-space transfer effect has been studied by Hess and the others[1] and exploited in some device applications.[2,3] The Monte Carlo technique has been widely used to analyze electron dynamics in such device structures.[4] A necessary input to Monte Carlo calculations is the transition rate for electrons confined in the quantum well to scatter into free states out of the well. Previously this scattering rate has been approximately taken as either the three-dimensional (3D) bulk scattering rate or the two-dimensional (2D) bound-bound scattering rate.[4] The 2D to 3D nature of the transition has been neglected. Recently, we reported a calculation of this real-space transfer rate which fully includes the 2D to 3D character of the transition.[5] Our results show that by choosing the final states as free, current-carrying states explicitly, one obtains real-space transfer rates which are much smaller than either the bulk scattering rate or the 2D confined scattering rate. Here we present the detailed calculation and explanation of this phenomenon and include an additional extension of the calculation to more general quantum well structures which include wells which have an asymmetric potential profile.[6]

Our calculation extends the work of Müller *et al.*,[7,8] who were primarily concerned with impedance-matching conditions which minimize the effect of quantum reflections. They focused on the consequences of effective mass differences in the well and barrier materials, a difference we neglect. In their treatment, the unbound three-dimensional states were approximated by the bound states of a very large well.

Section II below develops the theory of polar optical phonon scattering out of quantum wells. In Sec. III we discuss our results for escape from square-well potentials. Section IV focuses on the directional preference which occurs in scattering out of potentials which are not symmetric. Our conclusions are summarized in Sec. V.

II. THEORY

We calculate polar optical phonon scattering rates using the Fermi Golden rule,

$$W = \frac{2\pi}{\hbar} \int |\langle \psi^f | \mathrm{H}_{e\text{-ph}} | \psi^i \rangle|^2 \delta(E_i - E_f) dS_f, \qquad (1)$$

where the integral is over all final states. The perturbation Hamiltonian describing the coupling of electrons to polar optical phonons is,[9]

$$H_{e\text{-ph}} = \frac{1}{\sqrt{N}} \frac{ee^*}{V_0 \epsilon_0} \sum_q \frac{q}{q^2 + q_0^2} (iQ_q e^{i\mathbf{q}\cdot\mathbf{r}} + \text{c.c.}) \qquad (2)$$

where e is the electron charge, e^* is the magnitude of effective charge on the atoms in the unit cell, V_0 is the volume of a unit cell, N is the number of unit cells, $\mathbf{q}$ is the phonon wave vector, Q_q is the normal coordinate, and c.c. stands for complex conjugate. It is well established that polar optical phonon scattering is the dominant scattering mechanism in III-V materials over a wide temperature range.[12]

We consider electrons which absorb or emit a phonon through the perturbation represented by Eq. (2). In so doing, an electron may transfer from a state which is bound in one direction by a potential well to a state which is completely free. We consider an electron which has made such a transition to have escaped from the well. In order to calculate the rate for an electron to escape from a bound state in the well to unbound states out of the well using Eq. (1), it is necessary to calculate the initial and final eigenstate wave functions. We assume a single-band, spherical, effective-mass model and neglect the difference between the effective masses in the well and boundary materials. The quantum well is in the z direction, extending from $z = 0$ to a. Both initial and final states consist of a product of plane waves in the x and y directions and a function of z, which is the solution of the one-dimensional Schrödinger equation for the effective heterostructure potential.

The initial state wave function is labeled by bound-state index n and a vector $\mathbf{k}$ in x-y plane, and written as

$$\psi_{n,k_x,k_y}(x,y,z) = Ae^{ik_x x + ik_y y} \psi_n(z). \qquad (3)$$

 0021-8979/90/161741-09$03.00 © 1990 American Institute of Physics

The final state wave function is written

$$\psi_{k'_x,k'_y,k'_z}(x,y,z) = Ae^{ik'_x x + ik'_y y}\psi_{k'_z}(z), \tag{4}$$

where $\psi_{k'_z}(z)$ is chosen to be a scattering state which carries current either to the right $(+z)$ or to the left $(-z)$. In both cases A is an overall normalization factor which depends on the current and carrier density. The z-dependent part of the wavefunction for the unbound states which carry current to the right (left) is denoted $\psi^{(R/L)}_{k'_z}(z)$:

$$\psi^R_{k'_z}(z) = \begin{cases} L_z^{-1/2}(e^{ik'_z z} + r_R e^{-ik'_z z}) & \text{if } z<0 \\ u_R(z) & \text{if } 0<z<a, \\ L_z^{-1/2}\ t_R e^{ik'_z z} & \text{if } a<z \end{cases} \tag{5}$$

$$\psi^L_{k'_z}(z) = \begin{cases} L_z^{-1/2}\ t_L e^{-ik'_z z} & \text{if } z<0 \\ u_L(z) & \text{if } 0<z<a. \\ L_z^{-1/2}(e^{-ik'_z z} + r_L e^{ik'_z z}) & \text{if } a<z \end{cases} \tag{6}$$

We normalize the wave function by "box normalizing" the *incoming* plane wave part of the wave function. Here L_z is the (arbitrary) normalization length in which there is unit probability of finding an incoming particle. The final state wave vector, k'_z, is always taken to be positive. The functions $u_R(z)$ and $u_L(z)$, and the reflection and transmission coefficients, r_R, t_R, r_L, and t_L, are all calculated by numerically solving the Schrödinger equation in the well region and matching to the solution outside the well.

The rate for an electron to escape from a bound state to an unbound state with current to the right (R) or left (L) is then given by

$$W^{(R/L)}(k_x,k_y,n) = \frac{2\pi}{\hbar}\int |M^{(R/L)}_{if}|^2 \delta(E_{k'_x,k'_y,k'_z} - E_{k_x,k_y,n} \pm \hbar\omega)\, dS_f, \tag{7}$$

where the square of the matrix element, $|M^{(R/L)}_{if}|^2$, is

$$|M^{(R/L)}_{if}|^2 = \frac{\hbar\omega_0 e^2}{2\epsilon_p V}\frac{1}{q^2} N^{\pm}(\omega_0)|M^{e(R/L)}_{if}|^2. \tag{8}$$

The upper and lower sign are for absorption and emission respectively. The phonon occupation factor is given by

$$N^{\pm}(\omega_0) = \frac{1}{\exp(\hbar\omega_0/k_B T) - 1} + \frac{1}{2} \pm \frac{1}{2},$$

and $1/\epsilon_p \equiv 1/\epsilon_\infty - 1/\epsilon_0$. $M^{e(R/L)}_{if}$, the electronic part of matrix element, is given by

$$|M^{e(R/L)}_{if}|^2 = \frac{2\pi}{L_x}\frac{2\pi}{L_y}\delta(\mathbf{k}_\parallel \pm \mathbf{q}_\parallel - \mathbf{k}'_\parallel) G^{2(R/L)}_{n,k'_z}(q_z). \tag{9}$$

Here the Dirac delta function represents conservation of momentum in the (x,y) plane. The momentum uncertainty perpendicular to the interfaces due to the localization inside the quantum well is represented by the effective matrix element $G^{2(R/L)}_{n,k'_z}(q_z)$, which is given by

$$G^{2(R/L)}_{n,k'_z}(q_z) = \left| \int_{-\infty}^{+\infty} [\psi^{(R/L)}_{k'_z}(z)]^* e^{iq_z z}\psi_n(z)\, dz \right|^2. \tag{10}$$

The integration over the all final states involves the wave vector $\mathbf{q}$ of the absorbed or emitted phonon and final electron wave vector $\mathbf{k}'$, that is

$$dS_f = \frac{L_x}{2\pi}\frac{L_y}{2\pi}\frac{L_z}{2\pi} d^3k' \frac{V}{(2\pi)^3} d^3q \tag{11}$$

The rate for an electron to escape from a bound state labeled by n, with total energy E, into a state with current to the right (R) or left (L) is therefore,

$$W^{(R/L)}_n(E) = \frac{L_z}{(2\pi)^3}\left(\frac{e^2\omega_0}{2\epsilon_p}\right) N^{\pm}(\omega_0)\int_0^\infty dk'_z \int d^3q \frac{1}{q^2} \times G^{2(R/L)}_{n,k'_z}(q_z)\delta(E' - E \mp \hbar\omega_0). \tag{12}$$

Notice that because the final state wavefunction includes the normalizing factor of $L_z^{-1/2}$, the factor of L_z in Eq. (12), which comes from the final density of states, is exactly canceled. As a result, the rate is manifestly independent of this normalizing factor, as it must be.

In order to carry out numerical calculations, the δ function and integral limits in Eq. (12) must be reexpressed. The integral over phonon wave vector $\mathbf{q}$ can be expressed in cylindrical coordinate system, that is $d^3q = q_\parallel dq_\parallel d\theta_q\, dq_z$. Here $\theta_{q_\parallel}$ is the angle between $\mathbf{q}_\parallel$ and $\mathbf{k}_\parallel$. Changing the integral variables k'_z and $q_\parallel$ to $E'_z, E_{q_\parallel}$, which are defined as

$$E'_z = \hbar^2 k'^2_z/2m^*, \tag{13}$$

$$E_{q_\parallel} = \hbar^2 q^2_\parallel/2m^*, \tag{14}$$

and denoting the integral over q_z by

$$\tilde{G}^{2(R/L)}_n(E'_z, E_{q_\parallel}) \equiv \int_{-\infty}^{+\infty} dq_z \frac{(L_z)G^{2(R/L)}_{nk'_z}(q_z)}{2m^* E_{q_\parallel}/\hbar^2 + q_z^2}, \tag{15}$$

we have

$$W^{(R/L)}_n(E) = \frac{1}{(2\pi)^3}\left(\frac{e^2\omega_0}{2\epsilon_p}\right) N^{\pm}(\omega_0)\frac{m^*}{2\hbar^2}\int \frac{dE'_z}{\sqrt{E'_z}} \times \int \frac{dE_{q_\parallel}}{\sqrt{E_{q_\parallel}}} q_\parallel d\theta_{q_\parallel} \tilde{G}^{2(R/L)}_n(E'_z, E_{q_\parallel}) \times \delta(E' - E \mp \hbar\omega_0). \tag{16}$$

We rewrite the delta-function in terms of $\theta_{q_\parallel}, E'_z$,

$$\delta(E' - E \mp \hbar\omega_0) = \delta(\alpha \pm \beta\cos\theta_{q_\parallel}) \tag{17}$$

where

$$\alpha = E'_z - E_n + E_{q_\parallel} \mp \hbar\omega_0, \tag{18}$$

$$\beta = 2\sqrt{(E - E_n)E_{q_\parallel}}. \tag{19}$$

The integral over $\theta_{q_\parallel}$ of the delta function will yield a factor of $2/\sqrt{\beta^2 - \alpha^2}$. Then the scattering rate becomes

$$W^{(R/L)}_n(E) = \frac{1}{(2\pi)^3}\left(\frac{e^2\omega_0}{2\epsilon_p}\right) N^{\pm}(\omega_0)\frac{m^*}{2\hbar^2} \times \int_0^\infty dE_{q_\parallel}\int_{E'_{z0}}^{E'_{z\max}} \frac{dE'_z}{\sqrt{E'_z}} \times \tilde{G}^{2(R/L)}_n(E'_z, E_{q_\parallel})\frac{2}{\sqrt{\beta^2 - \alpha^2}}, \tag{20}$$

where the integral range $(E'_{z0}, E'_{z\max})$ over E'_z is determined by the restrictions

$$-1<\frac{\alpha}{\beta}<1, \quad E'_z>0. \tag{21}$$

Further manipulations yield

$$W_n^{(R/L)}(E) = W_0\left(\frac{m^*(\hbar\omega_0)^{1/2}}{(2\pi\hbar)^2}\right)N^{\pm}(\omega_0) \times\int_0^\infty dE_{q_\parallel}\int_{E'_{z0}}^{E'_{z\max}} dE'_z \times\frac{\tilde{G}_n^{2(R/L)}(E'_z,E_{q_\parallel})}{\sqrt{E'_z}\sqrt{(E'_z-E'_{z\min})(E'_{z\max}-E'_z)}}, \tag{22}$$

where

$$E'_{z\max} = 2\sqrt{(E-E_n)E_{q_\parallel}} + E_n - E_{q_\parallel} \pm \hbar\omega_0, \tag{23}$$

$$E'_{z\min} = -2\sqrt{(E-E_n)E_{q_\parallel}} + E_n - E_{q_\parallel} \pm \hbar\omega_0, \tag{24}$$

$$E'_{z0} = \max(0,E'_{z\min}), \tag{25}$$

and W_0 is the unit of basic rate, which is given by

$$W_0 = 2\alpha_{cp}\omega_0. \tag{26}$$

Here

$$\alpha_{cp} = \frac{e^2}{4\pi\hbar\epsilon_p}\left(\frac{m^*}{2\hbar\omega_0}\right)^{1/2} \tag{27}$$

is the dimensionless electron-phonon coupling constant ($\alpha_p \approx 0.07$, $\omega_0 \approx 5\times10^{13}$ Hz, and $W_0 \approx 7\times10^{12}$ Hz for GaAs).

Our procedure is then as follows. We first solve the one-dimensional effective-mass Schrödinger equation numerically, using an Airy function technique,[13] to obtain the bound-state wavefunctions $\psi_n(z)$ and the final state wave functions $\psi_{k_z}^{(R/L)}(z)$. We then calculate the effective matrix element $G_{n,k_z}^{2(R/L)}(q_z)$ from Eq. (10). The scattering rate is then calculated using Eqs. (15) and (22). Similarly, by taking initial and final states to be both bound we can calculate the bound-bound scattering rate. Throughout we use an effective mass of $m^* = 0.063\ m_0$ and $\hbar\omega_0 = 0.036$ eV for the optical phonon energy. All the calculations are at $T = 300$ K.

III. SQUARE WELL POTENTIALS

We consider first escape of electrons bound in a square well potential. Because of symmetry, the electron escape rates (and effective matrix elements) to final states with current to the right and left are identical and we simply sum the two contributions.

$$W_n(E) = W_n^R(E) + W_n^L(E). \tag{28}$$

A. The weakly bound limit

In the limit of a very shallow quantum well, the scattering rate should approach the rate in bulk material. Since at least one bound state always exists, we calculate the escape rate from a shallow well and the bound-bound scattering which leaves the electron in the well. We examine a well with 300-Å width and 0.005-eV depth for which only one bound state exists. The electron, initially in the bound state, is scattered by emitting or absorbing a polar optical phonon. The final state after scattering may be either a free state out of the well (escape) or bound state in the well with a different parallel momentum (bound-bound scattering). We calculate the total rate, which includes the escape rate and the bound-bound rate using Eq. (10), (15), and (22). The results are shown in Fig. 1. Also plotted (dotted line) is the bulk rate for polar optical phonon scattering given by[9]

$$W(E) = W_0\left(\frac{E}{\hbar\omega_0}\right)^{-1/2}\left[n(q)\sinh^{-1}\left(\frac{E}{\hbar\omega_0}\right)^{1/2} + [n(q)+1]\sinh^{-1}\left(\frac{E}{\hbar\omega_0}-1\right)^{1/2}\right]. \tag{29}$$

Here the first and the second terms in the square brackets are absorption and emission rates, respectively, and it is understood that the second term is zero if $E<\hbar\omega_0$. The rates are shown in unit of basic rate W_0. Figure 1 illustrates that, as expected, the total rate is very close to the bulk rate. This serves as a check on the calculation. Notice that even for this very shallow well, scattering within in the well is almost as likely as scattering out of the well.

B. Depth dependence

We examine the dependence of the escape rate on well depth by considering quantum wells with 300-Å width but various depths from 0.005 to 1.0 eV. For simplicity we always choose initial states which are in the lowest energy bound state of the well. We calculate $W_1(E)$ using Eq. (22). We will denote this rate as simply $W(E)$. The total escape rates (emission and absorption) for this series of wells of different depths are shown in Figs. 2 and 3. These rates are also normalized to W_0 as in Fig. 1.

As these figures illustrate, the escape rates decrease dramatically as the well depth increases. The rate for the well of 0.2-eV depth has already become an order of magnitude smaller than the bulk rate. When the well depth increases to 1.0 eV, the escape rate drops by almost a factor of 100. However, if the escape rate is added to the bound-bound rates

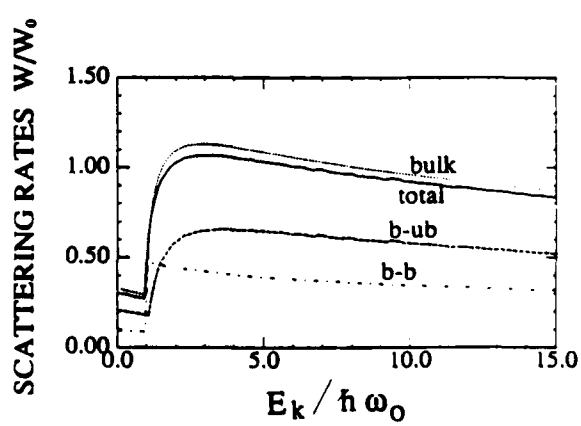


FIG. 1. Optical phonon scattering rates for a very shallow well. The well has a width of 300 Å and a depth of 0.005 eV. Shown are the bound-to-bound rate (dot-dashed curve), the bound-to-unbound (i.e., escape) rate (dashed curve). The total rate, bound-to-bound plus bound-to-unbound, is the solid curve. For comparison the bulk scattering rate is shown with a dotted curve

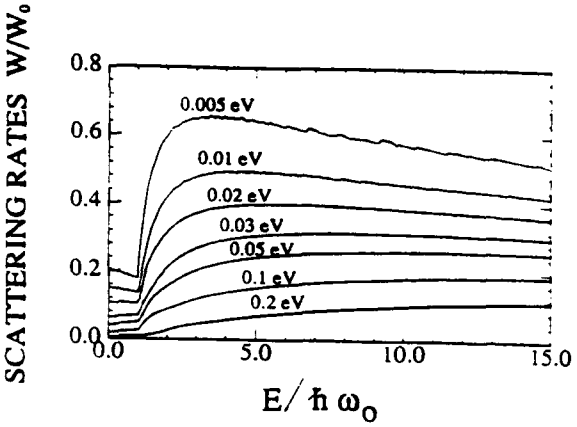


FIG. 2. Escape rates for scattering out of a well by polar optical phonon scattering. The well has a width of 300 Å and the various depths shown. The energy scale is the initial electron energy in units of the optical phonon energy (in this case, 0.036 eV). Initially the electron is in the lowest bound state of the well.

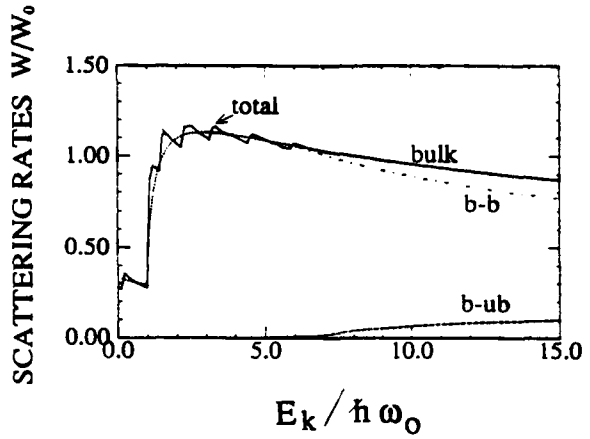


FIG. 4. Optical phonon scattering rate for a deep well. The well is 300-Å wide and 0.2 eV deep. Shown are the bound-to-bound rate (dot-dashed curve), the bound-to-unbound (i.e. escape) rate (dashed curve). The total rate, bound-to-bound plus bound-to-unbound, is the solid curve. For comparison the bulk scattering rate is shown with a dotted curve.

which includes the rates from the lowest bound state to all other bound states, the total rate is still very close to the bulk rate (see Fig. 4). The total rate does exhibit a sawtooth structure due to the separation of bound state energy levels. This feature of the bound-bound rate has been discussed by Ridley and Riddoch.[10,11] From Fig. 4, it is clear that most of the contribution to the total rate comes from the bound-bound transitions.

Note that the escape rate is small even for electrons with more than enough energy to surmount the well walls. An electron energy of $E = 15\hbar\omega_0$ corresponds to 0.54 eV. Figure 3 illustrates that the escape rate for such an electron in a well 0.4 eV deep is still more than an order of magnitude smaller than the bulk or bound-bound scattering rates. We discuss the explanation for this rapid decrease in the escape rate with increasing well depth below.

C. Magnitude of the escape rates

The most obvious feature of the escape rates from quantum wells is that they are so much smaller than either bulk scattering rates or bound-bound scattering. The escape rate is small even when the electron initially has a much larger energy than the well depth. This suppression of the escape rates can be understood as the consequence of two facts: (1) only phonons with momenta larger than a critical value can contribute to escape from the quantum well, and (2) the nature of the electron-phonon coupling results in a $1/q^2$ dependence of the scattering rate on the phonon wave vector q.

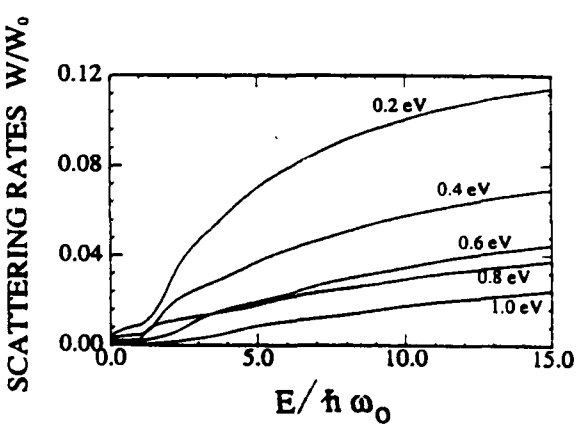


FIG. 3. Escape rates for scattering out of a well by polar optical phonon scattering. The initial state is the lowest bound state of the well. The physical parameters are the same as in Fig. 2 with the well depth increasing as indicated. Some continuum resonance effects are apparent at low energies.

Consider the effective matrix element defined in Eq. (10), for the case of an electron initially in the ground state of the well:

$$G^2_{1,k'_z}(q_z) = \left| \int_{-\infty}^{+\infty} \psi'^*(z) e^{iq_z z} \psi_1(z) dz \right|^2. \tag{30}$$

The initial state $\psi_1(z)$ we can treat approximately as a sine wave with wavelength $2a$ inside the well and neglect the exponential tail which penetrates the well walls (this approximation is for purposes of discussion only — in the calculation we use the exact wavefunction). The final state $\psi^{(R/L)}_{k'_z}(z)$ consists of a combination of reflected and transmitted plane waves. If we consider a final state which is just barely free, κ'_z will be small. This corresponds to a plane wave with a long wavelength outside the well region. However, the contribution to the integral in (30) will be dominated by $\psi_{k'_z}$ evaluated inside the well because the initial-state wave function is almost zero outside the well. In this region the final-state wave function is a combination of plane waves with wave vector

$$\kappa'_z = \sqrt{2m^*(E'_z + V_0)}/\hbar, \tag{31}$$

which corresponds to a much smaller wavelength if V_0 is non-negligible. The integrand in (30) is therefore the product of the slowly varying initial state wave function with oscillating plane waves $e^{\pm i\kappa'_z z}$ and $e^{iq_z z}$. This rapidly oscillating product will make the whole integral nearly vanish unless

$$q_z = \pm \kappa'_z. \tag{32}$$

If the well depth V_0 is large κ'_z is large even for small E'_z, i.e.,

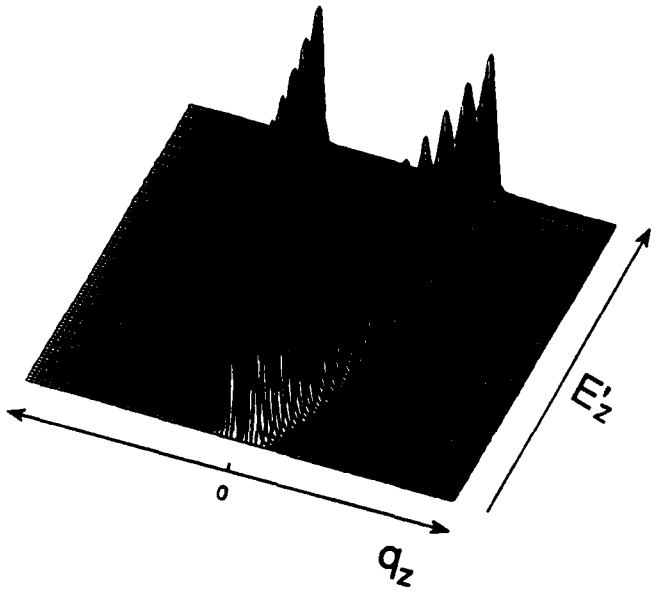


FIG. 5. The effective matrix element $G^2_{1,k'_z}(q_z)$ as a function of q_z and $E'_z = (\hbar k'_z)^2/2m^*$ for the potential well with a depth of 0.005 eV.

for electrons which are barely free.

The above argument is supported by Figs. 5 and 6, in which the effective matrix element $G^2_{1,k'_z}(q_z)$ is plotted as a function of q_z and E'_z for well potentials $V_0 = 0.005$ and 0.8 eV. It is apparent that the matrix element peaks at larger values of q_z in the deeper well.

An electron initially bound in the well has no momentum in the z direction ($\langle \hat{p}_z \rangle = 0$). To be free it needs to acquire only a vanishingly small z component of momentum. Nevertheless, to escape from the well, it must emit or absorb a large-momentum phonon — one with wavevector roughly equal to κ'_z. Because the electron-phonon coupling is

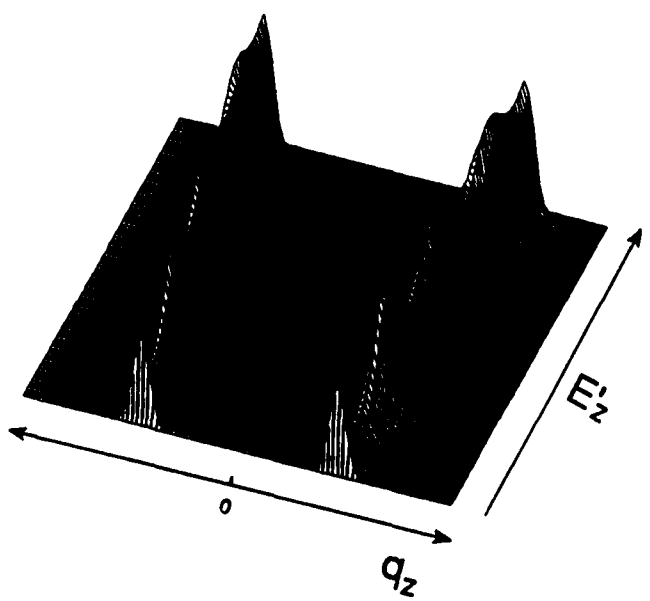


FIG. 6. The effective matrix element $G^2_{1,k'_z}(q_z)$ as a function of q_z and $E'_z = (\hbar k'_z)^2/2m^*$ for the potential well with a depth of 0.8 eV.

inversely proportional to q, as seen in Eq. (2), the golden-rule scattering rate (1) suppresses large-q phonon scattering events. Whereas in bulk scattering the small-q phonons dominate, scattering out of a well relies on large-q phonons and so is correspondingly reduced in magnitude.

D. Energy dependence

Another important feature of the escape rate is the dependence on initial electron energy. There is a significant change in this energy dependence with different well depths. In bulk material the total scattering rate reaches a maximum value when the initial electron energy is around two optical phonon energies. It then drops roughly as $E^{-1/2}$ at higher energies. From Fig. 4, it can be seen that the scattering rates in the shallow quantum well have a similar behavior. There is a transition to qualitatively different behavior as the well depth increase to a value comparable to the optical phonon energy. For wells deep compared with $\hbar\omega_0$, the escape scattering rate increases monotonically with electron energy. The increase is fairly slow after the initial jump in which phonon emission becomes possible.

This change in energy dependence is due to an increase in the number of phonons which can participate in the scattering proces. In bulk scattering the matrix element between initial and final electrons states yield a delta function which enforces momentum conservation. In scattering from a well, the momentum conservation in the z direction is relaxed because of the Heisenberg Uncertainty associated with the localization of the electron in the well. This is reflected in the effective matrix element $G^2_{n,k'_z}(q_z)$ defined in Eq. (10), which would be a delta function if the initial state were a completely delocalized plane wave. Figures 5 and 6 show plots of G^2 for a shallow and deep well. The deep well results in a considerable broadening both as a function of q_z and $E'_z = (\hbar k'_z)^2/2m^*$. This broadening represents the relaxation of momentum conservation in the z direction and results in many more phonons being able to contribute to the scattering into a particular final electron state. The final-state integration includes the integration over all values of q_z and values of E'_z up to a maximum given by energy conservation (Eq. 23). For the shallow well, because G^2 is much more sharply peaked, this contribution increases only linearly as the range of final-state integration is increased. The other factors in Eq. (22) reduce the linear increase to a slow decrease. For the deep well, however, the very broad peaks yield a rapidly increasing contribution to the integral with higher energy (see Figs. 5 and 6). Note that the *magnitude* of the matrix element is much larger for the shallow well than for the deep well, as discussed in the previous section. The broadening of the momentum-conserving peaks in the matrix element explains the change in the energy dependence of the escape scattering rates.

E. Resonance effects

In addition to confined bound states, a quantum well also produces resonances in the continuum of states with energies above the well walls. These resonances are the consequence of interference in the well region due to reflections

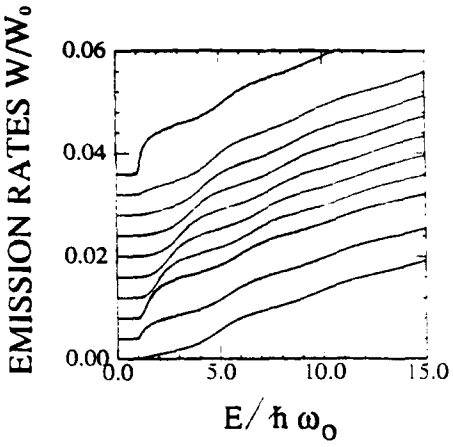


FIG. 7. Escape rates for scattering out of the wells by polar optical phonon emission. The well widths are 300 A and the depths are 0.80 eV (top curve), 0.82, 0.84, 0.86, 0.88, 0.90, 0.92, 0.94, 0.96, and 0.98 eV (bottom curve), respectively. The curves have the zeros offset on the vertical scale for clarity. The enhancement in scattering rates near continuum resonances is clearly evident.

at the well boundaries. In a square well the resonance condition occurs when

$$n\frac{\lambda}{2} \equiv n\,(\pi/\kappa_z') = a, \quad n = 1,2,3,\ldots, \tag{33}$$

where a is the well width, λ is the deBroglie wavelength in the well region, and κ_z' is given by Eq. (31).

The matrix element is influenced by these resonant states and is enhanced at values of E_z' corresponding to this condition. Because the wave function amplitude is enhanced by being resonant with the well, oscillations of the matrix element with genuine maxima and minima are observed.[14,15] These can be seen in the plot of the effective matrix element in Fig. 6. The scattering rates should also exhibit oscillatory structure when additional resonant states are included in the set of accessible final states. For shallow wells, these resonance effects are weak, but as the well becomes deeper, the resonance effects become stronger. The escape rates of electrons scattered from the lowest bound state of a quantum well by polar optical phonon emission are shown in Fig. 7 for a sequence of wells with increasing depths. The wells are all 300-Å wide and have depths from 0.82 to 0.9 eV. We choose rather large well depths to make more obvious the resonance effects. Rate enhancements occur near the resonance condition. Since the maxima and minima of the matrix element are smeared out by the integration of final state over parallel momentum components, the scattering rates exhibit the steplike structures shown rather than true maxima and minima. The energy of the resonant enhancements decreases as the well is deepened, in agreement with Eq. (33).

If the well depth, V_0, is fixed, changing the well width a also changes the energy for resonance with the well. Figure 8 shows the escape rate for scattering from the lowest bound state of the well via polar optical phonon emission for a group of wells with the same depth of 1.0 eV but increasing well width from 100 to 550 Å. The oscillation associated with transitions to continuum resonances is evident.

IV. ASYMMETRIC WELLS

We now consider escape from quantum wells with an asymmetric potential profile. Several asymmetric quantum wells shown in Figs. 9(a)–9(d) are investigated. All of these potential wells have the same total width of $D = 300$ Å and maximum depth of $V = 0.2$ eV. The left side of each well is always an abrupt barrier but the right wall has a different form in each case. We focus on the effect of the asymmetric potentials on the escape rate due to polar optical phonon scattering. Again, for simplicity we always choose initial states which are in the lowest energy bound-state of the well. We calculate $W_1^{(R/L)}(E)$ using Eq. (22). We will denote this rate as simply $W^{(R/L)}(E)$. Of course, the symmetry in the potential means that $W^R(E)$ and $W^L(E)$ need no longer be equal.

Figure 10 depicts the escape rates W^L and W^R for the triangular potential shown in Fig. 9(a). At every energy the scattering rate out of the well to the left is greater than the corresponding rate to the right. Recall that "to the right" here means "into an eigenstate with a well-defined net current in the positive z direction." The asymmetry in the escape rates is in this case quite substantial. At some energies it is greater than 50%, and it is at least 20%. The question

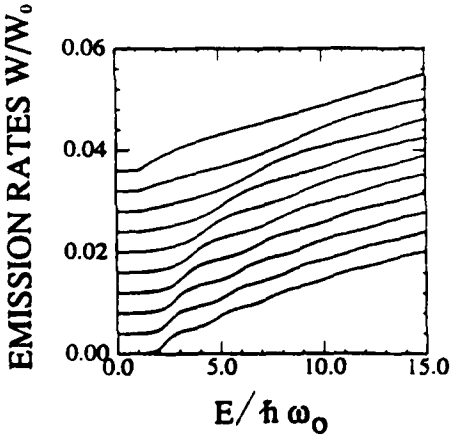


FIG. 8. Escape rates for scattering out of the wells by polar optical phonon emission. The well widths are 100 Å (top curve), 150, 200, 250, 300, 350, 400, 450, 500, and 550 Å (bottom curve), respectively. Each well is 1.0-eV deep. The curves have the zeros offset on the vertical scale for clarity.

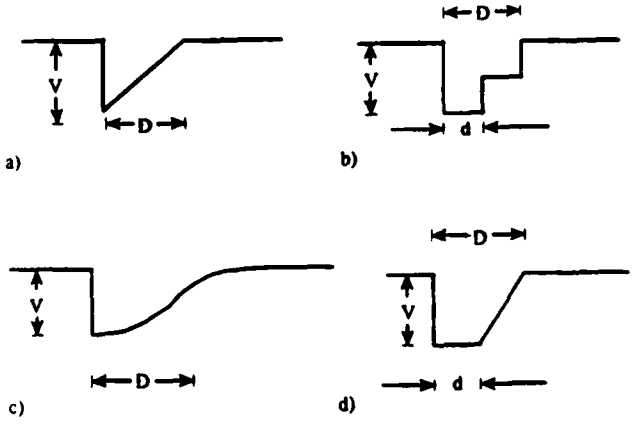


FIG. 9. Potential energy profiles for several asymmetric quantum wells.

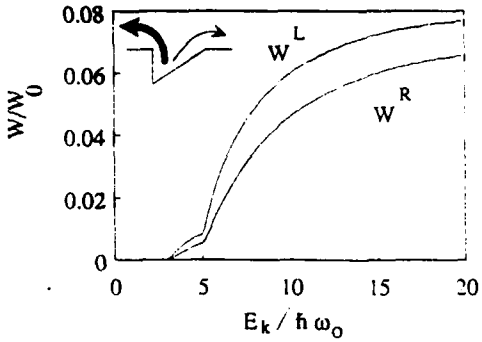


FIG. 10. The normalized escape rates for the potential profile shown in Fig. 9(a). The rates are shown for scattering into final state with current going to the right (W^R) and to the left (W^L).

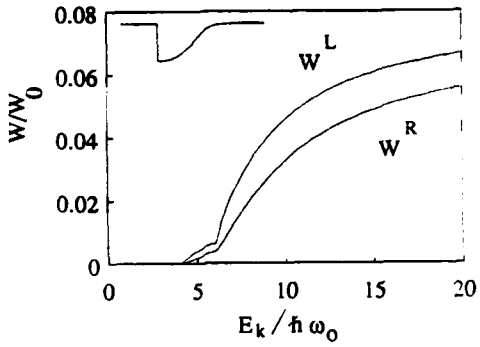


FIG. 12. The normalized escape rates for the potential profile shown in Fig. 9(c).

raised by this result is whether the asymmetry in the escape rate is caused by the geometrical asymmetry in the well shape itself, or by the difference in abruptness of the two walls.

The stepped potential profile shown in Fig. 9(b) shares the asymmetric nature of the triangular profile but retains some of the abruptness of the barriers in the original square well. The corresponding escape rates for this structure are shown in Fig. 11. Clearly the asymmetry in the escape rates, while still present, has been considerably reduced. This suggests that it is the difference in the abruptness of the interfaces which is the key to the preferential scattering out of the well in one direction over the other.

The potential profile in Fig. 9(c) has a very gentle interface on the right edge of the well and an abrupt interface on the left. The gradual interface is modeled by a Fermi-type function,

$$V(z) = 2V_0\{[\exp(-z/D)+1]^{-1}-1\}.$$

The resulting escape rates, shown in Fig. 12, display a large asymmetry. This would seem to confirm the connection between the abruptness of the interfaces and the preferred direction for escape.

In order to investigate this hypothesis, we examine a series of potential wells, shown in Fig. 9(d), with the left wall abrupt and the right wall gradual. The overall width D is kept constant at 300 Å while the width of the well bottom d is varied, thus changing the slope α of the right side of the well. The slope α represents the steepness of the interface. Because wells with different α's have different ground-state energies we calculate the scattering rate as a function of the initial kinetic energy of the electron, E_k. We quantify the asymmetry in the escape rates by defining, somewhat arbitrarily,

$$\frac{\Delta W}{W} = \left.\frac{W^L(E) - W^R(E)}{W^R(E)}\right|_{E=10\hbar\omega_0}. \tag{34}$$

We calculate the escape rates for this family of wells and plot $\Delta W/W$ as a function of the slope α in Fig. 13. After some initial structure (having to do with continuum resonances) $\Delta W/W$ decreases smoothly as the steepness of the right interface increases. As α goes to infinity, the bottom width d approaches the top width D, the well becomes a square well, and any asymmetry must disappear.

We would like to understand the origin of the directional preference illustrated in Fig. 13. The process of scattering out of the well to the right has the same initial state as the process of scattering out to the left. The scattering Hamiltonians are the same for the two processes also. The difference in the rates must be due to the difference in final state wave functions.

We can make the following simple argument for why one would expect the amplitude of the final "left-going"

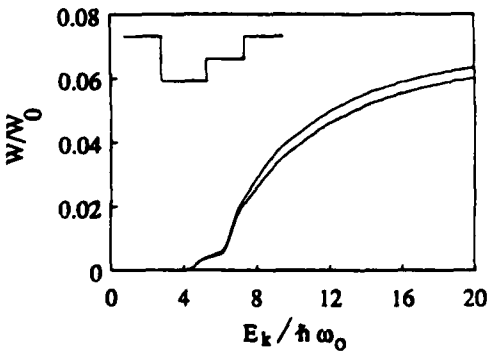


FIG. 11. The normalized escape rates for the potential profile shown in Fig. 9(b).

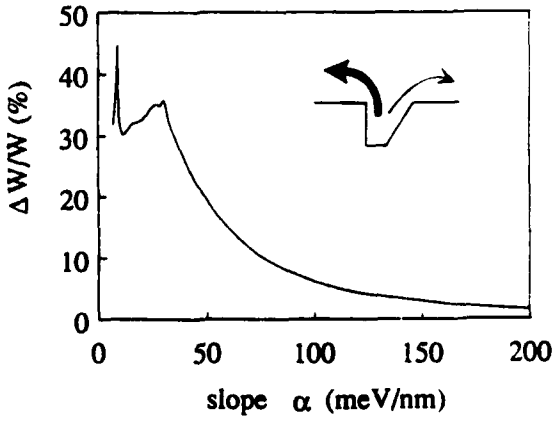


FIG. 13. The percent difference between the escape rates for scattering to the right and left for wells with profiles as shown in Fig. 9(d). The difference is plotted as a function of the slope of the graded interface. The total well width is 300 Å, and the maximum depth is 0.2 eV.

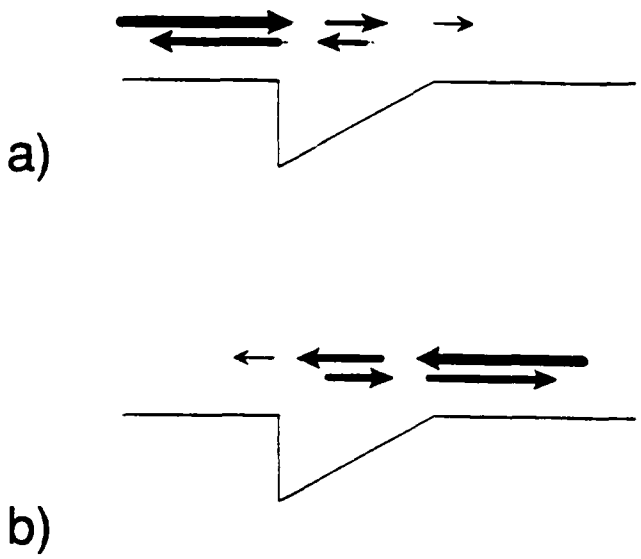


FIG. 14. Schematic illustration of the origin of the directional dependence of the escape rates. The final, unbound state is illustrated for the case of net current in the positive direction (a), and net current in the negative direction (b). Because of the stronger reflections at the abrupt interface, the final-state amplitude in the well region is greater in (b). This enhances the matrix element for scattering into states with current to the left.

wave function to be larger than the amplitude of the final "right-going" wave function in the well region. The final state wave functions are defined by Eqs. (5) and (6). ψ^L can be thought of as resulting from a plane wave incident from the right on the well structure. The wave first hits the gentle, graded interface which reflects only a small amount of the incoming wave. Most of the wave is reflected at the abrupt interface on the left side of the well. This results in a fairly large amplitude in the well region itself. By contrast, ψ^R consists of a wave incident from the left, which is strongly reflected by the abrupt interface which it encounters first. Thus, a smaller amplitude is expected over the well region. This is illustrated in Fig. 14. In each case, the total reflected amplitude is the same (one can show rigorously that $T_L = T_R$ and $R_L = R_R$[16]). The difference is whether the incoming wave is reflected before or after it reaches the inside of the well.

In Fig. 15, we show an example of the absolute square of

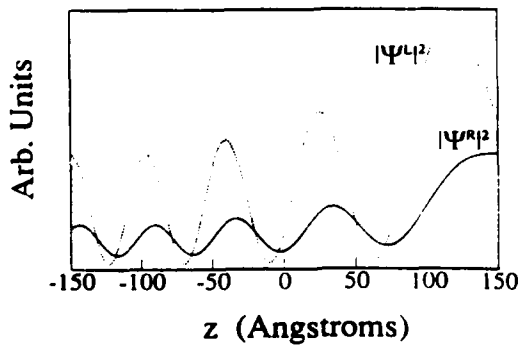


FIG. 15. The absolute square of wavefunctions for two possible unbound final states for the potential shown in Fig. 9(d). Both states have the same energy but one carries current to the right and the other carries current to the left.

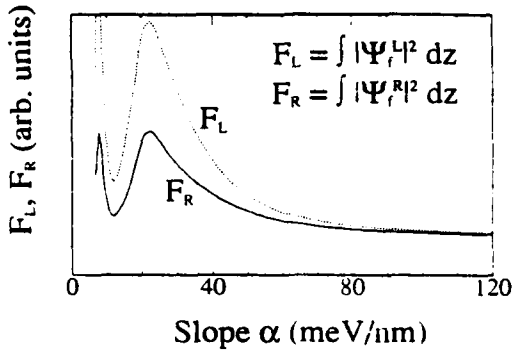


FIG. 16. The integral over the well region of probability density as functions of the slope α of the graded interface.

two final states, ψ^R and ψ^L, in the well region for the well shown in Fig. 9(d) with $d = 60$ Å and an energy of 0.01 eV. We can see that ψ^L indeed has a much larger amplitude than ψ^R.

Because the wavefunction of the initial bound state decays very fast outside the well region, the integral over the well region dominates the matrix element. The rate asymmetry can therefore be characterized by the difference in the final state amplitude over the well region.

To explore further the effect of the well asymmetry on final state wavefunction ψ^L and ψ^R, we define F_L and F_R as

$$F_L = \int_0^a |\psi_{k_z}^L(z)dz|^2 \tag{35}$$

and

$$F_R = \int_0^a |\psi_{k_z}^R(z)dz|^2. \tag{36}$$

These correspond to the effective matrix elements [Eq. (10)] in the rather crude approximation that the initial state wavefunction is a constant in the well region and zero elsewhere, and that the phonon momentum q_z is zero. We calculate F_L and F_R for the family of the wells depicted in Fig. 9(d). These wavefunctions are all evaluated at energy $E_z' = 0.01$ eV. Figure 16 shows F_L and F_R as a function of α, the slope of the graded interface. Figure 17 shows the relative difference between F_L and F_R as a function of α. The clear similarity between Figs. 17 and 13 indicates that, in-

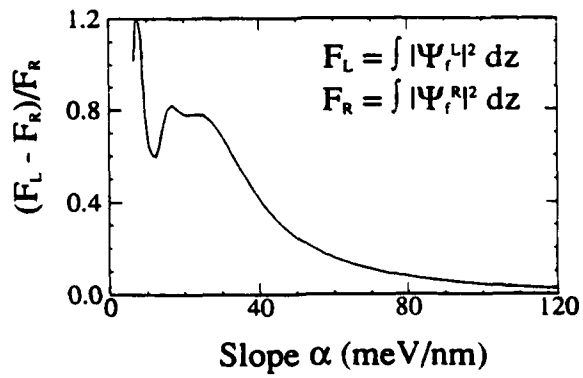


FIG. 17. The relative difference between the integrated probability densities for final states with current in opposite directions as a function of the slope of the graded interface of the well shown in Fig. 9(d).

deed, the asymmetry in scattering rates can be explained as a final-state amplitude effect. This makes it clear why the electrons are scattered preferentially in the direction of the more abrupt well barrier.

V. CONCLUSION

We have calculated the real-space transfer rate at which electrons are scattered out of a quantum well by absorbing or emitting a polar optical phonon. We use genuinely free, three-dimensional final states with a well-defined current. Our results indicate that the magnitude of this escape rate is much smaller than the three-dimensional bulk rate or the two-dimensional bound-bound scattering rate. The rate depression is primarily due to the large phonon momenta required in these processes. We also find that in asymmetric wells, electrons are scattered out of the wells preferentially into states which carry current in the direction from the less abrupt interface to the more abrupt interface. We show that it is the asymmetry in the amplitudes of final state wavefunctions that is the origin of the directional dependence in the escape scattering.

ACKNOWLEDGMENTS

The authors wish to thank Wolfgang Porod and Alfred Kriman for helpful discussions. This work was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-88-0096, and the National Science Foundation under Grant No. EET87-07628. This work was partially supported by the National Science Foundation under Grant No. ECS890025 through the National Center for Computational Electronics, and utilized the Cray-2 at the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign.

[1] K. Hess, J. Phys. (Paris) **42**, 3 (1981).
[2] S. Luryi, A. Kastalsky, A. C. Gossard, and R. H. Hendel, IEEE Trans. Electron Devices **ED-31**, 832 (1984).
[3] A. Kastalsky, S. Luryi, A. C. Gossard, and R. H. Hendel, IEEE Electron Device Lett. **EDL-5**, 57 (1984).
[4] K. F. Brennan, and Duke H. Park, J. Appl. Phys. **65**, 1156 (1989).
[5] Craig S. Lent, Lie Liang, and Wolfgang Porod, Appl. Phys. Lett. **54**, 2315 (1989).
[6] See also Craig Lent and Lie Liang (unpublished).
[7] J. F. Müller, A. F. J. Levi, and S. Schmitt-Rink, Phys. Rev. B **38**, 9843 (1988).
[8] J. F. Müller, S. Schmitt-Rink, and A. F. J. Levi, Appl. Phys. Lett. **52**, 236 (1988).
[9] B. K. Ridley, *Quantum Processes in Semiconductors* (Oxford, New York, 1982).
[10] B. K. Ridley, J. Phys. C **15**, 5899 (1982).
[11] F. A. Riddoch and B. K. Ridley, J. Phys. C **16**, 6971 (1983).
[12] E. M. Conwell, "High Field Transport in Semiconductors", in *Solid State Physics, Advances in Research and Applications*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1967), Vol 9, p. 156.
[13] Wayne W. Lui and Masao Fukuma, J. Appl. Phys. **60**, 1555 (1986).
[14] Wolfgang Porod and Craig S. Lent, Solid-State Electron. **31**, 359 (1988).
[15] Craig S. Lent and Wolfgang Porod, Superlattices and Microstructures **4**, 77 (1988).
[16] D. D. Coon and H. C. Liu, Appl. Phys. Lett. **47**, 172 (1985).

INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING, VOL. 29, 1527–1537 (1990)

THE NUMERICAL SIMULATION OF ELECTRON TRANSMISSION THROUGH A TWO-DIMENSIONAL QUANTUM DEVICE BY THE FINITE ELEMENT METHOD

DAVID J. KIRKNER,* CRAIG S. LENT† AND SRINIVAS SIVAPRAKASAM†

University of Notre Dame, Notre Dame, IN 46556 U.S.A.

SUMMARY

In this paper we present a method for numerically solving the Schrödinger equation for the problem of electron transmission through a quantum device defined on a two-dimensional domain. An important aspect of our formulation is the treatment of the boundary conditions at the contact–device interfaces, allowing the problem to be discretized on the device domain only. With the FEM aproach that we employ, the nature of the potential field does not effect the computational expense (for smoothly varying fields). We examine a sample problem of a symmetric cavity in a wave-guide structure.

INTRODUCTION

Techniques for numerical simulation of electronic devices governed by semiclassical transport theory are well developed.[1] However, a great deal of contemporary research is directed toward the development of a class of semiconductor devices fabricated at a scale where quantum effects dominate.[2] In this regime the wave-like nature of the electron must be accounted for, and simulations must be based on solutions of the time-independent Schrödinger's equation:

$$\frac{-\hbar^2}{2m^*}\nabla^2\psi(x, y, z) + V(x, y, z)\psi(x, y, z) = E\psi(x, y, z)$$

The complex wavefunction ψ specifies the state of the electron and $|\psi|^2$ is interpreted as a probability density.[3] The Schrödinger equation describes an electron with an effective mass m^* which moves in a potential V.

The solutions to the Schrödinger equation are of two types, bound or unbound, depending on whether the electron is localized in a particular region of space (bound), or delocalized with a non-zero probability density arbitrarily far away from the structure (unbound). This distinction is reflected in the boundary conditions for the Schrödinger equation. Bound electrons are described by a wavefunction which vanishes at infinite distances. The wavefunction for an unbound electron, by contrast, does not vanish at infinite distances but typically assumes the form of a travelling wave which carries current. The unbound states are sometimes called scattering states because they correspond to an electron beam impinging on a potential and being

* Department of Civil Engineering
† Department of Electrical and Computer Engineering

0029–5981/90/071527–11$05.50
© 1990 by John Wiley & Sons, Ltd.

Received 15 March 1989
Revised 13 November 1989

scattered outward. For the bound-state problem, the Schrödinger equation becomes an eigenvalue equation and the numerical solution scheme is fairly straightforward. Both bound and unbound electron states occur in quantum devices. The bound states represent charge trapped in the device and the unbound states represent the current passing through the device. We focus here on the unbound states both because they have the key role of carrying current in devices and because bound-state solutions present fewer numerical difficulties.

Semiconductor device fabrication technology has until recently been limited to creating structures with features that are small in one spatial dimension. This has nevertheless produced an explosion of new device ideas and interest in electronic behaviour of quantum-scale structures. This development of ultra-small devices has been greatly facilitated by the insight provided by solutions to the one-dimensional Schrödinger equation. The solution for the case of a simple trapezoidal-shaped potential subject to incident plane waves is easily found in terms of Airy functions. If a potential barrier of arbitrary shape is approximated as piecewise linear, the Airy function solution along with a transfer matrix approach can be used to find an approximate solution very efficiently.[4] Transmission coefficients and currents are then readily determined. This solution approach, however, is limited to one-dimensional problems.

Recent advances in fabrication and crystal growth techniques[5,6] are now making feasible the creation of structures quantized in two and three spatial dimensions. This has created the need for more sophisticated modelling capabilities, both to understand device behaviour and to guide the invention of new devices. For the current carrying (unbound) states, the use of standard numerical solution procedures, such as finite differences or finite elements which are commonly employed in problems of higher spatial dimensions, is hampered by the extended nature of the wavefuntion.[7,8] This is a reflection of the fact that a quantum device is an open system, i.e. a problem with at least some of the boundary conditions at infinity. In this paper we give particular attention to the treatment of the conditions at the interface of the device and the contacts. Boundary conditions are developed which account properly for reflection and transmission of waves. Thus, we are able to cast the problem as a boundary value problem on the domain described by the boundary of the device. We then develop a finite element solution strategy for the problem of electron transmission through a potential of arbitrary shape. We give results for the case of a resonant cavity, and examine the behaviour of the transmission coefficient versus energy curve as we vary the device geometry.

PROBLEM STATEMENT

In this paper we consider the transmission of electrons through a quantum structure defined by a two-dimensional domain, as shown schematically in Figure 1. We call Ω the device domain and Ω^L and Ω^R the left and right contacts respectively. For simplicity in presentation both contacts have width d and lie on the x_1 axis. The analysis following is not restricted to this geometry, however.

Since Ω^L extends to $x_1 = -\infty$ and Ω^R extends to $x_1 = +\infty$ we have a so-called open problem. The form of the complex wavefunction is known in Ω^L and Ω^R, however, and thus we will reduce the problem to a boundary value problem on the device domain Ω. Following the determination of the wavefunction on Ω, the transmission coefficient can be calculated at the device–right contact interface.

