ISSN 0268-1242

IoP

Semiconductor Science and Technology

Volume 13 Number 8A August 1998

2nd International Workshop on Surfaces and Interfaces of Mesoscopic Devices, Hawaii, 7–12 December 1997

Online: http://www.iop.org

A Journal Recognized by
The European Physical Society

EPS

Institute *of* **Physics** PUBLISHING

Semiconductor Science and Technology

Institute of Physics Publishing is a not-for-profit learned society publisher with a reputation for quality and high standards. It has a comprehensive range of products serving the physics and physics-related communities and in particular is an established leader in the world of scientific journals. Authors of journal articles are supplied with 25 free offprints of their articles and there are no page charges. Readers benefit from the rigorous refereeing procedures, prompt publication times and rapid response to research developments, ensuring that the journals are timely, topical and fully validated.

Internet Services

Institute of Physics Publishing also provides an extensive range of Internet services to support its publication programme. These services are available via the World Wide Web (http://www.iop.org) and e-mail.

Services include:

- *Electronic Journals* - all Institute of Physics Publishing journals are also available online; institutional subscribers to a print journal receive the electronic version for free. You can search, browse, view and print full articles online - a choice of formats, including PDF, is available.
- *Physics Express Letters* - free access to the letters and rapid communications from a range of our online journals. Letters are available in HTML, PDF and PostScript formats.
- *CoDAS Web* - an alerting service containing abstract information from more than 60 journals in condensed matter and materials science. This collaborative project has the support and backing of five major scientific publishers and is co-owned by Elsevier Science and Institute of Physics Publishing.
- *Author Enquiry Service* - track your article through the assessment and production process.
- *Referee Services* - submit your personal information and reviews online.
- *Physics World Jobs* - you can now search for physics posts everywhere and also choose to receive an e-mail alert.
- *PEERS* - our worldwide e-mail directory for physicists enables you to search for colleagues and other useful contacts in your field of interest.
- *PhysicsNet* - a fully indexed and searchable online buyer's guide.
- *Institute of Physics* - find out what the Institute can do for you.
- *Conference Diary* - your online guide to what's on, where it is and when.

We are constantly developing and enhancing our online services. Why not bookmark our home page (http://www.iop.org) and visit regularly to see what's new?

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*

2. REPORT DATE
4-2-99

3. REPORT TYPE AND DATES COVERED
Final Oct. 97 - Sept. 98

4. TITLE AND SUBTITLE
Semiconductor Science and Technology: 2nd International Workshop on Surfaces and Interfaces of Mesoscopic Devices, Hawaii, 7-12 December 1997.

5. FUNDING NUMBERS
N00014-98-1-0042
Mod A00001

6. AUTHORS
J. P. Bird, Editor

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Arizona State University
Department of Electrical Engineering
Tempe, AZ 85287-5706

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5660

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

The Second International Workshop on Surfaces and Interfaces of Mesoscopic Devices (SIMD'97) was held among the luxurious surroundings of the Hyatt Regency Hotel, on the Hawaiian island of Maui, from 7-12 December 1997. Some seventy researchers from around the world gathered in this idyllic setting to report key advances in the physics of mesoscopic devices and their future technological applications. The atmosphere of the workshop was highly stimulating, with all contributions presented in oral format and sufficient time allocated for lively question and answer sessions. The scope of the workshop was also ambitious, although without being over-diluted. Major topics covered included: fundamentals of electron transport in quantum dots and quantum dot arrays, organics on semiconductors, impurities in heterostructures, fabrication of nanostructures, single electron devices, quantum wells and nanodevices.

As Guest Editor of this special issue of Semiconductor Science and Technology, I would like to take this opportunity to thank all of the workshop committee members for their hard work and to gratefully acknowledge the financial support for SIMD'97 from the Arizona State University and the Office of Naval Research. Here's looking forward to doing it all again!

14. SUBJECT TERMS
Semiconductor transport; quantum dots; single electron devices; nanostructures.

15. NUMBER OF PAGES
183

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT
Unclassified

18. SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. SECURITY CLASSIFICATION OF ABSTRACT
Unclassified

20. LIMITATION OF ABSTRACT
UL

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. Z39-1
298-102

DTIC QUALITY INSPECTED 4

19990413 235

Published monthly as twelve issues per annual volume by Institute of Physics Publishing, Dirac House, Temple Back, Bristol BS1 6BE, UK.

Subscription information—1998 volume

For all countries, except the United States, Canada and Mexico, the subscription rate is £829.00 per volume. Single-issue price £63.80 (except conference issues/supplements–prices available on application). Delivery is by air-speeded mail from the United Kingdom to most overseas countries, and by airfreight and registered mail to subscribers in India.

Orders to:

Order Processing Department
Institute of Physics Publishing
Dirac House
Temple Back
Bristol BS1 6BE, UK

For the United States, Canada and Mexico, the subscription rate is US$1686.00 per volume. Delivery is by transatlantic airfreight and onward mailing.

Orders to:

American Institute of Physics Subscriber Services
500 Sunnyside Boulevard
Woodbury
NY 11797-2999, USA

Back issues

Orders and enquiries for the previous volume should be sent to the subscription addresses given above, and for earlier volumes to:
Dawson UK Ltd, Cannon House, Folkestone
CT19 5EE, UK.

United States Postal Identification Statement

Semiconductor Science and Technology (ISSN: 0268-1242) is published monthly by Institute of Physics Publishing, Dirac House, Temple Back, Bristol BS1 6BE, UK in association with the American Institute of Physics, 500 Sunnyside Boulevard, Woodbury, NY 11797, USA. Periodicals Postage Paid at Woodbury, NY 11797, and additional mailing offices. POSTMASTER: Send address changes to: *Semiconductor Science and Technology*, American Institute of Physics, 500 Sunnyside Boulevard, Woodbury, NY 11797, USA.

Copyright ©1998 by IOP Publishing Ltd and individual contributors. All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the written permission of the publisher, except as stated below. Single photocopies of single articles may be made for private study or research. Illustrations and short extracts from the text of individual contributions may be copied provided that the source is acknowledged, the permission of the authors is obtained and IOP Publishing Ltd is notified. Multiple copying is permitted in accordance with the terms of licences issued by the Copyright Licensing Agency under the terms of its agreement with the Committee of Vice-Chancellors and Principals. Authorization to photocopy items for internal or personal use, or the internal or personal use of specific clients, is granted by IOP Publishing Ltd to libraries and other users registered with the Copyright Clearance Center (CCC) Transactional Reporting Service, provided that the base fee of $19.50 per copy is paid directly to CCC, 222 Rosewood Drive, Danvers, MA 01923, USA.

Printed in the UK by William Gibbons & Sons Ltd, Wolverhampton WV13 3XT.

∞™ The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences — Permanence of Paper for Printed Library Materials, ANSI Z39.48-1992.

Semiconductor Science and Technology

Devoted exclusively to semiconductor research and applications, this journal's multidisciplinary approach reflects the far-reaching nature of this topic.

The Materials and Device Reliability section helps to link the research oriented interests of those working on physical aspects of materials and device preparation with the technical interest of those in industry looking to optimize device perfomance.

Editor

G Parry, *Imperial College, London, UK*

Editorial Board

I bar-Jossef, *Weizmann Institute of Technology, Israel*
B Etienne, *CNET, Bagneux, France*
C T Foxon, *University of Nottingham, UK*
V Kulakovskii, *Russian Academy of Sciences, Chernogolovka, Russia*
N Ledentsov, *Ioffe Institute, St Petersburg, Russia*
J C Maan, *University of Nijmegen, The Netherlands*
A Ourmazd, *Institut für Halbleiterphysik, Frankfurt an der Oder, Germany*
K Ploog, *Festkorperelectronik Paul Drude Institut, Berlin, Germany*
S Porowski, *Polish Academy of Science, Warsaw, Poland*
E Rosencher, *Thomson-CSF, Orsay, France*
F Rossi, *Università di Modena, Italy*
L Viña, *Universidad Autónoma, Madrid, Spain*
K von Klitzing, *Max Planck-Institut für Festkörperforschung, Stuttgart, Germany*

American Sub-Board

F Capasso (Chairman), *Bell Labs, Murray Hill, NJ, USA*
S Datta, *Purdue University, West Lafayette, IN, USA*
S Forrest, *Princeton University, NJ, USA*
L Kolodziejski, *Massachusetts Institute of Technology, Cambridge, USA*
U K Mishra, *University of California, Santa Barbara, USA*
U Ravaioli, *University of Illinois at Urbana-Champaign, USA*
K Vahala, *California Institute of Technology, Pasadena, USA*

Japanese Sub-Board

Y Shiraki (Chairman), *University of Tokyo, Japan*
C Hamaguchi, *Osaka University, Japan*
N Sawaki, *Nagoya University, Japan*
S Tarucha, *NTT, Ibaraki, Japan*
N Yokoyama, *Fujitsu Laboratories Ltd, Atsugi, Japan*

Materials and Device Reliability Sub-Board

H E Maes (Section Editor), *IMEC, Leuven, Belgium*
M Ciappa, *ETH Zentrum, Zürich, Switzerland*
H Hartnagel, *Technische Hochschule, Darmstadt, Germany*
C Hu, *University of California, Berkeley, USA*
J W McPherson, *Texas Instruments Inc., Dallas, USA*
E Takeda, *Hitachi Ltd, Tokyo, Japan*
E Wolfgang, *Siemens AG, Munich, Germany*

Publishing Editor

S Quin

Desk Editor

N Couzin

Consultant Editor

V Grigor'yants
IOPP Editorial Office
Ioffe Physico-Technical Institute
26 Polytechnicheskaya
194021 St Petersburg, Russia

Publishing Office

Institute of Physics Publishing
Dirac House, Temple Back
Bristol BS1 6BE, UK

Tel: +44 (0)117 929 7481
Fax: +44 (0)117 929 4318
Internet: prod1@ioppublishing.co.uk

Advertisement Sales

Jack Pedersen/John Irish
D A Goodall Ltd
17 Blossom Street
London E1 6JH

Tel: +44 (0)171 375 1155
Fax: +44 (0)171 375 1115

Semiconductor Science and Technology

Journal scope

Experimental and theoretical studies of the structural, electrical, optical and acoustic properties and the doping of bulk, low-dimensional and amorphous semiconductors; computational semiconductor physics; interface properties, including the physics and chemistry of heterojunctions, metal-semiconductor and insulator-semiconductor junctions; all multi-layered structures involving semiconductor components.

Physics of semiconductor devices, including theoretical modelling and experimental demonstration; all aspects of the technology of semiconductor device and circuit fabrication; growth and preparation of materials, including both epitaxial (e.g. molecular beam and chemical vapour methods) and bulk techniques; the structural, electrical and optical characterization of material and device structures.

Also included are appropriate aspects of surface science; the influence of growth kinetics and chemical processing on layer and device properties; growth models; relevant areas of `molecular electronics' and semiconductor structures incorporating Langmuir-Blodgett films; resists, lithography and metallization where they are concerned with the definition of small geometry structure.

Articles in *Semiconductor Science and Technology* are abstracted in: INSPEC Information Services; Bulletin Signaletique; AEA (UK INIS/IEA Energy Technology Data Exchange); Chemical Abstracts; Engineering Information (Compendex★Plus); INIST (PASCAL); ISI (Science Citation Index, SciSearch, Research Alert, Materials Science Citation Index, Current Contents/Physical, Chemical and Earth Sciences); CoDAS Web; Applied Science and Technology Abstracts; Applied Science and Technology Index.

A brief guide for authors

A submission to *Semiconductor Science and Technology* must be the original work of the author(s) and must not be published elsewhere or under consideration for another publication in its submitted or a substantially similar form in any language. Research Papers (up to 8500 words) and Letters to the Editor (up to 2500 words) will be considered. They may be in English, French or German, but an abstract, title and list of figure and table captions in English must be provided. The following material should be submitted.

1. Three copies of the typescript, at least one of which must be single-sided, each including:
    - (a) Title page with title of article, name(s) of author(s) and address(es) of establishment(s) where work was carried out.
    - (b) Abstract.
    - (c) Text with double line spacing on A4 (210 mm x 297 mm) or similar paper.
    - (d) List of references given in either the numerical or the Harvard (alphabetical) system (on separate sheet).
    - (e) List of captions for illustrations.
    - (f) Set of copies of illustrations.
2. One complete set of illustrations suitable for reproduction. The quality of labelling, line thickness and other detail should be adequate to maintain acceptable size, density and clarity after an appropriate reduction in size. Photographic illustrations should be supplied as glossy prints (not negatives or slides). Graphics files may also be supplied (see below) but good quality originals are still required in case the electronic files cannot be used.
3. Copies of any unpublished or obscure references that may be necessary for the refereeing process.
4. A completed Assignment of Copyright form. Authors unable to include this with their submissions will be sent a copy for signature.

Colour reproduction

Colour reproduction of illustrations is available in four colours or spot colour, for which the author will be asked to pay the additional reproduction costs incurred.

Electronic submissions

Articles for consideration by the journal can now be submitted without the need to send a hard copy. The text of the article can be prepared using Microsoft Word or any common variant of TeX (including LaTeX, REVTeX, AMS-TeX, etc). Figures should be submitted as separate files, preferably in Encapsulated PostScript (EPS) or TIFF formats. The text and figure files should be packaged together into one archive and compressed using a common utility such as PKZip, tar+gzip or Stuffit. Further information, including details of additional formats which can be used after your article has been accepted, is available from http://www.iop.org/Journals/nfa and from the booklet *Notes for Authors* (see below).

Address for submissions

The typescripts and figures should be sent to:

Managing Editor
Semiconductor Science and Technology
Institute of Physics Publishing
Dirac House, Temple Back
Bristol BS1 6BE, UK

Further details

Authors who are submitting to *Semiconductor Science and Technology* for the first time, or who require more details on presentation and style, should consult the booklet *Notes for Authors*, obtainable free of charge from the Journals Publishing Department at the Publishing Office. E-mail requests should be addressed to notes4au@ioppublishing.co.uk. *Notes for Authors* is also available in electronic format via the World Wide Web server (http://www.iop.org).

Foreword

The Second International Workshop on Surfaces and Interfaces of Mesoscopic Devices (SIMD'97) was held among the luxurious surroundings of the Hyatt Regency Hotel, on the Hawaiian island of Maui, from 7 to 12 December 1997. Some seventy researchers from around the world gathered in this idyllic setting to report key advances in the physics of mesoscopic devices and their future technological applications. The atmosphere of the workshop was highly stimulating, with all contributions presented in oral format and sufficient time allocated for lively question and answer sessions. The scope of the workshop was also ambitious, although without being over-diluted. Major topics covered included: fundamentals of electron transport in quantum dots and quantum dot arrays, organics on semiconductors, impurities in heterostructures, fabrication of nanostructures, single electron devices, quantum wells and nanodevices.

As Guest Editor of this special issue of *Semiconductor Science and Technology*, I would like to take this opportunity to thank all of the workshop committee members for their hard work and to gratefully acknowledge the financial support for SIMD'97 from the Arizona State University and the Office of Naval Research. Here's looking forward to doing it all again!

Jonathan Bird
Guest Editor

Semicond. Sci. Technol. **13** (1998) A1–A3. Printed in the UK PII: S0268-1242(98)91858-2

Electronic properties in quantum dots with asymmetric confining potential

T Ezaki, Y Sugimoto, N Mori and C Hamaguchi

Department of Electronic Engineering, Osaka University, 2-1 Yamada-oka, Suita,
Osaka 565, Japan

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. N-electron eigenstates in circular, elliptical and triangular quantum dots are calculated by numerically diagonalizing the N-electron Hamiltonian. The addition energy spectrum for a circular quantum dot shows good agreement with the experimental results reported by Tarucha *et al*. In elliptical quantum dots, the degeneracy of the single-particle states is removed, resulting in a transition of the ground state for a quantum dot with four electrons from the spin-polarized half-filled configuration to the spin-singlet state. The states with 3, 6 and 9 electrons in a triangular quantum dot are found to be slightly more stable compared with a circular quantum dot, which is interpreted in terms of a geometrical effect.

1. Introduction

In quantum dots (QDs) containing a few electrons, the Coulomb interaction plays an important role in determining the energy states, and thus many-particle energy levels depend strongly on the number of electrons in the QDs [1]. Numerical calculations, therefore, require more sophisticated treatments such as an exact diagonalization method [2–8] for the electronic states in QDs to be obtained. In the present work, we calculate N-electron eigenstates in QDs by numerically diagonalizing the N-particle Hamiltonian [8]. We especially focus on the effects of the symmetry of the confining potential on the electronic states.

2. Model and method

We consider a QD fabricated with the $In_{0.05}Ga_{0.95}As$/AlGaAs heterostructure [9, 10]. Electrons are assumed to be confined by an infinite square potential well, $H(z)$, of width W along the z direction. In order to consider anisotropic QDs, we model the confining potential in the x–y plane, $V(x, y)$, by the following form:

$$V(x,y) = \tfrac{1}{2}m^*(\omega_x^2x^2 + \omega_y^2y^2)(1 + \alpha\tfrac{2}{7}\cos 3\phi) \qquad (1)$$

where ω_x and ω_y are the confining energies along the x and y directions, respectively, α (= 0 or 1) is a parameter to specify the shape of the confining potential and ϕ is the angle with respect to the specific axis in the x–y plane. When $\alpha = 0$, the lateral confining potential becomes an ellipse (for $\omega_x \neq \omega_y$) or a circle (for $\omega_x = \omega_y$). A triangularly shaped confining potential can be obtained by setting 1 for the parameter α and putting $\omega_x = \omega_y$. In the present calculation we consider QDs with the average confining energy $\omega_0 \equiv (\omega_x + \omega_y)/2 = 3$ meV and well width of $W = 12$ nm. The material parameters for $In_{0.05}Ga_{0.95}As$ are computed by the linear interpolation of the parameters between InAs and GaAs, which are the effective mass of an electron of $m^* = 0.065m_0$ and the static dielectric constant of $\epsilon = 12.9\epsilon_0$.

The N-electron eigenstates are expanded by the N-electron Slater determinants composed from single-electron eigenfunctions for the two-dimensional isotropic and harmonic system by dividing the lateral confining potential $V(x, y)$ into two parts: one is an isotropic and parabolic potential and the other is the remaining asymmetric part. The confining potential in the polar coordinates (ρ, ϕ) is then given by the following equation:

$$V(\rho,\phi) = \tfrac{1}{2}m^*\tilde{\omega}^2\rho^2 + V'(\rho,\phi) \qquad (2)$$

where the first term on the right-hand side is the isotropic and harmonic potential ($\tilde{\omega} \equiv (\omega_x^2 + \omega_y^2)^{1/2}$) and the second term is the asymmetric part which is given by $V'(\rho,\phi) = \frac{1}{4}m^*(\omega_x^2 - \omega_y^2)\rho^2\cos 2\phi$ for an elliptic QD and $V'(\rho,\phi) = \frac{2}{7}\tilde{\omega}^2\rho^2\cos 3\phi$ for a triangular QD. The N-electron Hamiltonian in the QD can be written as follows:

$$\mathcal{H} = \sum_{i=1}^{N} h_0(\boldsymbol{r}_i) + \sum_{i=1}^{N} V'(\boldsymbol{r}_i) + \sum_{i<j}\frac{e^2}{4\pi\epsilon|\boldsymbol{r}_i - \boldsymbol{r}_j|} \qquad (3)$$

where $h_0(\boldsymbol{r}) = p^2/2m^* + \frac{1}{2}m^*\tilde{\omega}^2\rho^2 + H(z)$ is the single-electron Hamiltonian for the isotropic and harmonic system.

3. Results and discussion

Figure 1 shows the chemical potential difference $\Delta\mu(N) \equiv \mu(N+1) - \mu(N)$, or addition energy, in a circular QD as a function of the number of electrons. Full circles show

0268-1242/98/SA0001+03$19.50

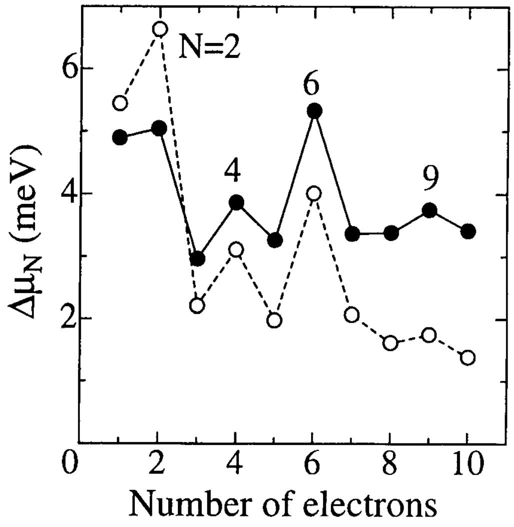


Figure 1. Addition energies in a circular QD as a function of the number of electrons: ●, calculated results for a circular QD with $\omega_0 = 3$ meV; ○, experimental results reported by Tarucha *et al* [9].

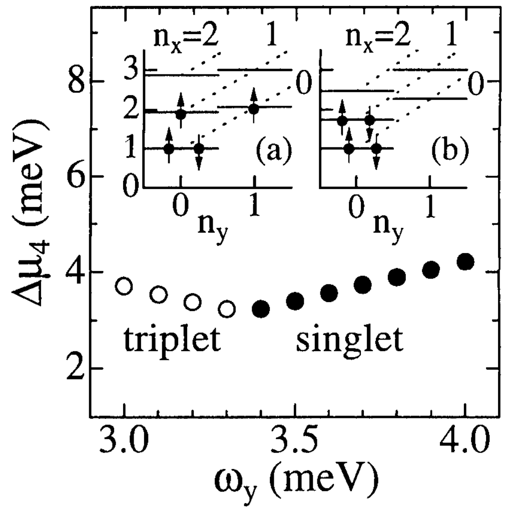


Figure 2. Addition energies for QDs with four electrons as a function of the confining energy ω_y with the average confining energy kept at a constant 3 meV: ○, ●, the spin-triplet and the spin-singlet ground state for the QD with four electrons, respectively. Insets (a) and (b) show schematic diagrams for the single-particle energy levels and electron configuration for the elliptical QDs containing four electrons with confining energies of $(\omega_x, \omega_y) = (2.8\text{ meV}, 3.2\text{ meV})$ and (2.2 meV,3.8 meV), respectively.

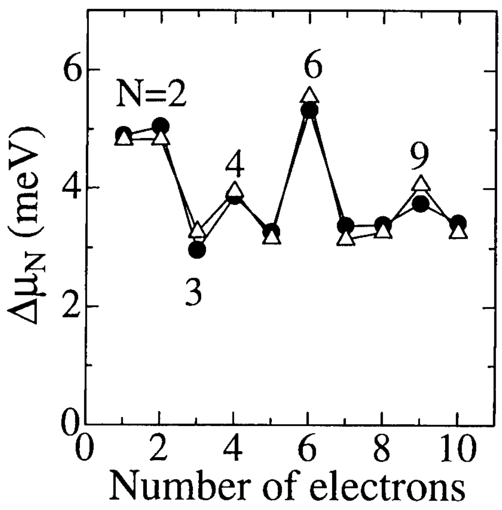


Figure 3. Chemical potential differences in a circular QD (●) and a triangular QD (△) as a function of the number of electrons for $\omega_0 = 3$ meV.

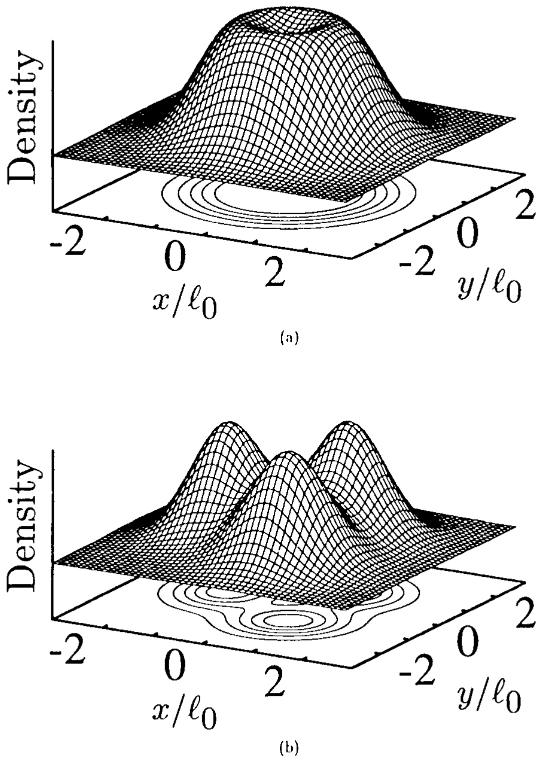


Figure 4. Density distribution of electrons in the QDs containing three electrons with (a) a circularly and (b) a triangularly shaped confining potential in the plane parallel to the heterointerface. l_0 is 20 nm.

the calculated results for the circular QD with a confining energy of $\omega_0 = 3$ meV and open circles the experimental results reported by Tarucha *et al* [9]. The chemical potential is defined by $\mu(N) \equiv E(N) - E(N-1)$ with $E(N)$ being the ground-state energy for N electrons. As seen in figure 1 the addition energy spectrum obtained by the exact diagonalization method agrees well with the experimental results. Addition energy spectra reveal relatively large peaks for $N = 2$ and 6 which are due to the completely filled shell structures [11]. In addition weak structures can be seen at $N = 4$ and 9, which correspond to the spin-polarized half-filled shell structure [9, 12].

Figure 2 shows the addition energy for the elliptical QDs containing four electrons as a function of the confining potential ω_y with the average confining energy being kept at a constant value of 3 meV. Open and full circles correspond to the spin-triplet and the spin-singlet ground states, respectively, for four electrons. For elliptical QDs with weak asymmetry (3 meV $< \omega_y \lesssim$ 3.3 meV), the spin-triplet states remain the lowest eigenstate, and the addition energy decreases with increasing asymmetry owing to the removal of the degeneracy of the single-particle states. For QDs with $\omega_y >$ 3.3 meV, the spin-singlet state becomes the ground state, and the addition energy increases with increasing ω_y. The transition between the spin-triplet and the spin-singlet ground states is explained as follows.

A schematic diagram for the single-particle configuration of the four-electron ground state is plotted in the inset of figure 2 for elliptical QDs with confining energies of (ω_x, ω_y) = (2.8 meV, 3.2 meV) (inset (a)) and (2.2 meV,3.3 meV) (inset (b)). For the elliptical QDs, the single-particle energy is given by $\varepsilon_{n_x n_y} = \hbar\omega_x(n_x + \frac{1}{2}) + \hbar\omega_y(n_y + \frac{1}{2})$ with n_x and n_y being the quantum numbers along the x and y directions, respectively. In the inset of figure 2, the horizontal and the vertical axes represent n_y and $\varepsilon_{n_x n_y}/\hbar\omega_0$, respectively, and the energy levels with the same quantum number n_x are connected by broken lines. For the elliptical QD with confining energies (ω_x, ω_y) = (2.8 meV, 3.2 meV), the energy difference $\Delta\varepsilon$ between the single-particle states $(n_x, n_y) = (0, 1)$ and $(1, 0)$ is $\Delta\varepsilon$ = 0.4 meV, which is slightly smaller than the energy reduction $\Delta_{\mathrm{xc}} \approx$ 0.44 meV due to the exchange-correlation effect, and therefore two electrons with parallel spins occupy the single-particle states $(0, 1)$ and $(1, 0)$ as seen in inset (a). For the elliptic QD with (ω_x, ω_y) = (2.2 meV, 3.8 meV), on the other hand, the energy difference $\Delta\varepsilon$ is 1.6 meV which exceeds Δ_{xc}. The ground state therefore becomes the spin-singlet state as seen in inset (b). Since the energy of the $(1, 0)$ state decreases with increasing asymmetry of the QD, the spin-singlet ground state becomes stable with increasing ω_y as shown in figure 2.

Figure 3 shows $\Delta\mu_N$ for the triangular QD of ω_0 = 3 meV together with $\Delta\mu_N$ for the circular QD (full circles). The addition energies for the circular and the triangular QDs are found to exhibit almost the same characteristics. The addition energy for the triangular QD with N = 3, 6, and 9, however, is found to be slightly larger than that of the circular QD. These features may be explained as follows. The density distribution of electrons in the QDs containing three electrons is plotted in figure 4(a) for the circular QD and figure 4(b) for the triangular QD. As seen in the figure, electrons in the circular QD form a rotationally symmetric distribution, and the electron density becomes small in the centre of the QD because of the electron–electron repulsion. On the other hand, in the triangular QD, each electron moves toward each corner of the triangle and forms a more stable state, giving rise to a slightly larger addition energy. For the same reason, electronic states for N = 6 and 9 in the triangular QD become slightly more stable compared with the case in the circular QD.

4. Conclusion

We calculated N-electron eigenstates in QDs with the vertical confinement of a square quantum well and with lateral confinement of circularly, elliptically and triangularly shaped potentials. The eigenstates are obtained by numerically diagonalizing the N-particle Hamiltonian including the Coulombic interaction. For a circular QD with N = 2 and 6 electrons, the addition energy exhibits a large value, which corresponds to the completely filled shell structure. For a QD with N = 4 and 9 electrons the spin-polarized half-filled shell structure becomes the slightly more stable state owing to the exchange effect between spin-parallel electrons. For elliptical QDs containing four electrons, we found a transition from the spin-triplet state to the spin-singlet state with increasing asymmetry of the QDs. In a triangular QD, slightly stable states resulting from the localization of the electrons at the corners are expected for N = 3, 6 and 9, which may be interpreted in terms of a geometrical effect.

References

[1] Ashoori R C, Stormer H L, Weiner J S, Pfeiffer L N, Baldwin K W and West K W 1993 *Phys. Rev. Lett.* **71** 613

[2] Maksym P A and Chakraborty T 1990 *Phys. Rev. Lett.* **65** 108

[3] Wagner M, Merkt U and Chaplik A V 1992 *Phys. Rev.* B **45** 1951

[4] Kinaret J M, Meir Y, Wingreen N S, Lee P A and Wen X G 1992 *Phys. Rev.* B **46** 4681

[5] Hawrylak P 1993 *Phys. Rev.* B **71** 3347

[6] Palacios J J, Martín-Moreno L, Chiappe G, Louis E and Tejedor C 1994 *Phys. Rev.* B **50** 5760

[7] Yang R and Ruden P P 1995 *J. Appl. Phys.* **78** 1798

[8] Ezaki T, Mori N and Hamaguchi C 1997 *Phys. Rev.* B **56** 6428

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

[10] Austing D G, Honda T and Tarucha S 1996 *Semicond. Sci. Technol.* **11** 388

[11] Macucci M, Hess K and Iafrate G J 1997 *Phys. Rev.* B **55** R4879

[12] Tanaka Y and Akera H 1997 *J. Phys. Soc. Japan* **66** 15

Semicond. Sci. Technol. 13 (1998) A4–A6. Printed in the UK PII: S0268-1242(98)91859-4

Intrinsic stable orbits in open quantum dots

**J P Bird†, R Akis†, D K Ferry†, J Cooper‡||, K Ishibashi‡,
Y Ochiai§, Y Aoyagi‡ and T Sugano‡**

† Center for Solid State Electronics Research, Arizona State University, Tempe,
AZ 85287-6206, USA
‡ Nanoelectronic Materials Laboratory, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama
351-01, Japan
§ Department of Materials Science, Chiba University, 1-33 Yayoi-cho, Inage-ku,
Chiba 263, Japan

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. The conductance of open quantum dot devices is shown to reveal a series of strikingly periodic oscillations as the negative bias applied to their defining gates is varied at zero magnetic field. These surprisingly regular oscillations persist with unaltered characteristics over a wide range of gate voltages, suggesting that the intrinsic transport properties of these devices are dominated by the selective excitation of a small number of stable orbits.

Recently, much interest has focused on the connection between the classical scattering properties of quasi-zero-dimensional quantum dot structures and their low-temperature electrical characteristics [1]. These devices are typically realized using standard surface gate techniques and are connected to external reservoirs by means of quantum point contacts (figure 1, left inset). With these openings configured to support one or more propagating modes, the dot is strongly coupled to the reservoirs and it has been suggested that its discrete level spectrum should be obscured by broadening effects [2, 3]. More recent studies have demonstrated, however, that in the presence of this coupling the main effect is in fact a selective excitation of dot eigenstates by the input point contact [4–9]. In other words, the introduction of coupling between the dot and the charge reservoirs filters the effective density of states that is probed in transport measurements [9]. The origin of this filtering effect in turn lies in the collimating action of the quantum point contacts, which inject electrons into the quantum dot in a highly directed beam [4–6, 9].

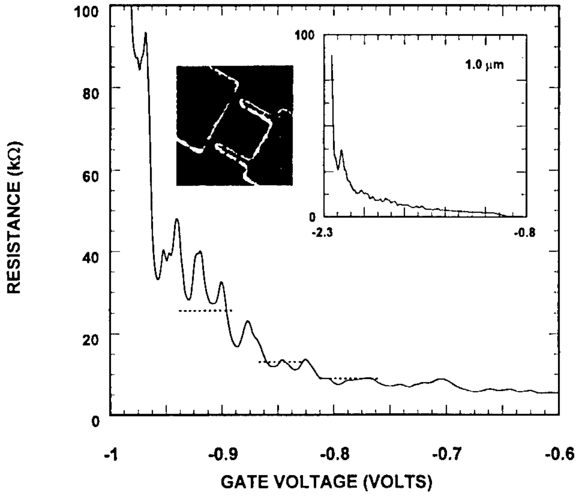


Figure 1. Variation of the resistance of a 0.4 μm quantum dot with gate voltage at zero magnetic field and at a temperature of 0.01 K. Broken lines are merely intended to indicate the expected positions of the resistance plateaux for two quantum point contacts in series which add ohmically. Left inset: SEM micrograph of a 1 μm quantum dot whose gate design is similar to that of all dots studied here. Right inset: variation of the resistance of a 1 μm quantum dot with gate voltage at zero magnetic field and at a temperature of 0.01 K.

Evidence for the selective nature of electron transport in open quantum dots is provided by the results of weak-field magnetotransport experiments. At sufficiently low temperatures, phase coherence of the electrons is maintained over long distances and the magnetoconductance reveals a series of quasi-periodic oscillations [7, 8]. Quantum mechanical simulations of electron transport through these devices reveal the presence of striking wavefunction scarring, the intensity of which oscillates in a similarly periodic manner as the magnetic field is varied [4, 5]. Semiclassically, the scarring is understood to result from the predominance of a few selectively excited orbits, which precess around the dot under the influence of the Lorentz force and which ultimately give rise to the non-uniform accumulation of probability density that the scarring implies. A natural question therefore concerns whether evidence for the

---

|| Present address: Cavendish Laboratory, Madingly Road, Cambridge CB1 0HE, UK.

0268-1242/98/SA0004+03$19.50 © 1998 IOP Publishing Ltd

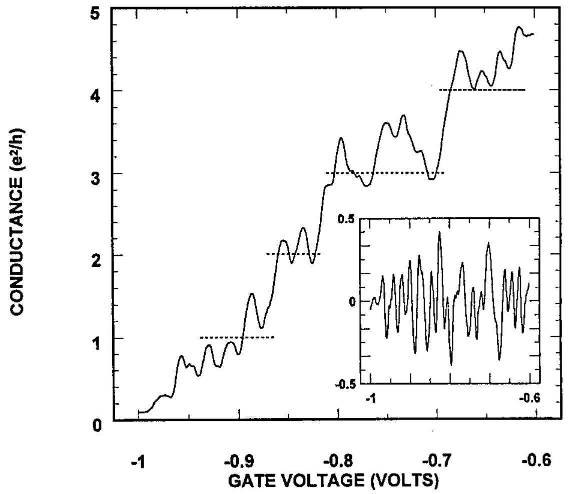


Figure 2. Conductance–gate voltage characteristic for the resistance data shown in figure 1. Broken lines again indicate the positions of expected plateaux for point contacts which add ohmically. Inset: subtraction of a monotonic background from the raw data shown in the main figure yields a strikingly regular, oscillatory component.

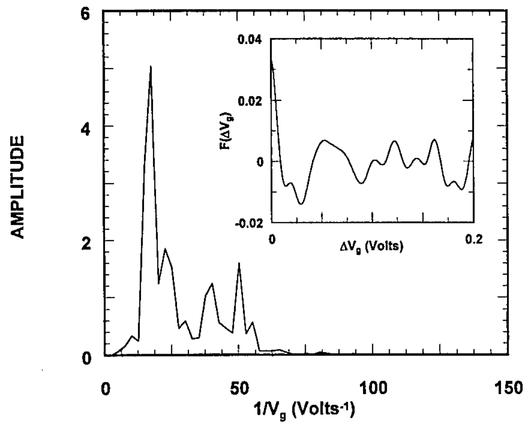


Figure 3. Autocorrelation function of the conductance oscillations shown in the inset to figure 2. In order to avoid the usual confusion associated with poor statistics in the tail, each value of the correlation function is obtained by averaging over a fixed number of data points [7]. Inset: Fourier power spectrum of the conductance oscillations shown in the inset to figure 2.

existence of these stable orbits can be inferred in the absence of an applied magnetic field. In this paper, we address this issue by studying the gate voltage dependent variation of electron transmission through open quantum dots at zero magnetic field. Reminiscent of previous magnetotransport studies [7, 8], the low-temperature conductance of the quantum dots is found to exhibit a series of regular oscillations as a function of gate voltage [10–15]. The oscillations persist over a wide range of gate biases and autocorrelation and Fourier analyses confirm their periodic nature, in turn implying that the zero-field transport properties of these devices are indeed dominated by a small number of stable orbits.

In this report, we summarize the results of studies of more than ten different quantum dots and, while space limitations prevent an exhaustive review, we emphasize that the selected results presented here may be taken as representative of typical device behaviour. The quantum dots were realized in GaAs/AlGaAs heterojunction material using the split-gate technique and ranged in lithographic size from 0.4 to 2 μm (figure 1, left inset) [1]. After illumination with a red LED at liquid helium temperatures, the Hall bars on which the gates were patterned were found to have carrier densities of (4–5) $\times 10^{15}$ m^{-2} and mobilities of 20–70 m^2 V^{-1} s^{-1}. The samples were clamped to the mixing chamber of a dilution refrigerator and their transport properties were measured at zero magnetic field. Standard lock-in techniques were employed and an excitation voltage of less than 3 μV was used for the current-biased measurements.

At temperatures of order a few kelvins or so, sweeping the gate voltage in experiment is found to give rise to a smooth variation of the resistance of the quantum dots, revealing clear evidence for the roughly quantized plateaux that are expected for pairs of quantum point contacts aligned in series [16] (not shown here). On lowering the temperature to well below 1 K, however, these smooth characteristics become obscured by reproducible, resonant-like fluctuations (figure 1). While it might be speculated that the fluctuations somehow arise from the presence of unintentional impurities within the quantum dots, we emphasize that they are found to be a generic feature of low-temperature measurements. A comparison of the results obtained before and after low-temperature illumination suggests that the oscillations do not result from an impurity scattering effect, as might be speculated, but instead reflect the fundamental properties of the quantum dots. In particular, prior to improving the electron mobility with illumination, the oscillations are found to be very strongly suppressed. The oscillations are ubiquitous to illuminated devices, however, and, since illumination is expected to result in improved screening of impurities, this in turn suggests that the oscillations are not related to the presence of unintentional impurities in these devices.

Closer inspection of the fluctuations obtained on sweeping gate voltage reveals a highly periodic nature that persists across a wide range of this parameter (figure 2). In figure 3, we show the results of autocorrelation and Fourier analyses of the conductance oscillations. The correlation function is plotted in the inset of this figure and shows periodic oscillations across its entire tail, as well as significant regions of negative excursion. Consistent with this, the Fourier power spectrum of the same data shows a strong fundamental peak with evidence for second and third harmonics (figure 3). Periodic conductance fluctuations were previously reported in weak-field magnetotransport studies of these devices and their correlation functions and Fourier spectra were found to reveal the same characteristics as those obtained here [7, 8]. In particular, the simple periodicity of the magnetically induced fluctuations was found to result from distinctive scarring of the electron wavefunction, produced by the remnants of a few semiclassical orbits [4, 5]. The strikingly similar characteristics obtained on sweeping either magnetic field or gate voltage suggest that these selected orbits persist in the absence of an applied magnetic field. In other words, the predominance of a few stable orbits in transport is an intrinsic characteristic of these quantum devices. While conductance oscillations as a function of gate voltage have been reported previously [10–15], the underlying periodicity of these and their connection to the selective excitation of stable orbits have not been widely appreciated until now.

At this point, it is instructive to consider the implications of our findings in further detail. When studying the transport properties of open quantum dots, it is frequently assumed that their discrete eigenspectrum is obscured by virtue of the coupling introduced to the external reservoirs [2, 3]. This in turn allows the assumption of ergodicity, in the presence of which the conductance may be expressed as a uniform sum over all possible semiclassical paths. Given such a model, the conductance is then expected to exhibit highly aperiodic fluctuations as a function of magnetic field or gate voltage, either of which may be used to modulate the phase of the electron wavefunctions. Such behaviour is clearly not apparent here (or in magnetotransport studies [7, 8]), indicating that the transport properties of these devices are dominated by the selective excitation of a small number of orbits. As is apparent from the inset to figure 2, the periodicity of the conductance oscillations is largely independent of gate voltage, suggesting these orbits are highly stable as was previously inferred from our magnetotransport studies [7, 8].

In conclusion, the zero-field conductance of open quantum dot devices exhibits strikingly periodic oscillations which persist unaltered across a wide range of gate voltages. The simple periodicity of the oscillations is confirmed by autocorrelation and Fourier analyses and their persistence over a wide range of lead openings suggests that the transport properties of these devices are dominated by the selective excitation of a small number of stable orbits, even in the absence of an applied magnetic field. Future studies will focus on the detailed comparison of these experimental results with realistic quantum mechanical simulations of the devices [17].

References

[1] Bird J P *et al* 1997 *Chaos, Solitons Fractals* **8** 1299, and other articles in the same issue

[2] Jalabert R A, Baranger H U and Stone A D 1990 *Phys. Rev. Lett.* **65** 2442

[3] Marcus C M, Westervelt R M, Hopkins P F and Gossard A C 1993 *Phys. Rev.* B **48** 2460

[4] Akis R, Ferry D K and Bird J P 1996 *Phys. Rev.* B **54** 17 705

[5] Akis R, Bird J P and Ferry D K 1996 *J. Phys.: Condens. Matter* **8** L667

[6] Akis R, Ferry D K and Bird J P 1997 *Phys. Rev. Lett.* **79** 123

[7] Bird J P, Ferry D K, Akis R, Ochiai Y, Ishibashi K, Aoyagi Y and Sugano T 1996 *Europhys. Lett.* **35** 529

[8] Bird J P, Akis R, Ferry D K and Sugano T 1997 *J. Phys.: Condens. Matter* **9** 5935

[9] Ferry D K, Akis R and Bird J P 1998 *Superlattices Microstruct.* **23** 611

[10] Smith C G, Pepper M, Ahmed H, Frost J E F, Hasko D G, Peacock D C, Ritchie D A and Jones G A C 1988 *J. Phys. C: Solid State Phys.* **21** L893

[11] Hirayama Y and Sadu T 1990 *Solid State Commun.* **73** 113

[12] Wu J C, Wybourne M N, Weisshaar A and Goodnick S M 1991 *Appl. Phys. Lett.* **59** 102

[13] Taylor R P, Sachrajda A S, Adams J A, Coleridge P T and Zawadzki P 1993 *Phys. Rev.* B **47** 4458

[14] Liang C-T, Castleton I M, Frost J E F, Barnes C H W, Smith C G, Ford C J B, Ritchie D A and Pepper M 1997 *Phys. Rev.* B **55** 6723

[15] Persson M, Pettersson J, von Sydow B, Lindelof P E, Kristensen A and Berggren K F 1995 *Phys. Rev.* B **52** 8921

[16] Wharam D A, Pepper M, Ahmed H, Frost J E F, Hasko D G, Peacock D C, Ritchie D A and Jones G A C 1988 *J. Phys. C: Solid State Phys.* **21** L887

[17] Akis R and Ferry D K 1998 *Semicond. Sci. Technol.* **13** A18

Semicond. Sci. Technol. **13** (1998) A7–A10. Printed in the UK PII: S0268-1242(98)91861-2

The roles of leads and periodic orbits in the conductance fluctuations of high-mobility quantum dots

I V Zozoulenko†, A S Sachrajda‡, P Zawadzki‡, K-F Berggren†, Y Feng‡ and Z Wasilewski‡

† Department of Physics and Measurement Technology, Linköping University, Linköping, Sweden
‡ Institute for Microstructural Sciences, National Research Council of Canada, Ottawa K1A 0R6, Canada

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. The fluctuations of a square quantum dot have been studied as a function of the dot width. Measurements were made at magnetic fields ranging from the zero-field to edge-state regimes. The experimental results are compared with numerical quantum calculations and semiclassical periodic orbit theory.

It is well established that at low temperatures the magnetoconductance of submicron quantum dots defined in high-mobility 2DEGS exhibits fluctuations and a weak localization peak originating from the wave interference of the phase-coherent ballistic electrons [1]. A great deal of attention both experimental and theoretical has been devoted to the statistical properties of the fluctuations [1–7]. There is, however, a second related question which is being addressed, namely whether one can predict the specific frequencies of the fluctuations from the lithographic

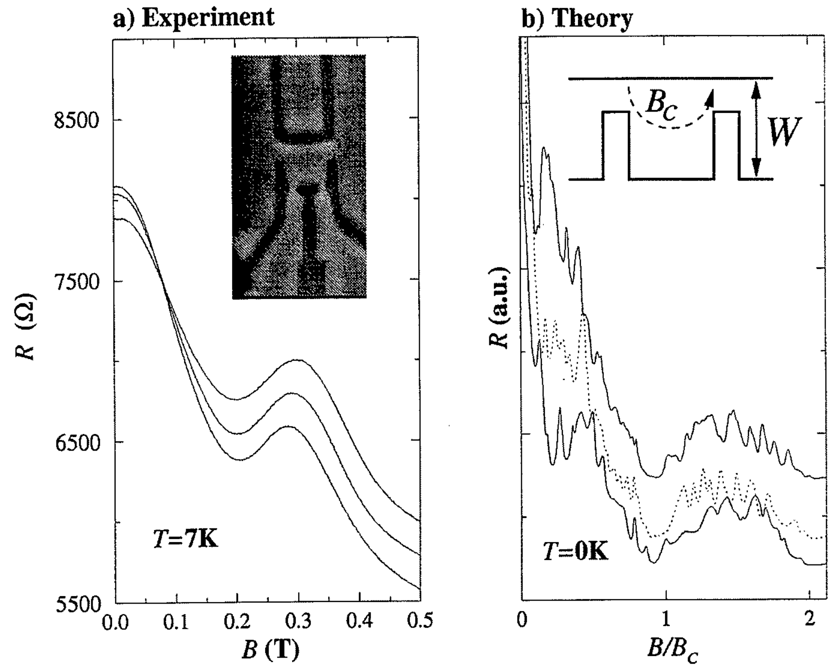


Figure 1. (*a*) Measured magnetoresistance of the dot at $T = 7$ K. The dot (450 nm by 490 nm) is defined by four independently controllable gates; see inset. Different traces correspond to different plunger gate voltages. (*b*) Calculated magnetoresistance of a rectangular dot (see inset) ($L = 500$ nm). The transverse side, $W = 800$, 600 and 400 nm (top to bottom); electron density, $n = 2 \times 10^{15}$ m^{-2}.

0268-1242/98/SA0007+04$19.50

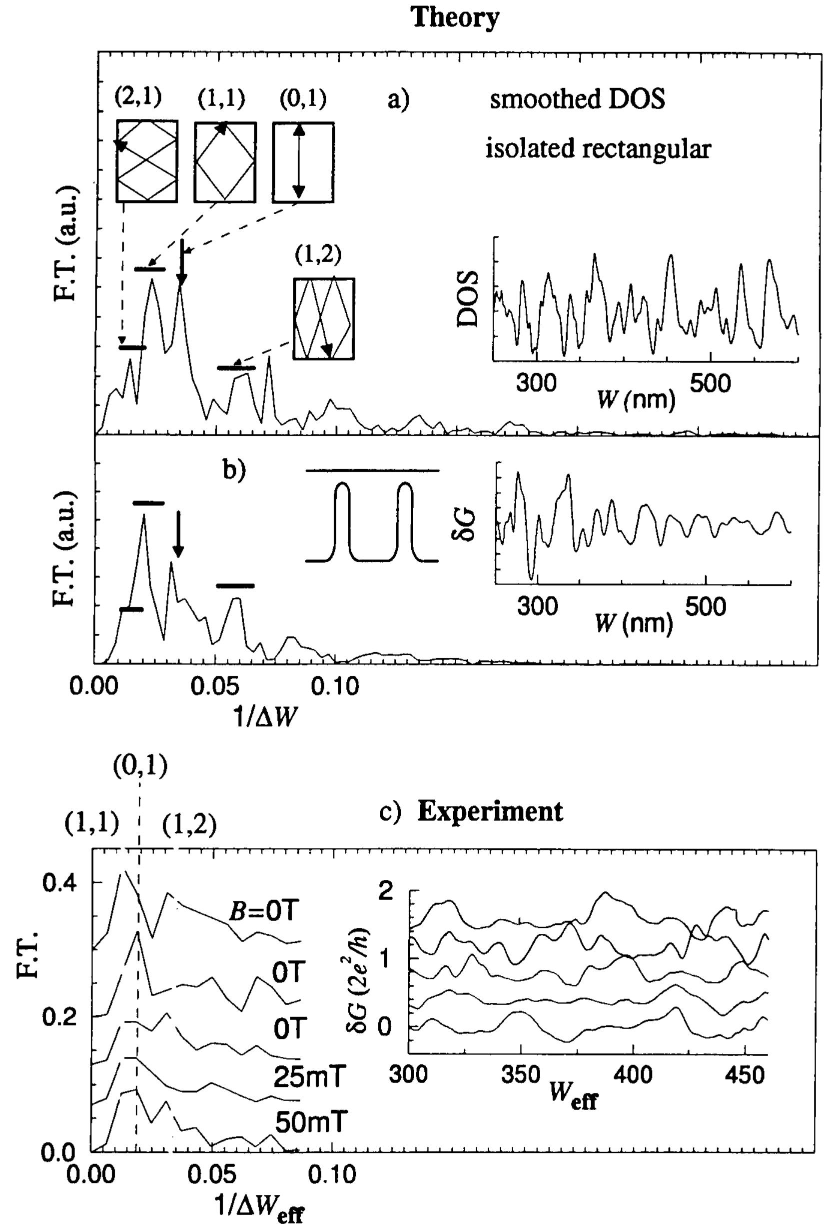


Figure 2. (*a*) Smoothed density of states of an isolated rectangular dot and its Fourier transform (FT) as a function of the transverse size, *W*. Horizontal bars and an arrow mark the expected contributions from several shortest periodic orbits which are schematically depicted in the figure (see text for details). (*b*) Calculated conductance oscillations (with a smooth background subtracted) and their FT for the open dot (depicted in the right inset); $B = 0$. (Parameters of the dots in (*a*), (*b*) are similar to those of figure 1(*b*); corners of the dot are rounded.) (*c*) Measured low-field conductance oscillations (with smooth background subtracted) and their FTs as a function of the effective transverse size, W_{eff}. The three upper curves correspond to different finger gate voltages setting two, three or four modes in the constrictions at $B = 0$ T.

shape of the device [5, 7–14]. In this regard the periodic orbit approach [15, 16] has proven to be very successful in the interpretation of the observed and calculated fluctuation frequencies. These fluctuations are usually observed experimentally as a function of magnetic field [1, 3–7]. This, however, leads to a problem of interpretation. As the magnetic field is raised the zero-field periodic orbits are quickly distorted, become less stable and may eventually vanish [14]. In this paper in contrast we perform measurements as a function of the dot width at *fixed* magnetic fields.

The device consists of a (lithographically) rectangular quantum dot whose geometry is defined by four independently controllable gates. The gates were deposited on a very-high-mobility AlGaAs/GaAs heterostructure. An SEM micrograph of a device similar to the one used is shown in the inset of figure 1(*a*). One gate defined the top edge of the dot. Two finger gates were used to complete the entrance and exit potentials of the dot while a fourth, the plunger gate, was used to control the width of the dot.

Fully quantum calculations are made to compare with the experimental data. These are performed within the Landauer formalism which relates the two-terminal conductance of the device, G, and the transmission coefficient by the formula $G = (2e^2/h)T$. A 'hybrid' recursive Green's function technique [17] (note that the magnetic field is included both in the leads and the dot regions) is used to compute T. In a few cases (e.g. figures 1(*b*) and 3(*b*)) a phenomenological imaginary optical potential is introduced to mimic the effects of temperature and inelastic broadening. Hard-wall confinement and a flat potential inside the dot are assumed for simplicity. Self-consistent calculations based on the Thomas–Fermi approximation have shown that this is a good approximation for the actual confinement of a square dot [18].

In order to achieve the required quantitative analysis of our experiment it was necessary first to calibrate extensively the parameters of the dot itself. The number of modes in the entrance and exit leads could be set accurately because of extensive edge-state backscattering experiments for different applied gate voltage combinations. The classical ballistic features which dominate the dot magnetoresistance at elevated temperatures (>2 K) have been used to extract information on the dot electron density. Figure 1(*a*) shows the experimental results taken at 7 K for three different plunger gate voltages. Results of numerical calculations at $T = 0$ are shown in figure 1(*b*). The gross features are in good agreement. The peak in conductance at B_c corresponds to a conductance enhancement related to the classical magnetic focusing trajectory illustrated in the inset of figure 1(*b*). Similar magnetic focusing features have been seen previously in various geometries [1, 11, 19, 20]. From the position of the peak $2r_c = L$ at $B = B_c$ we estimate the dot density, n, to be 0.67×10^{15} m^{-2}. We used a value of $L = 400$ nm and assumed a constant density. It is important to note that the uncertainty of L had surprisingly little effect on the zero-field interpretation of the characteristic frequencies of oscillations while the above calibrations confirmed that, over our small plunger gate voltage range, the density changed by only a small amount. Measurements made at higher fields (which will be described in detail elsewhere) in the edge-state regime as well as Coulomb blockade measurements were used to relate the variation of the plunger gate voltage to changes in dot size.

Previous related studies have shown that transport through open dots is mediated by the resonant states associated with eigenstates of the corresponding closed structure [9, 11, 12]. This implies that conductance oscillations can be related to the variation of the oscillating part of the quantum mechanical density of states of the isolated dot. In turn, a semiclassical Gutzweiller periodic orbit theory [15, 16] provides a powerful tool to explain and analyse the density of states of an isolated device in terms of several shortest periodic orbits of the corresponding classical system. Consider first, therefore, figure 2(*a*) where we plot the results of a quantum calculation of the smoothed density of states of an *isolated* rectangular dot as a function of its width, W, at $B = 0$. The FT confirms that the density of states contains several components which are directly related to specific families of periodic orbits. These are identified in the figure (note that the orbits can be classified in the usual way by the number of collisions with the vertical and horizontal walls [16]). Of course, certain periodic orbits, e.g. (1, 0), do not contribute to the fluctuations since they are not to first order affected by the plunger gate. Orbits other than the (0, 1) orbit are quasi-periodic since the period varies slightly with the width. However, it is important to stress that the ranges of quasi-periodicity for specific families of periodic orbits do not, as a rule, overlap (the ranges are shown as bars in the figure). Figure 2(*b*) shows the theoretical results for the conductance fluctuations for an open dot and reveals that these are dominated by the same periodic orbits as the isolated dot. The FT of the conductance fluctuations, with a smooth background removed, is found to contain the identical peaks. Finally these results are compared with experiment in figure 2(*c*). The top three curves show the conductance fluctuations and related FTs for three different lead conditions (containing two, three and four modes) at $B = 0$. It can be seen that the experimental results are in good agreement with the theoretical model (note that the same periodic orbits are responsible for the fluctuations). For the calculations, L, and the density, n_s, are $L = 500$ nm and $n_s = 2 \times 10^{15}$ m^{-2}, which are slightly different from the experimental parameters. This is the only reason why the FT peaks appear at slightly different positions for the experimental traces.

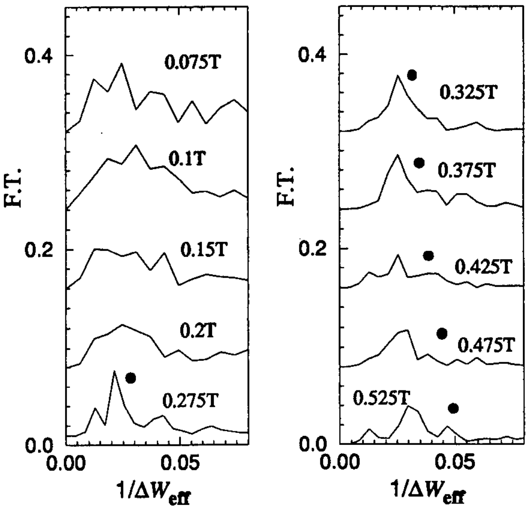


Figure 3. FTs of the measured conductance oscillations in the field region 0.075 T < B < 0.525 T. ●, frequency of the Aharonov–Bohm oscillations.

The fact that the behaviour is very similar for different numbers of modes in the leads is important since it indicates that the fluctuations are fundamentally a property of the dot and not simply a reflection of the coupling of the leads to the eigenstates of the dot. The lower two curves are taken at 25 mT and 50 mT. They reveal that for these very low fields the same periodic orbit governs the fluctuations. At such low fields the periodic orbits are only slightly distorted from their zero-field counterparts. This behaviour only occurs, however, for very low fields. In figure 3 we plot the FT as the magnetic field is increased. Between 0.1 T and 0.2 T the fluctuations have a much smaller amplitude and the FTs become much smoother, containing no significant peaks. While periodic orbit theory is still valid in this regime it is likely that the stability of the orbits is dramatically reduced resulting in a less significant contribution of individual periodic orbits to the fluctuations (i.e. a smaller Liuponov exponent).

At magnetic fields above 0.25 T (i.e. $B > B_c$) a sharp peak develops in the FT. This peak moves gradually to a larger value of $1/\Delta W_{eff}$ as the magnetic field is increased. The full circles mark the position of the period predicted by the Aharonov–Bohm formula ($\Delta W L = 1/B\Phi$). It is clear that the peak follows the Aharonov–Bohm formula. It is important to note that at these fields the transport is still not mediated by edge states and yet the Aharonov–Bohm formula describes the periodicity of the fluctuations convincingly. Finally, as the edge-state regime is reached at higher fields the period obtained is simply that predicted from the Aharonov–Bohm period.

In this paper we have studied conductance fluctuations in a square dot as a function of the width of the dot at fixed magnetic fields. We have discovered different behaviours in a variety of magnetic field regimes.

References

[1] Marcus C M, Rimberg A J, Westervelt R M, Hopkins P F and Gossard A C 1992 *Phys. Rev. Lett.* **69** 506

[2] Jalabert R A, Baranger H U and Stone A D 1990 *Phys. Rev. Lett.* **65** 2442

[3] Berry M J, Katine J A, Marcus C M, Westervelt R M and Gossard A C 1994 *Surf. Sci.* **305** 495

[4] Berry M J, Katine J A, Westervelt R M and Gossard A C 1994 *Phys. Rev.* B **50** 17 721

[5] Persson M, Pettersson J, von Sydow B, Lindelof P E, Kristensen A and Berggren K F 1995 *Phys. Rev.* B **52** 8921

[6] Bird J P, Ishibashi K, Ferry D K, Ochiai Y, Aoyagi Y and Sugano T 1995 *Phys. Rev.* B **52** 8295

[7] Bird J P, Ferry D K, Akis R, Ochiaj Y, Ishibashi K, Aoyagi Y and Sugano T 1996 *Europhys. Lett.* **35** 529

[8] Reiman S M, Person M, Lindelof P E and Brack M 1996 *Z. Phys.* B **101** 377

[9] Berggren K-F, Ji Z-L and Lundberg T 1996 *Phys. Rev.* B **54** 11 612

[10] Schwieters C D, Alford J A and Delos J B 1996 *Phys. Rev.* B **54** 10 652

[11] Zozoulenko I V, Schuster R, Berggren K-F and Enslin K 1997 *Phys. Rev.* B **55** R10 209

[12] Zozoulenko I V and Berggren K-F 1997 *Phys. Rev.* B **56** 6931

[13] Wirtz L, Tang J-Z, Bungdörfer J 1997 *Phys. Rev.* B **56** 7589

[14] Blaschke J and M Brack 1997 *Phys. Rev.* A **56** 182

[15] Gutzwiller M C 1991 *Chaos in Classical and Quantum Mechanics* (New York: Springer)

[16] Brack M and Bhaduri R K 1997 *Semiclassical Physics* (Reading, MA: Addison-Wesley)

[17] Zozoulenko I V, Maaø F A and Hauge E G 1996 *Phys. Rev.* B **53** 7975; 1996 *Phys. Rev.* B **53** 7987

[18] Ouchterlony T, Zozoulenko I V, Wang C-K, Berggren K-F, Gould C and Sachrajda A S to be published

[19] Linke H, Christensson L, Omling P and Lindelof P E 1997 *Phys. Rev.* B **56** 1440

[20] Taylor R P, Sachrajda A S, Adams J A, Coleridge P T and Zawadzki P 1993 *Phys. Rev.* B **47** 4458

Semicond. Sci. Technol. **13** (1998) A11–A14. Printed in the UK

PII: S0268-1242(98)91866-1

The effect of mode coupling on conductance fluctuations in ballistic quantum dots

D P Pivin Jr†, J P Bird†, R Akis†, D K Ferry†, Y Aoyagi‡ and T Sugano‡

† Center for Solid State Electronics Research, Arizona State University, Tempe, Arizona 85287-6206, USA
‡ Nanoelectronics Materials Laboratory, Frontier Research Programme, RIKEN, 2-1 Hirosawa, Wako-shi, Saitama 351-01, Japan

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We study the effect of mode coupling on the magnetoconductance fluctuations of a square ballistic quantum dot fabricated in AlGaAs/GaAs heterostructure material and defined by Schottky gates. Structures prepared allow for the independent variation of the number of propagating input and output modes and their coupling to the main cavity. While adjusting the opening of the input and output point contacts changes the number of modes allowed to propagate into the dot, the excitation from a quantum wire creates an additional selection rule.
The autocorrelation and Fourier analyses of the fluctuations demonstrate that the formation of the quantum wire has a strong effect, not only on the average conductance of the dot but also on the excitation of the semiclassical orbits within it.

1. Introduction

Several groups have recently presented magnetoconductance studies in open quantum dots which have been interpreted surprisingly well using semiclassical arguments [1–3]. Recently, Akis *et al* have used classical and quantum mechanical simulations to enhance the semiclassical interpretations of experiments [4]. These simulations provide striking images of patterns, or scars, in the wavefunctions, which further motivate the application of semiclassical theories. The use of periodic orbit theory [5] for closed systems in magnetic fields has been thought to explain qualitatively the magnetoconductance fluctuations and energy level structure [6, 7]. The emphasis is clear in these recent works; the semiclassical, intuitive notion of classical trajectories dominating the often counterintuitive quantum transport is surprisingly valid.

Even more recently, the focus has been on the striking differences which remain between the predictions of theory for closed systems and the experimental and numerical results for open dots. It has been shown that only a few of the many available classical trajectories are selected for transport by the quantum point contact leads [4, 8–11], owing to the momentum matching conditions imposed on the electrons entering the dot. This manifests itself in the wavefunction as collimation. To some, the imposition of leads on the closed system may be perceived as a nuisance. To others, it presents a further challenge of our understanding and its application to future devices and circuits.

Here, we investigate the role of the mode coupling in the quantum point contact by changing the input from a two-dimensional electron gas (2DEG) to the modes of a quantum wire. Thus, we can modify explicitly the coupling from the environment to the modes of the lead and hence to the dot itself. Both the lead orientation and the lead opening have been varied in recent experiments and have been shown to affect the results [9, 12]. Similar behaviour is expected in these experiments.

2. Experiment

Samples are fabricated on AlGaAs/GaAs heterostructure material with a patterned bulk mobility of 10 m^2 V^{-1} s^{-1} and carrier density of 1.5×10^{15} m^{-2} at 10 mK. Mesa isolation, ohmic contacts and interconnect metallization are patterned and processed using standard fabrication techniques. The final Schottky barrier gates, defining the dot and wire structure, were patterned by electron beam lithography and liftoff. The dot was lithographically designed to be 0.8 μm square with 0.1 μm openings. The wire was designed to be 1.2 μm wide and 6.0 μm long. After depletion, the actual geometry of the dot was inferred from measurement to be 0.6 μm. The device is shown in figure 1.

The characteristics of the wire and dot lead opening are determined at zero magnetic field. The dot openings and wire width are varied and their independent conductances are used as a rough guide for the number of propagating modes at each stage of the structure. Measurements

0268-1242/98/SA0011+04$19.50

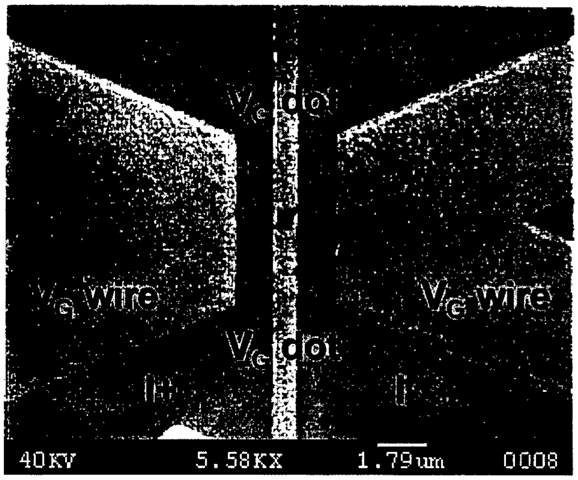


Figure 1. Scanning electron micrograph of the device. Superimposed are labels of the electrical connections used for the experiment.