Since we anticipate a finite element method (FEM) solution we find it convenient to consider the complex wavefunction as a real-valued vector quantity, $\psi = (\psi_1, \psi_2)^T$, where ψ_1 and ψ_2 are the real and imaginary parts of the wavefunction respectively. We first give the classic or strong

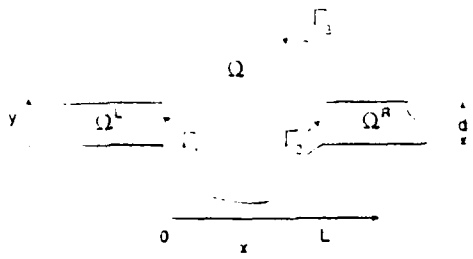


Figure 1. Geometry of the transmission problem. Within the region $\Omega \cup \Omega^L \Omega^R$ the potential V is finite. Outside this region the potential is infinite

statement of the governing equations: Find $\psi \in C^2(\Omega)$, $\psi^L \in C^2(\Omega^L)$, $\psi^R \in C^2(\Omega^R)$ such that

$$-\nabla^2\psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}) = E\psi(\mathbf{x}), \quad \mathbf{x} \in \Omega$$

$$-\nabla^2\psi^L(\mathbf{x}) = E\psi^L(\mathbf{X}), \quad \mathbf{x} \in \Omega^L$$

$$-\nabla^2\psi^R(\mathbf{x}) = E\psi^R(\mathbf{x}), \quad \mathbf{x} \in \Omega^R$$

$$\left.\begin{aligned} \psi &= \psi^L \\ \partial_{x1}\psi &= \partial_{x1}\psi^L \end{aligned}\right\} \quad \text{on} \quad \Gamma_1 = \{\mathbf{x}: x_1 = 0, 0 \leqslant x_2 \leqslant d\}$$

$$\left.\begin{aligned} \psi &= \psi^R \\ \partial_{x1}\psi &= \partial_{x1}\psi^R \end{aligned}\right\} \quad \text{on} \quad \Gamma_2 = \{\mathbf{x}: x_1 = L, 0 \leqslant x_2 \leqslant d\}$$

$$\psi = 0 \quad \text{on} \quad \Gamma_3 = \partial\Omega - (\Gamma_1 + \Gamma_2)$$

$$\psi^L = 0 \quad \text{on} \quad x_2 = 0, x_2 = d, x_1 \leqslant 0$$

$$\psi^R = 0 \quad \text{on} \quad x_2 = 0, x_2 = d, x_1 \geqslant L$$

$$\psi^L, \quad \text{bounded as} \quad x_1 \to -\infty$$

$$\psi^R, \quad \text{bounded as} \quad x_2 \to +\infty$$

In the above ψ^L, ψ^R are the complex wavefunctions in Ω^L and Ω^R respectively, $\mathbf{x} = (x_1, x_2)^T$ is a vector containing the Cartesian co-ordinates of a point, $V(\mathbf{x})$ is the real-valued potential assumed to be a continuous, bounded function on Ω, E is the total energy and $\partial\Omega$ is the boundary of Ω. Also, for convenience the factor $\hbar^2/2m^*$ is subsumed in the definitions of $V(\mathbf{x})$ and E.

The solution of Schrödinger's equation in the contacts satisfying the boundedness conditions above consists of travelling harmonic waves and standing, exponentially decaying waves. These latter modes are herein referred to as evanescent modes. We assume a single incoming mode from $x_1 = -\infty$ and no incoming waves from $x_1 = +\infty$. Again, this latter assumption is primarily for simplicity in presentation and the analysis is easily extended to allow multiple mode input from either direction. The solution of Schrödinger's equation in the contacts satisfying the conditions above has the following form:

$$\psi^L(\mathbf{x}) = \mathbf{R}_m(x_1)\mathbf{A}_m \sin\left(\frac{m\pi x_2}{d}\right) + \sum_{n=1}^{N} \mathbf{R}_n^T(x_1)\mathbf{b}_n \sin\left(\frac{n\pi x_2}{d}\right) + \sum_{n=N+1}^{\infty} \mathbf{b}_n e^{k_n x_1} \sin\left(\frac{n\pi x_2}{d}\right) \quad (1)$$

$$\psi^R(\mathbf{x}) = \sum_{n=1}^{N} \mathbf{R}_n(x_1)\mathbf{a}_n \sin\left(\frac{n\pi x_2}{d}\right) + \sum_{n=N+1}^{\infty} \mathbf{a}_n e^{-k_n x_1} \sin\left(\frac{n\pi x_2}{d}\right) \quad (2)$$

where

$$\mathbf{R}_n(x_1) = \begin{bmatrix} \cos(k_n x_1) & -\sin(k_n x_1) \\ \sin(k_n x_1) & \cos(k_n x_1) \end{bmatrix} \tag{3}$$

$$k_n = \sqrt{\left|E - \left(\frac{n\pi}{d}\right)^2\right|} \tag{4}$$

$$\mathbf{a}_n = (a_{n1}, a_{n2})^{\mathrm{T}} \tag{5}$$

$$\mathbf{b}_n = (b_{n1}, b_{n2})^{\mathrm{T}} \tag{6}$$

$$\mathbf{A}_m = (A_{m1}, A_{m2})^{\mathrm{T}} \tag{7}$$

The solutions are combinations of travelling waves which are oscillatory in x_2 and evanescent modes which decay exponentially away from the device region. We choose to define the wavevector k_n so that it is a real quantity for both oscillatory and exponential waves. The number of modes which represent travelling waves is given by

$$N = \max n, \text{ such that } E > \left(\frac{n\pi}{d}\right)^2. \tag{8}$$

Thus the total energy as well as the contact geometry determines the number of current carrying modes reflected and transmitted. As the energy increases, more modes become travelling modes which carry current.

The vector $\mathbf{A}_m$ represents the amplitude of the incoming mode. Specifying $\mathbf{A}_m$ is equivalent to specifying the amplitude and complex phase of the incoming wave.

We now use the continuity of the wavefunctions at the interfaces Γ_1, Γ_2 and the orthogonality of the sine functions to express the constants, $\mathbf{b}_n$, $\mathbf{a}_n$ in terms of the wavefunction ψ on Γ_1 and Γ_2, i.e.

$$\mathbf{b}_n = \frac{2}{d}\int_0^d \psi(0, x_2)\sin\left(\frac{n\pi x_2}{d}\right)\mathrm{d}x_2 - \mathbf{A}_m\delta_{mn} \tag{9}$$

$$\mathbf{a}_n = \frac{2}{d}\int_0^d \mathbf{R}_n^{\mathrm{T}}(L)\psi(L, x_2)\sin\left(\frac{n\pi x_2}{d}\right)\mathrm{d}x_2, \quad n \leqslant N \tag{10}$$

$$\mathbf{a}_n = \frac{2\mathrm{e}^{k_n L}}{d}\int_0^d \psi(L, x_2)\sin\left(\frac{n\pi x_2}{d}\right)\mathrm{d}x_2, \quad n \geqslant N \tag{11}$$

where δ_{mn} is the Kronecker delta and we have used $\mathbf{R}_n(0) = I$, the identity matrix. Taking the derivative of equations (1) and (3) with respect to x_1 and using the continuity of the normal derivative of ψ at the interfaces yields

$$\partial_{x1}\psi(0, x_2) = -k_m\mathbf{Q}_m(0)\mathbf{A}_m\sin\left(\frac{m\pi x_2}{d}\right) - \sum_{n=1}^{N} k_n\mathbf{Q}_n(0)^{\mathrm{T}}\mathbf{b}_n\sin\left(\frac{n\pi x_2}{d}\right) + \sum_{n=N+1}^{\infty} k_n\mathbf{b}_n\sin\left(\frac{n\pi x_2}{d}\right) \tag{12}$$

$$\partial_{x1}\psi(L, x_2) = -\sum_{n=1}^{N} k_n\mathbf{Q}_n(L)^{\mathrm{T}}\mathbf{a}_n\sin\left(\frac{n\pi x_2}{d}\right) - \sum_{n=N+1}^{\infty} k_n\mathbf{a}_n\mathrm{e}^{-k_n L}\sin\left(\frac{n\pi x_2}{d}\right) \tag{13}$$

where

$$\mathbf{Q}_n(x_1) = \begin{bmatrix} \sin(k_n x_1) & \cos(k_n x_1) \\ -\cos(k_n x_1) & \sin(k_n x_1) \end{bmatrix}$$

Now, using equations (9) to (11) in equations (12) and (13) allows us to write formally,

$$\partial_{x1}\psi(0, x_2) = \mathbf{f}^{\mathrm{L}}(\psi(0, x_2), x_2) \tag{14}$$

$$\partial_{x1}\psi(L, x_1) = \mathbf{f}^{\mathrm{R}}(\psi(L, x_2), x_2) \tag{15}$$

where

$$\begin{aligned} \mathbf{f}^{\mathrm{L}}(\psi(0, x_2), x_2) = &- k_m \mathbf{Q}_m(0)\mathbf{A}_m \sin\left(\frac{m\pi x_2}{d}\right) \\ &- \sum_{n=1}^{N} k_n \mathbf{Q}_n(0)^{\mathrm{T}} \left[\frac{2}{d}\int_0^d \psi(0, x_2)\sin\left(\frac{n\pi x_2}{d}\right)\mathrm{d}x_2 - \mathbf{A}_m\delta_{mn}\right]\sin\left(\frac{n\pi x_2}{d}\right) \\ &+ \sum_{n=N+1}^{\infty} k_n \left[\frac{2}{d}\int_0^d \psi(0, x_2)\sin\left(\frac{n\pi x_2}{d}\right)\mathrm{d}x_2\right]\sin\left(\frac{n\pi x_2}{d}\right) \end{aligned} \tag{16}$$

and

$$\begin{aligned} \mathbf{f}^{\mathrm{R}}(\psi(L, x_2), x_2) = &- \sum_{n=1}^{N} k_n \mathbf{Q}_n(L)^{\mathrm{T}} \left[\frac{2}{d}\int_0^d \mathbf{R}_n^{\mathrm{T}}(L)\psi(L, x_2)\sin\left(\frac{n\pi x_2}{d}\right)\mathrm{d}x_2\right]\sin\left(\frac{n\pi x_2}{d}\right) \\ &- \sum_{n=N+1}^{\infty} k_n \left[\frac{2\mathrm{e}^{k_n L}}{d}\int_0^d \psi(L, x_2)\sin\left(\frac{n\pi x_2}{d}\right)\mathrm{d}x_2\right]\mathrm{e}^{-k_n L}\sin\left(\frac{n\pi x_2}{d}\right) \end{aligned} \tag{17}$$

Equations (14) and (15) are the boundary conditions at the device–contact interfaces which allow us to restate the boundary value problem on the device domain only. Note that these boundary conditions are a generalization of the usual mixed or third-type boundary condition relating the primary unknown to its normal derivative on the boundary. The normal derivative of ψ is not directly related to ψ, but rather to a superposition of Fourier components of ψ. The boundary condition has a non-local character in that to apply it at any one point on the boundary requires a knowledge of the wavefunction over the entire boundary.

We can now restate the boundary value problem on the domain Ω: Find $\psi \in C^2(\Omega)$ such that

$$-\nabla^2\psi(\mathbf{x}) + V(\mathbf{x})\psi(\mathbf{x}) = E\psi(\mathbf{x}),\ \mathbf{x} \in \Omega$$

$$\psi = 0 \quad \text{on} \quad \Gamma_3$$

and such that the boundary conditions given by equations (14) and (15) are satisfied.

Prior to stating a variational equivalent of the above problem we define some notation. Let

$$H_3^1(\Omega) = \left\{v: \int_\Omega \left[\left(\frac{\partial v}{\partial x_1}\right)^2 + \left(\frac{\partial v}{\partial x_2}\right)^2 + (V - E)v^2\right]\mathrm{d}A < \infty,\ v = 0 \text{ on } \Gamma_3\right\}$$

and we say $\psi \in (H_3^1(\Omega))^2$ if $\psi_1, \psi_2 \in H_3^1(\Omega)$. Also, let

$$\nabla\psi = \begin{bmatrix} \partial/\partial x_1 & 0 \\ 0 & \partial/\partial x_1 \\ \partial/\partial x_2 & 0 \\ 0 & \partial/\partial x_2 \end{bmatrix} \begin{Bmatrix} \psi_1 \\ \psi_2 \end{Bmatrix}$$

We now give a weak variational problem equivalent to the above problem: Find $\psi \in (H_3^1(\Omega))^2$ such that

$$(\nabla\bar{\psi}, \nabla\psi) + (\bar{\psi}, (V - E)\psi) = (\bar{\psi}, \mathbf{f}^R)_{\Gamma_2} - (\bar{\psi}, \mathbf{f}^L)_{\Gamma_1}, \ \forall \bar{\psi} \in (H_3^1(\Omega))^2 \tag{18}$$

where

$$(\nabla\bar{\psi}, \nabla\psi) = \int_\Omega \nabla\bar{\psi}^T \nabla\psi \, dA$$

$$(\bar{\psi}, (V - E)\psi) = \int_\Omega (V - E)\bar{\psi}^T \psi \, dA$$

$$(\bar{\psi}, \mathbf{f}^R)_{\Gamma_2} = \int_0^d \bar{\psi}(L, x_2)^T \, \mathbf{f}^R(\psi(L, x_2), x_2) \, dx_2$$

$$(\bar{\psi}, \mathbf{f}^L)_{\Gamma_1} = \int_0^d \bar{\psi}(0, x_2)^T \, \mathbf{f}^L(\psi(0, x_2), x_2) \, dx_2$$

The equivalence of the two problems is shown by integrating the gradient term, $(\nabla\bar{\psi}, \nabla\psi)$, by parts and then using the fact that $\bar{\psi}$ is an arbitrary element of $(H_3^1(\Omega))^2$.

FEM SOLUTION ALGORITHM

Given a regular discretization of Ω, we express the FEM approximation to ψ on this mesh as

$$\psi(\mathbf{x}) = \mathbf{N}(\mathbf{x})\mathbf{u}$$

where

$$\mathbf{u} = (\psi_1(\mathbf{x}_1), \psi_2(\mathbf{x}_1), \psi_1(\mathbf{x}_2), \psi^2(\mathbf{x}_2), \ldots, \psi_1(\mathbf{x}_M), \psi_2(\mathbf{x}_M))^T$$

$\mathbf{N}(x)$ is the $2 \times 2M$ matrix of global shape functions,

$$\mathbf{N}(\mathbf{x}) = \begin{bmatrix} \phi_1(\mathbf{x}), & 0, & \phi_2(\mathbf{x}), & 0, & \ldots, & \phi_M(\mathbf{x}), & 0 \\ 0, & \phi_1(\mathbf{x}), & 0, & \phi_2(\mathbf{x}), & \ldots, & 0, & \phi_M(\mathbf{x}) \end{bmatrix}$$

$\mathbf{x}_j$ means the co-ordinates of the jth nodal point, $\phi_i(\mathbf{x})$ is the global shape function associated with node i and M is the number of nodes in $\Omega \cup \Gamma_1 \cup \Gamma_2$.
We can express the gradient of ψ as

$$\nabla\psi = \mathbf{B}(\mathbf{x})\mathbf{u},$$

where $\mathbf{B}(\mathbf{x})$ is a matrix containing derivatives of the shape functions. Thus the left hand side of the variational equation can be written

$$(\nabla\bar{\psi}, \nabla\psi) + (\bar{\psi}, (V - E)\psi) = \bar{\mathbf{u}}^T \left(\int_\Omega (\mathbf{B}^T\mathbf{B} + (V - E)\mathbf{N}^T\mathbf{N}) dA \right) \mathbf{u}$$

$$\equiv \bar{\mathbf{u}}^T \mathbf{A} \mathbf{u} \tag{19}$$

This relation defines the 'stiffness' matrix A. We now give a fairly detailed treatment of the boundary contributions. Assume that there are M_1 nodes on Γ_1 and M_2 nodes on Γ_2. Let $\mathbf{u}_i$ be the vector of length $2M_i$ which is the restriction of $\mathbf{u}$ to Γ_i, and let $\mathbf{N}_i(x_2)$ be the $2 \times 2M_i$ matrix of shape functions such that

$$\psi(0, x_2) = \mathbf{N}_1(x_2)\mathbf{u}_1 \tag{20}$$

$$\psi(L, x_2) = \mathbf{N}_2(x_2)\mathbf{u}_2 \tag{21}$$

Using the FEM approximation to ψ with equation (10), (11) and (13) allows us to write

$$(\bar{\psi}, \mathbf{f}^R)_{\Gamma_2} = -\bar{\mathbf{u}}_2^T\left(\sum_{n=1}^{N} \frac{2k_n}{d}\mathbf{N}_{2n}^T\mathbf{Q}_n(L)\mathbf{R}_n(L)^T\mathbf{N}_{2n} + \sum_{n=N+1}^{I} \frac{2k_n}{d}\mathbf{N}_{2n}^T\mathbf{N}_{2n}\right)\mathbf{u}_2$$

where

$$\mathbf{N}_{2n} = \int_0^d \mathbf{N}_2(x_2)\sin\left(\frac{n\pi x_2}{d}\right)dx_2$$

Note that the total number of modes is truncated at a finite number, I. The required value of I must be determined experimentally. In general we have found that six or seven modes above N are sufficient. Now note that

$$\mathbf{Q}_n(x_1)\mathbf{R}_n^T(x_1) = \mathbf{D} \equiv \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

therefore,

$$\begin{aligned}(\bar{\psi}, \mathbf{f}^R)_{\Gamma_2} &= -\bar{\mathbf{u}}_2^T\left(\sum_{n=1}^{N} \frac{2k_n}{d}\mathbf{N}_{2n}^T\mathbf{D}\mathbf{N}_{2n} + \sum_{n=N+1}^{I} \frac{2k_n}{d}\mathbf{N}_{2n}^T\mathbf{N}_{2n}\right)\mathbf{u}_2 \\ &\equiv -\bar{\mathbf{u}}_2^T\mathbf{C}_2\mathbf{u}_2\end{aligned}$$

Likewise,

$$\begin{aligned}(\bar{\psi}, \mathbf{f}^L)_{\Gamma_1} &= \bar{\mathbf{u}}_1^T\left(k_m\mathbf{N}_{1m}^T\mathbf{D}^T\mathbf{A}_m + \left(\sum_{n=1}^{N} \frac{2k_n}{d}\mathbf{N}_{1n}^T\mathbf{D}\mathbf{N}_{1n} + \sum_{n=N+1}^{I} \frac{2k_n}{d}\mathbf{N}_{1n}^T\mathbf{N}_{1n}\right)\mathbf{u}_1\right) \\ &\equiv \bar{\mathbf{u}}_1^T(\mathbf{p}_m + \mathbf{C}_1\mathbf{u}_1),\end{aligned}$$

We now define $2M \times 2M$ matrices $\hat{\mathbf{C}}_1$, $\hat{\mathbf{C}}_2$ and a vector $\hat{\mathbf{p}}_m$ of length M such that

$$\bar{\mathbf{u}}_i^T\mathbf{C}_i\mathbf{u}_i = \bar{\mathbf{u}}^T\hat{\mathbf{C}}_i\mathbf{u} \tag{22}$$

$$\bar{\mathbf{u}}_i^T\mathbf{p}_m = \bar{\mathbf{u}}^T\hat{\mathbf{p}}_m \tag{23}$$

Thus the discrete form of the variational problem is

$$\bar{\mathbf{u}}^T\mathbf{A}\mathbf{u} = \bar{\mathbf{u}}^T(\hat{\mathbf{p}}_m - (\hat{\mathbf{C}}_1 + \hat{\mathbf{C}}_2)\mathbf{u})$$

or since the elements of $\bar{\mathbf{u}}$ are arbitrary,

$$(\mathbf{A} + \hat{\mathbf{C}}_1 + \hat{\mathbf{C}}_2)\mathbf{u} = \hat{\mathbf{p}}_m \tag{24}$$

Note that the matrices $\hat{\mathbf{C}}_1$ and $\hat{\mathbf{C}}_2$ are not explicitly formed; equations (22) and (23) simply represent the assembly process. Also note that the real and imaginary parts of the wavefunction are coupled only through the boundary conditions, equations (14) and (15), which, in the FEM procedure are effected by the matrices $\mathbf{C}_1$ and $\mathbf{C}_2$. These matrices are not symmetric and thus the final system to be solved is unsymmetric.

TRANSMISSION COEFFICIENT

The current in Ω^L or Ω^R due to electron transmission is (ignoring a multiplicative constant which does not effect the transmission coefficient)

$$J = \int_0^d (\psi^*\partial_{x1}\psi - (\partial_{x1}\psi)^*\psi)dx_2$$

where ψ is the complex wave function, $\psi = \psi_1 + i\psi_2$, and ψ^* is the complex conjugate of ψ. Using equation (3) we can express the current in the region Ω^R as follows:

$$J^R = \sum_{n=1}^{N} k_n \mathbf{a}_n^T \mathbf{a}_n$$

Note that the evanescent modes in the contacts do not contribute to the current. If we calculate the incoming current (i.e. ignoring the reflected modes) in Ω^L we get from equation (1)

$$J_0 = k_m \mathbf{A}_m^T \mathbf{A}_m$$

Therefore the transmission coefficient is

$$T(E) = \frac{J^R}{J_0} = \frac{1}{J_0} \sum_{n=1}^{N} k_n \mathbf{a}_n^T \mathbf{a}_n$$

where $T(E)$ indicates the energy E is used as a parameter. Using equations (10) and (21) we can express $T(E)$ in terms of the nodal values of ψ on the interface Γ_2, i.e.

$$T(E) = \left(\frac{4}{J_0 d^2}\right) \mathbf{u}_2^T \left(\sum_{n=1}^{N} k_n \mathbf{N}_{2n}^T \mathbf{N}_{2n}\right) \mathbf{u}_2 \tag{25}$$

Thus after the wavefunction is determined on the device domain, i.e. after equation (24) is solved, the transmission coefficient is readily computed from equation (25).

Note that it only requires a modification of the vector $\hat{\mathbf{p}}$ to include a superposition of incoming modes, which involves only minor computation.

SAMPLE PROBLEM

To illustrate our solution procedure we will analyse the resonant cavity shown in Figure 2. Although inclusion of an arbitrary potential in the device is an easy matter, we present the resonant cavity results simply to restrict the number of parameters. We assume only mode one is incoming from the left and we take $A_{11} = 1$, $A_{12} = 0$. For all simulations the effective mass, m^*, is taken to be $0{\cdot}05\, m_0$. Furthermore, grids for each simulation were generated using square, bilinear elements, 5 Å on a side. We found all results were sufficiently converged using this mesh for graphical purposes.

Figure 3 shows transmission coefficient versus energy curves for a family of devices. For all cases the width of the contacts, d, and the length of the device, L, are taken to be 100 Å. The transmission coefficient curves shown are for device widths, D, varying from 200 to 250 Å. The minima in these curves correspond to bound states in the cavity. For this simple device geometry the device bound-state energies are easily determined. At these energies the electron is trapped in the device and this explains the resulting reduction in the value of the transmission coefficient. As the device width, D, is increased we expect the energy levels of the bound states in the cavity to decrease. We see this effect exhibited by a shift to the left of the minima. In general, the more separated the bound state energies the broader the band over which full transmission, or close to full transmission, occurs.

It should be noted that the occurrence of *minima* at energies of bound states is a unique feature of the two-dimensional problem. In one dimension, it is more typical to see transmission *maxima* associated with bound-state, or more precisely, quasibound-state resonances. The characteristic sharp minima and broad peaks have been noted by Sols and co-workers[8] in their work on the stub-tuner transistor.

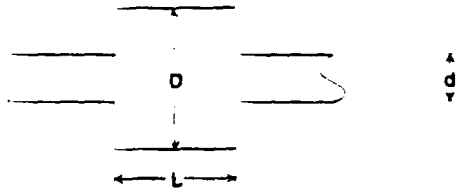


Figure 2. Geometry for the sample problem of the resonant cavity

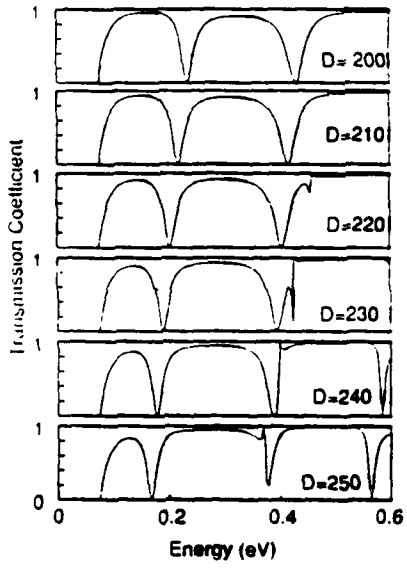


Figure 3. The transmission coefficient versus energy for the resonant cavity shown in Figure 2

Figure 4 shows a plot of the real or the imaginary part of the wavefunction, for the 200 Å wide cavity at an energy of 0·43 eV, which corresponds to the first excited bound state of the cavity. For energies corresponding to the bound states of the cavity the real and the imaginary parts of the wavefunction are identical in shape and simply scale; thus we show only one plot for this case. The wavefuntion at this energy is essentially a standing wave corresponding to the bound-state eigenfunction.

It is interesting to note the shape of the wavefunction at the exit boundary. This tells us the form of the wavefunction in the contact. Since there is basically no transmission at 0·43 eV, the solution should contain almost no participation of the travelling modes in the contact. For this simulation the first evanescent mode is the third mode (since $N = 1$ and the problem is symmetric; and the third mode is precisely the shape which is predominant at this boundary. Thus the wavefunction is essentially an evanescent mode in the contact. Higher modes have little participation. Although we infer the shape of the wavefunction in the contact from its shape at the exit boundary, we could easily determine the explicit expression for ψ in the contact since the amplitudes of the wave forms are easily computed as shown in equation (25). However, we are primarily interested in its behaviour in the device and at the contact interface.

Figures 5 and 6 show the real and imaginary parts respectively of the wavefunction at an energy of 0·31 eV. This energy corresponds to almost full transmission ($T(E) = 0{\cdot}90$). Note that the

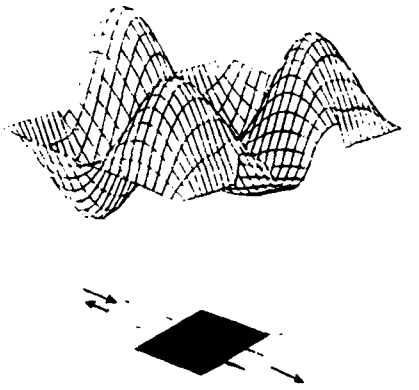


Figure 4. The wavefuntion for transmission through the resonant cavity with $D = 200$ Å at an energy of 0·43 eV

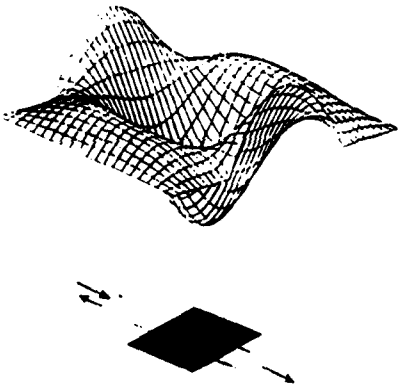


Figure 5. The real part of the wavefunction for transmission through the resonant cavity with $D = 200$ Å at an energy of 0·31 eV

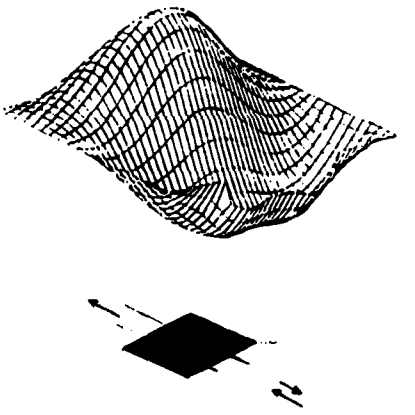


Figure 6. The imaginary part of the wavefunction for transmission through the resonant cavity with $D = 200$ Å at an energy of 0·31 eV

wavefunctions are basicaly travelling waves in the cavity. We also see in this case that the shape of mode one, a travelling wave, is predominant at the boundary. Thus in the contact we have a mode one travelling wave with little effect of the evanescent modes close to the device for this case of near full transmission. The role of the evanescent modes in the contacts becomes important in multiple devices if the devices are 'close' together.

CONCLUSION

In this paper we have presented a method for numerically solving the problem of electron transmission through a two-dimensional quantum device. The boundary conditions at the conact–device interfaces given by equations (14)–(17) are important in that they allow the problem to be discretized on the device domain only. With the FEM approach that we employ, the nature of the potential field does not affect the computational expense (for smoothly varying fields). Also we can incorporate more contacts and multiple mode inputs with little additional expense. Such studies are presently underway. An extension of our formulation to allow an efficient analysis of multiple devices is also being studied.

This work was supported by the Air Force Office of Scientific Research under grant number AFOSR-88-0096.

REFERENCES

1. S. Selberherr, *Analysis and Simulation of Semiconductor Devices*, Springer-Verlag, Wien, 1984.
2. H. L. Grubin (ed.), *The Physics of Submicron Structures*, Plenum, New York, 1984.
3. Eugene Merzbacher, *Quantum Mechanics*, Wiley, New York, 1970.
4. Wayne W. Lui and Masao Fukuma, 'Exact solution of the Schrödinger equation across an arbitrary one-dimensional piecewise-linear potential barrier', *J. Appl. Phys.*, **60**, 1555 (1986).
5. J. M. Gaines, P. M. Petroff, H. Kroemer, R. J. Simes, R. S. Geels and J. H. English, 'Molecular-beam epitaxy growth of tilted GaAs/AlAs superlattices by deposition of fractional monolayers on vicinal (001) substrates', *J. Vac. Sci. Technol.*, **B6**, 1378 (1988).
6. G. Bernstein and D. K. Ferry, 'Negative differential conductivity in lateral surface superlattices', *J. Vac. Sci. Technol.*, **B5**, 964 (1987).
7. The finite element method has been employed in atomic scattering problems. For a review of the chemical literature see J. Linderberg, 'Finite elements in quantum mechanics', *Comp. Phys. Reports*, **6**, 209 (1987).
8. A Green's function technique has recently been proposed. See Fernando Sols, M. Macucci, U. Ravaioli and Karl Hess, 'On the possibility of transistor action based on quantum interference phenomena', *Appl. Phys. Lett.*, **54**, 350 (1989).

The quantum transmitting boundary method

Craig S. Lent
Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556

David J. Kirkner
Department of Civil Engineering, University of Notre Dame, Notre Dame, Indiana 46556

(Received 13 December 1989; accepted for publication 30 January 1990)

A numerical algorithm for the solution of the two-dimensional effective mass Schrödinger equation for current-carrying states is developed. Boundary conditions appropriate for such states are developed and a solution algorithm constructed that is based on the finite element method. The utility of the technique is illustrated by solving problems relevant to submicron semiconductor quantum device structures.

I. INTRODUCTION

Recent advances in submicron lithography have made it possible to create very small structures in which electrons in the two-dimensional electron gas formed at a semiconductor heterostructure interface are further confined by metal gates. A negative voltage applied to the gate depletes electrons in the region beneath the gate. Several device structures have been proposed and realized using this scheme.[1-5] For sufficiently small scales, electrons in such structures are ballistic and coherent over the entire device region. For many purposes, electronic transport in this regime is governed by the effective-mass Schrödinger equation. Understanding and modeling the performance of these device structures requires solving the two-dimensional Schrödinger equation.

Several well-established techniques have been employed for obtaining numerical solutions of the two-dimensional Schrödinger equation for bound states. Bound-state solutions, however, are not of primary interest in understanding quantum devices. States which carry current, sometimes called "scattering states" are more important in understanding current flow through small coherent regions. In one dimension, the current carrying states can be solved for rather straightforwardly because the Schrödinger equation can be simply integrated from one side of the device to the other. In two dimensions the problem is more difficult.

Current-carrying states are composed of solutions in a device region (the scattering region) and solutions that extend to infinity along the input and output leads. Because of the regularity of the potential in the leads, the form of the solution is known in the lead regions. This can usually be expressed as a sum of traveling and evanescent modes in the leads. The complex amplitudes of each of these modes is unknown, however. The ampitudes are a result of the problem's solutions, commonly expressed in terms of transmission and reflection coefficients. Therefore, although the asymptotic form of the solution may be known, the values of the wave function and its derivative on a boundary region are only known in terms of coefficients which are to be determined. Numerical solution requires knowledge of the wave function or its derivative on some boundary region. In one dimension, this boundary value problem can be treated as an initial value problem using Numerov-type methods. In two dimensions it cannot.

We formulate the boundary conditions appropriate for the numerical solution of the Schrödinger equation for current-carrying states on a two-dimensional domain. We employ these boundary conditions in developing a numerical solution algorithm based on the finite element method. We call this the quantum transmitting boundary method (QTBM). In Ref. 6, we formulated the technique for the rather restricted case of two colinear leads with infinite square-well cross sections and identical widths. Here we develop the completely general formulation, capable of solving the general problem of the transmission and reflection properties of a region with an arbitrary potential, including multiple multimoded input leads of arbitrary cross section and orientation. The QTBM yields not only the transmission and reflection coefficients, but the full wave function in the device region.

Section II contains the formal statement of the problem and the development of the boundary conditions for each lead. This section is quite general and does not depend on any particular numerical technique. The numerical algorithm is developed in Sec. III. Section III C contains a description of the basic features of the QTBM algorithm and a comparison with other techniques. In Sec. IV we present a few short examples of the application of the QTBM to some interesting geometries for quantum structures. Section V states our conclusions.

II. FORMULATION OF THE BOUNDARY CONDITIONS

A. Problem statement

The region of interest is partitioned into a "device" region Ω_0, and several lead regions, $\Omega_1, \Omega_2, \ldots, \Omega_n$, which extend to infinity. We want to solve the two-dimensional effective-mass Schrödinger equation on $\Omega \equiv \Omega_0 \cup \Omega_1 \cup \Omega_2 \ldots \cup \Omega_n$. The boundary of the region Ω_0 we denote Γ. The boundary between a lead region, Ω_i, and the device region, Ω_0, we call Γ_i. This lead boundary can, without loss of generality, be taken to be a straight line. The rest of the boundary Γ, which is not a lead boundary, is denoted Γ_0 (see Fig. 1).

The problem we wish to solve can be stated as follows:

Given: The total energy E, the potential energy in each region, $V_i(x,y)$, $i = 0,1,2\ldots,n$, and the complex amplitudes of each incoming mode in each lead,

 0021-8979/90/106353-07$03.00 © 1990 American Institute of Physics

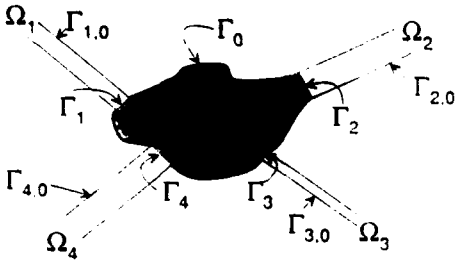


FIG. 1. The problem geometry.

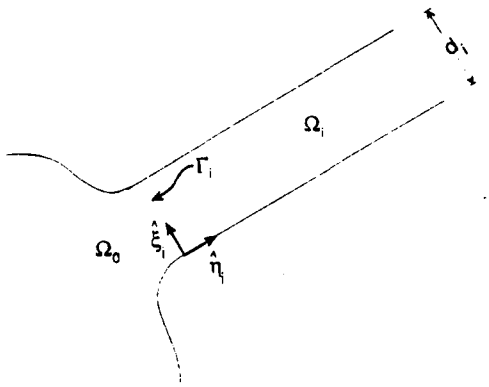


FIG. 2. Local coordinate system for lead i.

Find:

$$\psi_0 \epsilon C^2(\Omega_0), \psi_1 \epsilon C^2(\Omega_1), \ldots, \text{ and } \psi_n \in C^2(\Omega_n),$$

such that

$$-(\hbar^2/2m^*)\nabla^2\psi_i(x,y) + V_i(x,y)\psi_i(x,y) = E\psi_i(x,y), \quad (x,y)\in\Omega_i, \tag{1}$$

and

$$\psi_0 = \psi_i \text{ on } \Gamma_i, \tag{2}$$

$$\nabla\psi_0 \cdot \hat{n}_{\Gamma_i} = \nabla\psi_i \cdot \hat{n}_{\Gamma_i} \text{ on } \Gamma_i, \tag{3}$$

$$\psi_0 = 0 \text{ on } \Gamma_0 \equiv \partial\Omega_0 - \sum_i \Gamma_i, \tag{4}$$

$$\psi_i = 0 \text{ on } \Gamma_{i0} \equiv \partial\Omega_i - \Gamma_i, \tag{5}$$

$$\psi_i \text{ bounded as } \sqrt{x^2+y^2} \to \infty. \tag{6}$$

We will develop a solution algorithm valid for any potential $V_0(x,y)$ in the device region. The condition that ψ_0 vanish on the boundary Γ_0 need not be restrictive in that the boundary can always be placed far enough from the active device region that the wave function has vanishingly small amplitude on the boundary. The device region must, in this sense, contain the wave function within Ω_0 except along the lead boundaries.

We require that the potential $V_i(x,y)$ in each lead be independent of the distance along the lead. This is what we mean by leads—channels that contain no longitudinal structure, although they may have a complicated cross section. In the next section we consider the form of the solutions in the leads themselves. Our goal is to formulate a boundary condition on the lead boundaries Γ_i that will allow us to specify the incoming flux in each lead and solve the Schrödinger equation only in the device region Ω_0.

B. Solution in the leads

By requiring the leads to be uniform along their lengths we can separate the problem in the lead regions into coupled one-dimensional problems. We define a local coordinate system in each lead i,

$$(\eta_i, \xi_i) = [\eta_i(x,y), \xi_i(x,y)], \tag{7}$$

such that $\hat{\eta}_i$ is parallel to the lead walls and points down the lead in the direction away from Ω_0 (see Fig. 2). The $\hat{\xi}$ direction is perpendicular to $\hat{\eta}$ and the boundary between Ω_i and Ω_0 can be written

$$\Gamma_i = \{(\eta_i, \xi_i) | \xi_i \epsilon (0, d_i), \eta_i = 0\}. \tag{8}$$

Here d_i is the perpendicular width of the lead. The potential in the lead is assumed to vary only across the lead:

$$V_i(\eta_i, \xi_i) = V_i(\xi_i). \tag{9}$$

For the simple case of square-well leads, $V_i(\xi_i) = 0$.

Because the potential in the leads is independent of η, the Schrödinger equation in the leads separates and we can write the general solution in lead i as

$$\psi_i(\eta_i, \xi_i) = \sum_{m=1}^{N^i} a^i_m \chi^i_m(\xi_i) e^{-ik^i_m \eta_i} + b^i_m \chi^i_m(\xi_i) e^{ik^i_m \eta_i} + \sum_{m=N^i+1}^{\infty} b^i_m \chi^i_m(\xi_i) e^{-k^i_m \eta_i}, \tag{10}$$

where $\chi^i_m(\xi_i)$ is the mth eigenstate of the one-dimensional Schrödinger equation,

$$-(\hbar^2/2m^*)[\partial^2\chi^i_m(\xi_i)/\partial\xi_i^2] + V_i(\xi_i)\chi^i_m(\xi_i) = E^i_m \chi^i_m(\xi_i). \tag{11}$$

The a^i_m's are the coefficients of the incoming traveling-wave states and are an input to the problem. N^i is the number of traveling-wave modes allowed in lead i. The b^i_m's in the first sum are the coefficients of outgoing traveling-wave states. The b^i_m's in the second sum are coefficients of evanescent (exponentially decaying) states. The coefficients b^i_m are unknowns that must be calculated as part of the problem solution. We can choose the eigenstates χ to be orthonormal:

$$\int_0^{d_i} [\chi^i_m(\xi_i)]^* \chi_n(\xi_i) d\xi_i = \delta_{m,n}. \tag{12}$$

For infinite square-well leads,

$$\chi^i_m(\xi_i) = \sqrt{2/d_i}\, \sin[(m\pi/d)\xi_i]. \tag{13}$$

The wave vector for the mth mode in the ith lead is given by

$$k^i_m = \sqrt{|(2m^*/\hbar^2)(E - E^i_m)|}. \tag{14}$$

The number of traveling waves N^i is the maximum m such that $E > E^i_m$.

C. Boundary conditions at the lead interfaces

At the boundary between the device region and lead i we require both continuity of the wave function and the normal derivative [Eqs. (2) and (3)]. The condition on the deriva-

tive can be written

$$\nabla\psi_0(\mathbf{r})\cdot\hat{n}_{\Gamma_i} = \nabla\psi_i(\mathbf{r})\cdot\hat{n}_{\Gamma_i} \text{ on } \Gamma_i$$

$$= \frac{\partial}{\partial\eta_i}\psi_i(\eta_i,\xi_i)\bigg|_{\eta_i=0}. \tag{15}$$

We calculate the normal derivative from the known form of the wave function given by Eq. (10):

$$\frac{\partial\psi_i}{\partial\eta_i}\bigg|_{\eta_i=0} = \sum_{m=1}^{N'} -ia^i_m k^i_m \chi^i_m(\xi_i) + ib^i_m k^i_m \chi^i_m(\xi_i)$$

$$- \sum_{m=N'+1}^{\infty} k^i_m b^i_m \chi^i_m(\xi_i). \tag{16}$$

We can evaluate the b^i_m's by using the orthogonality of the χ's at $\eta_i = 0$:

$$b^i_m = \int_0^{d_i} \chi^i_m(\xi_i)\psi_i(\eta_i = 0,\xi_i)d\xi_i - a_m. \tag{17}$$

So

$$\frac{\partial\psi_i}{\partial\eta_i}\bigg|_{\eta_i=0} = \sum_{m=1}^{N} ik^i_m \chi^i_m(\xi_i)$$

$$\times\left(-2a^i_m + \int_0^{d_i} \chi^i_m(\xi_i)\psi_i(\eta_i = 0,\xi_i)d\xi_i\right)$$

$$- \sum_{m=N+1}^{\infty} k^i_m \chi^i_m(\xi_i)$$

$$\times\left(\int_0^{d_i} \chi^i_m(\xi_i)\psi_i(\eta_i = 0,\xi_i)d\xi_i\right). \tag{18}$$

The derivative boundary condition then becomes

$$\nabla\psi_0(\mathbf{r})\cdot\hat{n}_{\Gamma_0} = f_i[\xi_i,\psi_i(\eta_i = 0,\xi_i)]$$

$$= \sum_{m=1}^{N'} ik^i_m \chi^i_m(\xi_i)$$

$$\times\left(-2a^i_m + \int_0^{d_i} \chi^i_m(\xi_i)\psi_i(\eta_i = 0,\xi_i)d\xi_i\right)$$

$$- \sum_{m=N'+1}^{\infty} k^i_m \chi^i_m(\xi_i)$$

$$\times\left(\int_0^{d_i} \chi^i_m(\xi_i)\psi_i(\eta_i = 0,\xi_i)d\xi_i\right). \tag{19}$$

We can now employ Eq. (2), the continuity of the wave function itself, and replace ψ_i's on the right-hand side by ψ_0's since they are evaluated only on the boundary. We then obtain a boundary condition on ψ_0 and its normal derivative:

$$\nabla\psi_0(\mathbf{r})\cdot\hat{n}_{\Gamma_0} = f_i[\xi_i,\psi_0(\eta_i = 0,\xi_i)]$$

$$= \sum_{m=1}^{N'} ik^i_m \chi^i_m(\xi_i)$$

$$\times\left(-2a^i_m + \int_0^{d_i} \chi^i_m(\xi_i)\psi_0(\eta_i = 0,\xi_i)d\xi_i\right)$$

$$- \sum_{m=N'+1}^{\infty} k^i_m \chi^i_m(\xi_i)$$

$$\times\left(\int_0^{d_i} \chi^i_m(\xi_i)\psi_0(\eta_i = 0,\xi_i)d\xi_i\right). \tag{20}$$

Equation (20) defines the functional f_i and is the boundary condition we require in order to formulate the problem for the current-carrying states as a boundary-value problem. Note that Eq. (20) relates the value of the wave function's normal derivative at a particular point to the values of the wave function at all the other points along the boundary.

D. Weak variational form of the Schrödinger equation

In this section we develop the weak variational form[7] of the Schrödinger equation in a way suitable for numerical discretization. Our goal is to discretize the wave function only in the device region Ω_0, and apply suitable boundary conditions to match wave function and derivative to the solution in the leads. We begin with the time-independent Schrödinger equation for the wave function ψ_0 in the device region:

$$-(\hbar^2/2m^*)\nabla^2\psi_0(x,y) + V(x,y)\psi_0(x,y) = E\psi_0(x,y). \tag{21}$$

This is multiplied by an arbitrary test function $\bar{\psi}$ and integrated over Ω_0:

$$-\frac{\hbar^2}{2m^*}\int_{\Omega_0} \bar{\psi}\nabla^2\psi_0 d^2r + \int_{\Omega_0} \bar{\psi}(V-E)\psi_0 d^2r = 0. \tag{22}$$

The test function is chosen so that it obeys the same essential boundary conditions as does ψ_0, i.e., $\bar{\psi} = 0$ on Γ_0.

Using Green's first identity we have,

$$\int_\Omega \phi\nabla^2\psi\, d^2r = -\int_\Omega \nabla\phi\cdot\nabla\psi\, d^2r + \oint_\Gamma \phi(\nabla\psi\cdot\hat{n}_\Gamma)d\Gamma, \tag{23}$$

so the Schrödinger equation becomes

$$\frac{\hbar^2}{2m^*}\int_{\Omega_0} \nabla\bar{\psi}\cdot\nabla\psi_0\, d^2r + \int_{\Omega_0} \bar{\psi}(V-E)\psi_0\, d^2r$$

$$= \frac{\hbar^2}{2m^*}\oint_\Gamma \bar{\psi}(\nabla\psi_0\cdot\hat{n})d\Gamma. \tag{24}$$

The integration around the boundary Γ must be performed in a counterclockwise direction. Since $\bar{\psi}$ is zero on Γ_0, the right-hand side can be rewritten as an explicit sum over these contact regions:

$$\frac{\hbar^2}{2m^*}\int_{\Omega_0} \nabla\bar{\psi}\cdot\nabla\psi_0\, d^2r + \int_{\Omega_0} \bar{\psi}(V-E)\psi_0\, d^2r$$

$$= \frac{\hbar^2}{2m^*}\sum_i \int_{\Gamma_i} \bar{\psi}(\nabla\psi_i\cdot\hat{n})d\Gamma_i. \tag{25}$$

E. The reformulated problem

The original problem statement can now be cast in the form of a weak variational statement with the boundary conditions given by Eq. (20). The geometry is the same illustrated in Fig. 1.

Given: The energy E, the potential energy $V(\mathbf{r})$, and the set $\{a^i_m\}$, where each a^i_m is the complex amplitude of the incoming wave in the mth mode of the ith lead,

Find: $\psi_0(\mathbf{r})$ for $\mathbf{r} \in \Omega_0$ such that

$$\frac{\hbar^2}{2m^*}\int_{\Omega_0} \nabla\bar{\psi}\cdot\nabla\psi_0\, d^2r + \int_{\Omega_0} \bar{\psi}(V-E)\psi_0\, d^2r = \frac{\hbar^2}{2m^*}\sum_i \int_{\Gamma_i} \bar{\psi} f_i[\xi_i,\psi_0]\, d\Gamma_i, \tag{26}$$

where

$$f_i[\xi_i,\psi_0] = \sum_{m=1}^{N'} ik'_m \chi^i_m(\xi_i) \times\left(-2a'_m + \int_0^{d_i} \chi^i_m(\xi)\psi_0(\eta_i=0,\xi_i)\, d\xi_i\right) - \sum_{m=N'+1}^{\infty} k'_m \chi^i_m(\xi_i) \times\left(\int_0^{d_i} \chi^i_m(\xi_i)\psi_0(\eta_i=0,\xi_i)\, d\xi_i\right), \tag{27}$$

and $\bar{\psi}$ is an arbitrary test function which is zero on Γ_0.

III. NUMERICAL SOLUTION

A. Finite element discretization

The discretization is on a mesh with m nodal points, $\mathbf{r}_1,\mathbf{r}_2,\ldots,\mathbf{r}_m$. Associated with each nodal point, $\mathbf{r}_i$, is a global shape function $\phi_i(\mathbf{r})$ which has the property that

$$\phi_i(\mathbf{r}_j) = \delta_{i,j}. \tag{28}$$

The wave function can be approximately expanded in the basis of these shape functions,

$$\psi_0(x) = \sum_i \psi(\mathbf{r}_i)\phi_i(\mathbf{r}) = \sum_i u_i\phi_i(\mathbf{r}) \tag{29}$$

or

$$\psi_0(x) = \mathbf{N}(\mathbf{r})\cdot\mathbf{u}, \tag{30}$$

where $\mathbf{N}(\mathbf{r})$ is the $(1\times m)$ matrix of global shape functions,

$$\mathbf{N}(\mathbf{r}) = [\phi_1(\mathbf{r}),\phi_2(\mathbf{r}),\phi_3(\mathbf{r}),\ldots,\phi_m(\mathbf{r})], \tag{31}$$

and $\mathbf{u}$ is the $(m\times 1)$ matrix of the (unknown) nodal values of ψ. Similarly, we can approximate the gradient of the wave function by an expansion on this basis set,

$$\nabla\psi_0(\mathbf{r}) = \begin{pmatrix} \dfrac{\partial\psi_0}{\partial x} \\ \dfrac{\partial\psi_0}{\partial y}\end{pmatrix} = \mathbf{B}(\mathbf{r})\cdot\mathbf{u}, \tag{32}$$

where

$$\mathbf{B}(r) = \begin{pmatrix} \partial_x\phi_1(\mathbf{r})\partial_x\phi_1(\mathbf{r})\cdots\partial_x\phi_m(\mathbf{r}) \\ \partial_y\phi_1(\mathbf{r})\partial_y\phi_2(\mathbf{r})\cdots\partial_y\phi_m(\mathbf{r})\end{pmatrix} \tag{33}$$

is the $(2\times m)$ matrix of derivatives of the global shape functions.

We can make similar approximate expansions for the test function $\bar{\psi}$:

$$\bar{\psi} = \bar{\mathbf{u}}^T\cdot\mathbf{N}^T \tag{34}$$

$$\nabla\bar{\psi} = \bar{\mathbf{u}}^T\cdot\mathbf{B}^T, \tag{35}$$

where $\bar{\mathbf{u}}$ is the vector of nodal values of $\bar{\psi}$.

Inserting these approximate expansions into the Schrödinger equation yields,

$$\bar{\mathbf{u}}^T\left(\int_{\Omega_0} \frac{\hbar^2}{2m^*}\mathbf{B}^T(\mathbf{r})\mathbf{B}(\mathbf{r})d^2r\right)\mathbf{u} + \bar{\mathbf{u}}^T \times\left(\int_{\Omega_0} [V(\mathbf{r})-E]\mathbf{N}^T(\mathbf{r})\mathbf{N}(\mathbf{r})d^2r\right)\mathbf{u} = \frac{\hbar^2}{2m^*}\sum_i\left(\int_{\Gamma_i} \bar{\psi}(\mathbf{r})f_i[\xi_i,\psi_i(0,\xi_i)]d\Gamma_i\right). \tag{36}$$

For the moment we leave the right-hand side in an undiscretized form. We define the $(m\times m)$ matrices $\mathbf{K}$, $\mathbf{T}$, and $\mathbf{V}$ as follows:

$$\mathbf{T} \equiv \int_{\Omega_0} \frac{\hbar^2}{2m^*}\mathbf{B}^T(\mathbf{r})\mathbf{B}(\mathbf{r})d^2r, \tag{37}$$

$$\mathbf{V} \equiv \int_{\Omega_0} [V(\mathbf{r})-E]\mathbf{N}^T(\mathbf{r})\mathbf{N}(\mathbf{r})d^2r. \tag{38}$$

The (partially) discretized Schrödinger equation is then,

$$\bar{\mathbf{u}}^T(\mathbf{T}+\mathbf{V})\mathbf{u} = \frac{\hbar^2}{2m^*}\sum_i\left(\int_{\Gamma_i} \bar{\psi}(\mathbf{r})f_i[\xi_i,\psi_0(0,\xi_i)]d\Gamma_i\right). \tag{39}$$

The discretization of the right-hand-side boundary term proceeds as follows.

Let $\mathbf{u}_i$ be the projection of $\mathbf{u}$ onto Γ_i. Let the length of $\mathbf{u}_i$ be M_i, the number of nodal points on the boundary (see Fig. 3). Let $\mathbf{N}_i(\xi_i)$ be the $(1\times M_i)$ matrix of global shape functions, such that

$$\psi_0(\mathbf{r}\in\Gamma_i) = \psi_0(\eta_i=0,\xi_i) = \mathbf{N}_i(\xi_i)\cdot\mathbf{u}_i. \tag{40}$$

Further, define a $(1\times M_i)$ matrix $\mathbf{N}_{i,m}$ by

$$\mathbf{N}_{i,m} \equiv \int_0^{d_i} \chi^i_m(\xi_i)\mathbf{N}_i(\xi_i)d\xi_i. \tag{41}$$

The boundary term then can be written,

$$\frac{\hbar^2}{2m^*}\sum_i\left(\int_{\Gamma_i} \bar{\psi}(\mathbf{r})f_i(\mathbf{r},\mathbf{u})d\Gamma_i\right) = \frac{\hbar^2}{2m^*}\sum_i \bar{\mathbf{u}}^T\left[\sum_{m=1}^{N'} -2ia'_m k'_m \int_0^{d_i} \mathbf{N}_i^T(\xi_i)\chi^i_m(\xi_i)d\xi_i + \sum_{m=1}^{N'} ik'_m\left(\int_0^{d_i} \mathbf{N}_i^T(\xi_i)\chi^i_m(\xi_i)d\xi_i\right)\left(\int_0^{d_i} \mathbf{N}_i(\xi_i)\chi^i_m(\xi_i)d\xi_i\right)\mathbf{u}_i - \sum_{m=N'+1}^{\infty} k'_m\left(\int_0^{d_i} \mathbf{N}_i^T(\xi_i)\chi^i_m(\xi_i)d\xi_i\right)\left(\int_0^{d_i} \mathbf{N}_i(\xi_i)\chi^i_m(\xi_i)d\xi_i\right)\mathbf{u}_i\right]. \tag{42}$$

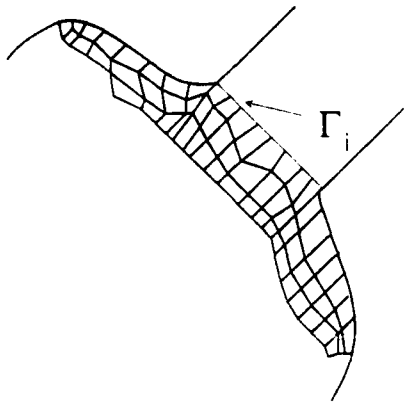


FIG. 3. Discretization of the interior region in the neighborhood of lead i.

Now using the definition of $\mathbf{N}_{i,m}$, we obtain,

$$\frac{\hbar^2}{2m^*}\sum_i\left(\int_{\Gamma_i}\bar{\psi}(\mathbf{r})f_i(\mathbf{r},\mathbf{u})d\Gamma_i\right)$$
$$=\frac{\hbar^2}{2m^*}\sum_i\bar{\mathbf{u}}^T$$
$$\times\left[\sum_{m=1}^{N^i}-2ia^i_m k^i_m\mathbf{N}^T_{i,m}+\left(\sum_{m=1}^{N^i}ik^i_m\mathbf{N}^T_{i,m}\mathbf{N}_{i,m}\right)\mathbf{u}_i\right.$$
$$\left.-\left(\sum_{m=N^i+1}^{\infty}k^i_m\mathbf{N}^T_{i,m}\mathbf{N}_{i,m}\right)\mathbf{u}_i\right]. \tag{43}$$

Define a vector $\mathbf{P}_i$ and matrix $\mathbf{C}_i$ as

$$\mathbf{P}_i\equiv-\frac{\hbar^2}{2m^*}\sum_{m=1}^{N^i}2ia^i_m k^i_m\mathbf{N}^T_{i,m} \tag{44}$$

and

$$\mathbf{C}_i\equiv-\frac{\hbar^2}{2m^*}\sum_{m=1}^{N^i}ik^i_m\mathbf{N}^T_{i,m}\mathbf{N}_{i,m}$$
$$+\frac{\hbar^2}{2m^*}\sum_{m=N^i+1}^{\infty}k^i_m\mathbf{N}^T_{i,m}\mathbf{N}_{i,m}. \tag{45}$$

The boundary term can then be written as

$$\sum_i\bar{\mathbf{u}}^T(\mathbf{P}_i-\mathbf{C}_i\mathbf{u}_i). \tag{46}$$

This can be further simplified if we define the embedded matrices $\hat{\mathbf{C}}$ and $\hat{\mathbf{P}}$ by

$$\sum_i\bar{\mathbf{u}}_i^T\mathbf{C}_i\mathbf{u}_i=\bar{\mathbf{u}}^T\hat{\mathbf{C}}\mathbf{u} \tag{47}$$

and

$$\sum_i\bar{\mathbf{u}}_i^T\mathbf{P}_i=\bar{\mathbf{u}}^T\hat{\mathbf{P}}. \tag{48}$$

The discretized Schrödinger equation then becomes

$$\bar{\mathbf{u}}^T(\mathbf{T}+\mathbf{V})\mathbf{u}=\bar{\mathbf{u}}^T\hat{\mathbf{P}}-\bar{\mathbf{u}}^T\hat{\mathbf{C}}\mathbf{u} \tag{49}$$

or

$$\bar{\mathbf{u}}^T(\mathbf{T}+\mathbf{V}+\hat{\mathbf{C}})\mathbf{u}=\bar{\mathbf{u}}^T\hat{\mathbf{P}}. \tag{50}$$

Since $\bar{\mathbf{u}}$ is arbitrary, this reduces to simply

$$(\mathbf{T}+\mathbf{V}+\hat{\mathbf{C}})\mathbf{u}=\hat{\mathbf{P}}. \tag{51}$$

This is simply a set of m algebraic equations for the m unknown nodal values of ψ in terms of the known quantities $\mathbf{T}$, $\mathbf{V}$, $\hat{\mathbf{C}}$, and $\hat{\mathbf{P}}$.