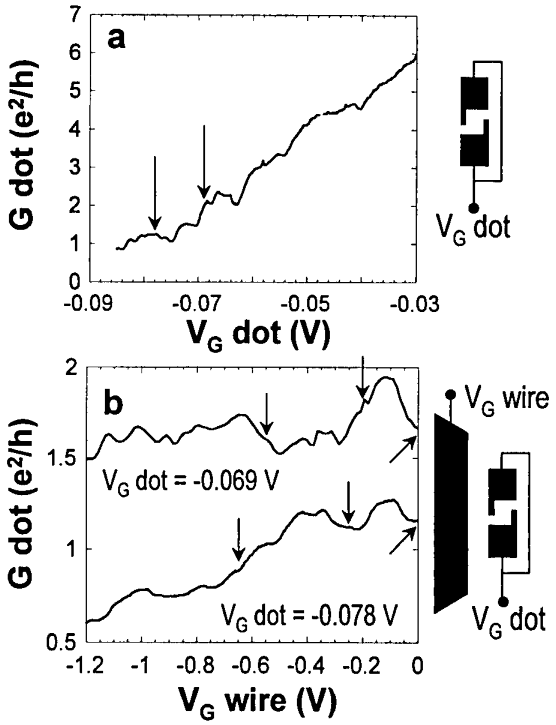


Figure 2. Dot lead opening and wire gate characteristics. Schematics of the measurements are at the right. (*a*) Dot conductance versus dot gate voltage without the wire defined. (*b*) The result of wire gate voltage variation at constant dot lead opening voltage for two values of dot voltage. Arrows in both (*a*) and (*b*) represent bias conditions for the magnetoconductance results presented in figure 3.

of the magnetoconductance of the dot are performed at various configurations of wire width and dot lead opening. For all experiments, at most only one of the quantum wires is defined, providing the equivalent of an input–output impedance mismatch. Measurements are made using standard lock-in techniques at 10 mK in a dilution refrigerator. The overall constant current excitation of the device is kept well below kT to avoid electron heating.

3. Results and discussion

The conductance of the bare dot, with the wire gate voltage set to zero, versus lead opening voltage is shown in figure 2(*a*). In addition to a gradual decrease in the conductance as the dot gate bias is made more negative, reproducible fluctuations are observed. This particular behaviour is discussed by Bird *et al* [13]. In this experiment, the lead opening gates are part of the dot-defining geometry and therefore it is impossible to control the effective geometry of the dot and the opening independently. The result is a situation where the mode number and spread of the collimation are changing as well as the effective area of the dot. Recent magnetoconductance experiments and simulations suggest that the collimation condition does not change significantly enough with lead opening to have a drastic effect on the selected periodic orbits [9]. The major cause of the fluctuations is more likely to be a modulation of the effective area.

At various dot lead voltages, the wire is additionally defined at one side, and its gate voltage is also varied. This produces dot conductance characteristics as shown in figure 2(*b*). The effect is a fairly regular modulation of the conductance as the input–output mode matching conditions of the dot are varied. This modulation depends on the overall conductance. As the dot leads are closed, the magnitude of the wire-induced modulation decreases and is always less than e^2/h. We may consider the wire to be modulating the transmission of each mode in the lead but never completely eliminating any. Therefore, the relative effect of the wire on the system is expected to be greatest when there are between two and five modes propagating in the quantum point contact.

Magnetoconductance studies are performed to characterize the effects of both the lead opening and the wire gate voltage on the dot. At several combinations of lead opening voltage, the effect of the wire on the magnetoconductance is observed. Lead gate voltages are varied so that an estimate for the propagating modes in the quantum point contact is from less than one to three. A representative example of the magnetoconductance results is shown in figure 3. The magnitude of the fluctuations is $\leq 0.2e^2/h$ and decreases with the average conductance of the dot. Note that both lead opening and wire gate voltages affect the nature of the fluctuations.

In most cases magnetoconductance traces are fitted to remove the background behaviour and this fit is subtracted from the data. Autocorrelation functions and Fourier transforms are typically derived from this result. For ballistic quantum dots, the dominant fluctuation periods are thought to be associated with particular periodic trajectories in the system, although the theoretical basis for this is not well developed. An important difficulty is the fact that the magnetoconductance fluctuations are characteristic of a modulation of the density of states at the Fermi energy. The nature of these oscillations, although quasi-periodic, is non-stationary, unlike Aharonov–Bohm oscillations, which

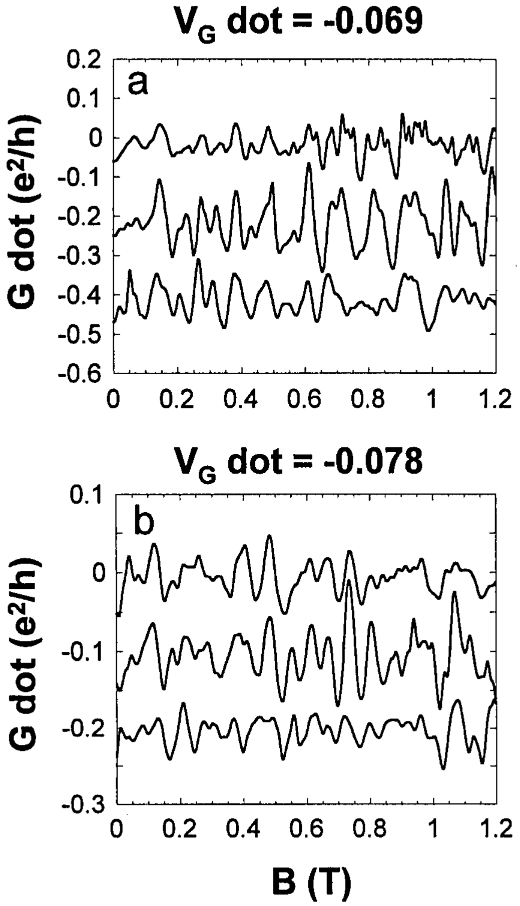


Figure 3. Magnetoconductance results for varying wire voltages at two lead opening voltages. The measurement configuration is shown in figure 1. Only one of the wires is defined at a time. In both (*a*) and (*b*) a polynomial fit was subtracted from the original data. In (*a*) (V_G(dot) = −0.069 V) the data are offset by $0.2e^2/h$; from top to bottom the wire gate voltages are 0, 0.2 and 0.55 V. In (*b*) (V_G(dot) = −0.078 V) the data are offset by $0.1e^2/h$; from top to bottom the wire gate voltages are 0, 0.25 and 0.65 V. Note that, for the narrower lead opening, the amplitude of the fluctuations is roughly halved.

are essentially independent of the absolute field over large ranges. In other words, the Lorentz force bends the classical periodic trajectories and must alter their associated fluctuation periods. Another difficulty is that there are several families of degenerate periodic orbits to consider. As a result, it is practically impossible to determine exactly which trajectories are contributing at a given magnetic field. Nonetheless, analyses such as this, on a finite window of the data, centred at a particular magnetic field have been the accepted tool for describing the fluctuations. The ideal, which is a useful and interesting interpretive result of periodic orbit theory investigations, is a decomposition of the fluctuations into the contributions of each periodic trajectory [6, 7]. This would provide an intuitive semiclassical connection to the experimental results in quantum ballistic semiconductor billiards.

As more detailed experiments are performed on these systems, the ambiguity of the analysis becomes unacceptable. In an effort to remove some of the ambiguity and it is hoped to identify subtle differences in the character of the fluctuations, a slightly more detailed analysis approach is taken here. After fitting the data to remove the background (with a low-order polynomial) autocorrelation functions are generated with a suitably large window centred at many magnetic field values. The Fourier transform of each segment is then assembled in a three-dimensional surface plot. In this way, the dominant periods in the data can be tracked through magnetic field. Although it is by no means a rigorous decomposition of the fluctuations, this analysis provides a useful 'fingerprint' for visualizing and interpreting the non-stationary phenomena.

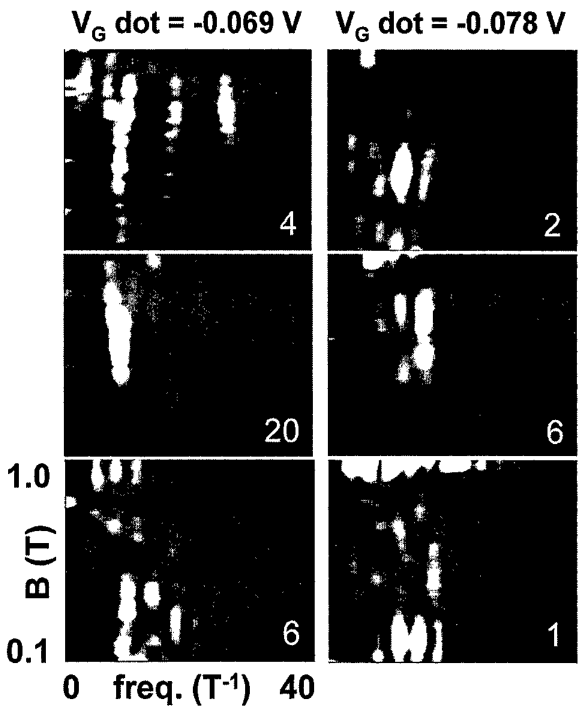


Figure 4. Fourier analysis of the autocorrelation functions for the data of figure 3. The vertical axis of each panel is the centre of the autocorrelation averaging window in magnetic field. Brightness corresponds to higher spectral power. Relative powers corresponding to white are shown in the lower right corner of each panel. The left and right columns correspond to V_G(dot) = −0.069 V (figure 3(*a*)) and V_G(dot) = −0.078 V (figure 3(*b*)) respectively. The rows, from top to bottom, correspond to increasing wire gate voltages as in figure 3. For this sample the ratio of the cyclotron radius to an estimated side of the dot (0.6 μm after depletion) is unity at about 0.1 T. A window of 0.2 T was used for the autocorrelation function centred at many points in magnetic field. The size of this window ensures that several periods in the fluctuations are sampled.

Results of the analysis for the magnetoconductance data of figure 3 are shown in figure 4. From this analysis it is clear that the selection of dominant periods varies with variations of both lead opening and wire width. The specific role of lead openings in the selection of these periods has been discussed previously [9, 10]. As shown in figure 4, the dominant periods can be slightly shifted with a change in the effective dot geometry. It is also clear that the mode matching condition, modulated by the wire, at a constant lead opening, affects the observed

Fourier spectrum. With a change of wire width, the characteristic periods do not apparently change, but their relative strength over the entire range of magnetic field is altered. Also, the overall magnitude of the fluctuations is damped with decreasing wire width. Similar to the situation of the gate characteristics in figure 2(*b*), this effect is greatest when the overall conductance is low. An important observation is that the dominant periods do not appear to vary smoothly with magnetic field. Although it is not at present understood, we emphasize that this effect is not an artefact of the analysis.

The current semiclassical theory for these structures does not normally account for these results. The lead collimation, and resulting selection of classical trajectories, might be accounted for by a modification of the sum of contributing periodic orbits [14]. Recent theoretical approaches have been proposed to account properly for the lead behaviour at zero magnetic field [15, 16]. However, it is unclear exactly what role the quantum mechanical mode matching condition, complicated by the magnetic field, has in the semiclassical picture.

4. Conclusion

In conclusion, we have outlined a visualization technique which offers a more comprehensive study of quasi-periodic conductance fluctuations. Using this technique, the character of the fluctuations can be followed through magnetic field to allow a more detailed interpretation based on semiclassical theory. From experiment we observe that, by varying the input–output mode matching conditions of a ballistic quantum dot, the selection of classical periodic trajectories is modified without changing the relatively open nature of the dot itself. This has bearing on the application of these dots in more complex arrangements, or circuits, where the interaction between them is crucial. This feature also further complicates the adaptation of an intuitive semiclassical theory for these engineered systems. Future experimental studies and improvements to the semiclassical theory to account for open systems will it is hoped enhance this approach.

Acknowledgments

The authors wish to thank their collaborators at the Japanese Institute for Physical and Chemical Research (RIKEN) for use of the low-temperature measurement facilities. Electron beam lithography at ASU is performed on a RAITH Inc ELPHY system. This work is supported in part by the Office of Naval Research and the Defense Advanced Research Projects Agency.

References

[1] Chang A M, Baranger H U, Pfeiffer L N and West K W 1994 *Phys. Rev. Lett.* **73** 2111
[2] Marcus C M, Rimberg A J, Westervelt R M, Hopkins P F and Gossard A C 1992 *Phys. Rev. Lett.* **69** 506
Marcus C M, Westervelt R M, Hopkins P F and Gossard A C 1993 *Phys. Rev.* B **48** 2460
[3] Bird J P, Ishibashi K, Ferry D K, Ochiai Y, Aoyagi Y and Sugano T 1995 *Phys. Rev.* B **52** 8295
[4] Akis R, Ferry D K and Bird J P 1996 *Phys. Rev.* B **54** 17 705
[5] Brack M and Bhaduri R K 1997 *Semiclassical Physics* (New York: Addison-Wesley)
[6] Richter K, Ullmo D and Jalabert R A 1996 *Phys. Rep.* **276** 2
[7] Blaschke J and Brack M 1997 *Phys. Rev.* A **56** 182
[8] Bird J P, Olatona D M, Newbury R, Taylor R P, Ishibashi K, Stopa M, Aoyagi Y, Sugano T and Ochiai Y 1995 *Phys. Rev.* B **52** 14 336
[9] Bird J P, Akis R, Ferry D K, Aoyagi Y and Sugano T 1997 *J. Phys.: Condens. Matter* **9** 5935
[10] Zozoulenko I V, Schuster R, Berggren K-F and Ensslin K 1997 *Phys. Rev.* B **55** 10 209
[11] Akis R, Bird J P and Ferry D K 1996 *J. Phys.: Condens. Matter* **8** L667
[12] Marcus C M *et al* 1997 *Chaos, Solitons, Fractals* **8** 1261
[13] Bird J P, Akis R, Ferry D K, Cooper J, Ishibashi K, Ochiai Y, Aoyagi Y and Sugano T 1998 *Semicond. Sci. Technol.* **13** to be published
[14] Ferry D K, Akis R and Bird J P 1998 *Superlattices Microstruct.* **23** 611
[15] Schwieters C D, Alford J A and Delos J B 1996 *Phys. Rev.* B **54** 10 652
[16] Wirtz L, Tang J-Z and Burgdörfer J 1997 *Phys. Rev.* B **56** 7589

Semicond. Sci. Technol. **13** (1998) A15–A17. Printed in the UK PII: S0268-1242(98)91871-5

Low-temperature magnetotransport in ballistic quantum dots and wires

Y Ochiai†, L-H Lin†, K Yamamoto†, K Ishibashi‡, Y Aoyagi‡, T Sugano‡, J P Bird§, D Vasileska§, R Akis§ and D K Ferry§

† Department of Materials Science, Chiba University, 1-33 Yayoi-cho, Inage-ku, Chiba 263, Japan
‡ Nanoelectronic Materials Laboratory, RIKEN, 2-1 Hirosawa, Wako, Saitama 351-01, Japan
§ Nanostructures Research Group, Arizona State University, Tempe, AZ 85287-6206, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We have studied the low-temperature magnetotransport in quantum dots and wires. Using the correlation function and electron focusing methods, we observed evidence for the presence of a stabilized orbit in the dot and wire. We have also studied a geometry-induced fractal behaviour of the magnetotransport in the quantum dot and wire, which confirms the regularity of the dominant orbits.

In submicron-sized quantum devices, electronic transport is considered to occur via only a few discrete quantum states so that the transport trajectory relates to a characteristic wavefunction 'scarring' [1]. The scarring constitutes a direct manifestation of classical mechanics in the quantum transport properties of the devices and can be probed experimentally through the application of a magnetic field [2]. This scarring may even be observed in an open structure such as corrugated quantum wires and is thought to be connected to the observation of backscattering peaks in their low-temperature magnetoresistance (MR) [3]. On the other hand, geometry-induced fractal behaviour was recently predicted for conductance fluctuations in mesoscopic systems and was attributed to the trajectory trapping effect in mixed (chaotic–regular) phase space [4–6].

In this report, we present results of magnetotransport studies performed in a ballistic quantum dot and wire based on the fractal behaviour in both of the systems by analysis of the fluctuations themselves.

All devices have been realized through the standard split-gate technique on the surface of a high-mobility GaAs/AlGaAs wafer. The wafer is patterned into a standard Hall bar geometry with a width of 80 μm with a voltage probe separation of 200 μm. The electron carrier density and mobility are respectively 4.1×10^{15} m^{-2} and 20 m^2 V^{-1} s^{-1} for the dot and are 2.7×10^{15} m^{-2} and 48 m^2 V^{-1} s^{-1} for wire. The quantum dot is patterned as a rectangular or a stadium-like geometry, consisting of one centre split gate and two quantum point contact (QPC) leads. The lithographic dimensions of the dot are roughly 0.8×0.8 μm^2 (see inset of figure 2), while the measured mean free path in the bulk wafer is 2.2 μm. Thus, the dot system is expected to be ballistic in nature. The designed length and width of the corrugated quantum wire were 3.0 and 1.0 μm, respectively. The corrugation tips with 0.2 μm height and 0.1 μm length were located at both of the wire walls with a period of 0.7 μm (see inset of figure 3). Then it can be considered as a dot array system.

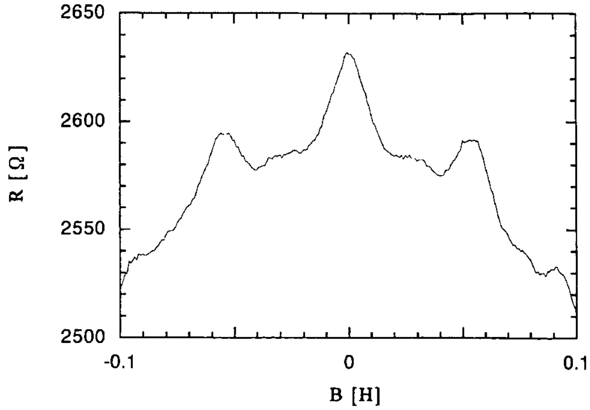


Figure 1. Low-temperature MR of the dot array system of the corrugated gated wire at low fields.

At temperatures of the order of 1 K and below reproducible and highly regular oscillations are observed in the MR of the dot as shown in figure 1. As we discuss below, in the region near zero magnetic field for both types of dots, we observe a sharp peak of the MR thought to be due to ballistic weak localization (WL) [7–9]. Reproducible fluctuations due to ballistic electron interference are also observed. The correlation function analyses are performed for the low-field MR. While the initial decay can be used to define a correlation field B_c, the dominant feature is the oscillatory behaviour. This periodic nature becomes stronger as V_p, the voltage of the gate bias at the QPC,

0268-1242/98/SA0015+03$19.50 © 1998 IOP Publishing Ltd

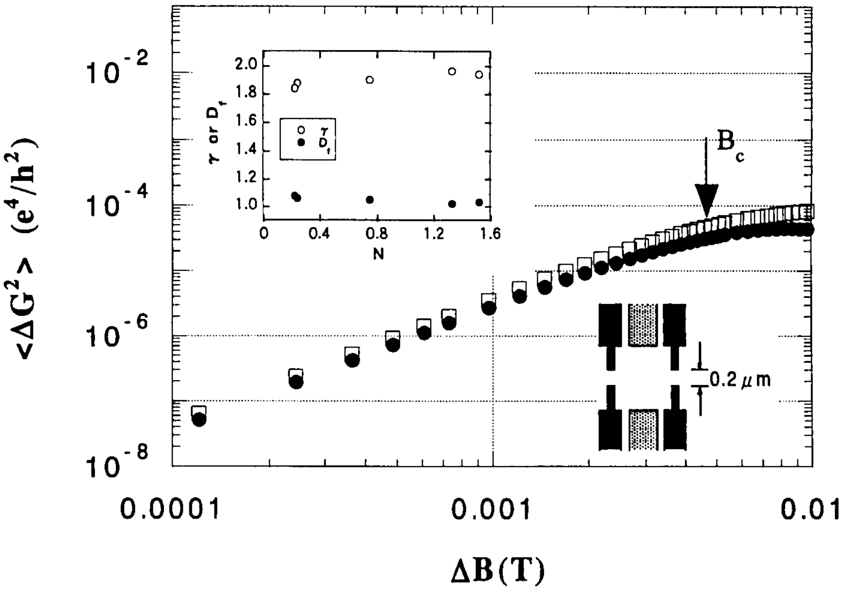


Figure 2. Variance, $\langle(\Delta G)^2\rangle$, for a single dot is plotted as a function of the field interval, ΔB. Inset shows the exponent γ and fractal dimension D_F as a function of the channel conductance at the centre gate voltage $V_c = -0.60$ V.

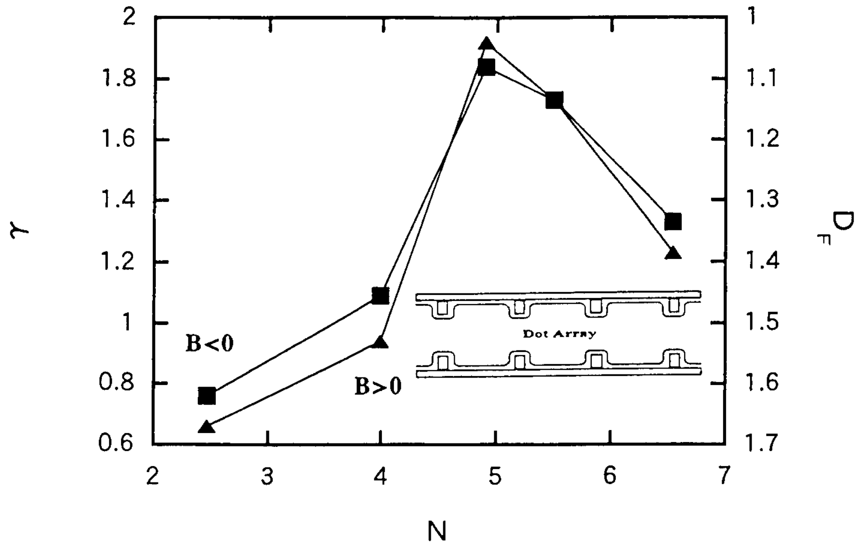


Figure 3. The exponent γ and fractal dimension D_F as a function of the channel conductance, N, for dot array.

increases, reducing the number of modes in the QPC [10, 11]. A preliminary analysis of the periodic oscillations at $V_p = -0.50$ V has been given previously [10], where it was shown that a 70 T^{-1} component was prominently observed in the FFT analysis for both the experiment and a theoretical simulation. If one applies a flux quantization criterion to this observed frequency, the average area of the orbit corresponds to 0.3 μm^2, one-half of the dot area, a result in keeping with earlier observations [2]. For only a few electron channels in the dot, highly regular oscillations are observed in the FFT analysis of the correlation function of the low-field magnetoconductance and are thought to be due to a regularly recurring wavefunction scarring [12]. However, the fundamental frequency of the oscillations is low and the estimated area is smaller than the dot area by one order of magnitude. This reduction can be explained by the flux cancellation effect due to a stabilized electron orbit with an hourglass-type trajectory [12].

Next, we focus on the fractal behaviour of the transport. For a Gaussian random process, electron transport described by a fractal Brownian motion has been shown to exhibit a variance of the magnetoconductance fluctuation $\langle(\Delta G)^2\rangle$ proportional to $(\Delta B)^\gamma$ [4, 5]. The exponent γ is in turn related to the fractal dimension through the relation $D_F = 2 - \gamma/2$ [4]. The experimental field resolution is ~0.1 mT. This provides a significant range of fields over which the fractal dimension can be estimated. The upper limit of the useful field is B_c, which is located near 1 mT for the dot. Below 1 mT, it may be seen that the slope of the resulting curve of $\langle(\Delta G)^2\rangle$ is almost 2 as shown in figure 2, although the exponent γ seems to deviate slightly from the quadratic relation as the gate voltage is varied. γ and D_F are shown in the inset of the figure as a function of the gate conductance, N, scaled by $2e^2/h$ for only a few electron channels in the dot. This result leads to a fractal dimension of $\sim$1, which is suggestive of linear transport along a classical *regular* orbit.

As for the dot array system (corrugated quantum wire), the results for both γ and D_F are shown in figure 3. Although we obtained $D_F \sim 1$ for the dot, scattered values of D_F from 1.0 to 1.6 are observed in the wire. In the wire, since the WL lineshape of the corrugated wire can be almost fitted with a Lorentzian, the stabilized trajectory seems to be difficult to realize in all cases. Near $N = 5$ in figure 3, D_F is almost unity as in figure 2 and electron focusing peaks are usually clearly observed in low-field MR [3]. At $N > 5$, the clarity decreases. When we control the gate bias, the shape of the dot array changes slightly; it reflects a large change of the trajectory of the transport. If the stability of the electron orbit is sensitive to the change of the scattering nature of the potential wall, the change of D_F is expected. Therefore the fractal nature strongly depends on the details of device geometry. Clear wavefunction scarring is also suggested in the corrugated quantum wire [3].

In conclusion, we have studied the low-temperature MR of a gated ballistic quantum dot and wire based on the fractal nature of the transport. Variation of the voltage applied to these split gates allows us to tune the shape of the quantum dot or wire and thus to investigate the stability of the orbits which are thought to be responsible for the fractal behaviour. The fractal behaviour in the dot and wire leads to $D_F \sim 1$ for a stable trajectory, confirming the regular nature of the orbits which are most likely to be stable orbits even in the stadium-shaped dot (where the classical dynamics is normally thought to be chaotic).

Acknowledgments

We are very grateful to Professor Richard Newbury and Dr Richard P Taylor for extensive discussions on fractal transport behaviour. Portions of this work were supported (Chiba University) by a grant in aid for the Science Research Project on the priority area of 'Single Electron Devices and Their High Density Integration' from the Ministry of Education, Science and Culture of Japan and (Arizona State University) by the Office of Naval Research and the Defense Advanced Research Projects Agency.

References

[1] Akis R, Ferry D K and Bird J P 1996 *Phys. Rev.* B **54** 17 705
[2] Bird J P, Akis R and Ferry D K 1996 *Europhys. Lett.* **35** 529
[3] Ochiai Y, Widjaja A W, Sasaki N, Yamamoto K, Akis R, Ferry D K, Bird J P, Ishibashi K, Aoyagi Y and Sugano T 1997 *Phys. Rev.* B **56** 1073
[4] Ketzmerick R 1996 *Phys. Rev.* B **54** 10 841
[5] Hegger H, Hecker K, Reckziegel G, Feimuth A, Huckestein B, Janssen M and Tuzinski R 1996 *Phys. Rev. Lett.* **77** 3885
[6] Akis R and Ferry D K 1997 *Proc. EP2DS-12 (Tokyo)*
[7] Taylor R P, Newbury R, Dunford R B, Coleridge P T, Sachradja A S and Adams J A 1995 *Phys. Rev.* B **51** 9801
[8] Zozoulenko I V and Berggren K-F 1996 *Phys. Rev.* B **54** 5823
[9] Ochiai Y, Widjaja A W, Sasaki N, Yamamoto K, Bird J P, Ishibashi K, Aoyagi Y, Sugano T and Ferry D K 1997 *Japan. J. Appl. Phys.* **36** 3971
[10] Okubo Y, Sasaki N, Ochiai Y, Bird J P, Ishibashi K, Aoyagi Y, Sugano T, Vasileska D, Akis R and Ferry D K *Physica* B, submitted
[11] Okubo Y, Ochiai Y, Vasileska D, Akis R, Ferry D K, Bird J P, Ishibashi K, Aoyagi Y and Sugano T 1997 *Phys. Lett.* A **236** 120
[12] Ochiai Y *et al Proc. EP2DS-12 (Tokyo)*

Semicond. Sci. Technol. 13 (1998) A18–A20. Printed in the UK PII: S0268-1242(98)91872-7

Closed versus open: to what extent do leads influence the magnetotransport in square quantum dots?

R Akis and D K Ferry

Center for Solid State Electronics Research, Center for Systems Science and Engineering, and Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-5706, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We have performed numerical simulations of the quantum mechanical transport and corresponding wavefunctions in square quantum dots. In order to obtain a better understanding of the role of leads in transport, we examine the conductance as a function of both magnetic field and energy for two different lead configurations and compare these results with a 'closed' dot connected to the external environment by two tunnelling barriers.

Recently, studies of magnetotransport fluctuations in circular quantum dots connected to leads have used simple tunnelling models in order to explain the results [1, 2]. The conclusion drawn was that the periodic nature of the fluctuations could be understood almost exclusively in terms of the density of states of the *closed* dot, even with several modes passing through the leads. On the other hand, studies of magnetotransport fluctuations in *square* dots have taken an entirely different point of view, suggesting that the leads are of primary importance [3–5]. In this regard, it has been found that the diffraction effect due to the quantization of modes in the leads causes the electrons to enter the dot in collimated beams and it is believed that this collimation is responsible for a highly selective excitation of dot states. In hand with this, '*scarred*' wavefunctions, which have amplitude maximized along a particular periodic classical orbit (in particular, a diamond-shaped orbit), have also been seen in simulations of open quantum dots [3–5]. These scars appear at resonances in the conductance and can occur periodically, with periodicities apparently coinciding with peaks in the Fourier power spectrum of the conductance fluctuations [3, 4].

In this paper, to clarify further the role of the leads, we employ a stabilized variant of the transfer matrix approach [6] to study the magnetotransport and corresponding wavefunctions of square quantum dots. Specifically, we examine the conductance, G, as a function of *both* Fermi energy, E, and magnetic field, B, for several different configurations, including one in which the dot is 'closed' and transport is allowed via two tunnelling barriers. As will become apparent, studying the three-dimensional function $G(E, B)$ provides a very graphical way to see first hand the periodicities that occur in the conductance. Among other results, we find that, when the input and output leads are aligned at the top of the dot, a very regular cross-hatched pattern results, particularly as the dot is progressively made more open. The grooves evident in these patterns in many cases coincide with points where the scars occur.

Because of space limitations, we refer the reader elsewhere for the mathematical details [4, 6] and begin simply with our results. In figure 1, we show several plots of $G(E, B)$ versus E and B, with darker shading corresponding to higher conductance. In each case, the quantum dot under study is 0.3 μm square, which is approximately the same size as a dot with top aligned leads that was studied experimentally and found to yield periodic conductance fluctuations [3]. The typical electron density found in the experimental dots was $\sim 4 \times 10^{11}$ cm^{-2}, which corresponds to the lower part of the Fermi energy range we examine here, 14–15.5 meV. In the simulations, the dots are connected to waveguides that extend outward to $\pm\infty$. Being 0.3 μm wide, these typically support $\sim$16 modes in our calculations over the given energy range. It should also be noted that in this energy range the ratio of the cyclotron radius to dot size is of the order of $\sim$1.4 for the highest fields considered here ($B = 0.25$ T).

In figure 1(a), we show G for the case where the dot is enclosed by two tunnelling barriers 0.05 μm thick, with the potential barrier height being 15.5 meV, which was sufficient to ensure that most of the transmission was of a resonant nature over the given energy range. Importantly, since transmission here occurs along the entire length of the barriers, there is no specially favoured point or angle of entry into the dot in this case. As one would expect in the

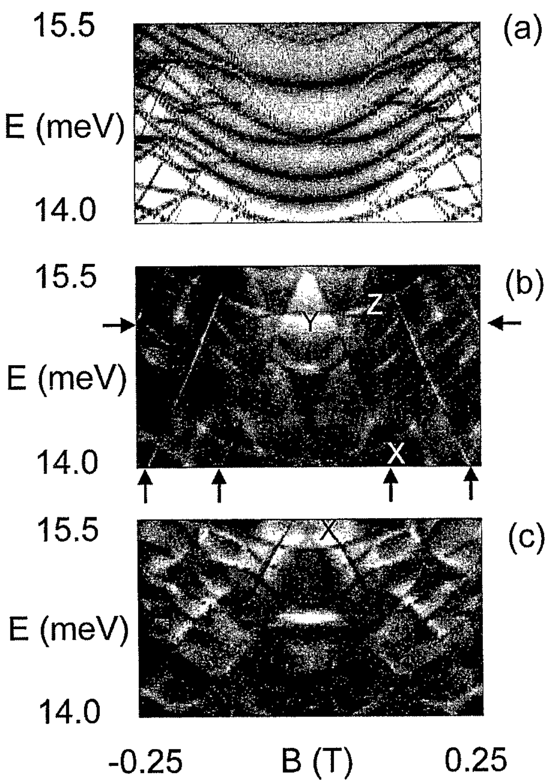


Figure 1. Grey-scale images of *G(E, B)* versus *E* and *B* are plotted for (a) a 0.3 μm square dot formed by two tunnelling barriers, (b) a dot with top aligned leads and (c) a dot with centrally aligned leads. Here, darker shading indicates higher conductance.

tunnelling regime, the 'spiderweb' pattern of intersecting lines in this picture closely resembles the energy level structure for a closed dot (see, for example, the calculations of Richter *et al* [7] and the experiments of McEuen *et al* [8], Tarucha *et al* [9] and Zhitenev *et al* [10]). It should be noted that the shading in this instance was done on a logarithmic scale, so that the fainter lines are actually several orders of magnitude lower in conductance than the darkest spots (the highest value of G here is $\sim 2e^2/h$).

In contrast, figure 1(b), shows results for the case where the dot is connected to the outer waveguides by leads that are at the top edge of the dot. These leads are formed by essentially infinitely high tunnelling barriers that have had the top section removed (the dot geometry can be seen in figure 2). Here, the width of these leads has been adjusted to permit between four and five modes to propagate over the given energy range. The conductance here ranges from $\sim 2e^2/h$ to $\sim 8e^2/h$ and the shading here is on a linear scale. Despite the fact that the dot is 'open', there is still resonant behaviour, as indicated by the strong lines or grooves that are superimposed on the conductance. As with the tunnelling picture, there are upward-bending parabolic lines of resonances, but superimposed on these are a series of almost parallel lines, tilted to the left and the right, forming a very regular cross-hatched pattern. Given this result, it is easy to understand why dots of this configuration would yield periodic conductance fluctuations. Figure 1(c) shows what occurs when the input and exit leads are shifted to the middle of the dot, with the leads now supporting three modes. In this case, the pattern does not appear as regular. Moreover, it is apparent that many of the resonant grooves are in different locations.

In order to obtain physical insight into the conductance features shown in figure 1, one must look at the dot wavefunctions. In figure 2(a) we plot $|\psi(x, y)|$ versus x and y for a dot with the leads on the top edge. Here, we have set $E = 14$ meV and $B = 0.128$ T, which corresponds to the 'X' drawn on figure 1(b). Evident here is the diamond scar seen in previous work [3, 4]. Also apparent is the diffraction effect of the exit lead, which causes the electrons to come out in downward-pointing beams (there is some bending done by the magnetic field, but that only partially accounts for the effect seen here). Note that the 'X' falls on one of the sharp, tilted grooves. Importantly, this scar occurs along the entire length of the groove, at least where it remains sharp (as one might expect, the scar fades as the groove dissipates). Moreover, similar diamond scars are observed along several sharp grooves that run parallel to it. The end points of these grooves are indicated by the arrows in figure 1(b). In figure 2(b), we plot $|\psi(x, y)|$ for $E = 15.1$ meV and $B = 0.0$ T, which corresponds to the 'Y' on figure 1(b), which lies on one of the parabolic grooves. Here the wavefunction shows a sine-wave pattern which uniformly fills the dot. A virtually identical pattern occurs in the same location in the tunnelling dot. The wavefunction corresponding to 'Z' ($E = 15.25$ meV and $B = 0.112$ T) is plotted in figure 2(c). Here we see a 'bouncing ball' pattern that bears some similarity to that in figure 2(b), which probably is not surprising since it falls on the same parabolic line and probably represents essentially the same state evolved by the magnetic field. Finally, in figure 2(d), we plot $|\psi(x, y)|$ for the centred lead dot, setting $E = 15.5$ meV and $B = 0.022$ T. This corresponds to the 'X' in figure 1(c), which runs along a groove. Interestingly, this picture looks like what might be described as an 'antiscar', with the amplitude *minimized* along the diamond pattern. Indeed, if we just look at a single point contact in this case, the wavefunction amplitude would exit that contact in a V-shaped jet owing to the diffraction effect of the lead [11] (such a jet exiting the dot is very faintly visible in the picture). That said, it is likely that the orbits that are responsible for this pattern are a pair of tilted rectangular orbits which bounce between the diagonal corners of the dot. These have been observed elsewhere [4, 12]. One may ask why the apparently scarred wavefunctions lie along the linear grooves while the parabolic grooves seem to yield wavefunctions that show a more uniform excitation. In closed *circular* dots, the energy level structure follows the Darwin–Fock spectrum [13], which depends on radial and angular momentum quantum numbers, n and l, respectively. Importantly, $E_{n,l}$ has terms both parabolic and linear with magnetic field, with the linear term being proportional to l. Thus, states with higher angular momentum will show more linear behaviour. Perhaps something similar is happening here in the square dot. It should be noted that a current flow diagram of the diamond scar clearly shows circulation along the scar in one direction [14].

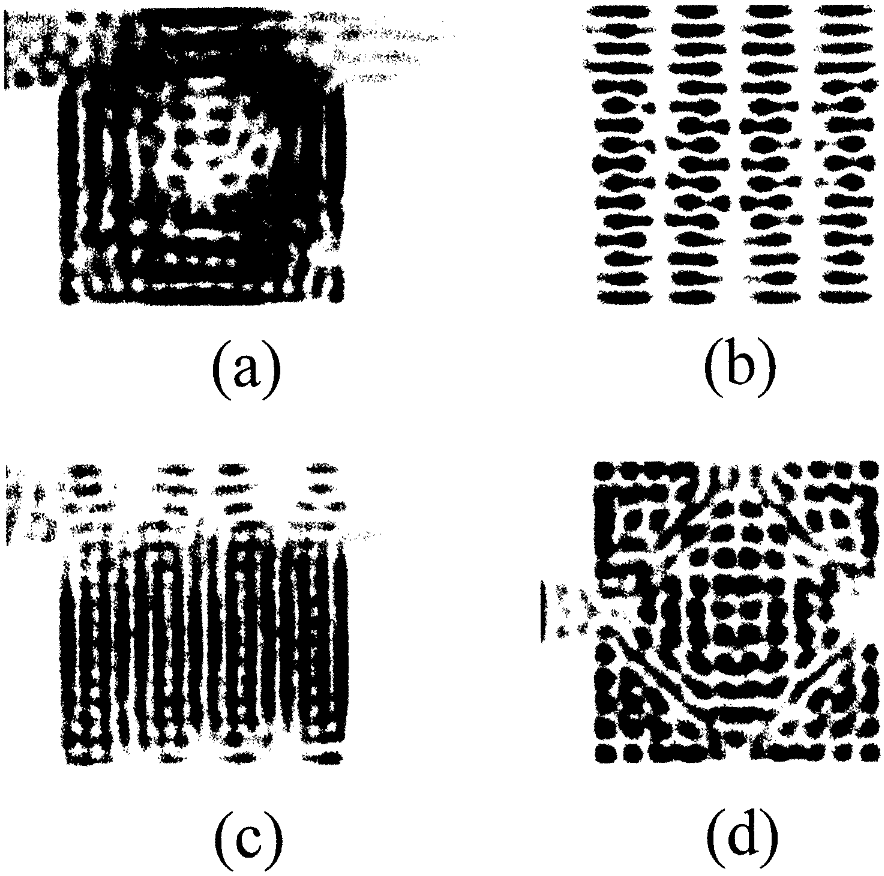


Figure 2. $|\psi(x, y)|$ versus x and y is plotted for (a) a top lead dot at the values of E and B indicated by the 'X' in figure 1(b), (b) a top aligned dot at the point marked 'Y' in figure 1(b), (c) a top lead dot at the point marked 'Z' in figure 1(b) and (d) a central lead dot at the point marked 'Z' in figure 1(b). Darker shading corresponds to higher amplitude.

What can we conclude from these results? Even with several modes propagating in the leads, the conductance for the open dots obviously shares many resonant features in common with the tunnelling dot, indicating that the closed dot density of states still plays an important role in determining the behaviour in the open system. On the other hand, from the shift in the positions of the resonant grooves with the different lead configurations, it is clear that the leads do play an important role as a spatial filter [15] which selects which dot states are resonantly excited in the open structures. Moreover, we find that certain resonance lines that appear in the conductance can be associated with scarring from a particular periodic classical orbit.

Acknowledgments

We acknowledge the financial support of ONR and DARPA and useful discussions with J P Bird.

References

[1] Persson M, von Sydow B, Lindelof P E, Kristensen A and Berggren K F 1995 *Phys. Rev.* B **52** 8921

[2] Berggren K F, Ji Z-L and Lundberg T 1996 *Phys. Rev.* B **54** 11 672

[3] Bird J P, Ferry D K, Akis R, Ochiai Y, Ishibashi K, Aoyagi Y and Sugano T 1996 *Europhys. Lett.* **35** 529

[4] Akis R, Ferry D K and Bird J P 1996 *Phys. Rev.* B **54** 17 705

[5] Zozoulenko I V and Berggren K F 1996 *Phys. Rev.* B **54** 5823

[6] Usuki T, Saito M, Takatsu M, Kiehl R A and Yokoyama N 1995 *Phys. Rev.* B **52** 8244

[7] Richter K, Ullmo D and Jalabert A 1996 *Phys. Rep.* **276** 1

[8] McEuen P L, Foxman E B, Meirav U, Kastner M A, Meir Y, Windgreen N S and Wind S J 1991 *Phys. Rev. Lett.* **66** 1926

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

[10] Zhitenev N B, Ashoori R C, Pfeiffer L N and West K W 1997 *Phys. Rev. Lett.* **79** 2308

[11] Akis R, Ferry D K and Bird J P 1997 *Phys. Rev. Lett.* **79** 123

[12] Akis R, Ferry D K and Bird J P 1997 *Japan. J. Appl. Phys.* **36** 3981

[13] Darwin C G 1930 *Proc. Cambridge Philos. Soc.* **27** 86
Fock V 1928 *Z. Phys.* **47** 446

[14] Bird J P, Akis R, Ferry D K, Aoyagi Y and Sugano T 1997 *J. Phys: Condens. Matter* **9** 5935

[15] Ferry D K, Akis R and Bird J P 1998 *Superlattices Microstruct.* **23** 611

Semicond. Sci. Technol. **13** (1998) A21–A23. Printed in the UK

PII: S0268-1242(98)91870-3

Size-dependent effects on the magnetotransport fluctuations of square quantum dots

N Holmberg, R Akis, D P Pivin Jr, J P Bird and D K Ferry

Center for Solid State Electronics Research, and Center for Systems Science and Engineering, Arizona State University, Tempe, AZ 85287-5706, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. Recently, experimental studies of magnetotransport in nominally square quantum dots have revealed a high degree of periodicity in the conductance fluctuations, with only a few apparently harmonically related frequencies dominating the power spectra. Analysis of dots of different sizes has revealed that there are unresolved issues regarding the scaling of the dominant frequency of the fluctuations. The experimental data appear to suggest that the dominant frequency scales with $A^{1/2}$, where A is the area of the dot. On the other hand, a semiclassical analysis of periodic orbits suggests that the scaling should be with area. In this paper, we attempt to resolve this issue by simulating dots of many different sizes using both quantum mechanical and classical approaches.

Electron interference is an important process in mesoscopic devices and strongly influences their resulting electrical behaviour. For example, disordered systems are characterized by diffusive transport, in which quantum interference gives rise to weak localization and universal conductance fluctuations. Recently, similar fluctuations have been observed in ballistic quantum dots [1]. In these ballistic quantum dots, the carriers scatter only from the confining walls of the cavity, and the quantum interference arises from interactions among the various trajectories. Transport can be expected to involve just a few eigenstates which are excited by the collimation effect of the entrance quantum point contact [2]. This behaviour leads to reproducible, periodic fluctuations in the magnetoconductance, which are described by a few, harmonically related frequencies in the Fourier spectra of the fluctuations [3–7].

In general, the eigenvalue spectrum of the dot arises from a series of δ functions located (in energy space) at each of the resonant eigenenergies of the dot cavity. Carrying this further, it is then possible to say that the density of states for the dot is

$$\rho(E) = \sum_n \delta(E - E_n). \qquad (1)$$

If the δ function is replaced by its Fourier representation (in energy space) via the Poisson summation formula, then the integrals can be evaluated by the saddle-point method, and the density of states can be expanded into the form [8]

$$\rho(E) = \frac{m^* a^2}{2\pi\hbar^2} \sum_{M_1,M_2=-\infty}^{\infty} J_0\left(\frac{1}{\hbar} S_{M_1,M_2}\right) - \frac{a}{4\pi\hbar}\left(\frac{2m^*}{E}\right)^{1/2} \sum_{M=-\infty}^{\infty} \cos\left[\frac{2Ma}{\hbar}(2m^*E)^{1/2}\right] \qquad (2)$$

for a two-dimensional square billiard of side a. Here, the action is given through

$$\frac{1}{\hbar} S_{M_1,M_2} = \frac{(2m^*E)^{1/2}}{\hbar} L_{M_1,M_2} = \frac{2(2m^*E)^{1/2}}{\hbar} a (M_1^2 + M_2^2)^{1/2}$$

$$I_i = \frac{1}{2\pi} \oint p_I \, \mathrm{d}x_I = n\hbar \qquad (3)$$

where M_1 and M_2 are the multiplicities of the primary periodic orbits defined through the quantization conditions I_i, $i = 1, 2$, L_{M_1,M_2} is explicitly defined within the equation, and p_i is the phase space momentum. The second term of equation (2) is a boundary correction.

The oscillatory density of states in a two-dimensional rectangular billiard in the presence of a weak magnetic can be written as [9]

$$\delta\rho(E, B) = \frac{m^* a^2}{\pi^{3/2}\hbar^2} \sum_{\Gamma} \sum_{n=1}^{\infty} \left(\frac{\hbar}{n S_\Gamma}\right)^{1/2} \cos\left(\frac{n\Phi}{2}\right) \times \cos\left(\frac{n}{\hbar} S_\Gamma - \frac{\pi}{4}\right)$$

0268-1242/98/SA0021+03$19.50

$$\Phi = \frac{a^2eB}{M_1M_2\hbar} \qquad (4)$$

where Γ sums over primitive periodic orbits and n sums over the repetition of these orbits, M_1M_2 must be an odd integer and the cyclotron radius must satisfy the condition $r_c \gg a$. It is clear from this equation that the enclosed Aharonov–Bohm-like area is one-half of the dot area (or less for larger M). However, the experiment suggests that the period scales not with $a^2 = A$ but with a itself [10]. In fact, this derivation is incorrect. Even beyond the fact that the cyclotron radius cannot be ignored because of its size in the experiments, equation (4) does not represent the proper physics. We return to this point below, after examing the simulations. By expanding into a lattice of vectors representing the trajectories in equation (2), the sum in equation (2) (or the sum over the trajectories in equation (4)) can be more easily seen [11], as each trajectory and action S_Γ leads to one lattice point, identified by the integers M_1 and M_2.

We have carried out simulations of the quantum transport through a square dot of sizes from 0.2 to 2.0 μm using a stable variant of the transfer matrix approach [2, 12]. In addition, we have calculated the classical transport through the device by a numerical integration of the structure, as previous work has shown that this gives results comparable with the quantum simulation [2]. Finally, for a few sizes, we have used a recursive Green's function approach as well. The 'area' can be found from the periodicity of the magnetoconductance. By Fourier transforming the fluctuation correlation function, we find the power density in the signal and then compare the dominant peaks as the dot size is varied. It should be noted that the Fourier transform of the correlation function is calculated rather than the transform of the raw data because the data are not stationary. The correlation function itself is

$$F(B_0, \Delta B) = \langle \delta G(B_0)\, \delta G(B_0 + \Delta B)\rangle \qquad (5)$$

In figure 1(*a*), we plot the power spectral density in the Fourier transform of the correlation function for a 0.2 μm dot, at $B_0 = 0.18$ T. In practice, we take the Fourier transforms for a large number of initial magnetic field values and then form a multidimensional plot with the axes B_0 and magnetic frequency (the transform of the ΔB variable) and with shading of the amplitude of the transform. Such a plot is shown in figure 1(*b*) for the 0.2 μm dot where a fixed window of 0.2 T is moved over the data (that is, the ensemble average in equation (5) is carried out over a 0.2 T range, and the plot shows the centre of this range). From this, one can clearly pick out the dominant peak, and determine how it moves with initial magnetic field.

However, there often is not a single dominant peak, and one must ascertain how one or a few peaks grouped together evolve as the dot size is changed. Moreover, there appears to be the onset of new families of orbit periods as the dot size is increased. Nevertheless, we have been able to determine largely the dominant groups of peaks (sometimes there are two groups), and the behaviour is shown in figure 2, where we plot the magnetic frequency as a function of the dot edge length. The classical simulation seems to lead clearly to a magnetic frequency that scales with the edge of the dot, i.e. with the square root of the area. However, there is a second dominant set of frequencies that comes in for dots $\geq$0.8 μm. The latter set is labelled as 'classical 2nd' and is at a higher magnetic frequency. The recursive Green's function (which is present for only a limited number of small dot sizes) and the stabilized mode matching computations appear to give a dependence that scales with the dot area. However, for the latter method, a second set of frequencies (labelled 'SMM 2nd') appears in the larger dots and seems to have values comparable with the classical ones, also scaling with the square root of the area. Finally, the experimental data are plotted as the full circles.

It seems to be clear that the experiments agree quite well with some of the calculated frequencies. Nevertheless, it seems to be the case that the frequency does scale with

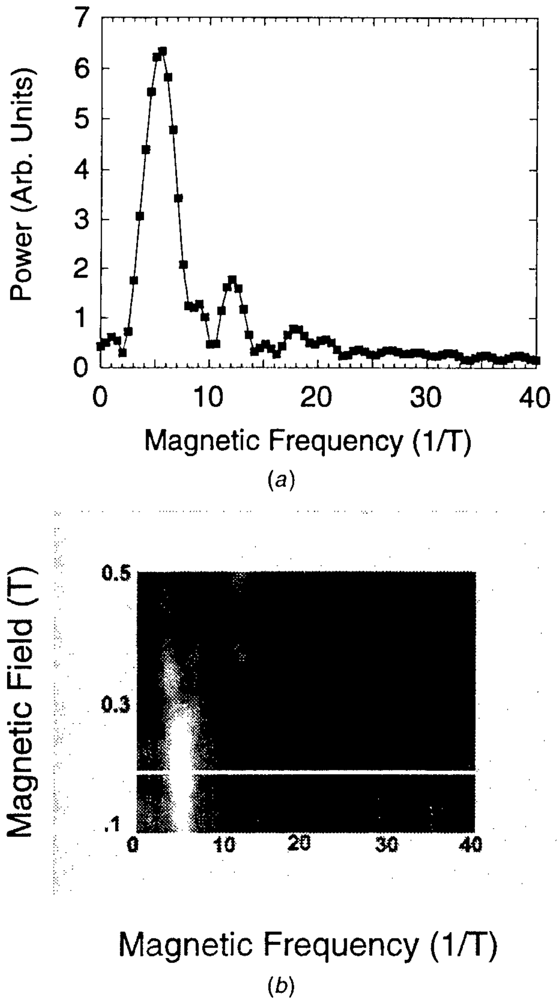


Figure 1. (*a*) The Fourier transform of the fluctuation correlation function for an initial magnetic field of 0.18 T in a 0.2 μm dot. (*b*) The map of multiple Fourier transforms, with the amplitude indicated by shading. The horizontal axis is ΔB and the vertical axis is B_0. A fixed window of 0.2 T is moved over the data to create this map. The horizontal cut, indicated by the white line, corresponds to the correlation function in (*a*).

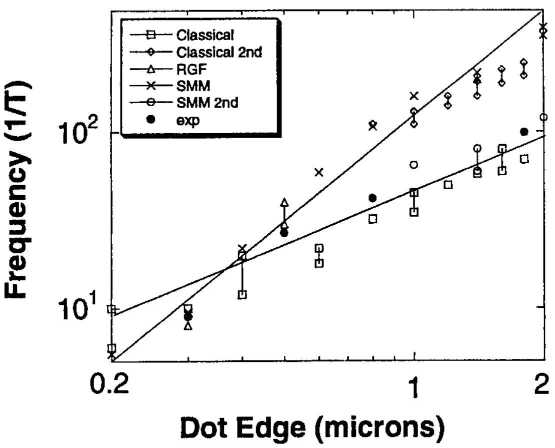


Figure 2. The magnetic frequency of the dominant peaks in the Fourier transform as a function of the dot edge length as determined from the map of multiple Fourier transforms. The fixed window and B_0 are varied inversely with the size of the dot. The various points are discussed in the text. The two full lines are guides to the eye and represent variation as *a* and $a^2 = A$.

the dot edge, that is with the square root of the area, for sufficiently large dots. For small dots, below 0.5 μm, the behaviour is not so clear, but an argument for variation of the frequency with the area could be made. This change in behaviour leads us to believe that there is a transition of modes when going from smaller to larger dot sizes as indicated in figure 2. It should be noted that the amplitudes of the individual modes are not equally excited because of to the collimation effect of the quantum point contact at the entrance of the dot. This leads to the error in equation (4). There are many trajectories, with the same set of indices M_1 and M_2, but with different enclosed areas. Normally, these are degenerate, according to equation (2). However, the magnetic field lifts this degeneracy, and any theory of the magnetic variation must include a sum over these different-area orbits. Equation (4) neglects this fact, assuming that all of these orbits would be equivalent. Each different area is expected to have a different magnetic periodicity, primarily because each of these different-area orbits (at $B = 0$) will have a different angular momentum for non-zero magnetic field. As the dot size becomes larger, the data of figure 2 seem to suggest that the dominant trajectory will shift to a smaller-area orbit (and probably the lowest-energy, high angular momentum state) with the smallest magnetic frequency. Although this is potentially a reason for the change in behaviour, it is clear that more work is required to provide more information for other possible changes in behaviour. Still, it is clear that the dots are dominated by regular orbits and are not described by an ergodic theory.

Acknowledgments

This work was supported in part by ONR and ARPA.

References

[1] Bird J P, Ishibashi K, Aoyagi Y, Sugano T and Ochiai Y 1994 *Phys. Rev.* B **50** 18 678

[2] Akis R, Ferry D K and Bird J P 1996 *Phys. Rev.* B **54** 17 705

[3] Bird J P, Ferry D K, Akis R, Ochiai Y, Ishibashi K, Aoyagi Y and Sugano T 1996 *Europhys. Lett.* **35** 529

[4] Bird J P, Akis R, Ferry D K, Ochiai Y, Ishibashi K, Aoyagi Y and Sugano T 1996 *Proc. Int. Conf. on Physics of Semiconductors (Berlin, 1995)* ed M Scheffler and R Zimmerman (Singapore: World Scientific) p 1553

[5] Ferry D K, Akis R, Pivin D P Jr, Connolly K M, Bird J P, Ishibashi K, Aoyagi Y, Sugano T and Ochia Y 1997 *Proc. Int. Conf. High Magnetic Fields in Semiconductors* vol 1 (Singapore: World Scientific) p 299
Bird J P, Akis R, Ferry D K, Aoyagi Y and Sugano T 1997 *J. Phys.: Condens. Matter* **9** 5935

[6] Bird J P, Ferry D K, Akis R, Newbury R, Taylor R P, Olatoona D M, Ochiai Y, Ishibashi K, Aoyagi Y and Sugano T 1996 *Superlattices Microstruct.* **20** 287

[7] Connolly K M, Pivin D P Jr, Ferry D K and Wieder H H 1996 *Superlattices Microstruct.* **20** 307

[8] Brack M and Bhaduri R K 1997 *Semiclassical Physics* (Reading, MA: Addison-Wesley) section 2.7

[9] Ullmo D, Richter K and Jalabert R A 1995 *Phys. Rev. Lett.* **74** 383

[10] Ferry D K, Bird J P, Akis R, Pivin D P Jr, Connolly K M, Ishibashi K, Aoyagi Y, Sugano T and Ochiai Y 1997 *Japan. J. Appl. Phys.* **36** 3944

[11] Ferry D K, Akis R and Bird J P 1998 *Superlattices Microstruct.* **23** 611

[12] Usuki T, Saito M, Takatsu M, Kiehl R A and Yokoyama N 1995 *Phys. Rev.* B **52** 8244

Semicond. Sci. Technol. **13** (1998) A24–A26. Printed in the UK
PII: S0268-1242(98)91853-3

Triangular ballistic quantum dots: classical, semiclassical and wave mechanical electron dynamics

H Linke†, K-F Berggren‡, L Christensson†, P E Lindelof§, A Löfgren†, P Omling†, M Yousefi† and I V Zozoulenko‡

† Solid State Physics and Nanometer Consortium, Lund University, Box 118, S-22100 Lund, Sweden
‡ Theoretical Physics Group, Department of Physics and Measurement Technology, Linköping University, S-581 83 Linköping, Sweden
§ Niels Bohr Institute, University of Copenhagen, Universitetsparken 5, DK-2100 Copenhagen 0, Denmark

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. The relation between classical electron orbits and quantum mechanical wavefunctions is discussed in two ways, using the example of triangular ballistic quantum dots. Firstly, the frequency of conductance fluctuations observed in the magnetoresistance is linked, in the spirit of the semiclassical periodic orbit theory, to a specific classical electron trajectory that is also important for ballistic commensurability effects. Secondly, we present initial results of a method to construct semiclassical particle density distributions inside the dot and compare the result with those of the classical and the quantum mechanical calculations of particle density distributions.

Ballistic quantum dots, sometimes also called electron billiards, are two-dimensional electron cavities smaller than the mean free path of electrons. At temperatures of a few kelvin, high enough that phase coherence is suppressed by electron–electron interaction, electron transport through such devices can be described in a classical billiard ball model. This is illustrated in figure 1 where the magnetoresistance of a triangular, equilateral quantum dot is shown. The major maxima of the resistance measured at 5 K (thin line) can be related to specific classical electron trajectories (insets in figure 1) that are commensurate with the shape of the cavity at the field values $B = B_c = 50$ mT and $B = 3B_c$. It has been found that these specific trajectories are particularly important for classical electron dynamics because they are short and relatively stable [1]. At temperatures below 1 K, phase coherence is maintained throughout the cavity, and electron interference gives rise to conductance fluctuations (CFs) (bold line in figure 1). One way of understanding the origin of the CFs is to view electron transport through the dot as a tunnelling process from the source to the drain reservoir via states in the dot [2–4]. When an external magnetic field is tuned, a diamagnetic shift of the shell structure of the density of states (DOS) gives rise to the fluctuations of the conductance.

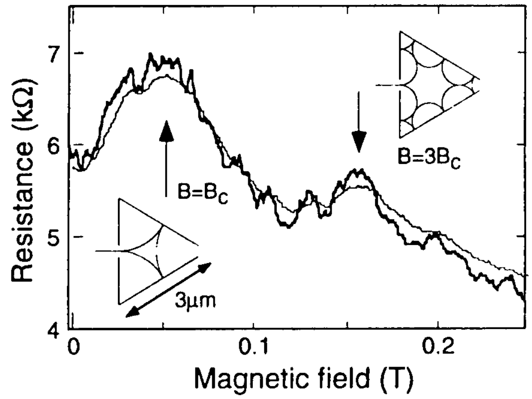


Figure 1. Magnetoresistance of a triangular electron billiard (side length 3 μm) at 5 K (thin line) and 0.3 K (bold line). The characteristic field B_c = 50 mT is defined by the field where the classical cyclotron orbit equals half the side length.

The coexistence of classical and interference effects makes ballistic quantum dots an ideal system for the study of how the billiard ball model can be linked to a wave mechanical description of electron transport. For the case of CFs this can be achieved by using the semiclassical periodic orbit theory (POT) which relates the DOS of a closed ballistic dot to classical periodic electron orbits [5, 6]. We will show in the first part of this report that the frequency spectrum of the CF observed in our experiment is in line with the prediction of the POT for the DOS in closed, triangular dots [5, 7]. In the second part, we will

0268-1242/98/SA0024+03$19.50

concentrate on the question concerning the extent to which the electron wavefunction inside the dot can be related to classical electron orbits. In particular, we discuss a simple, semiclassical approach that aims at constructing the quantum mechanical density distribution inside the dot from the classical trajectories.

We start by considering the results of the POT concerning the DOS in a triangular dot. The basic idea of the POT can be understood from the Bohr–Sommerfeld quantization condition, which for classical periodic orbits of electrons in a magnetic field can be written as

$$\hbar k_F L + 2\pi e B A = nh \qquad n = 0, 1, 2, \ldots \qquad (1)$$

Here, k_F is the electron wavevector, L and A are the length and directed area of the periodic orbit, respectively, B is the magnetic field and h is Planck's constant. Equation (1) expresses the idea that a particular classical periodic orbit makes a periodic contribution to the DOS as a function of k_F or B. For the case of triangular dots it has been shown that, at $B = 0$, only the three shortest (families of) orbits, shown in figure 2, need to be considered in order for us to be able to construct almost exactly the oscillating part of the density of states as a function of k_F [5]. Since the members of each family have the same length, the frequency spectrum of the DOS as a function of k_F consists of only three dominant frequencies. In our case, however, the magnetic field is the experimental parameter and we are interested in the periodicity of the DOS at fixed energy and variable magnetic field, such that the area of the orbits, instead of the length, needs to be considered. Arguing qualitatively, we expect that only the isolated orbit (1) in figure 2 makes a contribution to the frequency spectrum, because the area of family (10) is zero, while the areas of each member of family (11) are different, such that their contribution will be smeared out. As will be shown in the following, our experimental results agree with this expectation.

The triangular billiard (side length $a \approx 3$ μm), from which the data shown here were obtained, was defined by shallow wet etching in a GaAs/AlGaAs two-dimensional electron gas material. The Fermi energy and mean free path in the unstructured areas of the Hall bars were about 10 meV and 11 μm, respectively, at 0.3 K. The two contact openings at the centre of one side and in the tip had an estimated electric width of about 100 nm, such that about three subbands were occupied.

Figure 2 shows a Fourier transform of the CFs (figure 1, bold line) in the range between 0 and 75 mT. Before the transformation, the slowly varying background was removed by subtracting the temperature-averaged data (thin line). The maximum in the frequency spectrum, indicated by the arrow in figure 2, corresponds to the area $0.04a^2$ enclosed by orbit (1) in figure 2 if the distortion by the magnetic field is taken into account (broken curve orbit), in agreement with our expectation.

Summing up our discussion so far, we have found that the same distorted orbit (1) can be related to a ballistic effect (the global maximum at $B \approx B_c$ in figure 1) and to the CFs induced by wave interference in the same magnetic field range. In view of these results it is interesting to ask whether also the wavefunction at $B = B_c$ has any similarity to this particular orbit. In figures 3(a) and (b) we show quantum mechanically calculated particle density distributions for $B = B_c$ and $B = 3B_c$, respectively, for a triangle of side 1 μm and Fermi energy 10 meV. For comparison, we show in figures 3(c) and (d) the corresponding classical particle distributions obtained from many classical billiard ball trajectories (for details of the simulation techniques, see [1]). While the distorted orbit (1) is clearly visible in the relevant classical plot (figure 3(c)), the quantum mechanical density (figure 3(a)), constructed as a linear combination of all incoming modes [8], does not show any resemblance to this orbit. At higher fields, however, where the wavefunction is more localized, the classical skipping orbit (figure 3(d)) is clearly visible also in the wavefunction (figure 3(b)).

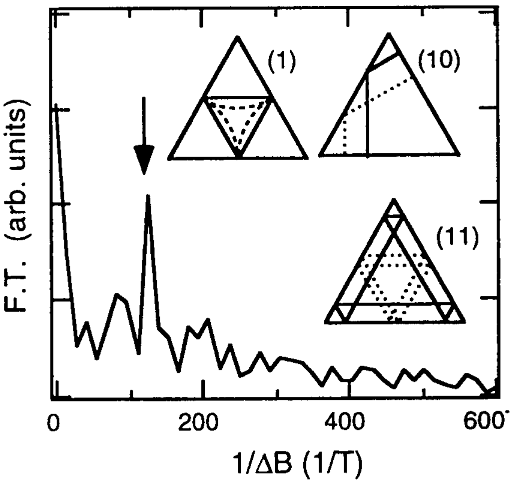


Figure 2. Fourier transform of the CFs in the range 0–75 mT (figure 1). The insets show the three shortest (families of) periodic orbits in equilateral triangular billiards (from [2]). The arrow indicates the fluctuation frequency that corresponds to the isolated orbit (1) in a magnetic field $B = B_c$ (broken curve orbit).

To learn more about how the classical orbits relate to the quantum mechanical density we are interested in the question of whether, and under which assumptions, the latter can be constructed from the former. In the following we therefore discuss initial results of a semiclassical approach in which a phase and an amplitude are added to the classical electron trajectories. We consider electrons that enter a triangular, hard-wall potential (side length 2 μm) via the base contact and assume that all electrons have initially the same phase. No impurity scattering or electron–electron interaction is taken into account. As opposed to the classical simulations, where an electron was assumed to leave the billiard as soon as it had accidentally passed one of the contact openings, we use here a contact reflectivity of 90% to increase the time during which the electrons build up an interference pattern in the dot. The electrons are given a finite range for interference by spatially smearing out the contribution of each trajectory with a Gaussian distribution of the width of one Fermi wavelength (λ_F = 50 nm). The resulting amplitudes (positive and negative) of all electrons at each point of

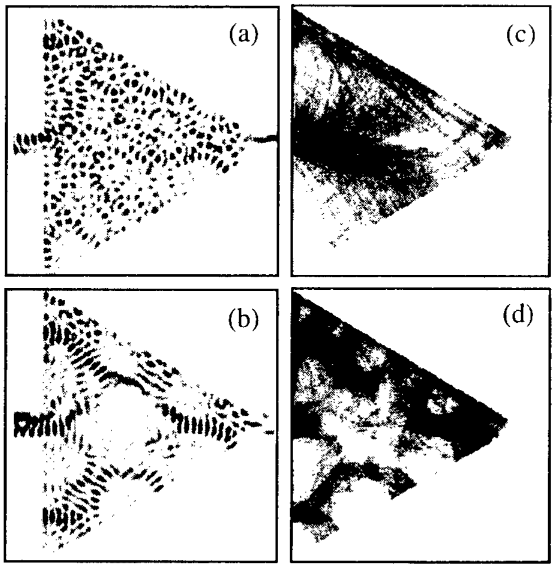


Figure 3. (a), (b) Quantum mechanical and (c), (d) classical electron density distributions inside an open, triangular dot at (a), (c) $B = B_c$ and (b), (d) $B = 3B_c$. For the classical case the same openings were used as for the quantum mechanical case.