B. Algorithm summary

We summarize the numerical solution algorithm as follows.

(i) Discretize the device region Ω_0 on a mesh $\mathbf{r}_i$, $i=1,2,\ldots,m$. The potential energy $V(x,y)$ should be known at least on these nodal points.

(ii) Choose a set of basis functions, $[\phi_i(x,y)]$, which obey Eq. (28).

(iii) Construct the $(m\times m)$ matrix $\mathbf{T}$ using Eq. (37) or

$$T_{i,j}=\frac{\hbar^2}{2m^*}\int_{\Omega_0}[\partial_x\phi_i(\mathbf{r})\partial_x\phi_j(\mathbf{r})-\partial_y\phi_i(\mathbf{r})\partial_y\phi_j(\mathbf{r})]d^2r. \tag{52}$$

(iv) Construct the $(m\times m)$ matrix $\mathbf{V}$ using Eq. (38) or

$$V_{i,j}=\int_{\Omega_0}[V(\mathbf{r})-E]\phi_i(\mathbf{r})\phi_j(\mathbf{r})d^2r. \tag{53}$$

(v) For each lead i, calculate the transverse eigenfunctions $\chi^i_m(\xi_i)$. Using the basis functions associated with M_i nodal points on the lead boundary Γ_i, calculate the vector $\mathbf{N}_{i,m}$ using Eq. (41).

(vi) Using $\mathbf{N}_{i,m}$, the given values of the a^i_m's, and k^i_m's from Eq. (14) calculate the M_i-by-M_i matrix $\mathbf{C}_i$ and the vector $\mathbf{P}_i$ for each lead using Eqs. (45) and (44). Embed these into the $(m\times m)$ matrix $\hat{\mathbf{C}}$ and the vector $\hat{\mathbf{P}}$ (this is usually done automatically in the assembly process).

(vii) Solve the system of linear equations

$$(\mathbf{T}+\mathbf{V}+\hat{\mathbf{C}})\mathbf{u}=\hat{\mathbf{P}} \tag{54}$$

for the unknown nodal values of the wave function

$$u_j=\psi_0(\mathbf{r}_j). \tag{55}$$

Because $\mathbf{T}+\mathbf{V}+\hat{\mathbf{C}}$ is a sparse banded matrix, it is not necessary to store the entire $(m\times m)$ complex array. We have used both band-storage and profile-storage schemes. Solution routines are also available which take advantage of the sparse character of the matrix.[7]

(viii) From the calculated wave function, compute transmission coefficients, currents, or other quantities of interest.

Equation (45) contains an infinite sum over all possible evanescent modes. In practice this sum must be truncated to a value less than half the number of nodal points on the lead boundary. Usually a much smaller number of modes is sufficient for convergence. No general rule applies, however, because the number of evanescent modes excited depends on the details of $V(x,y)$ in the device region. If V varies rapidly, more evanescent modes need to be included in the sum.

C. Features of the algorithm

The principle features of the QTBM are as follows:

(i) The current-carrying states are solved for directly, yielding the wave function over the entire problem domain. Any other physical quantities of interest, e.g., the electron density, current distribution, or transmission coefficients, can then be extracted from the knowledge of the wave function. In particular, because the electron density is immedi-

ately available, extension of the method to include Poisson self-consistency should be straightforward.

(ii) Solution times are independent of the shape of $V(x, y)$ in the device region. Other than tabulating the values of the potential at nodal points, no additional computations are required to handle an arbitrary $V(x, y)$ than to handle a flat potential.

(iii) Multiple input and output leads are included naturally. Leads can have any size and extend away from the device region in any direction. Nonsquare-well leads are also handled simply.

(iv) For a device region with N_p nodal points, the algorithm requires the solution of one $2N_p \times 2N_p$ system of linear equations for the $2N_p$ unknown nodal values of the real and imaginary parts of the wave function.

(v) The algorithm is easy to implement using standard finite element programs. The current-carrying boundaries can be included as simply another element type. The new element subroutine associated with a boundary element computes **C** and **P**. The normal assembly process then performs the embedding of these into $\hat{\mathbf{C}}$ and $\hat{\mathbf{P}}$.

Presently, the most popular way of dealing with current-carrying states is to solve the Dyson equation on a tight-binding lattice for the real-space Green function.[1] This method is easiest to employ if the solution domains are very regular and the exact Green function is known in each of several connected regions. It is considerably more costly for arbitrary shapes and potentials. More important, it does not yield the wave function itself in the device region. The considerable physical insight which can be gained by having the full function is lost. Because of this, it may prove difficult to develop self-consistent solutions of the Schrödinger and Poisson equations. It should be noted that the Green function's may have an advantage in extending the current methods to include dissipation, techniques based on the Schrödinger equation will probably never be able to accomplish.

Recently Frohne, McLennan, and Datta developed a solution algorithm based on the boundary element method.[8] Like the QTBM, this method is based on the real-space Schrödinger equation and yields the full wave function over the entire device domain. Its most important limitation is that it requires that the full set of eigenfunctions be known for a region which includes the device region, but may be larger and have more regular boundaries. It is particularly ill-equipped to handle variations of the potential within the device region, i.e., situations with $V(x, y) \neq 0$ over the device region because exact eigenfunctions are then difficult to obtain at the outset.

IV. EXAMPLES

We present two examples of the QTBM applied to fairly simple two-dimensional geometries. The first is a resonant cavity in a quantum waveguide. The geometry is illustrated in Fig. 4(b). For simplicity we assume here that the potential is zero inside the leads and in the interior of the cavity. Outside the leads and cavity the potential is taken to be infinite. The cavity has width W and length (in the direction of current flow) L. The width of both leads is d. The wave

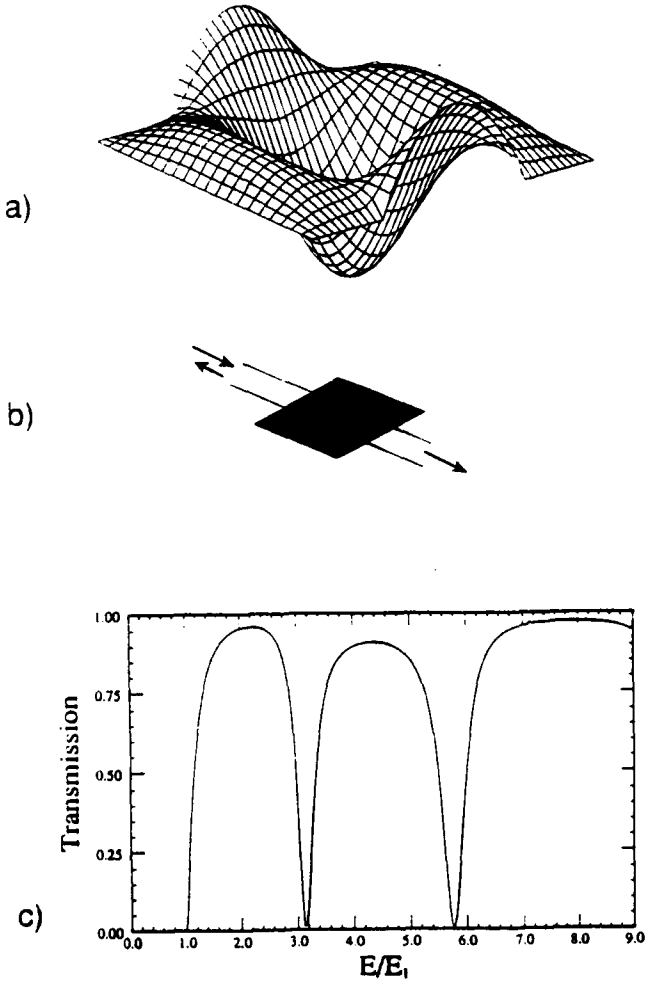


FIG. 4. Transmission through a rectangular cavity in a quantum waveguide. The real part of a typical wavefunction in a region with high transmission is shown in (a) The scattering geometry is shown in (b) which also illustrates the region (dark) over which the wave function is calculated. The transmission coefficient as a function of energy for the case of current incident in the lowest mode of the waveguide is shown in (c). The energy is normalized to E_1, the minimum energy for traveling waves to exist in the leads.

function for a state with an incident wave coming from the left, in the lowest mode of the waveguide, is solved for over the region of the cavity (the device region in this case) for various values of the energy E. In this example we take $L/d = 1$ and $W/d = 2$. Figure 4(a) shows a typical wave function obtained from the QTBM calculation. The transmission coefficient as a function of energy is shown in Fig. 5(c). The energy scale is normalized to E_1, the lowest transverse excitation of the waveguide. For energies lower the E_1, only evanescent states can occur in the channel. The minima in the transmission coefficient occur near the energies of bound states of the cavity. Transmission maxima are also observed. A discussion of the precise nature of these resonances will be given elsewhere.[9]

As a second example we consider transmission through a circular arc in a quantum waveguide. This problem is of interest if true quantum-waveguide devices are to be connected in more than a linear arrangement. The width of the waveguide is d and the central radius of curvature is r.

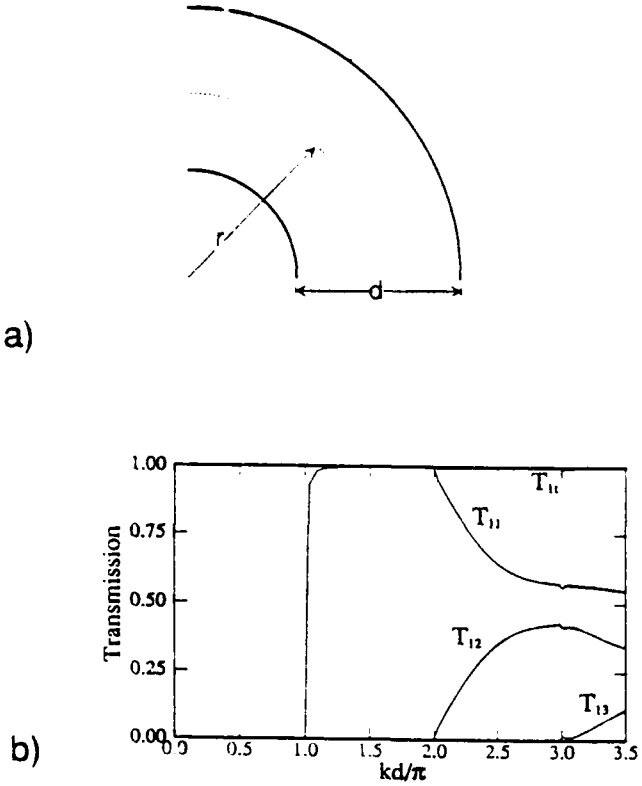


FIG. 5. A circular bend in a quantum channel. The geometry is shown in (a). The transmission coefficient for current incident in the lowest mode of the channel is displayed in (b) when $r/d = 0.7$. The transmission from the first mode into the first three modes is plotted vs kd/π where $k = \sqrt{2mE}/\hbar$.

Again, for simplicity the potential is assumed to be zero inside the waveguide and infinite outside. The arc is a full right-angle turn. The geometry of the waveguide is shown in Fig. 5(a). Unless the r is quite small the transmission is essentially unity for all values of the energy. However, the incoming wave in the first mode may be transmitted as a combination of higher-order modes. This mode mixing is the essential feature of such an arc. Figure 5(b) illustrates the mode-to-mode transmission coefficient Γ_{ij} for $i = 1$ (the incoming mode) and several outgoing modes when $r/d = 0.7$. This example illustrates the utility of the QTBM in nonrectilinear geometries. The elements used in this calculation are not rectangular but are formed by generating a regular mesh in ρ and θ, the cylindrical coordinates natural for this problem.

Both the resonant cavity and the circular arc problems will be treated at greater length elsewhere.[9] Transmission through a double cavity is discussed in Ref. 10. The examples here serve to illustrate the power of the QTBM approach. The circular arc geometry illustrates how easily nonrectilinear boundaries can be handled. In each case, a more complicated interior potential can easily be added.

V. CONCLUSIONS

We have developed a new technique for calculating numerically the solutions of the two-dimensional Schrödinger equation for current-carrying states. The quantum transmitting boundary method is based on our formulation of the boundary conditions appropriate for such states and an implementation of the finite element method. The technique is general enough to handle arbitrarily shaped device regions with complicated internal potentials. No *a priori* assumptions about the solution in the device region are required. Multiple contact leads with differing widths and various directions are handled naturally by the technique. We have demonstrated its utility in two cases of interest for quantum-waveguide devices.

[1] F. Sols, M. [illegible] .cci, U. Ravaioli, and K. Hess, Appl. Phys. Lett. **54**, 350 (1989).
[2] S. Datta, Superlattices Microstructures **6**, 83 (1989).
[3] M. Heiblum and C. Umbach, in *Nanostructure Physics and Fabrication*, edited by M. A. Reed and W. P. Kirk (Academic, Boston, 1989).
[4] A. M. Kriman, G. H. Bernstein, B. S. Haukness, and D. K. Ferry, *Superlattices Microstructures (in press)*.
[5] G. Bernstein and D. K. Ferry, J. Vac. Sci. Technol. B **5**, 964 (1987).
[6] David J. Kirkner and Craig S. Lent, International Journal of Numerical Methods in Engineering (in press).
[7] T. J. R. Hughes, *The Finite Element Method* (Prentice Hall, Englewood Cliffs, NJ, 1987).
[8] H. Rob Frohne, M. J. McLennan, and S. Datta, J. Appl. Phys. **66**, 2699 (1989).
[9] C. S. Lent and S. Sivaprakasam (unpublished).
[10] C. S. Lent, S. Sivaprakasam, and D. J. Kirkner, in *Nanostructure Physics and Fabrication*, edited by M. A. Reed and W. P. Kirk (Academic, Boston, 1989).

Solid-State Electronics Vol. 32, No. 12, pp. 1137–1141, 1989
Printed in Great Britain. All rights reserved

0038-1101/89 $3.00 + 0.00
Copyright © 1989 Pergamon Press plc

A TWO-DIMENSIONAL HOT CARRIER INJECTOR FOR ELECTRON WAVEGUIDE STRUCTURES

Craig S. Lent, Srinivas Sivaprakasam

Department of Electrical and Computer Engineering
and

David J. Kirkner

Department of Civil Engineering
University of Notre Dame
Notre Dame, Indiana 46556

ABSTRACT

The current-voltage characteristics for a constriction in a quantum waveguide channel are calculated. The constriction forms an effective barrier which can be employed as a tunneling injector. We find that such a structure may be useful in providing high-energy electrons in a single mode of the waveguide. We also examine the current in the far-from-linear response regime. Away from the linear region the current through the constriction saturates and the conductance falls to zero.

KEYWORDS

Hot carriers; Electron waveguides; Two-dimensional electron gas

INTRODUCTION

Improvements in epitaxial growth techniques have made possible a large number of quantum devices which are one-dimensional in the sense that carriers are confined in only one spatial dimension. Recent advances in lithography have begun to produce structures which are quantized in two and three spatial dimensions. Of particular interest are quantum waveguide devices in which carriers are confined to narrow channels which act analogously to microwave channels.

We investigate the ballistic transport properties of a constriction in a two-dimensional electron waveguide channel which can be used as a hot-electron injector. Electrons are confined in a channel of width d which narrows abruptly to a width d_c in the constriction itself. The constriction is a low conductivity region that forms an effective barrier between the contacts on either side. A voltage is applied across the constriction resulting in ballistic injection of carriers into the outgoing channel.

The behavior of the waveguide constriction is considered in two regimes. First, we examine a very small geometry for which only a few transverse modes are accessible. The constriction acts like a potential barrier. We compare the transmission characteristics of such a structure with a more conventional tunnel barrier. The focus here is on the utility of such a structure as a single-mode hot-electron injector. The ability to inject electrons at high energies in a single mode is an important first step in the development of a new class of quantum waveguide devices. The first wave of such devices includes the stub-tuner transistors proposed by Sols and coworkers (1989) and Datta (1989). We also examine a larger geometry in which more incoming modes are important. We find a surprising result for voltages beyond the linear response regime. The current through the constriction saturates at a value which is roughly linearly dependent on the constriction width. This current saturation results in a current-voltage characteristic remarkably similar to an FET. We discuss device possibilities briefly.

The geometries we examine are *not* directly analogous to the constrictions explored experimentally by Van Wees and coworkers (1988)

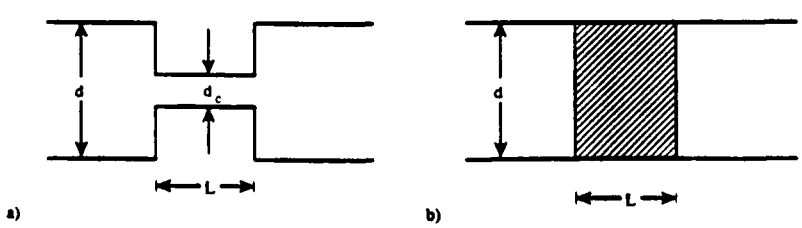


Fig. 1. The geometry of the quantum waveguide constriction is shown in a). A waveguide with no constriction, but a potential barrier is illustrated in b).

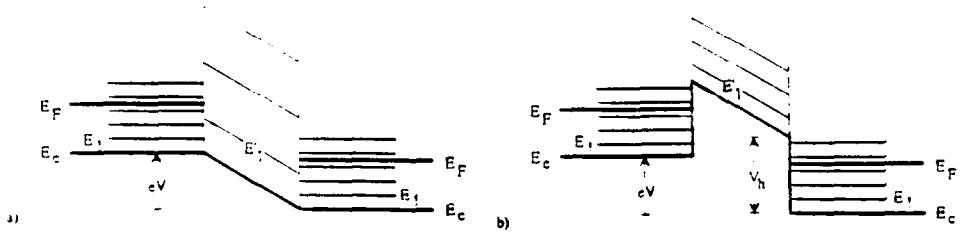


Fig. 2. The energy band diagram for the constriction, a), and the barrier, b), shown in Figure 1.

and theoretically by others (Kirzcenow, 1989; Szafer and Stone, 1989). We consider a case in which the width of the constriction is not smaller than a tenth the width of the incoming leads and both are assumed to be narrow in that only a few transverse modes are accessible. Furthermore, in our geometry the constriction is rectangular and abrupt. This resembles rather more the experimental geometry of Brown and coworkers.

THEORY

The constriction geometry is shown in Figure 1a. The incoming and outgoing leads are semi-infinite and the walls are assumed to be infinite barriers. The x direction is the direction of current flow and the waveguide is narrowed in the y direction. We suppose that quantization in the z direction, perpendicular to the plane of the figure, is complete and that only one bound state in the z direction is occupied. Note that if the potential is independent of z, the index of the bound state in the z direction is a conserved quantum number and our analysis would be correct for each bound state separately. A geometry such as that in Figure 1a can be realized by creating a high-mobility two-dimensional electron gas (2DEG) at a heterojunction interface. We envision the confining potentials on the lead walls and the constriction originating electrostatically from metal layers above the heterojunction (Sols, 1989; Datta, 1989).

Figure 2a shows an energy band diagram for the constriction. A conduction band edge is shown as a reference level and to indicate the variation in the electrostatic potential across the constriction. The confinement in the y direction raises the minimum energy of any electron above this level to E_1, the energy of the first transverse mode. The Fermi energies, as measured from the conduction band edge, are assumed to be equal in the left and right leads. The value of the Fermi energy is determined by the density in the leads. We assume all the applied voltage appears as a potential drop across the lower conductance constriction. This assumption must be checked for consistency with the final results. Note that unlike the usual one-dimensional heterostructure barriers, the contact regions here are highly conducting 2DEG's. Therefore, very little band bending in the contacts is expected. The potential is further assumed to vary linearly in the constriction region. An improved analysis would include the effect of the injected charge density, both in the constriction and in the contacts.

The wavefunction for an incoming wave in the contact can be written,

$$\psi_{n,k}(x,y) = Ae^{ik_n x} sin(n\pi y/d). \tag{1}$$

The total energy is given by

$$E(n,k_n) = \frac{\hbar^2}{2m^*}\left[k_n^2 + (\frac{n\pi}{d})^2\right]$$

and the minimum energy of each transverse mode is the infinite square well result.

$$E_n = \frac{\hbar^2}{2m^*}\left(\frac{n\pi}{d}\right)^2.$$

Figure 2a shows the energies of each mode in the contacts and in the constriction. The first mode in the constriction is higher than the first mode in the left contact, creating an effective barrier. The effective barrier differs from a real potential barrier in that it blocks higher modes from propagating more than lower modes. Although loosely speaking the barrier can be said to be higher for higher modes, it is not possible, in general, to define a single number which is an effective barrier height for a particular mode. Because mode number is not a conserved quantity, a higher mode in the lead can couple to a lower mode in the constriction. In Figure 2, for example, the electrons in the third mode in the lead cannot travel in the third mode in the construction but can couple to the first constricted mode. The differences in energy for modes in the two regions cannot alone account for the differences in the contributions of each mode to conduction. The transmission coefficients, which couple modes together must be included. For convenience we can define the effective barrier height for the first mode (which cannot couple to any lower modes) as simply the difference between the first mode energy in the constriction and in the incoming lead. Note that because the structure is symmetric, parity is a good quantum number, and even modes cannot couple to odd modes.

The current is calculated at each value of the voltage by a modified Esaki formula (Datta, 1989).

$$I(V) = \left(\frac{2e}{h}\right)\int[f(E-E_F-eV) - f(E-E_F)]\sum_{i,j} T_{i,j}(E,V)dE \tag{2}$$

The transmission coefficient $T_{i,j}$ is the ratio of the outgoing current in mode j to the incoming current in mode i (Duke, 1969). This formula reduces to the Buttiker-Landauer formula (for a review see Stone and Szafer, 1988) if the difference in the Fermi functions is expanded to first order in eV and evaluated at zero temperature. The use of equation (2) extends the treatment of Szafer and Stone (1989) beyond the linear

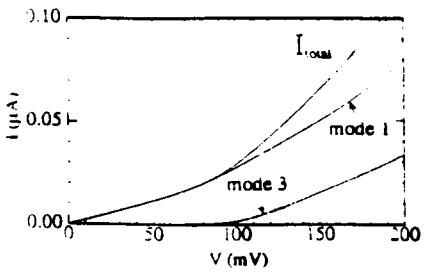


Fig. 3. The current-voltage relationship for a 200 Å channel with a 60 Å constriction. The total current and the components injected into the first and third mode are shown.

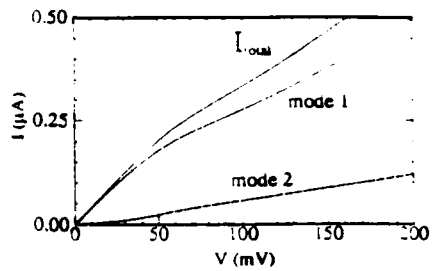


Fig. 4. The current-voltage relationship for a 200 Å channel with a 200 meV barrier. The total current and the components injected into the first and second mode are shown.

response regime. For each value of the applied voltage, the transmission coefficients are calculated by solving the effective-mass Schrodinger equation in the region of the constriction. We use a Finite Element discretization scheme. The problem of including current-carrying boundaries is treated by the Quantum Transmitting Boundary technique (Kirkner, Lent, and Sivaprakasam, 1989).

SINGLE-MODE WAVEGUIDE INJECTOR

The idea of using the waveguide constriction as a single-mode injector is to exploit two features of the constricted region. First, the effective barrier formed by the constriction can be employed as a tunneling barrier and used to inject hot electrons into the outgoing lead. Second, because it blocks higher modes, the constriction acts as a mode filter. Together the effect is to enhance single-mode injection.

We choose a geometry with $d_c/d = 3/10$ and and $L/d = 1/4$. The Fermi energy $E_F = E_2$, the energy of the second mode in the wide contact leads. Only the first two incoming modes are appreciably occupied. The temperature is taken to be 77 K and we use an effective mass $m^* = 0.05m_0$. If we take $d = 200$ Å, then the the effective first-mode barrier height for this constriction is 200 meV. The current is calculated as a function of voltage using Equation (2). The results are shown in Figure 3. The total current is plotted, as well as the components of the current in modes 1 and 3 in the outgoing lead. Higher modes cannot propagate at these energies. Mode 2 cannot propagate in the constriction and is totally reflected. There is no component of mode 2 in the outgoing lead because mode 1 in the constriction can only couple to symmetrical modes in the leads.

For comparison we calculate the current-voltage relation for a square potential barrier in a 200 Å channel with the same height (200 meV) as the effective barrier of the constriction. The geometry is shown in Figure 1b and an energy level diagram for this real barrier is shown in Figure 2b. For the real barrier mode number is conserved so the current can be broken into components in mode 1 and mode 2, which are the same in all regions.

Comparing Figures 3 and 4, several observations can be made. The current is higher in magnitude for the real barrier. This is partially because the mode 2 component in the incoming lead is completely reflected by the constriction. In addition, the fraction of the current through the constriction is limited by simply the geometrical factor of the ratio of the channel widths. Roughly speaking, that fraction of the incoming wave that hits the infinite barriers on either side of the entrance to the constriction is completely reflected. On the other hand, because the constriction effectively blocks the second mode from propagating, the transmitted current is nearly completely single-moded below 100 mV. After that, there is a growing admixture of the third mode. For some device applications, the ability to inject essentially single-moded electrons, with a high energy (100 meV), may be worth the somewhat lower current output.

INJECTED CURRENT SATURATION

For structures of the size discussed above, the current is a rapidly increasing function of voltage at all reasonable values of the applied voltage. Higher values would tend to transfer electrons into other valleys in the underlying bandstructure. If we scale up the size, however, we can observe the far-from linear response regime at accessible voltages.

Figure 5 shows the current-voltage characteristic for a constriction with $d_c/d = 0.3$ and $L/d = 0.25$. The total current is shown broken down into the components due to each incoming mode. The Fermi level for this case is chosen to be $35E_1$ (where E_1 is the first mode in the contact channels). Five incoming modes are occupied, although the contribution to the current of the fifth mode is very small. The current is nearly linear with voltage for small voltages, rolls off in a transition region, and saturates to a constant value. As the figure shows, each mode saturates at a slightly different voltage. Saturation occurs for the most important modes at about the point where the applied voltage reaches E_F/e.

To understand this behavior, consider first the current, as given by Equation (2), for small voltages (less than E_F/e). The voltage dependence occurs in both the transmission coefficients and in the difference of Fermi factors. There are some oscillations in the transmission coefficients, but these tend to be smeared out by the sum over all the modes. The dominant voltage dependence is from the difference in Fermi factors. This difference is linear in the voltage to first order (the linear response result), and the resulting linear behavior is seen in the current-voltage characteristics shown in Figure 5.

When the potential drop across the channel is greater than the Fermi voltage the current from the right lead back to the left becomes

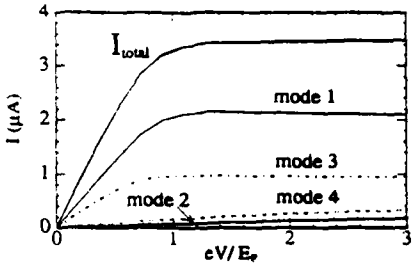


Fig. 5. The current-voltage characteristic for a 30% constriction ($d_c/d = 0.3$). The total current and the contribution of the first four incoming modes is shown.

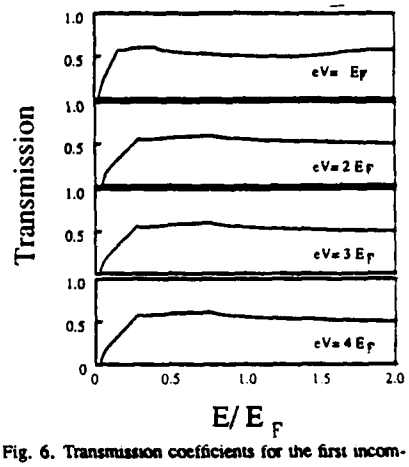


Fig. 6. Transmission coefficients for the first incoming mode at various values of applied voltage.

insignificant. Considering Equation (2), this means that the second Fermi factor in the integrand becomes negligible. The integral then extends from the conduction band edge in the left lead up to several $k_B T$'s above the Fermi energy in the left lead. The Fermi factor term looses any dependence on the voltage and only the voltage dependence of the transmission coefficients remain. The transmission coefficients are increasing as the effective barrier is bent down lower and lower. Because the effective barrier is small, and now thin, the transmission coefficients saturate. The transmission coefficients for the first incoming mode, $T_{1,t}$ are shown as a function of energy in Figure 6 for four values of applied voltage. For $eV = E_F/e$ the transmission is still changing with applied voltage, albeit weakly. For voltages of 2, 3, and 4 times the Fermi voltage, the transmission coefficients are nearly identical functions of energy. They saturate, for this first mode, at a value of about one half. This value is roughly the fraction of the incoming flux that "hits" the constriction instead of the walls on either side. Higher modes saturate at different values. It is this saturation of the transmission after the Fermi function term becomes voltage independent that is behind the current saturation.

The current-voltage characteristic for constrictions of various widths are shown in Figure 7. The saturation current varies nearly linearly with the constriction width. We also show in Figure 8 the result for $d_c/d = 1$, no constriction at all. This curve is only physically valid at small values of the potential but it does serve as an important check on the validity of our initial assumptions. At a value of the current corresponding to saturation in one of the constrictions, this shows how much of the voltage would be dropped by the incoming leads. Consider the 30% constriction. At the saturation current of the constriction, about 3 μA, the voltage across the leads would be about $V_F/5$. This is only 5 – 10% of the applied voltage across the constriction in the saturation regime. Within these limits, we are justified in assuming that the applied voltage appears across the constriction. Note that we are careful not to implicitly invoke Ohm's law here, but use the calculated $I(V)$ functions.

The low-voltage conductance of the unconstricted channel should correspond to the linear response result. Conductances are shown in Figure 9 for the various values of the constriction. The conductance of the unconstricted channel, in which 5 modes can propagating, is 5 times the fundamental conductance, $2e^2/h$, confirming the agreement with linear response theory for small voltages.

At very low temperature, the injected current actually diminishes after a saturation peak. Figure 10 shows a region of negative differential resistance which occurs in the low temperature case. Here we take $d = 1400$ Å, $d_c = 280$ Å, $L = 350$ Å, and $E_F = 10$ meV. This agrees qualitatively with the one-dimensional treatment of Kelly and the experimental measurements of Brown and coworkers.

The family of curves shown in Figure 7 illustrate the device potential of such a constricted geometry beyond that of a hot-electron injector.

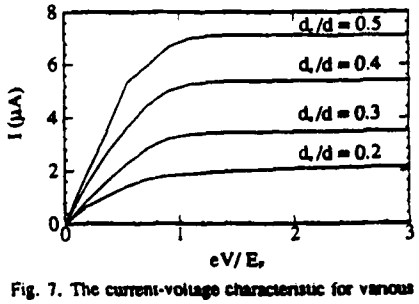


Fig. 7. The current-voltage characteristic for various size constrictions.

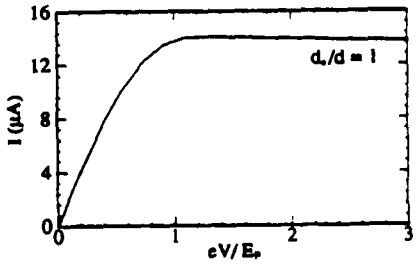


Fig. 8. The current-voltage characteristic for a channel with no constriction is shown for comparison with Figure 6.

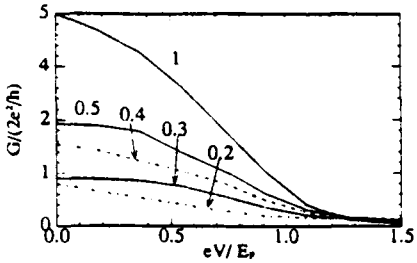


Fig. 9. The conductance as a function of applied voltage for a channel and various constrictions. The value of d_c/d is shown for each curve. There are five propagating modes in the incoming waveguide.

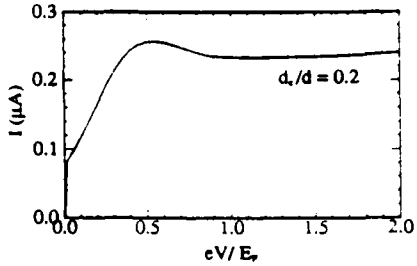


Fig. 10. The current-voltage characteristic for a narrow channel at very low temperature ($T = 100$ mK).

It is possible to construct a geometry in which a gate voltage varies the width of the channel in a manner similar to the tuning stub in the quantum the stub-tuner transistor (Sols, 1989; Datta 1989). It would then be possible to gate the saturation current with a third terminal.

We note finally that the transmitted currents involved here in the saturation regime are quite large. It is likely that the limits of strict validity of Equation (2) are being exceeded. In particular, the derivation of this equation assumes that the transmitted and reflected currents do not alter occupation statistics or the Fermi levels in the leads. In the saturation region, a large fraction of electrons which impinge on the constriction from the left are transmitted to the right so this assumption undoubtedly needs correcting. The very recent work of Bandara and Coon (Bandara and Coon, 1989) proposes just such a correction for the three-dimensional electron gas case.

CONCLUSIONS

We have examined the current-voltage relationship for a constriction in a quantum waveguide. We find that for small geometries the constriction may have advantages over barrier structures as a single-mode hot-electron injector to be used in conjunction with other wave-guide devices. We also examined the far-from-linear response regime for larger structures and find a surprising, and possibly useful current saturation effect.

ACKNOWLEDGEMENTS

The authors are grateful for preprints made available by M. J. Kelly and R. J. Brown, as well as helpful discussions.

The work was supported by the Air Force Office of Scientific Research under grant number AFOSR-88-0096, and the National Science Foundation under contract number EET87-07628.

REFERENCES

Bandara, K. M. S. V., and D. D. Coon, J. Appl. Phys. **66**, 693 (1989).

Brown, R. J., M. J. Kelly, M Pepper, H Ahmed, D. G. Hasko, D. C. Peacock, J. E. F. Frost, D. A. Ritchie and G. A. C Jones (unpublished).

Datta, S., Superlattices and Microstructures **6**, 83 (1989).

Duke, C. B., in E. Burstein and S. Lundqvist (Eds.) *Tunneling Phenomena in Solids*, (Plenum Press, New York, 1969) pp. 31-46.

Kelly, M. J., (unpublished).

Kirkner, D. J., C. S. Lent, and S. Sivaprakasam, (unpublished)

Kirczenow,G., Phys. Rev. **B39**, 10452 (1989).

Lent, C.S., D. J. Kirkner, and S. Sivaprakasam, in M. A. Reed and W. P. Kirk (Eds.) *Nanostructure Physics and Fabrication*, (Academic Press, Boston, 1989).

Sols, F., M. Macucci, U. Ravaioli, and K. Hess, Appl. Phys. Lett. **54**, 350 (1989).

Stone, A. D., and A. Szafer, IBM J. Res. Develop. **32**, 384 (1989).

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

Van Wees, B. J., 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).

CALCULATION OF BALLISTIC TRANSPORT IN TWO-DIMENSIONAL QUANTUM STRUCTURES USING THE FINITE ELEMENT METHOD [1]

Craig Lent
Srinvivas Sivaprakasam
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

David J. Kirkner
Department of Civil Engineering
University of Notre Dame
Notre Dame, Indiana 46556

I. INTRODUCTION

Current fabrication technology permits the construction of ultra-small semiconductor structures in one dimension, usually the direction of crystal growth. This capability has spawned a wealth of experiments and theory describing transport in the ballistic regime. This development has been aided greatly by the fact that ballistic transport can be understood, qualitatively at least, by simply solving the one-particle Schrödinger equation. The one-dimensional form of the Schrödinger equation is fairly easily solved so that transmission coefficients and currents can be calculated.

Fabrication technology is becoming increasingly sophisticated and is now beginning to create structures quantized in two and three spatial dimensions. The leap to two dimensions makes the solution of the Schrödinger equation considerably more challenging. Analytic textbook solutions become inadequate for guiding intuition and design. Unbound states which carry current require particularly careful analysis in two dimensions. It is important in improving our understanding of ballistic transport in two-dimensional electron wave-guide devices that sufficiently powerful and flexible numerical methods be developed.

We have used the Finite Element Method (FEM) to solve the single-particle Schrödinger equation for two-dimensional potentials. While calculations of bound

[1] Supported by the Air Force Office of Scientific Research under grant no. AFOSR-88-0096 and by an IBM Faculty Development Award.

Copyright © 1989 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-585000-X

state wavefunctions have been done previously [3], this represents the first method to yield wavefunctions for states which carry current. We present solutions for the transmission coefficients of double-cavity electron wave-guide structures. The FEM provides a very flexible, elegant way of handling boundary conditions for very complex structures. Ultimately self-consistent solutions, at least in the Hartree approximation, are required. The method presented here lends itself well to such an extension because it yields the wavefunctions directly. As demonstrated in the double-barrier resonant tunneling problem, single-particle solutions can nevertheless reveal most of the important transport features.

II. THEORY

There are several ways in which two-dimensional electron waveguide structures might be fabricated. A metal pattern deposited on an AlGaAs-GaAs heterojunction can be used to create channels in the two-dimensional electron gas formed at the heterojunction interface. This technique has been used by Bernstein and Ferry in making very fine grid structures on a FET gate [1]. The technique might be refined by using a quantum well instead of the heterojunction potential to confine carriers in the plane. Another technique which may prove useful involves etching and regrowth of lithographically defined patterns in quantum well layers. We do not concern ourselves here with the exact method used, but look instead for the basic transport features such structures would exhibit.

We consider a system in which electrons are confined in the xy-plane by some potential which is such that only the ground state z-eigenfunction is ever occupied. The potential in the xy-plane, defined by some lithographic means, is assumed to take the form of rectangular waveguides which act as input and output leads, connected to a device region. The geometry for the double-cavity structure is shown in Figure 1. In this case the device region is simply the two rectangular cavities and the short channel which connects them. For simplicity

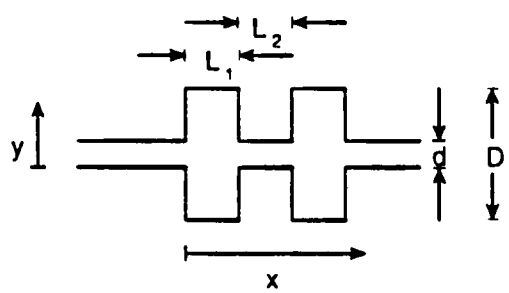


Figure 1: Geometry for the double resonant cavity. The two cavities are here assumed to have the same width D and length L_1.

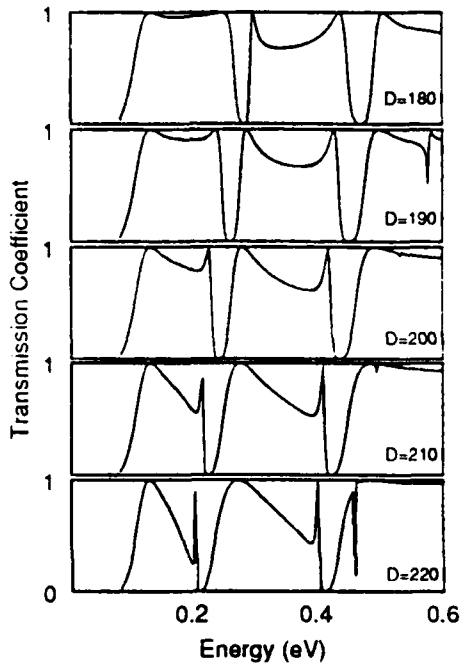


Figure 2: The transmission coefficient versus energy for the double cavity shown in Figure 1 with $d = 100$ Angstroms, $L_1 = 100$ Angstroms, $L_2 = 50$ Angstroms, and various values of D.

we assume the rectangular waveguides of width d have infinite potential walls so the wavefunctions in the leads consist of plane waves in the x-direction and sine functions in the y-direction. We assume an incoming plane wave from the left, which produces a reflected wave and a transmitted plane wave. The wavefunction in the left ($x < 0$) contact is then,

$$\begin{aligned} \psi^L(x,y) &= A_m e^{ik_m x} sin\left(\frac{m\pi y}{d}\right) \\ &+ \sum_{n=1}^{N} b_n e^{-ik_n x} sin\left(\frac{n\pi y}{d}\right) + \sum_{n=N+1}^{\infty} b_n e^{k_n x} sin\left(\frac{n\pi y}{d}\right). \end{aligned} \tag{1}$$

In the right contact the wavefunction is

$$\psi^R(x,y) = \sum_{n=1}^{N} a_n e^{ik_n x} sin\left(\frac{n\pi y}{d}\right) + \sum_{n=N+1}^{\infty} a_n e^{-k_n x} sin\left(\frac{n\pi y}{d}\right), \tag{2}$$

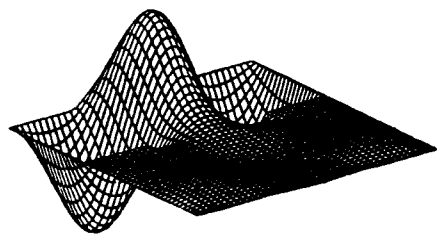


Figure 3: The real part of the wavefunction for an energy of 0.23 eV, for the double cavity problem. The geometry is the same as for Figure 2 with $D = 200\text{Å}$.

where

$$k_n = \sqrt{|E - \frac{\hbar^2}{2m^*}\left(\frac{n\pi}{d}\right)|}$$

The incoming wave has amplitude A_m and represents an excitation of the m-th transverse mode of the input waveguide. N is the number of traveling modes possible in the waveguide leads. The modes with $n > N$ are evanescent modes with complex wavenumber and carry no current.

We solve the two-dimensional effective-mass Schrödinger equation in the device region using the Finite Element Method (FEM). The details of the method are presented elsewhere [4]. The FEM enables us to straightforwardly include the condition that the wavefunction and its normal derivative match the analytical form of equations (1) and (2) as an additional set of constraint equations. The region is discretized into small elements on which the wavefunction is approximated by bilinear shape functions. This discretization yields a set of algebraic equations for the values of the wavefunction at the nodal points. For simplicity we assume infinite potential barriers at boundary walls. The boundary conditions are the that the wavefunction be zero on the cavity walls and that the wavefunction and its first derivative match those for the analytical expressions given in equations (1) and (2) above at the interface between the leads and the device region. The incoming and outgoing current are calculated directly from the wavefunction and the transmission coefficient is obtained from their ratio.

III. RESULTS

The transmission coefficient as a function of energy is shown for the double-

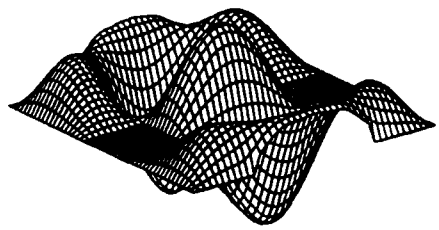


Figure 4: The real part of the wavefunction for an energy of 0.27 eV, for the double cavity problem. The geometry is the same as for Figure 2 with $D = 200$Å.

cavity structure in Figure 2. It has been calculated for several values of D, the cavity width. The input and output channels are 100 Å wide and each cavity is 100 Å in length with a 50 Å long channel connecting them. The FEM discretization is accomplished on square 5 Å elements. Results were sufficiently converged at this mesh size that further refinements had no significant effect on the results.

The two main minima, at energies slightly above 0.2 eV and 0.4 eV, are due to excitations of bound states of a single cavity. The real part of the wavefunction near such a minimum is shown in Figure 3. The plateau regions of high transmission between these minima owe their shape to the coupling between the two cavities. The maxima at the high-energy and low-energy ends of these plateaus are due to states which excite both wells. A sample wavefunction is shown in Figure 4.

It is interesting to note that the shape of the transmission coefficient vs. energy curves is quite different for cavity structures than for constrictions. In a constricted geometry, one observes sharp peaks, corresponding to the modes allowed in the narrow region, and broad valleys. In the cavity resonance case examined here, the transmission dips sharply at resonances and is generally high between them.

The energy dependence of the transmission coefficients shown in Figure 2 suggests possible applications of such cavities in waveguide devices. For some implementations of lithographically defined structures, such as the metal-on-heterojunction technique, the cavity dimensions can be changed dynamically by applying voltages to the metal overlayers and shrinking or enlarging depleted regions. This effectively would sweep the transmission characteristic through a family of curves such as those shown in the figure. This presents the possibility of transistor action in such devices [2]. In addition, as the cavity width increases,

a large region of negative differential resistance occurs as the high-transmission plateaus slope downward. This region, while not especially steep, has the advantages of being broad and tunable (through changes in D).

References

[1] G. Bernstein and D.K. Ferry, J. Vac. Sci. Technol. **B5** 964 (1987).

[2] Fernando Sols, M. Macucci, U. Ravaioli, and Karl Hess, Appl. Phys. Lett. **54**, 350 (1989).

[3] Steven E. Laux and Frank Stern, Appl. Phys. Lett. **49**, 91 (1986).

[4] Craig S. Lent, Srinivas Sivaprakasam, and David J. Kirkner, submitted to the International Journal of Numerical Methods in Engineering.

Scattering matrix analysis of electron transport in disordered Aharonov–Bohm interferometers and ballistic constrictions

M. Cahay
Nanoelectronics Laboratory, Department of Electrical and Computer Engineering, University of Cincinnati, Cincinnati, Ohio 45221-0030

S. Bandyopadhyay
Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556

H. R. Frohne
School of Engineering, Walla Walla College, College Place, Washington 99324

(Received 29 May 1990; accepted 25 July 1990)

We present a fully quantum-mechanical analysis of phase-coherent electron transport in disordered semiconductor nanostructures. The analysis is based on a scattering matrix formalism which allows us to simulate the effects of interface roughness scattering, as well as scattering from point defects and defect clusters. Using this technique, we have studied quantum conduction in electrostatic Aharonov–Bohm interferometers and narrow ballistic constrictions of submicron dimensions.

I. INTRODUCTION

Recent advances in nanolithography have made it possible to fabricate "mesoscopic structures" with dimensions smaller than the phase-breaking length of electrons at cryogenic temperatures.[1] In these structures, many novel quantum interference effects have been observed; they include the Aharonov–Bohm effect,[2] universal conductance fluctuations,[3] nonlocal voltage fluctuations,[4] quantized conductance of narrow ballistic constrictions,[5] and "stub-tuning" in electron waveguides.[6] In all of these effects, the quantum-mechanical wave nature of electrons plays a critical role so that the analysis of these phenomena require a full quantum-mechanical description of electron transport.

In this paper, we have used a fully quantum-mechanical formalism—termed the "generalized scattering matrix approach"—to study electron transport in electrostatic Aharonov–Bohm interferometers and narrow ballistic constrictions. The generalized scattering matrix approach was discussed by us earlier.[7] It has been shown to be highly successful in reproducing several results of the theory of weak and strong localization, as well as universal conductance fluctuations.[8] The advantages of this technique are that unlike in the case of most other techniques, it is easy to include the effect of evanescent states in the analysis. These states have a significant influence on transport if the structure under study is severely disordered.[7,9] Moreover, the scattering matrix method is ideal for treating quantum conduction through semiconductor nanostructures that have large-scale disorder[7,8] and offer an alternative approach to the more widely used Green's function and transfer matrix techniques to study mesoscopic systems.[10]

The organization of this paper is as follows. In Sec. II we describe the theoretical formalism. We then present results from our study of electrostatic Aharonov–Bohm interferometers and ballistic constrictions in Sec. III. Finally, Sect. IV contains concluding remarks.

II. THEORETICAL FORMALISM

In analyzing quantum conduction through a semiconductor nanostructure, we are essentially interested in the current response of the structure to a vanishingly small bias. In other words, we are interested in the linear response conductance which is given by the finite-temperature multichannel Landauer formula

$$G=\frac{e^2}{2hkT}\int dE\sum_i^M\sum_j^M |t_{ij}|^2\,\mathrm{sech}^2\left(\frac{E-E_F}{2kT}\right), \qquad (1)$$

where t_{ij} is the transmission amplitude for an electron incident from the left contact in mode i and exiting at the right contact in mode j, M is the total number of occupied modes in the contacts, E is the energy of the electron, and E_F is the Fermi energy.

The easiest way to evaluate the transmission amplitudes t_{ij} is to obtain them directly from the scattering matrix that describes the entire structure, including the contacts. The scattering matrix relates the amplitudes of all incoming electron waves to those of the outgoing electron waves. One can construct the overall scattering matrix for any arbitrary structure by cascading three scattering matrices—the first describing infusion of electrons from the left contact, the second describing propagation through the structure, and the third describing exit into the right contact. The elements of the first and the last scattering matrices that describe the entry and exit can be found by matching the electronic wave function and its first derivative along the boundaries between the structure and the contacts. For the case of narrow ballistic constrictions, we have calculated these two matrices exactly by using the so-called boundary element method described in Ref. 11. For the case of an Aharonov–Bohm interferometer (which is a doubly connected structure), the calculation of these matrices is somewhat complicated. Therefore, for the sake of simplicity, we have replaced them by the so-called Shapiro matrix[12] which has been widely used to describe a three-way splitter. The Shapiro matrix

 © 1990 American Vacuum Society

relates the amplitudes of the incoming and outgoing waves according to (see Fig. 1)

$$\begin{pmatrix} A^- \\ B_1^+ \\ B_2^+ \end{pmatrix} = \begin{pmatrix} -(a+b) & \sqrt{\epsilon} & \sqrt{\epsilon} \\ \sqrt{\epsilon}^* & a & b \\ \sqrt{\epsilon}^* & b^* & a \end{pmatrix} \begin{pmatrix} A^+ \\ B_1^- \\ B_2^- \end{pmatrix}, \tag{2}$$

where A and B are column vectors whose elements are the amplitudes of the incoming and outgoing waves in various modes as shown in Fig. 1. The superscript " + " refers to waves traveling to the right and " − " refers to waves traveling to the left. The superscript * denotes Hermitean conjugate. If there are M occupied modes in the contacts, then each element of the Shapiro matrix is a $M \times M$ complex matrix.

Because of the unitarity of the Shapiro matrix, the following relations hold:

$$a = \tfrac{1}{2}(\sqrt{1-2\epsilon} - 1), \tag{3}$$

$$b = \tfrac{1}{2}(\sqrt{1-2\epsilon} + 1). \tag{4}$$

The Shapiro matrix therefore has only one free parameter for every mode, namely the parameter ϵ. This parameter is a measure of the transmittivity of the mode from a contact into the structure, or from the structure into a contact. In other words, the higher the value of ϵ (its maximum value is 0.5), the more is the transmission coefficient for the mode.

The scattering matrix that describes propagation through the structure is constructed as follows.[8] The structure to be simulated is first broken up into a number of sections along its length such that each section either contains one elastic scatterer or is a region of ballistic propagation between two adjacent scatterers. For each such section, a scattering matrix can be calculated analytically. In the case of a point scatterer, we assume the scattering potential to be a delta function which approximates a heavily screened Coulomb potential for an ionized impurity. The prescription for constructing the scattering matrix for a delta potential has been given in Ref. 8.

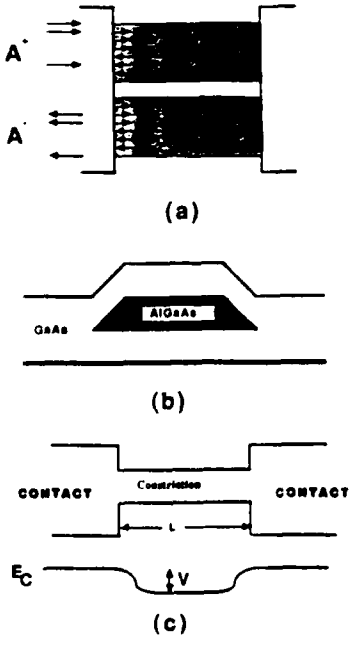


FIG. 1. (a) The incoming and outgoing waves at the junction between a contact and a doubly connected structure such as an Aharonov–Bohm interferometer. (b) A possible realization of an Aharonov–Bohm interferometer utilizing a GaAs-AlGaAs split quantum well. (c) A ballistic constriction and the bulk conduction-band-edge profile E_c along the length of the constriction. The built-in potential is denoted by V.

For simulating interface roughness scattering in a quantum well, or the effect of a defect cluster, we utilize the following procedure. Interface roughness can be viewed as a random variation of well width on a monolayer scale. This variation causes a variation in the potential energies in the well. Sudden discrete changes in the well width therefore causes a sudden discrete potential step. It may be noted that a defect cluster would also produce the same effect. Therefore, the scattering potential for either interface roughness scattering, or a defect cluster, is simulated by a constant step potential of finite but varying extent.

In Fig. 2 we elucidate the effect of interface roughness or a defect cluster. This figure shows the dispersion curves $E(k_x,k_y)$ for the mode or subband closest to the Fermi level. The left curves corresponds to a region of ballistic propagation (defect-free region) and the right curve corresponds to region containing a potential barrier due to a defect cluster. Note that the mode which was propagating (i.e., the bottom of the corresponding subband was below the local quasi-Fermi level) becomes evanescent when it enters the region with the defect cluster. In that case, this mode can no longer travel through the structure and contribute to current, unless it does so by tunneling. Since evanescent states cannot exhibit the electrostatic Aharonov–Bohm effect, it is conceivable then that such regions may have a dramatic effect on the characteristics of an Aharonov–Bohm interferometer. We shall examine such effects later in Sec. III.