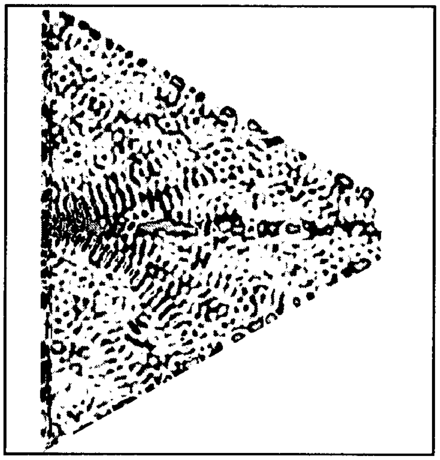


Figure 4. Semiclassical electron density distribution at $B = B_c$, obtained by adding an amplitude and a phase to classical trajectories as they are shown in figure 3(c).

the triangle were summed and then squared to give a value that corresponds to the probability of finding an electron at this point (for more careful semiclassical calculations of quantum dot properties based on classical trajectories, see also [9—11]).

In figure 4 the result of such a simulation is shown for $B = B_c$ and for $a/\lambda_F = 40$ (different from the quantum mechanical calculation in figure 3(a) where $a/\lambda_F \approx 20$ was used). While there is still some resemblance to the classical case (figure 3(c)), the semiclassical plot also shows a standing wave pattern as in the quantum mechanical case. In this sense the semiclassical result is truly intermediate between the classical and the wave mechanical results. The unrealistically high values of the particle density along the boundaries and along the symmetry axis indicate that our simple approach, of shifting the phase by 180° on each boundary collision, is only a first, rough approximation. In a more correct approach, boundary and magnetic effects should be incorporated as indicated in [9]. When fully developed, we believe that a semiclassical technique like the one presented here will be useful to study the influence of, for instance, phase breaking, escape time or the dot size on the relation between classical trajectories and wave mechanical states.

Acknowledgments

We should like to acknowledge financial support from the Swedish Research Councils for Natural and for Engineering Sciences and by the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine. We are indebted to Claus B Sørensen, the III–V NANOLAB, and CNAST for the MBE-grown structures.

References

[1] Linke H, Christensson L, Omling P and Lindelof P E 1997 *Phys. Rev.* B **56** 1440
[2] Berggren K-F, Ji Z-L and Lundberg T 1996 *Phys. Rev.* B **54** 11 612
[3] Zozoulenko I V, Schuster R, Berggren K-F and Ensslin K 1997 *Phys. Rev.* B **55** R10 209
[4] Zozoulenko I V and Berggren K-F 1997 *Phys. Rev.* B **56** 6931
[5] Brack M and Badhuri R K 1997 *Semiclassical Physics* (Reading, MA: Addison-Wesley)
[6] Reimann S M, Persson M, Lindelof P E and Brack M 1996 *Z. Phys.* B **101** 377
[7] Brack M, Blaschke J, Creagh S C, Magner A G, Meier P and Reimann S M 1996 *Z. Phys.* D **40** 276
[8] Christensson L, Linke H, Omling P, Lindelof P E, Berggren K-F and Zozoulenko I V 1998 *Phys. Rev.* B **57** 12 306
[9] Blaschke J and Brack M 1997 *Phys. Rev.* A **56** 182
[10] Schwieters C D, Alford J A and Delos J B 1996 *Phys. Rev.* B **54** 10 652
[11] Wirtz L, Tang J-Z and Burgdörfer J 1997 *Phys. Rev.* B **56** 7589

Semicond. Sci. Technol. **13** (1998) A27–A29. Printed in the UK

PII: S0268-1242(98)91869-7

Nonsymmetric conduction induced by the shape of electron billiards

H Linke†, W Sheng†, H Xu†, P Omling† and P E Lindelof‡

† Solid State Physics and Nanometer Consortium, Lund University, Box 118,
S-221 00 Lund, Sweden
‡ Niels Bohr Institute, University Copenhagen, Universitetsparken 5,
DK-2100 Copenhagen Ø, Denmark

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We find that the longitudinal resistance of triangular electron billiards, which have no symmetry axis perpendicular to the direction of the current, is in the nonlinear regime not symmetric on current reversal. Our experiments indicate that the effect is inherent in the nonsymmetric shape of the cavities and is related to electron wave interference. Quantum mechanical modelling of nonlinear electron transport in open triangular cavities confirms that nonsymmetric effects such as those observed can be induced by the cavity shape alone.

When finite voltages are used in measurements of the longitudinal resistance of small devices, nonlinear effects can be observed. These can be nonsymmetric with respect to zero voltage if the conductor lacks a symmetry axis perpendicular to the current direction and if, at the same time, the conductor is smaller than the electron phase coherence length so that the local description of transport breaks down [1]. In previous observations of nonsymmetric resistance (NSR) effects in mesoscopic devices, the symmetry of the scattering potential was broken because of the random position of impurities [2–7]. In the present work we study a more controlled situation in which the scattering potential is determined by the shape of the sample. Using two-dimensional electron billiard cavities with a size smaller than the electron mean free path, we observe an NSR that appears to be related to the nonsymmetric shape of the triangular cavities used. In fact, theoretical modelling of electron transport in triangular cavities in the nonlinear regime shows that the nonsymmetric shape of the cavities alone can give rise to an NSR similar to that observed experimentally.

Electron billiards of various shapes were defined in GaAs/AlGaAs two-dimensional electron gas material using electron beam lithography and shallow wet etching. The Fermi energy and mean free path in the unstructured areas of the Hall bars were about 10 meV and 11 μm at 0.3 K, respectively. The lithographic dimension of the billiards was 3 μm or less, so that transport inside all billiards was predominantly ballistic. In order to study the symmetry of the resistance on current reversal, a tunable and reversible dc bias voltage was added to the ac component used for the resistance measurement ($I_{ac} < 5$ nA rms).

Figure 1 shows the variation of the resistance with dc bias voltage of three different billiards at zero magnetic field. The three devices, two triangular in shape and one rectangular (inserts in figure 1), were coupled in series on the same wafer and the data shown were recorded simultaneously. The two triangular billiards were oriented in opposite directions on the Hall bar, and the sign of the voltage is defined in such a way that, for negative (positive) bias voltage, the probe electrons are injected into the billiards from the left (right). Clearly, the resistance of the two triangular billiards, which have no symmetry axis perpendicular to the direction of the current, is, at a temperature of 0.3 K (thick lines) and at small voltages ($|U| < 1$ mV), not symmetric with respect to the zero bias voltage. In contrast, the rectangular (symmetric) billiard does not exhibit such an NSR. We have studied in total about ten devices (triangular and rectangular cavities as well as point contacts) and found strong experimental indications that the observed NSR is (i) inherent in the nonsymmetric shape of the electron cavities and (ii) related to electron wave interference. In support of the latter, we note that the NSR is suppressed when the temperature is increased above a few kelvin (thin lines in figure 1), where electrons are dephased by electron–electron interaction. Further, the NSR is limited to small bias voltages ($|U| < 1$ mV), small enough that phase coherence is not destroyed by inelastic scattering among the nonequilibrium electrons in the cavities [8,9]. We have also found that the NSR is strongly influenced by a magnetic flux of the order of h/e through the area of the device [10] and that details of the NSR change on thermal cycling of the sample, as is generally observed in experiments in which electron wave interference is important.

One can think of a number of nonlinear effects related to electron interference that could, in principle, cause an NSR in a nonsymmetric billiard. For instance, the rate of phase breaking inside the dot due to current heating or due to inelastic scattering among the nonequilibrium electrons

0268-1242/98/SA0027+03$19.50

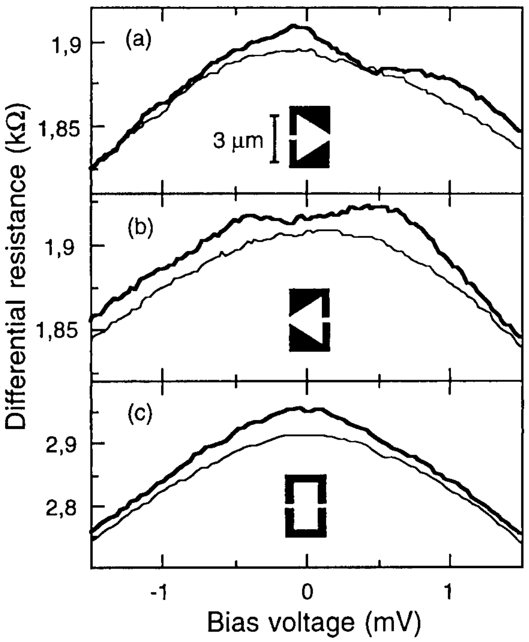


Figure 1. Variation of the resistance, $R(U)$, with bias voltage for three different electron billiards with the shapes indicated, measured simultaneously at zero magnetic field. Only the resistance of the triangular billiards at 0.3 K shows an NSR, and this depends on the orientation of the triangles with respect to the direction of the current (thick lines). At 4.6 K almost no NSR remains and $R(U)$ is qualitatively the same for all three billiard geometrices (thin lines).

in the cavity depends on the voltage drop at the source contact. In a triangular billiard, the two point contacts will in general have different conductances and, therefore, the rate at which conductance fluctuations or ballistic weak localization is suppressed by a bias voltage may depend on the sign of the voltage [9, 11]. In this report, however, we disregard these effects and focus solely on the importance of the shape of the billiard. In the following, we will consider electron transport in a triangular cavity at zero magnetic field in the nonlinear regime. We will show that some of our observations can be understood within the formalism of the nonlinear ballistic transport theory [12].

Our theoretical model is based on a quantum mechanical calculation of the differential conductance by assuming that the billiard is connected via point contacts to two-dimensional electron gas reservoirs which are in local equilibrium. We define the device structure by hard walls and assume that the reservoirs have the electrochemical potentials μ_1 and μ_2, such that $\mu_1 - \mu_2 = eV_{sd}$, when a source–drain bias voltage V_{sd} is applied. Electron transport inside the cavity is assumed to be ballistic, and phase-breaking processes occur only in the reservoirs. In general, the electrostatic potential inside the device may vary as a function of the position when a bias voltage is applied, and the voltage drop at the two contacts may be different. However, because we are here interested in the effect of the billiard shape alone, we simply assume a flat potential inside the device and equal voltage drops at the two quantum point contacts. Below, only a brief description of the theory is given, while details can be found elsewhere [12].

When a Fermi–Dirac distribution is assumed for the electrons in the two reservoirs, the total current through the device, at a finite temperature T, can be written as

$$I(V_{sd}) = \int_0^{\infty} \mathrm{d}\epsilon\, \{f[\epsilon - (\mu_F + eV_{sd}), T] \\ - f(\epsilon - \mu_F, T)\} J(\epsilon, V_{sd}) \qquad (1)$$

where μ_F is the electrochemical potential of the system at zero source–drain voltage and $J(\epsilon, V_{sd})\,\mathrm{d}\epsilon$ is the total current summed over all incident electrons in the energy range between ϵ and $\epsilon + \mathrm{d}\epsilon$. The differential conductance is given by

$$G(V_{sd}) = \frac{\partial I(V_{sd})}{\partial V_{sd}} \qquad (2)$$

which, in the limit of very small source–drain voltages V_{sd}, is equal to the linear response conductance [12]. In this model, deviations from the linear response conductance arise from the energy and voltage dependence of the current density $J(\epsilon, V_{sd})$. In order to calculate the current density we make use of the scattering matrix method [13]. For computational reasons we choose the side length of the triangles to be 1 μm, i.e. smaller than that in the experimental devices, while the size of the point contacts is chosen to be 100 nm, comparable with those in the real device. In all calculations presented below, a temperature of $T = 0.3$ K was used.

In figure 2(*a*) we show the calculated differential resistance as a function of the source–drain voltage for a triangular cavity with $\mu_F = 9$ meV. Clearly, the nonlinear signal is not symmetric with respect to zero voltage. In contrast, the calculated differential resistance of a rectangular (symmetric) billiard of comparable size (figure 2(*c*)) is perfectly symmetrical with respect to zero voltage, as is expected from symmetry arguments.

It should be noted, however, that the details of the nonlinear resistance depend strongly on the Fermi energy used. This can be seen by comparing figure 2(*b*) (where $\mu_F = 10$ meV was used in the calculation) with figure 2(*a*). Clearly, the calculated data for the higher value of μ_F are significantly different from those for the lower value of μ_F and show no resemblance to the experimental results, similarly to the case of a symmetric billiard where also theory (figure 2(*c*)) and experiment (figure 1(*c*)) show no resemblance. The reason for this strong dependence of the transmission on the Fermi energy is that the size of the cavity is about 20 times larger than the Fermi wavelength. As it is not possible to determine the exact Fermi energy inside the structure to an accuracy of better than 1 meV experimentally, a more detailed comparison of experiment and theory therefore requires experimental devices with tunable Fermi energy [14]. The important result at this stage, however, is that the existence of an NSR in electron billiards can be explained by the nonsymmetric shape of the billiards alone.

In general, nonsymmetric conductance effects can be observed if the scattering potential depends on the sign of the voltage. For instance, observations of nonsymmetric fluctuations of the conductance of ballistic point contacts

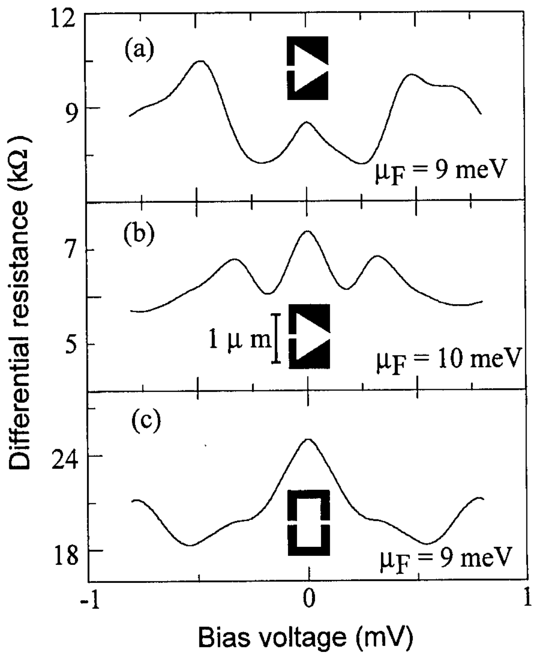


Figure 2. Quantum mechanically calculated differential resistance for triangular and rectangular electron billiards at T = 0.3 K as a function of the bias voltage, based on a model that assumes a flat potential inside the structure and equal potential drops at the two quantum point contacts. For clarity, the calculated curves were smoothed to suppress weak oscillations.

can be explained by the fact that the interfering electrons scatter with the different sets of impurities on the two sides of the point contact, depending on the sign of the voltage [5, 6]. In a ballistic cavity, however, the injected electrons scatter for both current directions with the same boundary of the cavity. Thus, the above simple, semiclassical argument cannot be directly applied to the observed NSR. To understand the physical mechanism behind the nonsymmetry in detail, it is, therefore, necessary to consider the effective scattering potential as a function of voltage, including the confining energy inside the cavity and charging effects.

In summary, we have observed an NSR of triangular electron billiards. Our theoretical results indicate that an NSR is inherent in the triangular shape of the billiards.

Acknowledgments

The authors acknowledge with thanks our fruitful discussions with C Pryor and S M Reimann. Financial support from the Swedish Research Councils for Natural (NFR) and for Engineering Sciences (TFR) and by the Göran Gustafsson Foundation for Research in Natural Sciences and Medicine is gratefully acknowledged. We thank Claus B Sorensen, the III–V NANOLAB and CNAST for the MBE-grown structures.

References

[1] Landauer R 1987 *Nonlinearity in Condensed Matter* ed A R Bishop *et al* (Berlin: Springer)
[2] Webb R A, Washburn S and Umbach C P 1988 *Phys. Rev.* B **37** 8455
[3] Kaplan S B 1988 *Surf. Sci.* **196** 93
[4] de Vegvar P G N, Timp G, Mankiewich M, Cunningham J E, Behringer R and Howard R E 1988 *Phys. Rev.* B **38** 4326
[5] Holweg P A M, Kokkedee J A, Caro J, Verbruggen A H, Radelaar R, Jansen A G M and Wyder P 1991 *Phys. Rev. Lett.* **67** 2549
[6] Ralph D C, Ralls K S and Buhrmann R A 1993 *Phys. Rev. Lett.* **70** 986
[7] Taboryski R, Geim A K, Persson M and Lindelof P E 1994 *Phys. Rev.* B **49** 7813
[8] Linke H, Omling P, Xu H and Lindelof P E 1997 *Phys. Rev.* B **55** 4061
[9] Linke H, Bird J P, Cooper J, Omling P, Aoyagi Y and Sugano T 1997 *Phys. Rev.* B **56** 14937
[10] Linke H, Omling P, Xu H and Lindelof P E 1996 *Proc. 23rd Int. Conf. On the Physics of Semiconductors* ed M Scheffler and R Zimmermann (Singapore: World Scientific) p 1593
[11] Linke H, Bird J P, Cooper J, Omling P, Aoyagi Y and Sugano T 1997 *Phys. Status Solidi* B **204** 318
[12] Xu H 1993 *Phys. Rev.* B **47** 15630
[13] Sheng W 1997 *J. Phys.: Condens. Matter* 1997 **9** 8369
[14] Linke H, Sheng W, Löfgren A, Omling P, Xu H and Lindelof P E to be published

Semicond. Sci. Technol. **13** (1998) A30–A32. Printed in the UK PII: S0268-1242(98)91852-1

Localized plasmons in point contacts

Henrik Bruus† **and Karsten Flensberg**‡

† Niels Bohr Institute, Blegdamsvej 17, DK-2100 Copenhagen, Denmark
‡ Danish Institute of Fundamental Metrology, Anker Engelunds Vej 1,
DK-2800 Lyngby, Denmark

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. Using a hydrodynamic model of the electron fluid in a point contact geometry we show that localized plasmons are likely to exist near the constriction. We attempt to relate these plasmons to the recent experimental observation of deviations of the quantum point contact conductance from ideal integer quantization. As a function of temperature, this deviation exhibits an activated behaviour, $\exp(-T_a/T)$, with a density-dependent activation temperature T_a of the order of 2 K. We suggest that T_a can be identified with the energy needed to excite localized plasmons, and we discuss the conductance deviations in terms of a simple theoretical model involving quasiparticle lifetime broadening due to coupling to the localized plasmons.

1. Introduction

Quantized conduction through a narrow point contact is one of the key effects in mesoscopic physics; the quantum point contact remains an important testing ground for the description of mesoscopic phenomena. Recently, significant deviations from the Landauer–Büttiker theory have been observed in quantum point contacts in the temperature dependence of the conductance quantization [1, 2] and as a so-called '0.7' structure or quasi plateau, appearing around 0.7 times the conductance quantum $2e^2/h$ [3]. Invoking a Luttinger liquid approach [4], the deviations have been discussed in terms of interaction effects [5–7]. However, firm conclusions have been difficult to obtain partly due to the narrow temperature range (0.1–4 K) in which the effect can be studied in conventional split gate quantum point contacts, where relatively close-lying one-dimensional subbands are formed.

Important progress was provided by the appearance of strongly confined GaAs quantum point contacts using a combination of shallow etching and a top gate [8]. In these new samples the conduction quantization can be followed up to around 30 K. In subsequent work [9] these samples were used to study the temperature dependence of deviations from perfect conductance quantization. At low temperature (~0.05 K) almost ideal quantized conductance is observed for the first conduction plateau, but deviations develop as the temperature is increased. The enlarged temperature range allowed for the observation of activated temperature dependence of these deviations: $\delta G(T) \propto \exp(-T_a/T)$. Furthermore, by changing the top gate it was found that T_a increases with increasing density. An explanation could not be found using the standard single-particle picture, and in the brief theory section of [9] we therefore suggested the inclusion of collective effects through plasmons. In short, we identified T_a as the energy needed to excite localized plasmons, and we discussed the conductance in terms of a simple theoretical model involving the additional effect of electrons scattering off the localized plasmons. In the present theoretical work we elaborate on that idea. In a point contact the charge is of course depleted. In order to study the collective excitations of such a system, we can approach it from two limits: (i) squeezing a homogeneous electron liquid, or (ii) connecting two spatially separated liquids. Below we follow the first approach, and we argue from a hydrodynamic model that localized plasmons may exist in realistic situations.

2. Plasmons of a homogeneous electron liquid in a cylinder

Following Fetter [10] we use a hydrodynamic model of a weakly damped, compressible charged electron fluid placed in a rigid, neutralizing positive background set to $+en_0$. The electron density is written as n_0+n, where n is a small perturbation, and the electronic velocity field is denoted $\boldsymbol{v}$. Finally, we include the electrostatic potential Φ and neglect radiation effects. The basic equations for the system are the linearized versions of the continuity equation and of Euler's and Poisson's equations [10]:

$$\partial_t n = -n_0 \nabla \cdot \boldsymbol{v} \tag{1}$$

$$\partial_t \boldsymbol{v} = -\frac{s^2}{n_0}\nabla n + \frac{e}{m}\nabla\Phi \tag{2}$$

$$\nabla^2\Phi = \frac{en}{\epsilon}. \tag{3}$$

Here $s = (\partial P/\partial n)/m = \sqrt{3/5}v_F$ is the sound velocity of the liquid. Combining equations (2) and (3) and introducing the plasma frequency $\omega_p = e^2 n_0/m\epsilon$ we obtain a wave equation for n:

$$-s^2\nabla^2 n + \omega_p^2 n = \omega^2 n. \tag{4}$$

0268-1242/98/SA0030+03$19.50

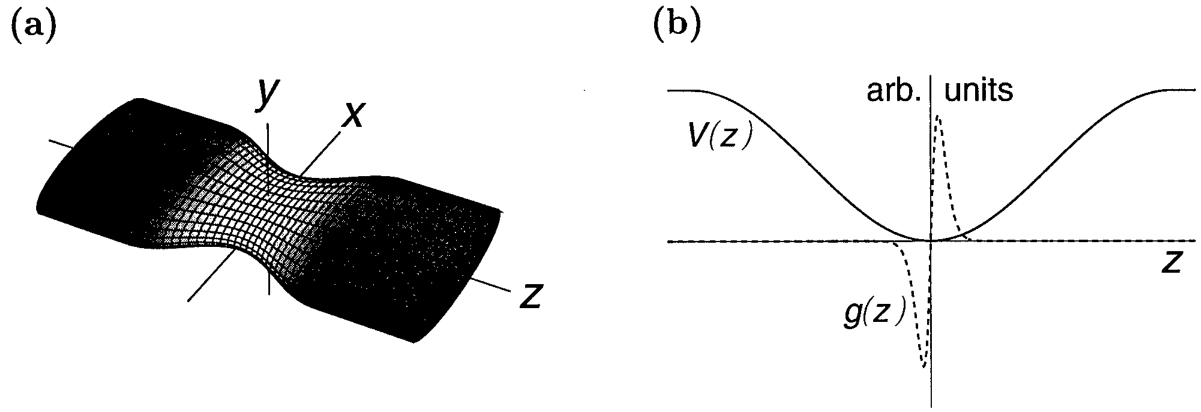


Figure 1. (a) The squeezed elliptical cylinder. High and low densities are represented by dark and light shading respectively. (b) The effective potential $V(z)$ (full curve) is determined by the parameter values of [9]. For this potential a solution of the wave equation (6) for $g(z)$ is found numerically (broken curve). The solution represents a localized plasmon with an energy of the order of 10 K.

For the case of a homogeneous electron liquid confined in a cylinder of radius R we let all fields have the dependence $f(r,\theta,z,t) = f_l(r)\exp[\mathrm{i}(l\theta + qz - \omega t)]$, with f_l being a Bessel function. Outside the cylinder $\Phi(r)$ must decay and fulfil equation (3) with $n = 0$, and so $\Phi_>(r) \propto K_l(qr)$. Inside the cylinder, $\Phi_<(r)$ can be either decaying, as $I_l(\kappa r)$, or oscillating, as $J_l(kr)$. The lowest lying modes are the decaying ones reminiscent of surface plasmons. The oscillation frequency ω is found by enforcing the boundary conditions that $\Phi(r)$ and its derivative are continuous at $r = R$, and that the normal component v_r of the velocity vanishes at the surface. The solution is

$$\omega^2 = qRI_l'(qR)\left(K_l(qR) - \frac{q}{\kappa}\frac{K_l'(qR)}{I_l'(\kappa R)}I_l'(\kappa R)\right)\omega_p^2 \\ \xrightarrow[q\to 0]{} \frac{e^2 n^{1\mathrm{D}}}{2\pi m\epsilon}q^2\ln\left(\frac{2}{qR}\right) \qquad (5)$$

where we also have given the 1D limit arising as $q \to 0$.

3. Plasmons of an inhomogeneous electron liquid in a squeezed elliptical cylinder

Next, to approach the point contact geometry we introduce two perturbations. First, the cylinder containing the inhomogeneous electron liquid is squeezed geometrically in a region of length 2Λ around $z = 0$, i.e. the radius becomes a function of z, say for example $R(z) = R_0 - \delta R[1 + \cos(\pi z/\Lambda)]\Theta(\Lambda - |z|)$. Similarly, a static z-dependent dip is imposed on the positive background charge density n_0 inside the squeezed cylinder, say $n_0(z) = n_0 - \delta n[1 + \cos(\pi z/L)]\Theta(L - |z|)$.

In the adiabatic limit where derivatives of $R(z)$ and $n_0(z)$ are neglected, the wave equation (4) remains separable in cylindrical coordinates, and we make the ansatz $n(r,\theta,z) = J_l(\kappa r)g(z)\exp[\mathrm{i}(l\theta - \omega t)]$, where J_l is a Bessel function and $g(z)$ an arbitrary function to be determined. The boundary condition $v_r(R(z)) = 0$ translates into a Neumann boundary condition $J_l'(\kappa R(z)) = 0$ and consequently the 'wavenumber' κ becomes a function of z, $\kappa = \kappa_{nl}(z) = \tilde{\gamma}_{nl}/R(z)$, with $\tilde{\gamma}_{nl}$ being the nth root of $J_l'(x)$. Furthermore, ω_p^2 also becomes a function of z, since $\omega_p^2(z) = e^2 n_0(z)/m\epsilon$, and similarly for the sound velocity, $s = s(z) \propto n_0(z)^{1/3}$. As a consequence the wave equation (4) for n is changed into the following eigenfunction equation for $g(z)$:

$$-s(z)^2\partial_z^2 g(z) + [s(z)^2\kappa_{nl}^2(z) + \omega_p^2(z)]\,g(z) = \omega^2 g(z). \qquad (6)$$

This is equivalent to Schrödinger's equation (with a position-dependent mass) as seen by the identifications $s^2 \leftrightarrow \hbar^2/2m$ and $s(z)^2\kappa_{nl}^2(z) + \omega_p^2(z) \leftrightarrow V(z)$. Since $\omega_p^2(0) < \omega_p^2(\pm\infty)$ bound states, i.e. localized plasmons, may exist. The 'effective potential' $V(z)$ is a sum of two terms; one, ω_p^2, is dipping down on the length scale L, the other, $s^2\kappa_{nl}^2$, is peaking on the length scale Λ. Depending on the relative strengths, shapes and length scales of the two terms the effective potential will appear rather differently. However, for realistic parameters, where the density variation dominates, we conclude that localized plasmons may exist in the squeezed, inhomogeneous cylindrical electron liquid as shown in figure 1.

The previous considerations dealt with a cylindrical geometry, but it is not difficult to approach the 2D case. The trick is simply to use elliptical coordinates (u, v, z) defined by $(x, y, z) = (\eta\cosh(u)\cos(v), \eta\sinh(u)\sin(v), z)$. The parameter η relates to the eccentricity of the ellipse. With these coordinates the wave equation separates as before. Instead of trigonometric functions of the angle θ we now obtain the Mathieu functions of the generalized angular variable v, and instead of Bessel functions we obtain the modified Mathieu functions of the generalized radial coordinate u. By letting the eccentricity η tend to infinity we end up with a 2D geometry close to the one realized in the quantum point contact experiments. The conclusions obtained for the circular cylinder can be restated for the elliptic cylinder, and thus localized plasmons are expected to exist in or near the constriction region of quantum point contacts.

4. Plasmon damping

So far we have treated only the undamped case. In real systems the collective plasmons are damped through their interaction with individual electron–hole pairs, the so-called Landau damping. This effect could be simulated by adding a damping term $-v/\tau$ to the right-hand side of the Euler equation (2). Instead we will leave the classical level of description and continue with a microscopic quantum treatment. The classical level is adequate for demonstrating the existence of the collective (almost classical) plasmon excitations, but fails when it comes to single-particle effects.

The point contact can be approximated by a 1D region, the constriction, connected at each end to 2D regions, the contacts. For this 2D–1D–2D model of the point contact we can estimate the frequency of the confined plasmon using our insight from the classical calculations: we calculate the dispersion relation for an infinite 1D-wire and insert the wavevector $q_c = 2\pi/L$, L being the length of the constriction and hence related to the size of the localized plasmon. The long-wavelength limit of the dispersion relation found in the random phase approximation (RPA) is

$$\omega_L^{1D} = \sqrt{v_f^2 + \frac{\gamma e^2 n^{1D}}{4\pi\varepsilon m^*}}\, q_c. \qquad (7)$$

Note how the second term under the square root resembles the classical result of equation (5). In [9] we used this formula successfully to fit the measured activation temperatures mentioned in the introduction.

The confined 1D plasmons will be Landau damped through their coupling to the 2D contacts outside the constriction. Inserting q_c in the RPA expression for the polarizability χ^{2D} we obtain the following rough estimate of the lifetime $\tau_p^{-1}/\omega_L^{1D}$ of the 1D plasmon coupled to the 2D contacts of the 2D–1D–2D model:

$$\frac{\tau_p^{-1}}{\omega_L^{1D}} \approx V^{2D}\mathrm{Im}\chi^{2D} \approx \frac{2\pi e_0^2}{q_c}\,\frac{m^*\omega_L^{1D}}{2\pi\hbar^2 v_F^{2D} q_c} \lesssim 1. \qquad (8)$$

The plasmons are seen to be damped, but not over-damped.

5. Quasiparticle lifetime

In the Landauer–Büttiker formalism the conductance is given by single-particle properties. Once a particle is launched in a given channel of the injecting lead the transmission probability amplitudes are governed by the elastic scattering matrix of the system. For a quasiparticle with a finite lifetime it is possible that a particle will decay before completing its traversal of the system. We propose that the observed deviation from perfect quantized conductance is indeed due to the finite lifetime of the quasiparticles. Furthermore we suggest that the main contribution restricting the lifetime comes from scattering against the localized plasmons. As demonstrated above, the localized plasmons provide a well defined finite energy $\hbar\omega_L$. Through the Coulomb interaction the electrons will scatter against the plasmons and hence the quasiparticle lifetime and the transmission properties are affected. The resulting lifetime and additional resistance is expected to exhibit an activated behaviour, $\tau^{-1} \propto \exp(-T_a/T)$, since a finite energy is needed to excite the localized plasmon. We are thus led to identify the activation temperature with the energy of the localized plasmon: $T_a = \hbar\omega_L/k_B$.

6. Conclusion

Using a hydrodynamic description of the electron fluid, we have shown that localized plasmons with a frequency ω_L are likely to exist near the constriction of a point contact. We have sketched how a more complete microscopic quantum calculation may account for a quasiparticle lifetime broadening τ^{-1} with a thermal activation behaviour $\tau^{-1} \propto \exp(-\hbar\omega_L/k_B T)$. We relate this broadening with conductance and are led to identify the recently measured activation temperature T_a for conductance deviations with the frequency ω_L of the localized plasmon.

Acknowledgments

It is a pleasure to thank our experimental colleagues Anders Kristensen, Poul Erik Lindelof and Jesper Nygård, at the Ørsted Laboratory of the Niels Bohr Institute, for many stimulating discussions. HB is supported by the Danish Natural Science Research Council through Ole Rømer Grant no 9600548.

References

[1] Tarucha S, Honda T and Saku T 1995 *Solid State Commun.* **94** 413
[2] Yacoby A, Stormer H L, Wingreen N S, Pfeiffer L N, Baldwin K W and West K W 1996 *Phys. Rev. Lett.* **77** 4612
[3] Thomas K J, Simmons M Y, Nicholls J T, Mace D R, Pepper M and Ritchie D A 1995 *Appl. Phys. Lett.* **67** 109
[4] Kane C L and Fisher M P A 1992 *Phys. Rev.* B **46** 15233
[5] Kawabarta A 1996 *J. Phys. Soc. Japan* **65** 30
[6] Shimizu A 1996 *J. Phys. Soc. Japan* **65** 1162
[7] Oreg Y and Finkelstein A M 1996 *Phys. Rev.* B **54** 14265
[8] Kristensen A *et al* 1998 *J. Appl. Phys.* **83** 607
[9] Kristensen A *et al* 1998 *Physica* B at press
[10] Fetter A L 1985 *Phys. Rev.* B **32** 7676

Semicond. Sci. Technol. **13** (1998) A33–A36. Printed in the UK PII: S0268-1242(98)91881-8

On properties of boundaries and electron conductivity in mesoscopic polycrystalline silicon films for memory devices

**G P Berman†, G D Doolen†, R Mainieri†, J Rehacek†,
D K Campbell‡, V A Luchnikov§ and K E Nagaev||**

† Theoretical Division and CNLS, Los Alamos National Laboratory, Los Alamos, NM 87545,USA
‡ Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green St, Urbana, IL 61801-3080, USA
§ Institute of Chemical Kinetics and Combustion, Siberian Branch of Russian Academy of Sciences, Institutskay 3 Street, Novosibirsk 630090, Russia
|| Institute of Radio-Engineering and Electronics, Russian Academy of Sciences, Mokhovaya Street 11, 103907 Moscow, Russia

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We present the results of molecular dynamics modelling of the structural properties of grain boundaries (GBs) in thin polycrystalline films. The transition from crystalline boundaries with low mismatch angle to amorphous boundaries is investigated. It is shown that the structures of the GBs satisfy a thermodynamical criterion suggested by Keblinski *et al* (1996 *Phys. Rev. Lett.* **77** 2965). The potential energy of silicon atoms is closely related to a geometrical quantity: tetragonality of their coordination with their nearest neighbours. A crossover of the length of localization is observed. To analyse the crossover of the length of localization of the single-electron states and properties of conductance of the thin polycrystalline film at low temperature, we use a two-dimensional Anderson localization model, with the random one-site electron charging energy for a single grain (dot), random non-diagonal matrix elements and random number of connections between the neighbouring grains. The results for the crossover behaviour of the localization length of the single-electron states and characteristic properties of conductance are presented in the region of parameters where the transition from an insulator to a conductor takes place.

1. Introduction

Recently it was demonstrated that thin polycrystalline silicon films are promising materials for future room-temperature single-electron devices [1–4]. The main reasons which make this material so attractive are the following. (I) Usually, the film's thickness varies from 1 to 5 nm, and the average lateral grain size is 10 nm or less. In this case, the energy of an electron in a single grain, E_e, is bigger than the thermal energy even at room temperature, $E_e > T = 300\,\mathrm{K}$. (II) It is believed that, in these films, the characteristic resistance, R_b, of the potential barriers between the grains is big enough, $R_b > R_c = h/e^2 \sim 25\ \mathrm{k}\Omega$. If both of these conditions are satisfied, an electron is strongly localized in the grain. At the same time, one can regulate (to some extent) the electron conductivity in these films by varying the gate voltage and creating a current channel. Those electrons which are stored in the grains (storage dots) create a Coulomb repulsion for those electrons which are involved in the current channel. This allows one to implement memory operations in these films at room temperature using a Coulomb blockade effect. Different implementations of these ideas have been discussed, for example, in [1–4]. To satisfy both of the above-mentioned conditions, the boundaries between the nanocrystalline grains in these films play an important role. For example, one of the most important characteristics of the electron transport in polysilicon films is connected with the distribution of crystalline and amorphous grain boundaries (GBs) [5]. At room temperature, the main factor which determines the structure of a GB is a mutual misorientation of the neighbouring crystalline grains. At present there does not exist a consistent theoretical approach for the description of the GBs in thin polycrystalline films.

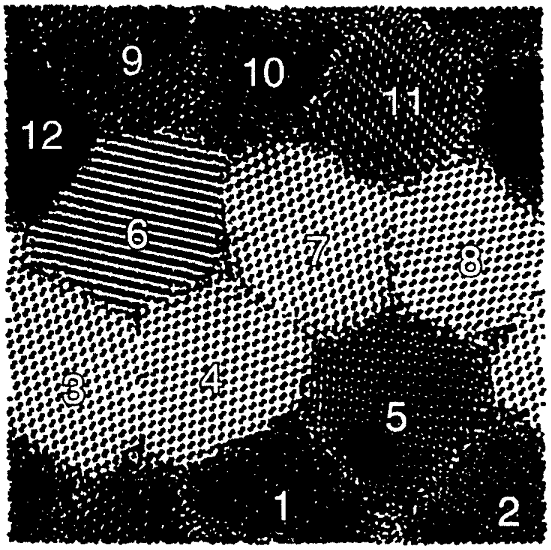


Figure 1. The (X, Y) projection of the polycrystalline film.

2. Molecular dynamics (MD) simulation of grain boundaries at room temperature

The initial configuration of the polysilicon film was generated using the Voronoi triangulation algorithm: first, the centres of the grains were defined; then each grain was obtained by filling the space, nearest to the centre of the grain, with the diamond lattice at a chosen orientation. The atoms with very high potential energies were eliminated from the GBs. The film had the dimensions 16.7 nm × 16.7 nm × 3.1 nm and contained 44 174 atoms in 12 grains (figure 1). The majority of the grains in our model were oriented randomly, so they had both tilt and twist misorientation. Four grains, namely grains 3, 4, 7 and 8, were generated so that the misorientation tilt angle between grains 3 and 4 was $\alpha_{3-4} = 11.0°$; between grains 4 and 7 $\alpha_{4-7} = 11.0°$ and between grains 7 and 8 $\alpha_{7-8} = 15.0°$. GB_{3-4} and GB_{7-8} are close to the $(1, 1, 1)$ interface, and GB_{4-7} is close to the $(0, 0, 1)$ interface. The $(1, 1, 1)$ planes in these four grains are perpendicular to the plane of the figure. The average linear size of a grain in our model, $d_{gr} \simeq 4.8$ nm, is of the same order as for the grains in the experimental polycrystalline silicon films obtained recently by thermal annealing of amorphous Si deposited by decomposing silane on an SiO_2 substrate [4]. Periodic boundary conditions were used in the X and Y directions, and free boundary conditions were used in the Z direction. It is shown that the potential energy of the atoms closely correlates with the degree of their tetragonal coordination with the nearest neighbours. To estimate the quality of the tetragonal coordination of the jth atom, we calculated the value $\mathcal{T}_j = \sum_{l<k}(l_{jl} - l_{jk})^2/15\bar{l}_j^2$, which is called the *tetragonality* [6]. Here l_{jk} is the length of the kth edge of the tetrahedron of general shape formed by the four nearest neighbours of the jth atom; $\bar{l}_j$ is the average length of the edge. By definition, $\mathcal{T}_j$ is equal to zero for an ideal tetrahedron (as it is in the case of the ideal diamond structure), and it increases if the shape of the tetrahedron is distorted. The value, $\mathcal{T}_j$, is correlated with the dispersion of the edge lengths of the tetrahedron. According to the criterion suggested in [6], the spatial figure formed by four points can be recognized as having the 'good tetrahedral shape' if $\mathcal{T}_j$ is less than or equal to $\mathcal{T}^{(c)} = 0.018$. The value of the tetragonality of an atom is very sensitive to the number of atoms in the first coordination shell, z. In the bulk amorphous phase of silicon, for which $z_a \simeq 4.04$, the average tetragonality is equal to $\mathcal{T}_{am} = 0.015$ which is slightly less than the critical value $\mathcal{T}^{(c)}$. In liquid silicon (in which the first coordination shell consists on average of $z_{lq} \simeq 4.46$ neighbours), the average value of tetragonality is $\langle \mathcal{T} \rangle_{lq} \simeq 0.034$. In table 1, we present the data for the potential energy per atom, density and average tetragonality for the longest GBs (for which it was possible to make measurements with reasonable accuracy). Densities of the tilt crystalline GBs between grains 3, 4, 7 and 8 are practically the same as, or slightly smaller than, the density of the bulk crystal, ρ_{cr}. The densities of other GBs are systematically larger than ρ_{cr}. In covalent materials with friable structure, such as silicon and germanium, some GBs, indeed, contract. This was observed in x-ray diffraction experiments for $(1, 0, 0)$ twist GBs in Ge [8] and in MD simulation of the $(1, 1, 1)$ twist boundaries in silicon [9]. It is remarkable that the contraction is observed for the twist boundaries. In such GBs, the covalent bonds are probably stretched rather than squeezed, and the system demonstrates a tendency to restore their lengths at the cost of a reduction of the GB's volume. MD simulations show that, in our model of polysilicon film, the majority of the GBs between the grains with random mutual orientation are disordered. The grains with small mutual misorientation are connected by crystalline GBs. The GBs with 'medium-angle' misorientation have a complex inhomogeneous structure. These boundaries consist of crystalline connections interspersed with disordered regions. In disordered regions, the misfit between the crystal structures of the neighbouring grains is compensated by non-crystalline arrangements of atoms, such as 5- and 7-fold rings. In general, the structure of the GBs satisfies the thermodynamical criterion suggested in [10]. The potential energy of the silicon atoms closely correlates with the tetragonality, $\mathcal{T}$, of their coordination with their nearest neighbours. The majority of atoms in the polysilicon grain boundary are well coordinated tetrahedrally, even if they are arranged in a non-crystalline manner in the high-angle GB. From the observed small values of the average tetragonality, it follows that the high-angle disordered GBs can be characterized as amorphous.

The results obtained are important for better understanding of structural properties of thin polycrystalline silicon films, which have been used recently for memory devices. However, to obtain more reliable data, larger polysilicon simulations should be considered.

3. Crossover behaviour of the localization length of single-electron states and the properties of conductivity

To model the properties of conductive electrons in a thin polycrystalline silicon film at low temperature, we

Table 1. Potential energy, density and average tetragonality of grain boundaries in the polysilicon model.

	U (eV/atom), ±0.04	ρ (g cm^{-3}), ±0.02	$\langle T\rangle$, ±0.0003
GB$_{3-4}$	−4.24	2.29	0.005
GB$_{4-7}$	−4.18	2.31	0.014
GB$_{7-8}$	−4.20	2.34	0.011
GB$_{3-6}$	−4.19	2.39	0.010
GB$_{6-7}$	−4.16	2.39	0.013
GB$_{4-1}$	−4.17	2.41	0.013
GB$_{1-5}$	−4.18	2.38	0.011
GB$_{5-2}$	−4.09	2.36	0.017
GB$_{9-10}$	−4.21	2.37	0.012
GB$_{10-11}$	−4.15	2.38	0.015
GB$_{9-12}$	−4.14	2.38	0.017
a–Si	−4.15[a]	2.29[a]	0.015[a]
cr–Si	−4.335	2.324	0

[a] Data for a–Si obtained from the model in [7].

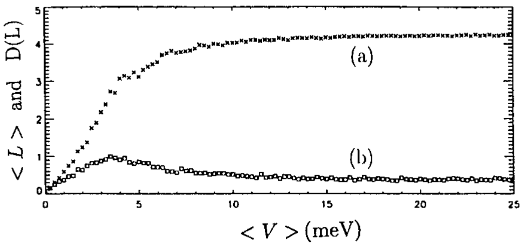


Figure 2. Demonstration of the crossover behaviour of the average localization length of the single-electron states. Dependence of the average localization length $\langle L\rangle$ (points a) and of the dispersion $D(L)$ (points b) on the average value of the matrix element, $\langle V\rangle$. $(M, N) = (18, 12)$; $\langle E\rangle = 50$ meV; $\delta E = 25$ meV; $V \in [0.9; 1.1]\langle V\rangle$.

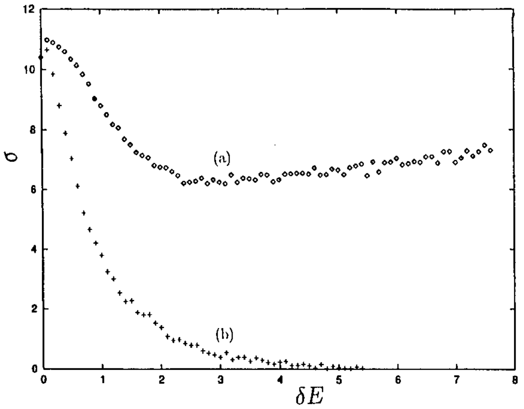


Figure 3. Dependence of the conductance, $\sum T_{\alpha,\beta}$, on the disorder parameter δE. The lattice size $(M, N) = (11, 11)$. 11 channels in the left and right electrodes were used. Dimensionless parameters: $V = 1$; $E_F = 0$. For points a, two of the boundaries are randomized: $(n = 1, m = 1, \ldots, M)$ and $(n = N, m = 1, \ldots, M)$. For points b, the whole active region is randomized.

use an approach based on the two-dimensional tight-binding Anderson model. Each nanosilicon grain (with coordinates (m, n) in the (X, Y) directions) is considered as an individual dot. The characteristic energy of quantization in the Z direction, E^z, for the lowest level, is of the order $E^z \sim$ 1–300 meV [1,2]. We assume that the total one-site electron energy is $E_{m,n} = \langle E\rangle + \delta E_{m,n}$, where $\delta E_{m,n}$ is a random variable (as the thickness of the film fluctuates). At low temperature, the electron conduction is governed by quantum tunnelling between the dots, which we describe by the matrix elements $V_{m,n}^{(m',n')}$. Below we consider only the simplest case of the neighbouring transitions: $V_{m,n}^{(m\pm1,n\pm1)} = \langle V\rangle + \delta V_{m,n}$, where $\langle V\rangle$ is the average value of the off-diagonal matrix elements, and $\delta V_{m,n}$ describes the fluctuations related to the random characteristics of the potential barriers between the neighbouring dots. Let us now estimate the number of conduction electrons which occupy an individual dot. The maximum value of the 2D electron density in the film can be chosen as $n_e \sim 10^{11}$–10^{12} cm^{-2}. Consider a subsystem of conduction electrons in a polysilicon film as a 2D one, with the size of the active region 100 nm × 100 nm [1,2]. Then, the number of conduction electrons, $\mathcal{N}$, in this region is of the order $\mathcal{N} = 10$–100. If the average size of the grain (dot) in the (X, Y) plane is of the order of 10 nm, then one has, on average, 100 dots in the active region. So, the average number of conductive electrons in a dot is of the order 0.1–1. In this paper, we consider the conduction electrons as non-interacting. To describe the electron subsystem at low temperature, we use a 2D Anderson tight-binding model with the Hamiltonian $\mathcal{H} = \sum_l E_l|l\rangle\langle l| + \sum_{l,k} V_{l,k}|l\rangle\langle k|$, where E_l describes the diagonal disorder. Figure 2 shows the dependences of the average length of localization, $\langle L\rangle$, and the dispersion of the length of localization, $D(L) = \langle (L - \langle L\rangle)^2\rangle$, of the single-electron states, $\Psi^{(i)}(m, n)$, as a function of the average value of the matrix element $\langle V\rangle$. (The size of the lattice is $(M, N) = (18, 12)$. Zero boundary conditions were chosen.) In figure 2, $\langle E\rangle = 50$ meV, $E_{m,n} \in [50$–$12.5; 50 + 12.5]$ meV and $\delta V_{m,n} \in [0.9; 1.1]\langle V\rangle$. For each state, $i = 1, \ldots, MN$, the length of localization was introduced as $L^{(i)} = (L_m^{(i)} + L_n^{(i)})/2$. We calculated $L_m^{(i)}$ as $L_m^{(i)} = [\sum_{m,n=1}^{M,N}(m - \bar{m}^{(i)})^2|\Psi(m, n)|^2]^{1/2}$, where $\bar{m}^{(i)} = \sum_{m,n=1}^{M,N} m|\Psi(m, n)|^2$. (Similar expressions were used to calculate $L_n^{(i)}$.) As one can see from figure 2, the fluctuations of the average localization length, $D(L)$, exhibit a characteristic maximum at $\langle V\rangle \approx \langle V_c\rangle = 3.5$ meV. For $\langle V\rangle < \langle V_c\rangle$, the system is essentially insulating. For $\langle V\rangle > \langle V_c\rangle$, the system exhibits metallic properties. We are at present investigating this crossover behaviour in connection with the metal–insulator transition. Finally, in figure 3, we present the characteristic results of numerical simulations of the dimensionless conductance in this system, $\sigma = \sum |T_{\alpha,\beta}|^2$, where $T_{\alpha,\beta}$ is the transmission amplitude from channel α to channel β. The method we used is based on a Green's function approach [11]. The lattice size is $(11, 11)$, with 11 channels in the left and the right electrodes. In figure 3, the horizontal axis is the dimensionless value of the diagonal disorder, δE. The dimensionless off-diagonal matrix element was chosen equal to $V = 1$ in the active region and in the electrodes. The results shown in figure 3 correspond to the dimensionless Fermi energy $E_F = 0$. In figure 3 the data indicated by diamonds (points a) correspond to the randomization of only two boundaries, $(n = 1, m = 1, \ldots, M)$ and $(n = N, m = 1, \ldots, M)$.

The data indicated by crosses (points b) corresponds to the randomization of the whole active region. One can see that increasing the diagonal disorder leads to a crossover behaviour of the conductance in the case when only the boundaries are randomized. The results presented in this paper can be useful for the performance of memory devices based on thin polycrystalline silicon films. Our further work is related to studying the influence of the Si–SiO$_2$ interfaces on the conductivity in polycrystalline silicon films and on the noise component.

Acknowledgments

We would like to thank D K Ferry and P Lomdahl for useful discussions. VAL and KEN thank the Center for Nonlinear Studies, Los Alamos National Laboratory, for hospitality. This work was supported by the Linkage Grant 93-1602 from the NATO Special Programme Panel on Nanotechnology and by the Defense Advanced Research Projects Agency.

References

[1] Yano K, Ishii T, Hashimoto T, Kobayashi T, Murai F and Seki K 1994 *Extended Abstr. Int. Conf. on Solid State Devices and Matter (1994)* pp 325–7
[2] Yano K, Ishii T, Hashimoto T, Kobayashi T, Murai F and Seki K 1994 *IEEE Trans. Electron Devices* **41** 1628
[3] Tiwari S, Rana F, Hanafi H, Hartstein A, Crabbé E F and Chan K 1996 *Appl. Phys. Lett.* **68** 1377
[4] Kamal A H M, Lützen J, Sanborn B A, Sidorov M V, Kozicki M N, Smith D J and Ferry D K 1997 *Semicond. Sci. Technol.* submitted
[5] Nomure S, Zhao X, Aoyagi Y and Sugano T 1996 *Phys. Rev.* B **54** 13974–9
[6] Medvedev N N and Naberukhin Yu I 1987 *J. Noncryst. Solids* **94** 402
[7] Luchnikov V A, Medvedev N N, Appelhagen A and Geiger A 1996 *Mol. Phys.* **88** 1337
[8] Lamarre P, Schmückle F, Sickafus K and Sass S L 1984 *Ultramicroscopy* **14** 11
[9] Phillpot S R and Wolf D 1995 *Philos. Mag.* A **72** 453
[10] Keblinski P, Phillot S P and Wolf D 1996 *Phys. Rev. Lett.* **77** 2965
[11] Berman G P and Nagaev K E in preparation

Semicond. Sci. Technol. **13** (1998) A37–A40. Printed in the UK PII: S0268-1242(98)91849-1

3D simulation of GaAs/AlGaAs quantum dot point contact structures

D Vasileska†, M N Wybourne‡, S M Goodnick† and A D Gunther†

† Center for Solid State Electronics Research, Arizona State University, Tempe, AZ 85287-5706, USA
‡ Department of Physics and Astronomy, Dartmouth College, Hanover, NH 03755-3528, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We present simulation results for the activation barrier ϕ_0 and the number of electrons N in a quantum dot structure that are obtained from the self-consistent solution of the 3D Schrödinger–Poisson problem. We observe an approximately linear rise in the electrostatic barrier ϕ_0 as a function of the gate bias once pinchoff occurs. There is an associated linear decrease in the number of the electrons in the dot region. Calculated values of ϕ_0 and N are in an agreement with those utilized in the energy balance analysis for similar structure investigated experimentally in connection with negative conductance behaviour. The simulation results also suggest that, if the dot area is too small, all particles may be depleted from the dot before the input and output barrier forms, thus preventing the bistable operation.

Devices that exhibit negative differential conductance are attractive candidates for potential application as oscillators because of their very short switching times. Several years ago, Hess *et al* [1] showed that a heterostructure hot-electron diode exhibits S-shaped negative differential conductance (SNDC) at low temperatures. Later, Higman *et al* [2] demonstrated that it is possible to observe the switching behaviour even at 300 K on a similar type of structure. More recently, Wu *et al* [3] reported that a multiterminal lateral hot-electron transistor (LHET) with a split-gate arrangement (figure 1) also exhibits pronounced SNDC. Contrary to the heterostructure diode case, the LHET has symmetric current–voltage (I–V) characteristics around the zero source–drain bias V_{ds} and can switch at lower voltages by varying the bias on the split gates.

For all three devices, a model based on the energy balance approach has been quite successful in explaining the experimental I–V characteristics [4, 5]. It assumes that all the carriers injected into the dot thermalize with the cold carriers in the dot which are isolated from the rest of the quasi-two-dimensional electron gas (Q2DEG) with two quantum point contacts (figure 1). According to this model

$$I\phi_0 = eN\left\langle\frac{\partial E}{\partial t}\right\rangle_{coll} \qquad (1)$$

where I is the current injected over the barrier, ϕ_0 is the gate voltage dependent barrier height, e is the elementary charge, N is the number of electrons in the dot and the brackets represent the average energy loss per electron in the dot due to acoustic and optical phonon scattering. Under the assumption that, at each of the quantum point contacts, I is limited by the thermionic emission over the barrier, a set of equations results, in which fitting parameters are the electron temperature in the dot T_e, ϕ_0, N and the source–drain bias V_{sd}. What physically occurs in the dot is that two stable operating points may exist simultaneously for a given value of the voltage, V_{sd}. The first one corresponds to a low-current state with the dot temperature close to the equilibrium lattice temperature. The second one corresponds to a state in which a relatively large current may flow with the effective electron temperature in the dot much higher than the lattice temperature.

To obtain quantitative results for the parameters ϕ_0 and N, we carried out numerical simulations using tools developed at Arizona State University. Using a quasi-3D model, we calculated the equilibrium conduction band profile for the LHET structure from figure 1 for various biases on the split gates, from which we extracted the magnitude of the activation barrier ϕ_0. We also investigated whether one can fine tune the number of electrons in the dot by varying the voltage on the plunger gate.

The LHET shown in figure 1 was realized in GaAs/Al$_x$Ga$_{1-x}$As technology using standard processing techniques. The split-gate structure consists of four electrodes that can be independently biased. The three narrow electrodes are externally connected together and biased as one. The plunger gate is either biased independently or held at the same potential as the other three electrodes. The MBE-grown heterostructure consists of 1 μm of undoped GaAs grown on a semi-insulating (SI) GaAs substrate, 85 nm of undoped Al$_x$Ga$_{1-x}$As ($x =$

0268-1242/98/SA0037+04$19.50

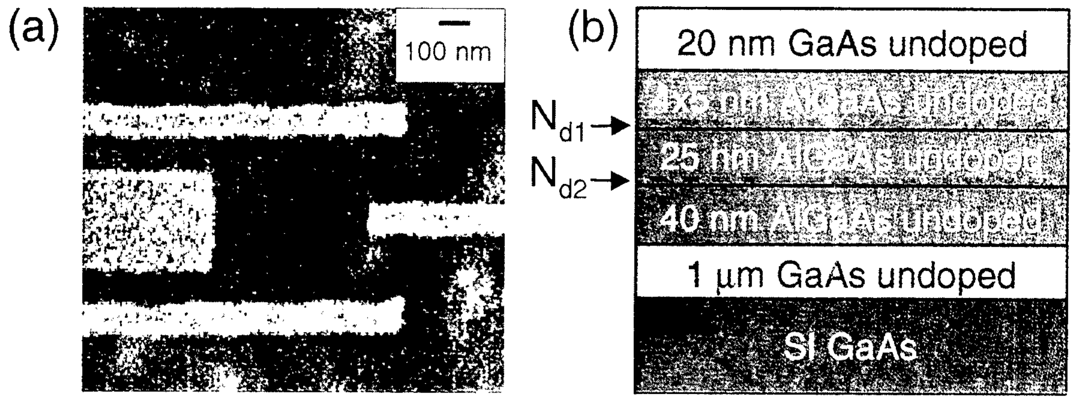


Figure 1. (*a*) Scanning electron micrograph of the top split-gate structure. (b) MBE-grown epitaxial layers.

0.25) layer which incorporates five Si delta-doped layers (four of which have $N_{D1} = 1 \times 10^{12}$ cm^{-2} and one has $N_{D2} = 3 \times 10^{12}$ cm^{-2}) and 20 nm of undoped GaAs cap layer. For the undoped GaAs layer we use standard parameters for the effective mass and relative dielectric constant. Following Adachi [6], the conduction band offset between the Al$_x$Ga$_{1-x}$As and GaAs layers is calculated as $\Delta E_c = 1.0075x + 0.2405x^2$. For the electron effective mass in the Al$_x$Ga$_{1-x}$As layer we use $m^* = (0.067 + 0.083x)m_0$. The relative permitivity of this layer is calculated as $\epsilon_r = 13.18 - 3.12x$. For the effective Schottky barrier height of the gate electrode to the GaAs cap layer we use $\Phi_b = 0.7$ eV, and for the binding energy of the deep donors in the Al$_x$Ga$_{1-x}$As we assume that $E_D = 25$ meV. To obtain realistic values for the confining potential in the dot, at the exposed surface we use the more physical Neumann boundary conditions [7].

Details about the calculation of the Q2DEG density in an ungated structure are given in [8]. To find the confining potential in the dot, we coupled our 3D Poisson equation solver with our 1D Schrödinger solver. The relatively large dot size justifies our treatment of electrons in the dot as a finite-extent Q2DEG. We use both the incomplete factorization scheme and the preconditioned Bi-CGSTAB method for the numerical solution of the matrix equation $\boldsymbol{Ax} = \boldsymbol{b}$ that results from the standard 7-point finite difference approximation scheme of the 3D Poisson equation in the presence of piecewise-constant dielectric constants [9]. Dirichlet boundary conditions are imposed at the Schottky contacts, whereas Neumann boundary conditions are assumed at the artificial boundaries.

In figure 2 we show the conduction band profile E_c and the energy of the ground subband for the ungated structure. Assuming that 50% of the Si atoms in the barrier layer are electrically active and using -2.36×10^{12} cm^{-2} for the surface-charge density yields a channel electron density of $N_s = 3.36 \times 10^{11}$ cm^{-2} at $T = 4.2$ K. Shubnikov–de Haas and Hall measurements revealed an electron density of 3.7×10^{11} cm^{-2}. The close agreement between the calculated and experimentally derived values for N_s confirms the adequacy of our model. Simulation results for the confining potential for the electrons in the plane 5.4 nm below the Al$_x$Ga$_{1-x}$As–GaAs interface (near the maximum of the electron charge distribution) and an applied bias of −3.6 V on all gates are shown in figure 3. The nonuniform finite-difference tensor-product grid used here has 50 × 180 × 54 points along the x-axis (length), y-axis (depth) and z-axis (width). We use the same surface-charge density as in the 1D case, which is assumed to be constant over the entire free surface, thus ignoring the charge reversal at the edges of the gates. It is obvious that the confining potential in the dot resembles a truncated parabola with a flat bottom because of the charge accumulation in the dot. An additional important feature is the presence of a finite potential energy barrier at the quantum point contacts that connect the dot with the input and output leads. Variation of both ϕ_0 and N as a function of V_g is shown in figure 4. We observe an approximately linear increase in the electrostatic barrier height ϕ_0 once pinchoff occurs. The opposite is true for N; it decreases approximately linearly with increasing V_g. In figure 5 we show the variation of ϕ_0 and N with the voltage on the plunger gates $V_{plunger}$, for fixed voltages on the narrow gates. It is obvious that one can fine tune the number of electrons in the dot by varying $V_{plunger}$. Because of the large dot size, the barrier heights are only slightly affected with the variation of $V_{plunger}$. This suggests that one can

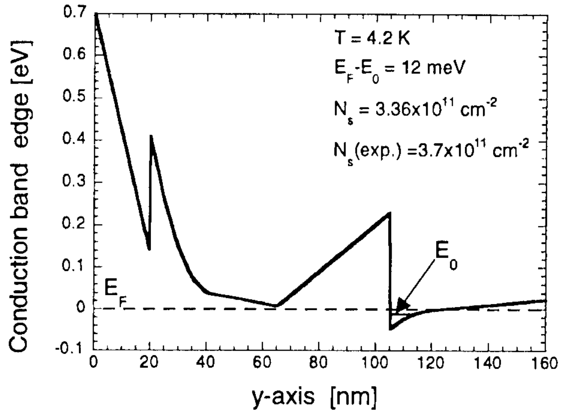


Figure 2. Conduction band profile along the growth direction (y-axis) and the energy of the ground quantized energy level. Fermi level E_F is the reference energy level.

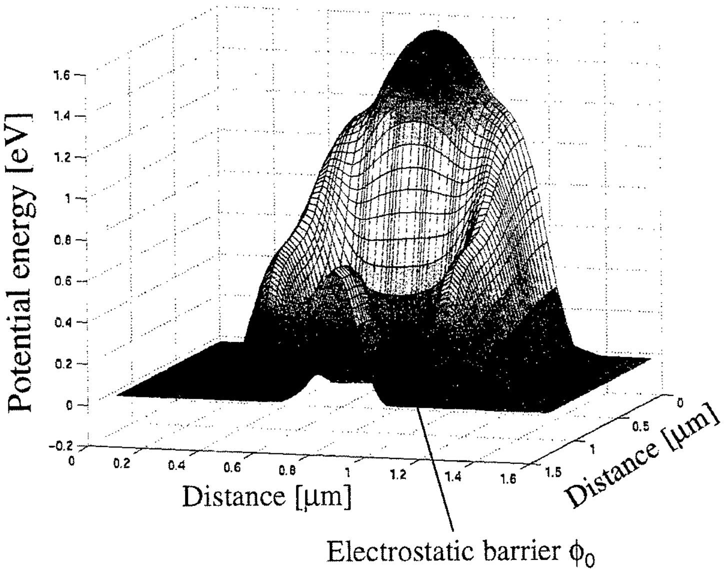


Figure 3. Conduction band profile in the x–z-plane.
(This figure can be viewed in colour in the electronic version of the article; see `http://www.iop.org`)

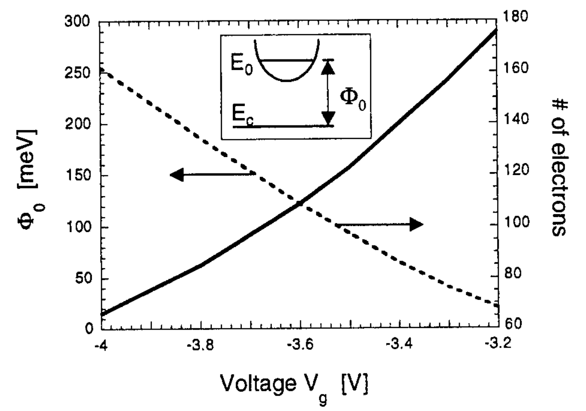


Figure 4. Variation of ϕ_0 and N as a function of the bias on the split gates V_g.

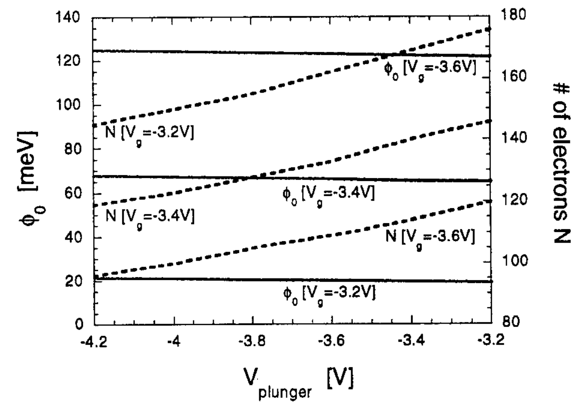


Figure 5. Variation of ϕ_0 and N as a function of $V_{plunger}$ for fixed values of the voltage V_g on the narrow gates.

investigate the influence of N on the I–V characteristics of the LHET by varying the voltage on the plunger gate without altering ϕ_0.

Finally, we want to comment on the relevance of our simulation results. For the same LHET, the energy balance model predicts that, at $T = 4.2$ K and for $V_g = -2.95$ V, $N = 155$ and $\phi_0 = 49$ meV [5]. According to Laux *et al* [10], the reversal of the surface charge density at the exposed surface leads to a transition region whose width is approximately 10% of the width of the gate opening. The inclusion of this will shift the gate voltage by about -0.2 V. In addition, the inclusion of space quantization in the lateral direction will further reduce N and the width of the quantum point contacts (QPCs) and, therefore, increase ϕ_0. This can be accounted for through an additional shift of the gate voltage of about -0.15 V [10]. From the results shown in figure 4 it follows that, for $V_g = -3.3$ V, $N = 157$ and $\phi_0 \approx 40$ meV. These values are in agreement with the energy balance model predictions. We want to comment that the uncertainty of the physical dot size and the strong dependence of ϕ_0 on the physical width of the QPC could be a reason for the small discrepancy in the results for ϕ_0.

In conclusion, our simulator gives good estimates for the number of electrons in the dot N and the electrostatic barrier height ϕ_0. However, further refinements of our simulation model are required to account for the quantization in the lateral direction.