Once the scattering matrices describing the individual sections have been evaluated, they are then cascaded according to the law of composition of scattering matrices[13] to yield the composite scattering matrix for the structure. Finally, the three scattering matrices—two for the contacts and one for the structure proper—are cascaded to yield the overall scattering matrix for the entire structure from whose elements the transmission amplitudes t_{ij} are obtained directly. From these amplitudes, the linear response conductance is calculated using Eq. (1).[14]

III. RESULTS

In this section, we first present results from our study of Aharonov–Bohm interferometers. An Aharonov–Bohm in-

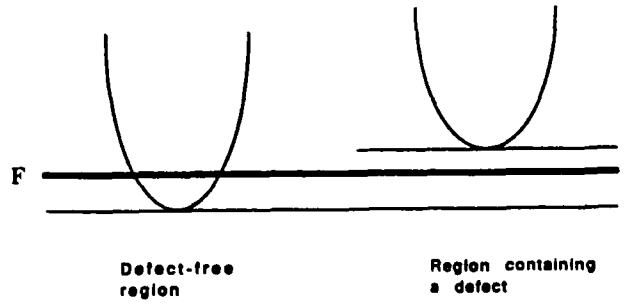


FIG. 2. Energy dispersion parabola for the subband closest to the quasi-Fermi level in a structure containing a potential step due to a defect cluster or interface roughness. Note that the bottom of the parabola goes above the quasi-Fermi level F in the region containing the defect so that the mode goes from propagating to evanescent. This figure elucidates the mode quenching effect.

terferometer is depicted schematically in Fig. 1(a) while Fig. 1(b) shows a possible realization utilizing a GaAs-AlGaAs double quantum well. The interferometer is basically a doubly connected structure consisting of two contiguous conducting paths connected at both ends. Electrons enter at one end and leave at the other. The quantum-mechanical phase shifts in the two paths can be controlled by an electrostatic potential impressed between the paths. Changing the potential modulates the interference between the two paths and therefore the net transmission through the structure. This is turn modulates the conductance causing it to oscillate as the potential is varied.

The feasibility of electrostatic Aharonov–Bohm interferometers have been investigated quite widely both theoretically[8,15,16] and experimentally.[17,18] In this paper, we study the conductance oscillations of these interferometers in the presence of interface roughness scattering or defect clusters. The conductance is calculated using Eq. (1) with the transmission amplitudes obtained from the overall scattering matrix as described in Sec. II. The scattering matrix for electrostatic Aharonov–Bohm interferometers has been described in Refs. 16 and 19.

In Fig. 3(a) we present the conductance oscillations in an ideal interferometer with no defect or interface roughness in either arm. The results are shown for various ambient temperatures. The two arms of the interferometer are each 1 μm long and the quantum wells constituting them are 50 Å wide, the material is GaAs, and the carrier concentration is 1.2×10^{6}/cm. Ten subbands are occupied in either arm at a temperature of 0 K. For this calculation, the elements of the Shapiro matrix were assumed to be different for different modes. We have assumed $\epsilon_i = 0.5 - 0.01i$, where ϵ_i is the free parameter in the Shapiro matrix for the ith mode or the ith subband. A higher value of ϵ for lower subbands mimics the fact that electrons with lower transverse energy and therefore higher momentum along the direction of current flow have a higher probability of transmitting through the structure.

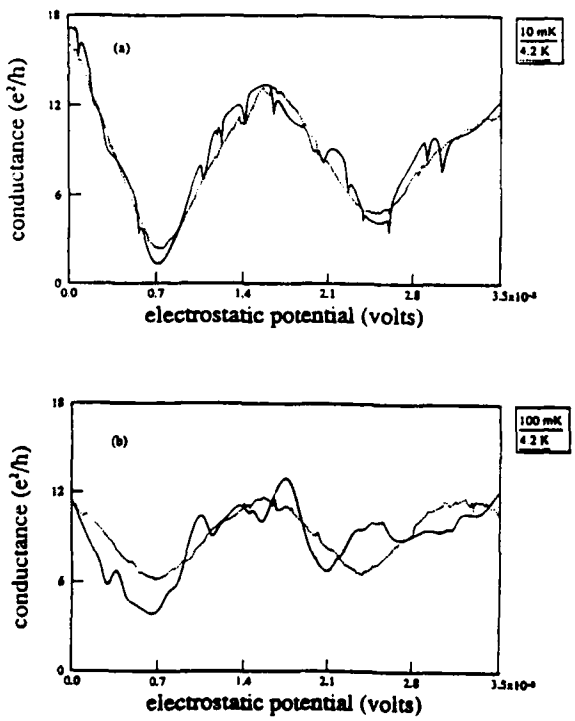


FIG. 3. (a) The electrostatic Aharonov–Bohm conductance oscillations at two different temperatures when the two arms of the interferometer are defect-free. (b) When there are two potential islands in one arm due to two regions of interface roughness.

In Fig. 3(b) we show the conductance oscillations when there are two potential islands in one of the two arms. The islands are each 200 Å long and are situated 2500 and 3500 Å from the left contact. They represent the situation that in these regions, the width of the well has decreased by a monolayer thickness (2.8 Å). This simulates interface roughness.

In Fig. 3(b) we find that the conductance oscillation exhibits sudden bumps at low temperatures. We believe that these are due to an interesting effect caused by the defect clusters. In Sec. II we discussed how a propagating mode could become evanescent and thus get quenched upon entering a region containing a defect cluster. The reverse effect can also occur in an Aharonov–Bohm interferometer. When a potential difference is applied between the two arms by holding the potential of one arm constant and lowering that of the other, an evanescent mode could become propagating as the bottom of the corresponding subband dips below the local quasi-Fermi level. Every such "mode regeneration" adds an amount $2e^2/h$ to the conductance and causes a bump in the oscillation characteristics. The exact position of these bumps in the oscillations will depend on the precise nature and locations of the defect clusters. These bumps can therefore be viewed as fingerprints of the defect configuration in much the same way as universal conductance fluctuations. The occurrence of these bumps will also be sample specific like universal conductance fluctuations since the defect configuration is sample specific.

In addition to the bumps, the conductance oscillations exhibit many kinks and glitches that smear out at elevated temperatures. We believe that this feature has the following origin. The conductance oscillation of a ring due to the electrostatic Aharanov–Bohm effect exhibits two distinct sets of minima arising from two different interference conditions.[16] One set of minima is caused by the usual destructive interference of transmitted electrons and the other arises due to the constructive interference of an electron traveling completely around the ring and interfering with itself at its point of entry into the ring.[16] The positions of the secondary minima in the oscillations are strongly mode dependent. In a multimoded structure, such as the one that we have studied, the varying positions of the secondary minima (for the different modes) impart to the oscillation pattern a random speckled nature that create the appearance of glitches. The secondary minima however bleach out very quickly with increasing temperature[16] so that the glitches disappear at elevated temperature.

It is obvious from Figs. 3(a) and 3(b) that interface roughness can almost wash out the oscillations and make them indiscernible. This is a serious deleterious effect which is somewhat mitigated at elevated temperatures. However, this does not mean that the performance of Aharonov–

Bohm interferometers with interface roughness will improve with increasing temperature, since the conductance modulation also decreases rapidly with increasing temperature.

We now proceed to discuss electron transport in narrow ballistic constrictions. Such a structure is depicted schematically in Fig. 1(c). The structure is usually patterned by electron-beam lithography on a two-dimensional electron gas such as an inversion layer or an accumulation layer. The layer width is assumed to be 100 Å and the background impurity concentration is $1.2\times10^{16}/\mathrm{cm}^3$ everywhere. The width of the constriction is 1250 Å and the wider contact regions are 5000 Å wide. The Fermi energy is 1.43 meV. Eleven subbands are occupied in the contact regions and two in the constriction.

In Figs. 4(a) and 4(b) we show the conductance of the constriction as a function of its length normalized to the deBroglie wavelength of the lowest-lying transverse mode in the wider contact region. This wavelength is approximately 1250 Å. Figures 4(a) and 4(b) show non-self-consistent and self-consistent results, respectively, where self-consistent results are those that account for space-charge effects or the Hartree correction. We have accounted for self-consistence approximately by incorporating a built-in potential (caused by space-charge effects) at the interface of the contact regions and the constriction. The built-in potential was calculated by requiring that the Fermi level be flat at equilibrium and the bulk carrier concentration (or the Fermi energy) be the same everywhere. The built-in potential was found to be -0.10251 meV, which is $\sim 7\%$ of the Fermi energy. The bulk conduction-band profile including the built-in potential is shown in Fig. 1(c).

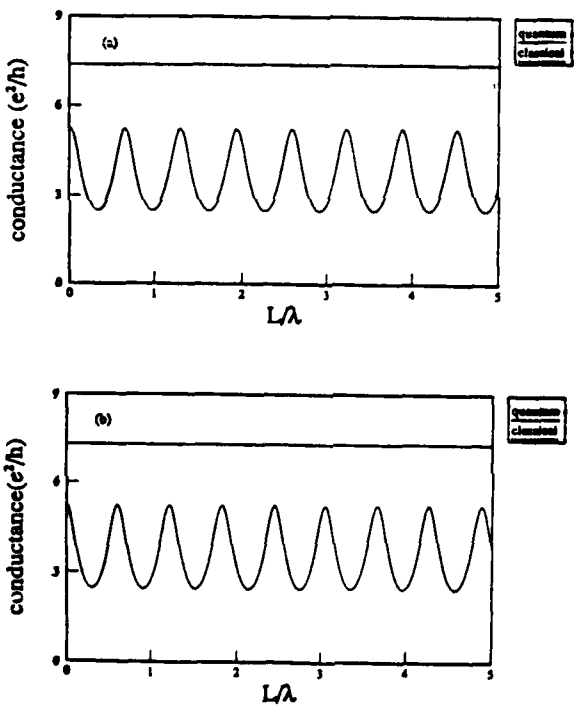


FIG. 4. Conductance of a ballistic constriction as a function of its length normalized to the deBroglie wavelength of the lowest transverse mode in the wide contact region. (a) Non-self-consistent result; (b) self-consistent result.

The broken lines in Figs. 4(a) and 4(b) are the results of a classical calculation obtained by cascading probability scattering matrices rather than amplitude scattering matrices.[8] The classical conductance is found to be independent of the length of the constriction instead of being inversely proportional to the length since transport within the constriction is purely *ballistic*. The only scattering (which gives the conductance the finite value of 7.37 e^2/h) is the scattering that takes place at the interface between the wide and narrow regions. The classical conductance is therefore purely a result (and a measure) of the contact resistance.

There are several interesting features in Figs. 4. First, the conductance calculated quantum mechanically is less than the classical result, and second, it oscillates as a function of the length of the constriction. These are both due to the fact that the constriction acts as a potential barrier to electrons impinging from the wider contact regions. Classically, every electron that arrives with a kinetic energy larger than the barrier height can transcend the barrier and contribute to current, but quantum mechanically there is a nonzero probability of reflection. Because of this reflection, the quantum-mechanically calculated current (or conductance) is always less than the classical result. In addition, the conductance oscillates because of continuum resonances.[20,21] Changing the length of the constriction is equivalent to changing the extent of the barrier which causes the transmission probability to oscillate. Finally, another significant feature to note in these plots is the importance of space-charge effects. It can be seen from Figs. 4(a) and 4(b) that the number of oscillation cycles within the period shown are different depending on whether space-charge effects are included or neglected. The difference is due to the built-in potential discussed before. This potential alters the energy and hence the wave vector of the incoming electrons thereby affecting the period of the oscillations.

IV. CONCLUSIONS

In this paper we have analyzed phase-coherent electron transport through disordered Aharonov–Bohm interferometers in the presence of interface roughness scattering and defect clusters. We pointed out the phenomenon of mode quenching and mode regeneration in these structures that give rise to well-resolved sample-specific bumps in the conductance oscillations. These bumps act as fingerprints of the internal defect configuration. We have also analyzed electron transport through ballistic constrictions and found that space-charge effects may be important in these structures.

ACKNOWLEDGMENTS

One of us (M.C.) acknowledges granting of supercomputing time by the Ohio Supercomputer Center. The work at Notre Dame was supported by the Air Force Office of Scientific Research under Grant No. AFOSR-88-0096 and, in part, by an IBM award.

[1] For a review, see Y. Imry, in *Directions in Condensed Matter Physics*, edited by G. Grinstein and G. Mazenko (World Scientific, Singapore, 1988), pp. 101–163.

[2] For a review, see S. Washburn and R. A. Webb, Adv. Phys. 35, 375

(1986).

[3] J. C. Licini, D. J. Bishop, M. A. Kastner, and J. Melngailis, Phys. Rev. Lett. **55**, 2987 (1985); W. J. Skocpol, P. M. Mankiewich, R. E. Howard, L. D. Jackel, D. M. Tennant, and A. D. Stone, *ibid.* **56**, 2865 (1986).

[4] C. P. Umbach, P. Santhanam, C. van Haesendonck, and R. A. Webb, Appl. Phys. Lett. **50**, 1289 (1987).

[5] B. J. VanWees, H. Van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. Van der Marel, and C. T. Foxton, Phys. Rev. Lett. **60**, 848 (1988); D. A. Wharam, T. J. Thornton, R. Newbury, M. Pepper, H. Ajmed, J. E. F. Frost, D. G. Hasko, D. C. Peacock, D. A. Ritchie, and G. A. C. Jones, J. Phys. C **21**, L209 (1988); For a review, see M. A. Reed and W. P. Kirk, eds., *Nanostructure Physics and Fabrication* (Academic, Boston, 1989), Chap. 6.

[6] D. Miller, R. K. Lake, S. Datta, M. S. Lundstrom, M. R. Melloch, and R. Reifenberger, in *Nanostructure Physics and Fabrication*, edited by M. A. Reed and W. P. Kirk (Academic, Boston, 1989), p. 165.

[7] M. Cahay, S. Bandyopadhyay, M. A. Osman, and H. L. Grubin, presented at the Fourth International Conference on Modulated Semiconductor Structures, Ann Arbor, MI, July 1989 [Surf. Sci. **228**, 301 (1990)].

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

[9] P. F. Bagwell, Phys. Rev. B **41**, 10354 (1990).

[10] S. Bandyopadhyay and M. Cahay, Proceedings of the National Center for Computational Electronics Workshop, Urbana, Illinois, May 1990.

[11] H. R. Frohne and S. Datta, J. Appl. Phys. **64**, 4086 (1988).

[12] B. Shapiro, Phys. Rev. Lett. **50**, 747 (1983).

[13] R. Redheffer, J. Math. Phys. **41**, 1 (1962).

[14] Equation (1) is the two-probe multichannel Landauer formula adapted from P. W. Anderson, Phys. Rev. B **23**, 4828 (1981). This is also the Tsu–Esaki formula in the limit of a vanishingly small bias.

[15] S. Datta, M. R. Melloch, S. Bandyopadhyay, and M. S. Lundstrom, Appl. Phys. Lett. **48**, 487 (1986); S. Bandyopadhyay, S. Datta and M. R. Melloch, Superlattice and Microstructures **2**, 539 (1986); S. Bandyopadhyay and W. Porod, *ibid.* **5**, 239 (1989); S. Bandyopadhyay and W. Porod, Appl. Phys. Lett. **53**, 2323 (1988).

[16] M. Cahay, S. Bandyopadhyay, and H. L. Grubin, Phys. Rev. B **39**, 12 989 (1989).

[17] S. Washburn, H. Schmid, D. Kern, and R. A. Webb, Phys. Rev. Lett. **59**, 1791 (1987).

[18] P. G. N. de Vegvar, G. Timp, P. M. Mankiewich, R. Behringer, and J. Cunningham, Phys. Rev. B **40**, 3491 (1989).

[19] M. Cahay, Ph.D. thesis, Purdue University, 1987.

[20] Y. Zohta, J. Appl. Phys. **57**, 2334 (1985).

[21] The presence of continuum resonances in the conductance modulation of quantum point contacts have also been investigated while varying the position of the Fermi level and the width of the constriction: A. Szafer and A. D. Stone, Phys. Rev. Lett. **62**, 300 (1989); G. Kirczenow, Solid State Commun. **68**, 715 (1988); J. Phys. Condens. Matter **1**, 305 (1989); Phys. Rev. B **39**, 10 452 (1989).

Transmission through a bend in an electron waveguide

Craig S. Lent
Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556

(Received 10 January 1990; accepted for publication 10 April 1990)

The transmission properties of a circular, right-angle bend in a two-dimensional electron waveguide are calculated. Significant reflections from such a bend would have serious consequences for the development of a quantum electron waveguide technology. The results show nearly perfect transmission around the bend, except for energies very close to the threshold for propagation in the channel. This is true even for rather sharp bends. A significant amount of mode mixing is found, however, for bends with a small radius of curvature.

Interest in the creation of nanometer-scale quantum devices has focused attention on the problem of the conduction properties of quantum channels in which electrons are confined in two spatial dimensions. Typically electrons are confined in one dimension by the self-consistent potential at a single semiconductor heterojunction interface or by the quantum well potential formed by two interfaces. The resulting two-dimensional electron gas (2DEG) can be confined further by imposing a potential variation in the plane of the electrons. This is often done by depositing a patterned metal gate to which a negative voltage is applied. The gate electrodes deplete the area beneath them, leaving only conducting channels in the 2DEG.[1,2]

Such constricted geometries are often thought of as electron waveguide structures in analogy with microwave devices. If the active region is small enough, electron scattering can be neglected and electronic transport in the region becomes completely ballistic and coherent. In this regime, steady-state transport can be described simply by the time-independent effective-mass Schrödinger equation.

The possibility of transistor action by purely coherent electron interference effects in waveguide structures has been pointed out by Sols *et al.*[2] They examined the properties of a stub-tuner type transistor, a variation of which has been fabricated by Datta and co-workers.[3] The author has pointed out the existence of transistor action in transmission through a constriction in a quantum waveguide.[4] Barker[5] has discussed many interesting electron waveguide geometries and their possible uses in novel device structures. Although exploration of this area is very new and many obstacles remain, it may prove possible to construct an integrated circuit technology based on the interference properties of electrons confined in these waveguide structures.

In contemplating a quantum device technology in which many nanoscale devices are interconnected, one is led to consider the feasibility of interconnections in which a quantum channel bends around a curve to connect two devices. If such a bend were to cause significant reflections, the interconnections would have to be considered as part of the devices themselves, greatly complicating the design and understanding of nanocircuits. In the worst case, one would be forced to avoid bends entirely, limiting the design to linear arrays only. Clearly, a bend with a sufficiently large radius of curvature would be expected to cause very little trouble. If the required radius were very large, however, the integrated designs would be severely constrained. This would be particularly problematic given that the entire active region (or subregion) of the structure must be smaller than an electron mean free path.

I report here a calculation of the transmission properties of circular, right-angle bends in quantum channels. The results are encouraging in that for even very tight corners very little reflection is seen. A close analysis reveals that mode mixing can be important.

The quantum waveguide problem is formally similar to the problem of microwave propagation in conducting waveguides. For a straight channel, both can be reduced to a scalar Helmholtz equation. However, the boundary conditions for the quantum wave function are analogous to those of TM modes in rectangular waveguides whereas it is the TE_{10} electromagnetic mode (which corresponds to different boundary conditions) which is the dominant mode. The electromagnetic problem of transmission around a circular bend is further complicated by the vector nature of the electric and magnetic fields.

The quantum calculation is done within a parabolic single-band effective-mass model with $m^* = 0.05m_0$. Electrons are assumed to be completely confined in the plane of the 2 DEG with only the ground-state wave function in the perpendicular direction accessible. The potential energy profile of the quantum channel depends on exactly how it is realized and is usually assumed to be either parabolic or a square well.[6] Here it is taken to be a square well of width d, with $V = 0$ inside the channel and V infinite outside. The bend itself is a circular arc with central radius r.

I consider here only waves incident on the bend which are in the fundamental mode of the channel. Reflected and transmitted waves may be in higher modes.

The wave function is calculated numerically by discretizing the Schrödinger equation in the region of the waveguide bend. A 41×41 nonrectangular mesh is used here. The boundary conditions are given by the requirements that the wave function vanish at the waveguide walls and that it match the allowed solutions in the leads at the input and output boundaries. Since the reflection and transmission coefficients are unknowns which depend on the solution in the interior region this is not a trivial procedure. Inclusion of these kinds of boundary conditions is accomplished using the quantum transmitting boundary method (QTBM) of Lent and Kirkner, described in detail elsewhere.[7] The com-

 0003-6951/90/252554-03$02.00 © 1990 American Institute of Physics

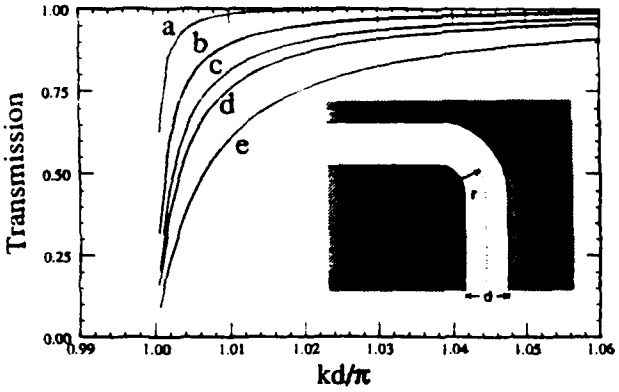


FIG. 1. Transmission coefficients for circular bends in a quantum waveguide with differing radii of curvature. The transmission is shown for $r/d = 2.0$ (a), 1.0 (b), 0.75 (c), 0.65 (d), and 0.5 (e). The inset shows the problem geometry.

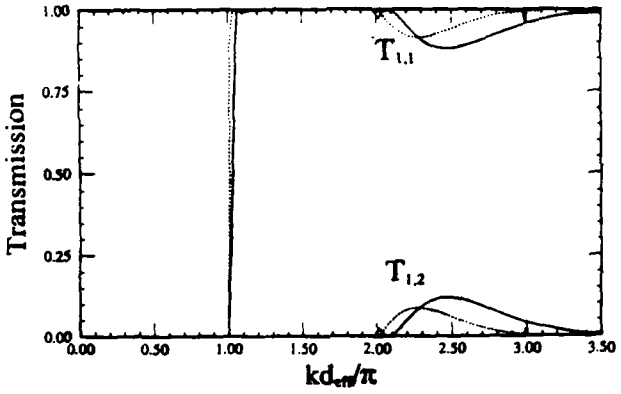


FIG. 3. Transmission coefficients of the circular bend with $r/d = 2.0$ for an incident wave in the fundamental mode of the waveguide. The solid line corresponds to the case of infinite potential barriers forming the channel walls. The dotted line corresponds to finite barriers of height $48E_1$, where E_1 is the energy of the fundamental channel mode. For the finite barrier case, and effective width d_{eff}, is defined by the condition that $kd_{eff}/\pi = 1$ at the cutoff energy.

plete solution yields the wave function in the region of the arc, and the transmission coefficients $T_{1,j}$, the ratio of the outgoing current in the jth mode to the incoming current in the first mode.

Figure 1 shows the transmission coefficient $T_{1,1}$ for energies just above the threshold where the fundamental mode can propagate. The transmission is shown as a function of kd/π, where $k = \sqrt{2m^*E/\hbar^2}$. Since all quantities in the problem scale with the de Broglie wavelength of the electron, plotting the transmission coefficients in this way gives results which are independent of the channel width d and depend only on r/d. The first mode of the channel can propagate for $kd/\pi > 1$, the second mode propagates for $kd/\pi > 2$, and so forth. Figure 1 shows the transmission coefficient for $r/d = 0.5$, 0.65, 0.75, 1.0, and 2.0.

The most striking feature of Fig. 1 is how rapidly the transmission function approaches unity, even for bends with a very small radius of curvature. Transmission is greater than 90% at values of kd/π only very slightly above the threshold for transmission. Notice that $r/d = 0.5$ corresponds to a bend with inner radius of zero, i.e., a sharp right angle. Even this abrupt turn produces little reflection except at energies just above the threshold. As is evident in the figure, a central radius of twice the channel width produces a very sharp transmission rise to essentially unity over a very small energy range. For design considerations, this seems encouraging.

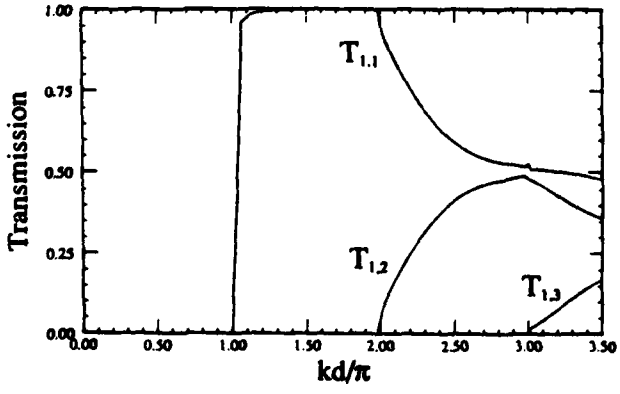


FIG. 2. Transmission coefficients of the circular bend with $r/d = 0.65$ for an incident wave in the fundamental mode of the waveguide.

At energies well above the threshold for the first mode, transport in higher order modes becomes possible. At these higher energies, an incoming wave which is purely in the first mode can be mixed into higher modes by the bend. Figure 2 shows the transmission coefficients from the first mode into the first three modes for $r/d = 0.65$. Clearly a significant amount of mode mixing is possible at energies which allow multimode transport. As one would expect, at larger radii of curvature, the mixing diminishes. Figure 3 shows the transmission from the fundamental mode into the first two modes for a bend with $r/d = 2.0$. Transmission into the second mode is greatly reduced and transmission into the third mode, while possible at the energies shown, is negligible.

Real waveguide bends would have a more complicated potential than the infinite square well assumed in this model. The channel walls would allow some penetration of the wave function into the potential barrier. The simplest model which includes this softening is a square well channel with finite barriers. Figure 3 also shows the results for transmission around a circular bend with $r/d = 2.0$ and a confining potential of $48\,E_1$, where E_1 is the energy of the lowest transverse mode. These results suggest that softening of the walls reduces reflection and mode mixing.

In summary, the transmission properties of a circular bend in an electron quantum waveguide have been calculated. Reflections from the bend, which could make the design of integrated quantum waveguide devices difficult, are found to be negligible except in an energy region very close to the threshold for traveling waves in the channel. At energies high enough for multimode transport, a bend can introduce significant mode mixing. This can be minimized by choosing bends with radii a few times the channel width.

The work was supported by the Air Force Office of Scientific Research under grant number AFOSR-88-0096. This work was partially supported by National Science Founda-

tion under grant number ECS890025 through the National Center for Computational Electronics, and utilized the Cray-2 at the National Center for Supercomputing Applications, University of Illinois at Urbana–Champaign.

Note added in proof: We have recently learned of a similar study by Sols and Macucci to appear in Physical Review B.

[1]G. Bernstein and D. K. Ferry, J. Vac. Sci. Technol. B 5, 964 (1987).
[2]Fernando Sols, M. Macucci, U. Ravaioli, and Karl Hess, Appl. Phys. Lett. 54, 350 (1989).
[3]S. Datta, Superlatt. Microstruct. 6, 83 (1989).
[4]Craig S. Lent, Srinivas Sivaprakasam, and David J. Kirkner, Solid-State Electron. 32, 1137 (1989).
[5]J. R. Barker, in *Nanostructure Physics and Fabrication*, edited by M. A. Reed and W. P. Kirk (Academic, Boston, 1989), p. 253.
[6]G. Kirczenow, Solid State Commun. 68, 715 (1988).
[7]Craig S. Lent and David J. Kirkner, J. Appl. Phys. 67, 6353 (1990).

To appear in the Proceedings of the
SPIE Nanostructures Conference
March 1990, San Diego

Quantum Electron Waveguides: Bends, Constrictions, and Cavities

Craig S. Lent and Srinivas Sivaprakasam

Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

ABSTRACT

We examine transport through several quantum electron waveguide geometries. Reflection and mode-mixing in transmission through bends in quantum waveguides are calculated and the impact on interconnections between devices discussed. We calculate the current-voltage relationship for a constriction in a quantum waveguide for applied voltages beyond linear-response regime. Strong resonance effects in waveguide cavities are found to persist even when cavity geometry is non-abrupt. We demonstrate the existence of current vortices in resonant cavities.

1. Introduction

There is ample experimental evidence that a two-dimensional electron gas (2DEG), created at a heterojunction interface can be further confined electrostatically by depositing a metal gate pattern above the interface [1,2]. The electrons are apparently ballistic and coherent over regions as large as a micron or more. By lithographically patterning the metal gate, electrons can be confined to travel in narrow quantum channels which act much like microwave waveguides. Since the nanometer-scale confinement geometry can be controlled by the voltage applied to the gate, many device applications are possible for such structures.

We explore some of these possibilities by calculating the transmission properties of several waveguide structures. Our approach is to solve the two-dimensional effective-mass Schrödinger equation directly in real space. Throughout we use an effective mass of $m^* = 0.05\ m_0$ and assume that the confinement in the direction perpendicular to the plane of the 2DEG is complete. We use a Finite Element discretization scheme and current-carrying boundaries are

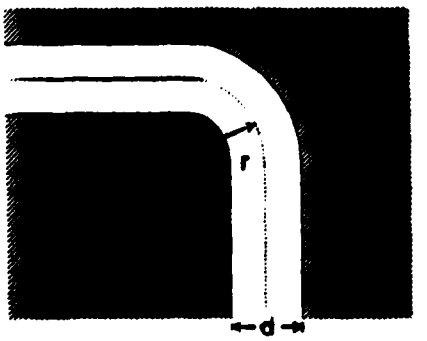


Fig. 1. Geometry of a right-angle circular bend in a quantum waveguide.

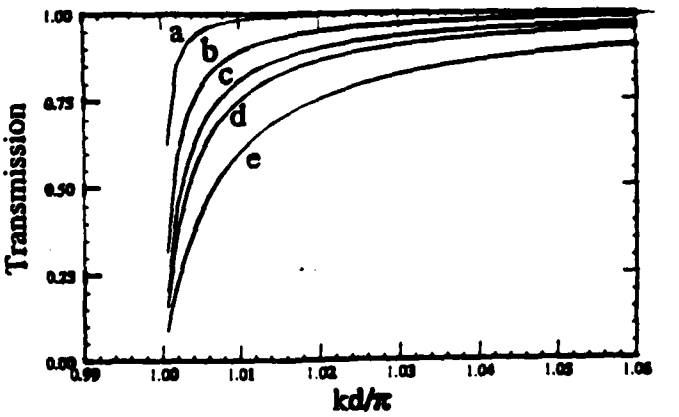


Fig. 2. Transmission coefficients for circular bends in a quantum waveguide with differing radii of curvature. The transmission is shown for $r/d = 2.0$ (a), 1.0 (b), 0.75 (c), 0.65 (d), and 0.5 (e).

included using the Quantum Transmitting Boundary Method [3].

In the next section we discuss transmission through bends in quantum waveguides, a question relevant for the problem of making interconnections between electron waveguide devices. Section 3 explores the transport through a constriction in a waveguide when a finite voltage is applied. Section 4 examines resonant cavities in quantum waveguides and the appearance of vortex excitations of the current through a cavity.

2. Bends

The construction of large-scale integrated nanoelectronics will require a basic circuit element, the role played by field-effect and bipolar junction transistors presently, and an architecture for connecting the basic elements in a useful way. While the choice for neither of these is clear today, it is certain that nanometer-scale devices will need to be connected to each other. The question of the transmission properties of bends in quantum waveguides becomes immediately important. If the physics of ballistic transport causes significant reflections in all but the most gentle of bends, design of an integrated technology will be seriously constrained. Most proposed quantum interference devices require single-moded transport in the channels for optimal effectiveness. Mode-mixing caused by bends also becomes an important consideration.

With this in mind, we have examined the transmission characteristics of circular bends in quantum waveguides [4]. The geometry of the bend is shown in Figure 1. The waveguide has a width d, which is maintained around the bend. The walls are taken to be infinite potential barriers with no potential in the channel itself. The radius of curvature of the center of the channel is r. We plot the transmission coefficient for the fundamental mode of several waveguide bends as a function of the the scaled energy in Figure 2. The plot is independent of d if we plot the transmission versus kd/π where $k = \sqrt{2m^*E}/\hbar$. The mth traveling mode of the waveguide becomes accessible at an energies corresponding to $kd/\pi \geq m$. The plot shows the energy region just above the cutoff for the first mode. The transmission coefficient is shown for several values of r/d ranging from 2.0 to 0.5. The case of $r/d = 0.5$ corresponds to a bend with an inner radius of curvature of 0, and therefore represents the most abrupt limit of a circular bend. Even for this extreme case, the transmission rises rather rapidly, reaching nearly 90% when $kd/\pi = 1.06$. For more gentle curves, of course, the turn-on is more rapid. For the case when $r/d = 2.0$, the transmission is essentially unity for kd/π just 2% above the cut-off.

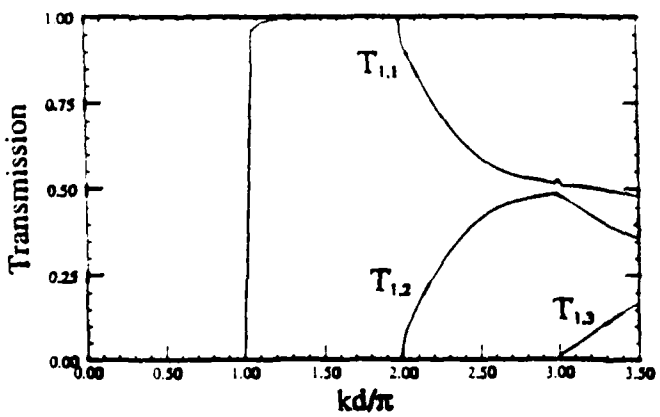


Fig. 3. The transmission coefficients of the circular bend with $r/d = 0.65$ for an incident wave in the fundamental mode of the waveguide.

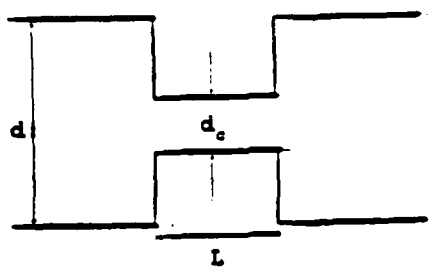


Fig. 4. Geometry of constriction in an electron waveguide.

As the energy increases to the point where the second waveguide mode becomes accessible, mode-mixing begins to occur as the wave traverses the bend. Figure 3 illustrates the transmission from mode 1 in the incoming channel into the first three modes of the outgoing channel for the $r/d = 0.65$ case. Mixing into the second mode is reduced to about 10% for $r/d = 2.0$

Our results indicate that circular bends in waveguides probably do not impose serious constraints on interconnections. Clearly, some care has to be taken to avoid mode-mixing if single-moded operation is desirable. That would most easily be achieved by keeping the injected electrons below the energy threshold for multi-moded transport.

3. Constrictions

Ballistic transmission through structures narrow enough to be called electron waveguides was first observed and has been widely studied in the "split-gate" geometry [1,2]. A narrow quantized channel connects two regions which are both two-dimensional in character. We focus our attention here on a related, but different geometry, a constriction in the narrow channel itself [6]. This is also referred to as the "wide-narrow-wide" geometry and has recently been studied experimentally [5]. The problem geometry is shown in Figure 4. The channel of width d is narrowed to $d_c < d$ for a length L and then widens again to its original width. For simplicity we take the transition from wide to narrow regions to be abrupt. We also have assumed infinitely hard side-wall barriers.

The constricted region presents a larger ballistic resistance to current flow. It is possible, therefore to maintain a relatively large voltage difference between the two sides of the constriction. If the voltage drop across the constriction is large compared to kT/e, the transport can no longer be described by linear response theory. Landauer theories in which the conductance is proportional to the transmission coefficient evaluated at the Fermi energy [7,8,9] are not appropriate in this regime. To calculate the current-voltage relationship we use a modified Esaki formula.

$$I(V) = \left(\frac{2e}{h}\right) \int [f(E - E_F - eV) - f(E - E_F)] \sum_{i,j} T_{i,j}(E,V) dE. \tag{1}$$

In this equation $T_{i,j}(E,V)$ is the transmission coefficient from mode i in the left channel into mode j in the right channel at an energy E, when a voltage difference of V is applied between the two wider regions. The functions $f(E)$ are Fermi-Dirac distribution functions. We approximate the effect of the applied voltage by assuming a linear potential drop across the constriction. No bending of the bands in the wider lead regions is included. This

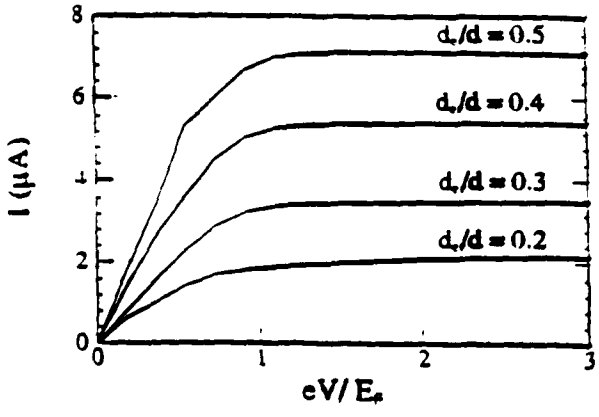


Fig. 5. The current-voltage characteristic for various constrictions sizes.

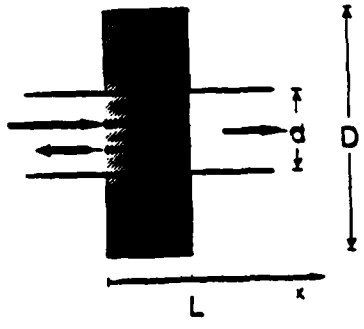


Fig. 6. Idealized cavity in a quantum waveguide.

approximation is reasonable because the resistance of the narrow regions is much higher than the the unconstricted channels.

We examine a geometry for which L/d=0.25 and we vary d_c/d between 0.2 and 0.5. We choose the Fermi level in the leads to be $E_F = 35\ E_1$, where E_1 is the energy of the lowest channel mode. Five incoming modes are then occupied, although the contribution of the fifth mode is very small. At energies corresponding to occupied states, we calculate the transmission coefficients required in equation(1) and integrate to obtain the current. The current voltage relationship obtained for various constriction widths is shown in Figure 5.

The most dramatic feature of these results is the saturation of the injected current at voltages greater than the Fermi voltage. This can be understood by examining Equation (1). At low voltages the dominant energy dependence in the integral is the difference of the Fermi functions. Variations in the transmission coefficients tend to be smeared out by the sum over the modes. For small voltages, the difference in the Fermi functions is a linear function of the voltage. This is simply the linear-response result. As the voltage increases beyond V_F, the contribution of the second Fermi function becomes negligible and the first term in the integral becomes independent of voltage. The voltage dependence is then due primarily to the transmission coefficients. However, as the voltage drop becomes steeper, the transmission coefficients tend to saturate at values proportional to the geometrical ratio d_c/d. Thus the total current saturates at a value determined by the constricted width. Since the width of the constriction can be controlled electrostatically, this geometry may have potential device applications.

4. Cavities

Two nanometer-scale quantum interference transistors have been proposed which are both based on the analogy between microwave waveguides and electron waveguides. The microwave stub-tuner design has been adapted to electron waveguide structures in somewhat different ways by Sols *et al.* [12], and Datta [10]. Realization of the structure proposed by Datta has recently been reported [11] Each device employs a resonant cavity which modulates the transmission through a waveguide channel. For abrupt potentials with infinitely hard walls (like the potentials discussed above), the transmission coefficient vanishes for certain energies [10,12]. By electrostatically controlling the dimensions of the resonant cavity, transistor action is obtained.

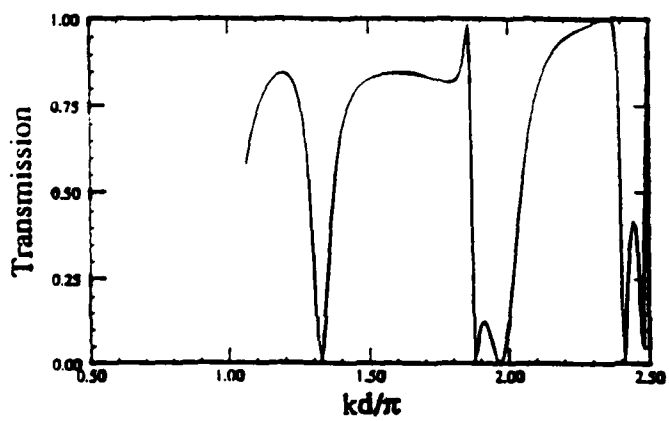


Fig. 7. Transmission coefficient for the first waveguide mode through the cavity depicted in Figure 6. The transmission is shown as a function of kd/π where $k = \sqrt{2m^*E}/\hbar$.

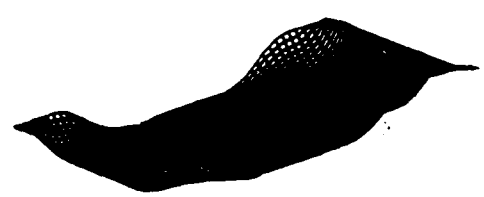


Fig. 8. Model potential for a resonant cavity in an electron waveguide.

In this context, we examine the phenomenon of transmission through resonant cavities in quantum electron waveguides. We consider a symmetric structure, the geometry of which is illustrated in Figure 6. The transmission coefficient is shown for a cavity with hard walls in Figure 7. We focus here on three questions concerning transport in such structures.

1. Does transmission through more realistic, rounded potentials display the same sharp vanishing of the transmission coefficient at resonances as is obtained in the idealized structure?
2. How does the charge in the cavity change as the electron wave moves on and off resonance?
3. How does the presence of the cavity affect the flow of current in the cavity region?

Figure 8 illustrates a symmetric cavity with a less abrupt geometry than the idealized case [14]. In the incoming leads the potential is zero across a width t and then increases quadratically. This form of the potential profile across a waveguide was suggested by the calculations of Stern, Frank and Laux [15], and the experiments of Wharam *et al.* [16]. In the cavity itself, the potential is zero in the center region and has quadratically increasing walls. The corners have also been rounded. This potential is not the result of a self-consistent calculation, but is rather a model which mirrors the important features of the actual potential.

The Schrödinger equation is solved first in the leads to obtain the lead eigenfunctions. These are then used in the construction of the boundary conditions for the problem of transmission through the cavity using the QTBM.

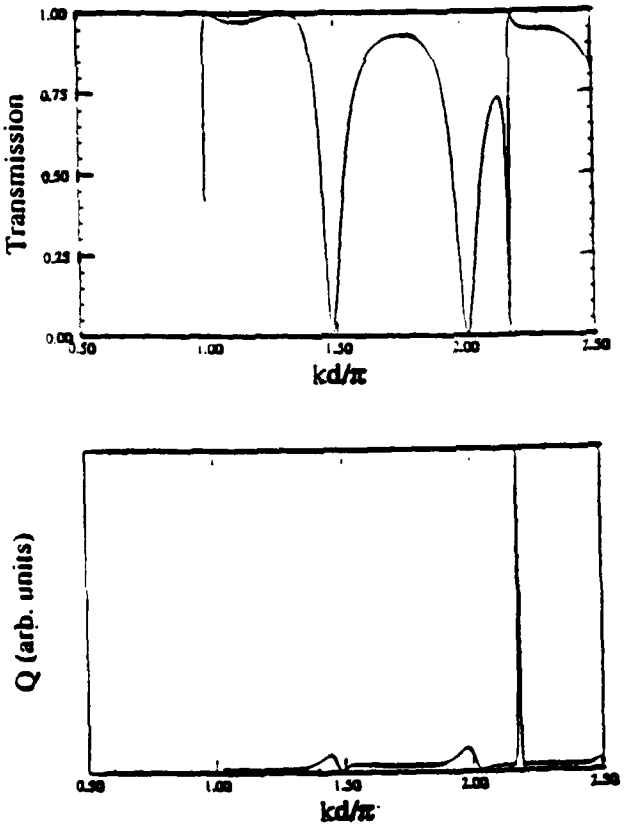


Fig. 9. The upper graph is the transmission coefficient for the first waveguide mode through the cavity depicted in Figure 8. The lower graph shows Q, the integral over the cavity of the electron probability density as a function of the scaled energy.

We choose cavity dimensions such that $L = 2t$, $D/t = 3$. The lowest mode in the channel is localized to a width of approximately $1.4t$ which we call d, the effective channel width (which now varies for different modes). The calculated transmission coefficient for the $m = 1$ mode incident on the cavity is shown in Figure 9(a). The features are certainly qualitatively similar to the hard-wall case shown in Figure 7. At resonance conditions the transmission coefficient vanishes. The answer to the question (1) above is that the phenomenon of complete quantum reflection at certain resonance conditions is preserved, even when the corners of the cavity are rounded and the walls are softened.

At each injection energy, we calculate a quantity Q defined by

$$Q = \int_{\text{cavity}} |\psi(x,y)|^2 dx\, dy \tag{2}$$

which is proportional to the total electronic charge in the cavity. Figure 9(b) shows the Q for various injection energies. At resonance conditions the charge in the cavity is enhanced. The extent of the enhancement is different for different resonance states. Over the range shown, the greatest increase in Q occurs at a transmission peak. Such peaks are related to the peaks in transmission over a one-dimensional well, at which energies the probability density in the well also increases. Further discussion of the nature of the resonances is presented elsewhere [18]. From this example it is clear that modulation of the charge in the cavity accompanies modulation in the transmission through the cavity. If the magnitude of the charge stored in the cavity is sufficient, it may be able to provide electrostatic modulation of nearby quantum structures, thus permitting a direct coupling between nanometer-scale devices.

Since we solve the Schrödinger equation for the wavefunction in the cavity region, we can calculate the current

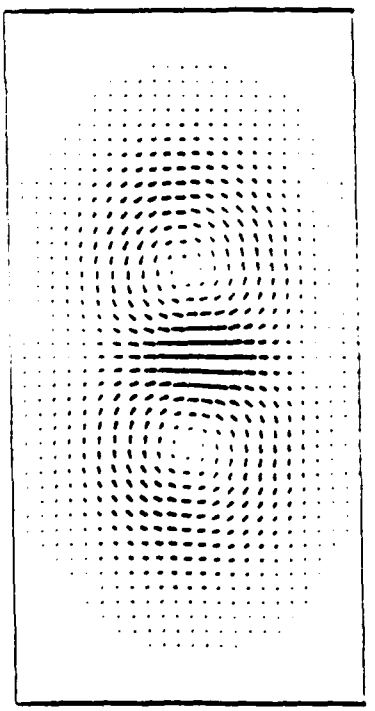


Fig. 10. The current density in the cavity for an energy such that $kd = 1.478$. This corresponds to the first zero in the transmission shown in Figure 9.

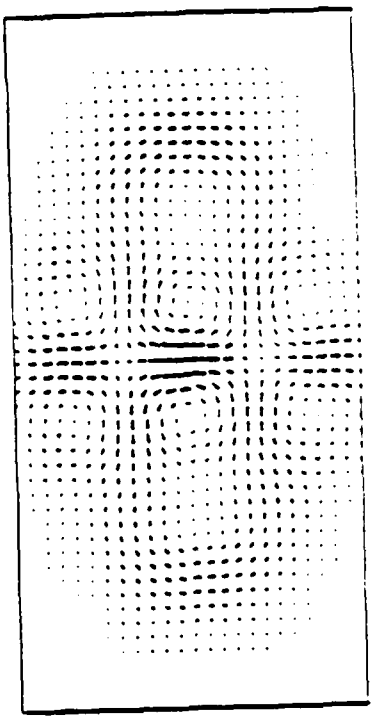


Fig. 11. The current density in the cavity for an energy such that $kd = 2.187$. This corresponds to a maximum in the transmission shown in Figure 9.

density directly from the wavefunction. The electron particle current density is given by,

$$\vec{j}(x,y) = \frac{i\hbar}{2m^*}\left(\psi(x,y)\vec{\nabla}\psi^*(x,y) - \psi^*(x,y)\vec{\nabla}\psi(x,y)\right). \tag{3}$$

Figure 10 shows the current density in the cavity for an energy corresponding to the first minimum (zero) in the transmission. Figure 11 illustrates the current density in the cavity for the maximum in transmission which occurs at $kd = 2.187$. Current vortices are excited in the cavity by the incoming wave. The vorticity observed here is distinct from hydrodynamic vortex formation in that there are no interparticle interactions in the model. Vortex formation in this system occurs in the context of a one-electron model. Perhaps the closest analogy is with vortex formation in superfluids [17].

The examples of vortex formation shown in Figures 10 and 11 display some features unique to this ballistic electron system. Note that for the case shown in Figure 10 the transmission coefficient is zero. The *net* current in both input and output leads is zero. The incoming wave is totally reflected by the cavity. The current enters the cavity, swirls around, and exits through the input lead. The vortices could not be excited if there were no incoming wave, of course, but the wave may be totally reflected. The case depicted in Figure 11 shows a situation where reflections by the cavity are minimal. Note, however, that at the center of the cavity the current is actually going in the reverse direction – toward the input lead. The primary current paths split on either side of the cavity center and merge again on the output side.

The excitation of current vortices in the cavity region is a feature of ballistic transport not limited to special resonance conditions. The rounded features of this model potential enhance vortex formation but are not essential – the abrupt, hard-wall potentials also show vortex excitations. Vortices need not occur in vortex-antivortex pairs if the potential is not symmetric. Very general arguments can be made to show vortex excitations are a basic feature of the two-dimensional electron gas system. A breaking of the time-reversal symmetry, in this case provided by the injected current, is required to expose them [18].

This work was supported by the Air Force Office of Scientific Research and by the National Science Foundation under grant number ECS890025 through the National Center for Computational Electronics, and utilized the Cray-2 at the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign.

References

[1] B. J. Van Wees, H. van Houten, C. W. J. Beenakker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marel, and C. T. Foxon, Phys. Rev. Lett. 60, 848 (1988).

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

[3] C.S. Lent and D.J. Kirkner, J. Appl. Phys. (accepted for publication).

[4] C.S. Lent submitted to Appl. Phys. Lett.

[5] R. J. Brown, M. J. Kelly, M Pepper, H Ahmed, B. Miller D. G. Hasko, D. C. Peacock, J. E. F. Frost, D. A. Ritchie and G. A. C Jones, Solid State Electronics 32, 1179 (1989).

[6] Craig S. Lent, Srinivas Sivaprakasam, and David J. Kirkner, Solid State Electronics 32, 1137 (1989).

[7] R. Landauer, IBM J. Res. Develop. 1, 223 (1957).

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

[9] A.D. Stone and A. Szafer, IBM J. Res. Develop. 32, 384 (1988).

[10] S. Datta, Superlattices and Microstructures 6, 83 (1989).

[11] S. Datta, in M. A. Reed and W. P. Kirk (Eds.) *Nanostructure Physics and Fabrication*. (Academic Press, Boston, 1989)

[12] F. Sols, M. Macucci, U. Ravaioli, and K. Hess, Appl. Phys. Lett. **54**, 350 (1989).

[13] F. Sols, M. Macucci, U. Ravaioli, and K. Hess, J. Appl. Phys. **66**, 3892 (1989).

[14] Craig S. Lent, Srinivas Sivaprakasam, and David J. Kirkner, in M. A. Reed and W. P. Kirk (Eds.) *Nanostructure Physics and Fabrication*. (Academic Press, Boston, 1989).

[15] S.E. Laux, D.J. Frank, and Frank Stern, Surface Sci. **196**, 101 (1988).

[16] D.A. Wharam, U. Ekenberg, M. Pepper, D.G. Hasko, H. Ahmed, J.E.F. Frost, D.A. Ritchie, D.C. Peacock, and G.A.C. Jones, Phys. Rev. B **39**, 6283 (1989).

[17] R.P. Feynman, *Statistical Mechanics* (Benjamin/Cummings, Reading, Mass., 1972).

[18] C.S. Lent (unpublished).

To appear in the Proceedings of the
National Center for Computational
Electronics 1990 Workshop

CALCULATION OF TRANSPORT THROUGH BALLISTIC QUANTUM STRUCTURES

Craig S. Lent
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

Abstract

Coherent ballistic transport through a circular resonant cavity in a quantum electron waveguide is examined. The circular cavity is weakly coupled to a rectangular channel. The current density in the cavity region is calculated from the solution to the effective-mass Schrödinger equation. Current vortex formation in the cavity and a related increase in probability density in the the cavity occur at certain resonant energies.

Electrons can be trapped in the narrow quantum well formed at the heterojunction interface between two semiconductors with differing bandgaps. The resulting two-dimensional electron gas (2DEG) can have a very high mobility. It is possible to confine the electrons further by making some areas of the plane have a higher potential than others. A common technique for imposing this additional potential is to deposit a lithographically defined metal pattern on the surface of the semiconductor. When a negative voltage is applied to the metal, the regions immediately beneath the metal become potential barriers for the electrons and are depleted of carriers. Electron channels which function as electron waveguides, in analogy with microwave

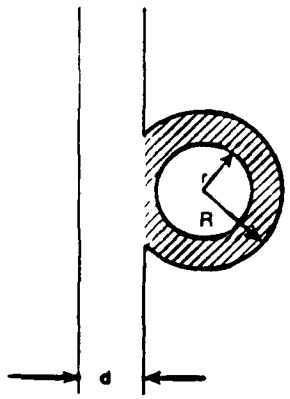


Figure 1: Geometry for the circular resonant cavity.

waveguides, can be formed in this way. More complicated potentials in the plane of the 2DEG can be created by suitably patterning the metal gate layer.

Two designs for creating a quantum-interference based transistor have recently been proposed. Sols *et al.* proposed a stub-tuner transistor consisting of a rectangular stub on a straight electron channel [1]. The wave traveling down the channel resonates with the stub cavity and produces strong maxima and minima in the transmission through the device. The modulations in the transmission become modulations in the current through the device. Since the length of the stub can be controlled electrostatically, it can act as a gate. Datta has discussed a variant of this design [2].

The nature of the resonances in such devices has not been fully investigated. The energies at which the resonances occur are not simply related to the resonances of bound states of the stub region. Elsewhere we have shown that essential in understanding these resonance effects is the recognition that they are associated with the formation of ballistic current vortices in the cavity (stub regions) [3]. Here we examine a resonant structure which is designed to maximize these vortex effects.

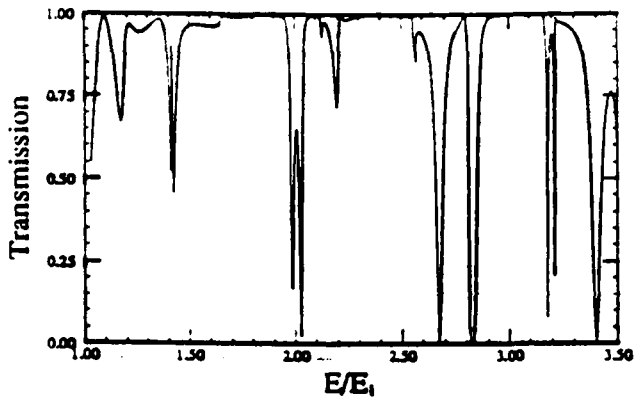


Fig. 2. Transmission through the circular cavity.

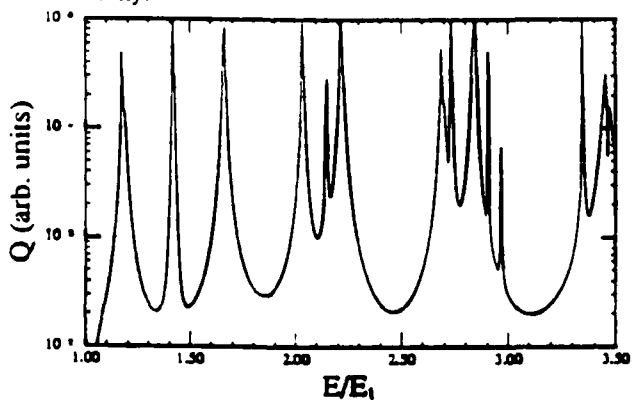


Fig. 3. Charging in the circular cavity.