References

[1] Hess K, Higman T K, Emanuel M A and Coleman J J 1986 *J. Appl. Phys.* **60** 3776

[2] Higman T K, Higman J M, Emanuel M A, Hess K and Coleman J J 1987 *J. Appl. Phys.* **62** 1495

[3] Wu J C, Wybourne M N, Berven C, Goodnick S M and Smith D D 1992 *Appl. Phys. Lett.* **61** 2425

[4] Higman T K, Miller L M, Favaro M E, Emanuel M A, Hess K and Coleman J J 1988 *Appl. Phys. Lett.* **53** 1623

[5] Goodnick S M, Wu J C, Wybourne M N and Smith D D 1993 *Phys. Rev.* B **48** 9150

[6] Adachi S 1985 *J. Appl. Phys.* **58** R1

[7] Chen M and Porod W 1995 *J. Appl. Phys.* **78** 1050

[8] Akis R, Vasileska D, Ferry D K, Bird J P, Okubo Y, Ochiai Y, Ishibashi K, Aoyagi Y and Sugano T *Physica* B in press

[9] Vasileska D, Gross W J, Kafedziski V and Ferry D K *VLSI Des.* in press

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

Semicond. Sci. Technol. **13** (1998) A41–A43. Printed in the UK PII: S0268-1242(98)90811-2

Scale factor mapping of statistical and exact self-similarity in billiards

A P Micolich†, R P Taylor†, J P Bird‡, R Newbury†,
T M Fromhold§, J Cooper‖, Y Aoyagi¶ and T Sugano¶

† School of Physics, University of New South Wales, Sydney, NSW 2052, Australia
‡ Center for Solid State Electronic Research, Arizona State University,
AZ 85287-6206, USA
§ Physics Department, Nottingham University, Nottingham, NG7 2RD, UK
‖ Cavendish Laboratory, University of Cambridge, Cambridge CB3 0HE, UK
¶ Frontier Research Program, RIKEN, 2-1 Hirosawa, Wako, Saitama, 351-01,
Japan

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We investigate the fractal magnetotransport properties of mesoscopic billiards. Employing three geometries, we find the self-similarity to be either statistical or exact. We use a correlation analysis to investigate the relationship between these two families of scaling behaviours.

1. Introduction

Magnetoconductance fluctuations (MCF) act as 'magnetofingerprints' of the statistics generated by electron trajectories colliding with the walls of mesoscopic semiconductor billiards [1,2]. At low temperatures, phase-coherent quantum interference processes produce a rich spectrum of MCF [1]. Mandelbrot introduced a 'fractal' dimension D_F to quantify structure occurring on increasingly fine scales. Fractal structure in a data trace has the property that $1 < D_F < 2$ [3]. Recently, MCF observed in billiards were predicted to be fractal [4] and experiments revealed D_F in the range 1.1–1.6 [5–7]. Scaling fractals are a subset where the structure is scale invariant [3]: for 'exact self-similarity' an intrinsic pattern repeats itself exactly at different scales, in contrast to 'statistical self-similarity' where statistical properties of the trace are invariant under scaling [4,6,7]. The form of self-similarity in fractal MCF is highly topical. Square [6] and the stadium [7] geometries reveal statistical self-similarity, while a Sinai geometry [8] displays exact self-similarity [5]. Whereas D_F quantifies the presence of MCF structure at different magnifications, a correlation function F has been introduced to assess self-similarity [9]. The correlation function is a powerful tool in mapping out the scaling factors of fractal MCF. In this paper, three distinct geometries of billiards are used to investigate the relationship between exact and statistical self-similarities and the conditions required for their observation.

2. Exact self-similarity

Figure 1 shows the three billiard geometries which we label A, B and C. The diagrams represent electrostatic gates on the surface of an AlGaAs/GaAs heterostructure [5,6]. In each case, the electronic mean free path ($\ell > 5\ \mu$m) is significantly larger than the billiard and the electronic motion is sensitive to the billiard's geometry [1,2]. In the hard-wall approximation, the square of figure 1(a) supports stable trajectories, while the presence of the curved wall (the Sinai diffuser) in figure 1(b) generates chaotic trajectories [8]. In figure 1(c), the central antidot is the Sinai diffuser and its presence can be tuned with gate bias, allowing a study of the transition from the 'empty' square to the Sinai geometry [5]. In a semiconductor billiard defined by soft walls, the classical phase space is mixed (regular and chaotic) [4] and the effect of the Sinai diffuser is of considerable interest.

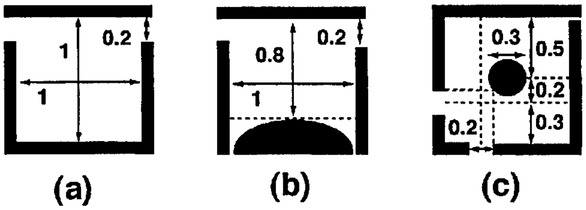


Figure 1. Schematic diagrams of the three billiard geometries considered. Dimensions are in microns.

Figure 2 shows MCF for the three billiards for magnetic fields applied perpendicular to the billiard plane. We first consider the MCF of device C (antidot activated) represented in figure 2(c) [9]. (We note that figure 2(c) is a damped Weierstrass fit to the MCF data of device C. The Weierstrass equation and the MCF raw data can be found in [9]. The correlation analysis is, however, applied to the MCF raw data.) The MCF clusters around two distinct scales: 'fine' (f) structure

0268-1242/98/SA0041+03$19.50

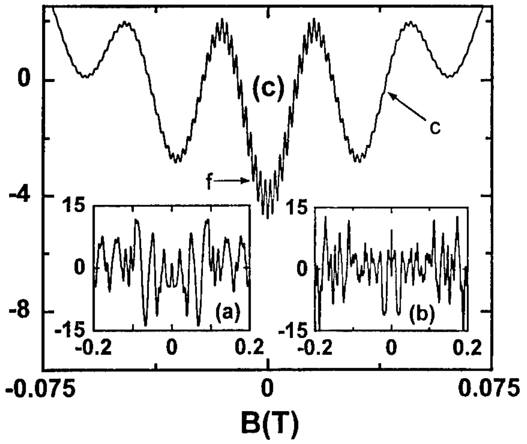


Figure 2. $\delta G(B) = G(B) - \langle G(B) \rangle$ versus B for (a) billiard A, (b) billiard B and (c) billiard C (with the antidot activated). In (c), the label 'f' points to the fine structure superimposed on the coarse 'c' structure.

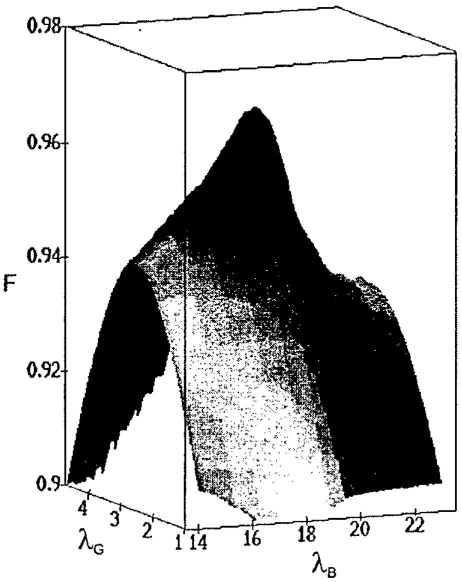


Figure 3. F versus (λ_B, λ_G) for device C (antidot activated).

superimposed on 'coarse' (c) structure with characteristic magnetic field scales $\Delta B_f \approx 0.5$ mT and $\Delta B_c \approx 10$ mT. For exact self-similarity in MCF, the f structure will be a scaled version of the c structure. Defining the c scale conductance as $\delta G_c(B) = G_c(B) - \langle G_c(B) \rangle$ (where $\langle\,\rangle$ represents an average performed over magnetic field) and the f scale conductance amplitude as $\delta G_f(B) = G_f(B) - \langle G_f(B) \rangle$, then conductance and magnetic field scaling factors, λ_G and λ_B, can be determined which give $\delta G_c(B) \approx \lambda_G \, \delta G_f(\lambda_B B)$. The correlation function F quantifies the similarity:

$$F = 1 - \frac{\langle [\delta G_c(B) - \lambda_G \, \delta G_f(\lambda_B B)]^2 \rangle^{1/2}}{N}. \tag{1}$$

N is a normalization constant which sets $F = 0$ when $\delta G_c(B)$ and $\lambda_G \, \delta G_f(\lambda_B B)$ are randomly related traces and $F = 1$ if the two traces are mathematically identical [9]. Figure 3 shows F as a function of λ_G and λ_B for the MCF data trace and reveals a maximum of $F = 0.97$ for $(\lambda_B, \lambda_G) = (18.6, 3.7)$. The single peak indicates a unique pair of scaling factors and the high value of F at the maximum confirms that the MCF has exact self-similarity. To be fractal, the MCF will continue to cluster at finer scales. An 'ultra-fine' (u) structure is superimposed on the f structure with a characteristic field scale ΔB_u determined by the common scaling factor $\Delta B_c / \Delta B_f = \Delta B_f / \Delta B_u = \lambda_B$. Assigning the c structure as the 'initiator' pattern with index number $n = 1$ [3], the MCF can be mapped out using $\Delta B_n / \Delta B_{n+1} = \lambda_B$. The conductance amplitudes of selected features ΔG_n are also related by a common scaling factor $\Delta G_n / \Delta G_{n+1} = \lambda_G$. As required for fractal behaviour [4, 5], the scaling factors are related by $\lambda_G = (\lambda_B)^{\beta}$, where $D_F = 2 - \beta$. Performing an equivalent analysis replacing ΔG_n with $\langle [\delta G(B) - \delta G(B + \Delta B_n)]^2 \rangle$ (the variance of $\delta G(B)$ with ΔB_n for the nth cluster), we obtain

$$\frac{\langle [\delta G(B) - \delta G(B + \Delta B_n)]^2 \rangle}{\langle [\delta G(B) - \delta G(B + \Delta B_{n+1})]^2 \rangle} = \left(\frac{\Delta B_n}{\Delta B_{n+1}} \right)^{\gamma}$$

$$\text{where } \lambda_B = \frac{\Delta B_n}{\Delta B_{n+1}} \tag{2}$$

and $D_F = 2 - \gamma/2$. Exact self-similarity is preserved when adjusting the bias V_o applied to the outer square gate and hence the number of conducting modes i in each port. (V_o, i, D_F) evolves as (−0.51, 7, 1.55), (−0.52, 5, 1.52) and (−0.55, 3, 1.45). Adjustment of the bias V_I applied to the inner gate, in order to remove the Sinai diffuser, results in a decrease in F [9] and a change to statistical self-similarity (see below).

3. Statistical self-similarity

In figures 2(a) and 2(b) the MCF data measured at 50 mK on devices A and B appear aperiodic, possessing a range of ΔB components [1]. They are predicted to obey statistics analogous to fractional Brownian motion [3] and hence are fractal [4]. For an arbitrary increment ΔB, the MCF are expected to follow

$$\langle [\delta G(B) - \delta G(B + \Delta B)]^2 \rangle \propto (\Delta B)^{\gamma} \tag{3}$$

where $D_F = 2 - \gamma/2$. Extracting D_F from linear fits to $\log\{\langle [\delta G(B) - \delta G(B + \Delta B)] \rangle\}$ versus $\log(\Delta B)$, we find (V_o, i, D_F) = (−0.60, 3, 1.2), (−0.47, 3, 1.1) and (−0.52, 5, 1.4) for devices A, B and C (antidot de-activated) respectively. A study of D_F versus i will be presented elsewhere. The scaling described by equation (3) is termed statistical self-similarity [4]. Selecting two arbitrary scales, ΔB_1 and ΔB_2, gives

$$\frac{\langle [\delta G(B) - \delta G(B + \Delta B_1)]^2 \rangle}{\langle [\delta G(B) - \delta G(B + \Delta B_2)]^2 \rangle} = \left(\frac{\Delta B_1}{\Delta B_2} \right)^{\gamma}$$

$$\text{where } \lambda_B = \frac{\Delta B_1}{\Delta B_2}. \tag{4}$$

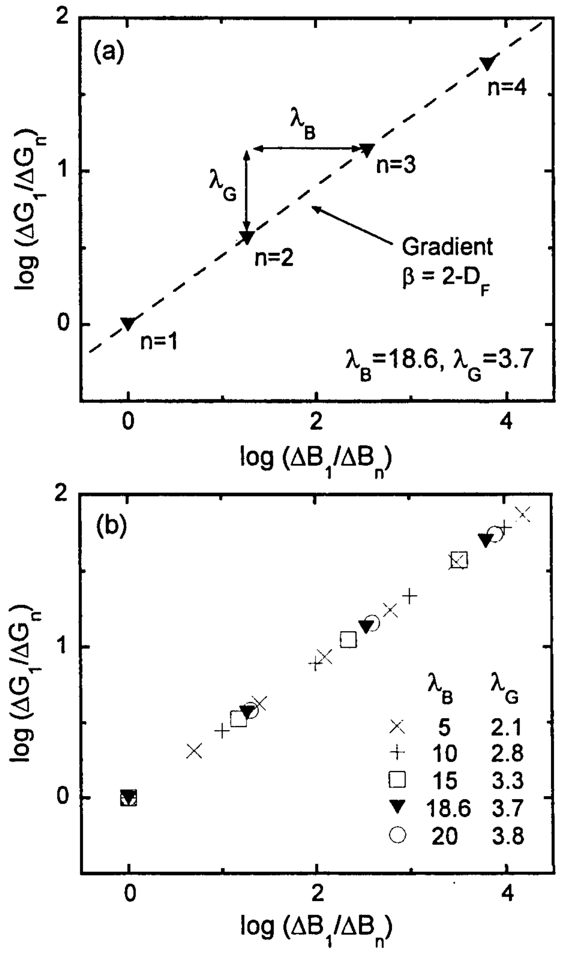


Figure 4. Schematic representation of (a) exact self-similarity and (b) statistical self-similarity. See text for details.

The difference between equations (2) and (4) is that the latter holds for a continuous range of λ_B while the former is only true for a single value. For both exact and statistical self-similarity, the MCF are described by a unique D_F which links each (λ_B, λ_G) pair by $\lambda_G = (\lambda_B)^{\beta}$ and $D_F = 2-\beta$. This proposed picture is summarized in the schematic of figure 4, where the antidot of billiard C is activated (figure 4(a)) and de-activated (figure 4(b)). In figure 4(a), the exact self-similarity is generated by a single hierarchy of points ($n = 1, 2, 3, 4, \ldots$) located at equal spacings determined by $(\lambda_B,\lambda_G) = (18.6,3.7)$ on a line with gradient $\beta = 2 - D_F$. In figure 4(b), de-activation of the antidot generates additional hierarchies at spacings determined by their (λ_B, λ_G) values. For simplicity we have shown just four additional hierarchies. For statistical self-similarity, a continuous range of hierarchies are expected. We are currently applying correlation analysis to the statistically self-similar traces to map out the ranges of scaling factors present in the billiards. We have also shown an identical gradient for figures 4(a) and 4(b), although in reality D_F is slightly less for the statistical case (see above). This is being investigated.

4. Conclusions

The two forms of observed fractal MCF can be understood within a common framework where exact self-similarity is generated by a unique pair of scaling factors compared with a range required for the statistical case. Geometries A, B, C (antidot de-activated) and the stadium [7] are all 'empty' billiards and exhibit statistical self-similarity. In contrast, introduction of the antidot at the centre of billiard C removes all but one pair of scaling factors from the system, generating exact self-similarity. We are currently using a soft-wall model to relate the scaling factors to stable electron trajectories.

References

[1] Marcus C M, Rimberg A J, Westervelt R M, Hopkins P F and Gossard A C 1992 *Phys. Rev. Lett.* **69** 506
[2] Jalabert R A, Baranger H U and Stone A D 1990 *Phys. Rev. Lett.* **63** 2442
[3] Mandelbrot B 1982 *The Fractal Geometry of Nature* (San Francisco, CA: Freeman)
[4] Ketzmerick R 1996 *Phys. Rev.* B **54** 10 841
[5] Taylor R P *et al* 1997 *Phys. Rev. Lett.* **78** 1952
[6] Micolich A P, Taylor R P, Newbury R, Bird J P, Wirtz R, Detlmann C P, Aoyagi Y and Sugano T 1998 *J. Phys.: Condens. Matter* **10** 1339
[7] Sachrajda A S, Ketzinerick R, Gould G, Feng Y, Kelly P J, Delage A and Wasilewski Z 1998 *Phys. Rev. Lett.* **80** 1948
[8] Sinai Ya G 1970 *Russ. Math. Survey* **25** 137
[9] Taylor R P, Micolich A P, Newbury R and Fromhold T M 1997 *Phys. Rev.* B **56** R12 733

Semicond. Sci. Technol. **13** (1998) A44–A46. Printed in the UK PII: S0268-1242(98)91891-0

Interfaces, boundaries and quantum device transport

H L Grubin† and D K Ferry‡

† Scientific Research Associates, Inc, Glastonbury, CT 06033, USA
‡ Center for Solid State Electronics Research, Arizona State University, Tempe, AZ 85287-6206, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. Simulations with recursive Green's functions with damping are applied to the study of a variety of nanostructure configurations, including quantum point contacts, lateral resonant structures and symmetric and asymmetric open quantum dots. The roles of fluctuations and length dependence are introduced and a comparison with experiments on multiple-point contact systems is discussed.

1. Introduction

In earlier studies [1], application of the quantum Liouville equation in the coordinate representation demonstrated the necessity of finding a suitable set of phase space boundary conditions to reflect the effects of device–device interaction and proximity effects on nanostructure transport. In particular it was shown that the self-consistent potential and density distributions are extremely sensitive to small variations in these boundary conditions. While these issues are present in multidimensional quantum confined systems, and the techniques used for the one-dimensional studies are in principle amenable to two and three dimensions, computational requirements may be excessive. Instead we have begun to apply recursive Green's functions (RGFs) techniques to the simulations of nanostructure transport. The emphasis in this study is on structure and how simple structural variations manifest themselves in changes, sometimes dramatic, in the conductance versus energy relation. The simplest structural variations include the *effects* of isolated impurities on the output of quantum point contacts, and in the case of open quantum dots the effects of placement and opening width on the conductance. In the latter case we find that structural variations can change conductance versus energy relations from the familiar plateau structure (see [2]) to one where the conductance displays oscillations. This study is limited in scope to rectangular structures of finite potential energy. The soft potential arising from self-consistency is not treated, and we expect the appearance of some oscillations which, while not spurious in a mathematical sense, would not appear experimentally.

We note that the recursive construction of Green's functions on a tight-binding lattice is not new. It was first introduced in the early 1980s (see [3] for a detailed discussion). As used below, a damping factor, $i\eta$, is included. An early illustration of the use of RGFs techniques for the computation of the conductance of an adiabatic quantum point contact (together with a discussion comparing the results with the mode matching methods with hard potential boundaries) is provided in [3].

Each of the examples discussed below involves the Kubo formalism for computing the conductance. For comparison with experiment there are two routes we take. (1) We specify the potential energy distribution and compute the conductance for a specific Fermi energy. We then vary the potential energy distribution in a way that mimics the variation of depletion width with bias and again compute the conductance. (2) Alternatively, for a specific potential energy distribution the conductance is computed as a function of Fermi energy. The latter is discussed below. There is equivalence when the transmission coefficient is a function of the difference between the Fermi energy and the subband energy.

2. Discussion

To set the stage for the results we refer to figure 1 for a structure in which the point contact opening on the left is smaller than that on the right. As discussed below we think that this calculation is relevant to the experiments of [2]. First we provide some details. The structure is defined on a 30×60 equally spaced rectangular array. The lattice spacing is 5 nm. The width of this structure is $W = 145$ nm, the widths $W_1 = W_2 = 50$ nm, $W_3 = W_4 = 20$ nm, the lengths $L = 295$ nm, $L_1 = 95$ nm, and $L_2 = 100$ nm. For this structure the potential energies on all of the boundaries as well as the shaded region were set at $6V_{aa}$, where $V_{aa} = \hbar^2/2ma^2$. For $a = 5$ nm and m the effective mass of GaAs electrons $V = 24.2$ meV.

The points in figure 1 display the conductance as a function of Fermi energy. The full line in the figure are obtained from the ideal conductance, $G = (2e^2/h)k_f W/\pi$, where the number of modes is given by $M \approx k_f W/\pi = (W/a\pi)(E_F/V_{aa})^{1/2}$, and W is the opening of the leftmost point contact. The results are clearly dominated by the quantum point contact with the smaller opening, a

0268-1242/98/SA0044+03$19.50

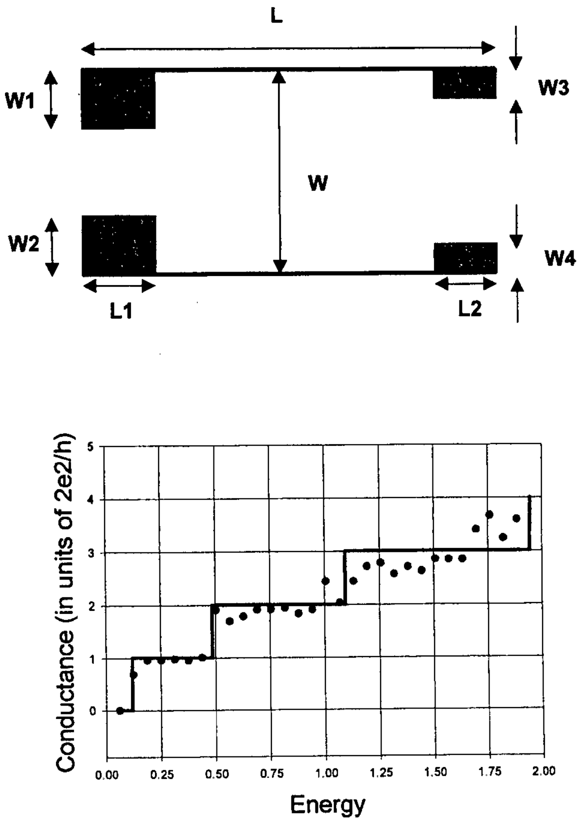


Figure 1. Conductance versus Fermi energy (in multiples of V_{aa}) for the structure shown.

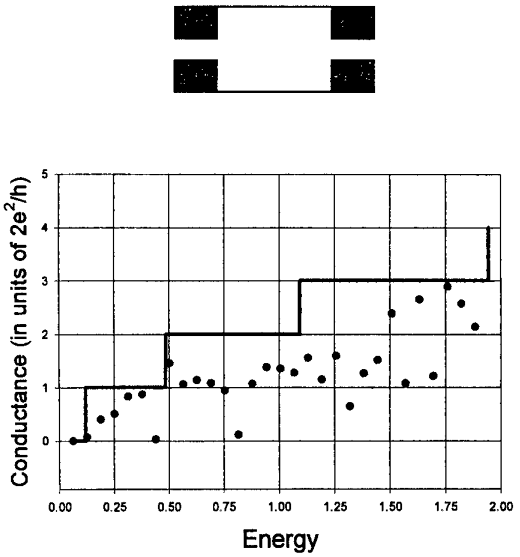


Figure 2. Conductance versus Fermi energy (in multiples of V_{aa}) for the symmetric structure shown.

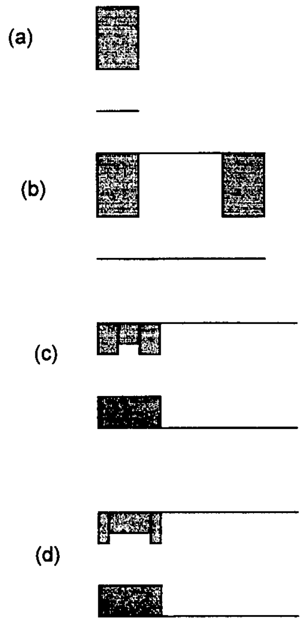


Figure 3. Various structures studied.

result that appears to be consistent with earlier discussions. (Note that in this formula we see an equivalence where the number of modes increases as the gate bias is made less negative, and the number of modes is increased by increasing the Fermi energy through an increase in the numbers of carriers.) This result is also likely to be relevant to the experiments of [2] for two isolated quantum point contacts in which only one of the gates was swept, the other being held at ground. In [2] the conductance displayed a series of plateaux, as in figure 1. Caution is advised here since the potential energy distribution in figure 1 is set at a fixed constant value. Additional calculations for an increased opening of the right-hand point contact did not qualitatively alter the calculated results.

The calculations take a strong turn, with the appearance of oscillations, when the opening of the second quantum point contact is reduced to that of the first one; see figure 2. These results remain qualitatively the same for structures in which the separation of the quantum point contacts is increased to 2 μm. The full line is figure 2 is the same as that in figure 1. It is apparent that the magnitude of the conductance would be better fitted if reduced by a factor of 2, where the inverses of the conductance are additive.

It is expected that the conductance variations are dependent on geometry. In addition to the above structures the conductance was computed for the structures shown in figure 3. Figure 3(a) showed conductance variations similar to those of figure 1. Figure 3(b), which is another representation of an open quantum dot, displayed a more regular set of conductance plateaux, suggesting the importance of geometry in the interpretation of conductance behaviour. Calculations with figures 3(c) and 3(d) were performed to determine the role of such things as isolated impurities in the conductance. The situation for figure 3(c) resulted in only marginal changes in the conductance from a 'perfect' point contact, while that for figure 3(d) where the change is over a longer distance resulted in severe distortions. Some of the latter are probably consequences of the sharpness of the boundaries.

3. Conclusions

The purpose of this study was to demonstrate how simple structural variations manifested themselves in changes, sometimes dramatic, in the conductance versus energy relation. The study demonstrated the effects of isolated impurities on the output of quantum point contacts, and in the case of open quantum dots the effects of placement and opening width on the conductance. In particular, we found that structural variations can change conductance versus energy relations from the familiar plateau structure to one where the conductance displays oscillations.

References

[1] Grubin H L, Caspar J R and Ferry D K 1997 *Phys. Status Solidi* **204** 365
[2] Taylor R P, Fortin S, Sachrajda A S, Adams J A, Fallahi M, Davies M, Coleridge P T and Zawadzki P 1992 *Phys. Rev.* B **45** 9149
[3] Ferry D K and Goodnick S M 1997 *Transport in Nanostructures* (Cambridge: Cambridge University Press) section 3.8.7

Semicond. Sci. Technol. **13** (1998) A47–A50. Printed in the UK PII: S0268-1242(98)91855-7

Probing the interactions of C_{60} on Si(100)-(2 × 1) using anisotropic molecular manipulation

P Moriarty†, Y-R Ma†, M D Upward†, P H Beton† and D Teehan‡

† Department of Physics, University of Nottingham, Nottingham NG7 2RD, UK
‡ Daresbury Laboratory, Daresbury, Warrington, Cheshire WA4 4AD, UK

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We have investigated the response of C_{60} molecules adsorbed on Si(100)-(2 × 1) to manipulation induced by a scanning tunnelling microscope. Our results show that the C_{60}–Si(100)-(2 × 1) interaction is greater than the C_{60}–C_{60} interaction. Attempts to move a molecular pair result in the transfer of one molecule across a dimer row due to barrier lowering caused by the intermolecular interaction. Our results suggest that C_{60} is chemisorbed and this is confirmed by Si 2p core-level photoemission spectra.

Lateral manipulation using a scanning tunnelling microscope (STM) may be used to probe the interactions experienced by atoms and molecules adsorbed on solid surfaces [1–3]. Following initial work by Eigler and Schweizer [1] this approach has been used to study the interaction of electrons at the surface of noble metals with adsorbed transition metal atoms [2] and more recently the anisotropy of a molecule–substrate interaction for a high-index surface [3]. The use of this technique to investigate molecular interactions has so far been restricted to metal substrates and a low-temperature environment since the adsorbate–substrate combinations studied to date have energy barriers which are too low to suppress diffusion at room temperature. The extension of this work to room-temperature operation requires a much larger diffusion barrier and therefore a much stronger adsorbate–substrate interaction and progress towards this goal has been limited by the difficulty of initiating manipulation of strongly adsorbed molecules. In a number of recent papers controlled manipulation of molecules at room temperature has now been reported [4–7]. However, progress to date has encompassed only placement of adsorbates and no investigations of the interactions of strongly adsorbed molecules have been reported.

In this paper we describe a series of experiments in which molecular manipulation is used to investigate the intermolecular and molecule–surface interactions of C_{60} adsorbed on the Si(100)-(2 × 1) surface. This adsorbate–substrate combination has recently been the focus of intense interest and several different models for adsorption have been proposed [8–10]. By using STM manipulation we are able to discriminate between these models. We show that the response of C_{60} to manipulation implies that C_{60} is chemisorbed [10] and this result is confirmed by synchrotron radiation photoemission experiments.

We use an ultrahigh vacuum (UHV) STM (Oxford Instruments SPM group, formerly WA Technology, Cambridge, UK) operating at room temperature. Electrochemically etched W tips cleaned in vacuum by heating were used throughout. The Si(100) samples were degassed at 750 °C overnight, flash annealed at 1200 °C for 20 s and then held at 800 °C for 3 min before being cooled to room temperature. C_{60} was sublimed from a Knudsen cell at a rate of 2×10^{-3} monolayers s^{-1}. Photoelectron spectroscopy (PES) experiments were performed at the Synchrotron Radiation Source, Daresbury, UK, using incident photon energies of 120 and 140 eV and a hemispherical analyser (resolution ~0.3 eV). For the PES experiments C_{60} was sublimed from a resistively heated Ta envelope.

Figure 1(*a*) shows an STM image in which C_{60} molecules appear as circular features and the rows of the Si(100)-(2 × 1) reconstruction run up the image. The rows arise from the dimerization of top-layer atoms [11]. The Si(100)-(2 × 1) surface and the adsorption site for C_{60} are shown schematically in figure 2. We find that C_{60} is adsorbed in the troughs between the dimer rows at the four-dimer site (labelled A in figure 2) in agreement with previous studies [9, 10].

Displacement of an adsorbate across a surface was first demonstrated by Eigler and Schweizer [1]. This original work was performed at low temperature but we have recently demonstrated that it may also be performed at room temperature [4, 5]. Following our work further demonstrations of room-temperature manipulation have been published [6, 7]. To induce manipulation on the Si(100)-(2 × 1) surface the tip–sample separation is first reduced by decreasing the gap resistance to a value ~1 GΩ. The tip is then moved across the surface in a controlled manner (typically through a total distance of 3 nm in steps

0268-1242/98/SA0047+04$19.50

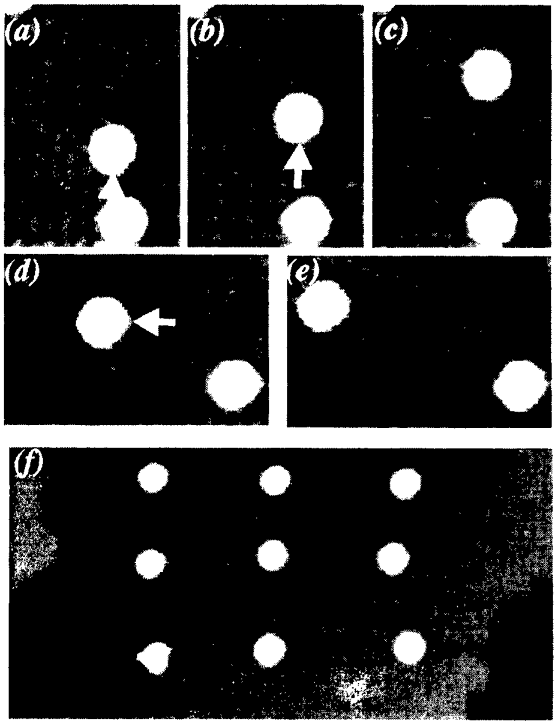


Figure 1. Demonstration of molecular manipulation: (*a*)–(*c*) parallel to the Si(100)-(2 × 1) dimer rows, scan area 7.5 × 10 nm^2; (*d*), (*e*) across dimer rows, scan area 10 × 6.3 nm^2. The arrows indicate the direction of tip displacement. (*f*) A 3 × 3 array of C_{60} molecules, scan area 40.5 × 19.7 nm^2. For (*a*)–(*f*): scan parameters, −3.5 V, 0.1 nA; manipulation parameters, −1.0 V, 1.0 nA.

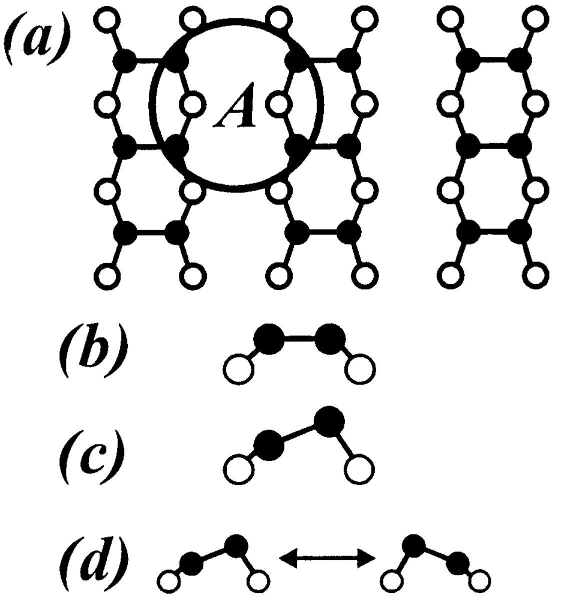


Figure 2. (*a*) Schematic of the Si(100)-(2 × 1) surface and the adsorption site for C_{60} labelled A; (*b*) symmetric and (*c*) buckled Si dimers; (*d*) hopping between buckled states which occurs at 300 K. ●, top layer Si atoms, ○, second layer Si atoms.

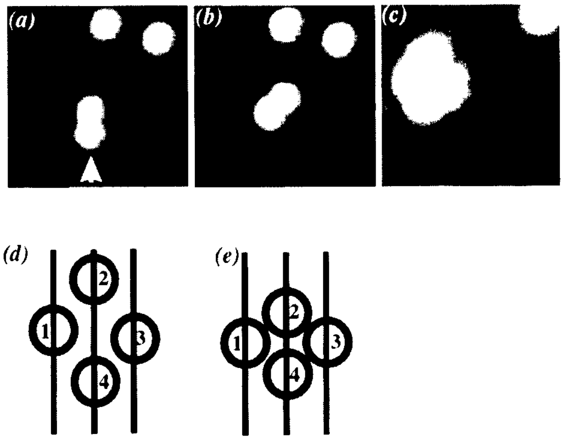


Figure 3. (*a*) A pair of C_{60} molecules (d = 1.15 nm) prior to manipulation. The result of an attempt to reduce the intermolecular separation by moving the lower molecule towards the upper molecule is shown in (*b*). The arrow in (*a*) indicates the direction of tip displacement. Manipulation parameters: −1.0 V, 1.5 nA. (*c*) A molecular cluster assembled using the STM tip. (*d*) Schematic diagram showing the position of the molecules in (*c*): ——, the minima of the troughs. (*e*) c(4 × 3) arrangement. According to a recent theory the cluster shown in (*c*) and (*d*) should spontaneously relax to the configuration in (*e*); however, this is not observed.

of 0.6 nm). This movement may be chosen to be either parallel or perpendicular to the dimer rows. A digital feedback loop is used to update the tunnel current after each step. The gap resistance is then increased to ~30–40 GΩ, causing the tip to withdraw from the surface, and the tip is returned to its original position. For the Si(100)-(2 × 1) surface we have not observed any difference in response to manipulation for forward and reverse bias.

For the Si(100)-(2 × 1) surface we observe a highly anisotropic response of C_{60} to STM manipulation. This is shown in figure 1 in which the response of C_{60} to manipulation either parallel (figures 1(*a*)–1(*c*)) or perpendicular (figures 1(*d*) and 1(*e*)) to the dimer rows is illustrated. In figures 1(*a*)–1(*c*) the effect of two successive 3 nm displacements of the tip is shown. For the manipulation perpendicular to the rows (figures 1(*d*) and 1(*e*)) five parallel lines separated by 6 Å were swept out forming an area of 3 × 3 nm^2. As shown in figure 1 it is possible to induce manipulation perpendicular to the rows although the success rate for this direction is ~15% as compared with ~95% for manipulation parallel to the dimer rows. Note also that following attempts at manipulation across dimer rows (figures 1(*d*) and 1(*e*)) we observe some displacement parallel to the dimer row. However, for displacement parallel to the rows the molecule is guided along the troughs on the Si(100)-(2 × 1) surface. This facilitates more precise and controllable placement. Figure 1(*f*) shows a simple pattern. We have undertaken further manipulation experiments to determine whether the manipulation results from a repulsive or attractive interaction. After application of a modified procedure, in which the tip is not retracted before returning to its original position, the C_{60} remains at the extreme of the tip excursion. This implies that the manipulation in figure 1 results from a repulsive interaction.

A model for the interaction of C_{60} with the Si(100)-(2 × 1) surface has recently been proposed [8]. According to this model a strong molecule–surface interaction causes the adsorption of C_{60} in the troughs between dimer rows. While the interaction with the surface causes the molecule to sit in troughs, the position of the molecule along the trough, which is determined by the component of forces parallel to the dimer rows, is assumed to be determined entirely by interactions with other adsorbed molecules. The intermolecular interaction is taken to be van der Waals [12] in origin, and the molecules are assumed to be physisorbed. Note that the equilibrium intermolecular separation according to this model is therefore given by $d_0 = 1.005$ nm. We have used STM manipulation of pairs and small groups of molecules to evaluate this model. Figures 3(*a*) and 3(*b*) show the result of an attempt to reduce the separation of a C_{60} pair by moving the lower molecule along the trough towards the upper molecule. The initial position of the tip was chosen so that it was sufficiently far away from the upper molecule so that no manipulation would result in the absence of the lower molecule, that is there should be no direct interaction between tip and upper molecule. Prior to manipulation the separation, d, of the C_{60} pair is 1.15 ± 0.02 nm (equal to $3a$, where a (= 0.384 nm) is the surface lattice constant). Figure 3(*b*) shows an STM image taken after application of the manipulation procedure and shows that the upper molecule has been forced to move across a dimer row into an adjacent trough. The final molecular separation is 1.09 ± 0.02 nm. Note also that both molecules move 'up' the image by ~0.8 nm.

We have repeated this experiment many times and the important aspects of the results shown in figure 3 are reproducible. In particular we have never observed a separation of molecules along a trough which is less than 1.15 ± 0.02 nm, a result which is inconsistent with [8] in which a value equal to the intermolecular separation d_0 is predicted. We have also assembled molecular clusters. According to the model described above the cluster in figure 3(*c*) should relax to an ordered (local c(4 × 3)) arrangement, but this is inconsistent with our data (see figures 3(*d*) and 3(*e*)).

There are two possible explanations for these differences. First, the molecules are not physisorbed, but chemisorbed, and the resulting charge exchange between adsorbate and substrate modifies the intermolecular potential leading to a different equilibrium separation. A second (and on the basis of our STM results more likely) possibility is that the positions of the molecules along the trough are determined by the molecule–surface interaction rather than intermolecular forces.

Both explanations imply a much stronger interaction between C_{60} and Si(100)-(2 × 1) than is consistent with physisorption. This is confirmed by photoemission spectra from the Si(100)-(2 × 1) surface before and after C_{60} deposition. Figure 4(*a*) shows the Si 2p core-level spectrum acquired from the clean surface. The spectrum is decomposed into a number of spin–orbit split Voight components and is in agreement with previous work [13]. The various components arise from Si atoms in different bonding configurations or chemical environments. The most intense peak (B) corresponds to atoms in a bulk configuration, peak C is related to second layer atoms and peak D has been identified as an electron energy loss feature [14, 15]. We focus on peak A, which arises from the 'up' atom of a buckled dimer.

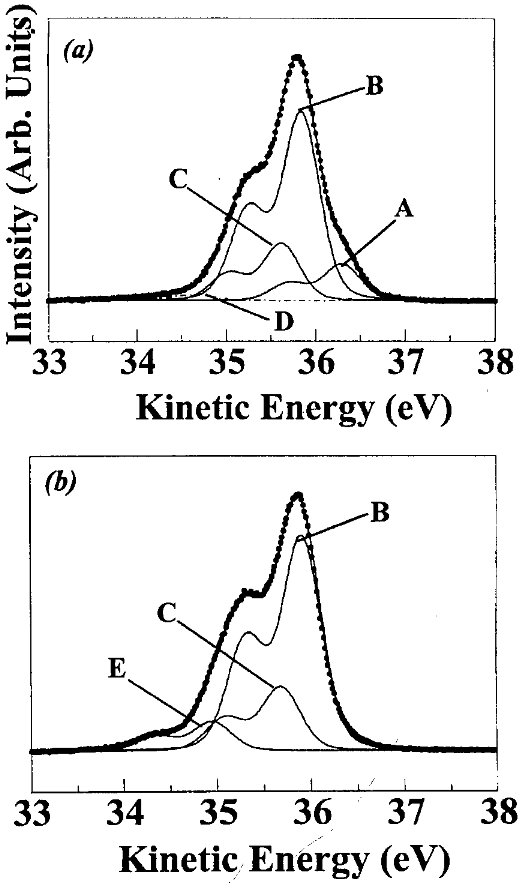


Figure 4. Si 2p core-level photoelectron spectrum from (*a*) the clean Si(100)-(2 × 1) surface and (*b*) Si(100)-(2 × 1) after deposition of ~0.2 monolayer C_{60}. The photon energy used in each case was 140 eV.

Buckling of dimers on the Si(100)-(2 × 1) is illustrated in figures 2(*b*)–2(*d*). A symmetric dimer (figure 2(*b*)) may adopt a lower energy configuration by buckling, that is raising one and lowering the other atom (figure 2(*c*)). The energy barrier for thermal activation between the two buckled configurations is low and transitions occur on a time scale less than the acquisition time of an STM image pixel [16]. The STM image is therefore an average of the two configurations and the dimers have a symmetric appearance, apart from a small number which are pinned by defects. However, the effects of buckling may be detected in photoemission which probes the surface on a much shorter time scale. Peak A is related to the 'up' atom of the buckled dimer [15].

Deposition of 0.2 monolayers of C_{60} induces distinct changes in the Si 2p spectrum (see figure 4(*b*)). Note that for physisorption we would expect no significant differences in the Si 2p spectrum following C_{60} deposition. Peak A is no longer present and a new peak (E) is resolved at lower energy. This shows that the bonding and net charge on the surface Si atoms have been modified by C_{60}, providing direct evidence for chemisorption of C_{60}. Our results may be explained in terms either of charge transfer into the lowest unoccupied molecular orbital of C_{60} or the formation of Si–C bonds. We believe that the latter explanation is unlikely as the core-level shift induced by C_{60} deposition (+0.9 eV) is much greater than expected for Si–C bond formation [17].

We have exploited an anisotropic surface reconstruction to investigate the interactions experienced by an adsorbed molecule on a solid surface. Our results imply that C_{60} is chemisorbed on Si(100) and this is confirmed by photoemission studies.

Acknowledgments

We thank the UK Engineering and Physical Sciences Research Council for financial support and Dr A A Cafolla for the software for analysis of core-level spectra. We are very grateful to Dr R Jones (University of Exeter) for helpful comments and discussions.

References

[1] Eigler D M and Schweizer E K 1990 *Nature* **344** 524
[2] Crommie M F, Lutz C P and Eigler D 1993 *Science* **262** 218
[3] Meyer G, Zöphel S and Rieder K-H 1996 *Phys. Rev. Lett.* **77** 2113
Meyer G, Zöphel S and Rieder K-H 1996 *Appl. Phys. Lett.* **69** 3185
[4] Beton P H, Dunn A W and Moriarty P 1995 *Appl. Phys. Lett.* **67** 1075
[5] Dunn A W, Beton P H and Moriarty P 1996 *J. Vac. Sci. Technol.* B **14** 943
[6] Jung T A, Schlittler R R, Gimzewski J K, Tang H and Joachim C 1996 *Science* **271** 181
[7] Cuberes M T, Schlittler R R and Gimzewski J K 1996 *Appl. Phys. Lett.* **69** 3016
[8] Klyachko D and Chen D M 1995 *Phys. Rev. Lett.* **75** 3693
[9] Chen D and Sarid D 1994 *Surf. Sci.* **318** 74
[10] Wang X-D, Hashizume T, Shinohara H, Saito Y, Nishina Y and Sakurai T 1993 *Phys. Rev.* B **47** 15 923
[11] Tromp R M, Hamers R J and Demuth J E 1985 *Phys. Rev. Lett.* **55** 1303
[12] Girifalco L A 1992 *J. Phys. Chem.* **96** 858
Girard Ch, Lambin Ph, Dereux A and Lucas A A 1994 *Phys. Rev.* B **49** 11 425
[13] Stober J, Eisenhut B, Rangelova G and Fauster Th 1994 *Surf. Sci.* **321** 111
Dufour G, Rochet F, Roulet H and Sirotti F 1994 *Surf. Sci.* **304** 33
[14] Farell H H, Stucki F, Anderson J, Frankel D J, Lapeyre G L and Levinson M 1984 *Phys. Rev.* B **30** 721
[15] Landemark E, Karlsson C J, Chao Y-C and Uhrberg R I G 1992 *Phys. Rev. Lett.* **69** 1588
[16] Wolkow R A 1992 *Phys. Rev. Lett.* **68** 2636
[17] Simons J K, Frigo S P, Taylor J W and Rosenberg R A 1996 *Surf. Sci.* **346** 21

Semicond. Sci. Technol. **13** (1998) A51–A54. Printed in the UK PII: S0268-1242(98)91854-5

Fullerene-derived molecular electronic devices

M Menon†, D Srivastava‡ and S Saini‡

† Department of Physics and Astronomy, University of Kentucky, Lexington, KY 40506, USA
‡ NASA Ames Research Center, MRJ, Mail Stop T27-A1, Moffett Field, CA 94035-1000, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. The carbon nanotube junctions have recently emerged as excellent candidates for use as the building blocks in the formation of nanoscale electronic devices. While the simple joint of two dissimilar tubes can be generated by the introduction of a pair of heptagon–pentagon defects in an otherwise perfect hexagonal graphene sheet, more complex joints require other mechanisms. In this work we explore structural and electronic properties of complex three-point junctions of carbon nanotubes using a generalized tight-binding molecular-dynamics scheme.

1. Introduction

The interest in pure carbon clusters received a tremendous boost from the discovery of the C_{60} molecule [1]. Following this discovery, quasi-one-dimensional structures made of carbon atoms, called carbon nanotubes, were experimentally observed by Iijima [2]. The carbon nanotubes consist of rolled-up graphene sheet with various chiralities. The electronic structure of these tubes can be either metallic or semiconducting, depending on both the diameter and the chirality which can be uniquely determined by the chiral vector (n, m), where n and m are integers [3–7].

The possibility of connecting nanotubes of different diameters and chiralities has generated considerable interest recently [8–12]. This is because of the possibility of the junctions being the building blocks of nanoscale electronic devices. The simplest way to connect two dissimilar nanotubes is found to be via the introduction of a pair of heptagon and pentagon in an otherwise perfect hexagonal graphene sheet [6]. The resulting structure still contains three-fold coordination for all carbon atoms forming the junction.

In reality nanotubes have finite lengths and, in most cases, tend to be closed with fullerene caps. The closure introduces a small gap in the electronic structures of these tubes. Also, formation of a tube junction results in considerable local strain which is relieved by the relaxation of the atoms. The relaxation is also expected to alter the electronic structure and local density of states of the atoms forming the junction and their neighbours. Although several theoretical models have been used in the study of nanotube heterojunctions, most of them tend to ignore the effects of relaxation altogether.

In this work we investigate the effects of full symmetry unconstrained relaxation of the nanotube heterojunctions on the structural and electronic properties using the generalized tight-binding molecular-dynamic (GTBMD) scheme of Menon *et al* [7]. The GTBMD makes explicit use of the non-orthogonality of the orbitals in treating interactions in covalent systems and has been found to be very reliable in obtaining good agreement with experimental and local density approximation results for the structural and vibrational properties of fullerenes and nanotubes [7, 13]. Additionally, GTBMD has been applied earlier to obtain equilibrium geometries for small carbon clusters [14], in good agreement with *ab initio* [15] results for the lowest-energy structures of carbon clusters of size up to $N = 10$ (for which *ab initio* results are available).

2. Three-point junctions

2.1. 'T-junctions'

Unlike the simple nanotube junctions, the three-point junctions of single-walled carbon nanotubes could be used as building blocks of nanoscale tunnel junctions in a 2D network of nanoelectronic devices. As a prototype of such junctions we study a 'T-junction' formed by fusing two nanotubes of different diameters and chiralities perpendicular to each other [16].

The T-junctions provide a challenge to the conventional rules applicable to tube bends. This is because, unlike the knee joint where one can clearly define the opposite sides of the joint as either the front or the behind, both sides are topologically equivalent. As a result, we can expect a net excess of heptagons over the pentagons at the junction. Furthermore, whereas the bend angles at a two-point tube junction depend on the tube parameters

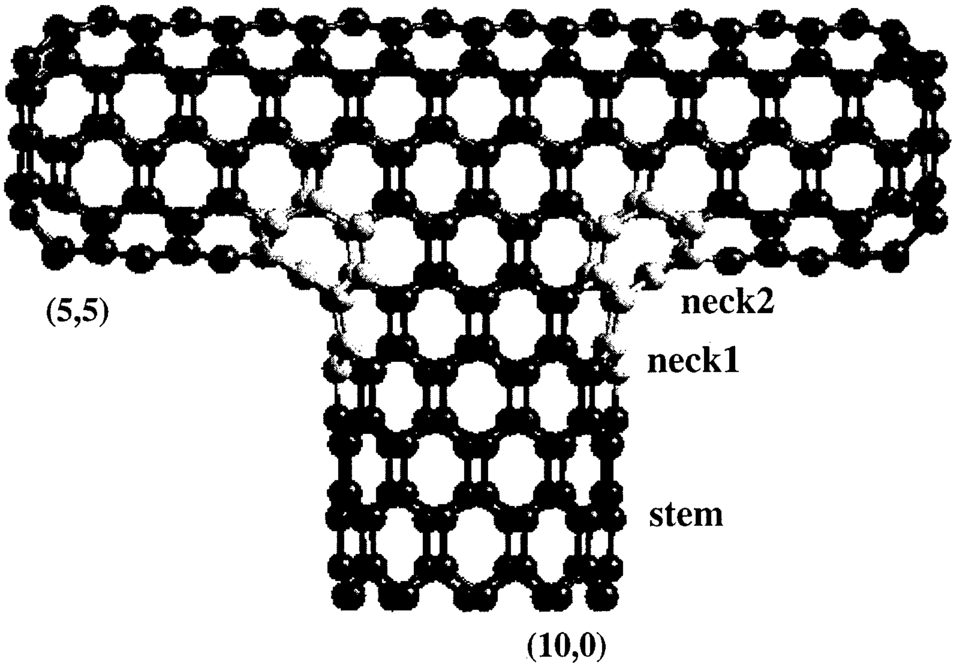


Figure 1. Fully relaxed (5,5)–(10,0)–(5,5) tube (T1). The structure contains six heptagons and no pentagons.

Table 1. Bond length analysis for the T-junction in figure 1. The GTBMD scheme has been used to relax the cluster. The bond lengths are averaged over the heptagons and hexagons.

Tube	Average bond length in heptagons (Å)	Average bond length in hexagons (Å)
T1	1.427	1.419

(diameter and chirality) of both the component tubes, no such dependence exists for the T-junction where the angle remains fixed at 90°.

We explore an alternative route to the formation of T-junctions that is not constrained by the usual heptagon–pentagon defect pair considerations. In particular, we examine a metal–semiconductor–metal T-junction, namely the (5,5)–(10,0)–(5,5) junction (figure 1). We designate this junction by the symbol T1. The structure T1 is composed of 314 atoms. The numbers chosen are sufficiently large to avoid the effects of the dangling π bonds at the edges on the junction. The (10,0) tube is semiconducting and the (5,5) tube is a semimetal. The (5,5) and (10,0) tubes have armchair and zig-zag configurations, respectively. In the armchair (zig-zag) configuration the tubes have C–C bonds perpendicular (parallel) to their axis. As seen in figure 1, in going across the junction in T1 from the (10,0) side to the (5,5) side, the orientation of the C–C bonds remains unchanged. Interestingly, T1 contains six heptagons and no pentagons at the junction (figure 1).

The geometry shown in figure 1 is the GTBMD-optimized structure for T1. The starting configuration of T1 has two-fold coordinated atoms at the ends of the armchair portion of the tubule within strong bonding interactions of each other. The GTBMD relaxation results in the closure seen in figure 1 on relaxation with three-fold coordination for all atoms at the armchair ends. No such closure results at the zig-zag ends, however, as the two-fold coordinated atoms are sufficiently far from each other. Table 1 summarizes the average bond lengths in the heptagons and hexagons at the junctions of T1. The average bond length in the heptagons is found to be longer than the one for the hexagons.

The Fermi level of the (5,5) nanotube lies within the gap of the (10,0) semiconducting tube. The T-junction forms a microscopic tunnel junction, made up entirely of carbon atoms, through which electrons can cross by quantum mechanical tunnelling. The tunnelling current can be controlled by an application of a potential difference that raises the chemical potential of one side with respect to the other. Since the tunnelling currents have been observed to obey Ohm's law, the T-junctions can thus form one of the smallest microscopic ohmic resistors. Furthermore, either *n*-type or *p*-type doping of the semiconducting portion of the T-junction should yield Schottky barrier type devices.

2.2. 'Y-junctions'

The Y-junctions pose a different kind of challenge when three-point junctions are formed [18]. As found in the case of T-junctions, the pentagon–heptagon defect pair rule is also found to be not applicable in the formation of the Y-junctions. We, therefore, explore the formation of large-angle bends through octagon–pentagon defect pairs and apply it to study Y-junctions which provide another route to the formation of a three-point junction of dissimilar tubes.

We show that the Y-junctions can be made with the incorporation of pentagons and octagons with no

Table 2. Bond length analysis for the Y-junction in figure 2. The GTBMD scheme has been used to relax the cluster. The bond lengths are averaged over the pentagons, octagons and hexagons.

Tube	Average bond length in pentagons (Å)	Average bond length in octagons (Å)	Average bond length in hexagons (Å)
Y	1.440	1.422	1.417

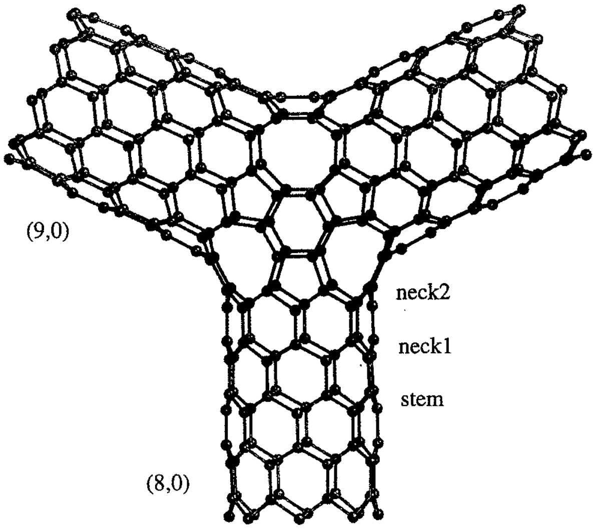


Figure 2. Fully relaxed Y-junction with two (9,0) arms and one (8,0) arm. The junction is formed by six octagons and six pentagons. The atoms forming pentagons and octagons are shown in dark colour.

assistance from heptagons and that incorporation of octagons in place of heptagons gives greater flexibility in the Y-junction formation involving similar or dissimilar nanotubes. Additionally, we find that the number of octagonal defects equals the number of pentagonal defects in the junction region, regardless of whether or not there is a change in chirality when crossing the junction. In particular, we examine a Y-junction made up of nanotubes in zig-zag configurations. The joint investigated consists of a Y-junction formed by two (9,0) nanotubes and one (8,0) nanotube.

The structure was fully optimized without any symmetry constraints using the GTBMD scheme.

All three arms of the structure have zig-zag configurations, but the (9,0) arms are metallic, while the (8,0) arm is semi-conducting. Thus, this structure forms a nanoscale metal–semiconductor–metal tunnel junction. The relaxed structure is shown in figure 2 and contains 272 atoms. The number of atoms chosen is sufficiently large to avoid the effects of the dangling π bonds at the edges on the junction. On relaxation, the structure was found to be stable with six octagons, six pentagons and, interestingly, no heptagons. The brunt of the curvature is borne by octagons and the hexagons at the junction while the pentagons are almost planar. Also, as seen in figure 2, all the pentagons are clustered near the centre of the junction. The atomic arrangement is radically different from the case of simple tube bends where both pentagons and heptagons participate in the tube bend. Table 2 summarizes the average bond lengths in the pentagons, octagons and hexagons at the junctions of the structure. As seen in the table, the average bond length in the octagons is found to be longer than the one for the hexagons but shorter than those in the pentagons.

In the light of the present work, serious considerations must be given to the possibility of the existence of octagonal defects in complex junctions. The octagonal defects may give greater flexibility for such junction formation involving similar or dissimilar nanotubes. The search for larger-angle bends or Y-junction formation in the existing experimental data on single- and multiwall nanotubes may point towards the possibility of octagonal defects in the nanotube-based pure carbon structures.

3. Summary

The implications of these findings are intriguing. Both T-junctions and Y-junctions defy the conventional arguments made in favour of equal numbers of heptagon–pentagon defect pairs for the stability of dissimilar tube joints. We have shown that T-junction joints can be made without the incorporation of pentagons. Our results also show that Y-junctions can be created with an equal number of pentagons and octagons and, surprisingly, no heptagons. Furthermore, the T-junctions can be used as 'universal joints' for forming a 2D network of tubes in which conduction pathways can be controlled. If produced, these junctions could be the prototypes of nanoscale tunnel devices.

Acknowledgments

Part of this research (MM) was supported by NSF Grant OSR 94-52895 and by the University of Kentucky Center for Computational Sciences. GTBMD relaxation and electronic structure computations were performed at NAS computational facility (DS) at NASA Ames Research Center.

References

[1] Kroto H W, Heath J R, O'Brien S C, Curl R F and Smalley R E 1985 *Nature* **318** 162
[2] Iijima S 1991 *Nature* **354** 56
[3] Mintmire J W, Dunlap B I and White C T 1992 *Phys. Rev. Lett.* **68** 631
[4] Saito R, Fujita M, Dresselhaus G and Dresselhaus M S 1992 *Phys. Rev.* B **46** 1804
[5] Hamada N, Sawada S and Oshiyama A 1992 *Phys. Rev. Lett.* **68** 1579
[6] Dunlap B I 1994 *Phys. Rev.* B **49** 5643
[7] Menon M, Richter E and Subbaswamy K R 1996 *J. Chem. Phys.* **104** 5875

[8] Chico L, Crespi V H, Benedict L X, Louie S G and Cohen M L 1996 *Phys. Rev. Lett.* **76** 971

[9] Charlier J C, Ebbesen T W and Lambin Ph 1996 *Phys. Rev.* B **53** 11 108

[10] Saito R, Dresselhaus G and Dresselhaus M S 1996 *Phys. Rev.* B **53** 2044

[11] Iijima S, Ichihashi T and Ando Y 1992 *Nature* **356** 776

[12] Lambin Ph, Fonseca A, Vigneron J, Nagy J B and Lucas A A 1995 *Chem. Phys. Lett.* **245** 85

[13] Rao A M *et al* 1997 *Science* **275** 187

[14] Menon M, Subbaswamy K R and Sawtarie M 1993 *Phys. Rev.* B **48** 8398

[15] Raghavachari K and Binkley J S 1987 *J. Chem. Phys.* **87** 2191

[16] Menon M and Srivastava D 1997 *Phys. Rev. Lett.* **79** 4453

[17] Kroto H W and Walton D R M 1993 *Chem. Phys. Lett.* **214** 353

[18] Menon M and Srivastava D 1998 *J. Mater. Res.* in press

Semicond. Sci. Technol. **13** (1998) A55–A58. Printed in the UK PII: S0268-1242(98)91873-9

Fluctuations in quantum dot charging energy and polarization

M Stopa

RIKEN (The Institute of Physical and Chemical Research) 2-1, Hirosawa,
Wako-Shi Saitama 351-01, Japan

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We discuss density functional and spin density functional calculations of semiconductor quantum dots. We show that spin-degenerate density functional theory tends to exhibit level clustering at the Fermi surface E_F owing to Coulombic energetic costs associated with occupying one spatial state fully as opposed to several partially. Within spin density functional calculations, this level clustering is replaced by a self-consistent tendency for levels to split, an effect which is enhanced by the exchange interaction. Consequently we predict that, for intermediate size dots with electron number $N \sim 100$, spontaneous spin polarization will occur. In addition, we find that the polarization fluctuates coincidentally with the charging energy E_C and the pattern of fluctuations is driven, in the case of chaotic quantum dots, by the presence of 'scars' in the eigenstates. The scarred states in small dots occupy a reduced area of the dot, Coulombically interact more strongly between spin up and down and so require more energy to double fill than more chaotic states. Both in the chaotic case and in the case of more regular dots we find that the polarization has a distinct fluctuating pattern which, in particular, shows polarization collapse in the vicinity of states with large diagonal Coulomb interaction.

The existence of 'scars' in the eigenstate intensity pattern of noninteracting, chaotic Hamiltonians has been recognized for some years [1]. Phenomenologically, scars emerge in, for example, the eigenstates of the stadium as narrow linear regions with enhanced intensity which appear to recapitulate periodic classical trajectories [2]. Although they are not fully understood, a heuristic explanation can be given in terms of a Gaussian wavepacket, launched along an unstable, periodic trajectory. The packet returns to the origin with frequency ω and a diminished amplitude characterized by the Lyanpounoff exponent λ of the classical trajectory [3]. The wavefunction so formed is thus not an exact energy eigenstate, but can be considered as a superposition of such eigenstates ϕ_n. As noted by Gutzwiller [3], a privileged set of eigenstates shows a marked preference for a given periodic orbit and the privilege is more exclusive as ω/λ increases.

We have recently shown that relatively small (electron number $N \sim 100$) quantum dot eigenstates, *computed self-consistently* within the density functional (DF) and spin density functional (SDF) theories, exhibit scars [4]. In these calculations, which are meant to model realistic GaAs–AlGaAs quantum dot structures, we restrict our consideration to eigenstates up to the Fermi surface or slightly above. Typically, in SDF theory, we calculate 100 states (of each spin) for a dot containing 120 electrons. The wavefunction in the z (growth) direction is generally taken as the lowest subband eigenfunction. In this paper we will discuss the effective two-dimensional (2D) eigenstates although it must be remembered that the calculation is fully 3D. We discuss the results from devices with two different gate patterns in this paper. The first is that employed by Sivan *et al* [5] in a recent publication. As noted in [4], the gate pattern is rectangular but, because of excess metal in one of the gates, the confining potential is very irregular and the dot is in the chaotic regime. This conclusion is substantiated by the distribution of level spacings which is close to the Wigner surmise showing substantial level repulsion.

The second gate pattern is a square and is intended to produce as symmetric a confining potential as possible. For this purpose, in this second case, we vary a backgate to change the electron number in the dot (thereby avoiding any change of dot shape as gate voltage V_g is varied). While lithographically square, the resultant potential at the 2DEG level is practically, but not perfectly, circular. A similar gate pattern (but using surface gates as 'plungers') was employed in extensive calculations discussed in [6]. The eigenfunctions exhibit the regularity expected from a nearly azimuthally symmetric potential. We are principally concerned in this paper with the chaotic eigenfunctions and will only use this symmetric device case for comparison. Thus the subsequent discussion of the eigenstates will refer to the first gate pattern unless otherwise specified.

The ratio of the Fermi wavelength to the dot size λ_F/L is of order 1/10. As a result we find that the unstable periodic orbits which show up as scars are only the simplest trajectories, often a simple line from one corner of the dot to another. The spacing, in energy, of the eigenstates is such that scars tend to be represented only by single eigenstates.

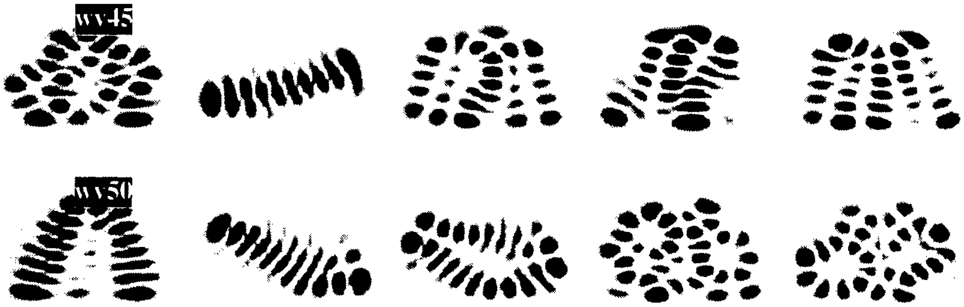


Figure 1. Series of eigenstates (moduli squared) for the dot with gate pattern from Sivan *et al* [5]. State index runs from 45, upper left, to 54, lower right. States $p = 46$, 50 and 51 are 'scarred' states which show large diagonal Coulomb matrix elements.

That is, in this regime, we find that the semiclassical approximation of taking a Gaussian wavepacket and setting it in motion along a particular classical trajectory produces a family of states, depending on initial momentum, each member of which is very close to an energy eigenfunction of the dot. Self-consistency, which is treated in a mean-field manner in DF and SDF theories, does not destroy the scarred nature of the states.

The remainder of the states (i.e. those that are not strongly scarred) appear to occupy the entire dot area in a more random manner. The result is that the spectrum can be decomposed into the scarred states, which are essentially one dimensional and follow specific periodic trajectories in the dot, and what we will simply call 'chaotic states', which tend to fill the full dot area. Figure 1 shows a sequence of states (moduli squared), ordered in energy, for a gate voltage such that $N \approx 100$. The eigenstates are negligibly affected by whether DF or SDF calculations are employed. In the figures, states $p = 46$, 50 and 51 are clear scars, as compared with, say, state $p = 47$, which occupies most of the dot area. We find that, as gate voltage and electron number are varied, the shape of the eigenstates only changes appreciably where anticrossing occurs. However, such anticrossings are common near the Fermi surface and the indicial location of the states (i.e. the energy ordering) changes frequently. In particular, the state 46 in figure 1 is at the Fermi surface when $N = 96$ (see discussion below) and is thus the 48th state at that point.

In our preceding paper [4] we discussed the import of the scar–chaotic distinction for the charging energy of quantum dots and the departure of its fluctuations from the predictions of the so-called 'random matrix theory plus constant interaction' model [5]. In particular we showed that the direct Coulomb matrix element between any state and itself W_{pp} (physically between a spin-up and a spin-down electron in the same spatial state) is uniformly greater than that between two arbitrary spatial states W_{pq}. Further, this 'diagonal' matrix element is particularly large for strongly scarred states owing to the compressed (one-dimensional) nature of the wavefunctions. The result is that large fluctuations in the charging energy occur when filling of the second spin state of a scarred wavefunction is forced.

Here we wish to look more closely at the relationship between the DF calculation, which is intrinsically spin degenerate, and the SDF calculation, which takes into account the difference in the exchange energy of electrons of different spin based on the local polarization

$$m(r) \equiv \frac{n_\uparrow(r) - n_\downarrow(r)}{n_\uparrow(r) + n_\downarrow(r)} \tag{1}$$

where $n_\uparrow(r)$ ($n_\downarrow(r)$) is the local density of spin-up (spin-down) electrons at r. In SDF theory the exchange-correlation potential differs between spin-up and spin-down electrons. Specifically,

$$\begin{aligned} V^{\sigma}_{XC}(r_s, m) = {} & V^{U}_{XC}(r_s) + f(m)[V^{P}_{XC}(r_s) - V^{U}_{XC}(r_s)] \\ & + [\epsilon^{P}_{XC}(r_s) - \epsilon^{U}_{XC}(r_s)][\mathrm{sign}(\sigma) - m]f'(m) \end{aligned} \tag{2}$$

where

$$f(m) = \frac{(1+m)^{4/3} + (1-m)^{4/3-2}}{2^{4/3} - 2} \tag{3}$$

with m given by equation (1). In equation (2), V^{U}_{XC} ($V^{P}_{XC}(r_s)$) is the exchange-correlation potential for a completely unpolarized (polarized) system of uniform density r_s, σ is the spin and $f(m)$ is an empirical interpolation function [7]. The crucial point to note is that the exchange term favours spin polarization (as in Hund's rule for atomic systems). Thus, if the dot is polarized spin up, then spin-up electrons are further lowered in energy by the difference in V_{XC} relative to spin-down electrons.

Figure 2 shows the spectrum (Kohn–Sham energies) for a small range of N for both the DF and the SDF cases. Here, the electrochemical potential of the dot is set to that of the leads and is the energy zero. As discussed in [6] and [4], we fill states according to a Fermi function and so N is actually a derived quantity and need not be an integer. Obviously the only physically meaningful states will be those where N is an integer (since the quantum point contacts are set so as to place the dot in the tunnelling regime). Nonetheless, the continuous variation of N is instructive.

In the DF case, there is a clear tendency for the states to cluster at the Fermi level. This results from the energetic benefit of spreading the added charge as smoothly as possible across the dot. In particular, double occupation of the same spatial eigenstate is very costly and double occupation of a scarred state is especially costly. In both the DF and the SDF results there is a gap at $N = 96$ which relates to a large fluctuation of the charging energy

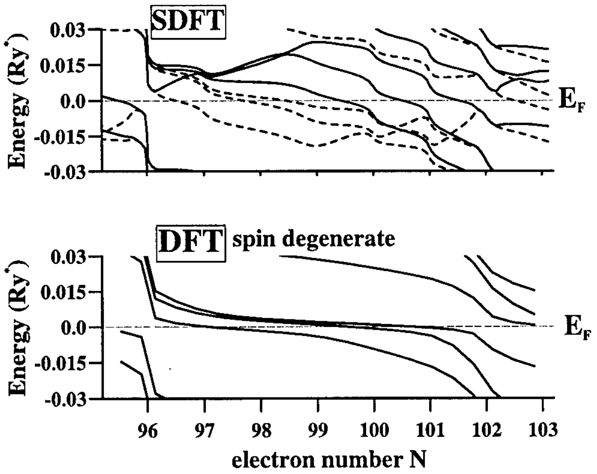


Figure 2. Close-up of spectrum near E_F (depth of Fermi sea is ≈1.2 Ry^*) for spin-degenerate DF calculation and SDF calculation. Level clustering at E_F in the former case is replaced by self-consistently driven spin polarization in the latter case. Two fluctuations in the charging energy (not shown) occur at $N = 96$ and $N = 102$, corresponding to double filling of strongly scarred states. Also, spin polarization collapses at these values.

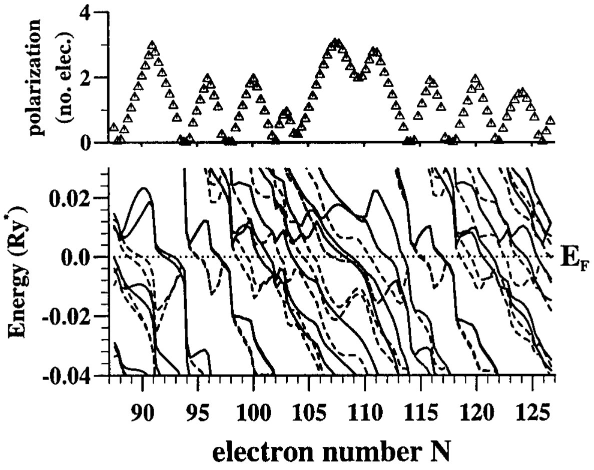


Figure 3. Spectrum over larger N range for a symmetric dot with corresponding total polarization. The lower Fermi level density of states near $N = 97$ and 117 is related to residual band structure of approximately parabolic potential.

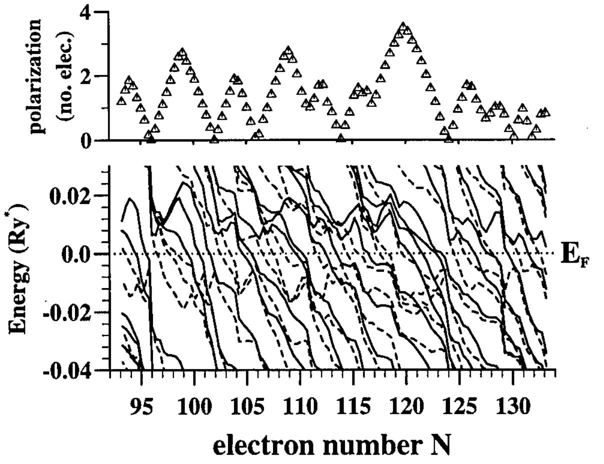


Figure 4. Spectrum and polarization of chaotic dot. Note that while spontaneous polarization due to the unfavourable energy condition of double occupancy is to be expected, the tendency for the polarization to return periodically to zero is unexpected.

there. As noted above, this is related to the scar state 46 in figure 1(a). Note in SDF results that all dot states are spin degenerate at $N = 96$. Thus filling of the second spin state of a scar only occurs when all other (spatial) states are either doubly filled or completely empty. This pattern is more fully exhibited between $N = 97$ and $N = 102$. Taking the broken lines as the 'up' spins, filling of $49_\uparrow$ pushes up the energy of $49_\downarrow$. The resulting nonzero polarization creates a difference in the exchange energy experienced by up and down electrons and thereby causes all states to spin split. Consequently the next state to go below E_F is again a spin-up state, which further increases the splitting. Finally $p = 51_\uparrow$, which is a strongly scarred state, fills. At this point the total dot polarization is maximum at 3 electrons.

Generally, strongly scarred states tend to produce a gap in the spectrum since they have a highly inhomogeneous spatial distribution. Empirically we find that the polarization tends to peak at the filling of the first spin state of a scar and drops to zero when it is the last polarized state to double fill. This occurs in figure 2 at $N = 102$. The $p = 51_\uparrow$ state undergoes anticrossings at $N \approx 100.5$ and $N \approx 101.4$, and rejoins its spin partner at $N = 102$, where the spin polarization goes to zero and all states become spin degenerate.

Clearly, the extent to which quantum dots will spin polarize depends on the excess of diagonal Coulomb matrix elements over off-diagonal elements (i.e. the cost to double occupy a state) as compared with the average bare level spacing Δ. Since Δ scales as L^{-2} and the Coulomb matrix elements scale as L^{-1}, one expects that for small dots spin polarization will be suppressed. This, however, does not take into account the regularity of the spectrum of small dots (due to, for example, approximately biparabolic confinement) which can produce level degeneracies. Also, while the typical Coulomb interaction goes as $1/L$, it is not clear that the excess of diagonal (same spatial state) over off-diagonal interactions also scales in the same way, particularly insofar as screening is likely to play an important role.

In figures 3 and 4 we show the spectra and net polarization P over a larger range of N for both symmetric and chaotic dots respectively (the two gate patterns mentioned above). The spectrum of the symmetric dot is characterized by large gaps near $N = 97$ and 117 which are the remnant of the degeneracy associated with a 2D parabolic potential. Several of the states in figure 3 have a parity-related degeneracy. The spectrum in figure 4 is, by contrast, much more regular (no large gaps). Nonetheless, the patterns of the evolving polarization are remarkably similar.

Note that it is by no means *a priori* obvious why the polarization should drop periodically to zero. One might expect, given a typical average difference between diagonal and off-diagonal Coulomb matrix elements and a typical level spacing, that some roughly constant, non-zero polarization would be favoured (say 1 or 2 electrons) and the plot of P versus N would show random fluctuations up and down from that average value. That polarization periodically collapses, a fact which above we have related to the unusually large diagonal matrix element of scarred states, still requires, we believe, a more elaborate explanation. In particular, the same pattern is found in the symmetric dot despite the fact that the qualitative distinction between scarred and chaotic states cannot be made and the eigenstates are much more regular.

In conclusion, we have examined the evolution of self-consistently induced, spontaneous polarization in a quantum dot as a function of electron number. We have shown that the filling of states in the chaotic dot is largely regulated by strongly scarred states which, by virtue of their quasi-1D nature, have large diagonal direct Coulomb matrix elements and are difficult to double fill. We have shown that the total polarization fluctuates similarly in both the chaotic and the symmetric dot cases despite the difference in the nature of the spectra. Finally, we have noted that a periodic collapse of polarization, observed in both gate pattern devices, is an unexpected result requiring further investigation.

Acknowledgments

I wish to thank Dr Uri Sivan for providing me with details of his devices and measurement. I have also had helpful conversations with O Agam, Y Alhassid and coworkers, S Das Sarma, F Stern, C Marcus, S Patel, N Wingreen Y Aoyagi and T Sugano. Support from Riken Computing Centre–Fujitsu VPP500 Supercomputer is gratefully acknowledged. This work supported by grants from the Japan Ministries of Education and Science and Technology.

References

[1] Heller E J 1987 *Phys. Rev.* A **35** 1360
[2] Heller E J, O'Connor P W and Gehlen J 1989 *Phys. Scr.* **40** 354
[3] Gutzwiller M C 1990 *Chaos in Classical and Quantum Mechanics* (New York: Springer)
[4] Stopa M, unpublished cond-mat/9709119
[5] Sivan U, Berkovits R, Aloni Y, Prus O, Auerbach A and Ben-Yoseph G 1996 *Phys. Rev. Lett.* **77** 1123
[6] Stopa M 1996 *Phys. Rev.* B **54** 13767
Stopa M 1996 *Phys. Rev.* B **77** 1123
[7] Radtke R J, Tamborenea P I and Das Sarma S 1996 *Phys. Rev.* B **54** 13832

Semicond. Sci. Technol. **13** (1998) A59–A62. Printed in the UK

PII: S0268-1242(98)91874-0

Intrinsic dopant correlations and transport properties of mesoscopic modulation-doped heterostructures

S Das Sarma and S Kodiyalam

Department of Physics, University of Maryland, College Park, MD 20742-4111, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. Calculations of *correlated* remote ion scattering limited mobility of modulation delta-doped GaAs–AlGaAs heterostructures are presented. The scatterers are modelled as a two-dimensional Coulomb gas whose distribution is determined directly through Monte Carlo simulations. Trends in the variation of the mobility with correlations in scatterer distributions (Coulomb gas temperature), density of the scatterers, density of electrons and the spacer width are discussed.

The maximum possible mobility of the two-dimensional electron gas (2DEG) at the interface of a modulation-doped GaAs–AlGaAs heterostructure has been the subject of both experimental [1] and theoretical interest [2–5]. In highly pure samples the mobility is limited by scattering from the ionized Si dopants (in the AlGaAs layer) that contribute electrons to the interface. It has been recognized that models of Coulomb scattering from these dopants, that assume a random distribution for them, underestimate the mobility [2]—some form of dopant correlation is needed to account for observed mobilities [1] of the order $\sim 10^7$ cm^2 V^{-1} s^{-1}. Previous attempts in this direction computed the ionized dopant correlations using thermodynamic arguments [4] and/or Monte Carlo (MC) simulations with various model interactions between dopant ions, for example, screened Coulomb [4, 6] or the hard sphere potential [5]. In the current work the spatial dopant distributions are determined through MC simulations using the long-range Coulomb interaction between the dopant ions. These distributions are then used to compute the Coulomb scattering limited mobility of the 2DEG. We find that in ultrapure samples it is possible to achieve mobilities which are substantially higher than the current values.

The standard expression for the mobility (μ) in terms of the transport lifetime (τ_t) (inverse momentum relaxation rate) is used [3] (the AlGaAs–GaAs interface is assumed to be the x–y plane with the AlGaAs in the region $z < 0$)

$$\mu = \frac{e}{m}\tau_t(q_i = k_F) \tag{1}$$

$$\tau_t^{-1}(\mathbf{q}_i) = \frac{2\pi}{\hbar}\sum_{\mathbf{q}_f \neq \mathbf{q}_i} \delta(E(\mathbf{q}_f) - E(\mathbf{q}_i))|\langle \mathbf{q}_f|V(\mathbf{r})|\mathbf{q}_i\rangle|^2 \times (1 - \cos\theta) \tag{2}$$

where k_F is the Fermi wavevector of the 2DEG ($= \sqrt{2\pi N_s}$; N_s being the 2DEG density), $V(\mathbf{r})$ is the effective scattering potential of the ionized dopants and θ is the angle between the incident ($\mathbf{q}_i$) and scattered ($\mathbf{q}_f$) electron wavevectors. Using the superposition of the potentials due to individual scatterers and that $E(\mathbf{q}) = \hbar^2 q^2/2m$, equation (2) may be rewritten as

$$\tau_t^{-1}(\mathbf{q}_i) = \frac{mN_i}{2\pi\hbar^3}\int_{0^+}^{2\pi^-} \mathrm{d}\theta\, s(\mathbf{q}_i - \mathbf{q}_f)|v_0(\mathbf{q}_i - \mathbf{q}_f)|^2 \qquad (q_i = q_f)$$

where $s(\mathbf{q})$ is the structure factor of the scatterer distribution:

$$s(\mathbf{q}) = \frac{1}{n_i}\left|\sum_{p=1}^{n_i} \mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{r}_p}\right|^2 = \tilde{s}(\mathbf{q}) + n_i\delta_{\mathbf{q}0}. \tag{3}$$

N_i (n_i) is the density (total number) of scatterers and $v_0(\mathbf{q})$ is the Fourier transform of the effective potential due to a single scatterer at $\mathbf{r} = 0$, $z = -W$, where W is the spacer width. It has been assumed that the thickness of the delta-doped layer is zero. This is a good approximation if the thickness is much smaller than W—as is expected to be the case in the best samples. $v_0(\mathbf{q})$, in the electric quantum limit together with screening by the 2DEG is given by [3]

$$v_0(\mathbf{q}) = v_0(q) = -\frac{2\pi e^2}{q\epsilon(q)}\mathrm{e}^{-Wq}\left(\frac{b}{b+q}\right)^3$$

where b is the variational parameter governing the z dependence of the wavefunction for an electron in the 2DEG and $\epsilon(q)$ is the static dielectric function of the 2DEG [3]:

$$b = \left(\frac{48\pi me^2(N_d + \frac{11}{32}N_s)}{\kappa\hbar^2}\right)^{1/3}$$

$$\epsilon(q) = \kappa\left\{1 + \frac{q_s}{q}\left[1 - u(k - 2k_F)\left(1 - \frac{4k_F^2}{q^2}\right)^{1/2}\right]\right\} \qquad (q_s = 2me^2/\kappa\hbar^2)$$

0268-1242/98/SA0059+04$19.50 © 1998 IOP Publishing Ltd

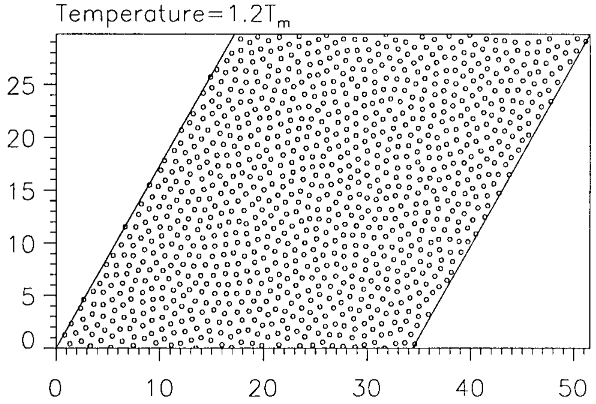


Figure 1. Ideal Coulomb gas configuration corresponding to a temperature of 1.2 × melting temperature (T_m). The system consists of 32^2 charges. Periodic boundary conditions replicate the parallelogram containing the system at integral multiples of its lattice constants. The axes are labelled in units of (particle density (N_i))$^{-1/2}$.

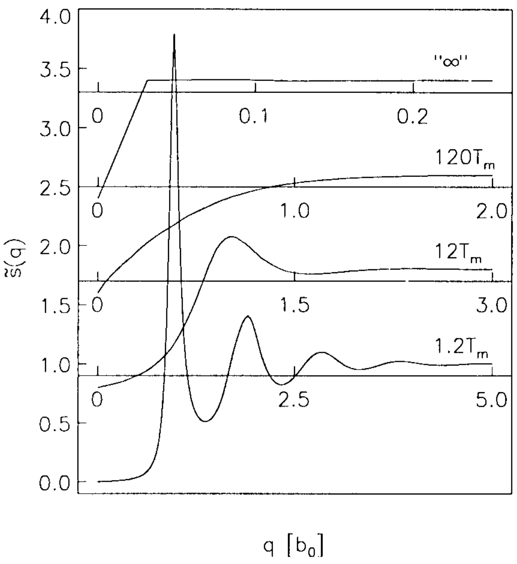


Figure 2. The curves show the structure factor $\tilde{s}$ of a system of particles interacting through the ideal Coulomb potential as a function of the wavevector magnitude q ($b_0 = 2\pi(N_i)^{-1/2}$). Dotted (full) curve(s) correspond to results for a scatterer system size of 42^2 (32^2). The corresponding temperature of the system of particles is marked in terms of the melting temperature of the ideal Coulomb lattice (T_m).

where $u(x) = 1(0)$ for $x > 0(<0)$. In this study, we set the depletion density N_d to zero, $m = 0.068m_e$ and $\kappa = 12.6$. Further *assuming* that the distribution of scatterers is rotationally invariant (i.e. $s(\mathbf{q}) = s(q)$) gives the following expression for $\tau_t^{-1}(\mathbf{q}_i)$:

$$\tau_t^{-1}(\mathbf{q}_i) = \tau_t^{-1}(q_i) = \frac{mN_i}{\pi\hbar^3}\int_0^{\pi} \mathrm{d}\theta[v_0(q)]^2\tilde{s}(q)(1-\cos\theta)$$
$$(q = 2q_i \sin(\theta/2)). \qquad (4)$$

Equations (1) and (4) are used to compute the mobility after determining $\tilde{s}(q)$ through MC simulations. For the maximum mobilities, all scattering, other than that by the remote dopants, is assumed negligible. We consider only the zero-temperature mobility here.

MC simulations were carried out with a system size of $n_i = m^2$ ($m = 32$ in most cases) particles (corresponding to ionized Si dopants) as illustrated in figure 1. These particles were allowed to occupy points on a uniform fine grid that had c^2 ($c = 42$ was used here) points per unit area ($= (N_i)^{-1/2}$). This discretization was adopted to enable the construction of a look-up table for the interaction potential between two particles. It also resulted in an upper bound on the wavevector values. The potential tabulated was the Ewald summed ideal Coulomb interaction [7], which was the result of applying periodic boundary conditions to the parallelogram in figure 1. This resulted in an exact six-fold symmetry together with additional reflection symmetry in the two-particle interaction that reduced the required size of the potential table by a factor of 12.

The simulations began with the particles in the ideal 2D crystal (triangular lattice) configuration. The Metropolis algorithm was used to evolve the system with attempts to move one particle after another. (Observed acceptance ratio of ≈0.65.) The temperatures considered were all beyond the melting temperature (T_m: $e^2(\pi N_i)^{1/2}/k_b T_m = 120$) of a two-dimensional Coulomb lattice [7]. Hence, at equilibrium the system would have to be rotationally invariant. Thus, the deviation from spherical symmetry of the pair distribution function $g(\mathbf{r})$ ($= A \times$ probability density of finding any two (not necessarily distinct) particles at a distance $\mathbf{r}$ from each other; A = total system area) is monitored to detect if the system has reached equilibrium. The dimensionless measure of this deviation was chosen to be

$$\Delta(g) = \int_A \left[g(r,\phi) - \frac{1}{\Phi(r)}\int_{\Phi(r)} g(r,\phi)\,\mathrm{d}\phi\right]^2 (N_i\,\mathrm{d}^2r)$$
$$(\mathbf{r} = (r,\phi))$$

where $\Phi(r)$ is the range of ϕ determined by r and the system boundaries. Equilibration occurred typically in $\sim 15\times10^3$ MC times steps (1 step = n_i Metropolis attempts). After equilibration the time-averaged $g(\mathbf{r})$ ($g_{\mathrm{av}}(\mathbf{r})$) is constructed using data from 15×10^4 MC steps which are spaced over $\sim 45 \times 10^4$ MC time steps to reduce temporal correlations. g_{av} constructed using the above procedure had $\Delta(g_{\mathrm{av}}) < 0.1$. Simulations were also carried out at 'infinite' temperature—in this case g_{av} is constructed from configurations generated independently by placing the particles randomly (however, particles were not allowed to be at the same position) and the averaging process is terminated when $\Delta(g_{\mathrm{av}}) < 0.1$. Finally, the structure factor $s(\mathbf{q})$ is computed as a two-dimensional (fast) Fourier transform of g_{av}:

$$s(\mathbf{q}) = \int_A g_{\mathrm{av}}(\mathbf{r})\,\mathrm{e}^{-\mathrm{i}\mathbf{q}\cdot\mathbf{r}}(N_i\,\mathrm{d}^2r).$$

Exploiting spherical symmetry ($\Delta(s) < 0.15$ in all cases), $s(\mathbf{q})$ is further averaged to construct $s(q)$ or, equivalently $\tilde{s}(q)$ through equation (3).

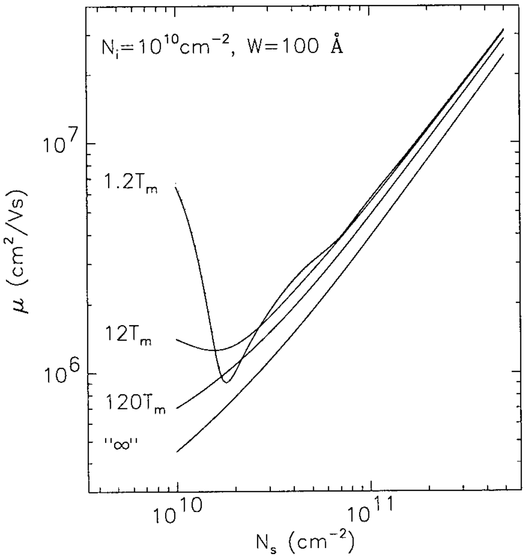


Figure 3. The curves show the computed Coulomb scattering limited mobility μ (as a function of the 2DEG density N_s) corresponding to the scatterer distributions of figure 2. N_i is the scatterer density and W is the spacer width. Dotted (full) curve(s) correspond to results for a scatterer system size of 42^2 (32^2).

Figure 2 shows the computed $\tilde{s}(q)$ for the different temperatures considered in this study. As expected, oscillations in $\tilde{s}(q)$ decay with increasing temperature and it becomes flat for the T = '∞' case. The variation in $\tilde{s}$ around $q = 0$ for this case is due to $\tilde{s}(q = 0)$ being zero by its definition (equation (3)). It illustrates the discretization of q due to finite system size. Errors (in the computed mobility at any temperature) due to such finite size effects can therefore be seen to become more severe with increasing spacer width W as then the values of $\tilde{s}(q)$ around $q = 0$ contribute more significantly to the integral in equation (4). By comparing results (for the mobility) from $\tilde{s}(q)$ corresponding to $n_i = 16^2$ and $n_i = 32^2$ it has been verified that an *interpolation* scheme that respects $\tilde{s}(q = 0) = 0$ is a poorer approximation than an *extrapolation* scheme that violates it. Hence, the extrapolation scheme is used throughout this study.

Mobility values have been computed by treating the parameters N_i, N_s and W (and in addition the scatterer distribution as determined by the equilibration temperature) as independent variables. This has been done to enable comparison with different experimental situations (such as the presence of gates that may affect the relationship between these parameters as determined by solving coupled Schrödinger–Poisson equations). The range of parameter values studied here are 10^{10} cm^{-2} $\leq N_i, N_s \leq 5 \times 10^{11}$ cm^{-2} and 100 Å $\leq W \leq$ 1400 Å. This is done for scatterer distributions corresponding to figure 2. Hence, computations corresponding to the larger system size ($n_i = 48^2$) provided an estimate of the finite size errors.

Figures 3, 4 and 5 illustrate some of the representative results. (It may be seen from these curves that finite size effects are negligible, although increasing with W.)

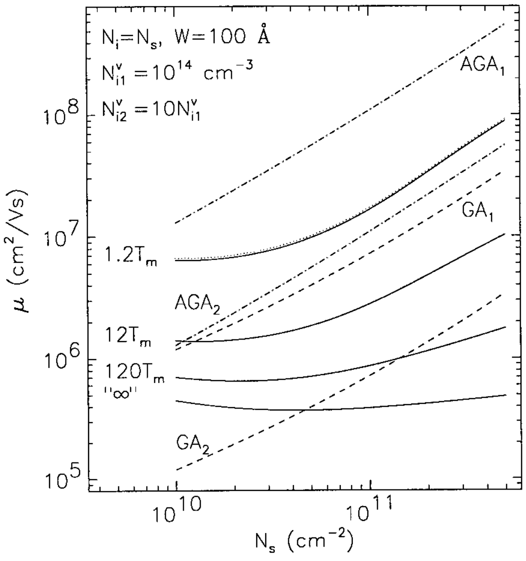


Figure 4. The curves show the computed Coulomb scattering limited mobility μ (as a function of the 2DEG density N_s) corresponding to the scatterer distributions of figure 2 with dotted (full) curve(s) corresponding to results for a scatterer system size of 42^2 (32^2). N_i is the scatterer density and W is the spacer width. For comparison, the mobility limited by impurities in the GaAs (GA$_{1,2}$) and AlGaAs (AGA$_{1,2}$) at volume densities $(N_i^v)_{1,2}$ is also shown.

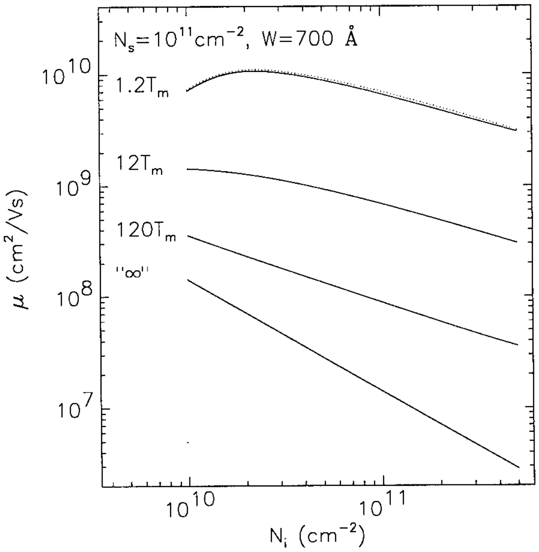


Figure 5. The curves show the computed Coulomb scattering limited mobility μ (as a function of the scatterer density N_i) corresponding to the scatterer distributions of figure 2. N_s is the 2DEG density and W is the spacer width. Dotted (full) curve(s) correspond to results for a scatterer system size of 42^2 (32^2).

These (and additional) results may be summarized by the following conclusions:

(i) Non-monotonic behaviour in the variations of the mobility (with N_s or N_i or both) is due to the oscillations in the structure factor of the corresponding scatterer distribution.

(ii) This non-monotonic behaviour decreases with increasing W and decreasing scatterer correlations.

(iii) The effect of scatterer correlations is more pronounced at larger W.

(iv) The concept of maximum possible mobility of a 2DEG is not particularly meaningful in the absence of scattering from the bulk (unintentional) impurities in the GaAs layer.

The mobility could be arbitrarily large depending on the equilibrium temperature of the scatterers and would therefore depend sensitively on the details of sample preparation and material growth. (As the depletion charge and scattering from it have been neglected the estimates are the most optimistic.) Our results, however, show conclusively that there is no limit in principle on the maximum achievable mobility because it increases with dopant correlations (and can be controlled by the equilibrium processing temperature). In particular, it should be possible to obtain a substantial enhancement (by as much as a factor of 10 perhaps) in the GaAs mobilities over the current best values [1] simply by using purer intrinsic GaAs material, which is still limited by scattering in the GaAs layer.

Acknowledgment

This work was supported by the US-ONR (Electronics Division, Larry Cooper).

References

[1] Pfeiffer L, West K W, Störmer H L and Baldwin K W 1989 *Appl. Phys. Lett.* **55** 1888
Pfeiffer L 1997 unpublished
[2] Stern F 1983 *Appl. Phys. Lett.* **43** 974
Saku T, Horikoshi Y and Tokura Y 1996 *Japan J. Appl. Phys.* **35** 34
[3] Das Sarma S and Stern F 1985 *Phys. Rev.* B **32** 8442
[4] Efros A L, Pikus F G and Samsonidze G G 1990 *Phys. Rev.* B **41** 8295
[5] Kawamura T and Das Sarma S 1996 *Solid State Commun.* **100** 411
[6] Shi J W, Koenraad P M, van de Stadt A F W, Peeters F W, Farias G A, Devreese J T, Wolter J H and Wilamowski Z 1997 *Phys. Rev.* B **55** 13 093
[7] Gann R C, Chakravarty S and Chester G V 1979 *Phys. Rev.* B **20** 326

Semicond. Sci. Technol. **13** (1998) A63–A66. Printed in the UK PII: S0268-1242(98)91876-4

Application of chemically enhanced vapour etching in the fabrication on nanostructures

M N Kozicki†, B Kardynal†, S-J Yang†, T Kim†, M V Sidorov‡ and David J Smith‡

† Center for Solid State Electronics Research, Arizona State University, Tempe, AZ 85287, USA
‡ Center for Solid State Science, Arizona State University, Tempe, AZ 85287, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. The technique of chemically enhanced vapour etching (CEVE) has been used in the fabrication of ultra-thin $CoSi_2$ wires on Si substrates. Essential aspects of the CEVE method are first summarized. We then demonstrate the application of the technique in the selective patterning of SiO_2 covered with a monolayer of organic acid. Parallel studies of thin Co layers directly deposited on (111) and (100) Si substrates were employed to optimize the $CoSi_2$ growth conditions. Conductance measurements of the $CoSi_2$ thin films and patterned wires are briefly discussed.

1. Introduction

As the sizes of semiconductor devices are reduced, the search continues for new lithography techniques which are capable of defining nanometre-size patterns. These techniques must have high sensitivity and resolution for good pattern definition and should also result in the formation of a robust surface mask. At present, the most commonly used resist in nanoscale device fabrication is polymethyl methacrylate (PMMA) [1], which can be patterned using electron beam irradiation. However, like other organic resists, PMMA is not particularly robust and therefore often has to be used for intermediate pattern transfer to another layer for 'harsh' processing steps such as reactive ion etching and silicidation. This limitation is not a consideration if inorganic materials such as SiO_2 are used directly as resists. SiO_2 is very attractive as a resist since it is extremely robust, does not introduce any additional contamination, and can even be left on the substrate as part of the final device, thereby reducing the number of processing steps. Moreover, it forms a very homogeneous layer with small molecular units, an attribute which is highly desirable for high-resolution lithography. On the other hand, direct patterning of silicon dioxide typically requires a very large irradiation dose, which in turn is likely to introduce damage into the substrate and make the lithography an impractically long process. In this paper we demonstrate the use of chemically enhanced vapour etching (CEVE) [2, 3] of SiO_2 covered with a monolayer of organic acid as a low-dose electron-beam lithographic technique for the fabrication of high-quality sub-micron metallic wires in Si substrates.

2. Description of method

Chemically enhanced vapour-phase hydrofluoric acid etching of silicon dioxide is possible in the presence of Brønsted bases, including water, on the surface of the oxide [4]. Etching of the silicon dioxide surface can be suppressed if the oxide surface is kept free from any molecules of such bases. In practice, this requirement means that the oxide must be clean and its temperature must be kept above the boiling point of water to avoid condensation on the surface. Care must also be taken to ensure the growth or deposition of high-quality oxide to avoid hydroxyl groups in its bulk. Local enhancement of the etching rate can then be achieved when water molecules or OH^- groups are captured on parts of the silicon oxide substrate in, for example, a thin layer (on the order of a monolayer) of organic molecules. This is the principle of the electron-beam lithographic technique that we use for patterning silicon oxide covered with undecylenic acid (10-undecenoic acid, $H_2C{=}CH(CH_2)_8COOH$) molecules for the production of sub-micrometre size features [5].

Sample preparation begins with the growth of a layer of thermal SiO_2 of the desired thickness. Oxide formation by chemical vapour deposition has also been successfully used. It is important to ensure that the oxide surface is terminated with a hydroxyl group, which then acts as an anchor for the head group of the monolayer molecule and results in a well-defined layer [4, 6]. Oxides grown at high temperature tend to be dehydroxylated [7], and an additional step is therefore needed to hydroxylate the surface in boiling water.

These substrates are subsequently cleaned and the layer of organic molecules can then be deposited. We

0268-1242/98/SA0063+04$19.50

used hydrocarbon contamination layers from laboratory air or vacuum chamber ambients in our earlier work and successfully demonstrated pattern formation in SiO_2 [8, 9]. However, optimum performance of the organic film requires molecules that are well ordered and small enough to obtain good resolution and which allow cross-linking to effectively take place. For this process, we chose undecylenic acid [10]. Molecules of the acid are deposited on the surface of the SiO_2 substrate by immersion in 1 mM solution of undecylenic acid kept at room temperature. Hydrogen bonding takes place between the molecules and the hydroxylated substrate and the samples are then ready for patterning.

Local modification of the layer of undecylenic acid molecules is achieved by exposure to electrons with energies of about 5 eV or more. The effect of the irradiation is two-fold. Firstly, it breaks carbon double bonds in the molecules resulting in cross-linkage between them, which then allows retention of the water molecules that are needed for CEVE. Secondly, species are formed during the etch step which, under suitable conditions, act as Brønsted bases and participate in the etching [4]. It is therefore possible, in principle, to use either scanning electron microscopy or scanning tunnelling microscopy as the lithographic tool. The latter method has an advantage in terms of resolution since there are no backscattered electrons which would broaden the lines by additional cross-linkage of the acid molecules. The disadvantages of the method are that the oxide thickness is limited to about 4 nm, and the Si substrate must be highly doped. These restrictions of the method narrow its use in electronic devices. The resolution of SEM-based lithography, on the other hand, is limited by backscattered electrons. For typical energies of the SEM electron beam of between 10 keV and 40 keV, backscattered electrons are capable of cross-linking the monolayer, producing a significantly broadened Gaussian-type distribution of the cross-linked molecules and of etch species.

Exposure of the monolayer is followed by selective removal of the uncross-linked molecules. Different schemes can be used to remove the monolayer, ranging from desorption of lighter molecules in oxygen or nitrogen ambient to various solution treatments. We investigated many methods for the removal of the uncross-linked monolayer and found that samples that had a monolayer removed by thermal treatment both in ambient air and oxygen also had a high etching rate in HF vapour and low selectivity, as defined by the ratio of the etching for exposed and unexposed areas. From studies of removal of the monolayer in solutions, we concluded that acetone cleaning in an ultrasonic bath was sufficient to break the hydrogen bonding between the molecules and silicon dioxide without damage to the cross-linked molecules. This method allowed good etch selectivity to be obtained.

The dependence of the selectivity on the electron beam dose is shown in figure 1. The etch rate of the silicon oxide at zero dose varies from 0.1 to 1.0 nm s^{-1} from sample to sample but the selectivity remains the same. Selectivity, which is proportional to the etch rate of the silicon dioxide in the CEVE, increases with the degree of monolayer cross-linking (the electron dose). This is a direct result of the dependence of the CEVE etching rate on the number of Brønsted base groups on the surface of the sample: more cross-linked areas will be etched faster than the less cross-linked ones, since they are able to generate and retain more H_2O. Dependence of the etch rate on the electron dose combined with proximity effects in SEM lithography discussed earlier, result in Gaussian etch profiles for trenches in SiO_2. An example is visible in the cross-sectional transmission electron micrograph (TEM) shown in figure 2.

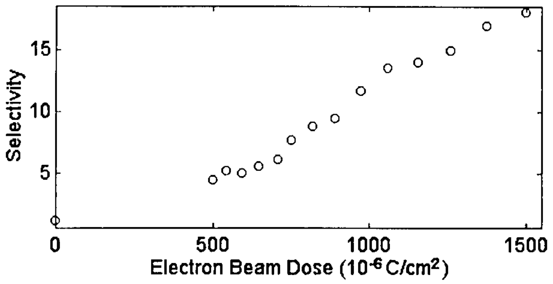


Figure 1. Dependence of etch selectivity on total dose of electron beam exposure after monolayer is removed in acetone. The etch rate of areas not exposed to the electron beam is about 1 nm s^{-1}.

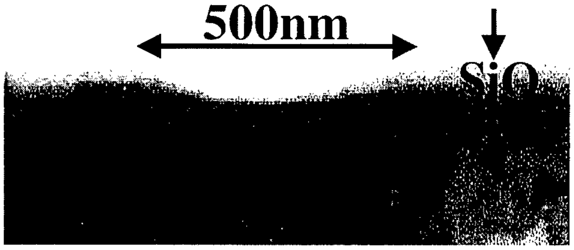


Figure 2. Etching profile of a trench in SiO_2 that was exposed to an electron beam of diameter 50 nm, passing only once along the wires. The width of the trench thus reflects the effect of backscattered electrons.

Although the selectivity increases with the beam dose, doses below 600 μC cm^{-2} can be used for pattern generation in silicon dioxide. This is very important since more damage is introduced in the silicon with increased electron beam dose, so our method does not suffer from the problem of damage known in direct patterning of SiO_2.

3. Application of the method

Once the desired pattern is made in SiO_2, it can be transferred into the Si substrate by selective etching, ion implantation, etc. Alternatively, the patterned oxide can be used to limit the thermal reaction between the silicon and a deposited layer of metal. In our study we used this lithographic technique to form cobalt disilicide ($CoSi_2$) wires. Cobalt is first deposited on the patterned substrate and capped with a thin layer of Si to avoid oxidation of the metal before $CoSi_2$ is formed [11]. Evaporation is performed in an electron-beam evaporator with a vacuum of less than 4×10^{-6} Torr. This bilayer is then subjected

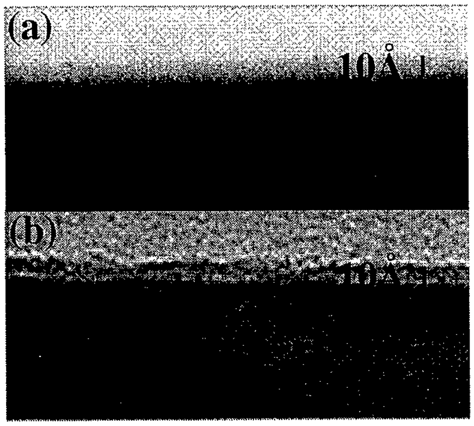


Figure 3. Electron micrograph of an as-evaporated Co/Si bilayer with a nominal thickness of 0.7 and 1 nm on (a) (111) and (b) (100) substrates.

to rapid thermal annealing in an N_2 ambient at 700 °C for 20 s. The final thickness of the $CoSi_2$ is determined from the initial thickness of the cobalt layer, which was varied between 1 and 10 nm. For the typical density of evaporated Co, the thickness of the resulting cobalt disilicide should be about 3.5 times the thickness of the cobalt layer [12]. The thickness of the Si capping layer is kept to about 1–2 nm to ensure that the cobalt reacts mostly with the Si substrate. After thermal annealing, samples are treated with $HCl:H_2O_2 = 3:1$ solution to remove non-reacted metal from the surface of the SiO_2 mask.

In parallel with these $CoSi_2$ lithographic studies, we also investigated the formation of $CoSi_2$ on (111) and (100) substrates using the same deposition methods. Cross-sectional electron micrographs of Co and Si bilayers deposited on Si (111) and (100) substrates, with nominal thicknesses of 0.7 and 1 nm respectively are shown in figure 3. These samples were prepared simultaneously to avoid variations between processes. The micrographs show that the Co layer is much more uniform on the (111) substrate. The reaction between Co and Si is completed during annealing at 700 °C for 20 s in the N_2 ambient, as shown in figure 4. In both cases, the reaction between Co and Si results in preferential alignment of the $CoSi_2/Si$ junction along (111) planes, thus favouring the formation of continuous thin layers of $CoSi_2$ on the Si(111) substrate, compared with formation of large discontinuous grains of $CoSi_2$ on the (100) substrate. In both cases, the resulting structure is polycrystalline, although a large percentage of the grains on the (111) substrate are epitaxially aligned with the substrate.

We assessed the electrical properties of the layers by four terminal measurements performed at 77 K. Results of the measurements performed on the macroscopic $CoSi_2$ resistors (10–100 μm wide with the separation between the voltage probes between 200 and 1000 μm) formed on Si (111) substrates (resistivity 12 Ω cm) are shown in figure 5. The parameters derived from these results agree well with the TEM study of the layers. The measurements were performed at 77 K to avoid parallel conductance of the Si substrate, an effect which is clearly visible in the room temperature data (figure 6). Contribution of the substrate conductance to the overall measured conductance was estimated from the measurements of Si with $CoSi_2$ Schottky contacts, and was at least six orders of magnitude larger than the resistance of the measured wires. By measuring the conductance between two independent wires we confirmed that the metal layer was completely removed from the SiO_2 in the course of processing.

As expected, the conductance per square of the film decreased almost linearly with the thickness of the film. The resistivity of the film subtracted from the data is

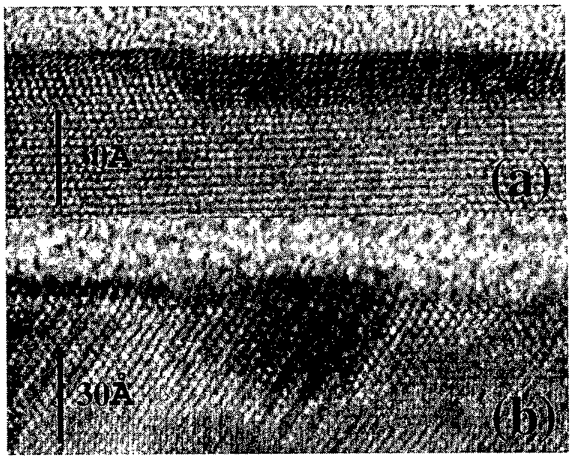


Figure 4. TEM micrograph of the $CoSi_2$ layer formed by rapid thermal annealing of an evaporated Co/Si bilayer on (a) (111) and (b) (100) substrates. The annealing took place at 700 °C for 20 s in N_2 ambient.

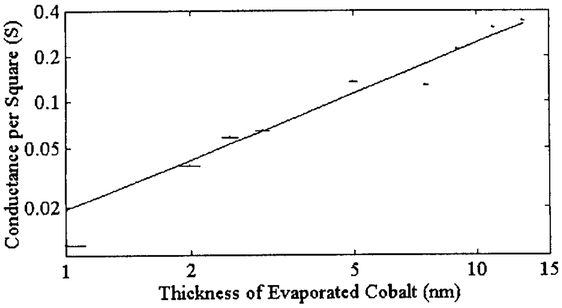


Figure 5. Conductance per square of $CoSi_2$ thin films as function of thickness of evaporated Co.

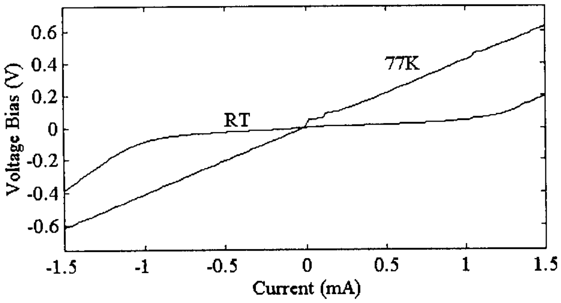


Figure 6. Typical example of room temperature and 77 K I/V characteristics taken on film fabricated from 1 nm Co evaporation.

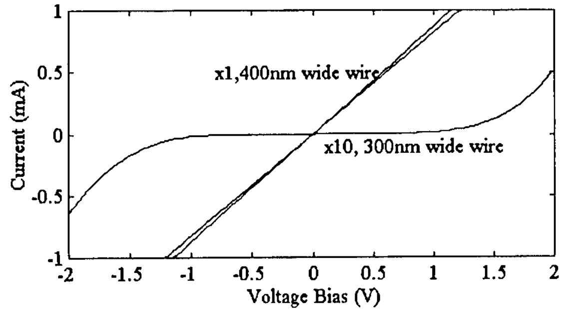


Figure 7. *I*/*V* characteristics of $CoSi_2$ wires patterned by electron-beam exposure of a monolayer on SiO_2 thin films followed by CEVE etching. The width of the wires is subtracted from the SEM micrograph of wire.

7.4×10^{-8} Ω m and is about twice as large as the bulk resistivity of the $CoSi_2$ at 77 K [12]. A similar result was observed for the resistivity of the thin $CoSi_2$ film at room temperature [13]. Measurements of the thin layers of $CoSi_2$ on Si (100) substrate showed that the silicide layer formed after evaporating less than 2 nm of cobalt did not conduct at 77 K.

It is worth noting that the thinnest layers of $CoSi_2$ are sufficiently thin that a decrease of the carrier concentration due to two-dimensional quantization of the energy levels should be considered [14]. Assuming carrier concentration of holes in bulk $CoSi_2$ to be 3.3×10^{28} m^{-3} and the effective hole mass to be equal to 1, then the hole Fermi wavelength is 0.93 nm. A uniform $CoSi_2$ layer of 3.5 nm thickness is populated by seven two-dimensional subbands and the carrier concentration is about 90% of the bulk carrier concentration. Although the layers are not uniform, our measurements are sensitive to the average thickness and the value given above is a good estimate of the average carrier concentrations in the layers.

Following the measurements of the thin cobalt disilicide layers of macroscopic size, we fabricated wires 5 μm long and varying width using the CEVE lithography technique with SEM exposure patterning of the monolayer. The wires were fabricated using 200 nm thick oxide as a mask and a 2 nm cobalt layer. The results of the measurement of these wires in a two-terminal configuration at room temperature are shown in figure 7. This figure shows that as the width of the wire, as estimated from the SEM micrographs, changes from 400 nm to 300 nm, the conductance of the wire changes from metallic to activated-type conductance. This change in behaviour can be explained on the basis of the etching profile of the trenches, as illustrated in figure 2. As is clear from the figure, due to the isotropic nature of the etch, the width of the trench at its base is much less than at the top. It is therefore likely that there is a residual SiO_2 layer in the trench which was used to form the 300 nm wide structure. Such residues stop formation of silicides and the structure is therefore discontinuous, i.e. it is likely that there is no silicide wire between the contact pads. The conductance of such a structure is therefore merely the conductance of two Schottky diodes back-to back, where the large contact pads for the 'missing' wire serve as Schottky metallization to the Si substrate.

4. Conclusions

We have demonstrated the possibility of using chemically enhanced vapour etching in the selective patterning of SiO_2 to form ultra-thin $CoSi_2$ wires. Studies involving the formation of $CoSi_2$ thin films on Si show that quasi two-dimensional layers of $CoSi_2$ can be formed on (111) Si but that there is a lower limit for layer thickness on (100) material. Based on above results, we fabricated ultra-thin $CoSi_2$ wires of sub-micrometre size and showed that they could be conductive at sub-10 nm thickness. Further decrease of the lateral dimensions can therefore produce devices which will show mesoscopic type behaviour. It is important to note that the fabrication of these devices is significantly simpler than commonly used techniques for fabrication of mesoscopic devices.

References

[1] Moreau W M 1988 *Semiconductor Lithography Principles, Practices, and Materials* (New York: Plenum) p 446
[2] Kozicki M N 1989 *US Patent* 4904338
[3] Kozicki M N, Ramesh V and Williams D 1988 *Proc. Interface '88* (KTI Corp.) p 283
[4] Whidden T K, Yang S-J, Jenkins-Gray A, Pan M and Kozicki M N 1997 *J. Electrochem. Soc.* **144** 605
[5] Pan M 1996 Electron-beam lithography using self-assembled monolayer on silicon dioxide *Master Thesis* Arizona State University
[6] Marshall K and Rochester C H 1988 *J. Chem. Soc. Faraday Trans.* **1 71(7-12)** 569
[7] Vansant E F, Van Der Voort P and Vrancken K C 1995 *Characterization and Chemical Modification of the Silica Surface* (New York: Elsevier)
[8] Allgair J, Rack M J, Whidden T K, Kozicki M N and Ferry D K 1996 *J. Vac. Sci. Technol.* A **14(3)** 1855
[9] Kozicki M N, Allgair J, Jenkins-Gray A, Ferry D K and Whidden T K 1996 *Physica* B **227** 318
[10] Pan M, Yun M, Kozicki M N and Whidden T K 1996 *Superlatt. Microstruct.* **20** 369
[11] Kamal A H M, Rack M J, Kozicki M N, Ferry D K, Lutzen J and Hallmark J A 1997 *J. Vac. Sci. Technol.* B **15** 899
[12] Maex K and van Rossum M (ed) 1995 *Properties of Metal Silicides* (London: IEEE) and references therein
[13] Maszara W P 1993 *Appl. Phys. Lett.* **62** 961
[14] Trivedi N and Ashcroft N W 1988 *Phys. Rev.* B **38** 12298

Semicond. Sci. Technol. **13** (1998) A67–A70. Printed in the UK

PII: S0268-1242(98)91877-6

Conductance oscillations in overgrown sub 100 nm InP/Ga$_{0.25}$In$_{0.75}$As quantum wires

I Maximov†, N Carlsson†, P Omling†, P Ramvall†||, L Samuelson†, W Seifert†, Q Wang†, S Lourdudoss‡, E Rodriguez Messmer‡, A Forchel§ and K Kerkel§

† Solid State Physics, Lund University, Box 118, S-221 00, Lund, Sweden
‡ Department of Electronics, KTH-Electrum 224, S-164 40, Kista, Sweden
§ Lehrstuhl für Technische Physik, University of Würzburg, Am Hubland, D-97074 Würzburg, Germany

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We report the observation of strong conductance oscillations as a function of Fermi energy in the quantized plateau region in epitaxially overgrown, heterostructurally defined, 90 nm wide InP/Ga$_{0.25}$In$_{0.75}$As quantum wires (QWRs). The oscillations are strong at liquid-helium temperatures but disappear either at temperatures around 10 K or if the excitation current is increased to 350 nA. We propose that the observed oscillations are caused by quantum interference of electron waves reflected at the sharp boundaries between the quantum wire and the wide 2DEG, defined by the heterostructure. Details of the fabrication of QWRs as narrow as 60 nm are also reported.

Quantized conductance in one-dimensional conductors has attracted a great deal of attention since its discovery in 1988 [1, 2]. If the width of the conductor is of the order of the Fermi wavelength of an electron in a two-dimensional electron gas (2DEG), a quantization of the conductance in units of $2e^2/h$ can be observed. The most common way to define the one-dimensional channel in a 2DEG is to deposit Schottky gates on top of an AlGaAs/GaAs modulation-doped heterojunction and to apply a negative voltage on the gates to deplete the electron gas underneath [3]. Using this concept, the definition of a quantum wire (QWR) is fairly straightforward. However, electrostatically defined QWRs are not suitable for device applications at elevated temperatures because of the weak lateral confining potential. Furthermore, they can be used only with materials on which good Schottky contacts can be formed.

Another approach includes the fabrication of a QWR by a geometrical definition of the narrow electron channel by high-resolution lithography and etching. A regrowth of the etched QWR would reduce electron scattering at the etched interface, allowing the formation of a high-quality, one-dimensional channel with the lateral confinement governed by the heterostructure band offset. The modulation-doped, InP/GaInAs quantum well system is most suitable for such an approach. The lower effective mass of this material system compared with the AlGaAs/GaAs one offers, apart from the larger quantization, a higher low-field electron mobility and a higher peak electron velocity because of the increased quantization, leading to a reduced intersubband scattering. The InP/GaInAs material system also has the advantage of not containing aluminium and thus not being susceptible to oxidation problems prior to overgrowth.

In this paper we demonstrate the fabrication of QWRs as narrow as 60 nm. We also have measured quantized conductance of 90 nm wires. In addition to the expected plateaux at integer steps of $2e^2/h$, and the plateaux suggested to be related to spin polarization at $0.2(2e^2/h)$, $0.7(2e^2/h)$ and $1.5(2e^2/h)$, strong conductance oscillations in the plateau regions were observed. These oscillations might be an indication of a quantum wire defined by sharp sidewalls [4–7], thus confirming the quality of our overgrown interfaces.

The samples were grown by low-pressure, metal organic vapour phase epitaxy (MOVPE) at 50 mbar on a semi-insulating (Fe-doped) (001) InP substrate. The following layer sequence was employed: a 2500 Å thick, not intentionally doped (NID) buffer layer, a 90 Å thick Ga$_{0.25}$In$_{0.75}$As quantum well layer, a 200 Å thick NID InP spacer layer, a 10 Å thick InP layer, Si doped to a density of 5×10^{18} cm^{-2}, and a 200 Å thick NID InP layer. The mobility (μ) and the sheet electron concentration (n_s) of the 2DEG after illumination, determined by the Hall effect at liquid-helium temperature, were typically $\mu = 4 \times 10^5$ cm^2 V^{-1} s^{-1} and $n_s = 5 \times 10^{11}$ cm^{-2},

---

|| Present address: RIKEN, The Institute of Physical and Chemical Research, Semiconductor Laboratory, 2-1 Hirosawa, Wako-shi, Saitama 351-01, Japan.

0268-1242/98/SA0067+04$19.50

respectively. The mobility is of the order of the highest achievable in a ternary material system.

Mesas were fabricated by conventional optical lithography and wet etching. As current and voltage probes, thermally deposited Au/Ge/Au ohmic contacts, alloyed at 440 °C for 2 min, were used. An 80 nm thick PMMA (950 K) layer was spin coated on the ready-made mesas, and a quantum well structure was defined by electron beam lithography using a scanning electron microscope (SEM) operating at 35 kV. The PMMA was exposed by a current of 100 pA and developed in a 1:3 mixture of methylisobutylketone and isopropanol for 60 s at 22 °C. The QWR was defined as a narrow channel, approximately 100 nm long, oriented along either the $[1\bar{1}0]$ or the [110] direction. For wet etching we used a 1:2:2 $HCl:CH_3COOH:H_2O_2$ solution at 15 °C, which etches both InP and GaInAs at a rate of about 9–13 nm s^{-1}. After a standard cleaning procedure, the QWR structures were overgrown either by low-pressure MOVPE with undoped InP or by hydride vapour phase epitaxy (HVPE) at atmospheric pressure and 685 °C with semi-insulating (Fe-doped) InP [8]. Both as-etched and overgrown QWRs were characterized by transport measurements, using a top gating technique to control the electron concentration in the 2DEG reservoirs and thus in the QWR [9]. Atomic force microscopy (AFM) and SEM were used to characterize the wire structures at the different stages of their fabrication.

Wet etching of InP/GaInAs produced a fairly smooth surface suitable for epitaxial regrowth without any additional treatment. Under nominally the same etching conditions, the variation of the etching rate of InP/GaInAs was typically about 20–30% [10]. Figure 1 shows an AFM image of an $InP/Ga_{0.25}In_{0.75}As$ QWR etched for 6 s. The 100 nm long channel is oriented along the $[1\bar{1}0]$ direction, in which a trapezoidal cross-section is expected (see below). In order to make sure that the $Ga_{0.25}In_{0.75}As$ layer was completely removed in the etched areas, the etching time had to be increased. However, backward-scattered electrons from adjacent exposed regions destabilize the resist mask, thus limiting the etching time to less than 10 s. The degradation of the PMMA masking properties is known to be due to increased porosity. Pores of the order of 1 nm in diameter are formed because of the release of volatile products of the main chain scission and excess free volume appears in the polymer films [11]. AFM images of as-etched QWRs indicated that the porosity of PMMA for any real pattern results in erosion of the InP cap layer. In order to minimize the underetching of the resist mask, an etching time of 6–7 s was used to produce a conducting, 90 nm wide QWR. For channels smaller than 90 nm, the unintentional thinning of the cap layer was so severe that it was impossible to define the QWR by wet etching without causing some damage.

These results confirm the assumption that the insufficient masking property of PMMA is the major limiting factor for decreasing the size of the QWR. To improve the masking ability of the resist in the wet etchant, we performed post-development baking of the PMMA above the glass transition temperature (T_g = 105 °C). This results in an increase of its masking ability during wet etching, but it also widens the mask. Experiments to determine the optimal baking conditions showed that the drift of PMMA is minimal at baking temperatures ranging from 110 °C to 130 °C. The masking properties of PMMA are also determined by the proximity dose and the baking temperature. After exposure and a post-development baking of PMMA at 130 °C for 30 min, the samples were etched for 7 s. SEM investigations of the surface after removal of PMMA did not show any etched holes in the InP. By contrast, the same dose on unbaked resist resulted in the formation of holes in the InP, indicating a degradation of the PMMA.

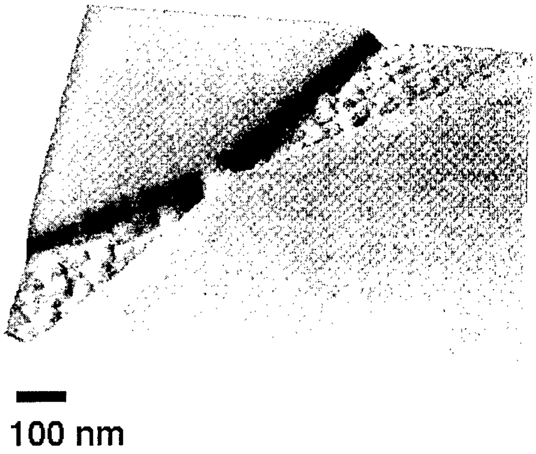


Figure 1. An AFM image of an as-etched, 90 nm wide and 100 nm long QWR defined in the $InP/Ga_{0.25}In_{0.75}As$ quantum well. The channel is oriented along the $[1\bar{1}0]$ direction. The etching depth is approximately 80 nm, as determined by AFM. The roughness of the etched surface is caused by resist residues remaining after development.

Investigations of the wet-etched $InP/Ga_{0.25}In_{0.75}As$ QWRs showed that the orientation of the channel along the $[1\bar{1}0]$ direction results in an etching profile which forms an angle of ≈130° between the sidewalls of the channel and the surface. Because of the trapezoidal shape, a decrease in width of the QWR in the $[1\bar{1}0]$ direction to less than 90 nm is not possible without thinning the InP cap layer. On the contrary, the [110] orientation allows the formation of almost vertical sidewalls. It was found that baking PMMA at 130 °C made it possible to fabricate QWRs as narrow as 60 nm (see figure 2). In addition, the resulting width of a QWR in the [110] direction was found to be sensitive to the baking temperature, and a change of the exposure dose by a few per cent led to a measurable change in the QWR width. From the processing point of view, it is therefore preferable to orient the QWR along the [110] direction.

The $[1\bar{1}0]$ 90 nm-wide and [110] 60 nm wide as-etched QW structures were overgrown either by MOVPE or HVPE. Both regrowth techniques showed good surface morphology, as observed by AFM and SEM. In addition, the cross-section of a cleaved test structure overgrown by MOVPE, investigated by SEM, showed no voids at the overgrown interface.

For transport measurements we used the 90 nm wide QWRs, because of the larger number of occupied one-dimensional subbands propagating in this structure

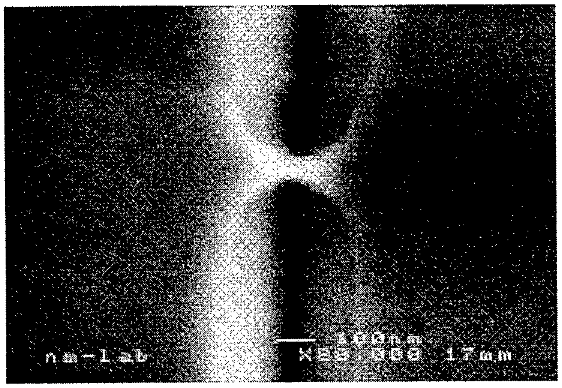


Figure 2. An SEM image of a 60 nm wide QWR formed after etching for 7 s. The channel is oriented in the [110] direction. The PMMA was baked at 130 °C for 30 min. Using unbaked PMMA under otherwise similar conditions did not result in the formation of a QWR because of poor masking ability.

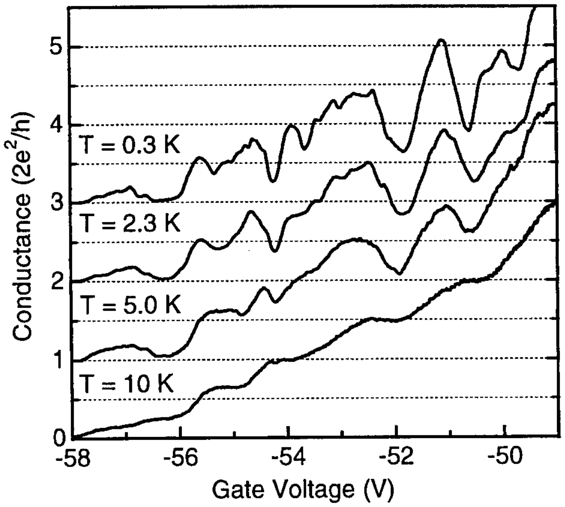


Figure 3. The conductance of the QWR as a function of the gate voltage. The amplitude of the oscillations decreases if the temperature is increased. For clarity, the curves are offset by $2e^2/h$.

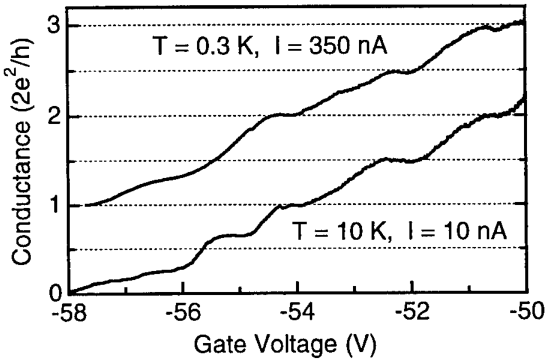


Figure 4. The oscillations in the conductance as a function of gate voltage can be quenched by increasing either the temperature or the current driven through the device. The two curves are offset by $2e^2/h$.

compared with a narrower one. Thus we expect to observe a larger number of conductance steps, which makes the interpretation of the data easier. Both as-etched and overgrown samples were characterized. A typical resistance of an as-etched QWR was 100 kΩ, which suggests an electric width smaller than 90 nm [2]. This can be readily explained by depletion from the sidewalls of the QWR caused by the pinning of the Fermi level on the etched surfaces. After regrowth with InP, the QWRs showed a much lower resistance, consistent with a low density of defects at the InP/GaInAs interface, i.e. less sidewall depletion. When the Fermi level of the overgrown samples was decreased by means of a top gate covering both the QWR and the wide 2DEG, the conductance did not decrease monotonically, but several steps were observed. Plateaux at integer steps of $2e^2/h$, as well as plateaux suggested to be related to spin polarization at $0.2(2e^2/h)$, $0.7(2e^2/h)$ and $1.5(2e^2/h)$, were observed. A detailed description of this measurement has been published elsewhere [9]. However, these steps were clearly observed only at elevated temperatures (see figure 3). At lower temperatures strong oscillations in the conductance as a function of the gate voltage are evident. The oscillations are apparently similar to those previously observed in a QWR formed in AlGaAs/GaAs by means of split gates [12]. The oscillations observed in the cited paper disappeared, however, already at 1 K. The origin of the oscillations may be a partial reflection of electron waves at both ends of the QWR [4, 12, 13] or to other conductance fluctuations [14].

The fact that quantized conductance plateaux are clearly observed suggests that very few impurities are present in the QWR region. In order to investigate this further, we have conducted measurements where the gate voltage (corresponding to the Fermi level) was kept constant while a magnetic field applied perpendicular to the QWR was swept over it. The QWRs for which quantized conductance could be observed did not show any weak localization or universal conductance fluctuations. This finding suggests that the oscillations observed as a function of gate voltage are caused by an interference of electron waves reflected at the sharp boundaries between the QWR and the wide 2DEG, defined by the heterostructure. The possibility of reflections at imperfections at the heteroepitaxially overgrown interface cannot yet be completely ruled out, however. Finally, it is worth noting that the oscillations of the conductance also decrease if the current driven through the device is increased (see figure 4), a result which is commonly observed for quantum interference effects.

To summarize, we have demonstrated the fabrication of epitaxially overgrown InP/$Ga_{0.25}In_{0.75}As$ quantum wires as narrow as 60 nm. The regrowth of the as-etched QWRs results in substantial improvement of interface properties, and quantum conductance and strong interference effects in the conductance are observed. The high quality of the overgrown wires suggests the InP/$Ga_{0.25}In_{0.75}As$ material system as a promising candidate for the fabrication of devices based on modal evolution in electron wave guides, such as the Y-branch switch [15].