We consider transmission through a rectangular quantum channel which is coupled through a barrier region to a circular cavity. The geometry is illustrated in Figure 1. The channel has width d and first-mode energy E_1. The potential is zero inside the channel and in the center section of the cavity. The walls of the cavity increase quadratically from zero at the inner radius r to $26E_1$ at the outer radius, R. The channel edge truncates this potential and a finite potential barrier is formed between the channel and the circular cavity. Electrons are incident from one end of the channel in the fundamental channel mode. For the range of energies considered here, only one propagating mode can exist in the channel.

The effective-mass Schrödinger equation is solved numerically over the device region using the Finite Element Method. Boundary conditions appropriate to traveling-wave states are included using the Quantum Transmitting Boundary Method [4]. This technique has been employed in the study of other quantum structures [5,6]. The calculated transmission coefficient is shown in Figure 2. To examine the build-up of charge in the cavity itself, we define a quantity Q which is proportional to the total cavity charge.

$$Q \equiv \int_{\text{cavity}} |\psi|^2 dx\, dy. \tag{1}$$

Figure 3 shows the increase in Q by more than two orders of magnitude which accompanies the resonances. Because of this charging of the cavity, further investigations will need to include self-consistent effects.

More information about the nature of the resonances can be gained by considering the current density. From the wavefunction, the ballistic probability current density can be computed using the relation

$$\vec{j} = \frac{i\hbar}{2m^*}\left[\psi\vec{\nabla}\psi^* - \psi^*\vec{\nabla}\psi\right]. \tag{2}$$

Both the current density and the probability density are shown in Figures 4 and 5 for the second resonance. The third resonance is depicted in Figure 6 and 7. The

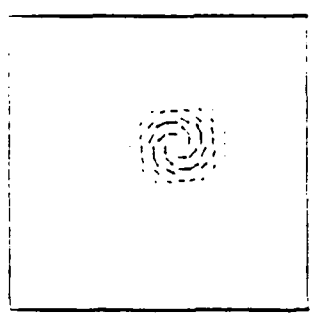


Fig. 4. Current density for the $E = 1.42E_1$ resonance.

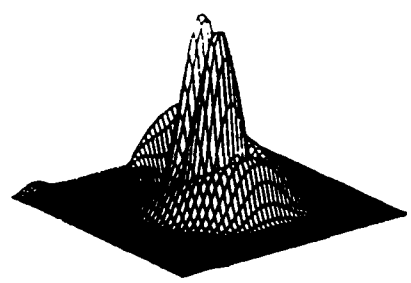


Fig. 5. Probability density for the $E = 1.42E_1$ resonance. For this resonance $(n, m) = (2, 1)$.

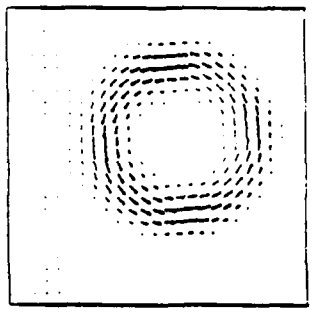


Fig. 6. Current density for the $E = 1.66E_1$ resonance

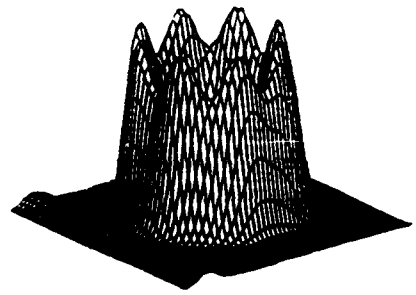


Fig. 7. Probability density for the $E = 1.66E_1$ resonance. For this resonance $(n, m) = (1, 4)$.

large current vortex formed in the cavity is apparent. The direction of current flow in the cavity is the same for all energies.

In the cavity, the wavefunction at these resonances is approximately that of the cavity in isolation. The resonances can be labeled by quantum numbers n and m which denote the number of radial nodes and the orbital angular momentum respectively. For higher energies, resonances occur which are associated with coupled cavity-lead excitations. These are not directly related to the cavity modes and are typically composed of many small vortices in the cavity and channel.

This work was supported by the Air Force Office of Scientific Research, an IBM Faculty Development Award, and by the National Science Foundation under grant number ECS890025 through the National Center for Computational Electronics, and utilized the Cray-2 at the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign.

References

[1] F. Sols, M. Macucci, U. Ravaioli, and K. Hess, Appl. Phys. Lett. 54, 350 (1989); F. Sols, M. Macucci, U. Ravaioli, and K. Hess, J. Appl. Phys. 66, 3892 (1989).

[2] S. Datta, Superlattices and Microstructures 6, 83 (1989).

[3] C.S. Lent unpublished.

[4] C.S. Lent and D.J. Kirkner, J. Appl. Phys. J. Appl. Phys. 67, 6353 (1990).

[5] Craig S. Lent, Srinivas Sivaprakasam, and David J. Kirkner, in M. A. Reed and W. P. Kirk (Eds.) *Nanostructure Physics and Fabrication*.

[6] Craig S. Lent, Srinivas Sivaprakasam, and David J. Kirkner, Solid State Electronics 32, 1137 (1989).

Ballistic current vortex excitations in electron waveguide structures

Craig S. Lent
Department of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana 56556

(Received 15 May 1990; accepted for publication 27 July 1990)

Coherent ballistic transport through a cavity in a quantum electron waveguide is examined. Electrons passing through the cavity excite vortices in the current density in the cavity. The circulating currents may be larger than the current in the waveguide leads. Vortex excitations are possible even when the incoming current is totally reflected and no net current passes through the cavity.

Quantum electron waveguide structures can be fabricated by further confining electrons in the two-dimensional electron gas (2DEG) formed at a semiconductor heterostructure interface. The additional confinement is usually accomplished by means of a metal pattern deposited on the surface and held at a negative potential, depleting the region immediately below it of carriers. With this technique, as well as others, it is possible to create channels with dimensions comparable to the de Broglie wavelength of the electrons. These channels then act as electron waveguides, in analogy with microwave waveguides.[1] Quantum resonant cavities can be formed by creating a small region in which the channel widens substantially. If the dimensions of the cavity are small compared with the electron mean free path, electron transport through the structure is both ballistic and coherent. Resonance effects strongly modulate the transmission through the cavity. This phenomenon provides the basis for proposed quantum interference transistors.[2-5]

The exact nature of these resonances has not been fully explored. The energies at which the resonances occur are not simply the bound-state energies of the cavity. The possible role of currents in the cavity has not been addressed. In this letter, resonance effects in a symmetric rectangular electron waveguide cavity will be examined and attention will be focused on the current density in the cavity itself.

Confinement in the plane of the 2DEG is assumed to be complete so that no perpendicular degrees of freedom are accessible. Transport through the cavity is calculated by solving the two-dimensional effective mass Schrödinger equation

$$-(\hbar^2/2m^*)\nabla^2\psi(x,y) + V(x,y)\psi(x,y) = E\psi(x,y). \tag{1}$$

Equation (1) is solved numerically for the wave function $\psi(x,y)$. The two-dimensional electron particle current density is calculated directly from the wave function through the relation

$$\mathbf{j}(x,y) = (-\hbar/m^*)\mathrm{Im}[\psi(x,y)\nabla\psi^*(x,y)]. \tag{2}$$

An effective mass of $m^* = 0.05\ m_0$ is assumed throughout. Since the electron motion is confined in the plane of the heterojunction interface, discontinuities in the effective mass need not be considered.

Figure 1 shows the problem geometry. Electrons in a long channel of width d are incident from the left on a cavity of width D and length L. Electrons traveling through the cavity exit on the right into a channel of the same width as the input lead. The calculations discussed here are for the case when $D/d = 3$ and $L = d$. The potential is assumed to be zero inside the cavity and the channels. Outside, the potential is infinite and the wave function is zero.

Quantum resonance effects are strongest if only one mode of the waveguide channel is excited. Most of the device structures proposed are designed to operate in this low-current, single-mode regime. The calculation here assumes that the wave incident from the left is in the first channel mode. Reflection and transmission into all other modes are included.

Equation (1) was discretized on the interior of the cavity region using the finite element method. A 49×49 mesh of nodal points was used. The boundary conditions at the input and output leads are established using the quantum transmitting boundary method (QTBM).[6] The QTBM permits the automatic inclusion of all reflected and transmitted modes, including evanescent modes, without requiring any discretization of the lead regions. Solution of the discretized Schrödinger equation yields the values of the wave function at the nodal points. The current into the right lead is calculated from the wave function directly and the transmission coefficient T, obtained as the ratio of the current transmitted to the incident current.

Figure 2 shows the transmission coefficient for the first

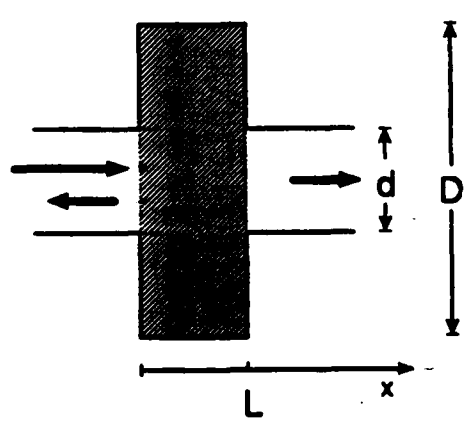


FIG. 1. Geometry of a resonant cavity in an electron waveguide. The hatched area indicates the region over which the Schrödinger equation is solved and for which the current density is displayed in Figs. 3 and 4.

 © 1990 American Institute of Physics

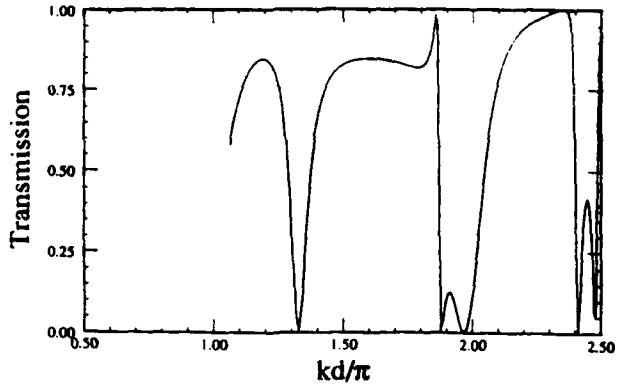


FIG. 2. Transmission coefficient for current incident in the first mode of the incoming channel. The transmission is shown as a function of kd/π where $k = \sqrt{2m^*E}/\hbar$.

channel mode as a function of energy. The energy scales with the electron wavelength and so is plotted here as kd/π, where $k = \sqrt{2m^*E}/\hbar$. The characteristic sharp minima and broad maxima which inspired the quantum interference devices are evident. The condition for propagation of

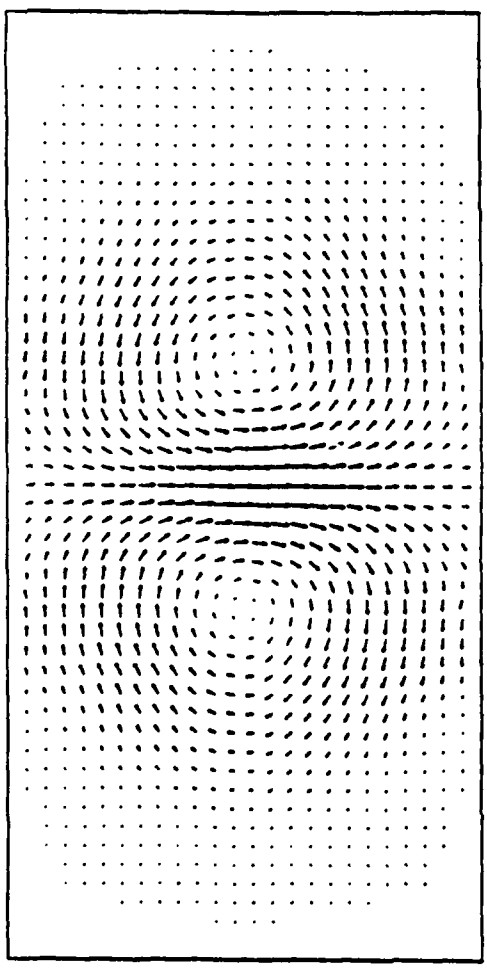


FIG. 3. Current density in the cavity for an energy such that $kd/\pi = 1.315$. This corresponds to the first minimum in the transmission shown in Fig. 2. The region displayed corresponds to the hatched area in Fig. 1.

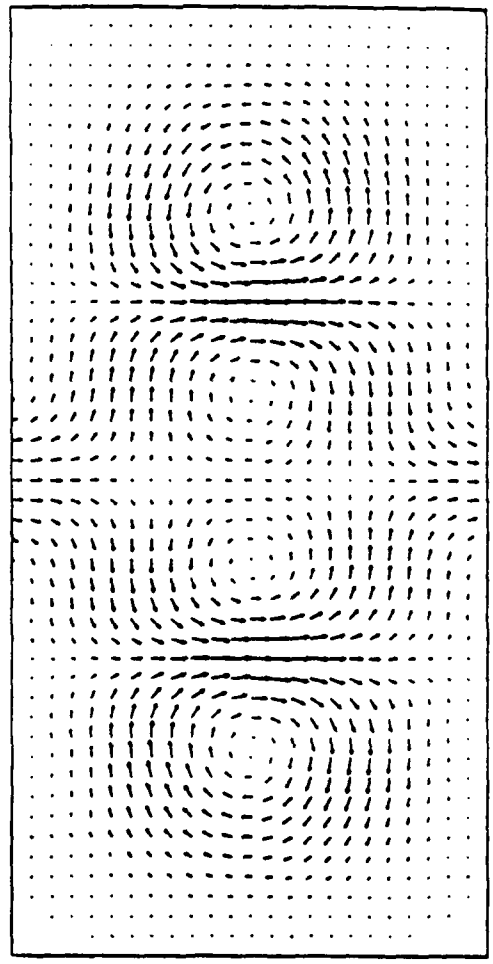


FIG. 4. Current density in the cavity for an energy such that $kd/\pi = 1.855$. This corresponds to a maximum in the transmission shown in Fig. 2. The region displayed corresponds to the hatched area in Fig. 1.

the nth mode in the channels is that $kd/\pi > n$. Since the cavity is symmetric, the parity of the wave function is conserved and the $n = 1$ incoming mode will couple only to modes with odd n. Over the range of energies shown in the figure, therefore, the first mode is the only current-carrying mode in either the input or output channels.

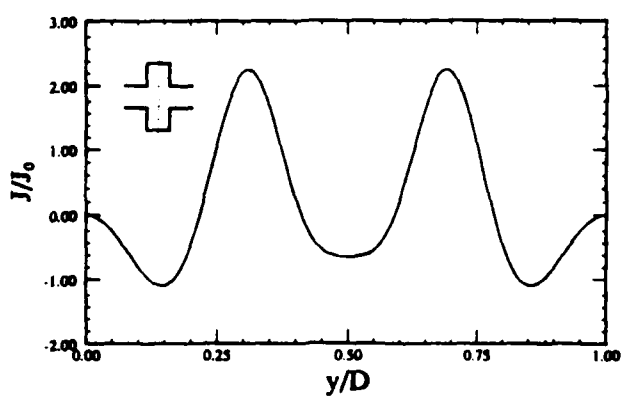


FIG. 5. Normalized current density across the mid point of the cavity at the energy shown in Fig. 4. At the mid point $j = j_x$. The graph shows the changes in the current density's sign and magnitude around the vortices.

For experiments in which the voltage across the device is small enough for linear response theory to hold, the transmission coefficient plotted in Fig. 2 is proportional to the conductance of the cavity.[7-9]

From the wave function in the cavity, the electron particle current is calculated using Eq. (2). Figure 3 shows the current density in the cavity region for an energy corresponding to $kd/\pi = 1.315$, the position of the first zero of the transmission coefficient. The appearance of ballistic current vortices is clear. At this energy a vortex-antivortex pair forms in the cavity. Figure 4 illustrates the current density in the cavity at the transmission maxima which occurs when $kd/\pi = 1.855$. Four vortices of alternating sign appear across the width of the cavity. Transmission at this energy is greater than 97%. Notice that in the center of the cavity, current flow changes sign and electrons actually flow back toward the input. The primary forward current path is split between two symmetric paths around the center of the cavity. At other energies different patterns of vortex excitations occur. Even in the limited energy range shown, excitations with up to eight vortices in the cavity appear. Excitation of the current vortices is not limited to special resonance energies. However, the magnitude of the vortex current is much larger at resonances.

The appearance of the strong vortex excitations is particularly interesting at the zeros of the transmission coefficient when the conductance of the cavity vanishes. The current flows in from the input lead, circulates in the vortex patterns, and then returns via the same lead. The transmitted current is zero and the *net* current in the input lead is zero. The incoming current is totally reflected, but drives circulating currents in the cavity. This type of excitation clearly has no hydrodynamic analogue. One might expect that the total charge in the cavity would increase due to the presence of the vortex currents. Indeed, it is shown elsewhere that the charge in the cavity increases significantly at the resonance conditions.[10]

Because the wave function can be multiplied by an arbitrary normalization factor, the magnitude of the currents in the cavity cannot be calculated directly without making some additional assumptions. However, the magnitude of the vortex current densities can be assessed by comparison to current densities in the input and output channels. Consider the case of unity transmission through the cavity (or equivalently, no cavity at all). Let the *peak* current density in the single-mode channel be J_0. The total current in the channel is then $I = -eJ_0d/2$. The current density across the center of the cavity, normalized to J_0, is shown in Fig. 5 for the same energy as depicted in Fig. 4. Across the center line of the cavity the component of the current in the y direction is zero. The figure shows the positive and negative values of j_x associated with opposite sides of the vortices. The magnitude of the current density around a vortex is clearly comparable or larger than the maximum in the leads. The total electron current across the cavity, computed by finding the area under the curve in Fig. 5, was calculated to be $J_0D/6$, in agreement with the requirement of current conservation.

A rough estimate of the magnetic field associated with a vortex can be made. Because of the large intrinsic resistance of a single-mode channel ($h/e^2 \approx 25$ kΩ), the channel current is typically 0.5 μA or less. If the vortex current were comparable and the current flow were approximated by a ring with a diameter as small as 5 nm, the associated magnetic field would be roughly 5×10^{-5} T. Thus, under rather optimistic assumptions the magnetic field generated by the current vortices is very small.

Hydrodynamic vortices occur because of interactions between particles. In this one-electron Schrödinger model, electron-electron interactions have been neglected. An improved theory would include electron-electron interactions at least in the Hartree sense, to account for the accumulation of charge in the cavity. Because current densities are so low in these systems, such an improvement is unlikely to alter the basic results.

The simple model of the confining potential used here is sufficient to illustrate the important features of current vortex formation. More realistic potentials involve less abrupt corners and walls with quadratic, rather than infinite, potentials.[11] The formation of current vortices persists, indeed is even enhanced, in softer and rounder potentials. A detailed comparison will be given elsewhere.

In summary, current transport through a cavity in a quantum electron waveguide has been examined. Current in the cavity circulates to form vortices. Vortex formation occurs even when the conductance of the cavity vanishes and all current is reflected.

I would like to thank Professor Gerald Arnold for helpful discussions. This work was supported by the Air Force Office of Scientific Research and by the National Science Foundation under grant No. ECS890025 through the National Center for Computational Electronics, and utilized the Cray-2 at the National Center for Supercomputing Application, University of Illinois at Urbana-Champaign.

[1] G. Timp, H. U. Baranger, P. deVegvar, J. E. Cunningham, R. E. Howard, R. Behringer, and P. M. Mankiewich, Phys. Rev. Lett. 60, 2081 (1988).
[2] F. Sols, M. Macucci, U. Ravaioli, and K. Hess, Appl. Phys. Lett. 54, 350 (1989).
[3] F. Sols, M. Macucci, U. Ravaioli, and K. Hess, J. Appl. Phys. 66, 3892 (1989).
[4] S. Datta, Superlatt. Microstructures 6, 83 (1989).
[5] A. B. Fowler, Workshop on Quantum Interference, Atlanta, Sept. 1988.
[6] C. S. Lent and D. J. Kirkner, J. Appl. Phys. 67, 6353 (1990).
[7] R. Landauer, IBM J. Res. Develop. 1, 223 (1957).
[8] R. Landauer, Philos. Mag. 21, 863 (1970).
[9] A. D. Stone and A. Szafer, IBM J. Res. Develop. 32, 384 (1988).
[10] C. L. Lent (unpublished).
[11] D. A. Wharam, U. Ebenberg, M. Pepper, D. G. Hasko, H. Ahmed, J. E. F. Frost, D. A. Ritchie, D. C. Peacock, and G. A. C. Jones, Phys. Rev. B 39, 6283 (1989).

Surface Science 228 (1990) 301-305
North-Holland

INFLUENCE OF EVANESCENT STATES ON QUANTUM TRANSPORT THROUGH AN ARRAY OF ELASTIC SCATTERERS

M. CAHAY *

Scientific Research Associates, Inc., Glastonbury, CT 06033, USA

S. BANDYOPADHYAY

Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, IN 46556, USA

M.A. OSMAN and H.L. GRUBIN

Scientific Research Associates, Inc., Glastonbury, CT 06033, USA

Received 31 May 1989; accepted for publication 9 August 1989

The problems of Anderson localization and universal conductance fluctuations have been treated in the past with theoretical techniques that did not usually include evanescent states in the analysis. In this paper, we have examined the effect of evanescent states on electron transport through a disordered structure using a novel generalized scattering matrix formalism. We find that evanescent states have a significant influence on the resistance of a disordered sample, especially when the degree of disorder is large. We also find that neglecting evanescent states in theoretical calculations can lead to a substantial error in the estimation of the localization length in realistic semiconductor structures.

1. Introduction

Electron transport in disordered mesoscopic systems has been modeled in the past with a variety of techniques that explicitly took into account the quantum-mechanical wave nature of electrons. They include the Anderson tight-binding Hamiltonian formalism [1], various Green's function methods [2] and scattering matrix approaches [3]. Most of these techniques however have neglected to include evanescent states in the analysis, even though these states are valid solutions of the Schrödinger equation governing transport. The rationale for this has been twofold: firstly, the evanescent states do not themselves carry current; therefore they are unlikely to affect a sample's conductance in a significant manner. Secondly, their amplitudes decay exponentially with distance, characteristically as $e^{-\kappa x}$. Consequently, their influence on the overall electronic wavefunction inside any disordered structure is expected to be negligibly small if $\kappa d \gg 1$ (where κ is the imaginary part of the electron's wavevector and d is the average spacing between elastic scatterers along the direction of current flow). While the second rationale certainly holds true for structures that are modulation doped (in which $\kappa d \gg 1$ even for the lowest lying evanescent states), it is unlikely to hold for either metallic structures, or Si MOSFET's, or even intentionally doped GaAs structures in which most of the pertinent experiments of quantum transport have been performed [4]. In the latter structures, it is quite likely that $\kappa d \leq 1$. Therefore, in the analysis of such structures, it is imperative to include evanescent states in the theoretical model. These states may not themselves carry current, but they can still influence the conductance of a sample by affecting

* Permanent address: Department of Electrical and Computer Engineering, University of Cincinnati, Cincinnati, OH 45221, USA.

0039-6028/90/$03.50 © Elsevier Science Publishers B.V. (North-Holland)

the phase-relationships between the various current-carrying states.

Recently, the inclusion of evanescent states in quantum transport models has been reported in the literature for both ballistic [5] and diffusive transport [6]. In this paper, we report an alternate technique for including evanescent states in transport calculations. Our approach is based on a generalized scattering matrix formalism which is an extension of the method reported in ref. [3]. In the next section, we describe our method which is generally applicable to any arbitrary structure.

2. Theory

We consider a quasi one-dimensional disordered resistor with an infinite confining potential in the y-direction (see fig. 1). The current flows in the x-direction. The quantum confinement in the y-direction gives rise to a set of transverse subbands in the left and right contacts whose wavefunctions are given by

$$\psi_m(x, y, t) = \phi_m(y)\, e^{\pm i k_m x}\, e^{-iEt/\hbar}, \tag{1}$$

where m is the subband index and the + or − sign refers to electron propagation to the right or left. The wavefunctions $\phi_m(y)$ are "particle-in-a-box" states for this problem.

We assume a parabolic dispersion relation for each subband

$$E = \epsilon_m + \hbar^2 k_m^2 / 2m^*, \tag{2}$$

where ϵ_m is the energy at the bottom of the mth subband. Any state at a given energy E is labeled by its wavevector k_m which can be either real or imaginary depending on whether the state is "propagating" ($E > \epsilon_m$) or "evanescent" ($E < \epsilon_m$).

At low enough temperatures and in the linear response regime, the 2- and 4-probe Landauer resistances of the structure are given by

$$R_{2\text{-probe}} = \left[2\frac{e^2}{h} \sum_{m=1}^{M} \sum_{m'=1}^{M} |t_{m',m}(E = E_F)|^2 \right]^{-1}, \tag{3}$$

$$R_{4\text{-probe}} = \left[2\frac{e^2}{h} \sum_{m=1}^{M} \frac{\sum_{m'=1}^{M} |t_{m',m}(E = E_F)|^2}{\sum_{m'=1}^{M} |r_{m',m}(E = E_F)|^2} \right]^{-1}, \tag{4}$$

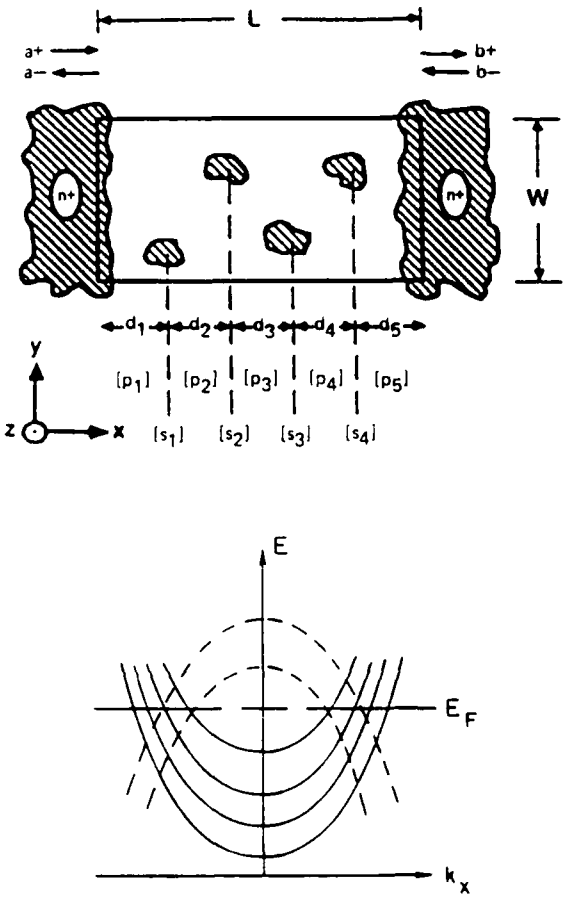


Fig. 1. A quasi-1D resistor with a random distribution of impurities. The transverse subbands are shown in the figure. The subbands that give rise to evanescent states at the Fermi level are denoted by dashed lines.

where $t_{m',m}(E = E_F)$ is the amplitude of transmission and $r_{m',m}(E = E_F)$ is the amplitude of reflection for an electron at the Fermi level E_F, traveling from a left subband with index m to a right subband with index m'. The quantity M in eqs. (3) and (4) is the total number of propagating states at the Fermi level. The summations over m and m' are carried out *only* over the propagating states. The evanescent states are not included in the sum since they do not themselves carry current and hence do not contribute directly to the conductance (or resistance). However these states are taken into account when calculating $t_{m',m}(E = E_F)$ and $r_{m',m}(E = E_F)$. They affect the transmission and reflection coefficients of the propagating

states, i.e. $t_{m',m}(E = E_F)$ and $r_{m',m}(E = E_F)$, and hence influence the conductance or resistance *indirectly*. As we shall see later, this indirect influence can be quite significant.

In order to calculate the resistance R, we simply need to calculate $t_{m',m}(E_F)$ and $r_{m',m}(E_F)$. For this purpose, we adopt the basic approach presented in ref. [3]. The transmission and reflection amplitudes $t_{m',m}$ and $r_{m',m}$ can be found from the overall scattering matrix $[S]$ describing propagation through the entire structure. The matrix $[S]$ is determined as follows. The structure is first decomposed into a number of spatial sections along the x-direction. Each section contains either an impurity (scattering center) or a region of free propagation between two adjacent impurities (see fig. 1). The scattering matrix for each section can be found analytically [3]. The overall scattering matrix $[S]$ is then found by cascading the scattering matrices for the individual sections according to the recipe presented in ref. [3]:

$$[S] = [p_1] \otimes [s_1] \otimes [p_2] \otimes [s_2] \otimes \cdots, \tag{5}$$

where $[s_i]$ is the scattering matrix associated with the ith scatterer along the x-direction and $[p_i]$ is the scattering matrix for the region of free-propagation between the ith and $(i+1)$th scatterer. Both $[s_i]$ and $[p_i]$ are $2(M+N) \times 2(M+N)$ matrices where M is the number of propagating modes and N is the number of evanescent modes at the Fermi level that are included in the analysis.

Ref. [3] gives explicit forms for the matrices $[s_i]$ and $[p_i]$. These forms are perfectly general and valid for both propagating and evanescent states. We have used the same forms in this paper.

Once the overall matrix $[S]$ is determined, $t_{m',m}(E_F)$ and $r_{m',m}(E_F)$ are found from appropriate elements of the matrix $[S]$. From these coefficients, one can evaluate the resistance R of the structure using eqs. (3) and (4). It should be noted that when evanescent states are included, the matrix $[S]$ is no longer unitary; however, current conservation still required that

$$\sum_{m'}^{M} |t_{m',m}(E_F)|^2 + |r_{m',m}(E_F)|^2 = 1 \tag{6}$$

for every value of $m \leq M$ where M is the number of propagating states.

In the next section we present some numerical results to illustrate the influence of evanescent states on the 2- and 4-probe resistances of quasi one-dimensional disordered samples in the strong and weak localization regimes.

3. Numerical example

We examine the length dependences of the resistances of two disordered GaAs samples in connection with the scaling theory of localization [7]. Each sample is 300 Å wide, with impurity concentrations of $N_1 = 10^{11}/\text{cm}^2$ and $N_2 = 10^{12}/\text{cm}^2$, respectively. In both samples, the Fermi energy E_F is made equal to the polar optical phonon energy in GaAs (= 36 meV). This makes the carrier concentration equal to $2.4 \times 10^{12}/\text{cm}^2$ and the number of propagating states at the Fermi level $M = 2$.

Fig. 2 shows the dependences of the 2- and 4-probe resistances on the sample length. For the sample with the lower impurity concentration of $10^{11}/\text{cm}^2$, we had to include 40 evanescent states in the calculation to achieve convergence of the results; while for the sample with the higher impurity concentration of $10^{12}/\text{cm}^2$, we needed 100 evanescent states. In both cases, the number of evanescent states required to achieve convergence exceeded by far the number of propagating states. In both samples, the imaginary part of the wavevector κ for the lowest evanescent state is $1.537 \times 10^6/\text{cm}$. For the first sample with impurity concentration N_1, the average spacing between the impurities is $d_1 = (N_1 W)^{-1} = 333$ Å whereas in the second sample, the average spacing is $d_2 = (N_2 W)^{-1} = 33$ Å. Therefore $\kappa d_1 \approx 5$ and $\kappa d_2 \approx 0.5$. In neither case is the condition $\kappa d \gg 1$ satisfied, so that we had to include a large number of evanescent states in both cases to achieve convergence. These two examples illustrate uniquely the importance of including evanescent states in theoretical calculations.

From fig. 2, we find that evanescent states have two effects on the resistance. Firstly, the fluctuations in the resistance are somewhat *reduced* – at least in the weak localization regime – which immediately implies that it is important to include

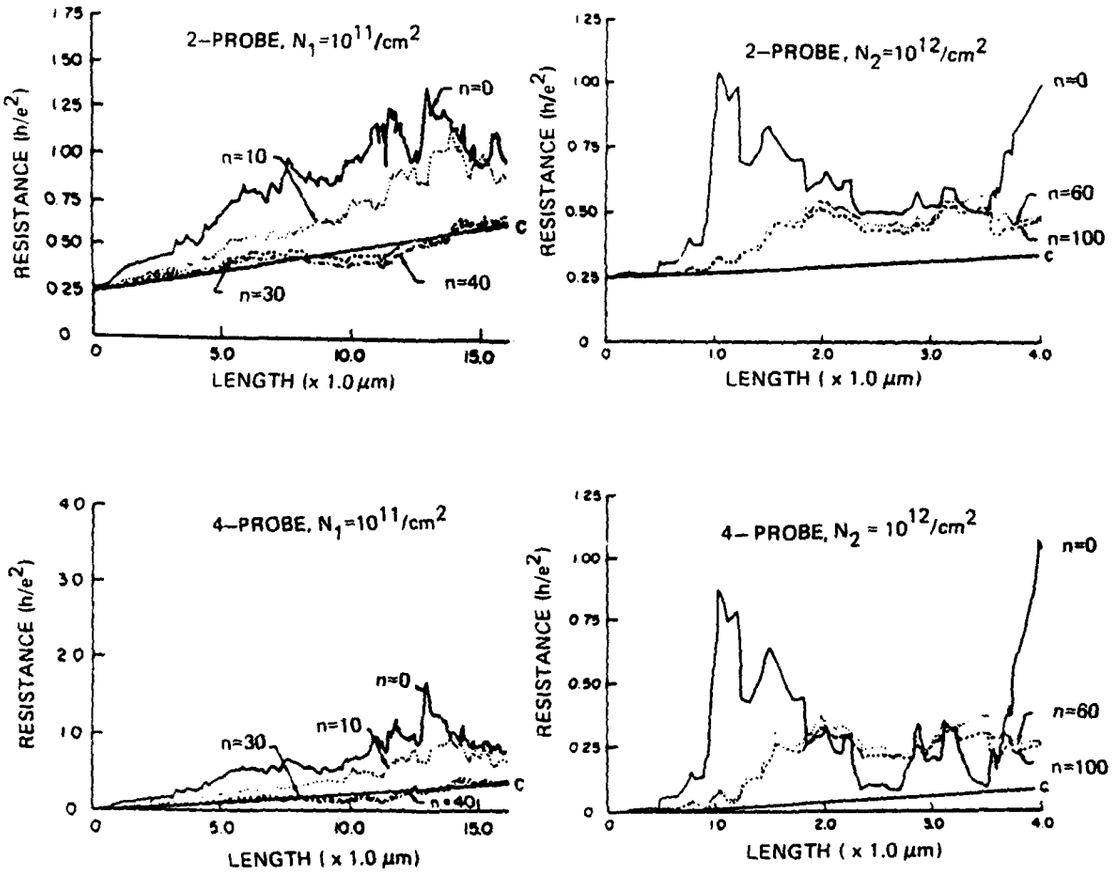


Fig. 2. Resistance versus length for a disordered quasi-1D GaAs structure. The fluctuations in the resistance are fingerprints of the impurity configuration and arise for the same reason as "universal conductance fluctuations". The number n is the number of evanescent states that were included in calculating the results. The curve labeled C is the classical result (Ohm's Law) which was calculated by cascading "probability scattering matrices" rather than "amplitude scattering matrices" as described in ref. [3]. The non-zero value of $R_{2\text{-probe}}$ at $L=0$ is the "contact resistance".

evanescent states in calculating the rms value of "universal conductance fluctuations". Secondly, on the average, the absolute value of the resistance is also *reduced*. We believe that this reduction is caused by tunneling between the impurity sites whose overall effect is to enhance the transmission probability and reduce the reflection probability thereby decreasing the resistance. The involvement of tunneling is further suggested by the fact that in the weakly disordered sample ($N_1 = 10^{11}/cm^2$), the quantum-mechanically calculated resistance is actually *lower* than the classical resistance at small lengths. Usually the quantum-mechanically calculated resistance is expected to be higher because of the Anderson localization effect or the so-called coherent backscattering effect [8]. But in the "quasi-ballistic" regime, the effect of Anderson localization is weak enough that tunneling can offset it and make the quantum-mechanically calculated resistance lower. A more detailed examination of this effect will be presented in a forthcoming publication.

Finally, another important issue that merits discussion is the effect of including evanescent states in estimating the localization length in disordered samples. We have estimated the localiza-

tion length L_{loc} for both samples by following the prescription of Thouless [9]. For the first sample, the value of L_{loc} was estimated to be ~ 4 μm before including evanescent states and ~ 14 μm after including evanescent states; whereas for the second sample, these two values were ~ 7000 Å and ~ 1.6 μm respectively. This shows that exclusion of evanescent states can lead to a significant error in the estimation of L_{loc}.

In the above two examples, we demonstrated the importance of including evanescent states in modeling electron transport through disordered structures. Further work is now in progress to examine the effect of evanescent states on the Aharonov–Bohm effect and universal conductance fluctuations.

Acknowledgements

The work at SRA was supported by the Air Force Office of Scientific Research under Contract No. F49620-87-C-0055. The work at Notre Dame was supported by the same agency under Grant No. AFOSR 88-0096 and by an IBM Faculty Development Award.

References

[1] P.W. Anderson, Phys. Rev. 109 (1958) 1492.
[2] D.J. Thouless and S. Kirkpatrick, J. Phys. C 14 (1981) 235; P.A. Lee and D.S. Fisher, Phys. Rev. Lett. 47 (1981) 882; A.D. Stone, Phys. Lett. 54 (1985) 2692.
[3] S. Datta, M. Cahay and M. McLennan, Phys. Rev. B 36 (1987) 5655;
M. Cahay, M. McLennan and S. Datta, Phys. Rev. B 37 (1988) 10125.
[4] For a review of quantum transport experiments in mesoscopic samples, see, for example, S. Washburn and R.A. Webb, Adv. Phys. 35 (1986) 375.
[5] R. Frohne and S. Datta, J. Appl. Phys. 64 (1988) 4086.
[6] A. Szafer and A.D. Stone, Phys. Rev. Lett. 62 (1989) 300.
[7] E. Abrahams, P.W. Anderson, D.C. Licciardello and T.V. Ramakrishnan, Phys. Rev. Lett. 42 (1979) 673.
[8] G. Bergmann, Phys. Rev. B 28 (1983) 2914.
[9] D.J. Thouless, Phys. Rev. Lett. 39 (1977) 1167.

Reprinted from COMPUTATIONAL ELECTRONICS: Eds. K. Hess, J. P. Leburton and U.Ravaioli, (Kluwer Academic Press) 1990

THE GENERALIZED SCATTERING MATRIX APPROACH: AN EFFICIENT TECHNIQUE FOR MODELING QUANTUM TRANSPORT IN RELATIVELY LARGE AND HEAVILY DOPED STRUCTURES

S. Bandyopadhyay[a] and M. Cahay[b]

[a] **Department of Electrical and Computer Engineering**
University of Notre Dame
Notre Dame, Indiana 46556

[b] **Nanoelectronics Laboratory and**
Department of Electrical and Computer Engineering
University of Cincinnati
Cincinnati, Ohio 45221

Over the past few years, a number of theoretical techniques have appeared in the literature for simulating phase-coherent electron transport through disordered mesoscopic structures. Notable among these are the Green's function methods [1] and transfer matrix approaches [2]. In this paper, we discuss an alternate technique - the generalized scattering matrix approach - which is ideal for studying transport through relatively *large* and *heavily doped* structures. Unlike the Green's function method which has a computational cost proportional to $(NL)^4$ and a storage requirement proportional to $(NL)^2$ (N is the number of dopants or scattering centers in the structure and L is the structure's length), the scattering matrix technique has a computational cost proportional to $(NL)^3$ and a storage requirement proportional to (NL) [3]. The reduced storage requirement is a highly desirable feature in a supercomputing environment since it decreases the number of small page faults and input/output operations which then reduces the real time of computation[1]. Consequently, the scattering matrix technique is optimal for treating those problems that require simulating transport in relatively *large* and *heavily doped* structures.

In the next section, we first briefly describe the scattering matrix technique. We then present results from our study of the scaling theory of 4-probe Landauer resistances of disordered mesoscopic structures which we have studied using this technique. Finally, we compare the scattering matrix technique with other computational techniques that are currently in vogue.

Scattering Matrix Technique: The details of this technique have been presented elsewhere [4]. We report here only the basic features. In this method, a disordered structure is first broken up into a number of sections along its length such that each section either contains an impurity (elastic scattering center) or is a region of ballistic propagation between two successive impurities. For each section, a scattering matrix is then constructed which relates the amplitudes of the incoming to the outgoing waves for every mode. It is possible to construct such a scattering matrix for any aribitrary shape of the impurity potential. These matrices are then cascaded according

[1] The transfer matrix technique also has a computational cost proportional to $(NL)^3$ and a storage requirement proportional to (NL), but it is unfortunately numerically singular for all but the smallest structures (with dimensions a few times the DeBroglie wavelength).

to the law of composition of scattering matrices [4] to yield the overall scattering matrix for the entire structure. From this overall scattering matrix, one immediately obtains the transmission amplitudes for all the propagating modes which can then be used in either the multichannel Landauer formula to obtain the resistance of the structure in the linear response regime, or in the Tsu-Esaki formula to obtain the current in the non-linear (but phase-coherent) regime.

In the next section, we present results from our study of the length dependence of the 4-probe Landauer resistances of GaAs mesas in the threshold regime corresponding to the transition from weak to strong localization. The mesas are 300 Å wide with an impurity concentration of 10^{12} cm^{-2} and a carrier concentration of also 10^{12} cm^{-2}. The impurities are assumed to be delta-scatterers. Two subbands are occupied in these structures at a temperature of 0 K.

Anderson localization in quasi one-dimensional structures: In Fig. 1 we plot the resistance of the structure as a function of its length. The solid lines are the results of a classical calculation (obtained by cascading probability scattering matrices rather than amplitude scattering matrices) and essentially reproduce Ohm's Law. The two broken lines in each figure are quantum-mechanical results obtained by cascading amplitude scattering matrices. The longdashed line corresponds to the case when no evanescent modes were included in the calculation and the shortdashed line corresponds to the case when 100 evanescent modes were included. In all cases, including 100 evanescent modes was sufficient to achieve convergence. Fig. 1(a) represents the case when the scattering potential is attractive while 1(b) represents the case when the potential is repulsive. In n-type GaAs structures, it is common to find both attractive and repulsive scatterers arising from donors and background acceptors [5].

From Fig. 1 we find that when the scattering potential is attractive, the localization length[2] is $\sim$ 1550 Å whereas for a repulsive scattering potential (of the same magnitude but opposite sign), it is $\sim$ 5500 Å. In localization theory, the elastic mean-free-path is given by the localization length divided by the number of propagating modes, so that the elastic mean-free-paths in the case of attractive and repulsive scatterers are 775 Å and 2750 Å respectively. The dependence of the elastic mean-free-path (and hence the mobility) on the *sign* of the scattering potential is an interesting result since it is not predicted by such formalisms as Fermi's Golden Rule[3].

One of the most important features in Fig. 1 is the effect of evanescent modes on the quantum-mechanically calculated resistances. The importance of these modes has been previously discussed by Bagwell [6] and us [7]. Note that for attractive potentials, the inclusion of evanescent modes increases the resistance; whereas for repulsive potentials, the resistance is decreased. It was shown in Ref. 6 that the effect of evanescent states is to renormalize the coupling between the propagating states in such a way as to increase the overall transmission in the case of repulsive scattering potentials and decrease it for attractive potentials. Consequently, evanescent modes increase the resistance and decrease the localization length for attractive scatterers while doing the opposite for repulsive scatterers.

[2] This is the length at which the resistance crosses the value h/e^2 and the length dependence changes from linear to exponential.

[3] This may have important implications for the mobilities of highly compensated structures that have both donor- and acceptor-type impurities

Comparison of the computational costs for the study of Anderson localization in wide structures: Before concluding this paper, we will discuss the computational costs involved in studying problems such as the scaling theory of Anderson localization in relatively *wide* structures. In doing so, we will compare the Green's function method with the scattering matrix method in terms of their relative computational efficiency. In the Green's function technique, one would solve the pertinent Schrödinger equation on a nearest-neighbor tight-binding lattice which requires the inversion of a N x N matrix for every impurity, N being the number of grid points along the width of the structure. The number of grid points is proportional to the width W so that one has to invert a matrix of size proportional to W (at a computational cost $\sim W^3$) as many times as there are impurities in the system. The number of impurities in the system is proportional to W^2. To see this, consider the fact that the localization length increases linearly with the number of propagating modes in the system which means that it is approximately proportional to the width. Therefore, the length of the structure that one needs to simulate in order to observe the onset of strong localization is also approximately proportional to W, and consequently the area (and hence the number of impurities) is proportional to W^2. The net computational cost is therefore proportional to W^3 x $W^2 = W^5$.

Let us now consider the scattering matrix technique. Here one has also to invert a matrix of size N x N as many times as there are impurities in the system, but in this case, the number N is the total number of modes (propagating + evanescent) required to achieve convergence. We will show that this number is independent of the width W. To see this, consider the fact that the convergence criterion is given by [6]

$$\kappa_n = Q\frac{m^*\gamma}{W\hbar^2} \tag{1}$$

where Q is a number much larger than unity, κ_n is the (imaginary) wavevector of the nth (highest lying) evanescent state included in the calculation and γ is the strength of the impurity interaction potential. We therefore have

$$n^2\frac{\hbar^2}{2m^*}\left(\frac{\pi}{W}\right)^2 = E_F + \frac{\hbar^2}{2m^*}\kappa_n{}^2 \approx \frac{\hbar^2}{2m^*}\left(\frac{Qm^*\gamma}{W\hbar^2}\right)^2 \tag{2}$$

This gives

$$n \equiv N = Integer\left[\frac{Qm^*\gamma}{\pi\hbar^2}\right] + 1 \tag{3}$$

which is independent of the width W. Consequently, the size of the matrix is independent of the width W (although the number of impurities is still proportional to W^2), so that the net computational cost is $\sim W^2$. The scattering matrix technique is therefore far superior to Green's function techniques for simulating transport through relatively wide structures.

REFERENCES

[1]. P. A. Lee and D. S. Fisher, *Phys. Rev. Lett.*, 47, 882 (1981).

[2]. E. Abrahams and M. Stephens, *J. Phys. C*, 13, L377 (1980); M. Ya Azbel, *Phys. Rev. B*, 28, 4106 (1983).

[3]. D. S. Fisher and P. A. Lee, *Phys. Rev. B*, **23**, 6851 (1981).

[4]. S. Datta, M. Cahay and M. McLennan, *Phys. Rev. B*, **36**, 5655 (1987); M. Cahay, M. McLennan and S. Datta, *Phys. Rev. B*, **37**, 10125 (1988).

[5]. R. J. Haug, K. v. Klitzing and K. Ploog, *Phys. Rev. B*, **35**, 5933 (1987).

[6]. P. F. Bagwell (preprint; also to appear in *Phys. Rev. B*).

[7]. M. Cahay, S. Bandyopadhyay, M. A. Osman and H. L. Grubin, (to appear in *Surface Science*.

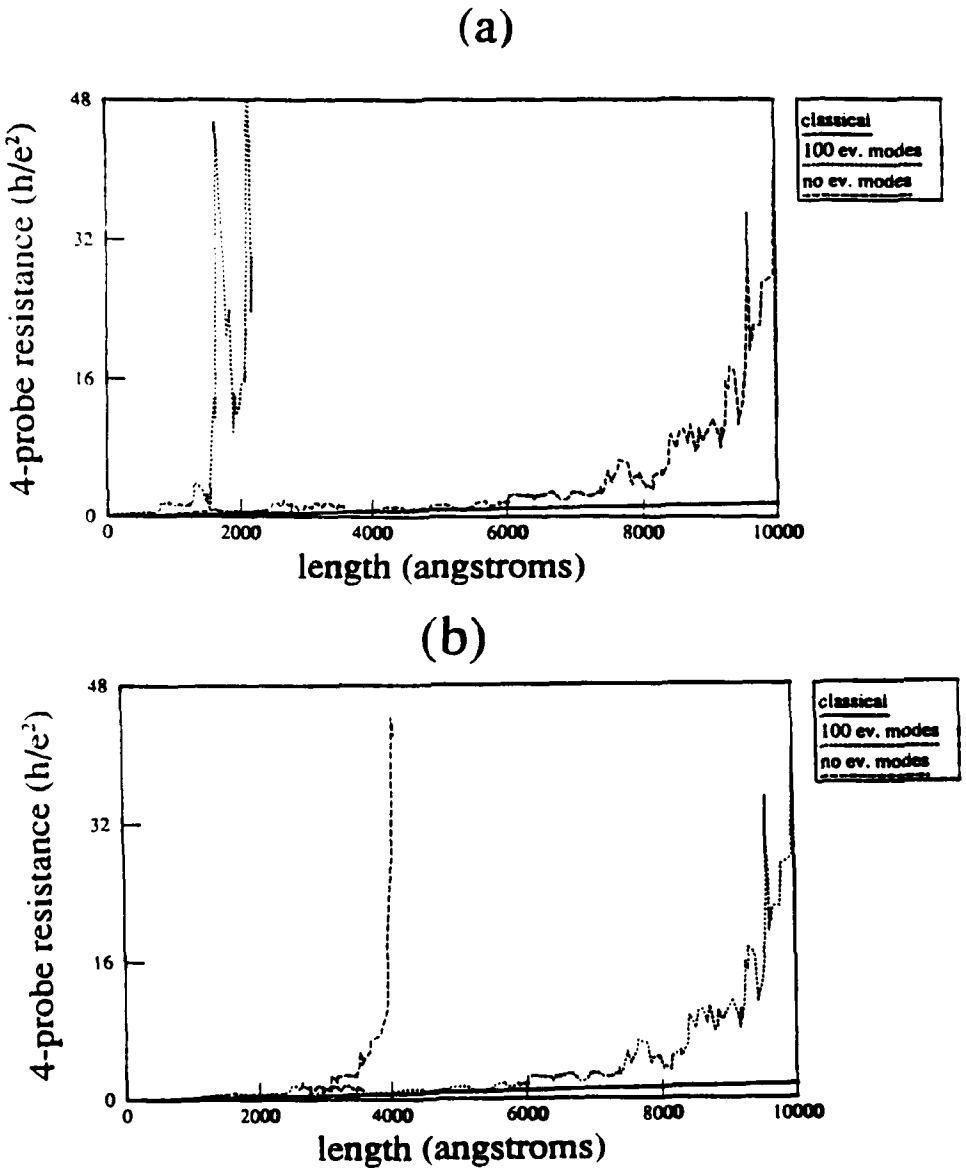


Fig 1: 4-probe Landauer resistance vs. length: (a) the scattering potential is attractive, and (b) the scattering potential is repulsive.

Reprinted from COMPUTATIONAL ELECTRONICS: Eds. K. Hess, J. P. Leburton and U. Ravaioli,(Kluwer Academic Press) 1990

NUMERICAL STUDY OF THE HIGHER ORDER MOMENTS OF CONDUCTANCE FLUCTUATIONS IN MESOSCOPIC STRUCTURES

M. Cahay[a], P. Marzolf[b] and S. Bandyopadhyay[b]

[a]Nanoelectronics Laboratory and
Department of Electrical and Computer Engineering
University of Cincinnati
Cincinnati, Ohio 45221

[b]Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

It is well-known that the conductance of a disordered mesoscopic structure exhibits sample-specific fluctuations in a magneuc field due to quantum interference effects [1]. The magnetic field changes the phase-relationships between the various Feynman trajectories in the structure thereby causing the conductance to fluctuate. In this paper, we have performed a numerical study of conductance fluctuations. Instead of using a magnetic field, we have induced the fluctuations by randomly altering the impurity *configuration* within the structure. Changing the impurity configuration has the same effect as applying a magnetic field since both alter the phase-relationships between the various Feynman paths which causes the conductance to fluctuate[1]. In addition to studying the fluctuations, we have also examined the probability distribution of the 2-probe Landauer conductances and calculated various moments of the distribution. As expected, we find that the second moment of the distribution (the standard deviation or the rms value of the fluctuations) is close to the claimed universal value of e^2/h [2], but its exact magnitude is somewhat affected by the presence of evanescent states.

In the next section, we first describe our simulation procedure. We then present results, and finally conclude with discussions about the probability distribution and its various moments.

Simulation Procedure: For calculating the conductance of a disordered structure, we have employed the scattering matrix technique [3]. The structure that was simulated is a 1.2 μm long and 2000 Å wide GaAs mesa. The Fermi energy is 36 mev and the impurity concentration is 10^{12} cm^{-2}. The impurities are assumed to be delta scatterers. Fifteen subbands are occupied in this structure at a temperature of 0 K.

For purposes of simulating samples with different impurity configurations, we have adopted the following numerically efficient scheme. Instead of varying the configuration throughout the entire structure, we first break up the structure into six equally long sections. We then cascade the scattering matrices for these six sections in all possible permutations. This procedure gives us a total of 6! = 720 different impurity configurations, or equivalently, 720 different samples to simulate. We have

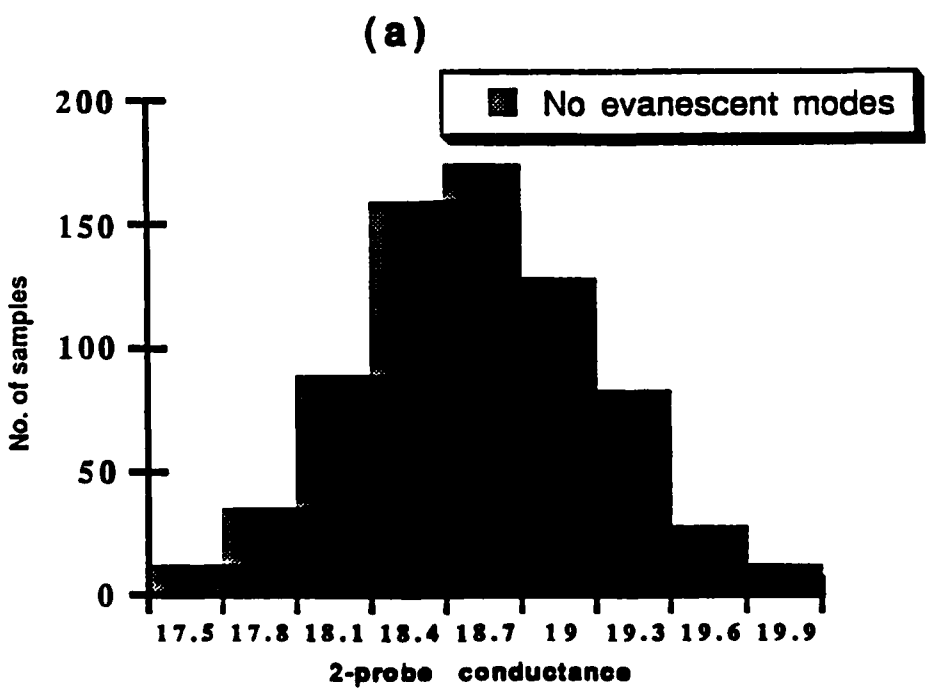
[1]The fluctuations caused by a varying magnetic field and by a randomly changing impurity configuration are statistically equivalent in all respects. This is known as the "ergodic hypothesis".

then constructed histograms from these 720 samples for the 2-probe Landauer conductances of the entire structure. From these histograms we have finally calculated the first six moments of the conductance distribution.

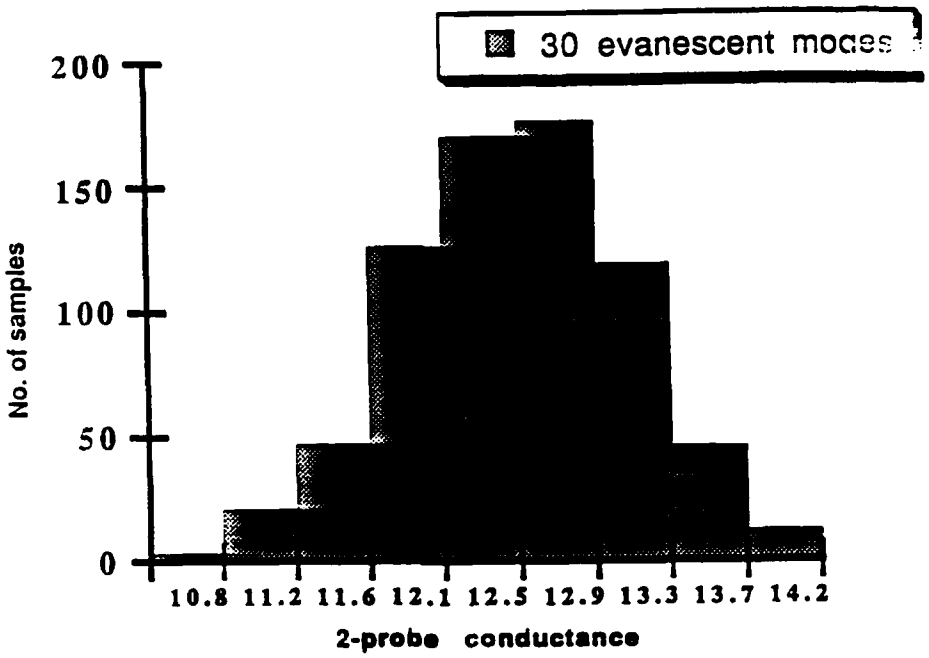
Results: Figs. 1(a) - 1(b) show the histograms of the 2-probe Landauer conductances when the scattering potential is attractive. The first figure corresponds to the case when no evanescent modes were included in the simulation and the second corresponds to the situation when 30 evanescent modes were included. Figs. 1(c) - 1(d) show the corresponding histograms for the case when the scattering potential is repulsive.

Discussion: The histograms are all approximately Gaussian. The table in the last page lists the various moments of the distribution. The moments are defined by the relation $M_n = (|\langle G^n - \langle G\rangle^n\rangle|)^{1/n}$, where M_n is the nth moment, G is the conductance, and the angular brackets denote ensemble average. Note that the effect of evanescent states is to decrease all the moments when the scattering potential is attractive. On the other hand, if the potential is repulsive, the odd moments are increased while the even moments are decreased. The origin of this feature is currently under investigation.

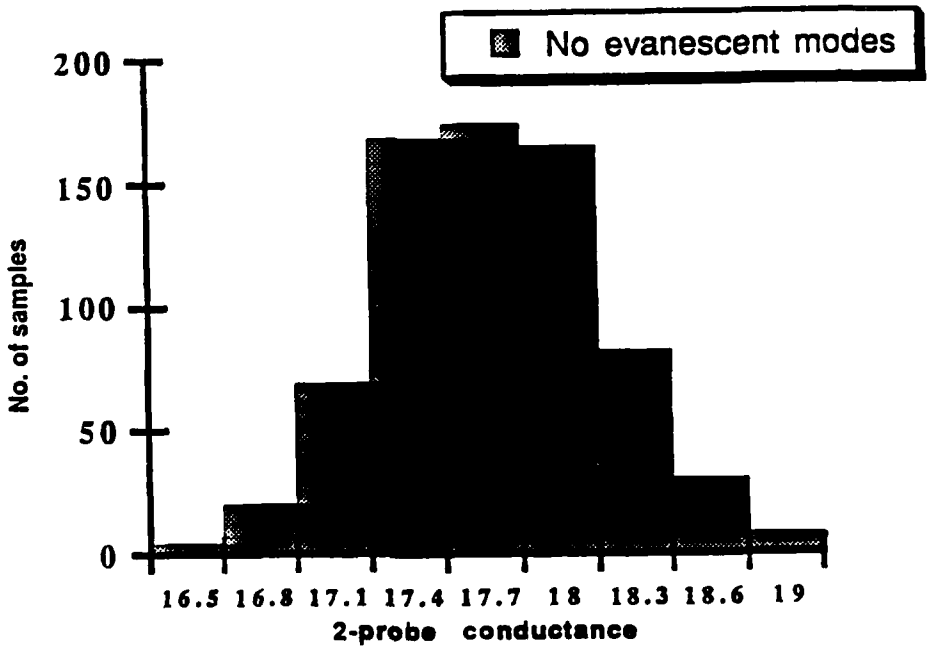
The second moment of the distribution has a value $\sim O\left(e^2/h\right)$ which agrees qualitatively with the universality theory [2]. It is interesting to note however that the exact value is somewhat affected by the presence of evanescent states. To our knowledge, this is the first systematic study of the higher order moments of the conductance fluctuations in realistic semiconductor structures.



(b)


30 evanescent modes
200
150
100
50
0
No. of samples
10.8 11.2 11.6 12.1 12.5 12.9 13.3 13.7 14.2
2-probe conductance

(c)


No evanescent modes
200
150
100
50
0
No. of samples
16.5 16.8 17.1 17.4 17.7 18 18.3 18.6 19
2-probe conductance

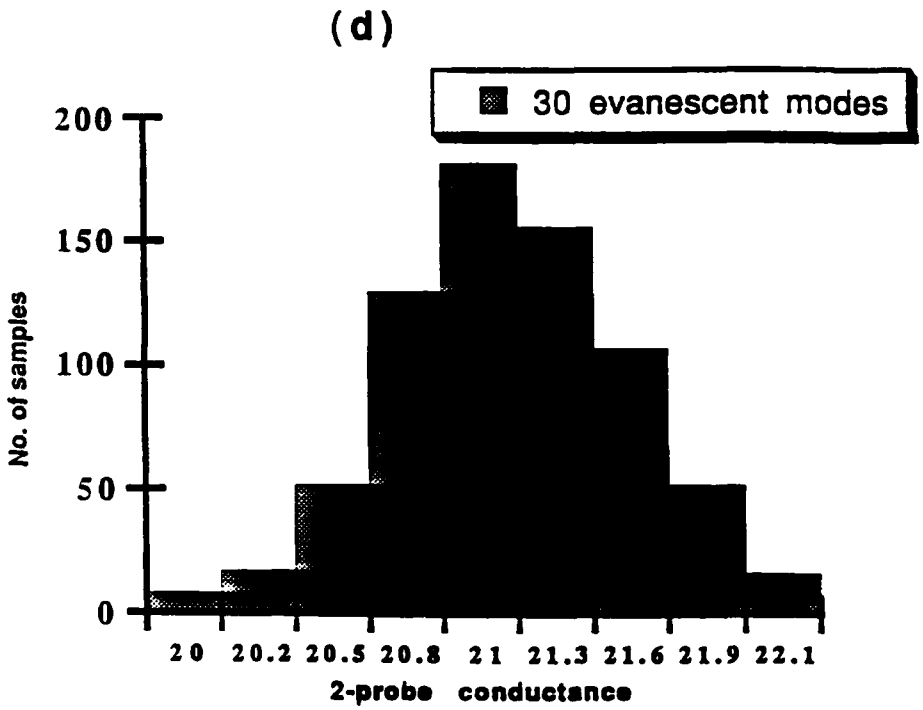


Fig 1: Histograms showing the probability distribution of 2-probe Landauer conductances. (a) and (b): the scattering potential is attractive; (c) and (d): the scattering potential is repulsive.

Table showing the moments of the conductance distribution

Scat. potential	Attractive	Attractive	Repulsive	Repulsive
No. of prop. modes	15	15	15	15
No. of ev. modes	0	30	0	30
$< G >$	18.84 e^2/h	12.86 e^2/h	17.90 e^2/h	21.30 e^2/h
2nd. moment	0.50 e^2/h	0.63 e^2/h	0.45 e^2/h	0.43 e^2/h
3rd. moment	0.14 e^2/h	0.31 e^2/h	0.16 e^2/h	0.18 e^2/h
4th. moment	0.65 e^2/h	0.83 e^2/h	0.60 e^2/h	0.56 e^2/h
5th. moment	0.29 e^2/h	0.64 e^2/h	0.29 e^2/h	0.43 e^2/h
6th. moment	0.76 e^2/h	0.98 e^2/h	0.70 e^2/h	0.66 e^2/h

REFERENCES

[1]. See, for example, B. L. Al'tshuler, *Pis'ma Zh. Eksp. Teor. Fiz.*, 41, 530 (1985) [JETP Lett., 41, 648 (1985)].

[2]. See, for instance, P. A. Lee, *Physica A*, 140, 169 (1986).

[3]. S. Bandyopadhyay and M. Cahay, these proceedings

Edge states in a circular quantum dot

Craig S. Lent
Department of Electrical Engineering, University of Notre Dame, Notre Dame, Indiana 46556
(Received 25 July 1990; revised manuscript received 3 October 1990)

The wave functions and currents in a circular quantum dot in a perpendicular magnetic field are calculated. The current in condensed (high-field) eigenstates is composed of concentric rings of current flowing in opposite directions. The current flow near the dot center flows in the direction opposite that expected from the Lorentz force. It is this inner circulation that is responsible for the "reverse" current flow associated with edge states. The correspondence between the quantum-mechanical currents and classical-particle trajectories is examined.

I. INTRODUCTION

Under a perpendicular applied magnetic field, the unconfined states of a two-dimensional electron gas (2DEG) form infinitely degenerate Landau levels whose energy increases linearly with the applied field. The Landau-level eigenstates can be thought of as localized states corresponding to the cyclotron orbits of classical electrons. If the 2DEG is confined further in the plane by a potential barrier, the energy of the states near the boundary will be altered. Further, it has long been recognized that states near the boundary produce a current which flows in a direction opposite to the circulation of inner orbits. These edge states, and their importance as a paramagnetic correction to the Landau diamagnetism, were discussed by Darwin,[1] who considered electrons in a parabolic confining potential for which analytic solutions exist. Further investigation of the effects of the edge states was done by Dingle,[2] and more recently, by Robnik.[3]

Interest in the behavior of small systems which are confined in three dimensions has been stimulated by the fabrication of individual dots and quantum-dot arrays in semiconductors.[4–7] Kumar, Laux, and Stern have solved the Schrödinger and Poisson equations self-consistently in three dimensions to obtain the electronic states for a quantum dot in a magnetic field.[8] Maksym and Chakraborty have examined the effects of electron-electron interactions in parabolic dots.[9]

In this paper we examine the one-electron states of a circular dot in a magnetic field. We take a simpler approach than Kumar, Laux, and Stern, using a fixed potential and assuming complete confinement in the plane of the 2DEG. Attention is focused here on the currents induced by the applied field and the correspondence between the quantum-mechanical results and classical cyclotron orbits. Our aim is to establish the precise nature of edge states and their relationship to the classical "skipping" orbits. The self-consistent potential obtained by Kumar, Laux, and Stern can be approximated by a flat potential with parabolic walls. After examining the hard-wall boundary case, we consider the effect of such soft walls. The numerical approach used here is similar to that of Stikova, Smrcka, and Isihara,[10] and Weisz and Berggren.[11] The calculation is performed in the framework of a single-band effective-mass model, neglecting electron-electron interactions and ignoring spin effects. Results for the circular dot which is completely confined in the plane are applicable directly to a cylindrical quantum wire. Fabrication of such structures has been reported by Reed and co-workers.[5]

II. THEORY

We consider the bound states of a particle in a two-dimensional circular quantum dot with an applied magnetic field described by a vector potential $\mathbf{A}$. The canonical momentum is given by

$$\mathbf{P}=m^*\mathbf{V}+q\,\mathbf{A}\ . \tag{1}$$

The effective-mass Hamiltonian for such a particle bound in an axially symmetric potential V_r is

$$H=\frac{1}{2m^*}(\mathbf{P}-q\,\mathbf{A})^2+V_r\ . \tag{2}$$

A completely unconstrained 2DEG in an applied magnetic field is described by the Landau Hamiltonian H_L,

$$H_L=\frac{1}{2m^*}(\mathbf{P}-q\,\mathbf{A})^2\ . \tag{3}$$

We examine the case of a perpendicular magnetic field $\mathbf{B}=\nabla\times\mathbf{A}=B_0\hat{\mathbf{z}}$ and use the symmetric gauge for the vector potential:

$$\mathbf{A}=(-B_0y/2,B_0x/2,0)\ . \tag{4}$$

The Hamiltonian can then be written in the form

$$H=\frac{1}{2m^*}(P_x^2+P_y^2)+\frac{\omega_c}{2}L_z+\frac{m^*\omega_c^2}{8}(X^2+Y^2)+V_r\ , \tag{5}$$

where L_z is the operator associated with the z component of angular momentum

$$L_z=XP_y-YP_x\ , \tag{6}$$

and

 ©1991 The American Physical Society

$$\omega_c = \frac{-qB_0}{m^*} \tag{7}$$

is the cyclotron frequency defined so that an electron has a positive ω_c corresponding to a counterclockwise orbit in the plane.

The bound states of the quantum dot (for any field) can be labeled by the radial quantum number n_r and the angular quantum number m. At zero magnetic field, the Hamiltonian is symmetric under time reversal so that $E(n_r,m)=E(n_r,|m|)$. The application of the magnetic field breaks this symmetry and the energy of positive-m states increase while negative-m states are lowered. The perturbation is dominated for low fields by the term linear in the field. The resulting splitting is due to the familiar paramagnetic interaction between a magnetic dipole and the applied field. As the field increases the (diamagnetic) quadratic term in (5), due to the induced current, becomes significant. Negative-m states which initially were reduced in energy, begin to increase and combine together, undergoing a transition from paramagnetic states to diamagnetic states at a field strength which depends on m. The negative-m states which have undergone this transition combine with n_L+1 states with positive or zero m (states that are always diamagnetic) to form the degenerate Landau levels (labeled by quantum number n_L). We refer to the coalescence of these energy levels as "Landau condensation" after Robnik.[3]

III. NUMERICAL RESULTS

A. Energy levels

We begin by examining the quantum dot with infinite hard walls. The potential V_r is zero if $r<R$ and infinite otherwise. The effective-mass Schrödinger equation for the Hamiltonian given by (2) was solved numerically using the finite element method. An effective mass of $0.05m_0$ was used. The symmetric gauge given by Eq. (4) was used. The discretization was performed on a square 49×49 node mesh. Eigenfunctions and eigenvalues were calculated using the subspace iteration technique for the lowest 20 eigenstates.

Figure 1 shows the eigenvalue spectrum as a function of applied magnetic field for the first 20 eigenstates. The calculation was performed for a dot with radius $R=500$ Å. Because all of the results scale with the de Broglie wavelength of the electron, they can be represented in a dimensionless form by appropriately scaling the energies and magnetic fields. The eigenenergies are scaled to E_0, the energy of the zero-field ground state. The magnetic field is plotted as the dimensionless quantity

$$\beta = \frac{eB\pi R^2}{\pi\hbar} . \tag{8}$$

For the 500-Å dot, $\beta=35$ corresponds to 9.21 T and $E_0=1.8$ meV.

The condensation of the bound states of the dot into degenerate Landau levels is clearly evident in the figure. At every value of the applied field, n_r and m remain good quantum numbers. The components which merge to

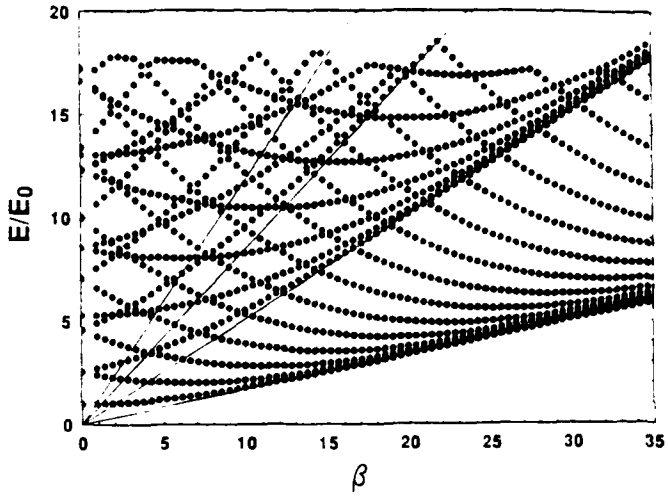


FIG. 1. Bound-state energy levels of a circular quantum dot as a function of applied magnetic field, $\beta=eB\pi R^2/\pi\hbar$. The energies are in units of E_0, the zero-field ground-state energy.

form the first three Landau levels are enumerated below:

$$\begin{aligned}
&n_L=0 ,\\
&n_r=0,\quad m=0,-1,-2,\ldots,-\infty ;\\
&n_L=1,\quad n_r=0,\quad m=1 ,\\
&\qquad n_r=1,\quad m=0,-1,-2,\ldots,-\infty ;\\
&n_L=2,\quad n_r=0,\quad m=2\\
&\qquad n_r=1,\quad m=1\\
&\qquad n_r=2,\quad m=0,-1,-2,\ldots,-\infty .
\end{aligned} \tag{9}$$

The condensation seen here does not occur in confining potentials which are parabolic. In such a potential the magnetic localization is never sufficient to isolate the electron from the walls. Parabolic potentials may be more appropriate for some dot structures examined experimentally.[12,13] Further, Demel *et al.*[13] have measured anticrossing of the energy levels in contrast to the crossing behavior in Fig. 1. They interpret these results as due to electron-electron interactions which are absent in our model.

B. Current flow

For each eigenstate, the particle (probability) current is calculated from the computed wave functions $\psi(x,y)$, using the relations,

$$\mathbf{j}=\mathbf{j}_0+\mathbf{j}_A , \tag{10}$$

$$\mathbf{j}_0=\frac{i\hbar}{2m^*}(\psi\nabla\psi^*-\psi^*\nabla\psi) , \tag{11}$$

$$\mathbf{j}_A=\frac{1}{m^*}e\,\mathbf{A}|\psi|^2 . \tag{12}$$

These relations are for electrons and the symbol e

represents the magnitude of the electronic charge. Writing the current this way separates the contribution of the wave function from that of the vector potential. This division, though gauge dependent, is helpful in understanding the nature of the eigenstates.

Figures 2–6 show the calculated current density in the dot and the probability density $|\psi|^2$ across the dot center for several eigenstates at a field of $B=5$ T ($\beta=19$). Figures 2 and 3 show the current for the $m=0$ and $m=-2$ states of the first Landau level. At this field, these states have already condensed (i.e., become degenerate with other states at the energy of a Landau level). Figure 4 shows the uncondensed $m=-7$, $n_L=0$ level. The uncondensed $n_L=1$, $m=0$, and $m=-3$ states are shown in Figs. 5 and 6.

The current flows depicted in the figures are not all easily explained by appealing to the usual classical-orbit notions. The condensed $m=0$ state shown in Fig. 2 corresponds to the classical picture of counterclockwise orbit caused by the Lorenz force on the electron. The negative-m state shown in Fig. 3 is somewhat more surprising. Rather than a central, counterclockwise current, we see a clockwise circulation in the center, surrounded by a counterclockwise outer current. This is characteristic of all the condensed negative-m states. The literature has frequently invoked the concept of "edge states" which corresponds to classical orbits that skip along the perimeter and thus carry the (particle) current clockwise, opposite that of the counterclockwise current induced in the bulklike central region. One might expect to see these edge states characterized by a clockwise current in the perimeter (edge) region of the dot. The $n_L=0$, $m=-7$ state depicted in Fig. 4 corresponds to this idea. The uncondensed $n_L=1$, $m=-3$ state in Fig. 6 should also be an edge state. Yet it has a weak perimeter current which is counterclockwise (bulklike) and an interior current which is clockwise. Below, we examine the quantum-mechanical argument for these, at first surprising, current-flow patterns. We then discuss the correspondence between the quantum wave functions and the classical orbits.

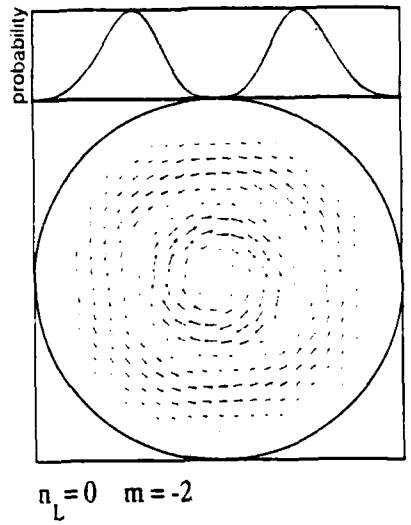


FIG. 3. Probability currents for the $n_L=0$, $m=-2$ eigenstate of a circular quantum dot. The magnetic field corresponds to $\beta=19.0$. The probability density across the dot center is shown in the upper portion of the figure.

The current $\mathbf{j}_0$ can be written in terms of ϕ, the complex phase of the wave function, and the probability density $n(\mathbf{r})=|\psi|^2$,

$$\mathbf{j}_0(\mathbf{r})=\frac{n(\mathbf{r})}{m^*}\hbar\nabla\phi(\mathbf{r}) . \tag{13}$$

Since each eigenstate of H is an eigenstate of L_z with ei-

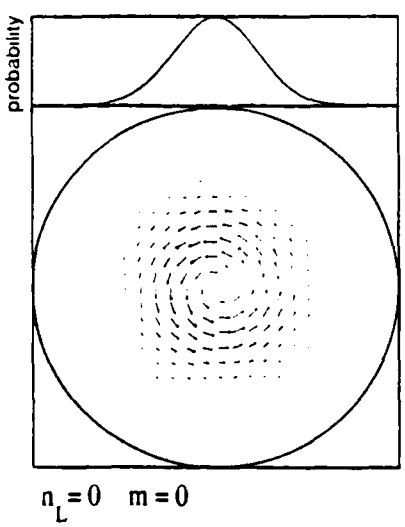


FIG. 2. Probability currents for the $n_L=0$, $m=0$ eigenstate of a circular quantum dot. The magnetic field corresponds to $\beta=19.0$ ($B=5$ T for a dot with 500-Å radius). The probability density across the dot center is shown in the upper portion of the figure.

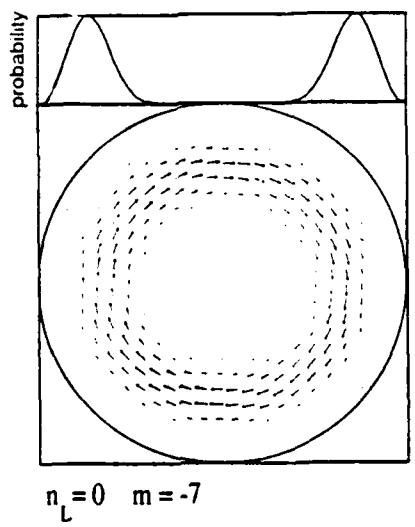


FIG. 4. Probability currents for the $n_L=0$, $m=-7$ eigenstate of a circular quantum dot. The magnetic field corresponds to $\beta=19.0$. The probability density across the dot center is shown in the upper portion of the figure.

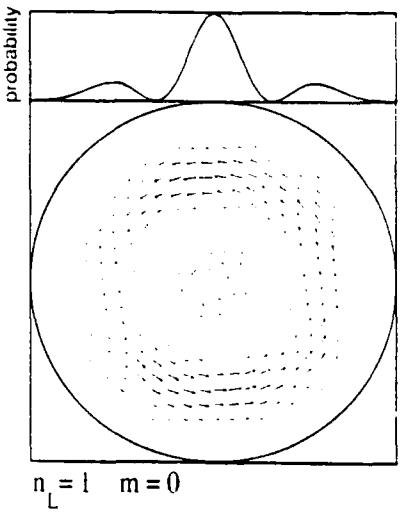


FIG. 5. Probability currents for the $n_L = 1$, $m = 0$ eigenstate of a circular quantum dot. The magnetic field corresponds to $\beta = 19.0$. The probability density across the dot center is shown in the upper portion of the figure.

genvalue $m\hbar$, we can write the complex phase in polar coordinates as $\phi(r,\theta) = m\theta$. Therefore,

$$\mathbf{j}_0(r) = \frac{n(r)}{m^*}\hbar m\hat{\theta} , \tag{14}$$

where $\hat{\theta}$ is the unit vector in the θ direction. The zero-field current of the eigenstate is simply a circulation around the origin which is proportional to the angular momentum. In polar coordinates, the explicitly field-dependent part of the current can be written,

$$\mathbf{j}_A(r) = \frac{n(r)}{2m^*} eB_0 r\hat{\theta} . \tag{15}$$

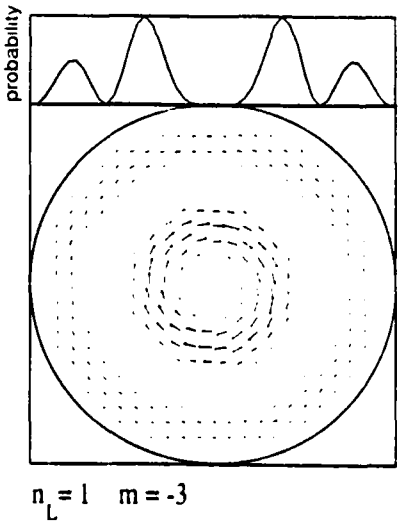


FIG. 6. Probability currents for the $n_L = 1$, $m = -3$ eigenstate of a circular quantum dot. The magnetic field corresponds to $\beta = 19.0$. The probability density across the dot center is shown in the upper portion of the figure.

For states with $m > 0$, both $\mathbf{j}_0$ and $\mathbf{j}_A$ are positive, i.e., the current flows in the counterclockwise direction. This is the direction in which a classical electron would circulate in response to the Lorentz force from the magnetic field. For states with negative m, however, $\mathbf{j}_0$ circulates in the clockwise direction. The nonzero $\mathbf{j}_A$ is in the opposite direction and increases with radial distance r. The competition between these two terms results in the concentric rings of current moving in opposite senses as seen in Figs. 3, 5, and 6. The r dependence in (15) means that the current always flows counterclockwise at large enough r, provided the probability density has not vanished. At small values of r, the $\mathbf{j}_0$ term must always dominate and leads to the current circulating in a clockwise sense near the dot center.

If the wall boundary at the dot perimeter were not present, all the negative-m states would look similar to the condensed states—a counterclockwise circulating outer ring and a clockwise inner ring. The presence of the wall reduces the probability density near the perimeter and effectively blocks the outer ring, leaving only the clockwise inner ring. At a high enough field, the magnetic field localizes the state closer to the dot center. This "restores" the outer ring of current by moving it inward from the wall region and results in bulklike behavior. The transition between edge states and bulk states does not occur by shrinking a current ring going the "wrong" direction[14] and then reversing it as it comes near the dot center. Rather it occurs by restoring a ring of current going the "right" direction which has been suppressed by its proximity to the outer edge. The circulation going the "wrong" direction is in the center and remains there in the condensed bulklike states.

C. Connection to classical orbits

In order to make clear the correspondence between the quantum-mechanical current flows and classical orbits, let us return to the case of unbound Landau levels described by the Hamiltonian H_L [Eq. (3) above]. We can define operators X_0 and Y_0 which correspond to the classical centers of the cyclotron orbits,[15]

$$X_0 = X - \frac{1}{\omega_c} V_y , \tag{16}$$

$$Y_0 = Y + \frac{1}{\omega_c} V_x . \tag{17}$$

The operators V_x and V_y are defined by Eq. (1). The operator Γ^2 is then defined to be the operator corresponding to the square of the distance from the origin to the orbit center,

$$\Gamma^2 = X_0^2 + Y_0^2 . \tag{18}$$

If we take the Landau-level wave functions $|n_L, m\rangle$ which are eigenvalues of L_z,

$$\Gamma^2 |n_L, m\rangle = \gamma^2 |n_L, m\rangle , \tag{19}$$

$$\gamma^2(n_L, m) = [2(n_L - m) + 1] L_H^2 , \tag{20}$$

where γ^2 is the eigenvalue of Γ^2 and $L_H = \sqrt{\hbar/eB}$ is the

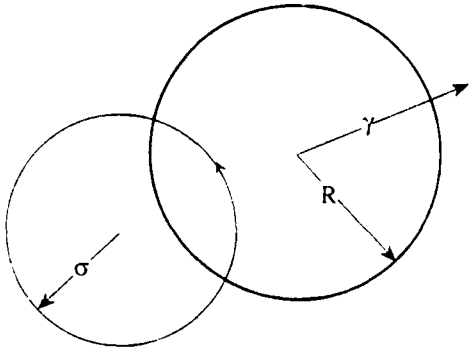


FIG. 7. The relationship between σ, the cyclotron radius, and γ, the radius of the orbit guide center. The origin is chosen to be the center of the quantum dot. The dot radius is R.

magnetic length. The classical cyclotron radius is represented by the operator

$$\Sigma^2=(X-X_0)^2+(Y-Y_0)^2 \tag{21}$$

$$=L_H^4\frac{2m^*}{\hbar^2}H_L\ , \tag{22}$$

so the eigenvalues are given by

$$\Sigma^2|n_L,m\rangle=\sigma^2|n_L,m\rangle\ , \tag{23}$$

$$\sigma^2(n_L,m)=L_H^4\frac{2m^*}{\hbar^2}(n_L+\tfrac{1}{2})\hbar\omega_c\ . \tag{24}$$

Figure 7 illustrates the relationship between γ, σ, and the origin at the dot center. The angular momentum operator L_z is related to Γ^2 and Σ^2 by

$$L_z=\frac{\hbar}{2L_H^2}(\Sigma^2-\Gamma^2)\ , \tag{25}$$

which is true classically as well.

In the unconfined Landau system, states with negative angular momentum correspond to classical orbits with centers displaced from the origin. Quantum mechanically, the position of the orbit center is not well defined since one cannot construct states which are simultaneously eigenstates of the operators X_0, Y_0, and H. The distance γ between the orbit center and the origin is a constant of the motion both classically and quantum mechanically. The cyclotron radius σ is also a good quantum number. The quantum wave function for a state with negative angular momentum corresponds, then, to all possible classical orbits of radius σ, which centers a distance γ from the origin. A circle of radius γ, centered on the origin, acts as the "guide center" for the classical orbits. As the magnetic field increases, both γ and σ become smaller. Another consequence of Eq. (20) is that states with $m=0,1,2,\ldots,n_L$ will also correspond to classical orbits with displaced orbit centers.

For the confined system, $H=H_L+V_r$, and Eqs. (20) and (24) are no longer strictly valid. For states which are already condensed and localized in the interior of the quantum dot, they will be very nearly true. We can *approximate* the value of the cyclotron radius for a particular eigenstate $|n_r,m\rangle$ of H by using Eq. (22) to define

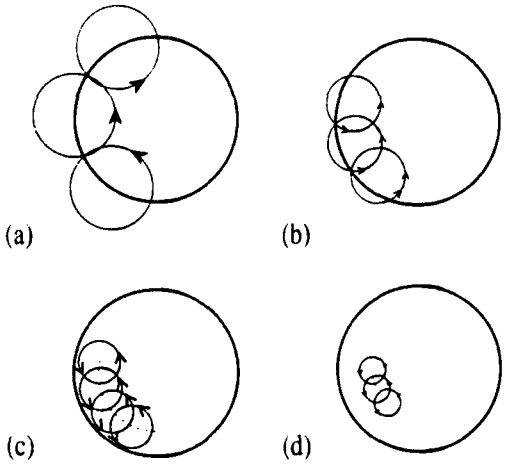


FIG. 8. Classical orbits for four increasing values of the magnetic field. The orbits (arrows), guide center (dotted), and dot wall (solid) are shown. In (a) the guide-center radius is equal to the dot radius and only a clockwise current results. As the field increases [(b)–(d)] the guide-center radius shrinks and a counterclockwise current develops around the dot perimeter, while the clockwise current becomes localized near the dot center. Orbits are chosen to correspond to specular reflection from the walls.

$$\sigma'(E)=L_H^2\sqrt{2m^*E/\hbar^2}\ , \tag{26}$$

where we use for E the *calculated* energy $E(n_r,m)$. The guide-center radius can then be obtained by using (25) to construct an approximate value,

$$\gamma'(E,m)=\sqrt{\sigma'^2(E)-2mL_H^2}\ , \tag{27}$$

where again we use the calculated value of the energy. The effect of the dot walls is then included in the raising

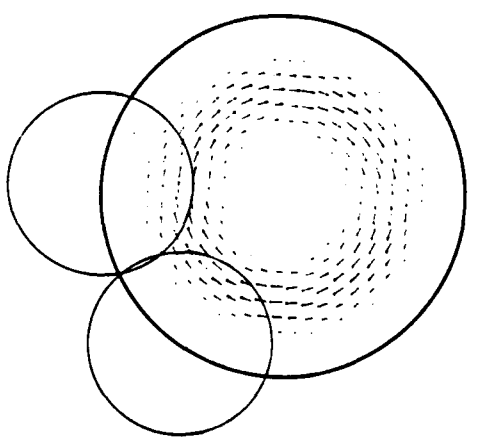


FIG. 9. Probability current for the $n_L=0$, $m=-4$ eigenstate of a circular quantum dot. The magnetic field corresponds to $\beta=11.4$ ($B=3$ T for a dot with 500-Å radius). Classical orbits calculated using Eqs. (37) and (38) are superimposed on the numerical solution.

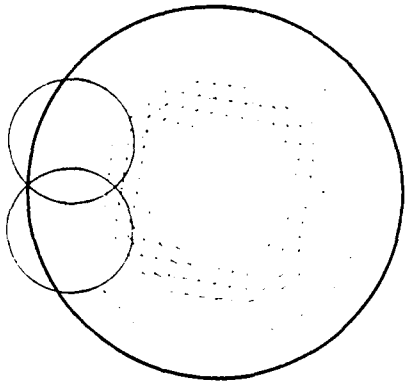


FIG. 10. Probability current for the $n_L=0$, $m=-4$ eigenstate of a circular quantum dot. The magnetic field corresponds to $\beta=15.2$ ($B=4$ T for a dot with 500-Å radius). Classical orbits calculated using Eqs. (37) and (38) are superimposed on the numerical solution.

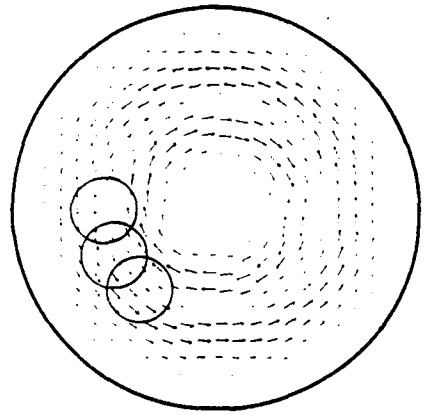


FIG. 12. Probability current for the $n_L=0$, $m=-4$ eigenstate of a circular quantum dot. The magnetic field corresponds to $\beta=30.4$ ($B=8$ T for a dot with 500-Å radius). Classical orbits calculated using Eqs. (37) and (38) are superimposed on the numerical solution.

of the eigenenergies for states which are not yet condensed in the center. Equations (26) and (27) reduce to (24) and (20) with the substitution $E=(n_L+\frac{1}{2})\hbar\omega_c$.

The classical orbits for the circular dot corresponding to quantum-mechanical eigenstates of L_z are illustrated in Fig. 8 at various stages of Landau condensation. In Fig. 8(a), the field is low and the guide-center radius $\gamma=R$. The orbits shown correspond to specular reflection off the dot walls. The current is dominantly clockwise and characteristic of a pure edge state. The guide circle (dotted) in Fig. 8(b) is inside the dot, but the orbit still corresponds to a skipping orbit, reflecting off the perimeter wall. Notice, however, that a counterclockwise current exist in the outer region between the guide circle and the dot wall. The clockwise current associated with the edge state in Fig. 8(a) has moved into the center of the dot. Figures 8(c) and 8(d) show the orbits as the magnetic field increases and the state becomes more localized and condensed. The state shown in Fig. 8(c) is not necessarily condensed, since its energy may still be elevated by proximity to the dot wall.

Figures 9–12 show the computed particle current for the $n_L=0$, $m=-4$ state at increasing magnetic fields. The values of the field are $B=3$, 4, 5, and 8 T corresponding to $\beta=11.4$, 15.2, 19.0, and 30.4. Across this range, the magnetic field transforms the state from purely edge-type, Fig. 9, to the nearly completely condensed bulk-type state shown in Fig. 12. Superimposed on the results of the Schrödinger solution are the classical orbits calculated using Eqs. (26) and (27). The relation between the classical current flows due to the orbits and the quantum results verifies the explanation above and demonstrates the utility of the quantities σ' and γ' in describing the confined states. Orbits computed using the unconfined 2DEG values σ and γ yield a much worse comparison with the calculated current patterns.

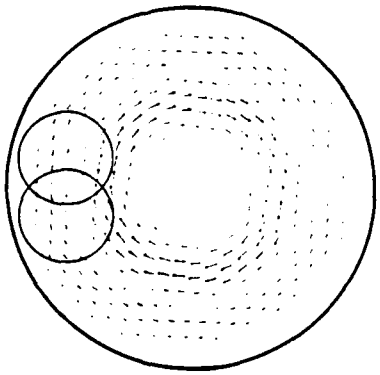


FIG. 11. Probability current for the $n_L=0$, $m=-4$ eigenstate of a circular quantum dot. The magnetic field corresponds to $\beta=19.0$ ($B=5$ T for a dot with 500-Å radius). Classical orbits calculating using Eqs. (37) and (38) are superimposed on the numerical solution.

IV. SOFT BOUNDARIES

Actual quantum dots or wires would not have abrupt hard-wall boundaries. The confining potential would be the result of the self-consistent solution of the Poisson equation for the band bending inside the semiconductor.

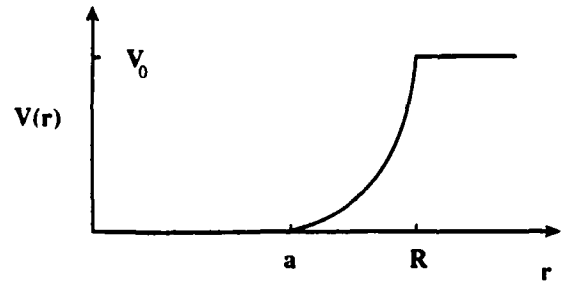


FIG. 13. Radial potential profile of quantum dot with parabolic walls.

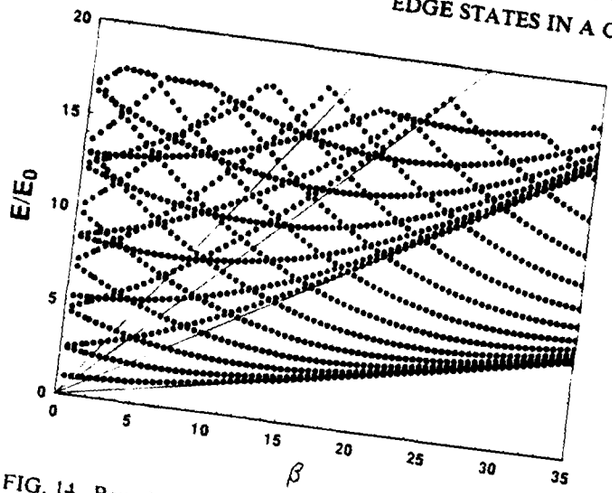


FIG. 14. Bound-state energy levels of a circular quantum dot with parabolic walls as a function of applied magnetic field. The radial potential profile is as depicted in Fig. 13 with $a=400$ Å, $R=500$ Å, and $V_0=200$ meV. $\beta=eB\pi R^2/\pi\hbar$ is a dimensionless measure of the magnetic field strength. The energies are in units of E_0, the zero-field ground-state energy.

Several calculations have shown that this potential can often be approximated by a flat ($V=0$) region and parabolically increasing walls.[8,16,17] We are particularly interested in eventually making a connection to the experiments of Reed *et al.*, after whose potential profiles we have roughly modeled ours. The effects of soft walls on the energy levels is considered here by calculating the eigenstates for a confined quantum dot with such a potential profile, as shown in Fig. 13. The radius of the inner, flat region is a, the outer radius is R, and the value of the potential at the outer rim ($r=R$) is V_0. In both of the cases examined here $R=500$ Å, $V_0=200$ meV. Figure 14 shows the energy levels as a function of applied magnetic field for $a=400$ Å. Figure 15 shows the results for $a=250$ Å. Clearly the primary effect of the walls coming in closer toward the center is to inhibit condensation until the wave functions are confined by the magnetic field to the central flat region. The qualitative discussion of the transformation of the states into Landau levels remains unchanged.

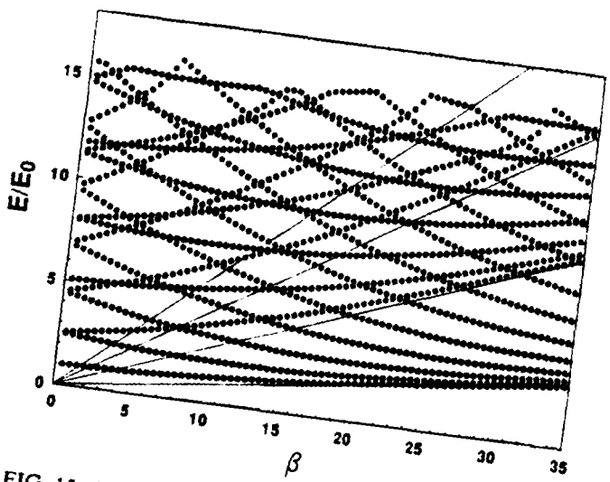


FIG. 15. Bound-state energy levels of a circular quantum dot with parabolic walls as a function of applied magnetic field. The radial potential profile is as depicted in Fig. 13 with $a=250$ Å, $R=500$ Å, and $V_0=200$ meV. $\beta=eB\pi R^2/\pi\hbar$ is a dimensionless measure of the magnetic field strength. The energies are in units of E_0, the zero-field ground-state energy.

V. CONCLUSION

In an unconfined 2DEG the degenerate Landau levels can be decomposed into eigenstates of angular momentum. States with negative angular momentum produce current distributions consisting of concentric rings with counter-circulating flows. The negative angular momentum states can be identified with classical orbits with orbit centers displaced from the origin. This identification was made by Darwin in his seminal paper, in which he examined states in a parbolic confining potential. In *that* context, he also correctly identified the negative-m states with edge states. In a parabolic potential Landau condensation never really occurs—the states always feel the effects of the potential.

In the hard-wall circular dot, negative-m states are pure edge states only at low magnetic fields. They are characterized by a single clockwise current flow near the dot perimeter. As the field increases, a counterclockwise current appears at the perimeter and the clockwise flow is squeezed inward. In this intermediate regime, the clockwise "skipping" current is not an edge current but rather a *central* current. At high fields both rings of opposing current are localized in the central region of the dot and the state becomes similar (asymptotically) to bulk Landau levels.

Numerical solutions of the Schrödinger equation for an electron in a quantum dot have been performed and the particle currents associated with particular eigenstates analyzed. The correspondence between the quantum current flows and the classical orbits has been established using approximate expressions for the guide-center radius and cyclotron radius which include the effect of the confining potential walls. The effects of softer, parabolic walls have been examined.

Note added in proof After completion of this work, I received a copy of work by Geerinckx, Peeters, and Devreese,[18] related to the effect of soft boundaries as discussed in Sec. IV.

ACKNOWLEDGMENTS

The author wishes to thank Wolfgang Porod for helpful discussions and Steven E. Laux for a copy of his calculation prior to publication. This work was supported by the Air Force Office of Scientific Research and by the National Science Foundation under Grant No. ECS890025 through the National Center for Computational Electronics and utilized the Cray-2 at the National Center for Supercomputing Applications, University of Illinois at Urbana-Champaign.

[1]C. G. Darwin, Proc. Cambridge Philos. Soc. **27**, 86 (1930).
[2]R. B. Dingle, Proc. R. Soc. London, Ser. A **212**, 47 (1952); **219**, 463 (1953).
[3]Marko Robnik, J. Phys. A **19**, 3619 (1986).
[4]W. Hansen, T. P. Smith, J. A. Brum, J. M. Hong, K. Y. Lee, C. M. Knoedler, D. P. Kern, and L. L. Chang, in *Vanostructure Physics and Fabrication*, edited by M. A. Reed and W. P. Kirk (Academic, Boston, 1989).
[5]M. Reed, J. Randall, J. Luscombe, W. Frensley, R. Aggarwal, R. Matyi, T. Moore, and A. Wetsel, Festkorperprobleme **29**, 267 (1989).
[6]A. Lorke, J. P. Kotthaus, and K. Ploog, Phys. Rev. Lett. **64**, 2559 (1990).
[7]C. T. Liu, K. Nakamura, D. C. Tsui, K. Ismail, D. A. Antoniades, and H. I. Smith, Appl. Phys. Lett. **55**, 168 (1989).
[8]Arvind Kumar, Steven E. Laux, and Frank Stern, Phys. Rev. B **42**, 5166 (1990).
[9]P. A. Maksym, and Tapash Chakraborty, Phys. Rev. Lett. **65**, 108 (1990).
[10]H. Stikova, L. Smrcka, and A. Isihara, J. Phys. Condensed Matter **1**, 7965 (1989).
[11]J. F. Weisz and K.-F. Berggren, Phys. Rev. B **41**, 1687 (1990).
[12]Ch. Sikorski and U. Merkt, Phys. Rev. Lett. **62**, 2164 (1989).
[13]T. Demel, D. Heitmann, P. Grambow, and K. Ploog, Phys. Rev. Lett. **64**, 788 (1990).
[14]The "wrong" circulation here means opposite that of a classical circular cyclotron orbit centered at the dot center.
[15]C. Cohen-Tannoudji, B. Diu, and F. Laloë, *Quantum Mechanics* (Wiley, New York, 1977), p. 742.
[16]K.-F. Berggren, T. J. Thornton, D. J. Newson, and M. Pepper, Phys. Rev. Lett. **57**, 1769 (1986).
[17]J. F. Weisz and K.-F. Berggren, Phys. Rev. B **40**, 1325 (1989).
[18]F. Geerinckx, F. M. Peeters, and J. T. Devreese, J. Appl. Phys. **68**, 3435 (1990).

Analysis of the device performance of quantum interference transistors utilizing ultrasmall semiconductor *T* structures

S. Subramaniam, S. Bandyopadhyay, and W. Porod
Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556

(Received 13 March 1990; accepted for publication 17 July 1990)

We present a theoretical study of a recently proposed class of quantum interference transistors that utilize quantum interference effects in ultrasmall semiconductor *T* structures. Our analysis reveals that the attractive features of these transistors are the very low power-delay product and multifunctionality; whereas the major drawbacks are extreme sensitivity of the device characteristics to slight structural variations, low gain, and low extrinsic switching speed in digital circuits caused by a large resistance-capacitance (*RC*) time constant arising from an inherently low current-carrying capability. The low switching speed of the transistors can however be improved dramatically by switching the device optically rather than electronically, using virtual charge polarization caused by optical excitation. This mode of switching (which is possible because of the small value of the threshold voltage) eliminates the *RC* time constant limitation on the switching time and results in an ultrafast optoelectronic switch.

I. INTRODUCTION

Recent advances in nanolithography have made it possible to fabricate "mesoscopic" structures with dimensions smaller than the phase-coherence length of electrons at cryogenic temperatures.[1] In such structures, electron transport is identical to microwave propagation through a waveguide so that these structures behave more like microwave or optical networks than ordinary circuit elements obeying Kirchoff's laws.[2] Recently, this feature of electron transport has found widespread applications in novel *electronic* devices that mimic the operation of well-known *optical* or *microwave* devices. One example of such a device is the recently proposed "quantum interference transistor" or "quantum reflection transistor"[3,4] which mimics the operation of a stub-tuned microwave *T* network.

The archetypal quantum interference transistor consists of a *T*-shaped semiconductor structure, with three terminals, in which the current between two terminals is varied by a voltage at the third terminal. Figure 1 shows a schematic representation of the structure. It can be delineated lithographically by patterning a *T*-shaped mesa (on a modulation-doped heterostructure or a quantum well) with longitudinal dimensions (i.e., dimensions along current flow) smaller than the phase-coherence length of electrons, and transverse dimensions smaller than the Fermi DeBroglie wavelength of electrons. At low enough temperatures, only the lowest subband is occupied everywhere in the structure so that each limb behaves as a true quantum wire or a "single-moded electron waveguide." Its operation as a transistor is elucidated below.

The *T*-shaped waveguide has three ports which we term (in conventional device parlance) the source (*S*), the drain (*D*), and the gate (*G*). A negative dc potential applied at the gate port will deplete a portion of the semiconductor under the gate-terminal thereby effectively controlling the conducting length of the gate arm. This is equivalent to inserting or withdrawing a "stub" (stub tuning) in a microwave *T* junction which modulates the transmission between the other two ports, namely, the source and the drain. It has been shown in Refs. 3 and 4 that modulating the potential at the gate indeed modulates the transmission (and hence the current) from the source to the drain which realizes the transistor action. Recently, the quantum interference effect underlying the operation of this transistor has also been demonstrated experimentally in a modulation-doped GaAs-AlGaAs heterostructure.[5]

In this paper, we have examined the device performance of this transistor using a scattering matrix approach. Our analysis assumes truly ballistic transport within the structure, with no scattering whatsoever—elastic or inelastic. Absence of inelastic scattering is required to preserve an electron's phase memory so that the interference effects are not destroyed. Absence of elastic scattering is also required in one-dimensional structures for two reasons. Firstly, the localization length in 1*d* structures is the elastic mean-free-path[6] so that any elastic scattering signals the onset of strong localization which renders the device inoperable. Secondly, even if strong localization does not set in, the presence of elastic scatterers will cause the conductance of the structure to exhibit sample-specific fluctuations with an rms value of $\sim e^2/h$—the so-called universal conductance fluctuations.[7] Since the maximum conductance of a one-dimensional structure is $2e^2/h$, universal conductance fluctuations will render the signal-to-noise ratio of a one-dimensional device too poor for practical applications.[8] Ballistic transport is therefore almost mandatory for the operation of one-dimensional devices. It may be pointed out that this may not be too difficult to attain at cryogenic temperatures in one-dimensional structures since elastic mean-free paths can become very long due to the one-dimensional confinement.[9]

II. ANALYSIS

The *T* structures that we have chosen to analyze are depicted schematically in Figs. 2(a) and 2(b). They are un-

 0021-8979/90/214861-10$03.00 © 1990 American Institute of Physics

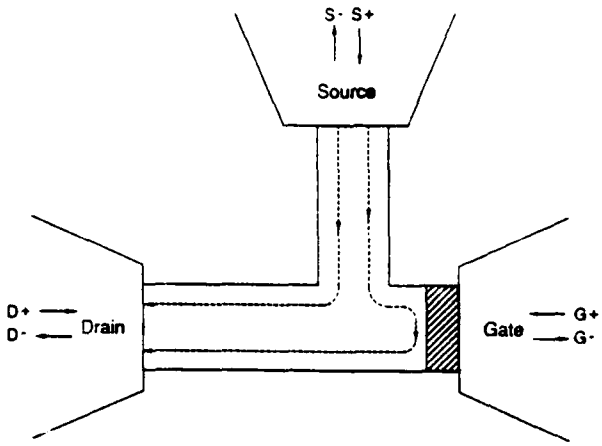


FIG. 1. Schematic representation of a T-structure transistor showing the incident and reflected waves at the three terminals. The figure also shows the two primary Feynman paths (dashed lines) whose interference results in the modulation of the source-to-drain current.

conventional transistor structures; the source is interposed between the gate and drain rather than having the gate in-between the source and drain as is conventional. Moreover, the source is much closer to the gate than to the drain for reasons that will be clarified later. In Fig. 2(b), the gate voltage which modulates the drain current is not applied to the gate termination as in Fig. 2(a), but instead is applied to a gate strip which creates a potential barrier of variable height underneath the strip. This particular structure does not operate as a conventional stub tuner. Instead, modulating the height of the barrier underneath the gate strip by a negative gate voltage modulates the electron's wave vector in that region and hence the phase-shift in traveling under the gate strip which controls the interference effect.[3] This latter structure is a modification of the original stub tuning version of Fig. 2(a) and could provide a much lower threshold voltage for turn-on and turn-off which translates into a significant advantage in terms of switching speed and power-delay product as we shall see later. We believe that the structures in Figs. 2(a) and 2(b) are optimal for the quantum interference transistor because of the following reason. It was pointed out in Ref. 3 that the modulation of the source-to-drain current is primarily due to the interference of two Feynman paths—one going directly from the source to the drain and the other going via the gate. These two paths are shown in Fig. 1. It is advantageous to make the magnitudes of the transmission coefficients of these two paths as equal as possible (over a wide range of electron wave vector) in order to obtain the best interference effect.[10] This can only be achieved by making the two paths as identical as possible. Note that in both Figs. 2(a) and 2(b), the two Feynman paths have almost the same path length (which was achieved by placing the source much closer to the gate than to the drain) and both paths have to bend around "corners" so that the transmission probabilities of these two paths will be approximately equal over a wide range of wave vectors.[11] This ensures the strongest possible interference effect.

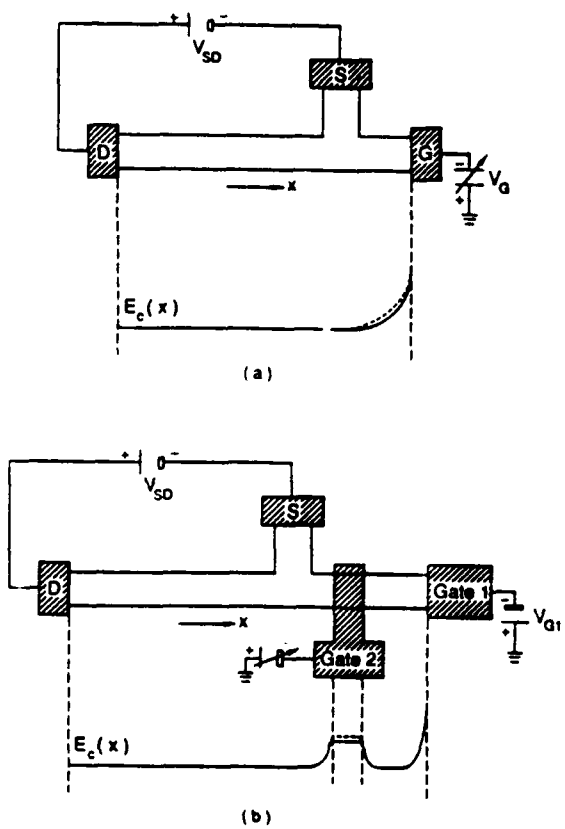


FIG. 2. (a) A single gate T-structure transistor adapted from Fig. 1. The depletion width under the gate terminal (and hence the phase-shift of an electron in traversing the gate arm) is varied by the gate potential. (b) A double-gate T-structure transistor. The purpose of gate 1 (maintained at a fixed potential) is to reflect an electron arriving at gate 1. The purpose of gate 2 is to create a localized potential barrier of variable height along the path of an electron in the gate arm. The figures also show the conduction band-edge profiles seen by an electron that probes the gate arm.