Acknowledgments

This work was performed within the Nanometre Structure Consortium in Lund and was financed by the Swedish research agencies NFR, TFR, NUTEK and SSF and the European Commission through ESPRIT IV Long Term Research, Q-SWITCH programme, project number 20960. The authors would like to thank A Bogdanov and H Xu for useful discussions.

References

[1] Wharam D A, Thornton T J, Newbury R, Pepper M, Ahmed H, Frost J E F, Hasko D G, Peacock D C, Ritchie D A and Jones G A C 1988 *J. Phys. C: Solid State Phys.* **21** L209
[2] Wees B J v, Houten H v, Beenakker C W J, Williamson J G, Kouwenhoven L P, Marel D v d and Foxon C T 1988 *Phys. Rev. Lett.* **60** 848
[3] Thornton T J, Pepper M, Ahmed H, Andrews D and Davies G J 1986 *Phys. Rev. Lett.* **56** 1198
[4] Kirczenow G 1988 *Solid State Commun.* **68** 715
[5] Szafer A and Stone A D 1989 *Phys. Rev. Lett.* **62** 300
[6] Tekman E and Ciraci S 1991 *Phys. Rev.* B **43** 7145
[7] Xu H 1993 *Phys. Rev.* B **47** 15 630
[8] Lourdudoss S, Hammarlund B and Kjebon O 1990 *J. Electron. Mater.* **19** 981
[9] Ramvall P, Carlsson N, Maximov I, Omling P, Samuelson L, Seifert W, Wang Q and Lourdudoss S 1997 *Appl. Phys. Lett.* **71** 918
[10] Maximov I *et al* 1997 *Proc. 9th Int. Conf. on InP and Related Materials (Hyannis, Cape Cod, MA, 1997)* p 145
[11] Ouano A C 1978 *Polym. Eng. Sci.* **18** 306
[12] Wees B J v, Kouwenhoven L P, Willems E M M, Harmans C J P M, Mooij J E, Houten H v, Beenakker C W J, Williamson J G and Foxon C T 1991 *Phys. Rev.* B **43** 12 431
[13] Hirayama Y, Saku T and Horikoshi Y 1989 *Phys. Rev.* B **39** 5535
[14] Stone A D 1985 *Phys. Rev. Lett.* **54** 2692
[15] Palm T and Thylén L 1992 *Appl. Phys. Lett.* **60** 237

Semicond. Sci. Technol. **13** (1998) A71–A74. Printed in the UK

PII: S0268-1242(98)91878-8

Compatibility of cobalt and chromium depletion gates with RPECVD upper gate oxide for silicon-based nanostructures

M J Rack†, A D Gunther†, M Khoury†, D Vasileska†, D K Ferry† and M Sidorov‡

† Center for Solid State Electronics Research, Arizona State University, Tempe, AZ 85287-6206, USA
‡ Center for Solid State Science, Arizona State University, Tempe, AZ 85287-1704, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. Three-dimensional confinement of electrons in silicon-based nanodevices may be achieved using a dual gate structure to confine carriers laterally in a 2D MOSFET inversion layer. We have investigated the temperature stability of cobalt and chromium for thin depletion gates, using remote plasma enhanced chemical vapour deposited (RPECVD) SiO_2 for the deposited dielectric. The thermal stability of the oxide/metal/oxide structure for various annealing regimes was studied by Auger electron spectroscopy sputter profiling and high-resolution cross-sectional transmission electron microscopy. Improvements to the RPECVD oxide for comparable annealing were characterized by electrical measurements on MOS capacitors made from deposited RPECVD oxide.

1. Introduction

In order to achieve 3D confinement of electrons in silicon-based nanodevices, additional gate structures may be employed to confine carriers laterally in a 2D MOSFET inversion layer. By locally depleting the inversion layer charge, these gates create confinement provided that they are very close to the inversion layer. The inverting gate may be more remote. The vertical dual gate structure shown in figure 1 illustrates the effect of such depletion gates on the inversion layer charge density and conduction band energy. This was simulated by our 3D Poisson solver [1] for 20 nm × 20 nm depletion gates separated by 50 nm, a 5 nm thermal oxide and a 50 nm deposited oxide. The source and drain are not shown. If the lower gates are negatively biased (−0.5 V), inversion charge density created by the positively biased top gate (13.75 V) is reduced and a potential barrier is created that may be used to confine electrons.

This implementation necessitates that a deposited oxide perform as a gate oxide. However, deposited oxides are generally inferior to thermally grown oxides and benefit from high-temperature annealing [2]. These anneals reduce bulk defects and interface charge in MOS structures fabricated from RPECVD oxides. If sufficiently high temperatures can be tolerated, a deposited oxide can function as well as a thermally grown gate oxide [3]. It is the thermal stability of the previously patterned depletion gates that may dictate the thermal budget allowed for improving the deposited oxide.

In this study, we investigated the use of cobalt and chromium for depletion gates and RPECVD SiO_2 for the deposited oxide. These metals were chosen for their high melting temperature, good adhesion and previously reported thermal stability [4,5]. The stability of the oxide/metal/oxide structure for various annealing regimes was studied by Auger electron spectroscopy (AES) sputter profiling and high-resolution cross-sectional transmission electron microscopy (HRXTEM). The improvements to the RPECVD oxide for comparable annealing were characterized by electrical measurements on MOS capacitors made from deposited RPECVD oxide.

2. Experimental details

The oxide/metal/oxide structures were composed of 20 nm thermal oxide on a silicon (100) substrate. Cobalt or chromium (15–20 nm) was evaporated on this oxide followed by 50 nm of RPECVD SiO_2 deposited in a microwave reactor. AES sputter profiles were obtained with a PHI 600 scanning Auger multiprobe. The energy region for the silicon signal was set at 82–97 eV in order to detect free silicon or any silicide that might have formed which would suggest that the metal had reacted measurably with the oxide, and to exclude silicon in SiO_2.

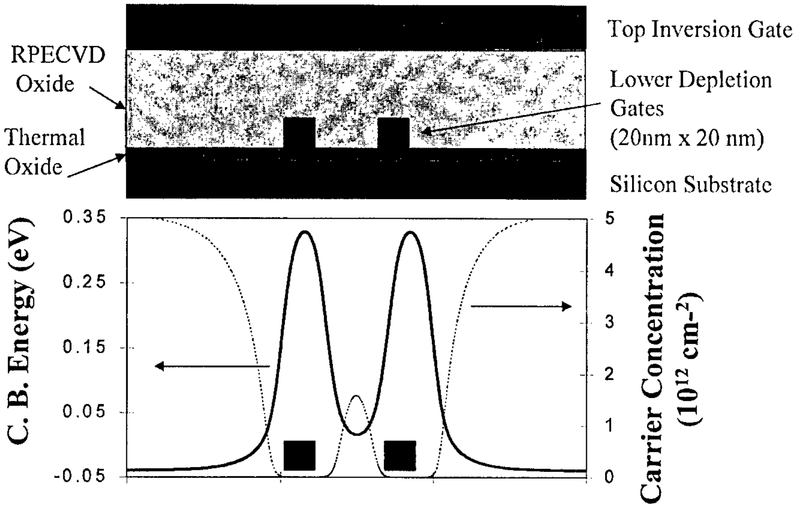


Figure 1. Effects of dual-gate structure on MOSFET inversion layer carrier density and conduction band energy.

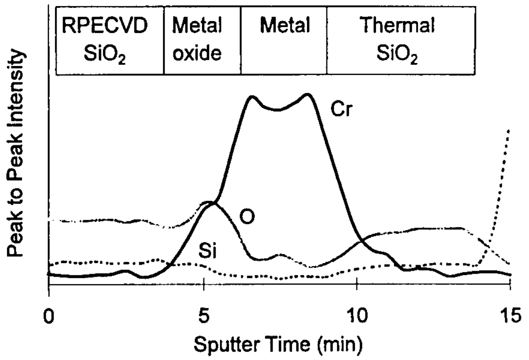


Figure 2. As-deposited chromium sample with unwanted metal oxide from RPECVD SiO$_2$ deposition step.

MOS capacitors were fabricated with shadow mask and optical lithography on boron-doped Si (100). The post-deposition anneals (PDAs), done in both the rapid thermal processor and a standard oxidation furnace, immediately followed oxide deposition. Aluminium gate contacts were used. A post-metallization anneal was performed at 450 °C for 15 min in forming gas.

3. Metal thermal stability

An AES sputter profile of the as-deposited SiO$_2$/Cr/SiO$_2$ 'metal sandwich' structure is shown in figure 2. Most of the top 50 nm deposited oxide has been pre-sputtered away to reveal the interfaces with more detail. We have found that the oxide deposition process will inevitably cause some oxidation of the metal because of the temperature and oxidizing ambient as can be seen in this profile.

Figure 3 is the AES profile and HRXTEM micrograph for the 700 °C anneal of both cobalt and chromium sandwich structures. The as-deposited metal oxide layer is beginning to disappear in the chromium sample and is no longer a separate layer in the cobalt system, with the oxygen content in the metal layer increasing as a result. The cobalt metal layer as seen in the HRXTEM image is no longer uniform, having bulging grains, and voids heavily concentrated at the thermal oxide–metal interface. Crystalline cobalt oxide was found in the metal layer in the form of grains (crystalline). The chromium layer is still intact at 700 °C. There is no indication of silicide or free silicon, which would suggest that the metal had reduced the SiO$_2$, and there are no bulging grains.

The effect of a 900 °C anneal on both chromium and cobalt oxide/metal/oxide structures is shown in the AES profiles and HRXTEM micrographs of figure 4. Small spheres can be seen in this micrograph in the chromium metal layer. The AES profile shows increased oxygen in the metal layer, a broadening of the metal layer thickness and some tailing of the chromium into the thermal oxide. An examination of the chromium MMM AES spectra at 35 eV indicates that more of the chromium is oxidized than in the as-deposited or 700 °C annealed metal layer. This implies that the chromium oxide that is formed during RPECVD SiO$_2$ deposition is incorporated into the metal layer during the high-temperature annealing. In the cobalt TEM micrograph, similar spheres are much larger and the metal layer is very deteriorated. This is evident in the AES sputter profile as well.

The instability of the Cr oxide is consistent with earlier reports that it could be reduced above 650 °C [6]. Clearly, the formation of oxide grains in a one-grain-wide metal wire would be detrimental to that wire's conductivity. In order to minimize the detrimental effects of the metal oxide that is created during the RPECVD oxide deposition, a 3 nm layer of silicon was evaporated on top of the chromium before deposition of the upper oxide. This layer successfully prevented oxidation of the Cr during the oxide deposition and protected the outer metal–oxide interface during the 900 °C PDA. The oxygen concentration in the metal layer after the 900 °C PDA was reduced by a factor of 2 compared with the unprotected sample.

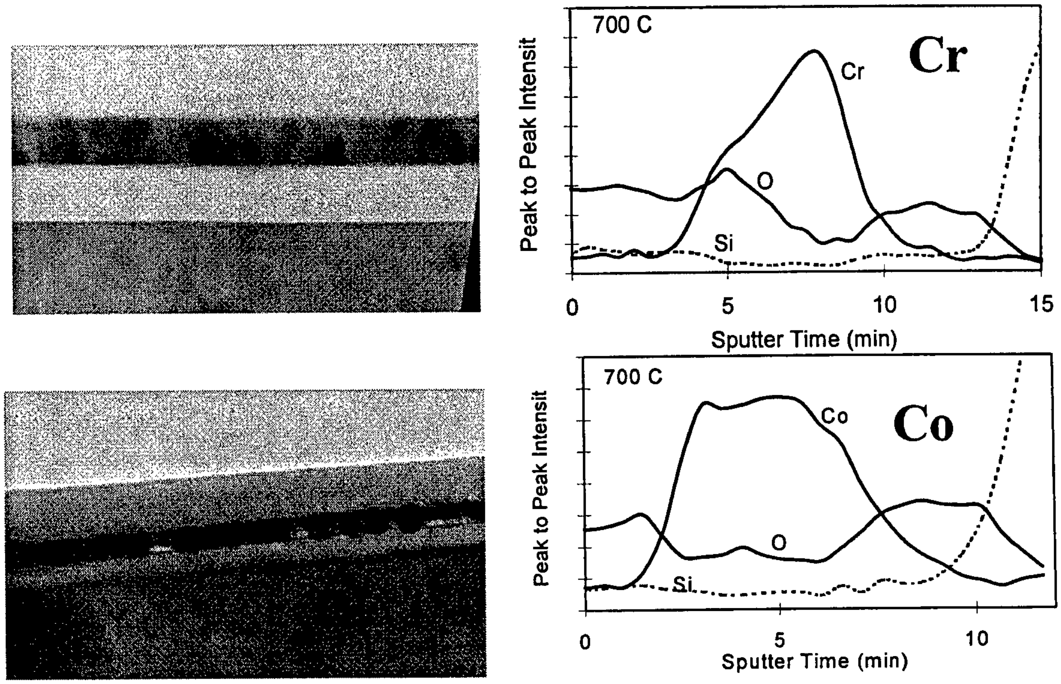


Figure 3. HRXTEM micrograph and AES sputter profile of 700 °C anneal of Cr (top) and Co (bottom) sandwich structures.

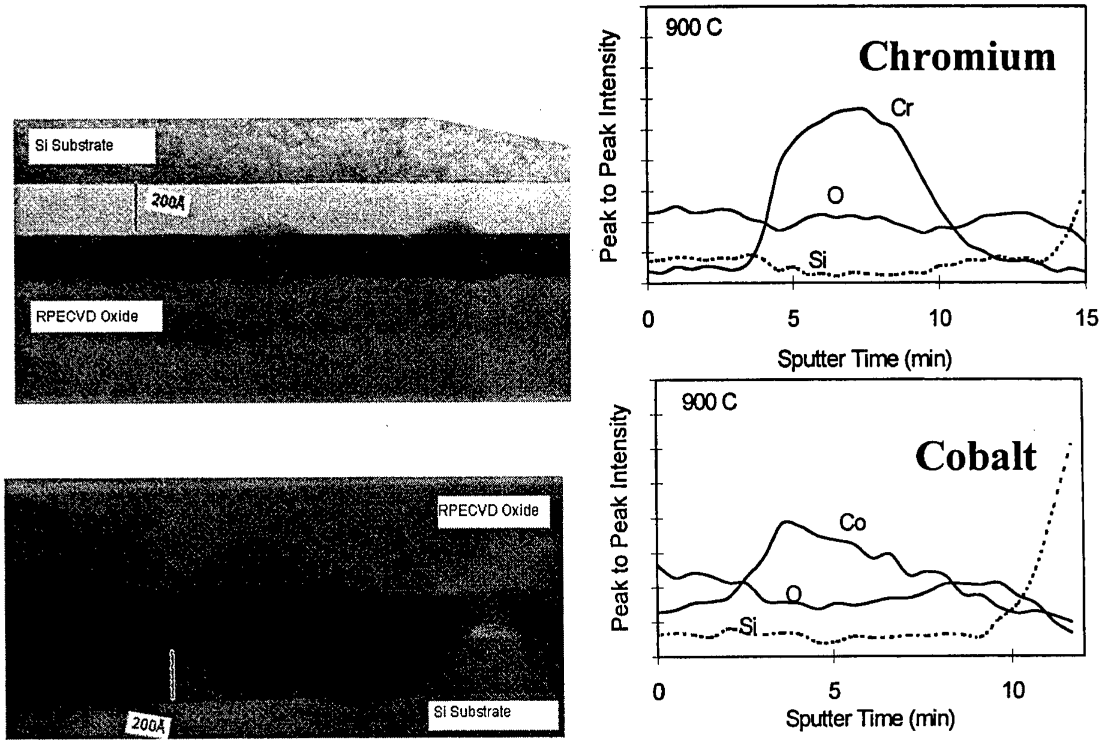


Figure 4. HRXTEM micrograph and AES sputter profile of 900 °C anneal of Cr (top) and Co (bottom) sandwich structures.

4. PDA effects on RPECVD oxide

The beneficial effects of high-temperature annealing on the RPECVD oxide electrical performance are determined from capacitance–voltage and current–voltage measurements made on MOS capacitors. The fixed, mobile and interface oxide charges were reduced by high-temperature annealing. All three were below 3×10^{10} cm^{-2} after the 900 °C anneal. Annealing did not affect the leakage current below a field strength of ~5 MV cm^{-1}, which was too small to measure,

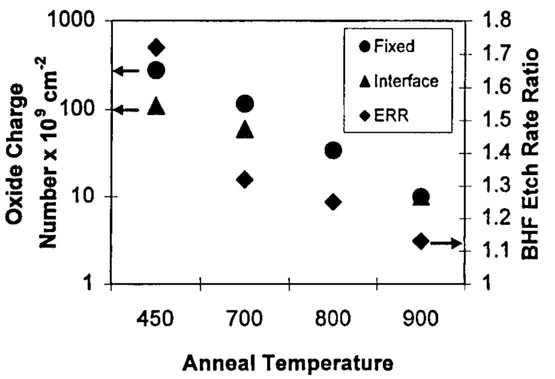


Figure 5. Benefits of high-temperature annealing on RPECVD oxide.

or the average breakdown field of ~8.7 MV cm^{-1}. The etch rate ratio of RPECVD SiO_2 to thermal oxide in 20:1 buffered HF, which is a more bulk-like indicator of oxide integrity, was also reduced with annealing as shown in figure 5.

5. Conclusions

In conclusion, chromium shows considerably more stability than cobalt for the purpose of depletion gates. A protective silicon layer is important to prevent the formation of a metal oxide layer during the RPECVD oxide deposition that might degrade the metal layer if annealed. From this work, it appears that a chromium depletion gate would tolerate a 700 °C PDA and that this anneal would improve the oxide.

References

[1] Vasileska D, Gross W J, Kafedziski V and Ferry D K *VLSI Des.* at press
[2] Batey J and Tierney E 1986 *J. Appl. Phys.* **60** 3136
[3] Lucovsky G, Niimi H, Koh K, Lee D R and Jing Z 1996 *The Physics and Chemistry of* SiO_2 *and the* SiO_2 *Interface—3* (Pennington, NJ: Electrochemical Society) p 441
[4] Wang S Q and Mayer J W 1991 *Thin Solid Films* **202** 105
[5] Pretorius R, Harris J M and Nicolet M-A 1978 *Solid-State Electron.* **21** 667
[6] Seo J M and Lee K H 1996 *Surf. Sci.* **369** 108

Semicond. Sci. Technol. **13** (1998) A75–A78. Printed in the UK

PII: S0268-1242(98)91833-8

A metal/oxide tunnelling transistor

E S Snow, P M Campbell, R W Rendell, F A Buot, D Park, C R K Marrian and R Magno

Naval Research Laboratory, Washington, DC 20375, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We have fabricated a new type of nanometre-scale transistor that operates by using a gate field to modulate the tunnelling probability of electrons through a lateral metal/oxide tunnel junction. Computer simulations predict that such a tunnelling transistor should have operating characteristics similar to those of a Si MOSFET but should be scalable to $\sim$10 nm gate lengths. The device is composed entirely of noncrystalline materials, thus facilitating fabrication on a variety of substrates and multilayer stacking of devices for three-dimensional circuit architectures. Our initial devices have a 40 nm wide Ti/TiO$_x$ tunnel junction on top of a planar Al$_2$O$_3$/Al buried gate. Application of gate bias results in an order of magnitude modulation of the source–drain tunnelling current at 77 K. However, the device transconductance is smaller than predicted by modelling, which we attribute to the gate field not fully penetrating to the active region of the tunnel junction.

The continuous downward size scaling of the conventional Si MOSFET, which has been the primary engine driving the remarkable progress of the microelectronics integrated circuit industry, is expected to fail at a gate length of $\sim$50 nm [1]. Below this size, quantum effects caused by restricted lateral dimensions become manifest in electron transport. Consequently, much effort has been devoted to investigating the utility of such quantum effects for new types of devices to replace conventional Si microelectronics and allow device size scaling beyond the Si MOSFET gate-length limit. Several novel size effects have been investigated, including interference phenomena, conductance quantization and Coulomb blockade effects [2]. Devices based on these phenomena have been fabricated and have been proposed as candidates for large-scale circuit applications [3]. However, even at the present limits of nanolithography, lateral device features are too large for either the quantized confinement energy (in the case of quantized conductance devices) or the capacitive charging energy (for Coulomb blockage) to dominate the thermal broadening ($\sim kT$) at practical operating temperatures, thus rendering them unfeasible for all but a few specialized applications. In addition, even if structures with sub 10 nm features could be made and room-temperature operation realized, device properties based on these principles would be far too sensitive to size variations of even a few atomic widths to be useful for large-scale integration, which requires large numbers of devices with precisely controlled and essentially identical operating characteristics.

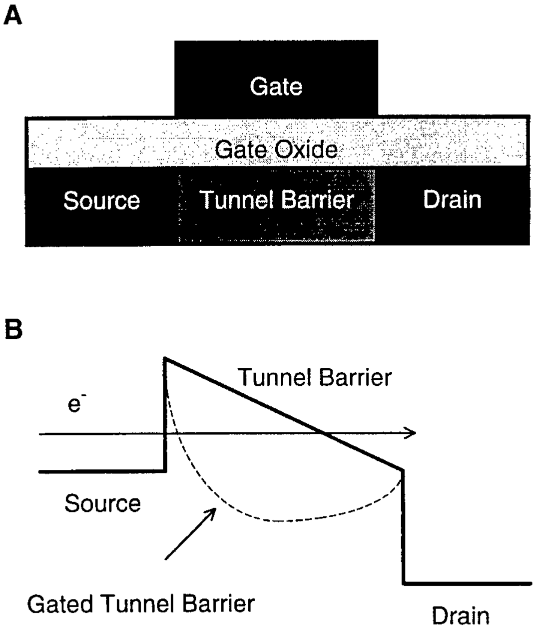


Figure 1. Schematic of (*a*) the device structure and (*b*) the energy diagram of a tunnelling transistor. The application of a gate voltage modulates the tunnel barrier potential and increases it transmission probability.

A promising solution to these size limit problems is a class of quantum effect devices that use the field generated by an applied gate bias to modulate the transmission probability through a tunnel barrier between the source and drain electrodes [4–6] (see figure 1). Because tunnelling is exponentially dependent on the height and width of the tunnel barrier, a small change in tunnel barrier shape induced by a moderate gate bias can result in a large change in transmission probability. A tunnelling device based on this mechanism should have several advantages over other quantum effect devices. One primary advantage derives from the fact that the maximum operating temperature of a tunnelling transistor is determined by the height of the

0268-1242/98/SA0075+04$19.50 © 1998 IOP Publishing Ltd

tunnel barrier and not by a lithographically defined lateral dimension, as is the case for other quantum effect devices. In the on state of the tunnelling transistor, the transmission through the tunnel barrier is controlled primarily by the applied gate voltage. Thus, the thickness of the gate oxide (whose growth or deposition can be controlled to a precision of a single monolayer) and not a lithographically defined lateral feature is the critical dimension that controls the tunnelling. Furthermore, numerical simulations predict that these devices should have operating characteristics similar to those of conventional Si MOSFETs, so that new circuit architectures are not required for their immediate implementation [4]. These features make the lateral tunnelling transistor a promising quantum effect device for integrated circuit applications, with the possibility of scaling well below the Si MOSFET size limit.

This concept of the gate-modulated tunnelling transistor has recently been realized in the form of a 1 μm gate-length Schottky-barrier MOSFET with metal silicides constituting the source and drain and the tunnel barrier formed by the Schottky barriers at the Si/silicide heterojunctions [7]. The current through the source Schottky barrier is modulated by a poly-Si gate. Device operation and numerical modelling predict performance characteristics (voltage thresholds, peak conductance, etc) similar to conventional Si MOSFETs, but the gate length should be scalable to ~10 nm [5]. This device design provides built-in isolation (with the carriers confined by the Schottky barriers), and the near-metallic conductivity of the source and drain permits much smaller contacts. Because most of the area in a conventional MOSFET is devoted not to active area but to isolation and ohmic contacts, the overall area of this device (compared with a conventional MOSFET with equivalent gate size and standard n^+ or p^+ source and drain) can be reduced dramatically, yielding a substantial economy of chip real estate of critical importance for high-density integration [8].

While the Schottky-barrier MOSFET has provided a proof of principle of the relevant device physics for the gate-modulated tunnelling transistor, replacing the silicide/Si Schottky barriers with a metal/oxide/metal tunnel junction offers several potential advantages. In the metal/oxide version of this device, a metal gate over a high-barrier dielectric modulates the transmission through a lateral metal/oxide/metal tunnel junction. Like the Schottky barrier MOSFET, this design also possesses the same economy of scale arising from the built-in isolation and high conductivity of the metallic source and drain, but offers several additional advantages over semiconductor-based devices. The most important advantage arises from the fact that the metal/oxide system is noncrystalline, which considerably relaxes the restrictions imposed on Si and other crystalline semiconductor device technologies, thus allowing an arbitrary choice of substrates and the possibility of multiple layers of active devices. Such devices could form the basis for three-dimensional circuits, for example, for neural network architectures or for reducing interconnect delays for highly integrated circuits.

In this paper we describe our initial effort to fabricate a metal/oxide lateral tunnelling transistor. We observe transistor action at 77 K, although the device transconductance is much smaller than our numerical modelling predicts. In addition, we identify the materials issues that limit the performance of the current device and offer direction for new research and development.

The devices reported in this paper were fabricated by using Al/Al_2O_3 for the gate/gate oxide and Ti/TiO_x the source–drain/tunnel barrier. Al_2O_3 was chosen for the gate oxide because of its large barrier for electron tunnelling, ~2 eV, and large dielectric constant, $\varepsilon_g = 10$. TiO_x was chosen for the lateral tunnel barrier because of its moderate barrier height (0.1–0.5 eV depending on x) and demonstrated performance as a lateral tunnel junction [9–12]. The first step of the fabrication was to deposit a 40 nm thick film of Al on a thermal SiO_2 layer grown on an Si wafer. Next, a gate oxide layer (gate oxide thickness 15 nm) was grown by anodic oxidation in an electrochemical cell over the entire wafer. Photolithography was then used to pattern a photoresist layer on the wafer surface in the shape of the gates and gate contact pads. This pattern serves as a mask for a second anodization in the electrochemical cell which oxidizes all the remaining Al unprotected by the photoresist. This step produces isolated Al gates buried beneath a planar gate oxide upon which we deposit the Ti source–drain wires.

We then patterned 10 nm thick Ti source–drain wires overlapping the gates by using optical lithography and metal liftoff. In the region of the gate contact pads the 15 nm thick Al_2O_3 layer was removed to expose the Al, and thick Au bond pads were deposited on the source, drain and gate contact pads. At this stage the wafer was sawed into chips, the chips attached to headers and the devices bonded.

The lateral TiO_x tunnel junctions were fabricated by using AFM anodic oxidation. For this step the bonded devices were mounted in an ambient AFM instrument and an electrically conducting Si AFM tip operated in ~40% humidity was used to anodize a 40 nm wide region completely through the Ti source–drain wire over the top of the Al gate [9–12]. This last step forms the lateral Ti/TiO_x/Ti tunnel junction and completes the device. A cross-sectional schematic of the device together with optical and AFM images of a completed device are shown in figure 2. Note in the AFM image that the gate layer is not completely planar. The anodization of the Al produces some swelling which delineates the submerged Al gate finger. This topography was used to align the AFM tip to the underlying gate electrode during the anodization.

Capacitance and temperature-dependent current–voltage measurements were used to determine the values of the dielectric constants ($\varepsilon_g = 10$ and $\varepsilon_t = 30$) and barrier heights ($\Phi_g = 2.0$ eV and $\Phi_t = 0.1$ eV) of the gate and tunnel barrier oxides, respectively. (The capacitance measurements were performed on large-area planar samples. The effective barrier height for Ti/TiO_x tunnel junctions ranges from ~0.1 to 0.5 eV depending on the method of fabrication and the post-fabrication anneal conditions.) The use of optical lithography to define the gate metal restricted the gate metal size to ~1 μm or greater, which is significantly larger than the 40 nm tunnel barrier width. This large buried gate metal finger produces a large parasitic gate capacitance; however,

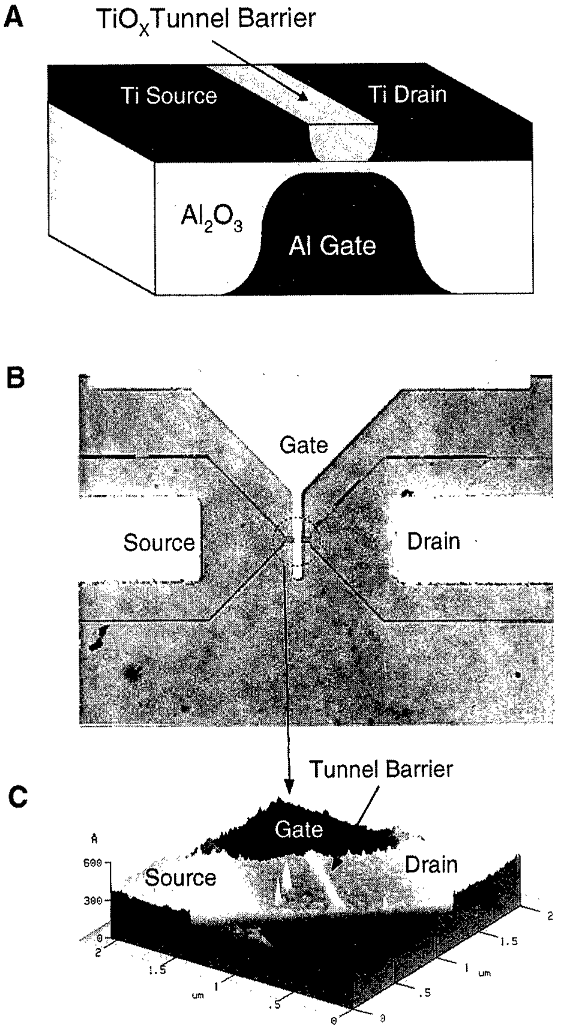


Figure 2. (*a*) Schematic of a Ti/TiO_x tunnel junction fabricated on top of a planar Al/Al_2O_3 gate. (*b*) An optical micrography of the top view of such a device. (*c*) A 2 μm × 2 μm AFM image of the tunnel junction region. The Al gate lies under the depressed region of the image.

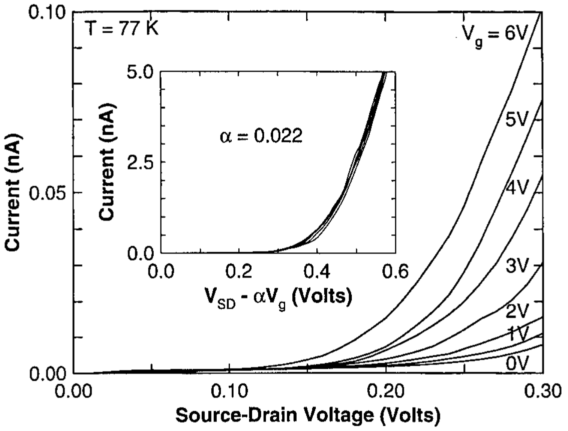


Figure 3. The source–drain current plotted as a function of source–drain voltage for various gate voltages. The inset to the figure shows the same *I–V* curves except that the voltage has been shifted by an amount αV_g, where $\alpha = 0.022$.

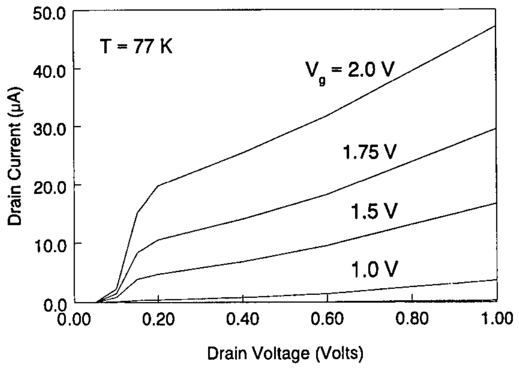


Figure 4. Calculated source–drain current plotted as a function of source–drain voltage.

this is not an issue here because in this paper we treat only the quasistatic device characteristics and not high-frequency operation. The devices discussed in this paper were optimized not for high-frequency performance but for ease of fabrication, with the primary goal of identifying and exploring problems with candidate material systems; other performance considerations were considered secondary to these goals.

Figure 3 shows representative quasistatic *I–V* characteristics of one such device measured at 77 K. The zero-gate-bias source–drain current is well fitted by Fowler–Nordheim tunnelling by assuming a tunnel barrier height of 0.1 eV [13]. Application of a positive gate bias sharpens the triangular Fowler–Nordheim potential and produces an increase in the source–drain current. As seen in figure 3, on application of a 5 V gate bias we observe an order of magnitude increase in the source–drain current with negligible gate leakage current. Under the same bias conditions about a two-to-one modulation of the source–drain current is observed at room temperature. The reduced ratio of current modulation at room temperature is due to an increase in parasitic (i.e. nongateable) thermionic emission over the source–drain tunnel barrier.

The measured characteristics of these devices clearly indicate gate modulation of the source–drain tunnel current; however, the magnitude of the source–drain current modulation is smaller than our numerical device modelling suggests. We have used the measured oxide properties and the known device geometry to calculate numerically the expected device characteristics. The electrostatic potential within the device under application of the source–drain and gate biases for the device geometry shown in figure 1 was calculated by using a fast Poisson solver which utilizes a Fourier analysis and cyclic reduction technique [14]. The calculated potential in the region of the tunnel barrier was used to obtain the tunnelling current. The tunnel current was calculated by using a barrier penetration probability based on WKB and Fermi functions appropriate for the Ti electrodes assuming a barrier height of 0.1 eV above the Fermi level [15]. The results, shown in figure 4, predict a much more effective modulation of the tunnelling current than is observed experimentally.

Evidence for the cause of the smaller than expected transconductance is gleaned from the following observation. In the inset to figure 3 we plot the same transistor curves, $I_{sd}(V')$, with the exception of the voltage, V', has been offset by an amount proportional to the gate voltage, i.e. $V' = V_{sd} + \alpha V_g$ where $\alpha = 0.022$. The good overlap of these curves indicates that application of a gate voltage V_g is equivalent to adding a drain bias of αV_g. Because the source–drain current is well fitted by Fowler–Nordheim tunnelling, the tunnel current is controlled by the electric field at the source electrode. An additional drain bias of αV_g would add an electric field, $F = \alpha V_g / L_t$, to the source electrode. Thus, the effect of the gate is to apply an electric field of magnitude $\alpha V_g / L_t$ to the source electrode. This field value is ~50 times smaller than the value predicted by the numerical simulation. This reduced value of the gate field, coupled with the exponential field dependence of Fowler–Nordheim tunnelling, accounts for the small transconductance of the initial devices.

There are a number of likely causes for the reduced effectiveness of the gate, including the possibilities that the gate field may be screened from the source electrode by non-anodized Ti, the device geometry may be different from what was expected or the active region of the tunnel junction may be more remote than expected from the gate dielectric interface. We are currently testing these hypotheses and devising solutions to each in order to increase the device transconductance.

While other quantum devices such as the single-electron transistor have received much attention from the nanoelectronics community, metals-based tunnelling transistors have not been explored in spite of their potentially unique and advantageous properties. The major obstacle faced by a metal/oxide-based device technology arises from the fact that very little research has been directed at materials appropriate for this type of device or towards the fabrication of high-quality lateral metal/insulator/metal tunnel junctions. Clearly, major issues on device design, fabrication and materials development must be addressed before the metal/oxide tunnelling transistor is considered a serious candidate for real applications. These challenges notwithstanding, the potential advantages of this device, combined with the lack of a suitable knowledge base in these critical areas, merit further research. We are currently exploring a variety of alternative techniques for fabricating lateral tunnel junctions (such as thermal oxidation etc) in order to optimize the performance of these devices. We expect that an optimum device design coupled with high-quality materials and fabrication processes will produce a quantum-effect transistor with room-temperature operation and characteristics similar to conventional Si MOSFETs but scalable to ~10 nm gate lengths, with the additional capability for three-dimensional integration.

Acknowledgments

This work was supported by the Defense Advanced Research Projects Agency and the Office of Naval Research.

References

[1] Taur Y *et al* 1997 *Proc. IEEE* **85** 486

[2] Cahay M and Bandyopadhyay S 1994 *Adv. Electron. Electron Phys.* **89** 93

[3] Averin D V and Likharev K K 1992 *Single Charge Tunneling* ed H Grabert and M Devoret (New York: Plenum) ch 9

[4] Tucker J R, Wang C and Carney P S 1994 *Appl. Phys. Lett.* **65** 618

[5] Wang C K, Zhang W E and Yang C H 1998 *IEEE Trans. Electron. Devices* **45** 842

[6] Fujimara K and Matsumura H 1996 *Japan. J. Appl. Phys.* **35** 2090

[7] Snyder J P, Helms C R and Nishi Y 1995 *Appl. Phys. Lett.* **67** 1420

[8] Tucker J R, Wang C and Shen T-C 1996 *Nanotechnology* **7** 275

[9] Snow E S, Park D and Campbell P M 1995 *Science* **270** 1639

[10] Snow E S, Park D and Campbell P M 1997 *Appl. Phys. Lett.* **69** 269

[11] Matsumoto K, Ishii M, Segawa K, Oka Y, Vartanian B J and Harris J S 1996 *Appl. Phys. Lett.* **68** 34

[12] Irmer B, Kehrle M, Lorenz H and Kotthaus J P 1997 *Appl. Phys. Lett.* **71** 1733

[13] Koehler M and Hummelgen I A 1997 *Appl. Phys. Lett.* **70** 3254

[14] Beard S J and Hockney R W 1985 *Comput. Phys. Commun.* **36** 25

[15] Duke C B 1969 *Tunneling in Solids* (New York: Academic)

Semicond. Sci. Technol. **13** (1998) A79–A82. Printed in the UK PII: S0268-1242(98)91834-X

Nanolithography by non-contact AFM-induced local oxidation: fabrication of tunnelling barriers suitable for single-electron devices

B Irmer, M Kehrle, H Lorenz and J P Kotthaus

Sektion Physik, Ludwig-Maximilians Universität München, 80539 München, Germany

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We study local oxidation induced by dynamic atomic force microscopy (AFM), commonly called tapping mode AFM. This minimizes the field-induced forces, which cause the tip to blunt, and enables us to use very fine tips. We are able to fabricate Ti–TiO$_x$ line grids with 18 nm period and well-defined isolating barriers as small as 15 nm. These junctions show a non-linear current–voltage characteristic and an exponential dependence of the conductance on the oxide width, indicating tunnelling as the dominant conduction mechanism. From the conductance–barrier width dependence we derive a barrier height of $\Phi = 178$ meV. Numerical calculations of the lateral field distribution for different tip geometries allow us to design the optimum tip for the most localized electric field. The electron beam deposition technique makes it possible to produce tips of the desired geometry.

Proximal probe based lithography has developed over recent years into a well-established tool for the fabrication of structures and electronic devices with nanometric dimensions. In particular, the tip-induced oxidation or, more generally, tip-induced local chemical reactions have been very successful [1–5] and appear to be one of the most promising approaches: they preserve the high lateral resolution of the scanned tip by omitting a subsequent etching step and thus avoid the problem of transferring the pattern into an underlying electron system, e.g. metallic thin films or heterostructures. Furthermore, this approach enables one to monitor the process *in situ* by measuring electronic properties, e.g. the conductance of a thin channel, defined and constricted by atomic force microscopy (AFM) induced oxide [6], or the formation of a barrier across a conducting channel [7]. To use this process optimally in nanofabrication requires (1) the understanding of the underlying electrochemical mechanism and the parameters that control it, (2) a tip, which is optimized for laterally focusing the electric field strength under the experimental conditions, and (3) a scanning technique which allows one to use these tips and to retain their properties. Here, we show that operating the atomic force microscope in a dynamic, non-contact mode is suitable for inducing local oxidation. Hereby the damage to the tip is reduced significantly and it allows us to address questions involving the importance of the tip radius and the overall geometry of the tip.

We start with thermally oxidized (250 nm SiO$_2$) n-type (10 Ω cm) Si (100) material, on top of which 30–50 Å titanium is thermally evaporated at high evaporation rates ($\approx$10 Å s^{-1}) and low background pressure ($p \leq 10^{-8}$ mbar). This metallic film is then patterned using optical lithography and an HF wet etch and finally wire bonded. Local oxidation is performed using a commercial atomic force microscope (Digital Instruments) and highly doped n$^+$-Si tips (NanoSensors), which we further sharpen by oxidation. Tip radii are well below 100 Å, typically around 50 Å. The relative humidity is kept constant at 40% during experiments shown here. The cantilever oscillates near its resonance frequency (approximately 250 kHz) with high amplitudes (10–100 nm). The applied tip bias for local oxidation induces additional charges on the tip, which bends the cantilever towards the surface. This force adds to the normal loading force and can easily damage either the tip or the surface. Moreover, in dynamic AFM the force gradient $\partial F/\partial z$ due to the electric field changes the force constant k to $k^\star = k - \partial F/\partial z$, shifting the resonance frequency to $\omega_0^\star = (k^\star/m)^{1/2}$. The driving bimorph oscillates unchanged at the fixed frequency $\omega < \omega_0^\star < \omega_0$, and therefore the oscillation amplitude is increased. The change in amplitude for a given tip bias can be easily measured from amplitude versus distance curves, which then can be used to readjust the working setpoint. As the feedback is enabled all the time and the damping of the amplitude does not change if only the setpoint is readjusted, the overall loading force

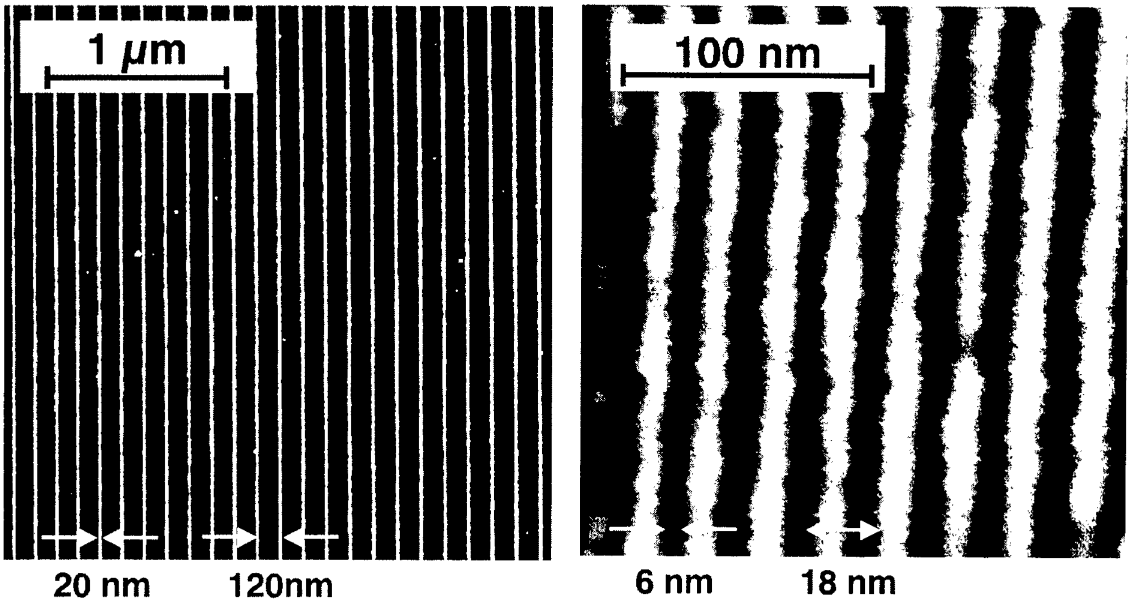


Figure 1. Two Ti–TiO$_x$ grids written by tapping mode AFM-induced local oxidation at −6.5 V tip bias and a scan speed of 300 nm s^{-1}. At room temperature, the resistance parallel to the lines is 100 kΩ and perpendicular to them ≫80 MΩ.

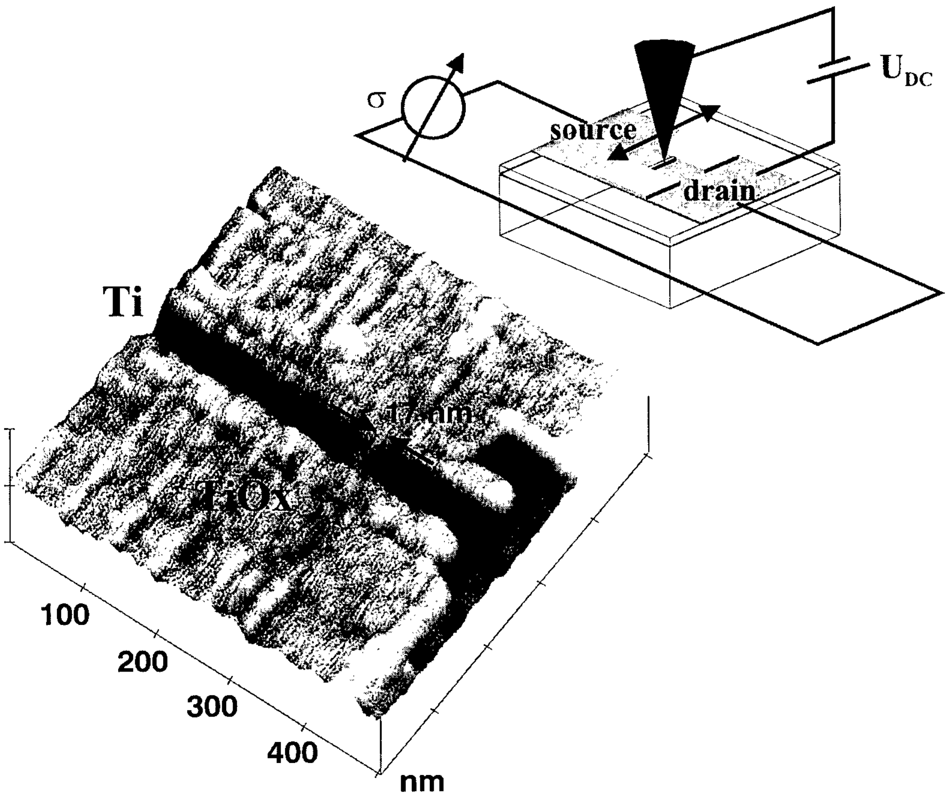


Figure 2. *In situ* control of the barrier formation. The source–drain conductance through the device is monitored while oxidizing. The tip is biased at −4 V and repeatedly scanned at 2 Hz across the 30 nm wide Ti channel to form a 17 nm wide barrier.

remains unchanged even for applied voltages up to 30 V [7]. In figure 1 we show two grids of oxide lines written at a rate of 300 nm s^{-1} at a tip bias of −6.5 V. The lines are very regular in width (18–20 nm) as well as in height, even for relatively large scan fields of 3 mm × 3 mm and above. If the period is reduced from 120 nm to 23 nm, parallel conducting wires of 6 nm linewidth are formed, which are still conductive along the channels (≈100 kΩ)

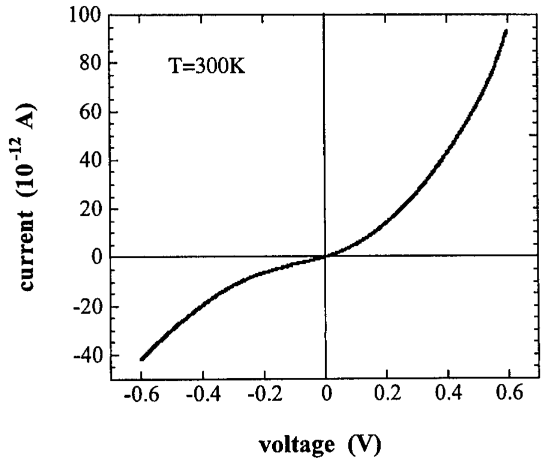


Figure 3. Room temperature current–voltage characteristic of a 20 nm wide barrier.

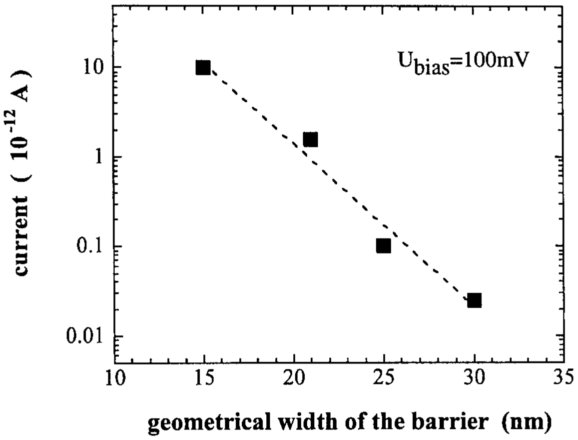


Figure 4. Dependence of the current on the geometrical barrier width. Four different devices are measured at 100 mV bias and $T = 300$ K. The current depends exponentially on the barrier width, indicating tunnelling as the dominant conduction mechanism.

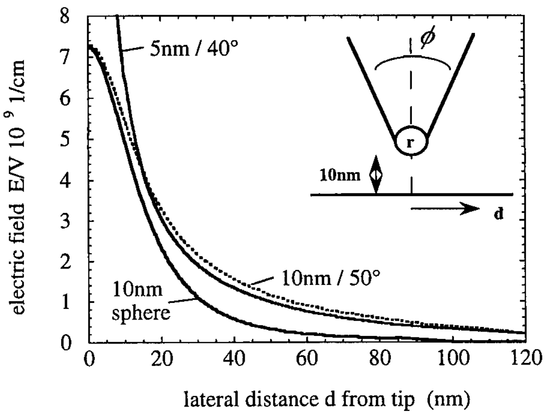


Figure 5. Calculated electric field distribution for three tips with different (r, ϕ) geometries, 10 nm in front of a conductive plane in vacuum: 10 nm sphere and pyramidal tips with 10 nm radius and 50° cone angle and 5 nm radius and 40° cone angle.

but isolating in the perpendicular direction ($\gg$80 MΩ, at room temperature). This demonstrates a complete oxidation process for the most of the average of the oxide lines. The observed oxide height is 3 nm, which agrees very well with what is expected from the change in density and molecular weight: $d_{TiO_2}/d_{Ti} = (\rho_{Ti}/\rho_{TiO_2})M_{TiO_2}/M_{Ti} \approx$ 3 nm/5 nm. It should be noted that, at a tapping frequency of $f_T = 250$ kHz and the observed damping of the amplitude, the contact time $t = 1/2f_T$ between the tip and the sample surface per cycle is below 10^{-3} ms. As the oscillation amplitude is very large (10–100 nm), it is unlikely that a stable water meniscus forms between tip and sample. Evidence for this is provided by force versus distance curves in contact AFM using the same cantilevers (data not shown). If the experiment is to be explained in a classical electrochemical set-up, wherein the tip acts as cathode, the water film as electrolyte and the sample as anode, the total exposure time is much shorter than in contact AFM. However, the total amount of oxidized material is very much the same as seen by contact AFM, e.g. by Avouris *et al* [8] for Si or by Sugimura *et al* [9] for Ti. We therefore conclude that corrosion has taken place at the Ti–TiO$_x$ interface, enhanced by the tip–sample electric field in the presence of humidity.

To define a tunnelling barrier we first constrict a predefined 1 mm Ti wire by oxidizing two large oxide pads, enclosing a 30 nm wide channel (figure 2). The barrier, perpendicular to the channel, is then oxidized at 2 Hz scanning frequency and a tip bias of −4.5 V. In order to avoid the formation of too thick barriers with a too small tunnelling probability by overexposure, we monitor the conductance along the channel. As soon as the conductance drops below the capacitive signal, oxidation is stopped. The quality of the AFM-induced oxide is characterized on wide barriers ($\approx$100 nm). Resistivities of $\rho = 2 \times 10^{11}$ Ω cm and maximum field strength $V_D = 2 \times 10^6$ V cm^{-1} are measured. These values are similar to those for macroscopic anodic oxides [10, 11].

At room temperature the devices show an asymmetric, non-linear I–V characteristic (figure 3). This may be understood in the picture of an asymmetrical, shallow barrier, which is no longer isotropic for forward and reverse biases. To determine the conduction mechanism for these devices, we investigate the dependence of the (tunnelling) current on the geometrical width of the barrier as obtained from AFM images. For four different devices, with barriers varying from 15 nm to 30 nm, the current decays exponentially with barrier width, indicating tunnelling as the dominant conduction mechanism (figure 4). We obtain here a barrier height of $\Phi = 178$ meV.

To determine the parameters that affect the lateral resolution of the oxidation and therefore to estimate the ultimate limit for this technique, we model the lateral field distribution for different tip geometries, namely tip radii and cone angles. In a first step we place the tip 10 nm in front of a conducting surface. The calculated electric field for a 10 nm sphere (radius in each case), a 10 nm spherical tip with 50° pyramidal cone and a 5 nm spherical tip with 40° cone are shown in figure 5. As expected, the pyramids widen the lateral field compared with the free-

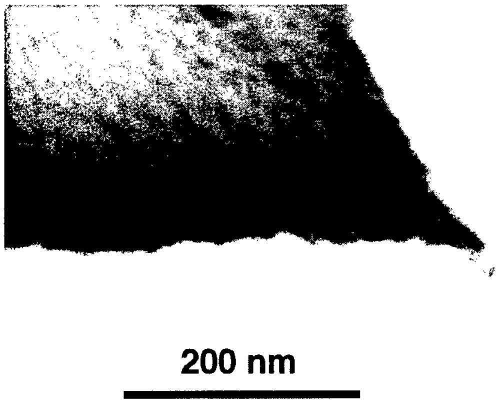


Figure 6. SEM image of an EBD tip, deposited onto a commercial Si tip coated with NiCr. From this image, the radius is ≈5 nm and the cone angle 5°. The tip shows a resistance of $R \leq 1$ MΩ, which is sufficient for local oxidation.

standing sphere, whereas smaller spheres increase the local field underneath the tip. At this stage we do not consider the growing oxide itself as well as the focusing effect of the water layer or meniscus because of its large ϵ. However, for an optimized focusing of the lateral fields, we would like to have a needle-like tip, which is still sufficiently conductive. So-called electron beam deposited (EBD) material is known to be suitable to define scanning tips with tip radii ≤ 5 nm and very small cone angles [12]. If deposited at high electron energies and low beam current densities, they appear to be conductive. Figure 6 shows an SEM image of an EBD tip deposited on top of an NiCr-coated Si tip. This tip shows an overall resistance $R \leq 1$ MΩ, which is sufficient for applications in electrochemical AFM and local oxidation and gives, in contrast to for example carbon nanotubes, the unique possibility to design the tip to exactly the requested geometry.

In summary, non-contact AFM has been used for locally oxidizing titanium thin films. In this mode, the tip–sample forces remain unchanged when a tip–sample bias is applied. This allows us to use oxide-sharpened Si tips, with which we are able to fabricate line grids with 6 nm structure sizes and 18 nm pitch. *In situ* electrical measurements give fine control over the lithographic process. In this way we fabricated tunnelling barriers as small as 15 nm. The current–voltage characteristic and the dependence of current on barrier width clearly indicate that tunnelling is the dominant transport mechanism in these devices. Numerical calculations of the lateral distribution of the tip-to-sample electric field indicate a further improvement in the lithographic resolution, if only needle-like tips with small radii and small cone angles are used. We show that 5 nm radius, 5° cone angle EBD tips are sufficiently conductive to be used for local oxidation.

Acknowledgments

The authors would like to acknowledge the contributions of S Manus and A Kriele. This work was supported financially by the BMBF and by the Volkswagen-Stiftung, which we gratefully acknowledge.

References

[1] Dagata J A, Schneir J, Harary H H, Evans C J, Postek M T and Benett J 1990 *Appl. Phys. Lett.* **56** 2001

[2] Snow E S, Campbell F M and McMarr P J 1993 *Appl. Phys. Lett.* **63** 749

[3] Snow E S and Campbell P M 1994 *Appl. Phys. Lett.* **64** 1932

[4] Matsumoto K, Ishii M, Segawa K and Oka Y 1996 *Appl. Phys. Lett.* **68** 34

[5] Shen T-C, Wang C, Lyding J W and Tucker J R 1995 *Appl. Phys. Lett.* **66** 976

[6] Snow E S and Campbell P M 1995 *Science* **270** 1639

[7] Irmer B, Kehrle M, Lorenz H and Kotthaus J P 1997 *Appl. Phys. Lett.* **71** 1733

[8] Avouris P, Hertel T and Martel R 1997 *Appl. Phys. Lett.* **71** 285

[9] Sugimura H, UchidaT, Kitamura N and Masuhara H 1993 *Appl. Phys. Lett.* **63** 1288

[10] Ghandi S K 1983 *VLSI Fabrication Principles* (New York: Wiley)

[11] Lide D R (ed) 1992 *CRC Handbook of Chemistry and Physics* 72nd edn (Boston, MA: CRC)

[12] Wendel W, Lorenz H and Kotthaus J P 1995 *Appl. Phys. Lett.* **67** 3732

Semicond. Sci. Technol. **13** (1998) A83–A85. Printed in the UK

PII: S0268-1242(98)91845-4

Measuring the mechanical resonance of a GaAs/AlGaAs cantilever using a strain-sensing field-effect transistor

R G Beck†, M A Eriksson†, R M Westervelt†, K D Maranowski‡ and A C Gossard‡

† Department of Physics and Division of Applied Sciences, Harvard University, Cambridge, MA 02138, USA
‡ Materials Department, University of California, Santa Barbara, CA 93106, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We have fabricated a cantilever with dimensions 0.5 μm × 11.4 μm × 65 μm from a GaAs/$\mathrm{Al_{0.3}Ga_{0.7}As}$ heterostructure containing a two-dimensional electron gas. A strain-sensing field-effect transistor (FET) integrated into the cantilever base acts as a displacement sensor via the piezoelectric effect. The FET was used to measure the cantilever mechanical quality factor $Q = 360$ and resonance frequency $f_{res} = 88.2$ kHz which is close to $f_{res} = 91$ kHz calculated from geometry. FETs can be integrated into GaAs/AlGaAs microelectromechanical systems as force and displacement sensors.

Micro-electromechanical systems (MEMS) provide a novel environment for basic sciences and device applications. Sensing systems for either strain or position are central to the functionality of many MEMS. The most widely used detection methods for strain include resonance frequency monitoring, piezoresistance and piezoelectrics, and for position include optical, thermal and capacitance. Recently, strain sensors have been developed composed of GaAs/AlGaAs heterostructure field-effect transistors (FETs) [1, 2]. We have integrated these transistors into GaAs/$\mathrm{Al_{0.3}Ga_{0.7}As}$ scanning probe microscope (SPM) cantilevers as low-noise deflection sensors [3]. The work presented here is intended to supplement that work by extending the previous results for static deflections to higher-frequency, dynamic operation of the FET. Dynamic mechanical response points toward the operation of SPM cantilevers in non-contact mode and also shows the ability of strain-sensing FETs to collect information on the mechanical response of small systems [4].

In this paper we describe the fabrication of a GaAs/$\mathrm{Al_{0.3}Ga_{0.7}As}$ cantilever with an integrated strain-sensing FET and the operation of the FET for the measurement of the cantilever mechanical resonance frequency. The FET, with a gate 5.8 μm wide by 0.7 μm long, has a small signal transconductance $g_m = 0.3$ mS and drain–source resistance $r_{DS} = 16$ kΩ at a drain current of $I_D = 73$ μA. FET noise power has a $1/f$ spectrum at low frequencies. At 90 kHz, the frequency of interest for this experiment, we measured a gate voltage noise $\delta V_g \sim 450$ nV Hz$^{-1/2}$ corresponding to a nominal gate charge noise $\delta q_g < 0.1$ electrons Hz$^{-1/2}$. The FET was integrated into the base of a GaAs/$\mathrm{Al_{0.3}Ga_{0.7}As}$ cantilever with dimensions 0.5 μm × 11.4 μm × 65 μm that was mounted on a piezoelectric bimorph in vacuum and driven through its mechanical resonance at $T = 4.2$ K. The strain-sensing FET was used to measure the cantilever mechanical quality factor $Q = 360$ and resonance frequency $f_{res} = 88.2$ kHz, which is lower than $f_{res} = 91$ kHz theoretically predicted from the cantilever geometry.

Figure 1(*a*) shows an SEM image of the GaAs/$\mathrm{Al_{0.3}Ga_{0.7}As}$ cantilever with a strain-sensing FET at its base. This structure was fabricated from a wafer containing a two-dimensional electron gas (2DEG) beginning 520 Å from the surface plus a 4000 Å sacrificial layer of AlAs beginning 5000 Å from the surface. The layers composing the cantilever shown in figure 1(*b*) in growth order are 50 Å GaAs, 4010 Å $\mathrm{Al_{0.3}Ga_{0.7}As}$, Si delta-doping layer, 220 Å $\mathrm{Al_{0.3}Ga_{0.7}As}$, 200 Å GaAs, 220 Å $\mathrm{Al_{0.3}Ga_{0.7}As}$, Si delta-doping layer, 250 Å $\mathrm{Al_{0.3}Ga_{0.7}As}$ and 50 Å GaAs. The cantilever and FET were fabricated using four aligned electron-beam lithography steps. First, AuNiGe contacts were thermally evaporated through a resist mask that was then lifted off allowing the contacts to be thermally annealed to make ohmic contact to the 2DEG. Second, the FET channel was partially defined by etch trenches 750 Å deep formed by etching through a patterned mask using a 10:1 solution of 50% citric acid and 30% hydrogen peroxide. Third, the FET gate was formed by thermally evaporating 250 Å Cr and 2000 Å Au through a patterned mask followed by a standard liftoff technique. The resulting gate-to-channel capacitance was calculated to be 9 fF based on geometry. The fourth resist pattern laterally defined the cantilever to the desired dimensions of 11.4 μm wide by

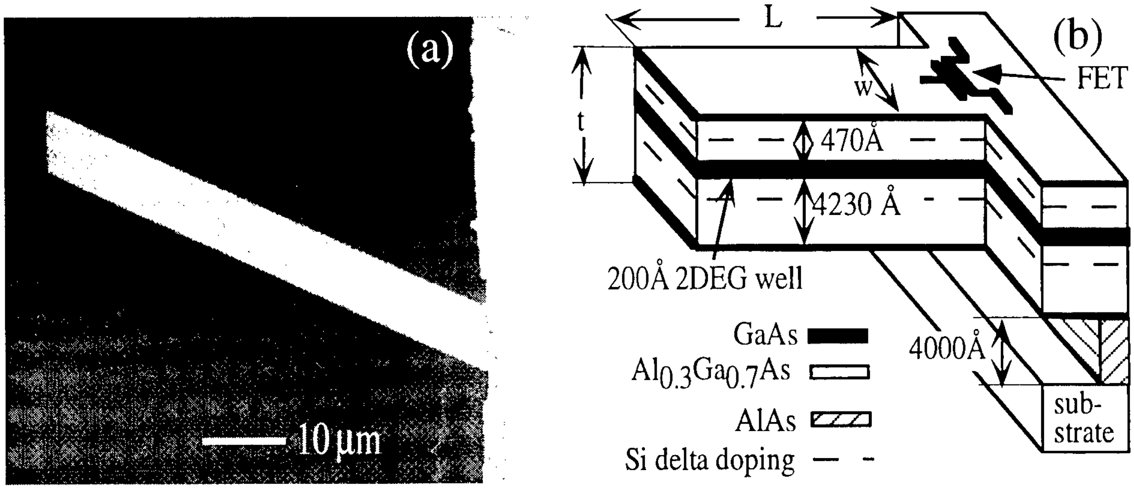
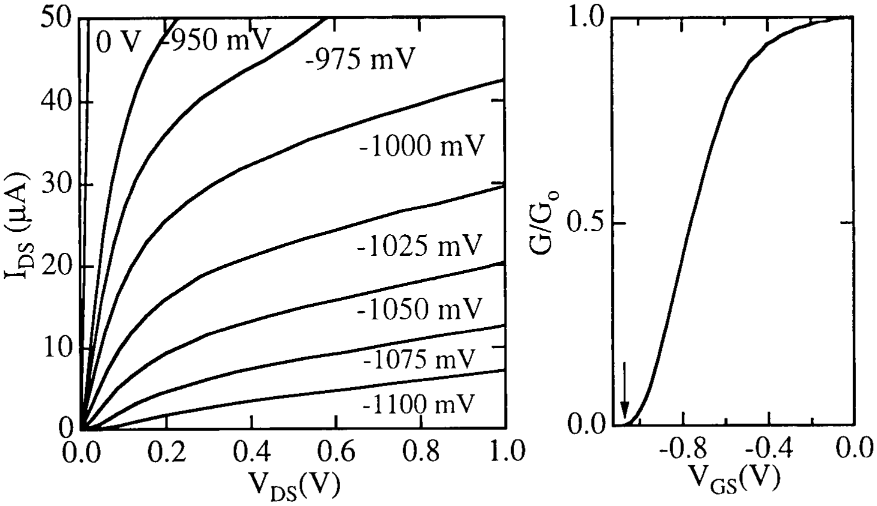


Figure 1. (*a*) Scanning electron microscope image of a GaAs/AlGaAs cantilever with an integrated strain-sensing FET. The FET gate can be seen at the cantilever base where the channel confines electrons to move on the suspended structure through the region of maximum strain. The alignment marker is on the substrate ~ 0.4 μm below the cantilever. (*b*) Schematic showing the layer structure of the cantilever and wafer.



(*a*) (*b*)

Figure 2. (*a*) Family of drain–source curves I_D versus V_{DS} for the on-cantilever FET taken at $T = 4.2$ K at the gate voltages indicated. (*b*) Cutoff graph of channel conductance normalized to its zero gate voltage value of 2.5 mS. The arrow indicates the cutoff voltage of $V_c = -1.05$ V.

65 μm long. With the cantilever and FET protected by resist, a 3 min reactive ion etch of 30 standard cm^3 min^{-1} of BCl_3 and 20 standard cm^3 min^{-1} of $SiCl_4$ at 50 mT pressure, 180 W power and −200 V dc bias was used to expose the underlying AlAs around the cantilever. The cantilever was then freed from the substrate by placing the sample in a 1:5 solution of 49% HF and water to etch selectively the underlying AlAs layer. In order to prevent surface tension from either breaking the cantilever or causing it to stick to the substrate a CO_2 critical point dryer was used to bring the sample through the liquid air interface.