To analyze the structures, we start with the scattering matrix for a three-port network. We assume that the transmission probabilities of the two Feynman paths discussed in the previous paragraph are equal. This is equivalent to assuming that the source-to-drain transmission is equal to the source-to-gate transmission, if the gate is perfectly reflecting. If this assumption holds, then without any loss of generality we can represent the scattering matrix of the structure by the so-called Shapiro matrix.[12] The Shapiro matrix relates the incident and reflected electron wave amplitudes at the source, gate and drain ports according to (see Fig. 1).

$$\begin{pmatrix} S^- \\ D^- \\ G^- \end{pmatrix} = \begin{pmatrix} -(a+b) & \sqrt{\epsilon}^* & \sqrt{\epsilon}^* \\ \sqrt{\epsilon} & a & b^* \\ \sqrt{\epsilon} & b & a \end{pmatrix} \begin{pmatrix} S^+ \\ D^+ \\ G^+ \end{pmatrix}, \quad (1)$$

where the asterisk denotes complex conjugate.

The requirements of current conservation and time-reversal symmetry mandate that the Shapiro matrix be unitary. This gives rise to the following relations between the elements of the matrix

$$b - a = e^{i\nu}, \quad (2)$$

$$\nu = \pi - \sin^{-1}\left[\sin\alpha \times \left(\frac{1 - |\epsilon| - \sqrt{1 - 2|\epsilon| - |\epsilon|^2 \tan^2\alpha}}{2}\right)^{0.5}\right] \quad (3)$$

$$|a|^2 = \frac{1 - |\epsilon| - \sqrt{1 - 2|\epsilon| - |\epsilon|^2 \tan^2 \alpha}}{2}, \quad (4)$$

where α is the phase of the element a in the Shapiro matrix and obeys the inequality

$$1 - 2|\epsilon| - |\epsilon|^2 \tan^2 \alpha \geq 0. \quad (5)$$

We have derived Eqs. (3)–(5) in the Appendix using only the unitarity of the Shapiro matrix.

If the wave amplitudes G^+ and G^- at the gate port are related according to

$$G^+ = RG^-, \quad (6)$$

where R is the reflection at the gate, then the transmission t from the source to the drain (which determines the source-to-drain current) is given by[3]

$$t = \sqrt{\epsilon} + b\,[R/(1 - Ra)]\sqrt{\epsilon}. \quad (7)$$

Using Eq. (2) to relate b to a, we obtain

$$t = \sqrt{\epsilon}\,[(1 + Re^{i\nu})/(1 - Ra)]. \quad (8)$$

We assume that the reflection coefficient R has a magnitude of unity. This will be true if the gate impedance is infinitely high. The gate impedance will be infinitely large if the depletion layer underneath the gate in Fig. 2(a) [or the gate termination in Fig. 2(b)] is wide enough that tunneling through the layer is negligible. The *variable* phase of the reflection coefficient R is denoted by θ which is controllable by the gate potential. For the structure in Fig. 2(a), θ is given by[3]

$$\theta = 2k\,[L_d(0) - L_d(V_G)]$$
$$= 2(\sqrt{2m^*E}/\hbar)\,[L_d(0) - L_d(V_G)], \quad (9)$$

where $L_d(V_G)$ is the gate-voltage-dependent depletion width under the gate. For a Schottky gate, it is given by[13]

$$L_d(V_G) = \sqrt{2\kappa(V_{bi} + V_G)/eN_c}, \quad (10)$$

where V_{bi} is the Schottky barrier height at the gate, N_c is the carrier concentration, κ is the dielectric constant and V_G is the gate potential

For the structure in Fig. 2(b),

$$\theta = 2[k(0) - k(V_G)]L_G$$
$$= 2\left(\frac{\sqrt{2m^*(E)}}{\hbar} - \frac{\sqrt{2m^*(E - V_G)}}{\hbar}\right)L_G, \quad (11)$$

where L_G is the width of the gate strip and $k(V_G)$ is the gate voltage dependent electron wave vector under the gate strip.

Substituting $R = e^{i\theta}$ in Eq. (8), we finally obtain an expression for the total transmission t through the structure

$$t = \sqrt{\epsilon}[(1 + e^{i(\nu + \theta)})/(1 - |a|e^{i(\alpha + \theta)})]. \quad (12)$$

Note that the transmission t is a function of the electron energy E as well as the gate voltage V_G since θ is a function of both these quantities.

From the transmission t, we can calculate the source-to-drain current I_{SD} as a function of the gate voltage V_G for various source-to-drain voltages V_{SD}. We use the Tsu–Esaki formula[14] which gives

$$I_{SD} = \frac{2e}{h}\int dE\,|t(E,V_G)|^2[f(E) - f(E + eV_{SD})], \quad (13)$$

where f is the Fermi–Dirac factor. The Fermi energy E_F is obtained by assuming that the Fermi wave vector k_F is related to the volume carrier concentration N_c as $k_F = \pi n_l/2 = \pi N_c A_c/2$, where n_l is the carrier concentration per unit length and A_c is the cross-sectional areas of the quantum wires constituting the T structure. In all our simulations, the area A_c is assumed to be 100 Å × 100 Å.

At low enough temperatures and for small source-to-drain biases, Eq. (13) reduces to the familiar single-channel 2-probe Landauer formula

$$G_{SD} = (2e^2/h)|t(E,V_G)|^2$$
$$= \frac{8\epsilon e^2}{h}\frac{\cos^2[(\theta + \nu)/2]}{1 - 2|a|\cos(\theta + \alpha) + |a|^2}, \quad (14)$$

where G_{SD} is the source-to-drain conductance in the linear response regime. In deriving the last equality, we have used Eq. (12) to substitute for the transmission t.

Equation (14) shows explicitly that G_{SD} is an oscillatory function of θ. By varying θ with a gate voltage we can make the conductance oscillate as the gate voltage is scanned. This realizes the transistor action. Figure 3 shows G_{SD} as a function of θ; the device is "on" when θ is an even multiple of π and it is "off" when θ is an odd multiple of π.

III. SENSITIVITY OF THE DEVICE CHARACTERISTICS TO SLIGHT STRUCTURAL VARIATIONS—IMPLICATIONS FOR INTEGRATED CIRCUITS APPLICATION

In this section, we examine the dependence of the source-to-drain conductance G_{SD} on various elements of the Shapiro matrix that characterize the structure. From Eqs. (14), (3), and (4), we find that G_{SD}, at any value of θ, depends on two basic parameters: ϵ and α. Note that ϵ is indicative of the transmission probability from the source to either the gate or the drain. Its maximum value of 0.5 corresponds to perfect transmission from the source to the drain and to the gate (i.e., a perfectly "transparent" source) and

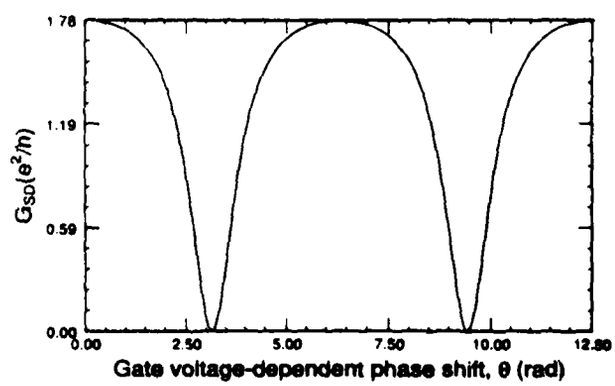


FIG. 3. The linear response source-to-drain conductance G_{SD} as a function of the gate voltage-dependent phase-shift θ. This curve was obtained from Eq. (14). The parameter ϵ was 0.5 and α was π.

its minimum value of 0 corresponds to no transmission from the source (i.e., a perfectly "opaque" source). The parameter α, on the other hand, is the phase of the internal reflection within the gate or the drain arm.

Both ϵ and α depend on the electron's wave vector or energy (which, in turn, depend primarily on the doping) as well as the precise dimensions and geometry of the structure. In other words, they are structural parameters characterizing the given structure. They can be calculated exactly by solving coupled Schrödinger–Poisson equations and calculations of this type have been performed in Ref. 3, 4, and 15. In this paper, instead of calculating these parameters exactly, we demonstrate instead how sensitive the device characteristics are to these parameters. This has implications for integrated circuits applications.

In Fig. 4, we show the amplitude of the source-to-drain conductance oscillation (which is the maximum "on-conductance" of the device) as a function of α and ϵ. A curious feature that emerges from Fig. 4 is that, contrary to intuitive reasoning, the maximum amplitude of the conductance modulation (or the maximum on-conductance) is not obtained with a perfectly transparent source (i.e., when $\epsilon = 0.5$). Instead, it is obtained when $\epsilon = 0.445$. This is more evident from the inset of Fig. 4 which shows G_{SD} (max) as a function of ϵ for $\alpha = \pi$. Since it is always desirable to have the "on-conductance" of the structure as large as possible in order to achieve the largest current carrying capability,[16] it appears that a perfectly transmitting source does not optimize the device performance. Instead, some amount of reflection at the source is desirable.

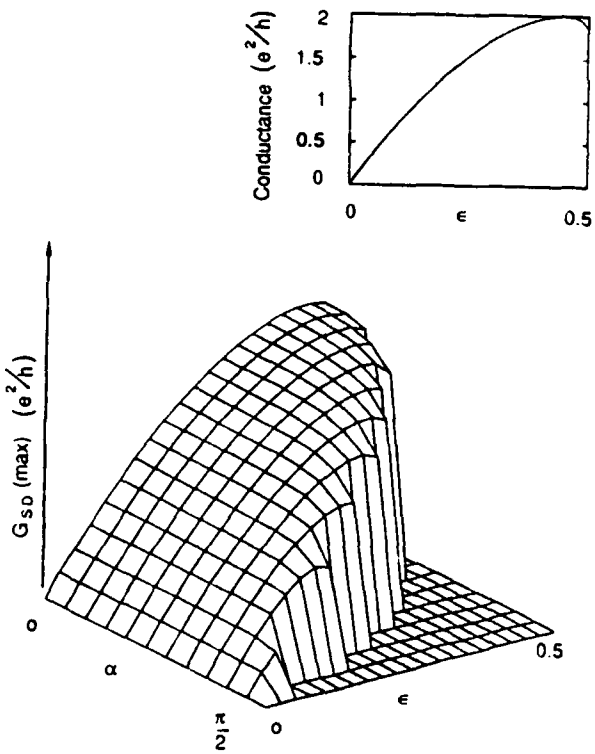


FIG. 4. The maximum source-to-drain conductance G_{SD} (max) (or the "on" conductance of the transistor as a function of ϵ, the transmission probability from the source to either the gate or the drain, and α, the phase of the internal reflection within the gate or the drain arm. The inset shows G_{SD} (max) as a function of ϵ for $\alpha = \pi$. The plateau region in this plot corresponds to those values of α and ϵ that are forbidden by the condition in Eq. (5). This plot is symmetric about the $\alpha = \pi/2$ plane.

A much more important feature to glean from Fig. 4 is the extreme sensitivity of G_{SD} to the structural parameter α. The on-conductance G_{SD} varies by almost 100% when α varies over a range of $\pi/2$. This can have a catastrophic effect in integrated circuits. The parameter α, which depends on the precise dimensions of the T structure and the electron wave vector, can vary significantly across a wafer. Consequently, different devices on a wafer will exhibit widely different behavior. For a typical carrier concentration of 1×10^6/cm, the Fermi wavevector is 1.57×10^6/cm, so that a variation of just ~ 50 Å in the dimensions of the structure alone can cause α to vary by $\pm \pi/2$. This then causes the on-conductance G_{SD} (and hence the output current for individual devices) to vary by almost 100% which renders integrated circuit implementation impossible. Suffice it to say then that the quantum interference transistor would be difficult to implement in integrated circuits where a ~ 50-Å variation in feature size from device to device is expected, even when the entire wafer is patterned by sophisticated techniques such as electron-beam or x-ray lithography.[17]

The extreme sensitivity of device characteristics to structural parameters is a serious drawback for some quantum devices and may at present preclude their application in integrated circuits.[18] It is probably true that quantum devices are not ideally suited for applications in integrated circuits given the present state of fabrication technology; however, "integrated circuits" do not constitute the only application of electronic devices. Some quantum devices are capable of performing exceedingly well as *discrete* elements for specific applications. In these applications, their performance can sometimes supercede the performance of more traditional classical devices. In the next sections, we examine the performance of a quantum interference transistor as a discrete circuit element by first studying its device characteristics and then identifying some promising applications.

IV. DEVICE CHARACTERISTICS OF THE TRANSISTOR

In this section, we examine the device characteristics of a quantum interference transistor. Figures 5(a) and 5(b) show the transfer characteristics (drain current versus gate voltage) for the structures in Figs. 2(a) and 2(b), respectively. The ambient temperature is assumed to be 4.2 K and the material is GaAs. These curves were obtained directly from Eq. (13). We assumed the drain bias V_{SD} to be 10 mV. At this bias, a ballistic electron arriving at the drain has an excess kinetic energy of 10 meV, which, although well below the threshold for polar optical phonon emission, will still raise the electron temperature to ~ 115 K. At this temperature, significant electron-electron scattering (which is the dominant phase-randomizing inelastic process in these structures) can occur. The mean time between electron-electron collisions in one-dimensional structures depends inversely on the square-root of temperature.[19] Electron-electron scattering times of ~ 1 ps have been measured in relatively heavily doped GaAs structures at 4.2 K.[20] We

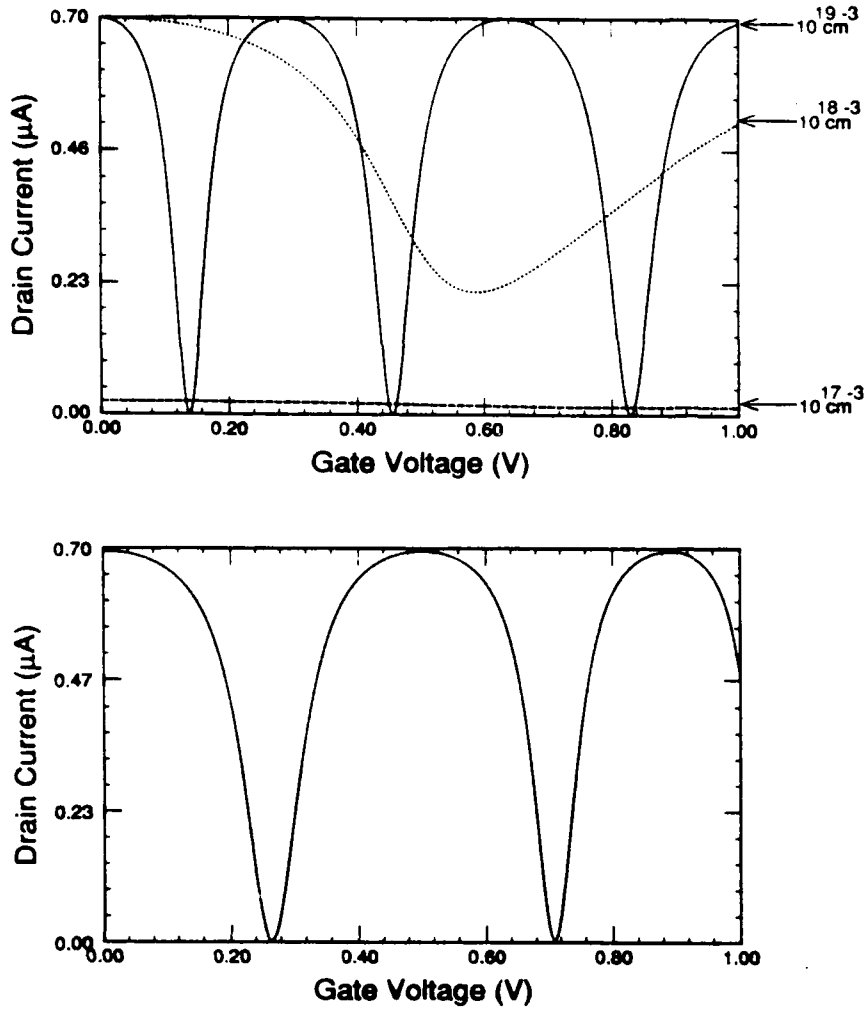


FIG. 5. The transfer characteristics (drain current vs. gate voltage) for (a) the structure in Fig. 2(a), and (b) the structure in Fig. 2(b). The characteristics in Fig. 2(a) are plotted for three different carrier concentrations. The ambient temperature is assumed to be 4.2 K and the source-to-drain voltage is 10 mV. For these plots, the parameter ϵ was assumed to be 0.5 and the phase $\alpha = \pi$.

therefore expect the scattering time at a temperature of 115 K to be $\sim 10^{-13}$ s. For a carrier concentration of 10^7 cm^{-1}, the Fermi velocity is $\sim 10^8$ cm/s, which gives an electron-electron scattering mean free path of ~ 1000 Å at a drain bias of 10 mV. This feature size is at the limit of present day lithographic capability, so that realistically, 10 mV is about the largest drain bias that can be applied in these structures.

Returning to Fig. 5(a), we find two salient features. Firstly, the gate voltage required to induce one cycle of the oscillation decreases with increasing carrier concentration N_c. This is easily understood from Eqs. (9) and (10). To induce one cycle of the oscillation, the phase-shift θ has to be changed by 2π. We find from Eqs. (9) and (10) that the quantity $\partial\theta/\partial V_G$ increases as $\sim\sqrt{N_c}$ since the wave vector k increases linearly with N_c while $L_d(V_G)$ decreases as $\sim 1/\sqrt{N_c}$. Consequently, a smaller gate voltage is required to induce a change in θ by 2π if N_c is larger. The second and perhaps the more important feature in Fig. 5(a) is that the peak-to-valley ratio of the drain current increases with increasing carrier concentration. Although not shown in Fig. 5(b), this is actually true of Fig. 5(b) as well. This can be understood as follows.

The current modulation in this structure is due to quantum interference and therefore depends critically on how tightly the phase-shift θ can be controlled. The phase-shift θ depends on the electron's energy. At nonzero temperatures, the thermal spread in the electron's energy introduces a spread $\Delta\theta$ in the phase-shift. If $\Delta\theta$ is large, the interference effect is washed out by thermal ensemble averaging. The smaller the value of $\Delta\theta$ at a given temperature, the stronger is the interference effect and the larger is the conductance modulation or the transconductance.

From Eq. (9), we find that for the structure in Fig. 2(a)

$$\begin{aligned}\Delta\theta &= 2\Delta k\,[L_d(0) - L_d(V_G)]\\ &\approx 2\,\frac{m^*\Delta E}{\hbar^2 k_F}\,[L_d(0) - L_d(V_G)]\\ &\approx 2\,\frac{2m^*kT}{\pi\hbar^2 n_l}\,[L_d(0) - L_d(V_G)],\end{aligned} \tag{15}$$

where ΔE is the thermal spread in energy ($\approx kT$) and n_l is the electron concentration per unit length.

For the structure in Fig. 2(b)

$$\Delta\theta = \frac{\sqrt{2m^*}kT}{2\hbar}\left(\frac{1}{\sqrt{E_F}} - \frac{1}{\sqrt{E_F - eV_G}}\right)L_G$$

$$= \frac{2m^*kT}{\pi\hbar^2 n_l}\left(1 - 1/\sqrt{1 - \frac{32m^*eV_G}{h^2 n_l^2}}\right)L_G. \qquad (16)$$

From Eqs. (15) and (16), we can see that for both structures, $\Delta\theta$ decreases with increasing carrier concentration n_l.[21] A larger carrier concentration therefore gives rise to a smaller $\Delta\theta$ at a given temperature and provides a larger current modulation or a larger peak-to-valley ratio of the drain current. This, combined with the fact that the gate voltage required to induce one cycle of the oscillation also decreases with increasing carrier concentration, means that the transconductance g_m ($=\partial I_{SD}/\partial V_G$) increases with increasing carrier concentration. It appears from Fig. 5 that an operating temperature of 4.2 K, we require a volume carrier concentration of 10^{19} cm^{-3} (which corresponds to a linear carrier concentration of 10^7 cm^{-1} assuming the cross-sectional area of the structure to be 100 Å × 100 Å) in order to obtain a sufficiently large transconductance. However, even at this large carrier concentration and low temperature, the actual value of the transconductance is rather small; it is only about 10^{-5} Siemens. This obviously has a deleterious effect on device performance and lowers both the small-signal gain and the bandwidth significantly. We will examine the cause for the small value of the transconductance later.

In Figs. 6(a) and 6(b), we plot the drain characteristics (drain current versus drain voltage) for the structures in Figs. 2(a) and 2(b) for various gate voltages. The most important feature to note here is that the drain characteristics do not saturate up to the maximum allowed drain voltage of 10 mV. This has very serious implications for device application. Because of this feature, it may be argued that it is not even meaningful to specify a transconductance for this device, since the transconductance is not constant over any appreciable range of the drain voltage or output voltage swing. More importantly, it also implies that this transistor is not suitable for conventional applications requiring signal amplification, since it is not possible to provide a constant amplification over any range of the input signal. This is a pathological problem with quantum devices that are constrained to operate in the linear response regime in order to avoid carrier heating. For signal amplification, a device must operate in the nonlinear response regime in the sense that the drain current must saturate. This can be achieved by lowering the carrier concentration in this structure (the drain characteristics do tend to saturate if the carrier concentration is less than 10^{17} cm^{-3}), but this also lowers the transconductance drastically. This means that even if the drain current can be made to saturate so that the differential drain resistance r_d ($=\partial V_{SD}/\partial I_{SD}$) is large, the transconductance g_m would still be very small so that the overall small-signal voltage gain a_v ($=g_m r_d$) will not be sufficiently large for signal amplification. In summary, this transistor is not suitable for applications requiring signal amplification.

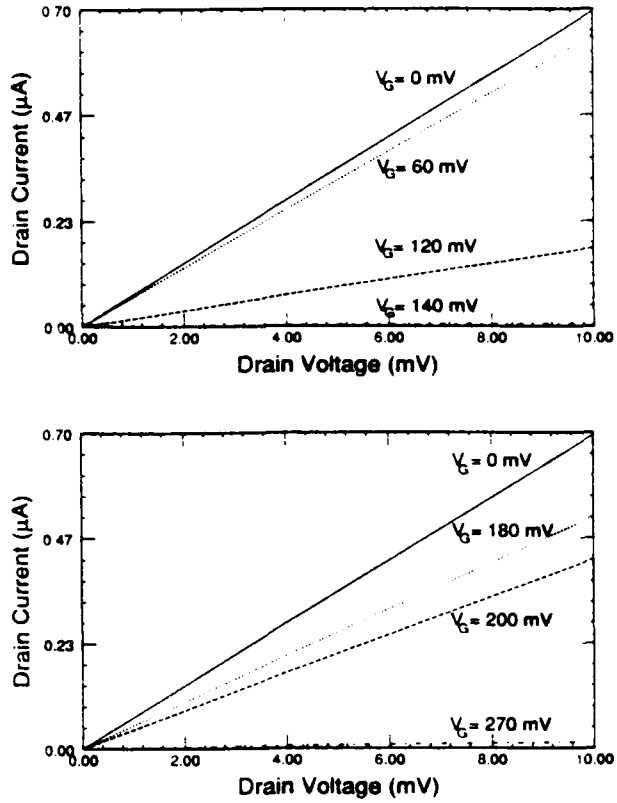


FIG. 6. The drain current vs drain voltage characteristics for various gate voltages plotted for (a) the structure in Fig. 2(a), and (b) the structure in Fig. 2(b). The ambient temperature is 4.2 K and the carrier concentration is 10^{19} cm^{-3}. For these plots, the parameter ϵ was assumed to be 0.5 and the phase $\alpha = \pi$.

V. ANALOG APPLICATIONS

Even though the T-structure transistor is not suitable for application as an analog amplifier, there are other applications for which it may be ideal. One obvious application is in frequency multiplication. Since the gate voltage swing required to make the drain current go through one cycle of oscillation is rather small, this transistor can be used as a *single-stage* frequency multiplier in the following way. If we apply a sinusoidal voltage of peak-to-zero amplitude V and frequency f at the gate, and the gate voltage swing required for one cycle of drain current oscillation is V_p, then the frequency of the drain current oscillation will be $(V/V_p)f$. This can be understood by referring to Fig. 7 where one cycle of the gate voltage swing results in three cycles of the drain current oscillation. The frequency multiplication factor N is simply given by

$$N = V/V_p. \qquad (17)$$

For a carrier concentration of 10^{19} cm^{-3}, $V_p \approx 300$ mV for the structure of Fig. 2(a) and 540 mV for the structure of Fig. 2(b) if the gate width is 100 Å. Hence for a gate voltage amplitude of 10 V, the frequency multiplication factor is ~20–30. Note that V_p scales inversely as the gate width for the structure in Fig. 2(b), so that we can decrease V_p and hence increase the frequency multiplication factor by making the gate width larger. However, this will make the entire

structure larger which will then necessitate decreasing the source-to-drain voltage in order to increase the electron mean-free-path.

Of course, there is a maximum limit on the frequency of the input signal that can be applied at the gate. The gate charges up in a time determined by the RC time constant associated with the charging of the gate capacitor so that the maximum input frequency is

$$f_{\text{limit}} = 1/(2\pi RC), \tag{18}$$

The RC time constant should be interpreted as the time required to charge the gate capacitor to a voltage V with a current I, i.e.,

$$V = \frac{1}{C_g}\int_0^{RC} I\,dt. \tag{19}$$

If we assume the current I to be time-independent, then

$$RC = C_g V/I \tag{20}$$

with C_g being the gate capacitance given by

$$C_g = A\sqrt{\epsilon\kappa N_c/2V_{bi}}. \tag{21}$$

The current I in Eqs. (19) and (20) is the gate current. For calculating the unity gain bandwidth, we set it equal to the drain current so that maximum value of I is

$$I = (2e^2/h)V_D(\max). \tag{22}$$

Combining Eqs. (18), (20), (21), and (22), we obtain

$$f_{\text{limit}} = \frac{e^2}{\pi hA}\sqrt{\frac{2V_{bi}}{\epsilon\kappa N_c}}\,\frac{V_D(\max)}{V}. \tag{23}$$

Finally, from Eqs. (17) and (23), we find that the maximum output frequency of the device (for unity current gain) is

$$f_T = f_{\max} = Nf_{\text{limit}} = (V/V_p)f_{\text{limit}}. \tag{24}$$

Using Eqs. (9), (10), and (11) to substitute for V_p in the above equation, we find that for the structure in Fig. 2(b), f_T is given by

$$f_T = f_{\max} = \frac{4m^*e^3L_G}{\pi h^3A(n_l)^{3/2}}\sqrt{\frac{2V_{bi}A_c}{e\kappa}}\,V_D(\max), \tag{25}$$

whereas for the structure in Fig. 2(a), f_T is given by

$$f_T = f_{\max} = \frac{e}{2\pi h}\frac{A_c}{A}V_D(\max) \sim \frac{I_{SD}(\max)}{e}, \tag{26}$$

where A_c is the cross-sectional area of the T structure.

Assuming $L_G = 100$ Å, $A = 100$ Å $\times$ 100 Å, $A_c = 100$ Å $\times$ 100 Å, $V_{bi} = 0.6$ V, $n_l = 10^7$ cm^{-1} and $V_D(\max) = 10$ mV, we obtain a value of $f_{\max} \approx 100$ GHz for the structure in Fig. 2(b) and ≈ 400 GHz for the structure in Fig. 2(a). These figures are certainly impressive, but not significantly larger than the largest bandwidths that have been obtained with more conventional devices such as InAlAs-InGaAs HBTs,[22] pseudomorphic InGaAs HEMTs,[23] GaAs MESFETs,[24] GaAs Gunn oscillators[25] and resonant tunneling diodes.[26] What limits the maximum frequency of the T-structure transistor ultimately is the low current carrying capability. This can be seen directly from Eq. (26). The low current carrying capability is caused by the rather low conductance of $2e^2/h$ ($= 7.8\times10^{-5}$ S) along with the rather low limit on the maximum drain voltage of $\sim$ 10 mV. In addition to decreasing the bandwidth of the transistor, the low current carrying capability also lowers the transconductance g_m and hence the transistor gain.

Although the bandwidth of this transistor is not spectacularly high, what is impressive about the device is the *multi-functionality* implicit in its use as a *single-stage* frequency multiplier. Conventional frequency multipliers require more than one active element, whereas this transistor provides frequency multiplication in a *single stage* thereby replacing an entire circuit by a single element. This multi-functionality, which is exhibited by many quantum devices,[27] is a very important attribute and has serious implications for the ultimate viability of these devices. Since quantum devices are not ideally suited for use in integrated circuits or circuits with many active elements, one can exploit this kind of multi-functionality to realize the function of a multi-element circuit without using more than one element. It is this multi-functionality that appears to be the most attractive feature of quantum devices at present.

VI. DIGITAL APPLICATIONS

We now examine the performance of the T-structure transistor for discrete logic applications. The non-saturating behavior of the drain characteristics poses a problem here as well since signal restoration at logic nodes, sufficient noise margin and sharp transitions between logic levels all require saturating (nonlinear) devices with preferably high transconductance. However, for very low-power logic and under rather low-noise operating conditions (low temperature), it may be possible to use the T-structure transistor as a discrete logic switch for digital applications. There are three basic requirements that such a switch has to satisfy: (1) The ratio of the "on" to "off" conductance must be sufficiently large so that there are well-defined logic levels. (2) The switching

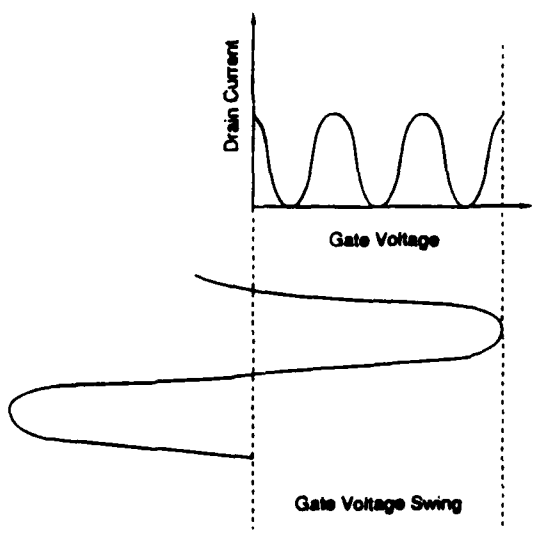


FIG. 7. Figure elucidating the application of a T-structure transistor as a single-stage frequency multiplier. A single swing of the gate voltage causes the drain current to swing over 3 cycles resulting in a multiplication of the input frequency by a factor of 3.

speed must be high. (3) The power-delay product must be low.

From Fig. 5 we find that in order to satisfy the first requirement, the carrier concentration must be at least 10^{19} cm^{-3}. We have therefore used this value of the carrier concentration to calculate the switching speed and power-delay product.

In calculating the switching speed, we first recognize that there are three time constants involved in switching. The gate charges up to the threshold voltage in a time determined by the RC time constant, the barrier in the device responds to the gate potential in the dielectric response time and the drain (output) current responds on timescales of the order of the transit time through the device. Of these three time constants, the dielectric response time is much smaller than the other two, so that the switching time is determined primarily by the other two time constants.

The transit time through the structure is essentially the time required to traverse the longer of the two Feynman paths in Fig. 1. Assuming this path length L to be ~ 1000 Å, we obtain

$$\tau_{\text{transit}} = L/v_F \approx 0.1 \text{ ps}, \tag{27}$$

where we assumed the Fermi velocity v_F to be $\sim 10^8$ cm^{-1} which is the limit set by Bragg reflection in GaAs.

The RC time constant is obtained as in the previous section

$$\tau_{RC} = C_g V_t / I. \tag{28}$$

It is important to note that the RC time constant depends on the threshold voltage V_t. The lower the value of V_t, the lower is the value of the RC time constant. The low values of threshold voltages in quantum devices partially offset the low values of the current I that can be sustained in these devices to yield a small enough RC time constant. From Figs. 5(a) and 5(b) we find that the drain current can be changed by more than 90% if the gate voltage is changed by ≈ 100 mV. Hence the threshold voltage $V_t \approx 100$ mV. Using this value of V_t along with the values of I and C_g from the previous section, we obtain a RC time constant τ_{RC} of ~ 0.15 ps. The overall *intrinsic* switching delay of the transistor is therefore ~ 0.1 ps. This is faster than the switching speeds of the fastest electronic devices that currently exist.[28]

The small value of the switching delay accrues primarily from the small threshold voltage and the small gate capacitance. The gate capacitance was calculated from Eq. (21) to be $\sim 10^{-18}$ F. While such small discrete capacitors have been realized in Coulomb blockade experiments,[29] it is unlikely that when interconnects are attached to the device, the overall gate capacitance (including that due to the interconnects) will be that small. A more realistic estimate for the overall capacitance (including the effect of interconnects) is $\sim 10^{-15}$ F. We therefore estimate that while the intrinsic switching delay is smaller than 1 ps, the *extrinsic* switching delay may actually exceed 100 ps. What limits the extrinsic switching speed in this case is the extrinsic RC time constant of the gate interconnects. In the next section, we propose a way to overcome this limitation by switching the device optically rather than electronically, thereby eliminating the RC time constant limitation associated with gate charging. This scheme of switching can provide an ultrafast optoelectronic switch.

Before concluding this section, we briefly examine the power-delay product of the T-structure transistor. The power-delay product is estimated as

$$PDP \approx C_g V_t^2. \tag{29}$$

For the structure of Fig. 2(b), this gives an intrinsic value of 5×10^{-19} J. Even if we include the effects of interconnects and raise the value of the gate capacitance to 1 fF, we still obtain a value of 5×10^{-16} J which is undoubtedly impressive and is comparable to that of Josephson junctions. This extremely low power-delay product accrues primarily from the small value of the threshold voltage. We therefore conclude that the T-structure transistor, despite its low switching speed, can still exhibit a power-delay product lower than that of most electronic devices that currently exist.

VII. OPTOELECTRONIC APPLICATIONS

In this section, we examine the viability of switching the T-structure transistor optically, rather than electronically, in order to eliminate the RC time constant limitation on switching and realize an ultrafast optoelectronic switch. Recently, it was pointed out that an optical radiation field, with a frequency lower than the band-gap frequency, can induce virtual charge polarization due to virtual transitions between electron and hole states in a quantum confined structure.[30] The field associated with this charge polarization may be sufficiently large to generate the small voltage required for switching a T-structure transistor. It is important to note that this voltage is not generated electronically, so that there is no RC time constant limitation on the switching.[31] Instead, the voltage is generated on time scales determined by the pulse width of the optical pulse and inherent time response of the virtual charge polarization mechanism which is expected to be ~ 100 fs.[30] This time scale is comparable to or smaller than the transit time through the T structure, so that the overall switching speed will be of the order of the transit time which can be made small enough (~ 1 ps) by making the structure short. Using this scheme, one can therefore realize an ultrafast optoelectronic switch.

We now examine any modifications that would be necessary to convert the T-structure transistor into an optoelectronic switch. It was calculated in Ref. 30 that for a reasonable laser pump power density of 10^8 W cm^{-2}, the screening field generated by virtual charge polarization is ~ 0.5 kV/cm for an optical detuning energy of 50 meV, when the structure is biased with a dc field of 10^5 V/cm along the direction in which the screening field is created. Since a laser spot can be focused to an area of 1 μm^2, the required pump power can be provided by a 1-W laser. For a T-structure with a vertical dimension of 200 Å, the required dc bias of 10^5 V/cm can be obtained by applying a constant voltage of 200 mV across the vertical dimension. With this arrangement, the voltage generated by the screening field will be 0.5 kV/cm $\times$ 200 Å = 1 mV. If the threshold voltage for switching can be lowered to this value, it will be possible to switch the transistor optically and realize an ultrafast optoelectronic switch.

To attain a threshold voltage of 1 mV while maintaining a carrier concentration of 10^{19} cm^{-3}, we need to increase the width of the gate strip in the structure of Fig. 2(b) to ~4000 Å [see Eq. (11)]. This will of course also necessitate increasing the source-to-drain separation, to about 1 μm, which in turn would necessitate reducing the drain voltage to about 1 mV in order to increase the mean-free-path for electron-electron scattering. This low value of the source-to-drain voltage does not present any special problems for switching applications since even with this small drain voltage, the drain current can swing over a range of $\sim (e^2/h) V_{SD} = 40$ nA which can be detected by sensing amplifiers.

We therefore conclude that it is possible to use a T-structure transistor to realize an ultrafast optoelectronic switch. A prototypical structure for such a device will consist of the configuration shown in Fig. 2(b) with the entire top surface covered with an optically opaque material leaving a transport slit of width ~4000 Å in the place of the gate strip. Since the vertical dimension of the structure is only about 200 Å, we can neglect the effect of diffraction through the slit and assume that incident radiation absorbed through the slit will create a localized potential barrier underneath the slit which can switch the transistor from an "on" to an "off" state or vice versa. The effective switching speed will be of the order of the transit time through the structure, which, for a 1-μm-long structure is ~1 ps. This is an attractively high switching speed.

VIII. CONCLUSION

In this paper, we examined the device performance of quantum interference transistors utilizing ultrasmall semiconductor T structures. Our analysis revealed both the shortcomings and the attractive features of these transistors. On the negative side, the major hurdle appears to be the lack of control over device characteristics (or their irreproducibility) which accrues from the extreme sensitivity of the device characteristics to small structural variations. At present, this may preclude the use of these devices in integrated circuits. However, this is a technological problem and newly emerging technologies such as scanning tunneling microscopy tools for lithography may make it possible in the future to exercise tighter control over device dimensions which will mitigate this problem to a large extent.

On the positive side, quantum devices exhibit multifunctionality and in some instances are capable of performing the functions of an entire circuit within a single element. This is a major advantage—both in terms of circuit size and speed—and needs to be investigated further. The power-delay product for these transistors can also be very low due to the small threshold voltages. Finally, the small value of the threshold voltage opens up the intriguing possibility of switching these devices optically rather than electronically thereby eliminating the RC time constant limitation on switching. In view of all these, it appears that although there are some serious drawbacks, there is still enough promise in quantum interference transistors to warrant serious consideration.

ACKNOWLEDGMENTS

The authors are indebted to Professor Supriyo Datta and Professor M. Yamanishi for many insightful discussions. This work was supported by the Air Force Office of Scientific Research under Grant No. AFOSR 88-0096 and by IBM.

APPENDIX

In this appendix, we derive the relationships between the various elements of the Shapiro matrix using only the unitarity of the matrix. The Shapiro matrix is given by

$$\begin{pmatrix} -(a+b) & \sqrt{\epsilon^*} & \sqrt{\epsilon^*} \\ \sqrt{\epsilon} & a & b^* \\ \sqrt{\epsilon} & b & a \end{pmatrix}.$$

From orthogonality of the various columns,

$$|\epsilon| + ab + ba^* = 0. \tag{A1}$$

Using the relation

$$b = a + e^{i\nu}, \tag{A2}$$

we obtain from Eq. (A1)

$$|\epsilon| + a(a + e^{i\nu}) + a^*(a + e^{i\nu}) = 0. \tag{A3}$$

Let $a = |a|e^{i\alpha}$. This reduces the above equation to

$$|\epsilon| + |a|^2(1 + e^{i2\alpha}) + |a|e^{i\nu}2\cos\alpha = 0, \tag{A4}$$

or

$$|\epsilon| + |a|^2[1 + \cos(2\alpha) + i\sin(2\alpha)] + |a|2\cos\alpha(\cos\nu + i\sin\nu) = 0. \tag{A5}$$

Equating real and imaginary parts to zero, we obtain

$$|a|\sin\alpha\cos\alpha + \sin\nu\cos\alpha = 0,$$

$$|\epsilon| + 2|a|^2\cos^2\alpha + 2|a|\cos\alpha\cos\nu = 0. \tag{A6}$$

Assuming $\alpha \neq \pi/2$, the first of the above two equations gives

$$-|a|\sin\alpha = \sin\nu, \tag{A7}$$

and the second gives

$$|\epsilon| + 2|a|^2\cos^2\alpha + 2|a|\cos\alpha\sqrt{1 - |a|^2\sin^2\alpha} = 0. \tag{A8}$$

Transposing and squaring Eq. (A8), we get

$$|\epsilon|^2 + 4|a|^4\cos^4\alpha + 4|\epsilon||a|^2\cos^2\alpha = 4|a|^2\cos^2\alpha(1 - |a|^2\sin^2\alpha), \tag{A9}$$

or

$$|a|^4 + |a|^2(|\epsilon| - 1) + |\epsilon|^2/4\cos^2\alpha = 0. \tag{A10}$$

The solution of the last equation is

$$|a|^2 = \frac{1 - |\epsilon| \pm \sqrt{1 + |\epsilon|^2 - 2|\epsilon| - |\epsilon|^2\sec^2\alpha}}{2}. \tag{A11}$$

Since $|a|$ cannot be greater than 0.5 (otherwise the unitarity of the Shapiro matrix will be violated), we can admit only the root with the negative value of the radical. Therefore,

$$|a|^2 = \frac{1 - |\epsilon| - \sqrt{1 - 2|\epsilon| - |\epsilon|^2\tan^2\alpha}}{2}. \tag{A12}$$

Combining Eqs. (A7) and (A12), we get

$$\nu = \sin^{-1}\left[\left(\frac{1-|\epsilon|-\sqrt{1-2|\epsilon|-|\epsilon|^2\tan^2\alpha}}{2}\right)^{0.5}\sin\alpha\right], \tag{A13}$$

or

$$\nu = \pi - \sin^{-1}\left[\sin\alpha \times\left(\frac{1-|\epsilon|-\sqrt{1-2|\epsilon|-|\epsilon|^2\tan^2\alpha}}{2}\right)^{0.5}\right]. \tag{A14}$$

The second solution for ν is the one that preserves the unitarity of the Shapiro matrix, so that ν is given by Eq. (A14).

Referring back to Eq. (A12), we find that $|a|$ (which is the "magnitude" of a quantity) has to be real, the phase α cannot be chosen arbitrarily. We therefore have to enforce the condition that the discriminant in Eqs. (A12) be positive, i.e.,

$$1-2|\epsilon|-|\epsilon|^2\tan^2\alpha \geqslant 0. \tag{A15}$$

The three pertinent relations relating the elements of the Shapiro matrix are therefore

$$|a|^2 = \frac{1-|\epsilon|-\sqrt{1-2|\epsilon|-|\epsilon|^2\tan^2\alpha}}{2}, \tag{A16}$$

$$\nu = \pi - \sin^{-1}\left[\sin\alpha \times\left(\frac{1-|\epsilon|-\sqrt{1-2|\epsilon|-|\epsilon|^2\tan^2\alpha}}{2}\right)^{0.5}\right], \tag{A17}$$

$$1-2|\epsilon|-|\epsilon|^2\tan^2\alpha \geqslant 0. \tag{A18}$$

[1] For a review of the physics of mesoscopic structures, see, for example, Y. Imry, *Directions in Condensed Matter Physics*, edited by G. Grinstein and G. Mazenko (World Scientific, Singapore, 1986), pp. 101–163.

[2] S. Datta, *Extended Abstracts of the 20th (1988 International) Conference on Solid State Devices and Materials* (Japan Society of Applied Physics, Tokyo, Aug., 1988), pp. 491–494; Talk at the Air Force Office of Scientific Research Workshop on Quantum Devices, Atlanta, Sept. 15–16 (1988).

[3] S. Datta, Superlattices and Microstructures **6**, 83 (1989).

[4] F. Sols, M. Maccuci, U. Ravaioli, and K. Hess, Appl. Phys. Lett. **54**, 350 (1989); *Nanostructure Physics and Fabrication*, edited by M. A. Reed and W. P. Kirk (Academic, Boston, 1989), p. 157; J. Appl. Phys. **66**, 3892 (1989).

[5] D. C. Miller, R. K. Lake, S. Datta, M. S. Lundstrom, M. R. Melloch, and R. Reifenberger, *Nanostructure Physics and Fabrication*, edited by M. A. Reed and W. P. Kirk (Academic, Boston, 1989), p. 165.

[6] N. F. Mott and W. D. Twose, Adv. Phys. **10**, 107 (1961).

[7] A. D. Stone, Phys. Rev. Lett. **54**, 2692 (1985); B. L. Al'tshuler, Pis'ma Zh. Eksp. Teor. Fiz. **41**, 530 (1985) [JETP Lett. **41**, 648 (1985)]; P. A. Lee, Physica **140A**, 169 (1986).

[8] This issue was discussed by R. Landauer, Talk at the Air Force Office of Scientific Research Workshop on Quantum Devices, Atlanta, Sept. 15–16 (1988).

[9] H. Sakaki, J. Vac. Sci. Technol. **19**, 148 (1981); S. Bandyopadhyay and W. Porod, Appl. Phys. Lett. **53**, 2323 (1988).

[10] If the two paths do not have equal transmission probabilities, perfect interference may still be obtained but only through the involvement of other multiply reflected paths. The more the number of paths that interfere to produce the minima, the narrower are the minima. If the minima are too narrow, thermal averaging at elevated temperatures may wash them out and effectively destroy the interference.

[11] Usually the drain terminal will be at a lower potential (for electrons) than the gate terminal so that the transmission probabilities of the two paths will not be exactly equal. This can be alleviated by applying a constant magnetic field perpendicular to the plane of the structure (oriented along the proper direction) so that it will exert a Lorentz force on the electrons thereby pushing them towards the gate and balancing the electrostatic attraction of the drain.

[12] B. Shapiro, Phys. Rev. Lett. **50**, 747 (1983). The Shapiro matrix has been widely used in the analysis of three-way splitters; see, for example, M. Büttiker, Phys. Rev. B **32**, 1846 (1985).

[13] S. M. Sze, *Physics of Semiconductor Devices*, 2nd ed. (Wiley and Sons, New York, 1981).

[14] R. Tsu and L. Esaki, Appl. Phys. Lett. **22**, 562 (1983).

[15] H. R. Frohne, M. McLennan, and S. Datta, J. Appl. Phys. **66**, 2699 (1989).

[16] A large current carrying capability is desirable for fast switching applications in digital circuits. See, for example, S. Bandyopadhyay, G. H. Bernstein, and W. Porod, *Nanostructure Physics and Fabrication*, edited by M. A. Reed and W. P. Kirk (Academic, Boston, 1989), p. 183.

[17] Control over feature sizes of ~50 Å may however be possible with the use of scanning tunneling microscope; see, for example, E. J. van Loenen, D. Dijkkamp, A. J. Hoeven, J. M. Lenssinck, and J. Dieleman, Appl. Phys. Lett. **55**, 1312 (1989).

[18] R. Landauer, *Nanostructure Physics and Fabrication*, edited by M. A. Reed and W. P. Kirk (Academic, Boston, 1989), p. 17; Physics Today, Oct. 1989, p. 119.

[19] P. Santhanam, S. Wind, and D. E. Prober, Phys. Rev. Lett. **53**, 1179 (1984).

[20] K. Ishibashi, K. Nagata, S. Ishida, K. Gamo, Y. Aoyagi, M. Kawabe, K. Murase, and S. Namba, Solid State Commun. **61**, 385 (1987).

[21] This is almost a generic feature of quantum interference effects in one-dimensional structures; see, for example, S. Bandyopadhyay and W. Porod, Appl. Phys. Lett. **53**, 2323 (1988).

[22] R. N. Nottenburg, Y. K. Chen, M. B. Panish, D. A. Humphrey, and R. Hamm, IEEE Electron Device Lett. **10**, 30 (1989).

[23] U. K. Mishra, A. S. Brown, and S. E. Rosenbaum, *Technical Digest of the IEEE International Electron Device Meeting* (IEEE, New York, 1988), p. 180.

[24] G. W. Wang, M. Fend, C. L. Lau, C. Ito, and T. R. Lepkowski, IEEE Electron Device Lett. **10**, 186 (1989).

[25] S. J. Jerome Feng and R. E. Goldwasser, IEEE Electron Device Lett. **10**, 412 (1989).

[26] E. R. Brown, T. C. L. G. Sollner, C. D. Parker, W. D. Goodhue, and C. L. Chen, Appl. Phys. Lett. **55**, 1777 (1989).

[27] See Ref. 16.

[28] T. Kobayashi, M. Miyaka, Y. Okazaki, T. Matsuda, M. Sato, K. Diguchi, S. Ohki, and M. Oda, *Technical Digest of the IEEE International Electron Device Meeting* (IEEE, New York, 1988), p. 881; S. Kotani, T. Imamura, and S. Hasuo, *Technical Digest of the IEEE International Electron Device Meeting* (IEEE, New York, 1988), p. 884; Y. Yamane, T. Enoki, S. Sugitani, and M. Hirayama, *Technical Digest of the IEEE International Electron Device Meeting* (IEEE, New York, 1988), p. 894.

[29] J. B. Barner and S. T. Ruggiero, Phys. Rev. Lett. **59**, 807 (1987).

[30] M. Yamanishi, Phys. Rev. Lett. **59**, 1014 (1987); D. S. Chemla, D. A. B. Miller, and S. Schmitt-Rink, Phys. Rev. Lett. **59**, 1018 (1987).

[31] M. Yamanishi, Superlattices and Microstructures, **6**, 403 (1989).

QUANTUM DEVICES BASED ON PHASE COHERENT LATERAL QUANTUM TRANSPORT[1]

S. Bandyopadhyay, G. H. Bernstein and W. Porod
Department of Electrical and Computer Engineering
University of Notre Dame
Notre Dame, Indiana 46556

Phase-coherent lateral transport phenomena hold the promise for many new types of "quantum devices" with vastly improved performance over conventional devices. In this paper, we address the performance of *lateral quantum devices* in which current flows parallel to the interfaces of a heterostructure. In particular, we focus on two specific devices, namely the Aharonov-Bohm interferometer and the recently proposed Quantum Diffraction Transistor, which have ultrafast *extrinsic* switching speed and a tremendous potential for *multi-functionality*.

I. INTRODUCTION

Over the past few years, numerous novel electronic devices have been proposed or demonstrated whose operations rely entirely on quantum transport phenomena. The most widely studied member of this class is the resonant tunneling diode in which electron transport occurs perpendicular to the interfaces of a heterostructure and the current depends on the interference of waves multiply reflected by heterobarriers. More recently, a different genre of devices has emerged (which we refer to as *lateral quantum devices*) in which current flows parallel to heterointerfaces. The inherent advantage of these devices is that the current levels in them can be much higher, which translates into a significant advantage in switching speed when the device is used in an integrated circuit. In "vertical quantum devices", such as the resonant tunneling diode, the current levels are typically low since the current is predominantly due to tunneling through large potential barriers caused by band-edge discontinuities. In contrast, there are no large potential

[1]Supported by the Air Force Office of Scientific Research under grant no. AFOSR-88-0096 and by IBM Faculty Development Awards.

Copyright © 1989 by Academic Press, Inc.
All rights of reproduction in any form reserved.
ISBN 0-12-585000-X

barriers in lateral quantum devices so that the current levels are generally much higher.

In an integrated circuit chip, the switching speed of a device is determined not so much by the intrinsic speed of the device, but rather by the time it takes to charge and discharge the interconnect and device capacitances. This time depends on the current that can be supplied to the capacitances and the (threshold) voltage levels to which the capacitances must be charged. Roughly speaking, the extrinsic switching time τ_s is given by

$$\tau_s = \frac{C_t V_t}{I}, \tag{1}$$

where C_t is the total circuit capacitance, V_t is the threshold voltage and I is the current[2].

It is advantageous to employ such quantum devices in integrated circuit chips that have very low threshold voltages and can carry relatively large currents so that the extrinsic switching speed is high. Lateral quantum devices, whose operations depend on phase-coherent *lateral* transport, are superior in this respect. They not only exhibit larger current carrying capability, but can also have very small threshold voltages[3]. Consequently, the extrinsic switching speed of lateral quantum devices is usually much higher than that of vertical quantum devices.

In the following Sections, we discuss the performance of two different classes of lateral quantum devices. They are the Aharonov-Bohm interferometer and the recently proposed Quantum Diffraction Field Effect Transistor [1].

A. The Aharonov-Bohm interferometer

In the Aharonov-Bohm interferometer, electrons in two contiguous paths are made to interfere by an external electrostatic potential which modulates the current. If the interferometer is *two-dimensional*, which means that each path is a two-dimensional structure (viz. a quantum well), then the current can be made arbitrarily large by increasing the transverse width of the structure. This may result in certain advantages, but not necessarily in the extrinsic switching speed, since increasing the width of the structure to increase current also increases the circuit capacitance. Besides, a more crucial drawback of two-dimensional interferometers is that they do not perform sufficiently well for device applications unless transport is truly ballistic [2]. Ballistic transport is not easy to achieve in devices with

[2]This limitation on the switching speed does not arise if an electronic device can be switched optically. An intriguing scheme for switching an Aharonov-Bohm interferometer optically (based on virtual charge polarization), instead of electronically, has been proposed by M. Yamanishi. (Proc. of the 4th. Intl. Conf. on Superlattices, Microstructures and Microdevices).

[3]An example is the Aharonov-Bohm interferometer (see Ref. 3).

present-day capability. It therefore behooves us to consider realistic disordered structures and examine device performance in the *diffusive* regime.

In diffusive transport, two-dimensional interferometers do not work well but one-dimensional interferometers (in which the interfering paths are quantum wires) work sufficiently well [2]. The primary reason for this is that in 2-d interferometers, there is a two-fold ensemble averaging - over the electron's energy E and the transverse wavevector k_t - whereas in 1-d interferometers, the averaging over k_t is absent. The latter averaging has disastrous results when elastic scattering is operative. Therefore, for real device applications, 1-d interferometers appear to be the inevitable choice, at least for the present.

In Fig. 1, we show the current modulation (due to the electrostatic Aharonov-Bohm effect) in a disordered 1-d GaAs interferometer in the weak localization regime at 77 K. The length of the structure is 1000 Å , the carrier concentration is 1.55 x 10^6 / cm and the impurity concentration is 10^5 / cm. The model for this calculation is the same as that employed in Ref. 4. The voltage over the structure is 36 meV which is the threshold for polar optical phonon emission (onset of strong inelastic scattering) in GaAs. Again, a $\sim$ 70 % modulation of the conductance is found at LN_2 temperature, which may be good enough for device applications.

From Fig. 1, we find that the maximum value of the current is 2.1 μA. This is very large for single-moded quantum devices whose typical cross-sectional area is 100 Å$\times$ 100 Å. The current level of 2.1 μA translates into an effective current density of more than 10^6 A/cm^2 which is about an order of magnitude higher than what can be achieved in resonant tunneling diodes. In addition, *the threshold voltage for switching* of such devices is also very low. The threshold voltage is $\sim$ 7 mV. Therefore, using Equation (1) and assuming that the total circuit capacitance is about 1 fF[4], we find that the extrinsic switching time τ_s is $\sim$ 3 ps. (The intrinsic switching speed of the device is the transit time of electrons which in this case is $\sim$ 230 fs). The 3 ps switching speed is comparable to that of the fastest GaAs and Silicon devices or even Josephson junctions.

Apart from a fast switching speed, Aharonov-Bohm quantum devices have the additional advantage of having high transconductance (even for nanometer feature sizes) which is advantageous for analog applications. For our prototypical structure, the maximum absolute transconductance was 0.45 mS. This is comparable to the highest transconductance that one could obtain with a 100 nm wide GaAs MODFET whose transconductance would rarely exceed 1 S/mm. For such a MODFET with a feature size of 1000 Å, the absolute transconductance will be 0.1 mS which is slightly lower than the transconductance obtainable with Aharonov-Bohm devices.

The most striking feature of the Aharonov-Bohm device however is its ex-

[4]The estimate of 1 fF is optimistic, but certainly realizable. Capacitances of 10^{-17} F have been obtained in "Coulomb Blockade" experiments.

tremely low power-delay product. This quantity is approximately given by

$$PDP \approx C_t V_t^2 \tag{2}$$

which gives a value of 5 x 10^{-20} Joules for a device with C_t = 1fF and V_t = 7 meV. This power delay product is a few orders of magnitude lower than what could be obtained with even Josephson junctions.