Figure 2(*a*) shows the drain–source characteristics I_D versus V_{DS} of the on-cantilever FET for the series of gate voltages V_{GS} indicated. We measure a small signal transconductance $g_m = 0.3$ mS and drain source resistance $r_{DS} = 16$ kΩ for a drain current of 73 μA. Figure 2(*b*) shows channel conductance normalized to its zero gate voltage value of 2.5 mS as the gate voltage V_{GS} is made more negative. The cutoff voltage $V_{GS} = 1.05$ V is the voltage for which the channel conductance has decreased by a factor of 100. The frequency response of the FET shows that transconductance is constant up to the measured frequency of 100 kHz. The calculated time constant for the FET gives a cutoff frequency [5] $g_m/2\pi C_g \sim 5.3$ GHz where C_g is the gate-to-channel capacitance.

The FET characterization measurements and the following cantilever measurements were taken at $T = 4.2$ K with the sample mounted in vacuum in an Infrared Labs dewar. In order to excite the cantilever through its resonance frequency, the sample was attached to a rigid Macor block that in turn was mounted to a piezoelectric bimorph. The sample was electrically shielded from the

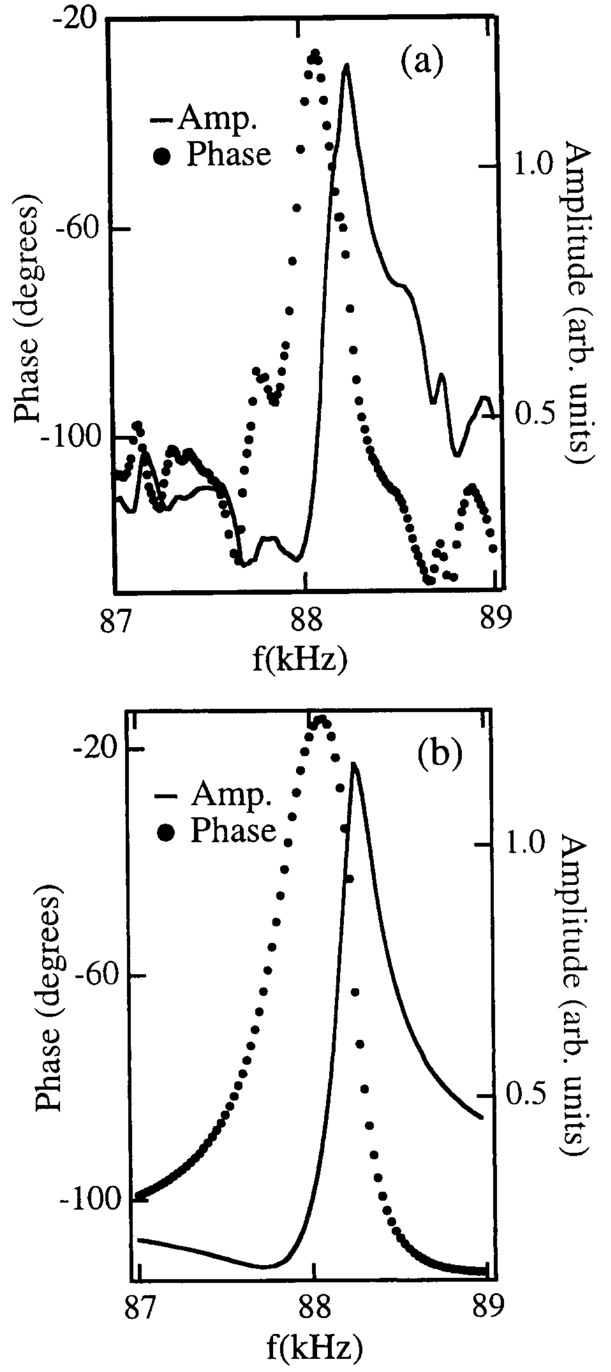


Figure 3. (*a*) Measured FET response as the cantilever is driven through its mechanical resonance. Both amplitude (——) and phase with respect to the piezoelectric driving voltage (•) were recorded. (*b*) Results of a model calculation explaining the shapes of the measured curves by including the cantilever resonance signal and the effect of the capacitive coupling between the bimorph driving circuit and the measurement circuit.

bimorph voltages by means of a grounded mount and protected from unwanted deformation by the rigid block. For the measurements below, the FET was biased with $V_{GS} = -1.0$ V and $V_{DS} = 0.20$ V resulting in a power dissipation of 4.4 μW. Drain current was monitored with an Ithaco 1211 current amplifier and a Stanford Research 530 digital lock-in amplifier. As the cantilever was driven through its mechanical resonance, both drain current amplitude and phase with respect to the bimorph driving voltage were measured.

Figure 3(*a*) shows a graph of FET drain current amplitude and phase with respect to the driving voltage as the cantilever is driven by the piezoelectric bimorph. The peaks in amplitude and phase indicate that the cantilever is passing through its mechanical resonance. The shapes of the curves can be understood with a model involving the strain-induced FET signal added to a constant term due to capacitive coupling between the bimorph driving circuit and the measurement circuit. Figure 3(*b*) shows results of a model calculation where the measured signal amplitude A_s and phase ϕ_s are modelled by

$$A_s \exp(i\phi_s) = \exp(i\phi_{amp})\{A_C \exp(i\pi/2) + A_{res}(\omega) \times \exp[i\phi_{bim} + i\phi_{res}(\omega)]\} \qquad (1)$$

where ϕ_{amp} is the phase shift due to the external amplifier, $A_C \exp(i\pi/2)$ is the capacitive term, $A_{res}(\omega)$ and $\phi_{res}(\omega)$ are the frequency-dependent amplitude and phase of the cantilever and ϕ_{bim} is the phase shift between bimorph and cantilever excitation due to driving the bimorph above its resonance frequency of 400 Hz. A reasonable choice for parameters in equation (1) leads to simulation results similar to measurements as seen in figure 3. Measurements indicate a mechanical quality factor $Q = 360$ and resonance frequency $f_{res} = 88.2$ kHz, which is lower than the theoretically predicted value of 91 kHz obtained from cantilever geometry. This measurement demonstrates that strain-sensing FETs can be used to monitor dynamically the mechanical response of a GaAs/AlGaAs cantilever.

Acknowledgments

The authors would like to thank J G E Harris, D D Awschalom and Joe Tien for help with cantilever fabrication. This work was supported at Harvard in part by the MRSEC program of the National Science Foundation under award DMR-94-00396, ONR award N00014-95-1-0104 and JSEP award N00014-89-J-1023 and at UCSB under award AFOSR F 49620-94-1-0158.

References

[1] Beck R G, Eriksson M A, Westervelt R M, Campman K L and Gossard A C 1996 *Appl. Phys. Lett.* **68** 3763
[2] Fung A K, Cong L, Albrecht J D, Nathan M I and Ruden P P 1997 *J. Appl. Phys.* **81** 502
[3] Beck R G, Eriksson M A, Topinka M, Westervelt R M, Maranowski K D and Gossard A C 1998 *Appl. Phys. Lett.* to be published
[4] Cleland A N and Roukes M L 1996 *App. Phys. Lett.* **69** 2653
[5] Sze S M 1981 *Physics of Semiconductor Devices* 2nd edn (New York: Wiley)

Semicond. Sci. Technol. 13 (1998) A86–A89. Printed in the UK PII: S0268-1242(98)91843-0

Dual-side electron beam lithography for independent submicron gating of double quantum well devices

J R Wendt, J A Simmons, J S Moon, W E Baca, M A Blount and J L Reno

Sandia National Laboratories, Albuquerque, NM 87185-0603, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We describe the first demonstration of dual-side electron beam lithography in achieving independent submicron gating in double quantum well devices. The technique utilizes the epoxy-bond and stop-etch process to remove the substrate material which allows the backside gates to be placed in close proximity (less than 1 μm) to the frontside gates. The use of electron beam lithography allows both the definition of submicron features and the precise alignment of the front and back features to each other. We have applied this technique to the fabrication of double quantum point contacts on coupled AlGaAs/GaAs double quantum wells. Low-temperature transport measurements clearly show the formation of coupled, independently controllable mesoscopic structures in each of the two quantum wells.

1. Introduction

Mesoscopic devices based on coupled double quantum wells (DQWs) are an area rich in the physics of electron transport and in potential electron device applications. Previous work has demonstrated a Coulomb-driven interwell correlation of electrons in a DQW under high magnetic field [1] and control of the coupling between a high-mobility QW and a closely spaced low-mobility QW [2]. The ability to gate two QWs independently, allowing independent ohmic contacts to each, has enabled the investigation of new physical phenomena such as 2D–2D tunnelling [3] and Coulomb drag [4] and has allowed the fabrication of a novel three-terminal quantum device, the double electron layer tunnelling transistor (DELTT) [5]. Earlier work utilized large-area back gates on thick substrates [6] but required large voltages (~100 V) to deplete the bottom quantum well. Furthermore, the thick substrate does not allow the formation of gate-defined submicron features in the bottom quantum well. The ability to backgate DQW structures on a submicron scale enables a wide range of new physics experiments on the interactions between mesoscopic structures defined independently in each quantum well. For example, giant magnetic field induced conductivity enhancements are predicted in coupled double quantum wires [7]. In the field of quantum electron devices, submicron gates with submicron alignment allow fabrication of DELTTs with low parasitics and hence very high speed.

We have previously addressed the problem of a thick substrate through the development of the epoxy-bond and stop-etch (EBASE) technique [8]. With this technique, the substrate is removed down to an AlGaAs stop-etch layer leaving as little as 300 nm of epitaxial layers epoxied to the host substrate, over areas of order 1 cm^2. Until now, the EBASE technique has been demonstrated only in conjunction with optical lithography and its accompanying resolution and alignment limitations. In this work, we demonstrate the first dual-side electron beam lithography in association with the EBASE technique to define submicron gates on each side of a 330 nm thick epitaxial layer. Sub 0.1 μm alignment accuracy between the front and back sets of gates is readily achieved. The combination of the EBASE technique and dual-side electron beam lithography provides a powerful tool for the fabrication of novel mesoscopic structures which were heretofore unavailable. We further demonstrate the utility of this technique by fabricating independently tunable double quantum point contacts (QPCs) in coupled AlGaAs/GaAs DQWs. Characterization of the QPC structures through low-temperature transport measurements clearly shows the formation of coupled, independently controllable mesoscopic structures in each of the two QWs.

2. Fabrication

The fabrication begins with conventional frontside processing using several levels of optical lithography and standard semiconductor processing to define device mesas and to deposit ohmic contacts and Schottky gates. The optically defined Schottky gates lead up to the device mesa and then terminate. The actual device gates are defined by electron beam lithography and are aligned to the optically defined gate leads.

The electron beam lithography was performed with a JEOL JBX-5FE thermal field emission system operating at 50 kV. The beam current used was 1 nA with a corresponding beam diameter of approximately 12 nm. The resist used was polymethylmethacrylate (PMMA) at a thickness of 400 nm. The PMMA was developed for 1 min in a 1:3 solution of methyl isobutyl ketone and isopropyl alcohol. The gates consist of 10 nm Ti/70 nm Au deposited by electron beam evaporation followed by liftoff in acetone. This completes the frontside processing.

The EBASE process has been described previously [8]. The two key steps are the epoxying of the active wafer, with the patterned frontside down, onto a host substrate, followed by the removal of the active wafer substrate down to an AlGaAs stop-etch layer. In this particular work, the AlGaAs stop-etch layer itself was also removed by a second selective etch which terminated on a 15 nm GaAs stop-etch layer. The second selective etch enabled attainment of a total epitaxial layer thickness of only 330 nm, as well as increasing the smoothness of the backside surface.

The key to the dual-side alignment is the use of the same set of metallic alignment marks for both the frontside and the backside electron beam patterning. As part of the frontside processing, a pair of electron beam alignment marks are defined optically. The alignment marks are recessed into the original front surface by wet chemical etching prior to the deposition of the Ti/Au marks. The depth of the recess is chosen such that the bottom of the mark lies close to the stop-etch interface so that the alignment mark will be at or near the surface of the sample at the end of the EBASE process. The exact depth of the recess is not critical in that the marks are easily detected by the 50 kV electron beam even through several hundred nanometres of epitaxial material. A scanning electron micrograph of the cross-section of a portion of an electron beam alignment mark is shown in figure 1. In this case, the epitaxial layer thickness is 1.2 μm and the mark is 300 nm below the sample surface after having been placed in a 900 nm recess. The use of the same set of alignment marks allows for sub 0.1 μm alignment between the frontside and backside patterns. A scanning electron micrograph of the cross-section of a fabricated dual split-gate test structure whose geometry corresponds to that of double quantum wires is shown in figure 2. The alignment can be seen to be nearly perfect for the 500 nm gaps aligned across a 1.2 μm thick epitaxial layer.

The double QPCs were fabricated from MBE-grown epitaxial material consisting of two 20 nm wide GaAs QWs separated by a 1.1 nm $Al_{0.3}Ga_{0.7}As$ barrier. Pairs of Si delta-doping layers of 5×10^{11} and 1×10^{12} cm^{-2} were placed symmetrically 70 and 80 nm, respectively, above and below the DQW, within 130 nm thick $Al_{0.3}Ga_{0.7}As$ layers. The ohmic contacts made electrical contact to both of the QWs. After the second selective etch, the remaining epitaxial layer is completely symmetric from front to back and therefore the final DQW structure is expected to be balanced. The QPCs are formed by mutually aligned pairs of sharply pointed split gates with point-to-point spacings ranging from 250 to 400 nm.

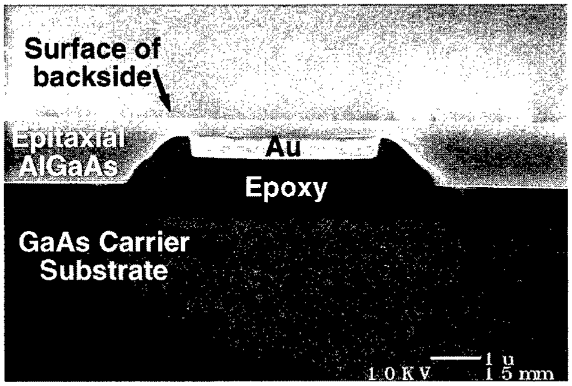


Figure 1. Scanning electron micrograph of the cross-section of a portion of the alignment mark used in electron beam writing. The mark is recessed into the original frontside of the active wafer to improve detectability from the backside.

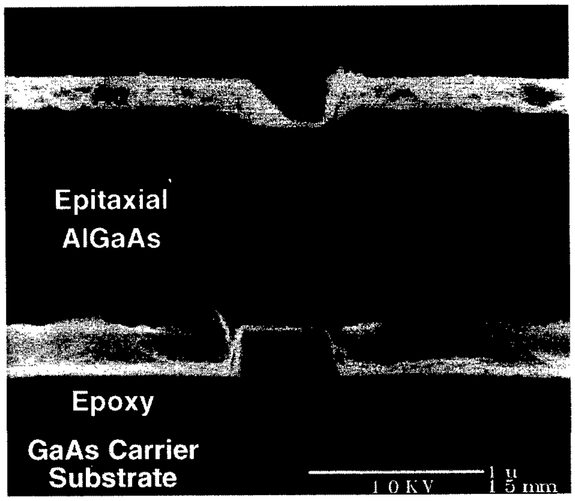


Figure 2. Scanning electron micrograph of the cross-section of a fabricated dual split-gate test structure having 500 nm gaps aligned across a 1.2 μm epitaxial layer. The rough upper edges of the lower Au gates are an artifact of cleaving.

3. Measurements

Electrical transport measurements on the double QPC samples were performed at 0.3 K in a top-loading pumped ^{3}He system using standard low-frequency lock-in techniques. The longitudinal resistance, R_{XX}, of the adjacent Hall bar was measured as a function of perpendicular magnetic field, B. The R_{XX} minima closely approached zero at selected integer Landau level filling factors and the strength of the minima follow the typical beating pattern observed in double-layer systems [9]. The Fourier power in $1/B$ of R_{XX} reveals two clear peaks at electron densities, n, of 2.9×10^{11} and 1.45×10^{11} cm^{-2} corresponding to the occupancies of the symmetric and antisymmetric subbands of the DQW, respectively. This

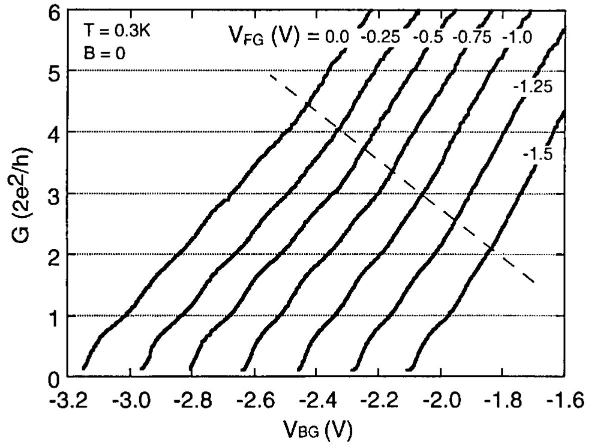


Figure 3. Plots of conductance versus front-gate pair bias, V_{FG}, at several values of back-gate pair bias, V_{BG}, for a double QPC with 400 nm gaps. The curves are not offset from one another. The dotted line indicates the position where the bottom QPC pinches off.

difference in density yields a symmetric–antisymmetric energy gap $\Delta_{SAS} = 5.2$ meV, in good agreement with the results of Hartree simulations of the growth structure [10]. The conductance as a function of in-plane magnetic field exhibited two strong anticrossing features whose positions are consistent with the density and Δ_{SAS} values given above [10]. The total mobility of the structure is quite high at 5.8×10^5 cm^2 V^{-1} s^{-1}, not significantly changed from the pre-EBASE processing value, which demonstrates that the EBASE technique does not degrade the quality of the two-dimensional electron gas.

We now discuss the behaviour of the 400 nm wide double QPC device. By adjusting the bias on the pair of front gates, V_{FG}, and on the pair of back gates, V_{BG}, four different situations can be realized: (1) both QPCs open, (2) only top QPC open, (3) only bottom QPC open and (4) neither QPC open. In figure 3 we show the conductance G of the device in units of $2e^2/h$ as a function of V_{BG} at several different V_{FG} values. (Because the growth structure is symmetric about the barrier, and the top and back gates are an equal distance from the DQW, we expect the front and back pairs of gates to have an equally strong effect on their own quantum well.) We first discuss the leftmost curve, for $V_{FG} = 0$. As V_{BG} is made increasingly negative, the QPCs form when the electrons beneath them are depleted at a few tenths of a volt. Both QPC channels will then be occupied until V_{BG} becomes sufficiently negative that the bottom QPC, which is closest to the backgates, is pinched off. This occurs at $V_{BG} = -2.43$ V, and results in a slight but definite change in the slope of the conductance curve. As V_{BG} is made further negative, the conductance continues to decrease as the top QPC now decreases in width, until finally the top QPC is also pinched off at $V_{BG} = -3.16$ V and the conductance goes to zero. The additional G versus V_{BG} curves shift towards less negative V_{BG} values as V_{FG} is made more negative. In addition, the change in slope moves to lower values of G, indicating that the bottom QPC is pinching off at less negative V_{BG} values. These two effects clearly demonstrate that the top and bottom QPC openings can be independently controlled via the two pairs of QPC gates.

As is obvious from figure 3, conductance steps, quantized in units of $2e^2/h$, do not appear when both QPCs are open, but rather only when the bottom QPC is pinched off. (Another set of conductance curves, this time as a function of V_{FG} for several V_{BG} values (not shown), exhibited similar behaviour, with conductance steps occurring only when the top QPC was open, as indicated by a change in overall slope.) The occurrence of well-defined conductance steps when only the top QPC is open can be attributed to the top QW having a much higher mobility, as is known to occur commonly in MBE-grown AlGaAs/GaAs DQW structures [11]. Because the two QWs are very closely coupled, when both QPCs are occupied, the wavefunction is delocalized across both QWs into symmetric and antisymmetric subbands. This causes the electron scattering rate to increase as a result of the scattering centres in the low-mobility bottom QW and hence the conductance steps to be obscured. (This picture is further supported by the fact that no conductance steps were observed when only the bottom QPC was open; not shown here.) The relative weakness of the conductance steps when only the top QPC is occupied is attributed to the fairly wide QPC opening and to the extremely sharp gate tips in this device, which decrease the adiabatic transport of electrons through the QPC.

Our dual-side electron beam patterning technique should enable several novel quantum transport experiments to be better performed, such as electron–hole Coulomb drag and 1D–1D tunnelling [12]. For the double QPC device, an area for future exploration is the effect of individually biasing all four of the QPC gates in such a manner as to shift the QPC channels laterally relative to one another, thus controlling the inter-QPC coupling *independently of channel occupation.*

4. Summary

We have described the first demonstration of dual-side electron beam lithography in achieving independent submicron gating in DQW devices. The technique utilizes the EBASE process to remove the substrate material allowing the backside gates to be placed in close proximity (less than 1 μm) to the frontside gates. The use of electron beam lithography allows both the definition of submicron features and the precise alignment of the front and back features to each other. The key to this alignment is the use of the same set of metallic alignment marks for both the frontside and the backside electron beam patterning. Double QPCs fabricated with this technique on coupled DQWs exhibited quantized conductance steps in the top QW which shift as a function of backgate bias, confirming the independent control of each channel. These results show the potential of this new technique in advancing the study and fabrication of mesoscopic devices.

Acknowledgment

Sandia is a multiprogramme laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under Contract DE-AC04-94AL85000.

References

[1] Boebinger G S, Jiang H W, Pfeiffer L N and West K W 1990 *Phys. Rev. Lett.* **64** 1793

[2] Palevski A, Beltram F, Capasso F, Pfeiffer L N and West K W 1990 *Phys. Rev. Lett.* **65** 1929

[3] Eisenstein J P, Pfeiffer L N and West K W 1991 *Appl. Phys. Lett.* **58** 1497

[4] Gramila T J, Eisenstein J P, MacDonald A H, Pfeiffer L N and West K W 1991 *Phys. Rev. Lett.* **66** 1216

[5] Simmons J A, Blount M A, Moon J S, Baca W E, Reno J L and Hafich M J 1997 *Tech. Dig. IEDM* 755

[6] Eisenstein J P, Pfeiffer L N and West K W 1990 *Appl. Phys. Lett.* **57** 2324

[7] Lyo S K 1996 *J. Phys.: Condens. Matter* **8** L703

[8] Weckwerth M V, Simmons J A, Harff N E, Sherwin M E, Blount M A, Baca W E and Chui H C 1996 *Superlattices Microstruct.* **20** 561

[9] Boebinger G S, Passner A, Pfeiffer L N and West K W 1991 *Phys. Rev.* B **43** 12 673

[10] Simmons J A, Lyo S K, Harff N E and Klem J F 1994 *Phys. Rev. Lett.* **73** 2256

[11] Simmons J A, Lyo S K, Klem J F, Sherwin M E and Wendt J R 1993 *Phys. Rev.* B **47** 15 471

[12] Eugster C C and Del Alamo J A 1991 *Phys. Rev. Lett.* **67** 3586

Semicond. Sci. Technol. 13 (1998) A90–A92. Printed in the UK PII: S0268-1242(98)91887-9

Stability and band offsets of AlN/GaN heterostructures: impact on device performance

J A Majewski, G Zandler and P Vogl

Walter Schottky Institute and Physics Department, TU Munich, 85748 Garching, Germany

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We have performed first-principles local density calculations of several interfaces between GaN and AlN and studied their formation enthalpy and band offsets and the charge pileup at polar interfaces. Monte Carlo studies of the electrical characteristics of submicron HFET structures reveal that the pyro- and piezoelectric moments of the nitrides are a key property for designing and optimizing future devices.

1. Introduction

III–V nitrides possess unique material properties. They have large band gaps in the near-ultraviolet regime (3–6 eV), large heterostructure band offsets (1–2 eV), small effective masses (typically 0.2 for electrons and 0.5 for holes), energetically high-lying satellite valleys in the conduction bands (offsets of 1.5–3 eV) and high optical phonon energies (600–900 cm^{-1}). All of these properties make them extremely attractive for optoelectronic and high-temperature, high-power microelectronic applications [1, 2].

Even though epitaxial growth is still hampered by a high density of threading and misfit dislocations, stacking faults and other extended native defects, these materials offer novel and unique possibilities for the engineering of their properties [3]. They possess, in their wurtzite phase, a spontaneous electric polarization and have the largest piezoelectric constants known among tetrahedrally bonded materials [4].

In this paper, we show that these large intrinsic electric moments of the nitrides are key properties for devices. We have performed systematic first-principles pseudopotential local density functional calculations of a variety of interfaces between GaN and AlN and studied their formation enthalpy, their band offsets and the charge accumulation at some of these interfaces. To assess the potential of these effects quantitatively, we have also performed self-consistent Monte Carlo studies of the electrical characteristics of submicron HFET structures.

2. Polarity of interfaces

The space group $P6_3mc$ of the wurtzite structure is compatible with a spontaneous polarization per volume along the hexagonal c-axis. In the case of GaN and AlN, the spontaneous polarization at zero macroscopic electric field has been predicted by first-principles calculations to be -0.029 and -0.081 C m^{-2}, respectively [4]. In perfect crystals with a surface or interface, these moments give rise to huge longitudinal electric fields $E = -4\pi P/\epsilon_s$ of the order of 1–10 MV cm^{-1}. Here, $\epsilon_s \approx 10$ is the static dielectric constant. In addition, GaN and AlN are highly piezoelectric, in both their wurtzite and their zincblende phase. The zincblende piezoelectric constants are given by $e_{14} = 0.73$ and 1.55 C m^{-2} in GaN and AlN, respectively, which is to be compared with 0.1 C m^{-2} in GaAs [4]. Since the epitaxially grown layers are usually strained, the piezoelectric moments can also produce very large electric fields of the order of a few MV cm^{-1} [5].

Using first-principles pseudopotential local density functional calculations [6], we have systematically studied several interfaces. In our calculations, the effects of lattice relaxation and electric polarization have been fully taken into account. Let us consider a single pseudomorphic interface A–B between nitride compounds A and B. This interface is called polar if the total charge density, integrated across the interface from minus to plus infinity, is nonzero. In the GaN/AlN system, we may introduce the following classification of interfaces.

- *Nonpolar interfaces between unpolarized phases*. Unstrained or biaxially [001] strained zincblende phase GaN and AlN possess neither a pyroelectric nor a piezoelectric polarization. Therefore, there is no charge pileup at the interface between these materials. All bonds at the interface are saturated and there is only a weak dipole moment at the interface that is reduced by the atomic relaxation near the interface. An example of such an interface is lattice-matched, strained GaN grown on an AlN substrate along the [001] direction. For the sake of brevity, we denote this interface by zb AlN–s-GaN [001].

0268-1242/98/SA0090+03$19.50

- *Nonpolar interfaces between polarized phases.* Currently available samples contain a large number of internal interfaces that deviate from the hexagonal [0001] or cubic [111] growth direction [3]. In cases such as zb AlN–s-GaN [110], wz AlN–s-GaN [$0\bar{1}10$] or wz AlN–s-GaN [$\bar{1}2\bar{1}0$], the bulk materials on both sides of the interface possess a spontaneous and/or strain-induced piezoelectric polarization. However, this polarization lies parallel to the interface and therefore does not give rise to a charge accumulation at the interface.
- *Polar interfaces between polarized phases.* In the case of lattice-matched wz AlN–s-GaN [0001], both semibulk materials possess a spontaneous polarization. Because of the strain in the GaN film, there is an additional piezoelectric contribution to the total polarization. Since the polarizations in both materials are different and lie perpendicular to the interface, one obtains an interface charge. Our first-principles calculations reveal that lattice relaxation near the interface reduces the monopole contribution of the interface charge approximately by a factor of 2 by redistributing the charge among a few atomic layers. There are situations where the nonpolar interfaces become polar solely because of the strain-induced piezoelectric moment in one of the lattice-matched materials. An example is zb AlN–s-GaN [111]. Here, the strain-induced piezoelectric moment in GaN yields a polarization P perpendicular to the interface and therefore a nonzero interface charge density. A stacking fault is another example that possesses a charged interface. Such a kind of defect is nothing but a homo-type interface between wurtzite and zincblende GaN, such as wz GaN [0001]–zb s-GaN [111]. In this case, both materials have polarizations perpendicular to the interface: one phase is spontaneously polarized and the other one has a strain-induced electric moment.

In figure 1, we show a schematic view of the polarizations across several types of interfaces. Horizontal (vertical) arrows indicate a polarization perpendicular (parallel) to the interface; the absence of arrows implies the absence of bulk polarization. The conduction and valence band bending induced by the electric field due to the interface charges is also shown schematically in figure 1(*a*). The quantitative value of the electric field can, in principle, be calculated from the surface charge density given below but will depend strongly on the detailed interface sequence in a sample.

3. Formation enthalpies and band offsets: predictions

By minimizing total energies of sufficiently long supercells that contain two interfaces, our first-principles calculations show that all nonpolar GaN–AlN interfaces are stable and have an exothermic formation enthalpy of typically $\delta H = -15$ meV/atom. Importantly, we find all lattice-relaxed polar interfaces shown in figure 1 to be only slightly less favourable with formation energies around $\delta H = 0$ within computational accuracy. Thus, these interfaces are likely to form in real materials, depending on the growth kinetics.

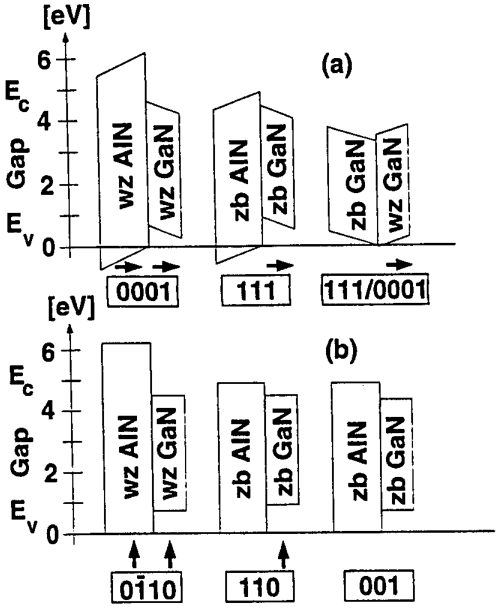


Figure 1. Predicted valence and conduction band offsets (in eV) for several AlN–GaN interfaces. In each case, the left-hand material is considered the unstrained substrate, whereas the right-hand material is lattice matched and strained. The growth direction is given in boxes beneath each interface. The arrows indicate the direction of the spontaneous or strain-induced piezoelectric polarization. (*a*) Polar interfaces that possess an interface charge. The corresponding electric field induces a band bending that is indicated schematically. The rightmost interface is a stacking fault and is also a type I heterostructure. (*b*) Nonpolar interfaces with no interface charge.

The calculated surface charge densities amount to -0.011 C m^{-2} for the wz AlN–s-GaN [0001] as well as for the zb AlN–s-GaN interface. In the latter case, this large value is somewhat surprising because the polarization is nonzero only on one side of the interface. The stacking fault wz GaN [0001]–zb s-GaN [111] has a slightly smaller but still appreciable interface charge density of 0.003 C m^{-2}.

In figure 1, we also show the calculated intrinsic valence and conduction band offsets between unstrained AlN and strained GaN. Only the leftmost value [7] in figrue 1(*a*) and the two last values [8, 9] in figure 1(*b*) have been determined previously and agree well with the published values. We have calculated these offsets from the asymptotic difference between the laterally averaged electrostatic potentials to the right and to the left of the interface after eliminating the monopole contribution to the potential [7]. It is interesting to note that all valence band offsets are of the order of 0.7 eV and are amazingly insensitive to the polar–nonpolar character of the interface.

All of the interfaces in figure 1 are of type I. This is particularly interesting for the stacking fault wz GaN [0001]–zb s-GaN [111] that has a valence band offset of -40 meV and a conduction band offset of -120 meV. Thus, confined regions of cubic GaN that are embedded within wurtzite material may act as excitonic traps and luminescence centres, depending on the electric fields present.

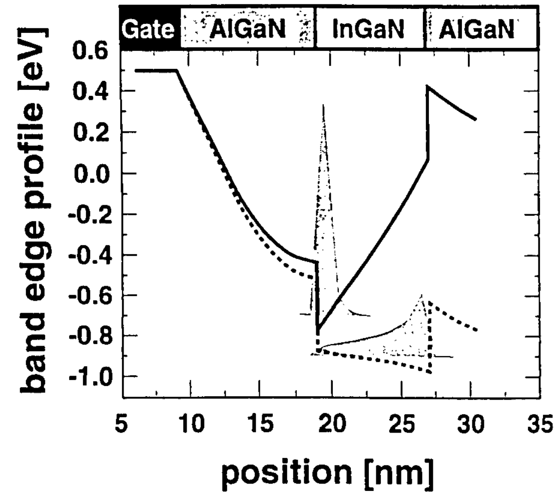


Figure 2. Self-consistent band edge profile and electron density for a 300 nm gate length $In_{0.05}Ga_{0.95}N/Al_{0.16}Ga_{0.84}N$ HFET with (——) and without (- - - -) strain-induced piezoelectric fields. The profile is taken at midgate position for a drain voltage of 20 V and gate voltage of +0.5 V.

4. Impact of polarization on HFET devices

We have carried out self-consistent Monte Carlo simulations for a planar short channel HFET with the following layer sequence [10]. Immediately beneath the contacts, we take a highly doped $Al_{0.16}Ga_{0.84}N$ supply layer of 10 nm. The next layer consists of an 8 nm $In_{0.05}Ga_{0.95}N$ channel with $n = 3 \times 10^{18}$ cm^{-3} and is followed by a 400 nm $Al_{0.16}Ga_{0.84}N$ buffer layer. In order to achieve acceptable off characteristics, we assume this buffer layer to be fully compensated or nearly intrinsic. For the gate contact, we assume a Schottky barrier height of 1 eV [11].

The lattice mismatch between the channel and the barrier material amounts to approximately 1%. This strain causes a macroscopic polarization of the channel in the growth direction. The resulting electric field can be calculated from the piezoelectric tensor and the elastic constants and amounts to 1.7×10^8 V m^{-1} for the HFET layer sequence given above. Experiments indicate that the polarization points towards the substrate [5, 12].

The effect of this huge polarization field on the self-consistent conduction band edge profile and the charge carrier density at midgate position is shown in figure 2 for gate and drain voltages of 0.5 V and 20 V, respectively. The gate length is 300 nm. One can see that the piezoelectric field pushes the electrons closer to the gate contact. Accordingly, the channel carrier density increases by about 20% for the open channel conditions given above. The high potential barrier that is formed by the piezoelectric field on the substrate side of the HFET (see figure 2) improves the response of the channel carrier density to a change in the gate voltage. For the device characteristics of a 300 nm gate HFET, this leads to an increase of the transconductance from 650 S m^{-1} to 980 S m^{-1}. Furthermore, this confining effect improves the turn-off behaviour of the device considerably, which becomes even more important at elevated temperatures. Our calculations predict average channel velocities exceeding 2×10^5 m s^{-1} for gate lengths below 200 nm [10].

References

[1] Morkoç H, Strite S, Gao G B, Lin M E, Sverdlov B and Burns M 1994 *J. Appl. Phys.* **76** 1363 and references therein
[2] Nakamura S and Fasol G 1997 *The Blue Laser Diode* (Berlin: Springer)
[3] Angerer H, Ambacher O, Dimitrov R, Metzger Th, Rieger W and Stutzmann M 1997 *MRS Internet J. Nitride Semicond. Res.* **1** 15
[4] Bernardini F, Fiorentini V and Vanderbilt D 1994 *Phys. Rev.* B **56** R10 024
[5] Takeuchi T, Sota S, Sakai H, Amano H, Akasaki I, Kaneko Y, Nakagawa S, Yamaoka Y and Yamada N 1997 *Proc. 2nd Int. Conf. on Nitride Semiconductors (ICNS 97) (Tokushima, 27–31 October 1997)* p 418
Yu E T, Sullivan G J, Asbeck P M, Wang C D, Qiao D and Lau S S 1997 *Appl. Phys. Lett.* **71** 2794
[6] Städele M, Majewski J A and Vogl P 1997 *Phys. Rev.* B **56** 6911
[7] Bernardini F, Fiorentini V and Vanderbilt M 1997 *Mater. Res. Soc. Symp. Proc.* **449** 923
[8] Albanesi E A, Lambrecht W R L and Segall B 1994 *J. Vac. Sci. Technol.* B **12** 2470
[9] Nardelli M B, Rapcewicz K and Bernholc J 1997 *Phys. Rev.* B **55** R7323
[10] Zandler G, Majewski J A, Städele M, Vogl P and Compagnone F 1997 *Phys. Status Solidi* b **204** 133
[11] Suzue K, Mohammad S N, Fan Z F, Kim W, Aktas O, Botchkarev A E and Morkoç H 1996 *J. Appl. Phys.* **80** 4467
[12] Martin G, Botchkarev A, Rockett A and Morkoç H 1996 *Appl. Phys. Lett.* **68** 2541

Semicond. Sci. Technol. **13** (1998) A93–A96. Printed in the UK PII: S0268-1242(98)91886-7

Electron logic devices and nanoinstrumentation based on laterally patterned interaction-free quantum measurement structures

J R Barker

Department of Electronics and Electrical Engineering, University of Glasgow,
Glasgow G12 8LT, UK

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. Recent photonic experiments on non-invasive or interaction-free measurement are discussed within the interpretative deterministic framework put forward by Bohm. It is proposed that analogous mesoscopic electronic structures utilizing spin polarization could support the non-invasive effect. The incorporation of a full Zeno effect leads to the possibility of building deterministic quantum switches.

1. Introduction

Many states in heterostructure systems have a strong influence on device characteristics but are notoriously difficult to measure on a nanometre scale. These include individual traps, surface and interface states, local soliton states in single electronic devices and local effects of etch damage. There is a clear requirement for instrumentation which can 'non-invasively' probe the surface and underlying structure of mesoscopic systems at high resolution without disturbing the probed structure. It is often assumed that measurements on a system cannot be made without disturbing it unless the system is already in a joint eigenstate with the measurement apparatus. However, Elitzur and Vaidman [1] were the first to raise serious prospects of interaction-free measurement in a photonics experiment involving a variant of a Mach–Zehnder interferometer. Their partial result was significantly improved on by Kwiat *et al* [2] in a demonstration of single-photon, interaction-free measurement by incorporating the quantum Zeno effect [3,4]. Kwiat *et al* [2] explain their results in terms of photons moving on paths through the system. This of course is a contradiction in the Copenhagen scheme but is easily explained if we use any of the different ontological interpretations of quantum mechanics put forward by Bohm and others [5–9], where the concept of trajectories exists for non-zero rest mass particles, or for the simplistic photon model.

2. Non-invasive measurement and its interpretation

The original idea behind interaction-free measurement [1] is sketched in figure 1(*a*), which is a schematic of a generic Mach–Zehnder interferometer comprising two arms A–M_1–BS_2 and A–M_2–BS_2. Quantons may be injected at A (photons if the arms are optical paths, electrons if we use quantum waveguides [10, 11]) and travel via perfect mirrors (photons) to a symmetric dual port output where they can be detected at D_1 or D_2. BS_1 and BS_2 are perfect 50–50 beam splitters (or junctions with appropriate S-matrices for electrons). The system has equal arm lengths so that there is total destructive interference along output BS_2–D_1. An incident quanton therefore reaches D_2 with probability 1 and reaches D_1 with probability 0. This wave picture (on the traditional view) vanishes if an absorbing object is placed in the upper path: figure 1(*b*). There is no interference at the splitter BS_2 (only one path). An incident quanton treated as a particle will enter the upper or lower branch with probability $P = 0.5$. If the lower path is taken there is a probability $P(D_1) = 0.25$ of exiting to detector D_1 and $P(D_2) = 0.25$ for exiting to D_2. If detection occurs at D_2 the event gives no information on the possibility of an object in the upper path. However, if the quanton is detected at D_1 there must be an object G in the upper path. Since only one quanton is injected and only one is detected at D_1 it is argued that the object is detected without interaction with the quanton. Evidently the probability for the quanton being absorbed by the object is $P(G) = 0.5$.

If we choose the simplest ontological interpretation, namely the deterministic Bohm picture [6], we consider the wavefunction field for the quanton penetrating both arms of the interferometer. The Bohm picture assumes that the quanton is a real, localizable particle which is sited at location r with probability density $\rho(r, t) = |\psi|^2 = R^2$, and has a local momentum $p(r) = \nabla S(r, t)$ where S is related to the phase of the wavefunction: $\psi = R\exp(iS/h)$. Substituting this form into the Schrödinger

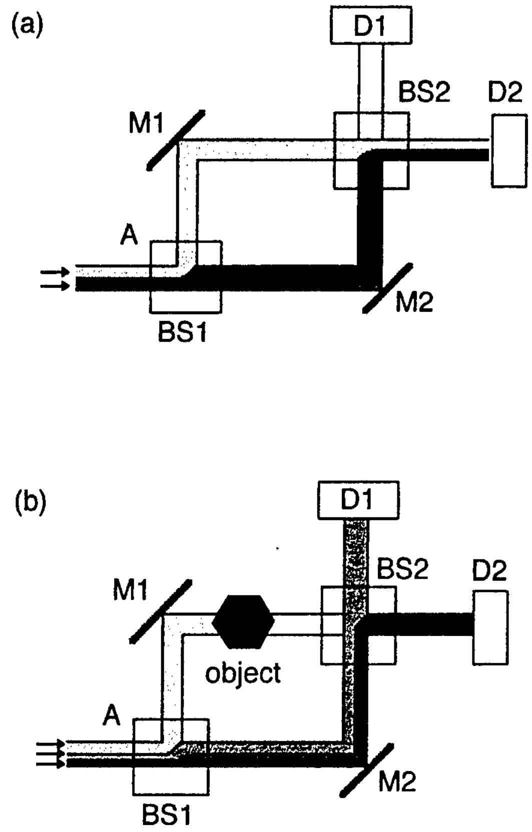


Figure 1. The Elitzur and Vaidman scheme: (*a*) free dual paths; (*b*) object in one path.

equation yields the continuity equation and a classical-like Hamilton–Jacobi equation for the action S:

$$\partial\rho/\partial t + \nabla\cdot\{\rho p/m\} = 0 \qquad (1)$$

$$-\partial S/\partial t = (\nabla S)^2/2m + \phi + \phi_Q$$
$$\phi_Q = -(\hbar^2/2mR)\,\nabla^2 R \qquad (2)$$

where the classical Hamiltonian is $p^2/2m + \phi$ and ϕ_Q is the Bohm quantum potential.

The possible family of deterministic trajectories $\boldsymbol{r}(t)$, $\boldsymbol{p}(t)$ may be found from equations (2) provided that we know R and S from the solution to Schrödinger's equation. The Bohm trajectories differ from classical trajectories in that they cannot cross. It is then very easy to see the topology of the quanton flow. In the case of figure 1(*a*), the quantons on the upper half of the incident path follow the upper arm, and the lower group follow the lower arm. At splitter BS_2 the two groups do not cross but run side by side out to D_2. In the case of figure 1(*b*), 50% of quantons (those on the upper half incident path) penetrate to the object where they are absorbed. The remaining 50% travel to BS_2 where half of them (those to the left) exit at D_1; the others exit at D_2, again with no crossing. There is no mystery in this picture: when a quanton is detected at D_1 in the case of figure 1(*b*), it must have avoided (i.e. not interacted with) the object. It carries information on the existence of the object because the full wave field reacts to the presence of the object and subsequently communicates with the quanton trajectories via the quantum potential. A good figure of merit for the interaction-free measurement is

$$M = P(D_1)/[P(D_1) + P(G)] \qquad (3)$$

where $P(D_1)$ is the probability for interaction-free measurement and $P(G)$ is the probability for strong interaction with the object G. Evidently, $M = 1/3$ for the case of figure 1. Kwiat *et al* [2] showed that by designing the beam splitter reflectivities efficiently (or in our case by optimizing the junction S-matrices) it is possible to achieve $P(D_1) = P(G)$ giving a maximum figure of merit of $M = 1/2$. The key result of Kwiat *et al* [2] was to show that by further exploiting the quantum Zeno effect [3, 4] the probability of interaction $P(G)$ can be reduced to zero and a maximum efficiency of $M = 1$ becomes possible.

3. Quantum Zeno effect

The quantum Zeno effect [3] uses repeated quantum measurements to inhibit the evolution of a quantum system. The original explanation relied heavily on the von Neumann projection postulate that the outcome of a measurement on a general quantum state results in the system evolving instantly to an eigenstate of the system, whereafter repeated measurements will produce the same result as the system remains in the originally (randomly) selected eigenstate. The effect is named after the classical Zeno 'paradox' and is conventionally interpreted to mean 'repeated observation stops the "motion"'. Kwiat *et al* [2] envisaged a sequence of N polarizing rotators, each rotating the photon polarization by $\pi/2N$. An incident horizontally polarized photon will thus emerge vertically polarized and therefore will be unable to be detected on the other side of an output horizontal polarizer H: $P(H) = 0$. However, if a horizontal polarizer is inserted into each stage the repeated measurement of the horizontal component leads to a strong probability of detecting the photon after the output horizontal detector H: $P(H) > 0$. A simple calculation shows that $P(H) = [\cos^2(\pi/2N)]^N$. Thus $P(H) \to 1$ as N becomes large ($P(H) = 0.67$ for $N = 6$).

In the simplest ontological picture [12] we have to consider the photon trajectory in the full configuration space which now includes polarization. Suppose the incident photon is horizontally polarized; within a short time t of the interaction with the first rotator the perturbed wavefunction will be proportional to t and is thus small. In order for a transition to take place the perturbed wavefunction must become sufficiently large to make a big contribution to the quantum potential which then moves the trajectories in configuration space. At short times, i.e. for rapidly repeated measurements, this cannot occur so no transition takes place.

A long sequence of N rotators and polarizing beam splitters is not actually necessary to obtain the Zeno effect: it is simpler to use a single rotator and a single horizontal polarizing beam splitter but to arrange for the photons to loop through the pair N times. By combining the Zeno effect arrangement with the optimized Mach–Zehnder interferometer, Kwiat *et al* [2] demonstrated an efficiency

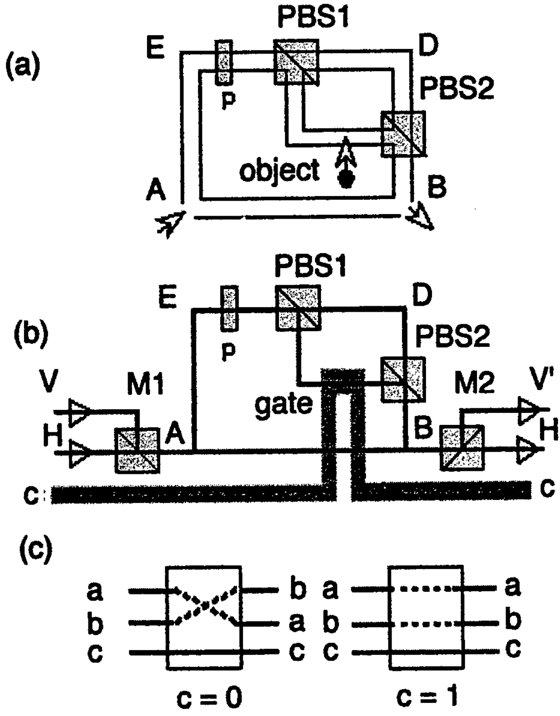


Figure 2. From non-invasive measurement to conservative logic (PBs, polarizing beam splitter): (*a*) the Kwiat scheme; (*b*) Fredkin gate version; (*c*) Fredkin gate modes.

for interaction-free measurement arbitrarily close to unity. Figure 2(*a*) illustrates the scheme: a horizontally polarized (H) photon is injected into the upper arm through a mirror at A. A polarizer P rotates the plane of polarisation by $\pi/2N$. The PBSs transmit H photons and reflect V (vertically polarized) photons. After N cycles through A–D–B–A, the photon is analysed at output B. With no object on the interferometer arm, and in the limit of large N, the output photon is vertically polarized (V). If an absorbing object is present the output photon is horizontally polarized (H) with probability → 1.0 as N is increased.

4. Electronic analogues?

Laterally patterned quantum waveguides [10] induced in two-dimensional electron gases provide a direct analogue with the photonic systems (figure 2). Mirrors may be replaced by graded bends in the wave guide; graded Y-junctions correspond to beam splitters. A sequence of Mach–Zehnder interferometers might be synthesized as a series of connected rings, or better as a multiport recursive re-entrant structure with a small lithographic footprint. To obtain the Zeno effect we need to introduce a further degree of freedom to provide a measurement in configuration space. One possibility is to inject spin-polarized electrons into a waveguide structure and to use a nanopatterned magnetic surface superlattice to act on the spins in the same fashion as the rotators and polarizers. This technology is just becoming available. The key ingredients required will be suitable injector and detector systems for single spin-polarized electrons. If such a structure were to be operated under Coulomb blockade conditions one might obtain well-correlated carrier flows [13, 14] which would permit the timing and synchronicity of the optical counterpart experiments. Evidently the net path length must be less than the electron coherence length which effectively limits the number of cycles N.

5. Deterministic quantum devices

Although very many quantum devices and indeed quantum computing systems have been studied [15], very little enthusiasm has remained for switches based on modulation of carrier interference. Apart from coherence difficulties there is one very serious obstacle to using interference devices switching small numbers of carriers. The channelling of carriers into the fringes of an interference pattern is typically a stochastic process: very many carriers must flow to build up the pattern. However, it follows from the Kwiat experiment that interference devices can be constructed with deterministic outcomes. Interaction-free measurement leads precisely to a probability of unity for one detection event and to zero for the complementary detection. There is still some stochasticity in the random time of arrival but the output is deterministic. This suggests that quantum interference devices could be re-examined within this new framework revealed by interaction-free measurement. The measured object is simply replaced by the action of a gate over the pathway. However, the output would be deterministic with a strong reduction in fluctuations, even for very small numbers of carriers. Figure 2(*b*) illustrates a simple modification to the Kwiat scheme in which H- and V-polarized single photons (or their electronic spin analogues) are injected in sequence into the structure and the 'object' is replaced by the action of a gate controlled by a signal c. With the gate off ($c = 0$) the input states are replaced by their complements. When the gate is 'on' ($c = 1$) the input and output states are the same. This is basically the action of a Fredkin gate in conservative logic [16]. Since the degree of determinism is controlled by the number of cycles N this scheme may provide a basis for the variable control of qbits in quantum computing.

References

[1] Elitzur A C and Vaidmam L 1993 *Found. Phys.* **23** 987
[2] Kwiat P, Weinfurter H and Zeilinger A 1995 *Phys. Rev. Lett.* **74** 4763–6
[3] Misra B and Sudarshan E C G 1977 *J. Math. Phys.* **18** 756–83
[4] Itano W M, Heinzen D J, Bollinger J J and Wineland D J 1990 *Phys. Rev.* A **41** 2295–300
[5] Bohm D and Hiley B J 1993 *The Undivided Universe* (London: Routledge)
[6] Holland P R 1993 *The Quantum Theory of Motion* (Cambridge: Cambridge University Press)
[7] Barker J R 1994 *Semicond. Sci. Tech.* **9** 911–17
[8] Mandel L 1966 *Phys. Rev.* **144** 1071–7
[9] Barker J R 1995 *Quantum Transport in Ultrasmall Devices* ed D K Ferry (New York: Plenum) pp 171–80
[10] Barker J R 1991 *Physics of Granular Electronic Systems* ed D K Ferry, J R Barker and C Jacoboni (New York: Plenum) pp 19–41

[11] Laughton M J, Barker J R, Nixon J A and Davies J H 1991 *Phys. Rev.* B **44** 1150–3

[12] Bohm D and Hiley B J 1987 *Phys. Rep.* **144** 323–48

[13] Barker J R and Babiker S 1995 *Quantum Transport in Ultrasmall Devices* ed D K Ferry (New York: Plenum) pp 217–26

[14] Barker J R, Babiker S and Roy S 1996 *Physica* B **227** 87–91

[15] Bandyopadhyay S and Roychowdhury V 1996 *Japan. J. Appl. Phys.* **35** 3350–62

[16] Barker J R 1995 *Introduction to Molecular Electronics* ed M Petty, D Bloor and M Bryce (London: Edward Arnold) ch 16, pp 345–76

Semicond. Sci. Technol. **13** (1998) A97–A99. Printed in the UK

PII: S0268-1242(98)91856-9

Relaxation rates of electrons in a quantum well embedded in a finite-size semiconductor slab

V I Pipa†‡, B A Glavin†‡, V V Mitin‡ and M Stroscio§

† Institute of Semiconductor Physics, Pr. Nauki 45, Kiev 252028, Ukraine
‡ Department of ECE, Wayne State University, Detroit, MI 48202, USA
§ US Army Research Office, PO Box 12211, Research Triangle Park, NJ 27709, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We calculated the momentum and energy relaxation rates of 2D electrons interacting with acoustic phonons in a quantum well situated close to the surface of a semiconductor slab in which the quantum well is embedded. These rates depend on the mechanical conditions at the surface of the slab and deviate substantially from the values corresponding to a quantum well situated in the bulk of an infinite crystal. At low enough temperatures the quantization of the acoustic modes becomes important and has an effect on the relaxation rates.

Recently, it has been established that if a quantum well (QW) is situated close to the surface of a slab in which the QW is embedded, the interaction between 2D electrons in the QW and acoustic phonons takes on new features [1,2]. These peculiarities arise as a result of the modification of the structure of the acoustic modes caused by the surface and are as follows: (i) interference between the incident and reflected acoustic waves, which gives rise to the formation of the nodes and antinodes of the potential of the electron–phonon interaction; (ii) appearance of a Rayleigh wave; (iii) reflection-induced mutual conversion of LA and TA phonons.

In [2] it has been shown that the effect of the proximity of the surface to the QW depends substantially on the mechanical conditions at free and rigid surfaces, the temperature, the 2D electron concentration and the distance between the QW and the surface. The influence of the vicinity of the surface on the magnitude of the electron relaxation rates is the strongest at low temperatures. At low enough temperature, the finite size of the slab in which the QW is embedded leads to new peculiarities in the properties of the phonons interacting with electrons. In fact, these peculiarities are due to the quantization of the acoustic phonons in the slab. As a result, at low temperatures the transport characteristics of the electrons for the cases of a slab of finite width and a semi-infinite sample are different.

To account for the finite size of the slab, we have calculated the acoustic modes confined in the slab for the cases of free and rigid boundaries. The elastic-continuum approach has been used. We have obtained the momentum, ν_p, and energy, ν_e, relaxation rates for electrons interacting with confined acoustic phonons via the deformation potential mechanism. The electron distribution function has been taken to be a displaced Fermi distribution and only the lowest QW level has been assumed to be populated. The former assumption is valid under conditions of strong electron–electron scattering, which is realized for the high electron concentrations.

The most important peculiarities of the phonon spectrum are the following. In the case of a slab with a free surface, the lowest branch of the phonon spectrum starts at the point $\omega = 0$, $q = 0$, where ω and q are the phonon circular frequency and wavevector. At high q, this branch corresponds to a Rayleigh wave, whose amplitude decays away from the surface. In contrast, for the case of a slab with a rigid surface, the lowest branch starts at $\omega = \omega_m$, $q = 0$. This corresponds to the absence of a Rayleigh wave in a slab with a rigid surface. As an order of magnitude estimate, $\omega_m \approx \pi s_t/b$, where s_t is the transverse sound velocity and b is the thickness of the slab. Another important peculiarity of the phonon mode structure is the surface-induced appearance of the nodes and antinodes of the lattice displacements. This feature arises because of the interference between the incident and reflected waves and leads to the formation of nodes and antinodes in the potential of the electron–phonon interaction. Since electrons usually interact with a great number of phonon modes, whose nodes and antinodes do not coincide, the resulting strength of the electron–phonon interaction does not depend on the QW position far from the surface. However, close to the surface *all* phonon modes behave similarly, because all modes obey the same boundary conditions at the surface. For the deformation potential mechanism of interaction one has nodes of the

0268-1242/98/SA0097+03$19.50

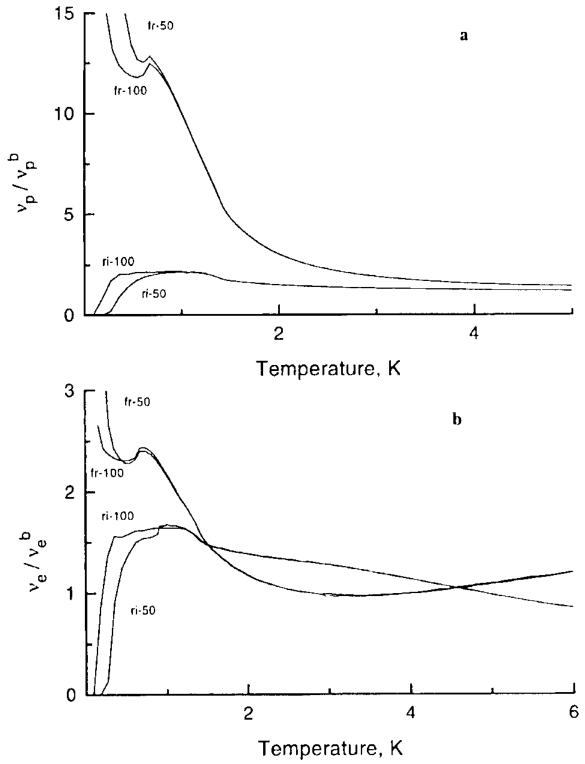


Figure 1. Temperature dependence of the relaxation rates of (*a*) the momentum and (*b*) the energy for a 3 nm QW situated at the surface of a slab. The curves 'fr-100' and 'fr-50' ('ri-100' and 'ri-50') are for the 100 nm and 50 nm slab with a free (rigid) surface respectively.

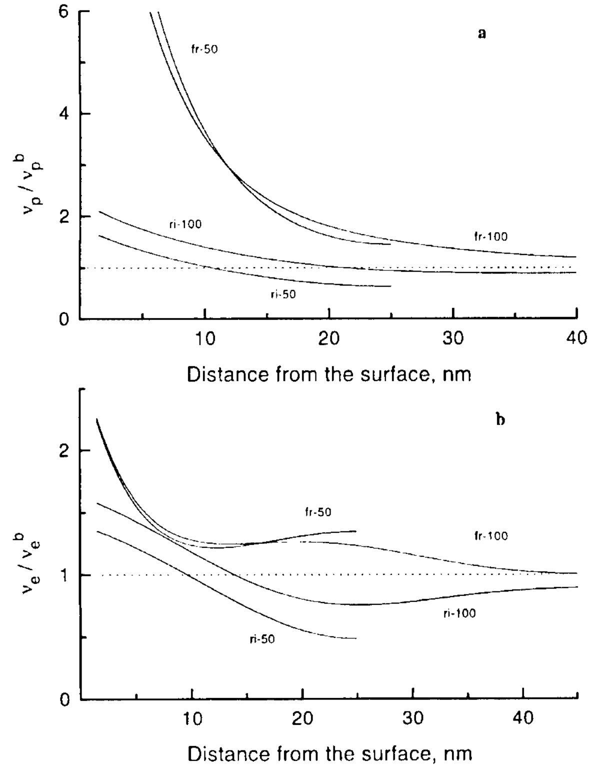


Figure 2. Dependence of the relaxation rates of (*a*) the momentum and (*b*) the energy on the distance between the QW and the surface for a temperature $T = 0.5$ K. The notations used are the same as in figure 1.

electron–phonon interaction potential at the surface for the case of a free surface and antinodes for the case of a rigid surface.

Clearly, many factors influence the electron–phonon interaction near the surface. Their contribution depends substantially on the temperature, which determines which phonons contribute most to the interaction. Two characteristic temperatures can be introduced [3]: $T_1 = 2\hbar s_t/a$ and $T_0 = 2p_F s_t$. Here a is the thickness of the QW and p_F is the Fermi momentum of the electrons. For $T \ll T_0$ the Bloch–Grüneisen regime of scattering occurs; this regime is characterized by the small-angle scattering. For $T \gg T_1$, electrons interact primarily with phonons propagating perpendicular to the QW and have a component of the wavevector perpendicular to the QW that scales in magnitude as $1/a$. Note that the intermediate range $T_0 \ll T \ll T_1$ can be realized for low electron concentrations and for narrow QWs.

In [2] the energy and momentum relaxation rates have been calculated and analysed in these temperature ranges for the case of a QW situated near the surface of a semi-infinite sample. Here, we present the results for the case of a QW embedded in a finite-size slab. In this case a new characteristic temperature should be introduced: $T_s = \pi\hbar s_t/b$. For $T \sim T_s$, the quantization of the phonons influences the values of the electron relaxation rates.

In figure 1 the dependences of the momentum and energy relaxation rates on temperature for a 3 nm QW are shown for 100 nm and 50 nm slabs. Relaxation rates are normalized to the corresponding values, ν_e^b, ν_p^b, for a QW in a bulk crystal . The electron concentration is 3×10^{15} m^{-2} and we use GaAs material parameters. Curves for the 100 nm and 50 nm slabs with free (rigid) surfaces are labelled as 'fr-100' and 'fr-50' ('ri-100' and 'ri-50') respectively. For such a system, $T_0 = 7$ K, $T_1 = 17$ K and $T_s = 1.2$ K for the 100 nm slab and 2.4 K for the 50 nm slab. Note that figure 1 represents results for the case $T < T_0$. As is illustrated, for $T > T_s$ the results do not depend on the slab width. In this case, the temperature dependences obtained are very similar to those obtained for the case of a QW near the surface of the semi-infinite sample [2]. Relaxation rates are enhanced for the slabs with both kinds of mechanical conditions at the surface. For the case of a rigid surface, this occurs as a result of the formation of antinodes in the electron–phonon interaction potential near the surface. For the case of a free surface, one has nodes of the interaction potential near the surface, but at $T < T_0$ scattering due to the Rayleigh wave becomes very important and the resulting relaxation rates are enhanced with respect to the rates for a QW in a bulk crystal.

For $T \sim T_s$, electrons are able to interact with only a few phonon modes. For $T \ll T_s$ and for a slab with a rigid surface, the energy corresponding to the minimum frequency of the acoustic modes, ω_m, is higher than the temperature and the relaxation rates decrease. In contrast, for the case of a semi-infinite sample with a rigid surface, one obtains an increase of the relaxation rate down to very low temperatures. For $T \ll T_s$ and a slab with a free surface, we obtain an increase in the relaxation rates at low temperatures since there is no minimum frequency for the phonon spectrum.

In figure 2 the dependences of the relaxation rates on the distance between the QW and the surface of the slab are shown for $T = 0.5$ K. For such a low temperature, the principal contribution to the scattering comes from a scattering with the lowest phonon mode and the distance dependence qualitatively traces the coordinate dependence of the potential of the electron–phonon interaction of these lowest modes. Note that for low enough temperatures the relaxation rates differ from the bulk values even if the QW is situated far from the surface. We see that, for $T = 0.5$ K and for $b = 50$ nm, the relaxation rates are different from the volume values even if the QW is situated in the centre of the slab. It is worth noting that, in the case of a slab with a free surface, the penetration length of the Rayleigh wave may be estimated by bT/T_s. This means that for low temperature the amplitude of the Rayleigh wave is high even near the centre of the slab.

In conclusion, we emphasize that the relaxation rates of the 2D electrons in a QW situated near the surface of a slab depend on the mechanical conditions at the surface of the slab. At low temperatures, where the effect is particularly strong, the finite width of the slab makes the phonon quantization important and gives rise to new peculiarities in the behaviour of the relaxation rates. Moreover, in this case the relaxation rates are different even if the QW is situated far from the slab surface.

Acknowledgment

This work was supported by US ARO.

References

[1] Badalyan S M and Levinson I B 1988 *Sov. Phys. Solid State* **30** 1592

[2] Pipa V I, Vasko F T and Mitin V V 1998 *J. Appl. Phys.* submitted

[3] Karpus V 1988 *Sov. Phys. Semicond.* **22** 268

Semicond. Sci. Technol. **13** (1998) A100–A103. Printed in the UK PII: S0268-1242(98)91857-0

Strong suppression of spin splitting in quantum wires due to correlations induced by edge state screening

O G Balev† and P Vasilopoulos‡

† Institute of Physics of Semiconductors, National Academy of Sciences,
45 Pr Nauky, Kiev 252650, Ukraine
‡ Concordia University, Department of Physics, 1455 de Maisonneuve Blvd O,
Montréal, Québec H3G 1M8, Canada

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. A modified local-density approximation is developed to treat self-consistently exchange and correlations in a quantum wire (QW) subjected to a strong perpendicular magnetic field. The influence of many-body interactions on the spin splitting between the two lowest Landau levels (LLs) is calculated within the screened Hartree–Fock approximation (SHFA), for filling factor $\nu = 1$, and the strong spatial dependence of the screening properties of electrons is taken into account. In comparison with the Hartree–Fock result, the spatial behaviour of the occupied LLs in a QW is strongly modified when correlations are included. Correlations caused by bulk screening are very weak whereas those caused by screening at the edges are very strong. The latter strongly suppress the exchange splitting and smoothen the energy dispersion at the edges. The theory is in reasonable agreement with the observed strong suppression of the spin splitting pertinent to the $\nu = 1$ quantum Hall effect (QHE) state and the destruction of this state in long QWs.

1. Introduction

One consensus of the theoretical work on the effects of electron–electron interactions on the edge state properties of a channel [1–4] and on the subband structure of quantum wires (QWs) [3, 5–7] is that this interaction must be treated self-consistently. Here we introduce a realistic model of a QW in a strong magnetic field B and self-consistently treat mainly the case when the lowest, spin-polarized, Landau level (LL) is occupied, i.e. when $\nu = 1$, in the interior part of a channel. Moreover, we consider submicron width channels with a rather steep confining potential that prevents the flattening [1] of edge states [2, 3, 6] in the vicinity of the Fermi level. To date we are aware of only the Hartree [3, 6] and Hartree–Fock [5] treatments of LLs in a QW, in a strong B field. We show that including correlations in the Coulomb interaction in a QW strongly modifies the spatial behaviour of the LLs and leads to the results announced in the abstract. The most important role played by correlations is related to screening by the edge states which in turn depends strongly on their (group) velocity v_g. The correlations can restore a smooth, on the scale of the magnetic length $l_0 = (\hbar/m^*\omega_c)^{1/2}$, dispersion of the single-particle energy as a function of the oscillator centre $y_0 \approx k_x l_0^2$ where ω_c is the cyclotron frequency. Because in typical experimental situations the *strong magnetic field limit* condition, $r_0 = e^2/\varepsilon l_0 \hbar\omega_c \ll 1$, is not satisfied, we propose a modified local-density approximation (MLDA) to treat self-consistently the effect of many-body interactions in a strong B, when $r_0 \lesssim 1$. For integer ν the self-consistent confining potential is nearly parabolic [6] . We consider a parabolic potential, of frequency Ω, and assume $\omega_c/\Omega \gg 1$. Our theory describes well the experimentally observed spin splitting in QWs [7].