Needless to say, the above performance figures that we have calculated are theoretical projections and one must wait for the realization of actual prototype devices to see if the predicted performance is approached. But more importantly, these devices have another intriguing characteristic; they exhibit multi-functionality. For instance, they can be used to realize unipolar complementary operations, single transistor static latches and single-stage differential amplifiers [3]. It is this multi-functionality that is the most attractive feature of quantum devices, especially in view of its impact on alternate architectures such as neural networks or cellular automata.

B. The Quantum Diffraction Field Effect Transistor

Another lateral quantum device that promises extraordinary multi-functionality s the recently proposed Quantum Diffraction Field Effect Transistor (QUADFET[1]. ts schematic is depicted in Fig. 2.

The QUADFET is basically a MODFET where the gate has a narrow slit efined by electron-beam exposure. Electrons incident from the source can diffract ırough the slit, and the diffraction pattern is viewed as the currents collected at arious fingers in a "drain" consisting of closely spaced fingers. Just like in an rdinary diffraction experiment, the diffraction pattern (and hence the currents ›llected at the fingers) can be changed by modulating the width of the slit. An

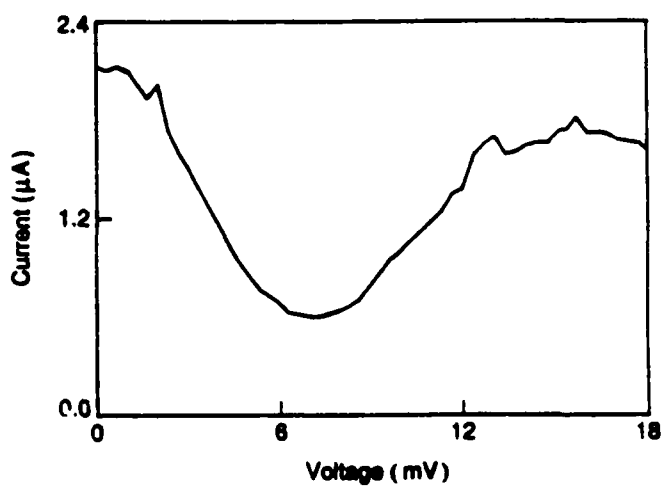


Fig. 1. The transfer characteristic of a disordered *1-d A-B interferometer*

analog voltage applied between the gate pads constricts the slit by extending the depletion layer surrounding the slit. This changes the slit width and alters the diffraction pattern so that the current levels in the fingers are changed. The current level in each finger can be made to represent a "bit" so that one can effectively transduce an analog signal between the gate pads into a bit pattern or a digital signal at the drain fingers. This suggests the operation of an A/D converter. The potential multi-functionality of this device exists in the replacement of an entire A/D circuit by a single transistor.

The problem with the QUADFET however is that it is inherently two-dimensional and therefore requires strictly ballistic transport[5]. In that sense, this device is somewhat futuristic, but assuming that future technology will routinely yield ballistic structures, we can examine the performance of this device, particularly to see if it can be operated at sufficiently high temperatures.

The condition for obtaining a minimum in the diffraction pattern, at any finger in the drain of the QUADFET, is expressed as

$$\frac{ak_F}{2\pi} sin\theta = n + \frac{1}{2} \tag{3}$$

where a is the slit-width.

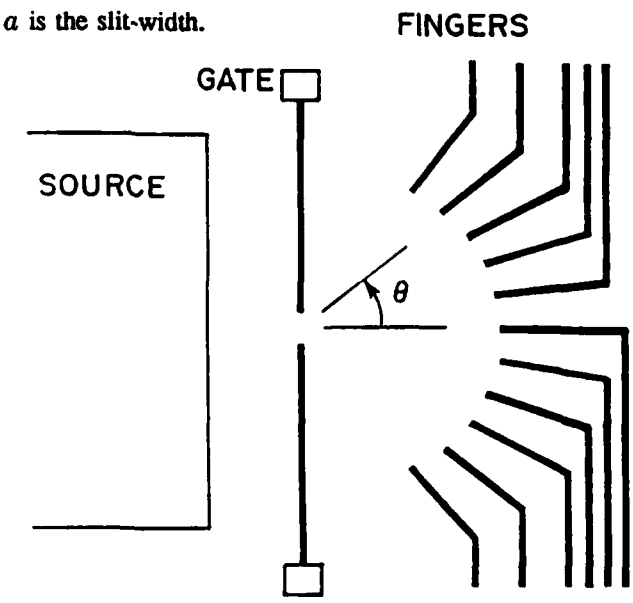


Fig. 2. Schematic of a Quantum Diffraction Field Effect Transistor

[5]The QUADFET requires ballistic transport for the same reason that a 2-d Aharonov-Bohm interferometer requires ballistic transport, namely the ensemble averaging over the electron's transverse wavevector.

In a 2-d structure, the spread in the transverse wavevector introduces a spread in θ, and the thermal spread in energy introduces a spread in the Fermi wavevector k_F, which together tend to make a minimum shallow. For a minimum to be clearly discernible, we require

$$\Delta\left(\frac{ak_F}{2\pi}sin\theta\right) = \frac{ak_F}{2\pi}\int_0^1 d(sin\theta) + \frac{a}{2\pi}sin\theta\int dk_F < \frac{1}{2} \qquad (4)$$

or

$$\frac{ak_F}{2\pi} + \frac{a}{2\pi}\frac{m^*kT}{\hbar^2 k_F} < \frac{1}{2}, \qquad (5)$$

where kT is the thermal spread in the Fermi energy and m^* is the effective mass of electrons. We have replaced $sin\theta$ by its maximum value of unity.

The minimum value of the left-hand-side is obtained when

$$E_F = \frac{kT}{2}; \quad or \quad n_s = \frac{2\pi kTm^*}{h^2}, \qquad (6)$$

where n_s is the carrier concentration. The surprising fact is that for a given carrier concentration, there is an optimum operating temperature.

It turns out coincidentally that in GaAs, for a typical carrier concentration of $\sim 10^{11}$ cm^{-2}, the optimum temperature is $\sim$ 77 K! Of course, for 77 K operation, it is necessary that the device dimensions be smaller than the mean-free-path at LN_2 temperature, but this is already not too far outside the capability of x-ray or electron-beam lithography. Also, for this temperature and carrier concentration, Equation (4) is satisfied with a slit-width $a < 200$ Å and this is achievable with present lithographic capabilities. It therefore appears that the QUADFET will become a viable device for electronic applications in the future.

In conclusion, lateral quantum devices are still in their infancy. But they are likely to play an increasingly important role in electronic circuits of the future.

REFERENCES

1. Kriman, A. M., Bernstein, G. H., Haukness, B. S., and Ferry, D. K. (1988). Presented at the 4th. Intl. Conf. on Superlattices, Microstructures and Microdevices.

2. Bandyopadhyay, S., and Porod, W. (1988). Appl. Phys. Lett. 53, 2323.

3. Bandyopadhyay, S., Datta, S., and Melloch, M. R. (1986). Superlattices and Microstructures. 2, 539.

4. Cahay, M., Bandyopadhyay, S., and Grubin, H. L. (1989), Proc. Intl. Symp. on Nanostructure Physics and Fabrication (this issue).

Properties of the Landauer resistance of finite repeated structures

M. Cahay
Nanoelectronics Laboratory and Department of Electrical and Computer Engineering, University of Cincinnati, Cincinnati, Ohio 45221

S. Bandyopadhyay
Department of Electrical and Computer Engineering, University of Notre Dame, Notre Dame, Indiana 46556
(Received 25 January 1990; revised manuscript received 12 April 1990)

Several properties of the Landauer resistance of finite repeated structures are derived. A theorem relating the energies of unity transmission through a finite repeated structure to the band structure of an infinite superlattice formed by periodic repetition of the finite structure [Vezzetti and Cahay, J. Phys. D 19, L53 (1986)] is generalized to the case of structures with spatially varying effective mass. We also establish a *sum rule* for the Landauer resistances of periodic structures formed by periodically repeating a basic subunit. Finally, we derive an analytical expression for the "boundary resistance" of a structure, as introduced by Azbel and Rubinstein in connection with pseudolocalization, and prove several properties of this quantity.

I. INTRODUCTION

The Landauer formula[1] for calculating the resistance of a dissipationless mesoscopic structure has been used quite widely in the study of quantum transport phenomena. The formula relates in a simple way the resistance of a structure (in the linear-response regime) to the probability of transmission of an electron through the structure. The usefulness of the formula lies in the fact that it reduces the problem of quantum mechanically calculating resistance—a rather difficult problem—to a much simpler problem of calculating just the transmission probability. In this paper, we prove several interesting properties of the Landauer resistance (i.e., the resistance in the linear-response regime) of a finite repeated structure such as a semiconductor superlattice. These properties are all derived from the properties of the transmission coefficient of an electron through a periodic potential of finite spatial extent.[2-5]

In Sec. II of this paper, we first employ a transfer-matrix technique to derive a general expression for the transmission probability of an electron through an arbitrary potential profile. We then extend this result in Sec. III to calculate the transmission probability $|T_N|^2$ of an electron through N subunits of a finite repeated structure. Using this expression, we extend an earlier result[5] relating the energies of unity transmission through a finite repeated structure to the energy–wave-vector relation for an infinite structure formed by periodically repeating the basic subunit of the finite structure. In Sec. IV, we prove a set of theorems that establish interesting and useful relationships between the transmission probabilities (and hence the Landauer resistances) associated with the subunits of a finite repeated structure. These theorems are all illustrated with numerical examples dealing with compositional and effective-mass superlattices. In Sec. V, we establish a *sum rule* for the Landauer resistances of periodic structures formed by successively repeating a basic subunit, and in Sec. VI, we derive an exact analytical expression for the "boundary resistance" of a structure as introduced by Azbel and Rubinstein in connection with pseudolocalization. Finally, in Sec. VII, we summarize our conclusions.

II. TRANSMISSION OF AN ELECTRON THROUGH AN ARBITRARY POTENTIAL

In this section, we first derive an expression for the transmission coefficient of an electron through an arbitrary *one-dimensional* potential of finite spatial extent. For the sake of generality, we allow for spatial variation of the electron's effective mass but assume it varies only in one direction. The time-independent Schrödinger equation describing the steady-state (ballistic) motion of an electron through such a potential is

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

where $E_c(z)$ is the one-dimensional potential that varies in the z direction and $m^*(z)$ is the spatially varying effective mass. In a semiconductor heterostructure, $E_c(z)$ is the conduction-band edge profile which incorporates any band bending due to space charges, variations due to compositional inhomogeneity, and also variations due to any external electric field.

Because the Hamiltonian in Eq. (1) is invariant in the x and y directions, the transverse wave vector $\mathbf{k}_t$ is a good quantum number. Furthermore, since the z component of the electron's motion is decoupled from the transverse motion in the x-y plane, the wave function ψ can be written as

$$\psi=\phi(z)e^{i\mathbf{k}_t\cdot\rho}\,, \quad (2)$$

 ©1990 The American Physical Society

where $\mathbf{k}_t=(k_x,k_y)$ and $\rho=(x,y)$.

The z component of the wave function $\phi(z)$ now satisfies the Schrödinger equation

$$\frac{d}{dz}\left[\frac{1}{\gamma(z)}\frac{d\phi}{dz}\right]+\frac{2m_c^*}{\hbar^2}\{E_p+E_t[1-\gamma(z)^{-1}]-E_c(z)\}\phi(z)=0\ , \quad (3)$$

where m_c^* is the effective mass of the electrons in the "contacts" sandwiching the region of interest (m_c^* is spatially invariant within the contacts and isotropic), $\gamma(z)=m^*(z)/m_c^*$, $E_t=\hbar^2k_t^2/2m_c^*$, and E_p is the kinetic energy associated with the z component of the motion in the contacts ($E_p=\hbar^2k_z^2/2m_c^*$).

The above equation cannot be solved exactly for an arbitrary potential $E_c(z)$. However, an approximate solution can be found by approximating the potential profile by a series of potential steps[6] (see Fig. 1) or by using a piecewise linear approximation for the potential.[7] In the former scheme, the region over which the potential varies is broken down into a finite number of intervals. Within each interval the potential and the effective mass are assumed to be *constant*. In that case, the wave function and its first derivative at the left and right edges of an interval are related through a so-called "transfer matrix," characteristic of that interval, whose elements do not depend on the z coordinate and can be determined analytically.

The transfer matrix for the nth interval $[z_{n-1},z_n]$ is defined according to

$$\begin{bmatrix}\frac{1}{\gamma(z_n^-)}\frac{d\phi}{dz}(z_n^-)\\ \phi(z_n^-)\end{bmatrix}=\begin{bmatrix}W_{11}^{(n)}W_{12}^{(n)}\\ W_{21}^{(n)}W_{22}^{(n)}\end{bmatrix}\begin{bmatrix}\frac{1}{\gamma(z_{n-1}^+)}\frac{d\phi}{dz}(z_{n-1}^+)\\ \phi(z_{n-1}^+)\end{bmatrix}, \quad (4)$$

where $W_{ij}^{(n)}$ are the elements of the transfer matrix, and z_{n-1}^+ and z_n^- stand for $z_{n-1}+\epsilon$ and $z_n-\epsilon$, respectively, with ϵ being a vanishingly small positive quantity. Explicit expressions for the elements of the transfer matrix are given in the Appendix.

Assuming continuity of $\phi(z)$ and $[1/\gamma(z)]/(d\phi/dz)$ everywhere in the structure, the overall transfer matrix W^{tot} describing the entire region $[0,L]$ (see Fig. 1) can be found by simply cascading (multiplying) the individual transfer matrices for the individual intervals:

$$W^{\text{tot}}=W^{(N)}\cdots W^{(1)}\ , \quad (5)$$

where $W^{(n)}$ is the transfer matrix for the nth interval as defined in Eq. (4).

The overall transfer matrix W^{tot} relates the wave functions and their first derivatives at the left and right contacts:

$$\begin{bmatrix}\frac{1}{\gamma(L^+)}\frac{d\phi}{dz}(L^+)\\ \phi(L^+)\end{bmatrix}=W^{\text{tot}}\begin{bmatrix}\frac{1}{\gamma(0^-)}\frac{d\phi}{dz}(0^-)\\ \phi(0^-)\end{bmatrix}. \quad (6)$$

In Eq. (6), $\phi(0^-)$ and $\phi(L^+)$ are the electronic states inside the left and right contacts. They are given by[1,6–10]

$$\phi(z)=\begin{cases}e^{ik_0z}+Re^{-ik_0z}, & z<0\\ Te^{ik_0(z-L)}, & z>L\end{cases} \quad (7)$$

where k_0 $[=(2m_c^*E_p/\hbar)^{1/2}]$ is the z component of the electron's wave vector in the contact and R and T are the overall reflection and transmission coefficients through the region $[0,L]$. Using these scattering states for the wave functions at $z=0^-$ and $z=L^+$ and noting that, by definition, $\gamma(L^+)=\gamma(0^-)=1$, we obtain from Eq. (6)

$$T\begin{bmatrix}ik_0\\ 1\end{bmatrix}=W^{\text{tot}}\begin{bmatrix}ik_0(1-R)\\ 1+R\end{bmatrix}. \quad (8)$$

Equation (8) finally gives us the two equations for the two unknowns T and R. From these two equations T and R can be found by straightforward algebra. Eliminating R gives

$$T=\frac{2ik_0(W_{11}^{\text{tot}}W_{22}^{\text{tot}}-W_{12}^{\text{tot}}W_{21}^{\text{tot}})}{ik_0(W_{11}^{\text{tot}}+W_{22}^{\text{tot}})+(W_{21}^{\text{tot}}k_0^2-W_{12}^{\text{tot}})}\ , \quad (9)$$

where W_{ij}^{tot} are the elements of the matrix W^{tot} that are found from Eq. (5).

Since W^{tot} is a unimodular matrix,[5] the term within parentheses in the numerator of Eq. (9) is unity. In addition (see the Appendix), W_{ij}^{tot} is always purely real. Therefore Eq. (9) gives

$$|T|^2=\frac{4k_0^2}{k_0^2(W_{11}^{\text{tot}}+W_{22}^{\text{tot}})^2+(W_{21}^{\text{tot}}k_0^2-W_{12}^{\text{tot}})^2}\ . \quad (10)$$

The above equation gives us a general expression for the transmission probability of an electron through an arbitrary potential. The transmission probability $|T|^2$ is, of course, related to the reflection probability $|R|^2$ according to the relation $|T|^2+|R|^2=1$ as required by current conservation.

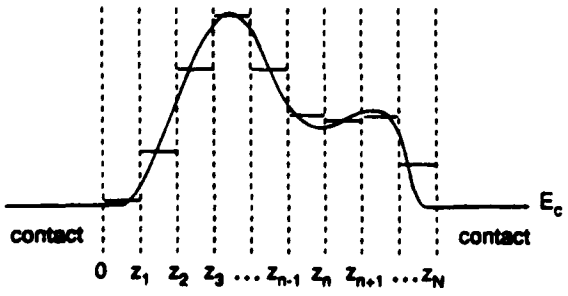


FIG. 1. An arbitrary potential profile approximated by a series of potential steps. Within each interval, the potential and effective mass are assumed to be spatially invariant.

III. TRANSMISSION OF AN ELECTRON THROUGH A FINITE REPEATED STRUCTURE

Having found a general expression for $|T|^2$, we now proceed to evaluate the transmission probability (and hence the Landauer resistance) associated with a finite repeated structure formed by the periodic repetition of a structure with arbitrarily varying potential.

Consider the potential profile in Fig. 2 formed by the

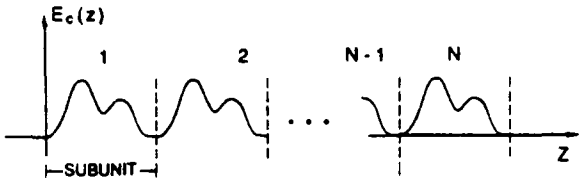


FIG. 2. The potential profile for a finite repeated structure formed by periodic repetition of a region with arbitrarily varying potential.

periodic repetition of an arbitrary potential. Every "period" in this structure has the same transfer matrix (say $\underline{W}$) characterizing that period and the grand overall transfer matrix $\underline{W}^{\text{tot}}$ describing the entire structure is, as before, obtained by cascading the transfer matrices for the individual periods. It is easy to see that for a structure with N periods with each period identical,

$$\underline{W}^{\text{tot}}=(\underline{W})^N \tag{11}$$

As shown in Ref. 10, the elements of the matrix $\underline{W}^{\text{tot}}$ can be expressed in terms of the elements of the matrix $\underline{W}$:

$$\underline{W}^{\text{tot}}=\underline{W}\frac{\sin(N\theta)}{\sin\theta}-\underline{I}\frac{\sin[(N-1)\theta]}{\sin\theta}\ , \tag{12}$$

where $\underline{I}$ is a 2×2 identity matrix and θ depends on the eigenvalues of the matrix $\underline{W}$ and is given by

$$\exp(i\theta)=\lambda_1=\lambda_2^{-1}=\frac{\mathrm{Tr}(\underline{W})}{2}+\left[\left[\frac{\mathrm{Tr}(\underline{W})}{2}\right]^2-1\right]^{1/2}, \tag{13}$$

where $\lambda_{1,2}$ are the eigenvalues of the 2×2 matrix $\underline{W}$ and the second equality follows from the fact that the matrix $\underline{W}$ is unimodular.

We can now find the overall transmission probability $|T_N|^2$ through a periodic structure with N periods. For this, we use Eq. (10) with the elements of $\underline{W}^{\text{tot}}$ now given by Eq. (12). This gives

$$|T_N|^2=\left\{[\sin^2(N\theta)]\left[\left[\frac{k_0^2W_{21}-W_{12}}{2k_0\sin\theta}\right]^2-1\right]+1\right\}^{-1}, \tag{14}$$

which is our main result.

The two- and four-probe (2-p and 4-p) Landauer resistances for a strictly one-dimensional repeated structure can now be found easily by substituting Eq. (14) for the transmission probability $|T_N|^2$ in the single-channel Landauer formula:

$$\begin{aligned}R_L^{2\text{-p}}&=\frac{h}{2e^2}\frac{1}{|T_N|^2}\\&=\frac{h}{2e^2}\left\{[\sin^2(N\theta)]\left[\left[\frac{k_0^2W_{21}-W_{12}}{2k_0\sin\theta}\right]^2-1\right]+1\right\},\\R_L^{4\text{-p}}&=\frac{h}{2e^2}\frac{1-|T_N|^2}{|T_N|^2}=R_L^{2\text{-p}}-\frac{h}{2e^2}\ ,\end{aligned} \tag{15}$$

where k_0 is the wave vector of the incident electron.

IV. TRANSMISSION THEOREMS FOR A FINITE REPEATED STRUCTURE

We now prove a set of theorems related to transmission through finite repeated structures. First, we prove a theorem that relates the energies of unit transmission (i.e., the values of the incident energy for which the transmission coefficient is exactly unity) through a finite repeated one-dimensional structure, to the band structure of the associated infinite lattice formed by periodic repetition of the one-dimensional structure. This theorem was stated for the first time in Ref. 5. A more detailed proof is given here with generalization to the case of a structure with a variable effective mass.

Theorem I. The transmission coefficient of a particle through a periodic structure, formed by N repetitions of a basic subunit, reaches unity at the following energies: (a) energies at which the transmission through the basic subunit is unity, and (b) $N-1$ energies in each energy band of the lattice formed by infinite periodic repetition of the basic subunit, where these $N-1$ energies are given by $E=E_i(k=\pm n\pi/NL)(n=1,2,3,\ldots,N-1$, and L is the length of a period). Here $E_i(k)$ is the energy–wave-vector relation (or the dispersion relation) for the ith band of the infinite lattice.

Part (a) of the theorem is actually fairly obvious. All it states is that by connecting identical structures of transmission unity, one always obtains unit transmission through the composite structure Although this is intuitive, we prove it nevertheless for the sake of completeness. For this, we first note from Eq. (14) that the transmission $|T_N|^2$ through N periods reaches unity when the term within the large square brackets vanishes. The term within the large square brackets vanishes when

$$\left[\frac{k_0^2W_{21}-W_{12}}{2k_0\sin\theta}\right]^2=1\ . \tag{16}$$

We now show that this corresponds to the condition that $|T_1|^2$ (i.e., the transmission through *one* period, or the basic subunit) is unity. Substituting $N=1$ in Eq. (14), we get that the condition for unit transmission through one subunit is given by

$$1=|T_1|^2=\left\{[\sin^2(\theta)]\left[\left[\frac{k_0^2W_{21}-W_{12}}{2k_0\sin\theta}\right]^2-1\right]+1\right\}^{-1}, \tag{17}$$

which, after simplification, reduces exactly to Eq. (16). This proves the first part of the theorem, viz., that the energies of unit transmission through one period are also the energies of unit transmission through all the N periods.

To prove the second part of the theorem, we note from Eq. (14) that the transmission $|T_N|^2$ also reaches unity for those values of θ that satisfy the conditions

$$\sin(N\theta)=0, \quad \sin(\theta)\neq 0 \; ; \tag{18}$$

i.e.,

$$\theta=\pm\frac{\pi}{N},\pm\frac{2\pi}{N},\pm\frac{3\pi}{N},\ldots,\pm\frac{(N-1)\pi}{N} \; . \tag{19}$$

We now have to prove that the above values of θ also correspond to the wave vectors $k=\pm n\pi/NL$ where L is the period. For this, we first apply the Bloch theorem to the infinite structure. The Bloch theorem gives

$$\phi(z+L)=\phi(z)\exp(ikL) \; , \tag{20}$$

where k satisfies the relation[11]

$$\det[W_{ij}-\delta_{ij}\exp(ikL)]=0 \; . \tag{21}$$

In the above equation, W_{ij} is the ijth element of the transfer matrix W describing one period and δ_{ij} is a Kronecker delta. From Eq. (21) we immediately see that $\exp(ikL)$ is the eigenvalue of the 2×2 unimodular matrix $\underline{W}$ and hence

$$\exp(ikL)=\lambda_1=\lambda_2^{-1}=\frac{\mathrm{Tr}(\underline{W})}{2}+\left[\left[\frac{\mathrm{Tr}(\underline{W})}{2}\right]^2-1\right]^{1/2} \; . \tag{22}$$

The right-hand sides of Eqs. (13) and (22) are identical so that their left-hand sides must also be identical. Therefore

$$\exp(ikL)=\exp(i\theta) \; , \tag{23}$$

or

$$kL=\theta(\mathrm{mod}\,2\pi) \; . \tag{24}$$

Consequently whenever $k=\pm n\pi/NL$, the quantity $\theta=\pm n\pi/N$. Thus the energies corresponding to $k=\pm\pi/NL,\pm 2\pi/NL,\pm 3\pi/NL,\ldots,\pm[(N-1)\pi]/NL$ are the energies corresponding to $\theta=\pm\pi/N$, $\pm 2\pi/N,\pm 3\pi/N,\ldots,\pm[(N-1)\pi]/N$, which, in turn, are the energies corresponding to unity transmission through the finite repeated structure with N periods as previously noted. Stated in other words, this means that the energies associated with unity transmission through an N-period structure are the band energies $E(k_n)$ corresponding to the wave vectors $k_n=\pm n\pi/NL$ in an infinite repeated structure. This gives us the $E(k_n)$-versus-k_n relation and proves the theorem.

The usefulness of theorem I lies in the fact that by evaluating the energies of unit transmission through a *finite* structure [which we can do from Eq. (14)], we can calculate the band structure of an *infinite* superlattice formed by the periodic repetition of the finite structure.

The locations of the band edges can be found directly from the following property, which we prove: The states characterized by wave vectors k for which $|\mathrm{Tr}(\underline{W})|>2$ are the evanescent states corresponding to the "stop band" of a finite repeated structure. The states characterized by wave vectors k for which $|\mathrm{Tr}(\underline{W})|<2$ are the propagating states corresponding to the "pass band" of the finite repeated structure.

To prove the property, we invoke Eq. (22). If $|\mathrm{Tr}(\underline{W})|>2$, then the right-hand side of Eq. (22) is purely real and *greater than unity*. In that case, the wave vector k must be purely imaginary which means that the state is an evanescent state corresponding to the "stop band" of the finite repeated structure. On the other hand, if $|\mathrm{Tr}(\underline{W})|<2$, the right-hand side of Eq. (22) becomes complex which permits k to be real. In the latter case, the state is a propagating state corresponding to the "pass band" of the structure. The values of wave vector k for which $|\mathrm{Tr}[\underline{W}]|=2$ evidently correspond to the edges between the pass bands and the stop bands.

Theorem II. At the energies of unity transmission through a finite repeated structure with N periods, the following equality holds: $|T_{N_1}|^2=|T_{N_2}|^2$ whenever $N_1+N_2=N$. Here $|T_{N_1}|^2$ and $|T_{N_2}|^2$ are the transmission probabilities through two subsections with N_1 and N_2 periods respectively.

As stated in the proof of theorem I, the transmission $|T_N|^2$ through N periods reaches unity under two conditions: (a) when the transmission through *each* of the N periods is unity, and (b) when

$$\theta=\pm\frac{n\pi}{N} \quad (n=1,2,3,\ldots,N-1) \; . \tag{25}$$

In case (a), the proof of theorem II is trivial. If the transmission through *each* period is unity, then, of course, the transmission through any arbitrary number of periods is also unity. In that case, obviously,

$$|T_{N_1}|^2=|T_{N_2}|^2=1 \; , \tag{26}$$

regardless of what N_1 and N_2 might be. This proves the theorem for case (a).

The proof for case (b) proceeds as follows. We first note that

$$\begin{aligned}\sin N_1\theta=\sin(N-N_2)\theta&=\sin(\pm n\pi-N_2\theta)\\&=(-1)^{2n+1}\sin N_2\theta \; ,\end{aligned} \tag{27}$$

where we used Eq. (25) to obtain the second equality. Using the above equality in Eq. (14), we immediately see that

$$|T_{N_1}|^2=|T_{N_2}|^2 \tag{28}$$

which proves case (b).

Theorem III. At the energies of unity transmission ($|T_N|^2=1$) through a finite repeated structure with N periods, the following equality holds: $|T_{N+M}|^2=|T_{N-M}|^2$ for all M such that $1\leq M<N$.

The proof of this theorem is very similar to that of theorem II and is therefore not presented.

A. Numerical examples

To illustrate theorem I, we show in Fig. 3 the construction of the energy-band diagram of an infinitely repeated structure whose basic subunit is shown in the inset. The points Q,Q' are the two lowest energies at which the transmission through two subunits is unity, whereas the points P,P' and R,R' are the two lowest energies for which the transmission through three subunits is unity. These points are on the two lowest-energy bands. Other points on the energy-band diagram can be found similarly by steadily increasing the number of periods and searching for the energies of unit transmission. Finally, the points P_1,P_2 and Q_1,Q_2 correspond to the band edges and are found from the condition $|\mathrm{Tr}(\underline{W})|=2$.

To illustrate theorems II and III, we provide the following numerical examples.

Example 1. We have calculated the transmission $|T_N|^2$ [using Eq. (14)] through a compositional superlattice consisting of rectangular wells and barriers in which the barrier and well thicknesses are 50 Å. The effective mass was assumed to be $0.067m_0$ everywhere and the barrier height was taken to be 0.3 eV. Figure 4 shows the transmission coefficient through one, two, and three barriers in the vicinity of the lowest resonant energy through two barriers. (Resonant transmission through two barriers has been studied extensively in connection with the double-barrier resonant tunneling diode.[8,12]) Figure 4 is a clear illustration of theorem III. It shows that when the transmission through two barriers is unity, the transmission through three barriers is equal to the transmission through one barrier, i.e., $|T_{N+M}|^2=|T_{N-M}|^2$ with $N=2$ and $M=1$. Figure 4 also shows that whenever $|T_3|^2=1$, $|T_1|^2=|T_2|^2$, illustrating theorem II for the case $N_1=1$, $N_2=2$.

Example 2. In Fig. 5 we show the transmission through an effective-mass superlattice[13] in which the conduction-band edges in the different layers are assumed to be aligned but the effective masses are different. We assume effective masses of $0.039m_0$ and $0.073m_0$, respectively, in two alternating layers. (These correspond to the effective masses of $In_{0.72}Ga_{0.28}As_{0.86}P_{0.14}$ and InP.[14]) The transmissions through one, two, and three layers were calculated from Eq. (14) at the resonant energy through three layers. Clearly, when $|T_3|^2=1$, $|T_1|^2=|T_2|^2$. This illustrates theorem II. Also when $|T_2|^2=1$, $|T_1|^2=|T_3|^2$ as stated in theorem III.

Theorem IV. If the Fermi energy of a finite repeated one-dimensional structure lies at the boundary between a "pass band" and a "stop band," then the four-probe Landauer resistance of N periods of the structure is equal to N^2 times the four-probe Landauer resistance of one period. This means that the four-probe Landauer resistance increases with the structure's length as L^2 instead

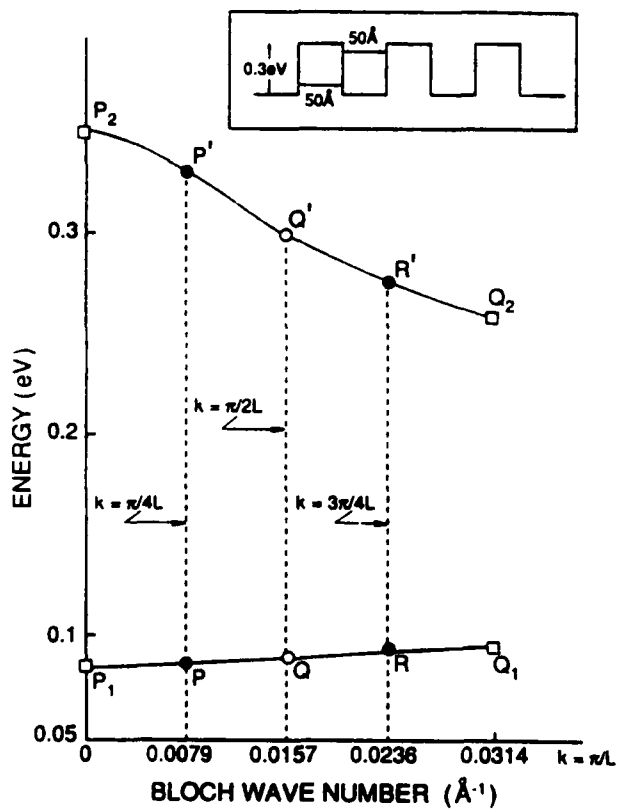


FIG. 3. Energy-band diagram of an infinitely repeated structure whose basic subunit is shown in the inset. The conduction band is constructed by numerically evaluating the energies at which the transmission through increasing number of periods go to unity. The points Q,Q' correspond to the two lowest energies at which transmission through two subunits is unity, whereas the points P,P' and R,R' correspond to the lowest energies for which transmission through three subunits is unity. The points P_1,P_2 and Q_1,Q_2 correspond to the band edges and are found from the condition $\mathrm{Tr}[\underline{W}]=2$.

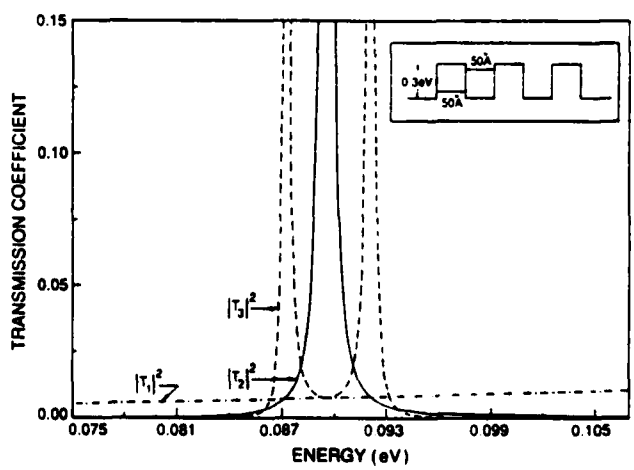


FIG. 4. Transmission coefficients through a periodic structure formed by repeating the subunit shown in the inset. The subunit consists of a GaAs well and an $Al_{1-x}Ga_xAs$ barrier both 50 Å thick. The barrier height is 0.3 eV and the effective mass is assumed to be $0.067m_0$ everywhere. Note that when $|T_3|^2=1$, $|T_1|^2=|T_2|^2$. Also whenever $|T_2|^2=1$, $|T_1|^2=|T_3|^2$. These illustrate theorems II and III, respectively.

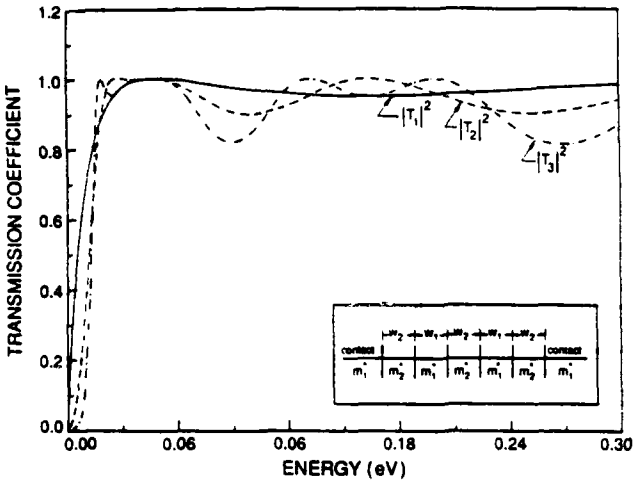


FIG. 5. Transmission coefficient through an effective-mass superlattice. The inset shows the superlattice composed of alternating layers of InP with effective mass $m_1^* = 0.073 m_0$ and $In_{0.72}Ga_{0.28}As_{0.86}P_{0.14}$ with effective mass $m_2^* = 0.039 m_0$. Each layer is 50 Å thick. These figures also illustrate theorems II and III. Note that as stated in theorem II, when $|T_3|^2 = 1$, $|T_1|^2 = |T_2|^2$. Also when $|T_2|^2 = 1$, $|T_1|^2 = |T_3|^2$ illustrating theorem III.

of as L in violation of Ohm's law.

To prove the theorem, we first show that the energies at the boundaries between the pass bands and stop bands of a structure correspond to $\theta = \pm n\pi$, where θ is defined from Eq. (13).

To show this, we first recast Eq. (13) as

$$\exp(i\theta) = \cos\theta + i\sin\theta = \frac{\mathrm{Tr}(\underline{W})}{2} + i\left[1 - \left[\frac{\mathrm{Tr}(\underline{W})}{2}\right]^2\right]^{1/2} . \quad (29)$$

Recalling that at the boundaries (band edges) $|\mathrm{Tr}(\underline{W})| = 2$, we see from the above equation that at the band edges, $\sin\theta = 0$ so that $\theta = \pm n\pi$.

We now obtain a general expression for the four-probe Landauer resistance of one period. Using Eq. (10) for the transmission $|T|^2$ through one period, we obtain

$$\begin{aligned} R_L^{4\text{-p}}(1) &= \frac{h}{2e^2}\frac{1-|T|^2}{|T|^2} \\ &= \frac{h}{2e^2}\left[\frac{(W_{11}+W_{22})^2 + (k^2W_{21} - W_{12})^2}{4k^2} - 1\right] \\ &= \frac{h}{2e^2}\left[\frac{4k^2\cos^2\theta + (k^2W_{21} - W_{12})^2}{4k^2} - 1\right] \\ &= \frac{h}{2e^2}\left[\frac{(k^2W_{21} - W_{12})^2 - 4k^2\sin^2\theta}{4k^2}\right] , \end{aligned} \quad (30)$$

where we used Eq. (29) to substitute for $(W_{11} + W_{22})$ in terms of $\cos\theta$.

From Eq. (15) we also find that the Landauer resistance of N periods is

$$\begin{aligned} R_L^{4\text{-p}}(N) &= \frac{h}{2e^2}\left[\left[\frac{\sin(N\theta)}{\sin\theta}\right]^2 \frac{(k^2W_{21} - W_{12})^2 - 4k^2\sin^2\theta}{4k^2}\right] \\ &= \left[\frac{\sin(N\theta)}{\sin\theta}\right]^2 R_L^{4\text{-p}}(1) . \end{aligned} \quad (31)$$

At the band edges when $\theta = \pm n\pi$, the above expression becomes

$$R_L^{4\text{-p}}(N) = \lim_{\theta \to \pm n\pi}\left[\frac{\sin(N\theta)}{\sin\theta}\right]^2 R_L^{4\text{-p}}(1) = N^2 R_L^{4\text{-p}}(1) , \quad (32)$$

which proves theorem IV. It also shows that since the resistance of N periods is N^2 times (instead of N times) the resistance of one period, the four-probe Landauer resistance increases with the square of the structure's length instead of with its length. This deviation from Ohmic behavior was pointed out in Ref. 15 where it was demonstrated for a periodic array of "δ potentials." In the present treatment, we have generalized it to any arbitrary potential profile.

The L^2 dependence of the resistance is an interesting feature. It is well known that in the pass band, where the states are extended, the resistance should be Ohmic and increase linearly with L, while in the stop band, where the states are localized, it should increase exponentially with L. The fact, that at the boundaries between the pass bands and stop bands the resistance increases as the *square* of L, can be used to identify the onset of *metal-insulator transition*[16] which occurs when the Fermi energy of a system crosses the boundary between a pass band and stop band.

It is also interesting to note from Eq. (31) that the Landauer resistance at the band edges goes to zero when

$$k^2W_{21} - W_{12} = 0 . \quad (33)$$

The above condition is in general not satisfied for any arbitrary potential. Specific cases when this condition is satisfied are discussed in Ref. 17.

V. SUM RULE FOR FOUR-PROBE LANDAUER RESISTANCES

In this section, we prove the following "sum rule" for the four-probe Landauer resistances associated with the subunits of a repeated one-dimensional structure.

The four-probe Landauer resistances of the various subunits of a repeated one-dimensional structure obey the following equality when evaluated at the energies of unit transmission through the structure corresponding to case (b) of theorem I:

$$\sum_{n=1}^{n=N-1} \frac{R_L^{(n)}(1)}{\sum_{m=1}^{m=N-1} R_L^{(n)}(m)} = 1 , \quad (34)$$

where $R_L^{(n)}(m)$ is the four-probe Landauer resistance of a subunit with m periods evaluated at the nth resonant energy (energy of unit transmission). The summation is carried out over all the $N-1$ resonant energies corresponding to case (b) of theorem I.

To prove the sum rule, we make use of Eq. (31). This equation is valid for all N. Replacing N by a running index m and then summing over m, we obtain

$$\sum_{m=1}^{m=N} \sin^2(m\theta) = \sum_{m=1}^{m=N} \frac{R_L(m)}{R_L(1)} \sin^2\theta \ . \tag{35}$$

Note that in the above equation we dropped the superscript "4-p" from $R_L^{4\text{-p}}$ for the sake of brevity. We will follow this convention in all following equations with the tacit understanding that the resistance being referred to is the four-probe rather than the two-probe resistance.

Making use of the trigonometric identity[18]

$$\sum_{m=1}^{m=N} \sin^2(m\theta) = \frac{N}{2} - \frac{1}{2} \frac{\sin(N\theta)\cos[(N+1)\theta]}{\sin\theta} \ , \tag{36}$$

and recalling from Eq. (18) that at the energies of unity transmission through N periods (i.e., at the resonant energies), $\sin(N\theta)=0$, we get from the above identity that at *any* resonant energy

$$\sum_{m=1}^{m=N} \sin^2(m\theta_r) = \frac{N}{2} \ , \tag{37}$$

where θ_r is a resonant value of θ, i.e., $\theta_r = n\pi/N$ where $n = 1,2,3,\ldots,N-1$.

Using the above result to substitute for the left-hand side in Eq. (35) we obtain

$$\begin{aligned} \frac{N}{2} &= \frac{\sum_{m=1}^{m=N} R_L^{(n)}(m)}{R_L^{(n)}(1)} \sin^2(\theta_r) \\ &= \frac{\sum_{m=1}^{m=N-1} R_L^{(n)}(m)}{R_L^{(n)}(1)} \sin^2(\theta_r) \\ &= \frac{\sum_{m=1}^{m=N-1} R_L^{(n)}(m)}{R_L^{(n)}(1)} \sin^2\left[\frac{n\pi}{N}\right] , \end{aligned} \tag{38}$$

where, in deriving the second equality above, we used the fact that at the energy of unity transmission through N periods, the Landauer resistance of N periods is zero, i.e., $R_L^{(n)}(N)=0$.

From Eq. (38), we obtain (by summing over the index n)

$$\begin{aligned} \frac{N}{2} \sum_{n=1}^{n=N-1} \frac{R_L^{(n)}(1)}{\sum_{m=1}^{m=N-1} R_L^{(n)}(m)} &= \sum_{n=1}^{n=N-1} \sin^2\left[\frac{n\pi}{N}\right] \\ &= \sum_{n=1}^{n=N} \sin^2\left[\frac{n\pi}{N}\right] - \sin(\pi) \\ &= \sum_{n=1}^{n=N} \sin^2(n\theta_{r1}) \ , \end{aligned} \tag{39}$$

where θ_{r1} is the value of θ_r at the first resonance, i.e., $\theta_{r1} = \pi/N$.

Comparing Eqs. (37) and (39), we finally obtain Eq. (34)

TABLE I. The sum of the three A's is unity, which satisfies Eq. (34) and, hence, the "sum rule." Note that the energies in the first column correspond to points R', Q', and P' in Fig. 3. $A \equiv (1-|T_1|^{-2})/(3-|T_1|^{-2}-|T_2|^{-2}-|T_3|^{-2})$.

Energy (eV) for which $\lvert T_4\rvert^2=1$	$\lvert T_1\rvert^2$	$\lvert T_2\rvert^2$	$\lvert T_3\rvert^2$	A
0.2798	0.1751	0.09595	0.1751	0.2499
0.3009	0.02357	1.0	0.02357	0.5
0.3304	0.3477	0.2105	0.3477	0.2499

which is the sum rule.

In Table I we provide a numerical example of the sum rule for the case $N=4$ and for the potential profile shown in the inset of Fig. 3. The right-hand side of Eq. (34) does become equal to unity within the numerical accuracy available.

VI. BOUNDARY RESISTANCE OF A FINITE REPEATED ONE-DIMENSIONAL STRUCTURE

While studying pseudolocalization, Azbel and Rubinstein[19] introduced the concept of the "boundary resistance" of a finite repeated one-dimensional structure. The "boundary resistance" of a structure with M periods is the limiting value of the "average resistance" of the structure defined as

$$R_{\text{av}}^M = \frac{1}{M} \sum_{m=1}^{m=M} R_L^{4\text{-p}}(m) \ , \tag{40}$$

where $R_L^{4\text{-p}}(m)$ is the four-probe Landauer resistance of a section composed of m subunits.

Reference 19 showed that in the case of uniformly spaced δ scatterers, the average resistance evaluated at the *resonant energies* of the structure converges to a nonzero constant value independent of the number of periods M (or the length of the structure), provided M is sufficiently large. This constant value was termed the "boundary resistance" since it arises from the effects of the boundaries that break the translational invariance of the structure. In this section we derive, for the first time, an *analytical* expression for the boundary resistance. We then prove two properties—one associated with the boundary resistance, and the other associated with the average resistance evaluated at the energies corresponding to the edges between the pass bands and stop bands of an infinitely periodic structure.

Property 1. The boundary resistance of a periodic structure is indeed independent of the number of periods M (or the length of the structure) and depends only on the potential profile within any one period.

The above result was demonstrated from numerical simulations (but not proved analytically) in Ref. 19 for the specific case of a periodic array of "δ-potentials." In this paper, we provide an *analytical* proof of this property which is valid for any arbitrary shape of the periodic potential.

Property 2. The average resistance of a periodic struc-

ture, evaluated at the edges between the "pass bands" and "stop bands" increases with the *square* of the length of the structure if the number of periods is large.

We first derive an analytical expression for the boundary resistance.

At the resonant energies [for case (b) of theorem I] we have from Eq. (38),

$$\sin^2(\theta_r)=\frac{(M/2)R_L(1)}{\sum_{m=1}^{m=M} R_L(m)} . \tag{41}$$

Therefore using Eq. (41) in Eq. (40) we obtain

$$R_{\rm av}^M(\text{resonance})=\frac{1}{2\sin^2(\theta_r)}R_L(1) . \tag{42}$$

Finally, using Eq. (30) to replace $R_L(1)$ in the above equation, we obtain

$$R_{\rm av}^M(\text{resonance})=\frac{h}{4e^2}\left[\left[\frac{k^2W_{21}-W_{12}}{2k\sin(\theta)}\right]^2-1\right]_{\text{resonance}} , \tag{43}$$

where the quantity in the right-hand side is evaluated at any one of the resonant energies for a structure with M periods.

We now have to prove that the right-hand side is independent of M if M is sufficiently large, i.e., if $M \gg 1$. This will prove property I.

Referring back to Eq. (14), we see that resonance conditions ($T_M=1$) are reached when either the term within the square brackets in the equation above reaches zero[20] or when $\sin(N\theta)=0$.

Case 1. For the former case, i.e., when the term within the square brackets is zero in Eq. (14), $R_{\rm av}^M$(resonance) is identically zero as seen from Eq. (43) and hence obviously independent of M. Thus, we have proved property 1 for this special case.

Case 2. When $\sin(N\theta)=0$ but the term within the large curly braces is nonzero, the value of $R_{\rm av}^M$(resonance) is not zero. The dependence of this nonzero value on the number of periods M enters through only four quantities—the matrix elements W_{21} and W_{12}, and also k_r^M ($=n\pi/ML$) and θ_r^M ($=n\pi/M$), where the last two quantities are the resonant values of the wave vector and the corresponding resonant values of θ for a structure with M periods. To prove property 1, we have to merely show that (1) the matrix elements W_{21} and W_{12} are continuous functions of energy, and (2) the difference between k_r^M and k_r^{M+1}, and also θ_r^M and θ_r^{M+1}, and hence $R_{\rm av}^M$(resonance) and $R_{\rm av}^{M+1}$(resonance), decreases continuously with increasing value of M. The former fact, namely that the matrix elements are continuous functions of energy, is obvious from the derivation of these elements given in the Appendix. The latter fact follows from the inequality

$$k_r^{M+1}-k_r^M=\frac{n}{L}\left[\frac{1}{M+1}-\frac{1}{M}\right]<\frac{1}{ML} , \tag{44}$$

$$\theta_r^{M+1}-\theta_r^M=n\pi\left[\frac{1}{M+1}-\frac{1}{M}\right]<\frac{\pi}{M} . \tag{45}$$

Hence the differences go to zero as $1/M$ which proves property 1.

We now proceed to prove property 2. When the Landauer resistances are evaluated at the edges between a pass band and a stop band,

$$\begin{aligned} R_{\rm av}^M(\text{band edges})&=\frac{1}{M}\sum_{m=1}^{m=M} R_L(m)\\ &=\frac{1}{M}\sum_{m=1}^{m=M} m^2R_L(1)\\ &=\frac{(M+1)(2M+1)}{6}R_L(1)\\ &\approx\frac{M^2}{3}R_L(1)\quad \text{if } M\gg 1 , \end{aligned} \tag{46}$$

where we used theorem IV to arrive at the second equality.

Hence the average resistance, evaluated at the band edges, increases as the square of the length of the structure when the number of periods in the structure is large. This proves property 2.

VII. CONCLUSION

In this paper we have proved several theorems related to the Landauer resistances of finite repeated structures. Of particular importance is the theorem that relates the energies of unity transmission through a finite, repeated one-dimensional structure to the energy–wave-vector dispersion relation for the associated infinite lattice formed by periodic repetition of the structure. This theorem is valid even for a structure with spatially varying effective mass and is therefore very useful in calculating the energy–wave-vector dispersion relation for any infinitely repeated structure.

ACKNOWLEDGMENTS

One of us (S.B.) was supported by the U.S. Air Force Office of Scientific Research under Grant No. AFOSR-88-0096 and by the IBM Corporation.

APPENDIX

In a region where both E_c and γ are constant (spatially invariant), the Schrödinger equation becomes [see Eq. (3)]

$$\frac{d}{dz}\left[\frac{1}{\gamma}\frac{d\phi}{dz}\right]+\frac{2m_c^*}{\hbar^2}\left[E-\frac{E_t}{\gamma}-E_c\right]\phi(z)=0 . \tag{A1}$$

To define the transfer matrix through a section of length L where both E_c and γ are constant, we look at solutions $u(z)$ of Eq. (A1) which satisfy the boundary conditions

$$u_1(0)=0, \quad u_1'(0)=1 , \tag{A2}$$

and

$$u_2(0)=1, \quad u_2'(0)=0 , \tag{A3}$$

where the prime denotes first derivative with respect to space. The solutions $u_{1,2}(z)$ are linearly independent solutions (their Wronskian is unity) and a general solution of Eq. (A1) can be written as

$$\phi(z)=A_1u_1(z)+A_2u_2(z) . \tag{A4}$$

The transfer matrix W is defined as follows:

$$\begin{bmatrix} \frac{\phi'}{\gamma}(L) \\ \phi(L) \end{bmatrix} = W \begin{bmatrix} \frac{\phi'}{\gamma}(0+) \\ \phi(0+) \end{bmatrix} . \tag{A5}$$

Using Equations (A2)–(A5), we obtain

$$W=\begin{bmatrix} u_1'(L) & \frac{u_{2'}}{\gamma}(L) \\ \gamma u_1(L) & u_2(L) \end{bmatrix} . \tag{A6}$$

The explicit forms for $u_{1,2}(z)$ are the following.

Case a. If $E>E_t/\gamma+E_c$,

$$u_1(z)=\frac{\sin\beta z}{\beta} , \tag{A7}$$

$$u_2(z)=\cos\beta z , \tag{A8}$$

where

$$\beta^2=\frac{2m^*}{\hbar^2}\left[E-\frac{E_t}{\gamma}-E_c\right] . \tag{A9}$$

Case b. If $E<E_t/\gamma+E_c$,

$$u_1(z)=\frac{\sinh(\kappa z)}{\kappa} , \tag{A10}$$

$$u_2(z)=\cosh(\kappa z) , \tag{A11}$$

where

$$\kappa^2=\frac{2m^*}{\hbar^2}\left[\frac{E_t}{\gamma}+E_c-E\right] . \tag{A12}$$

[1]R. Landauer, IBM J. Res. Dev. **1**, 223 (1957); Philos. Mag. **21**, 863 (1970).

[2]E. A. Pshenichnov, Fiz. Tverd. Tela (Leningrad) **4**, 1113 (1962) [Sov. Phys.—Solid State **4**, 819 (1962)].

[3]J. Heading, J. Atmos. Terr. Phys. **25**, 519 (1963); Proc. Cambridge Philos. Soc. **74**, 161 (1973).

[4]J. M. Kowalski and J. L. Fry, J. Math. Phys. **28**, 2407 (1987).

[5]D. J. Vezzetti and M. Cahay, J. Phys. D **19**, L53 (1986).

[6]M. Cahay, M. McLennan, S. Datta, and M. S. Lundstrom, Appl. Phys. Lett. **50**, 612 (1987); M. Cahay, S. Bandyopadhyay, S. Datta, and M. S. Lundstrom, *Proceedings of the Second International Conference on Semiconductor Devices and Processes* (Pineridge, Swansea, UK, 1986), pp. 58–67.

[7]H. Ohnishi, T. Inata, S. Muto, N. Yokoyama, and A. Shibatomi, Appl. Phys. Lett. **49**, 1248 (1986).

[8]R. Tsu and L. Esaki, Appl. Phys. Lett. **22**, 562 (1973).

[9]B. Ricco and M. Ya, Azbel, Phys. Rev. B **29**, 1970 (1984); **29**, 4356 (1984).

[10]A. M. Kriman, N. C. Kluksdahl, and D. K. Ferry, Phys. Rev. B **36**, 5953 (1987).

[11]F. Abeles, Ann. Phys. (Paris) **5**, 777 (1950).

[12]L. Esaki, IEEE J. Quantum Electron **QE-22**, 1611 (1986).

[13]A. Sasaki, Phys. Rev. B **30**, 7016 (1984); Surf. Sci. **174**, 624 (1986).

[14]M. Cahay, M. A. Osman, H. L. Grubin, and M. J. McLennan, in *Nanostructure Physics and Fabrication*, edited by M. A. Reed and W. P. Kirk (Academic, Boston, 1989), p. 495.

[15]M. Ya. Azbel, Solid State Commun. **37**, 789 (1981).

[16]The term "metal-insulator transition" is not used here in connection with the Mott transition. Rather it is used in analogy with the fact that if the Fermi energy of a periodic system lies in the stop band (forbidden band), then the resistance is very high (since current can flow only by tunneling) and the structure behaves like an insulator. On the other hand, if the Fermi level lies within a pass band, the states are extended and metallic conduction can take place.

[17]H. A. Kramers, Physica **2**, 483 (1935); see Ref. 4 for some numerical examples.

[18]H. B. Dwight, *Table of Integrals and Other Mathematical Data*, 4th ed. (MacMillan, New York, 1961), pp. 82 and 92.

[19]M. Rubinstein and M. Ya Azbel, Phys. Rev. B **27**, 6484 (1983); M. Ya Azbel and M. Rubinstein, *ibid.* **27**, 6530 (1983).

[20]It is easy to see from Eq. (10) that this condition corresponds to the situation when the transmission through each period is unity.