In section 2 we present the formalism and in section 3 we show how strong correlations result from screening at the edges. In section 4 we propose an MLDA and compare our theory with the experiment [7]. We conclude in section 5.

2. Basic relations

We consider a zero-thickness two-dimensional electron gas (2DEG) confined in a narrow channel, in the (x, y) plane, of width W and of length $L_x = L$. In the absence of exchange and correlation effects we take the confining potential along y as parabolic: $V_y = m^*\Omega^2 y^2/2$, where m^* is the effective mass. We use the Landau gauge $\boldsymbol{A} = (-By, 0, 0)$ for a field B applied along the z-axis. V_y is valid especially for $W \lesssim 0.3$ μm; see [5] and [7]. However, most of our results hold for the more realistic one V_y', if $W \gtrsim 3$ μm, which is flat in the interior of the channel and parabolic at its edges.

0268-1242/98/SA0100+04$19.50

The one-electron eigenvalues are

$$\epsilon_\alpha \equiv \epsilon_{n.k_x.\sigma} = \hbar\tilde{\omega}(n+1/2) + \frac{\hbar^2 k_x^2}{2\tilde{m}} + \sigma g_0 \mu_B B/2 \qquad (1)$$

and the eigenfunctions $|\alpha\rangle \equiv |nk_x\rangle|\sigma\rangle = \mathrm{e}^{\mathrm{i}k_x x}\Psi_n[y - y_0(k_x)]|\sigma\rangle/L^{1/2}$. Here, g_0 is the bare Landé g factor, μ_B the Bohr magneton and $\sigma = \pm 1$ for spin $\uparrow$ $(+)$ or $\downarrow$ $(-)$. Further, $\tilde{\omega} = (\omega_c^2 + \Omega^2)^{1/2}$, $\omega_c = |e|B/m^* c$, $\tilde{m} = m^*\tilde{\omega}^2/\Omega^2$, $y_0(k_x) = \hbar\omega_c k_x/m^*\tilde{\omega}^2$, $\Psi_n(y)$ is a harmonic oscillator function, $|\sigma\rangle \doteq \psi_\sigma(\sigma_1) = \delta_{\sigma\sigma_1}$ is the spin wavefunction, $\sigma_1 = \pm 1$, and $\tilde{l} = (\hbar/m^*\tilde{\omega})^{1/2}$.

In the strong magnetic field limit [2,5], the condition $r_0 = e^2/\varepsilon l_0 \hbar\omega_c \ll 1$ should be satisfied [8], where ε is the background dielectric constant. However, in the most important experiments r_0 is of order unity [4–8]. Nevertheless, it is believed that calculations for $r_0 \ll 1$ provide a useful framework also for $r_0 \sim 1$ [2,5].

Considering exchange to first order in r_0, in the screened Hartree–Fock approximation (SHFA) the exchange and correlation contribution to the total single-particle energy $E_{0.k_x.1} = \epsilon_{0.k_x.1} + \epsilon^{ec}_{0.k_x.1}$ is

$$\epsilon^{ec}_{0.k_x.1} = -\frac{1}{8\pi^3}\int_{-k_F}^{k_F} \mathrm{d}k'_x \int_{-\infty}^{\infty} \mathrm{d}q_y \int_{-\infty}^{\infty} \mathrm{d}q'_y \, V^s(k_-, q_y; q'_y) \times (0k_x|\mathrm{e}^{\mathrm{i}q_y y}|0k'_x)(0k'_x|\mathrm{e}^{\mathrm{i}q'_y y}|0k_x) \qquad (2)$$

$k_F = (\tilde{\omega}/\hbar\Omega)(2m^*\,\Delta E_{F\uparrow})^{1/2}$ $(|k_x| \le k_F)$, $k_\pm = k_x \pm k'_x$ and $\Delta E_{F\uparrow} = E_F - \hbar\tilde{\omega}/2 - g_0\mu_B B/2$. $V^s(q_x, q_y; q'_y)$ is the Fourier transform of the screened Coulomb interaction.

If we neglect the screening of the interaction between two electrons, at (x, y) and (x', y'), by all other electrons in the QW, then $V^s(q_x, q_y; q'_y) \equiv V^s_0(q_x, q_y; q'_y) = (4\pi^2 e^2/\varepsilon q)\delta(q_y + q'_y)$, $q = (q_x^2 + q_y^2)^{1/2}$. For $\tilde{k}_F \gg 1$ we have

$$\epsilon^{ex}_{0.k_x.1} \approx \epsilon^{ex}_0 = -(\pi/2)^{1/2}(e^2/\varepsilon\tilde{l})R(\Omega/\omega_c) \qquad (3)$$

if $\tilde{k}_F - |\tilde{k}_x| \gg 1$ where $\tilde{k}_{x.F} = k_{x.F}\tilde{l}$; for $\tilde{k}_x = \pm\tilde{k}_F$, we have $\epsilon^{ex}_{0.\pm k_F.1} = \epsilon^{ex}_0/2$. Here $R(x) = (1 + x^2)^{1/2}F(1/2, 1/2, 1; -x^2)$ and $F(\ldots)$ is the hypergeometric function. Equation (3) differs from equation (11) of [5] by the factor $R(\Omega/\omega_c)$. Notice that, if only one LL is occupied, $\tilde{k}_F \gg 1$ implies $\omega_c/\Omega \gg 1$ and $R(\Omega/\omega_c) \approx (1 + \Omega^2/4\omega_c^2) \approx 1$. That is, we can approximate R by 1 only for $\omega_c/\Omega \gg 1$. For $\nu = 1$ the pure exchange or correlation contribution to $E_{0.k_x.-1}$ vanishes independently of the value of r_0; the exact result is $E_{0.k_x.-1} = \epsilon_{0.k_x.-1}$.

As is usual in the SHFA [9], we treat the screened Coulomb interaction V^s in the static limit. However, an essential difference is that we take into account the spatial inhomogeneity of the 2DEG along the y direction. We calculate $V^s(q_x, q_y; q'_y)$ within the random phase approximation (RPA). Consider the statically screened, by the 2DEG, potential $\varphi(x, y; x_0, y_0)$ of an electron charge at (x_0, y_0), $e\delta(\boldsymbol{r} - \boldsymbol{r}_0)$. Its Fourier transform $\varphi(q_x, q_y, q'_y)$ obeys the integral equation

$$\varphi(q_x, q_y; q'_y) = \frac{e}{\varepsilon q}\left[4\pi^2\delta(q_y + q'_y) + \frac{e}{L}\sum_{\alpha\beta}^{\infty} F_{\beta\alpha}\int_{-\infty}^{\infty} \mathrm{d}q_{y1}\,\varphi(q_x, q_{y1}; q'_y)MN\right] \qquad (4)$$

$M = \langle\alpha|\mathrm{e}^{\mathrm{i}q_{y1}y}|\beta\rangle$, $N = \langle\beta|\mathrm{e}^{-\mathrm{i}q_y y}|\alpha\rangle$, $F_{\beta\alpha} = (f_\beta - f_\alpha)/(\epsilon_\beta - \epsilon_\alpha + i\hbar/\tau)$ and f_α is the Fermi function. For flat LLs, i.e. for $\Omega \to 0$, fixed W and $\nu = 1$, equation (4) leads to the wide-channel RPA dielectric function [9] and shows that the screening is weak.

For a narrow channel we solve equation (4) by iteration. This results in a power series in r_0, $r_0 \ll 1$. Writing

$$V^s(q_x, q_y; q'_y) = \sum_{j=0}^{\infty} V^s_j(q_x, q_y; q'_y) \qquad (5)$$

and using equation (4) with $V^s(\ldots) = e\varphi(\ldots)$, we see that $V^s_0(q_x, q_y; q'_y)$ is the same as that given above. Substituting $V^s(q_x, q_y; q'_y)$ given by equation (5) in equation (2) we obtain

$$\epsilon^{ec}_{0.k_x.1} = \sum_{j=0}^{\infty} \epsilon^{(j)ec}_{0.k_x.1} \qquad (6)$$

where $\epsilon^{(0)ec}_{0.k_x.1} \equiv \epsilon^{ex}_{0.k_x.1}$. The result for $\epsilon^{(1)ec}_{0.k_x.1}$ can be written as $\epsilon^{(1)ec}_{0.k_x.1} = \epsilon^{(1)ec}_{I}(k_x) + \epsilon^{(1)ec}_{II}(k_x)$, with

$$\epsilon^{(1)ec}_{I}(k_x) = -(e^4/4\pi^2\varepsilon^2)\int_{-k_F}^{k_F} \mathrm{d}k'_x\, \mathrm{e}^{-a^2k_-^2\tilde{l}^2} \times \int_{-\infty}^{\infty} \mathrm{d}k_{x\alpha}\, R_I(k_{x\alpha}, k_-)\, M_I^2(K) \qquad (7)$$

$$\epsilon^{(1)ec}_{II}(k_x) = (e^4/4\pi^2\varepsilon^2\hbar\tilde{\omega})\int_{-k_F}^{k_F} \mathrm{d}k'_x\, \mathrm{e}^{-a^2k_-^2\tilde{l}^2}\int_{-\infty}^{\infty} \mathrm{d}k_{x\alpha} \times \sum_{n_\alpha=1}^{\infty} \frac{f_{0.k_{x\alpha}-k_-}}{n_\alpha n_\alpha!}[|M^+_{n_\alpha}(K)|^2 + |M^-_{n_\alpha}(K)|^2]. \qquad (8)$$

Here $a = \omega_c/\tau_0$, $K = \{k_{x\alpha}, k_x, k'_x\}$, $M_I(\ldots)$ and $M^\pm(\ldots)$ are integrals of M, N and $R_I(k_x, q_x) = (f_{0.k_x-q_x} - f_{0.k_x})/(\epsilon_{0.k_x-q_x} - \epsilon_{0.k_x})$. Since $\epsilon^{(1)ec}_{I.II}(k_x) \equiv \epsilon^{(1)ec}_{I.II}(-k_x)$, we consider only the right-hand half of the channel $k_x \ge 0$.

3. Inhomogeneous screening

In equation (7) we can approximate $R_I(k_{x\alpha}, k_-)$ by $(-2\tilde{m}/\hbar^2)\delta(k_{x\alpha}^2 - k_F^2)$. Then, for fixed k_F or W, $\epsilon^{(1)ec}_I(k_x)$ is determined only by screening at the edges of the channel. The latter depends practically only on the slope of the energy dispersion, i.e. on the group velocity v_g of the edge states [10], whose Hartree value in the strong B limit can be well approximated by $v_g^H = \hbar^{-1}(\partial\epsilon_{0.k_x.1}/\partial k_x)_{k_x=k_F}$. Finally, for fixed W further analytical and numerical calculations show that $\epsilon^{(1)ec}_I(k_x) \propto (1/v_g^H)$. This behaviour follows from the factor $R_I(k_{x\alpha}, k_-)$ in equation (7). Then, for fixed k_F (or W) and $r_0 \lesssim 1$, the self-consistent v_g can be significantly different from v_g^H. We obtain $\epsilon^{(1)ec}_I(k_x) \propto 1/v_g$ if the energy dispersion is smooth on the scale of $1/l_0$. The most important correlations are related to the strong screening by the edge states.

In the inner region of the QW we may assume $\tilde{k}_F - \tilde{k}_x \gg 1$ and obtain, with $\Delta\tilde{E}_{F\uparrow} = \Delta E_{F\uparrow}/\hbar\tilde{\omega}$,

$$\epsilon^{(1)ec}_I(k_x) = (\mathrm{Ry}^*/2\,\Delta\tilde{E}_{F\uparrow})(1 - k_x^2/k_F^2)^{-1} \qquad (9)$$

where $\mathrm{Ry}^* = e^4 m^*/\hbar^2\varepsilon^2$ is the effective rydberg.

In the middle of the channel we have $k_x^2/k_F^2 \ll 1$ and equation (9) gives $\epsilon^{(1)ec}_I \approx Ry^*/2\,\Delta\tilde{E}_{F\uparrow}$. We see that in the inner region of the QW the correlations induced by screening near the channel edges are strong and they become considerably stronger as the edges of the QW are approached. From equation (9) for fixed k_F (or W) and k_x but variable Ω, we have $v_g^H \propto \Omega^2$ and $\epsilon_I^{(1)ec}(k_x) \propto 1/v_g^H$.

At the right edge of the QW we have

$$\epsilon_I^{(1)ec}(k_F) \approx (\mathrm{Ry}^*/16\,\Delta\tilde{E}_{F\uparrow})\;[1 + 4(2/\pi)^{1/2}\tilde{k}_F]. \qquad (10)$$

It follows that $\epsilon_I^{(1)ec}(k_F)/\epsilon_I^{(1)ec}(0) \approx \tilde{k}_F/(2\pi)^{1/2}$. The first term in the brackets of equation (10) is related to screening at the left edge and is $4(2/\pi)^{1/2}\tilde{k}_F$ times weaker than the second term describing screening at the right edge.

The edge velocity, when only the exchange contribution $\epsilon^{ex}_{0,k_x,1}$ is taken into account and correlation contributions are neglected, is logarithmically divergent:

$$v_g^{ex}(k_x \approx k_F) \approx (e^2/2\pi\hbar\varepsilon)\ln[8/(\tilde{k}_0^2 + 4\pi^2 l_0^2/L^2)] \qquad (11)$$

here $\tilde{k}_0 = \tilde{k}_x - \tilde{k}_F$ and $\tilde{k}_0^2/4 \ll 1$. We introduced the very small value $4\pi^2 l_0^2/L^2$ to avoid the divergence for $|\tilde{k}_0| \to 0$. In this case a finite and relatively small v_g, for $\omega_c \gg \Omega$, is obtained without explicit appeal to many-body interactions, when $\Delta E_{F\uparrow}$ is independent of Ω, as $v_g = v_g^H(k_F) = \hbar k_F/\tilde{m} \propto \Omega/\tilde{\omega}$, which is negligibly small compared with v_g^{ex} given by equation (11). That is, any finite $v_g^H(k_F)$ will lead to a divergent $v_g^{ex}(k_F)$ when only exchange is taken into account. For $|\tilde{k}_0| \gg 1$ equation (8) gives

$$\epsilon_{II}^{(1)ec}(k_x) \approx \mathrm{Ry}^*[1 + \Phi(\tilde{k}_F - \tilde{k}_x)]/4 \qquad (12)$$

where $\Phi(x)$ is the probability integral. The correlations here are related to the screening within the channel only. Equation (12) leads to a negative contribution to the total v_g.

4. MLDA

4.1. Self-consistent treatment

For $2\,\Delta\tilde{E}_{F\uparrow} = 1$ and $\Omega/\omega_c \ll 1$, the ratio of $\epsilon_I^{(1)ec}(0)$, from equation (9), to $|\epsilon^{ex}_{0,0,1}|$, from equation (3), is $(2/\pi)^{1/2} r_0$. Further, the above treatment shows that correlations induced by the screening can be so strong near the edges that in the strong magnetic field limit we must have not only $r_0 = \mathrm{Ry}^*/(e^2/\varepsilon l_0) \ll 1$ but also that $10 r_0$ essentially less than 1. In this limit we obtain $\epsilon^{ec}_{0,k_x,1} \approx \epsilon^{ex}_{0,k_x,1} + \epsilon^{(1)ec}_{0,k_x,1} = \epsilon^{ex}_{0,k_x,1}(1 + \epsilon^{(1)ec}_{0,k_x,1}/\epsilon^{ex}_{0,k_x,1})$. This can be rewritten as

$$\epsilon^{ec}_{0,k_x,1} \approx \epsilon^{ex}_{0,k_x,1}/(1 - \epsilon^{(1)ec}_{0,k_x,1}/\epsilon^{ex}_{0,k_x,1}) \qquad (13)$$

since $\epsilon^{(1)ec}_{0,k_x,1}/|\epsilon^{ex}_{0,k_x,1}| \sim r_0 \ll 1$; the total energy is $E_{0,k_x,1} = \epsilon_{0,k_x,1} + \epsilon^{ec}_{0,k_x,1}$. Then the RHS of equation (13) well approximates the main contributions to $\epsilon^{ec}_{0,k_x,1}$ related to exchange and correlations. We further assume that equation (13) is also valid for $r_0 \sim 1$ and $\epsilon^{(1)ec}_{0,k_x,1}/|\epsilon^{ex}_{0,k_x,1}| \gtrsim 1$. For the general case we assume that $\epsilon^{(1)ec}_{0,k_x,1}$ is given by

$$\epsilon^{(1)ec}_{0,k_x,1} = \epsilon_{II}^{(1)ec}(k_x) + (v_g^H/v_g)\epsilon_I^{(1)ec}(k_x) \qquad (14)$$

where $\epsilon_{II}^{(1)ec}(k_x)$ is given by equation (12) and v_g^H/v_g takes into account the slope $v_g = \hbar^{-1}(\partial E_{0,k_x,1}/\partial k_x)_{k_x=k_F}$ of the total energy dispersion. Equation (14) can be well justified for strong B and $r_0 \le 1$ when the dependence of $\epsilon^{ec}_{0,k_x,1}$ on k_x is smooth on the scale of $1/l_0$. As for equation (13), its validity for $r_0 \sim 1$ is not obvious beforehand because both $\partial\epsilon^{(1)ec}_{0,k_x,1}/\partial k_x$ and $\partial\epsilon^{ex}_{0,k_x,1}/\partial k_x$ are logarithmically divergent at $k_x = k_F$. However, numerical calculations show that for $r_0 \sim 1$, $\epsilon^{ec}_{0,k_x,1}$, as given by equation (13), depends smoothly on k_x even in the edge region $k_x \approx k_F$.

In the strong magnetic field limit we have $v_g^H/v_g \approx 1$ if the confining potential is not too smooth. In the general case we should solve the equation

$$v_g = v_g^H + (\partial/\hbar\,\partial k_x)[\epsilon^{ex}_{0,k_x,1}(1 - \epsilon^{(1)ec}_{0,k_x,1}/\epsilon^{ex}_{0,k_x,1})^{-1}]_{k_x=k_F} \qquad (15)$$

for v_g and then substitute it in equation (14).

Equations (13)–(15) are based essentially on (i) the assumption that the energy dispersion is smooth on the scale of $1/l_0$ as well as for $k_x \approx W/2l_0^2 \approx k_F$, (ii) the fact that $\epsilon_{I,II}^{(1)ec}(k_x)$ is practically independent of the changes of the eigenfunctions, when the smooth, on the scale of l_0, confining potential changes while W is fixed, and (iii) on the strong dependence of $\epsilon_I^{(1)ec}(k_x) \propto 1/v_g$ on v_g.

In the spirit of the local-density approximation (LDA) we assume that the energy dispersion relation given by equations (13)–(15) can be obtained approximately by solving the single-particle Schrödinger equation (for $\sigma = 1$) with the Hamiltonian $\hat{h} = \hat{h}^0 + V_{XC}(y)$, where the self-consistent exchange-correlation potential is

$$V_{XC}(y) = \epsilon^{ec}_{0,y/l_0^2,1} \qquad (16)$$

and the function $\epsilon^{ec}_{0,x,1}$ is determined by equations (13)–(15). Assuming that $V_{XC}(y)$ is smooth on the scale of l_0 we find that the corresponding energy dispersion is given again by equations (13)–(15) for the lowest occupied spin-polarized LL. This confirms the self-consistency of our approximate study of the present many-body problem. However, in contrast with the LDA, our $V_{XC}(y_1)$ depends essentially on the slope $\mathrm{d}[V_y + V_{XC}(y)]/\mathrm{d}y$ at the edges $y \approx \pm y_0(k_F)$. Because of that we refer, for a strong magnetic field, $r_0 \lesssim 1$, to the approximation involved in equations (13)–(16) as a modified LDA.

4.2. Comparison with the experiment

Now we apply our theory to the experiments of [7] in GaAlAs/GaAs QWs for which $g_0 = -0.44$. The estimated QW parameters [7] for sample 1 are $W \approx 0.3$ μm, $\hbar\Omega \approx 0.65$ meV, a linear density $n_L = n_S W \approx 7 \times 10^6$ cm^{-1}, where n_S is the strong B 2D electron density, and the $\nu = 1$ plateau is absent. For sample 2 they are $W \approx 0.33$ μm, $n_L \approx 5 \times 10^6$ cm^{-1}, and $\hbar\Omega \approx 0.26$ meV. In this sample, with smaller n_L, $\hbar\Omega$, but almost the same W, a wide $\nu = 1$ plateau develops and is centred at $B = 7.3$ T; this gives $\omega_c/\Omega \approx 45$ and $r_0 \approx 1.0$.

In figure 1 we show the energies, measured from the bottom of the ($n = 0, \sigma = -1$) LL assumed empty, for sample 2. In curve 1 we plot $E_{0,k_x,-1}/\hbar\omega_c = \tilde{k}_x^2/2(25)^2$ for the $\sigma = -1$ LL. Curve 2 shows $E_{0,k_x,1}$ obtained

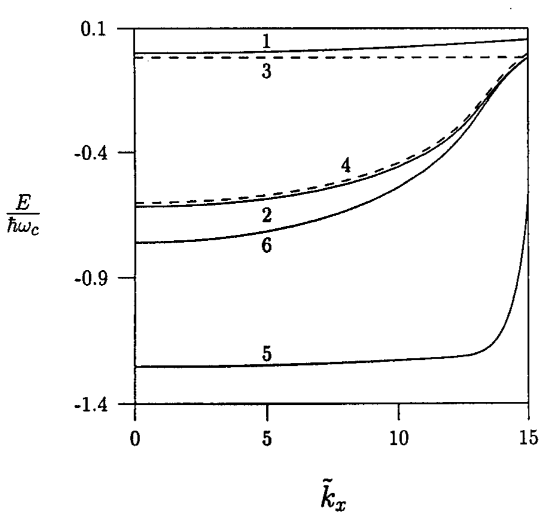


Figure 1. Energies as a function of $\tilde{k}_x$ for the parameters of sample 2 of [7] with $\nu = 1$. The various curves are explained in the text.

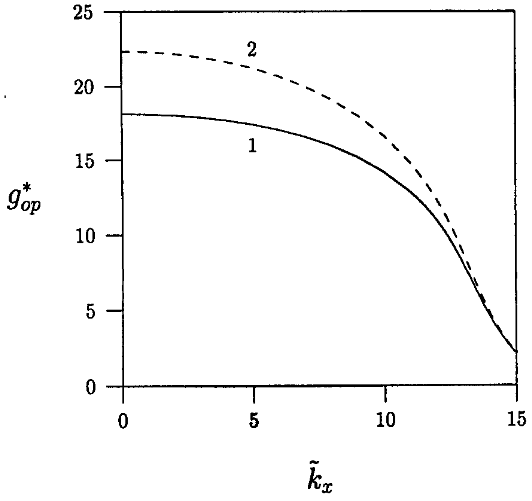


Figure 2. Effective, spatially inhomogeneous g factor g^*_{op} as a function of $\tilde{k}_x$.

from equations (13)–(15), i.e. in the MLDA in which $v_g^H/v_g \approx 0.2$; curve 4 shows the same quantity without the bare spin splitting $g_0\mu_B B$. Curve 3 is the Fermi level when only the states with $\tilde{k}_x \le \tilde{k}_F = 15$ in curve 2 are occupied. Curve 5 represents the $\sigma = 1$ LL without correlations, i.e. $E^{(0)}_{0,k_x,1} = \epsilon_{0,k_x,1} + \epsilon^{ex}_{0,k_x,1}$. In the MLDA we calculate a finite gap between the $\sigma = 1$ LL and the $\sigma = -1$ LL. In contrast, for sample 1 we find, as observed, no gap despite the assumption that the $\sigma = -1$ LL is empty. This contradiction is strong and implies that the $\sigma = -1$ LL must be at least partially occupied for a thermodynamically stable state to exist. This in turn indicates that the $\nu = 1$ quantum Hall effect (QHE) state cannot be realized in such a system. Such an argument has been fully developed in [11]. As for curve 6, it shows $E_{0,k_x,1}$, as curve 2, but without correlations related to screening in the bulk, $\epsilon_{II}^{(1)ec}(k_x)$. It is seen that $\epsilon_{II}^{(1)ec}(k_x)$ has a weak influence on $E_{0,k_x,1}$ especially near the edges.

Finally, in figure 2 we plot the effective, spatially inhomogeneous g factor $g^*_{op} = (E_{0,k_x,-1} - E_{0,k_x,1})/\mu_B B$ as a function of $\tilde{k}_x$. Curve 1 is obtained from curves 1 and 2 in figure 1 and curve 2 from curves 1 and 6. We call the g factor $g^*_{op}(k_x) \equiv g^*_{op}[y_0(k_x)]$ 'optical' because it involves spin-split states with the same k_x. It is seen that g^*_{op} is essentially spatially inhomogeneous and near the edges it can be suppressed very strongly. This g^*_{op} can be substantially different from that deduced from the activated behaviour of the conductance [7].

5. Concluding remarks

The above treatment was mainly devoted to QWs with width $W \lesssim 0.3$ μm. However, the main results can be directly extended to the regions close to the edges of substantially wider channels. This holds when the confining potential, without many-body interactions, can be approximated by V_y'. For such channels the g factor g^*_{op} is essentially spatially inhomogeneous in the range of many l_0 from the channel edge. It is also strongly suppressed in this region owing to strong correlations.

In the experiment of [7] a well-defined gap develops between the two spin states of the lowest LL for sample 2 but not for sample 1. Again, this agrees with our findings, detailed further in [11], and indicates, within the model, that the correlations can destroy the QHE. Certainly, the strength of the correlations depends on the filling factor ν as that of the screening does. In this regard it will be interesting to see their effect on the energetics for higher values of ν, e.g. $\nu = 2$ and $\nu = 3$.

Acknowledgments

This work was supported by the Ukrainian SFFI Grant 2.4/665 (OGB) and by the Canadian NSERC Grant OGP0121756 (PV).

References

[1] Chklovskii D B, Shklovskii B I and Glazman L I 1992 *Phys. Rev.* B **46** 4026
[2] Dempsey J, Gelfand B Y and Halperin B I 1993 *Phys. Rev. Lett.* **70** 3639
[3] Brey L, Palacios J J and Tejedor C 1993 *Phys. Rev.* B **47** 13 884
[4] Muller G, Weiss D, Khaetskii A V, von Klitzing K, Koch S, Nickel H, Schlapp W and Losch R 1992 *Phys. Rev.* B **45** 3932
[5] Kinaret J M and Lee P A 1990 *Phys. Rev.* B **42** 11 768
[6] Suzuki T and Ando T 1993 *J. Phys. Soc. Japan* **62** 2986
[7] Wrobel J, Kuchar F, Ismail K, Lee K Y, Nickel H, Schlapp W, Grobecki G and Dietl T 1994 *Surf. Sci.* **305** 615
[8] Kallin C and Halperin B I 1984 *Phys. Rev.* B **30** 5655
[9] Ando T, Fowler A B and Stern F 1982 *Rev. Mod. Phys.* **54** 437
[10] Balev O G and Vasilopoulos P 1994 *Phys. Rev.* B **50** 8727
Balev O G and Vasilopoulos P 1993 *Phys. Rev.* B **47** 16 410
[11] Balev O G and Vasilopoulos P 1997 *Phys. Rev.* B **56** 6748

Semicond. Sci. Technol. **13** (1998) A104–A106. Printed in the UK
PII: S0268-1242(98)91851-X

A.c. transport and collective excitations in a quantum point contact

I E Aronov†, N N Beletskii‡, G P Berman§, D K Campbell‖, G D Doolen§, S V Dudiy‡ and R Mainieri§

† Sage Technology Inc, 835 Franklin Court, Suite A, Marietta, GA 30067, USA
‡ Institute of Radiophysics and Electronics, National Academy of Sciences of Ukraine, 12 Acad Proskura St, 310085, Kharkov, Ukraine
§ Theoretical Division and CNLS, Los Alamos National Laboratory, Los Alamos, NM 87545, USA
‖ Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green St, Urbana, IL 61801-3080, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We study the a.c. transport through a two-dimensional quantum point contact (QPC) using a Boltzmann-like kinetic equation derived for the partial Wigner distribution function. An integral equation for a potential inside a QPC is solved numerically. It is shown that the electric field inside a QPC is an inhomogeneous function of the spatial coordinate, with a characteristic scale equal to the distance between the electron's turning points. A dependence of the admittance on the frequency of the a.c. field is found in the frequency range, $\omega \approx 0$–50 GHz. The contribution to the imaginary part of the admittance due to the open and closed channels is numerically calculated. It is shown that the crossover of quantum capacitance and quantum inductance from localized behaviour to distributed behaviour takes place at $\omega \sim 10$ GHz. A transition from 2D plasmons to quasi-1D plasmons is analysed as a function of two dimensionless parameters: $k_x d_0$ (where k_x is the longitudinal wavevector and d_0 is the width of the QPC) and the number of open electron channels, N.

Recent technological progress in manufacturing nanoscale solid-state structures has made it possible to fabricate semiconductor devices which operate in the quantum ballistic regime. One system which has attracted much attention is the two-dimensional ballistic quantum point contact (QPC) [1–5]. A QPC (see figure 1) is one of the basic elements of integrated nanocircuits. It provides the interfaces between nano-devices. A QPC displays unusual behaviours and can itself be considered to be a nano-device. For example, by variation of the gate voltage, one can change the boundaries (shape) of the QPC, causing step-like oscillations of the quantum conductance [4, 5]. Recently, it was discovered that, at low enough frequency, a QPC demonstrates not only the quantum peculiarities in the d.c. transport but also has quantum inductance [3] and quantum capacitance [2, 3]. This allows one to consider a QPC (in the low-frequency regime) as an extremely small elementary circuit, ~ 0.1–$10\ \mu$m. In this paper, we present the results of numerical simulations of the a.c. transport through a QPC (in the x-direction in figure 1), for a wide range of frequencies. We also investigate the crossover behaviour of collective excitations in a QPC, from 2D plasmons to quasi-1D plasmons.

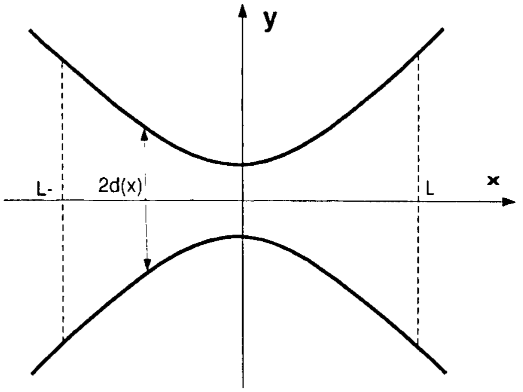


Figure 1. The geometry of the microconstriction. The width is denoted by $2d(x)$, the narrowest width is $2d_0$, and the effective length is $2L$.

In an adiabatic geometry which is smooth on the scale of the Fermi wavelength, the longitudinal and transverse motions of electrons can be separated [1]. In this case, the transverse energy, $\varepsilon_n(x) = \pi^2 n^2 \hbar^2 / 8 m d^2(x)$, plays the role of the effective potential energy for 1D longitudinal motion in the nth channel. (For the numerical simulations we assume a QPC boundary of the form $d(x) = d_0 \exp(x/\tilde{L})^2$.) At low temperatures ($T \ll \mu$, where μ is the chemical

0268-1242/98/SA0104+03$19.50 © 1998 IOP Publishing Ltd

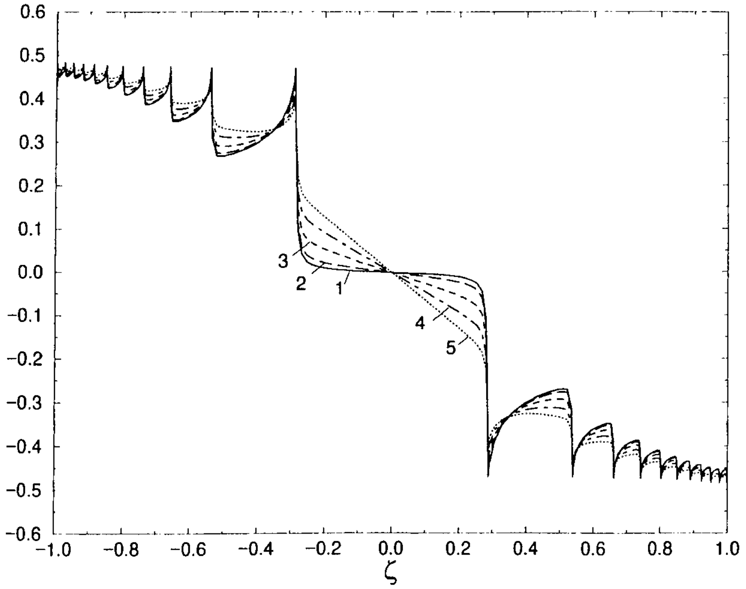


Figure 2. Distribution of the real part of the electric potential, $\Phi'(\zeta)$ (normalized by V) inside a QPC ($\zeta = x/L$), for five values of the frequency (curves 1–5). The values of parameters are as follows: $q = 1.7$; $N = 1$; $\tilde{N} = 11$; $L/\tilde{L} = 1.4$; $\tilde{L} = 10$ μm; $d_0 = 1.5 \times 10^{-6}$ cm; $k_F = 1.5 \times 10^{6}$ cm^{-1}; $v_F = 2.6 \times 10^{7}$ cm s^{-1}; dielectric constant $\epsilon = 13$; curve 1, $\omega = 0$ GHz; curve 2, $\omega = 9$ GHz; curve 3, $\omega = 19$ GHz; curve 4, $\omega = 28$ GHz; curve 5, $\omega = 37$ GHz.

potential), one has a separation between propagating (open) channels ($\varepsilon_n(0) < \mu$), and reflecting (closed) channels ($\varepsilon_n(0) > \mu$) [3]. It was shown in [3] that in the low-frequency regime the trajectories belonging to the open channels give a contribution both to the real part of the admittance (conductance) and to the inductance. The trajectories belonging to the closed channels may have turning points inside the constriction shown in figure 1 which are determined by the equation $\varepsilon_n(x_n) = \mu$. Under some conditions [3] these trajectories give a contribution to the capacitance. The number of open channels is $N = [q]$, where $q = 2k_F d_0/\pi$, and $\hbar k_F = (2m^*\mu)^{1/2}$ (k_F and m^* are the Fermi wavevector and the effective electron mass). The integer, $\tilde{N} = [q \exp(L/\tilde{L})^2] - N$, is the number of closed channels.

In this paper we use the formalism developed in [3] to calculate the admittance, $Y = I_{tot}/V$, in a wide range of frequencies ($\omega \approx$ 0–50 GHz), where I_{tot} is the total current through the QPC and V is the potential difference between the left ($x = -L$) and the right ($x = L$) edges of the constriction shown in figure 1. It is assumed that, inside the QPC, the electric field has the form $E(x,t) = -[\partial\Phi(x)/\partial x]\exp(-\mathrm{i}\omega t)$. The potential, $\Phi(x)$, inside the QPC is determined by the solution of the integral equation. (A procedure for deriving this equation will be discussed elsewhere [6].) As one can see from figure 2, the real part of the potential has characteristic peaks at the turning points $\zeta_n = x_n/L$ corresponding to the closed channels. The real part of the admittance (conductance), y', is shown in figure 3(a) ($Y = (2e^2/h)y$). Curves 1–3 correspond to different values of the parameter q. For all cases shown in figure 3, the number of open channels is $N = 2$. One can see that, at $\omega = 0$, the conductance assumes a familiar form, $y' = 2$. When the frequency increases, the dependence $y'(\beta)$ ($\omega = (v_F/L)\beta = 26$ GHz $\times\ \beta$, where v_F is the Fermi velocity) becomes nonlinear owing to the contribution of both open and closed channels. More complicated behaviour $y'(\beta)$ can be observed in the neighbourhood of q approximately equal to an integer. We shall discuss this behaviour of the admittance elsewhere [6]. In figure 3(b), we demonstrate the dependence of the imaginary part of the admittance, y'', on frequency. One can see that the low-frequency approximation is valid in the region $\omega < \omega_c \approx 10$ GHz. In this frequency region, and for q not close to an integer, a QPC can be considered as an effective circuit [3]. For $\omega > \omega_c$, one should take into consideration the nonlinear dependence $y''(\beta)$. The contribution to the admittance, $y(\beta)$, of both open and closed channels is significant in the whole frequency region. One can see from figure 3(b) that the effects connected with the imaginary part of the admittance can be large. For example, at $\beta = 1$ $y''/y' \approx 0.8$. Effects connected with the nonlinear behaviour $y(\beta)$ are of significant importance in systems with integrated nano-devices. We hope that observation of these effects can be experimentally realized soon.

Finally, for a QPC of a strip form with the width d_0, we have found that a crossover of the surface plasmon spectrum from 2D to quasi-1D is described by two dimensionless parameters: (1) the number of open channels, N; (2) the dispersion parameter, $k_x d_0$, where k_x is the plasmon wavenumber. In a quasi-2D case, when $N \gg 1$, and in the long-wave regime ($k_x d_0 \ll 1$), the spectrum of plasmons is of an acoustic type. In the short-wave regime ($k_x d_0 \gg 1$), the spectrum of plasmons exhibits a square-root behaviour. In the quasi-1D case ($N \geq 1$), the spectrum of plasmons is of an acoustic type. A detailed analysis of the crossover of the surface plasmon spectrum from 2D to quasi-1D is presented in [7].

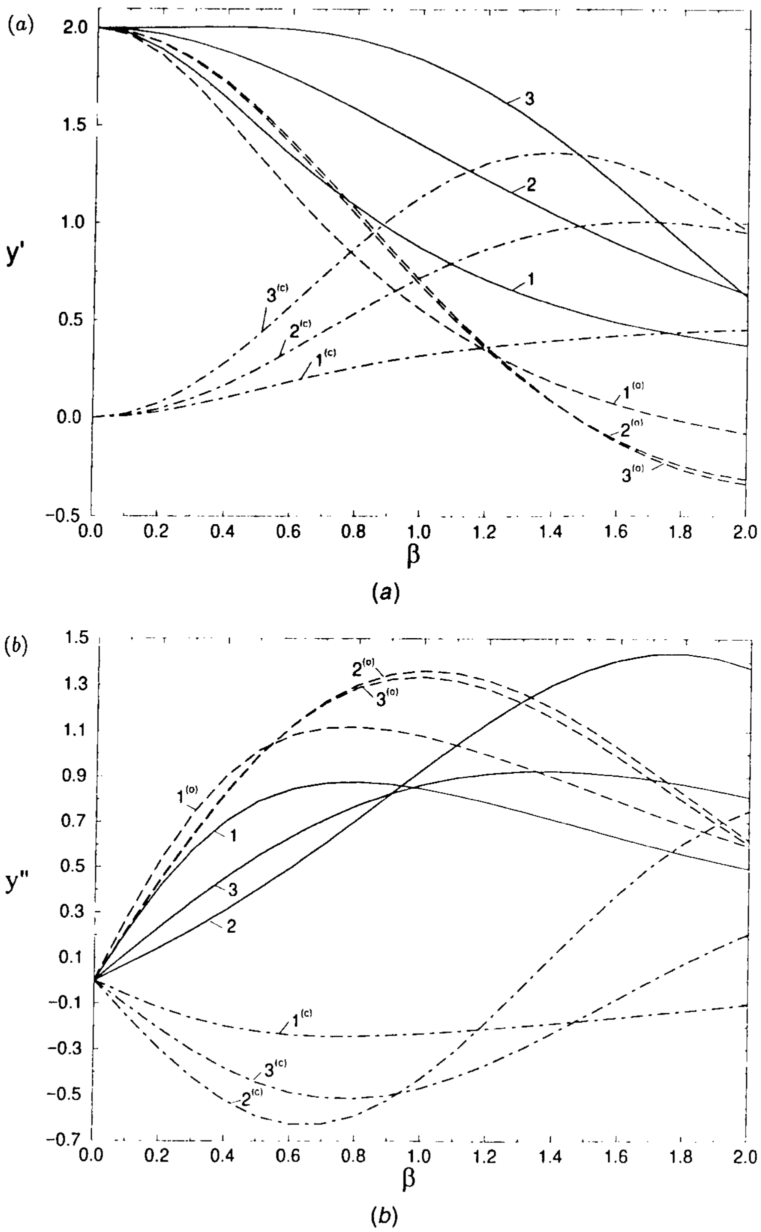


Figure 3. Frequency dependence of the admittance $y = (h/2e^2)Y$: (*a*) y', real part, and (*b*) y'', imaginary part of the admittance, for different values of q and $\tilde{N}$. $N = [q] = 1$; $L = \tilde{L} = 10$ μm; curves 1, $q = 1.1$, $\tilde{N} = 2$; curves 2, $q = 1.7$, $\tilde{N} = 3$; curve 3, $q = 1.75$, $\tilde{N} = 3$. The superscripts (0) and (*c*) indicate the contributions to the admittance of open and closed channels; $\omega = (v_F/L)\beta = 26$ GHz × β.

Acknowledgments

We are grateful to D K Ferry for fruitful discussions. This research was supported in part by the Linkage Grant 93-1602 from the NATO Special Programme Panel on Nanotechnology. Work at LANL was supported by the Defense Advanced Research Projects Agency.

References

[1] Glazman L I, Lesovik G B, Khmelnitskii D E and Shekhter R 1988 *JETP Lett.* **48** 238

[2] Christen T and Büttiker M 1996 *Phys. Rev. Lett.* **77** 143

[3] Aronov I E, Berman G P, Campbell D K and Dudiy S V 1997 *J. Phys.: Condens. Matter* **9** 5089

[4] van Wees B J, van Houten H, Beenakker C W J, Williamson J G, Kouwenhoven L P, van der Marel D and Foxon C T 1988 *Phys. Rev. Lett.* **60** 848

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

[6] Aronov I E, Beletsky N N, Berman G P, Campbell D C, Doolen G D and Dudiy S V 1998 *Phys. Rev.* B accepted

[7] Aronov I E, Campbell D K, Berman G P, Doolen G D and Dudiy S V 1988 *Physica* B accepted

Semicond. Sci. Technol. **13** (1998) A107–A110. Printed in the UK PII: S0268-1242(98)91889-2

Single-electron tunnelling device with variable environmental impedance

F Wakaya†, F Yoshioka†, S Iwabuchi‡, H Higurashi§, Y Nagaoka‖ and K Gamo†¶

† Department of Physical Science, Graduate School of Engineering Science, Osaka University, 1-3 Machikaneyama-cho, Toyonaka, Osaka 560, Japan
‡ Department of Physics, Faculty of Science, Nara Women's University, Kitauoya-Nishimachi, Nara 630, Japan
§ Advanced Semiconductor Devices Research Laboratories, Research and Development Centre, Toshiba Corporation, 1 Komukai Toshiba-cho, Saiwai-ku, Kawasaki 210, Japan
‖ Faculty of Engineering, Kansai University, 3-3-35 Yamate-cho, Suita, Osaka 564, Japan
¶ Research Centre for Materials Science at Extreme Conditions, Osaka University, 1-3 Machikaneyama-cho, Toyonaka, Osaka 560, Japan

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. A single-electron tunnelling (SET) device consisting of a double junction and an environmental impedance is considered. Current–voltage characteristics are calculated using a self-consistent treatment of the environmental impedance. It is observed experimentally that Coulomb gaps become wide for large environmental resistances, which is in agreement with the theory. This modulation of the Coulomb gap by the environmental resistance can be a new operation principle of SET devices.

1. Introduction

A single-electron tunnelling (SET) device [1] is a promising future electron device which may be used for logic and memory devices. Many researchers, therefore, have made efforts to reveal the physics of the SET and to realize new electron devices based on SET events.

It is well known that the Coulomb blockade, which is the origin of SET events, is strongly affected by an environmental impedance (electromagnetic environment effect) [2, 3]. Transport properties of SET devices, therefore, should be understood through a self-consistent treatment of charged states of the island [4, 5].

Higurashi *et al* have developed a self-consistent treatment of charged states of the island with an inductive environmental impedance [5]. It is, however, necessary to introduce an ohmic resistance to realize the feedback of the electromagnetic environment effect in practical experimental setups. In [6] we extended the theory to include an ohmic resistance as well as the inductance, which is important for observing the effect of the environmental impedance modulation experimentally, and suggested the possibility of an SET device which utilizes the environmental impedance modulation as a control method for the SET (Z-SET).

Although there are some experimental reports on results in high environmental impedances [7–9], most experimental research on SET devices has been performed in low-impedance environments. The reason is that the environmental impedances in the actual experiments become low very often owing to the stray capacitance of leads which are connected to the SET devices. In [10], we discussed the effects of the stray capacitance on the environmental impedance modulation and estimated the upper limit of the stray capacitance for the impedance modulation.

The purpose of this paper is to report the experimental results of the current–voltage (I–V) characteristics of the Z-SET, the environmental impedance of which is tunable by applying a gate voltage, and to discuss them with the theory.

2. Environmental impedance modulation

Figure 1(*a*) shows a schematic of the Z-SET consisting of a double-tunnel junction with junction resistances $R_T^{(i)}$ and capacitances C_i and an environmental impedance $Z(\omega)$. Here μ_i is the chemical potential of the ith electrode and Q_i is the charge of the ith tunnel junction. The island charge q can be defined as $Q_2 - Q_1$. In the present work, the environmental impedance $Z(\omega)$, which consists of an inductance L_{env} and a resistance R_{env}, is tunable to control the SET in the Z-SET. The effect of the capacitive

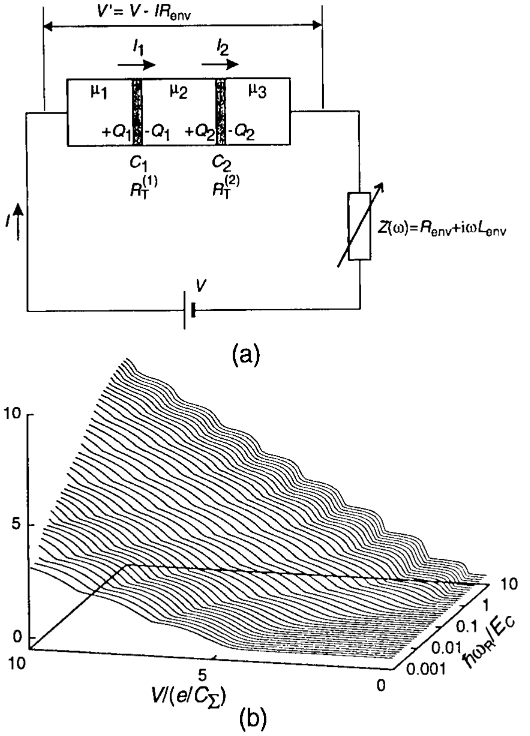


Figure 1. (*a*) A schematic of the Z-SET. Environmental impedance $Z(\omega) = R_{env} + i\omega L_{env}$ is used to control the single-electron tunnelling. (*b*) Calculated current–voltage characteristics for various environmental resistances. $U/k_BT = 25$, $C_1/C_2 = 10$, $R_T^{(1)}/R_q = 1000$, $R_T^{(2)}/R_q = 10$ and $\hbar\omega_L/E_c = 1$ were tentatively employed.

environmental impedance is not taken into account in this study and is discussed elsewhere [10].

Following [5] and [6], the tunnelling current flowing through the ith tunnel junction can be calculated as

$$I_i = \frac{1}{eR_T^{(i)}}\left[\kappa_i eV' + 2U\eta_i\left\langle\frac{q}{e}\right\rangle_q - \Phi_i(eV')\right] \tag{1}$$

where

$$\left\langle\frac{q}{e}\right\rangle_q = \left\lfloor n_c + \frac{1}{2}\right\rfloor + \sum_{n=1}^{\infty}\frac{\sinh(2\beta U\delta n_c)}{\cosh(2\beta U\delta n_c) + \cosh[2\beta U(n-1/2)]} \tag{2}$$

$\langle\cdots\rangle_X$ denotes the trace over the observable X, $\kappa_i \equiv C/C_i$, $\eta_i \equiv (-1)^i$, $U \equiv e^2/2C_\Sigma \equiv (C/C_\Sigma)E_c$, $C \equiv C_1C_2/C_\Sigma$, $C_\Sigma \equiv C_1 + C_2$, $\beta \equiv 1/k_BT$, $n_c \equiv (-\mu_2 + \mu_1 + \kappa_1 eV')/2U$ and $\delta n_c \equiv n_c - \lfloor n_c + 1/2\rfloor$ ($\lfloor\cdot\rfloor$ is Gauss' notation). The function $\Phi_i(eV')$ is defined as

$$\Phi_i(eV') \equiv \frac{i\hbar}{2\pi}\int_{-\infty}^{+\infty} dt \left(\frac{\pi}{\beta\hbar}\,\mathrm{cosech}\,\frac{\pi t}{\beta\hbar}\right)^2 \times[\tilde{\mathcal{F}}_\varphi(it,\kappa_i) - \tilde{\mathcal{F}}_\varphi(-it,\kappa_i)] \times\left\langle \sin\left[\frac{t}{\hbar}\left(\kappa_i eV' + 2U\eta_i\frac{q}{e}\right)\right]\right\rangle_q \tag{3}$$

where

$$\tilde{\mathcal{F}}_\varphi(\tau,\kappa_i) \equiv \exp\left[\kappa_i^2 J(\tau) - \frac{U\tau}{\hbar}\right] \tag{4}$$

$$J(\tau) \equiv \int_0^\infty \frac{d\omega}{\omega}\frac{\mathrm{Re}\,Z_t(\omega)}{R_q} \times\left[\coth\frac{\beta\hbar\omega}{2}(\cosh\omega\tau - 1) - \sinh\omega|\tau|\right] \tag{5}$$

$R_q \equiv \pi\hbar/e^2$, $1/Z_t(\omega) \equiv i\omega C + 1/Z(\omega)$ and $Z(\omega) \equiv R_{env} + i\omega L_{env}$.

If the current continuity condition $I_1 = I_2$ is imposed, the current of the device I is given by

$$I = \frac{1}{eR_\Sigma}[eV' - \Phi_1(eV') - \Phi_2(eV')] \tag{6}$$

where $R_\Sigma \equiv R_T^{(1)} + R_T^{(2)}$. Note that $\Phi_i(eV')$ must be evaluated using n_c determined by the current continuity condition.

Using these equations, we can numerically determine I–V characteristics for arbitrary environmental impedance. Figure 1(*b*) shows the calculated I–V characteristics. $U/k_BT = 25$, $C_1/C_2 = 10$, $R_T^{(1)}/R_q = 1000$ and $R_T^{(2)}/R_q = 10$ were tentatively employed. The inductive environmental impedance is fixed to be low ($\hbar\omega_L/E_c = 1$) while the resistive environmental impedance ($\hbar\omega_R/E_c$) is changed to control the current through the device. Here $\omega_L \equiv 1/(L_{env}C)^{1/2}$ and $\omega_R \equiv 1/R_{env}C$ are frequencies of the environmental modes in the system. As shown in the figure, the size of the Coulomb gap increases as the environmental impedance increases. This is due to the feedback of the electromagnetic environment effect onto the charged states of the island, by which we can control the SET of the Z-SET [6, 10]. The slope of the I–V curve outside the Coulomb gap linearly reflects the total resistance $R_T^{(1)} + R_T^{(2)} + R_{env}$ except for the variation peculiar to the Coulomb staircases, since we have calculated the current within the lowest order of the tunnelling. Such a lowest-order calculation is valid for relatively small bias voltages.

Assuming $C_1 = 500$ aF, $C_2 = 50$ aF, $R_T^{(1)} = 1000R_q$, $R_T^{(2)} = 10R_q$, which are typical for devices using GaAs/AlGaAs and metal Schottky gates, the actual values for important parameters in figure 1(*b*) are $T \simeq 67$ mK, $L \simeq 3$ nH, $e/C_\Sigma \simeq 0.3$ mV and $e/R_\Sigma C_\Sigma \simeq 22$ pA. Concerning the other horizontal axis in the figure, $\hbar\omega_R/E_c = 10$ corresponds to $R_{env} \simeq 822\ \Omega$ and $\hbar\omega_R/E_c = 0.001$ corresponds to $R_{env} \simeq 8.22$ MΩ. These values suggest that we can observe experimentally the environmental impedance modulation.

3. Experiments

In order to observe I–V characteristics controlled by resistive environmental impedance, we fabricated a Z-SET device whose micrograph taken by secondary electron microscopy (SEM) is shown in figure 2. The starting material was a modulation-doped GaAs/AlGaAs heterostructure. The mobility and the density of the two-dimensional electron gas at $\simeq$50 mK are 1.1×10^5 cm^2 V^{-1} s^{-1} and 4.1×10^{11} cm^{-2}, respectively. The ohmic contacts were formed by evaporating Au/Ni/AuGe and annealing at 460 °C for 240 s. Conventional electron beam lithography at an energy of 50 keV, AuPd evaporation

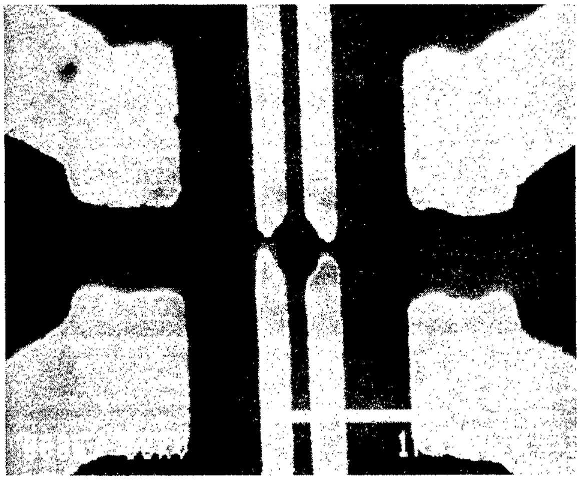


Figure 2. SEM micrograph of the fabricated device.

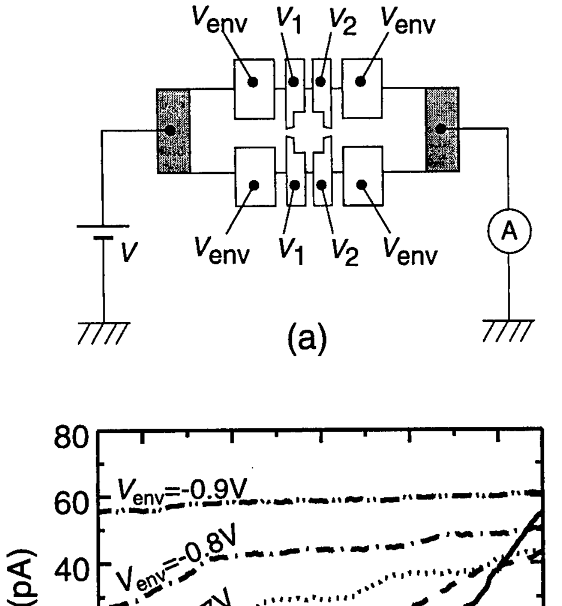


Figure 3. (*a*) Schematic top view of the fabricated device and measurement setup. (*b*) Observed current–voltage characteristics for various environmental resistances, which are controlled by applying V_{env}. The currents are offset (multiples of 15 pA) for clarity. V_1 and V_2 are fixed at −0.621 V and −0.662 V, respectively, during this measurement.

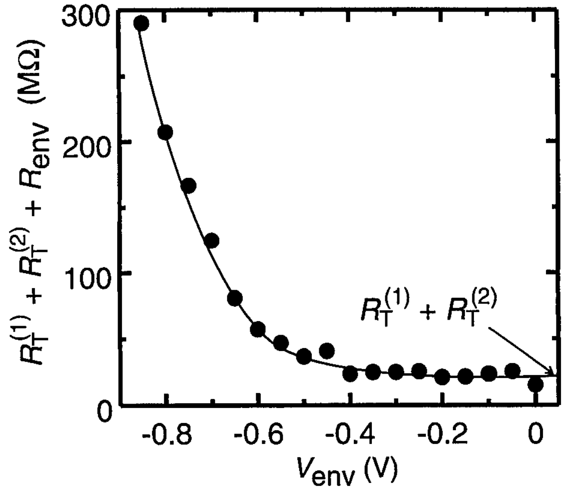


Figure 4. Total resistances $R_T^{(1)} + R_T^{(2)} + R_{env}$, which are estimated from I–V curves, as a function of V_{env}.

and a successive lift-off process are performed to fabricate the Schottky gates.

The experimental setup for low-temperature measurement is shown in figure 3(*a*). One of the ohmic contacts was biased by a dc voltage source, while the other is grounded through a current amplifier. The voltages V_1, V_2 and V_{env} were applied to each Schottky gate relative to the ground level. Here, the negative biases V_1 and V_2 are applied to form a quantum dot and V_{env} is applied to vary the environmental impedance. The measurement was performed at $\simeq$30 mK using a dilution refrigerator.

Figure 3(*b*) shows measured I–V characteristics for several environmental impedances, which were controlled by applying a negative gate voltage V_{env}. The voltages V_1 and V_2 were fixed at −0.621 V and −0.662 V, respectively. The Coulomb gap becomes wide when the environmental impedance increases, i.e. $|V_{env}|$ increases, as predicted.

As mentioned in the previous section, the slope of the I–V curve outside the Coulomb gap for the small bias voltages reflects the total resistance $R_T^{(1)} + R_T^{(2)} + R_{env}$. In order to estimate $R_T^{(1)} + R_T^{(2)} + R_{env}$ experimentally, we read the slope of I–V curves. The results obtained are shown in figure 4. As shown in the figure the environmental impedance is $\simeq$0 in the voltage range -0.4 V $\lesssim V_{env} < 0$ V. The resistance $R_T^{(1)} + R_T^{(2)}$ is, therefore, $\simeq$21 MΩ, which is high enough for Coulomb blockade to be observed. Subtracting $R_T^{(1)} + R_T^{(2)} \simeq 21$ MΩ from the estimated total resistance shown in figure 4, we can obtain R_{env} as a function of the applied gate voltage V_{env}. The environmental resistance R_{env} can be transformed to $\hbar\omega_R/E_c$ using the relation $\hbar\omega_R/E_c \equiv 2R_q/\pi R_{env}$. Thus we can obtain the size of the Coulomb gap as a function of $\hbar\omega_R/E_c$ as shown in figure 5. This figure should be compared with the shape of the Coulomb gap region in figure 1(*b*). The fluctuation of the data in figure 5 is probably due to the offset charge in the island, which is assumed to be zero in the calculation. Comparing figures 1(*b*) and 5, the shapes or tendencies of both figures are seen to be the same, while the absolute values are different. The reason of this deviation is probably as

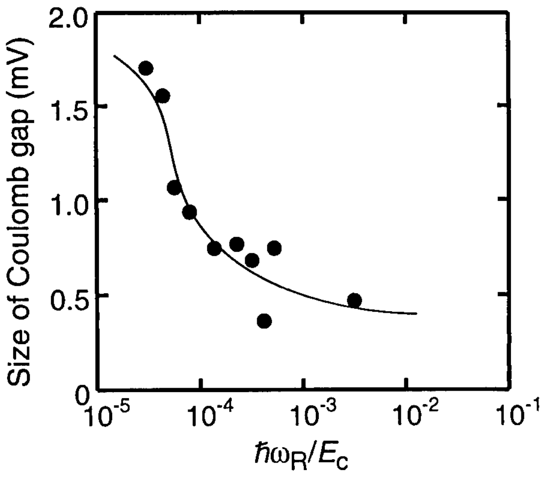


Figure 5. Sizes of Coulomb gaps, which are estimated from I–V curves, as a function of $\hbar\omega_R/E_c$.

follows. It is difficult to estimate precisely parameters such as C_1/C_2, $R_T^{(1)}/R_q$ and $R_T^{(2)}/R_q$ in the present measurement. These parameters, therefore, are not the same values between the calculation and the experiment. Moreover, the stray capacitance may affect the results. The estimation of the above parameters and the calculation using them are left for future study.

In the Coulomb gap, there are no currents in principle. In the present experiment, however, we observed a current of $\lesssim 1$ pA in the Coulomb gaps, which is due to the finite temperature and co-tunnelling. The environmental resistance R_{env} in the high-impedance case is of the order of 100 MΩ. Since these values lead to a voltage drop of 0.1 mV at R_{env}, the observed enlargement of the Coulomb gap is partially attributed to this voltage drop, which is several 0.1 mV at most. Nevertheless, we believe that the environmental impedance modulation effect was observed in the experiment since the observed enlargement of the Coulomb gap is $\simeq 1$ mV.

4. Summary

We have calculated I–V characteristics of an SET device with an environmental impedance (Z-SET) using the self-consistent microscopic theory of Coulomb blockade and suggested a new control method for SET devices by the environmental impedance modulation. We have observed experimentally that the Coulomb gap becomes wide when the environmental resistance increases as predicted by the theory. This is due to feedback of the electromagnetic environment effect onto the charged state of the island. This modulation of the Coulomb gap by the environmental resistance can be a new operation principle for SET devices. This control method reduces the total capacitance of the island and leads us to high-temperature operation of SET devices, since a control gate which is capacitively coupled to the island is not necessary. Furthermore, introduction of this new control method makes the design of SET circuits more flexible in combination with already known single-electron control methods such as capacitive and resistive couplings to the island.

Acknowledgments

The authors express their gratitude to Dr T Sakamoto, Fundamental Research Laboratories, NEC Corporation, for technical advice on the measurement system. This work was supported, in part, by a Grant-in-Aid for Scientific Research in Priority Areas, 'Quantum Coherent Electronics', from the Ministry of Education, Science, Sport and Culture (MESSC). The authors are grateful for the support given by the Industry–University Joint Research Programme 'Quantum Nanoelectronics' through MESSC.

References

[1] Grabert H and Devoret M H (eds) 1992 *Single Charge Tunneling* (New York: Plenum)
[2] Devoret M H, Esteve D, Grabert H and Ingold G-L 1990 *Phys. Rev. Lett.* **64** 1824
[3] Ingold G-L and Grabert H 1991 *Europhys. Lett.* **14** 371
[4] Iwabuchi S, Higurashi H and Nagaoka Y 1993 *Proc. 4th Int. Symp. on Foundations of Quantum Mechanics (ISQM) (Tokyo, 1992) Japan. J. Appl. Phys.* series 9, p 126
[5] Higurashi H, Iwabuchi S and Nagaoka Y 1995 *Phys. Rev.* B **51** 2387
[6] Wakaya F, Iwabuchi S, Nagaoka Y and Gamo K 1997 *Abstracts of Silicon Nanoelectronics Workshop (Kyoto)*
[7] Cleland A N, Schmidt J M and Clarke J 1991 *Phys. Rev.* B **45** 2950
[8] Popović D, Ford C J B, Hong J M and Fowler A B 1993 *Phys. Rev.* B **48** 12 349
[9] Shimazu Y, Ikehata S and Kobayashi S 1995 *J. Phys. Soc. Japan* **64** 2738
[10] Wakaya F, Iwabuchi S, Higurashi H, Nagaoka Y and Gamo K *Solid State Electron.* to be published

Semicond. Sci. Technol. **13** (1998) A111–A114. Printed in the UK PII: S0268-1242(98)91890-9

Room-temperature Coulomb-blockade-dominated transport in gold nanocluster structures

L Clarke†, M N Wybourne‡, L O Brown§, J E Hutchison§, M Yan§, S X Cai§ and J F W Keana§

† Department of Physics, University of Oregon, Eugene, OR 97403, USA
‡ Department of Physics and Astronomy, Dartmouth College, Hanover, NH 03755, USA
§ Department of Chemistry, University of Oregon, Eugene, OR 97403, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. In this paper we discuss the near-room-temperature electrical transport characteristics of structures made from ligand-stabilized metal clusters. The structures show threshold behaviour, nonlinear current–voltage characteristics and radio-frequency-induced plateaux consistent with Coulomb-blockade-dominated transport in disordered arrays of metal dots. Samples having triphenylphosphine and octadecanethiol ligand shells are found to have a 3 orders of magnitude difference in current above threshold. We discuss a possible explanation for this observation.

1. Introduction

One single-electron phenomenon that can be understood from classical electrostatics is the Coulomb-blockade effect in metal systems. For a metallic object of capacitance C, the energy associated with the transfer of one electron from a reservoir to the object is $E_C = e^2/2C$. When this energy is large compared with the available thermal energy, kT, and the tunnel resistance between the object and its surroundings, R_T, is much greater than the quantum resistance h/e^2, the transport properties of the system become strongly influenced by the discrete nature of the electron charge. Coulomb-blockade effects have been observed at room temperature in very-low-capacitance semiconductor devices [1]. A metallic system that offers both the small capacitance and the resistive isolation needed for room-temperature Coulomb blockade is ligand-stabilized metal clusters that contain a small number of atoms [2–5].

One nanocluster that has received particular attention is dodeca(triphenylphosphine)hexa(chloro)pentaconta gold, $Au_{55}[P(C_6H_5)_3]_{12}Cl_6$. This cluster has a metal core diameter of 1.4 nm and a total diameter of 2.1 nm [2]. Using the classical expression for capacitance of an isolated conducting sphere, $C = 4\pi\varepsilon\varepsilon_0 r$, with a dielectric constant $\varepsilon < 5$ the Coulomb charging energy of the Au_{55} core is estimated to be much greater than kT at room temperature. Furthermore, the ligand shell is expected to meet the requirement that $R_T >> h/e^2$.

A useful feature of ligand-stabilized nanoclusters is the lability of the ligand shell, which may be used to tailor transport properties of the system [6]. For instance, when exposed to octadecanethiol ($SC_{18}H_{37}$) groups, the $[P(C_6H_5)_3]_{12}Cl_6$ ligand shell undergoes substitution to form a cluster with approximately 20 stabilizing thiol ligands. Investigations have shown the octadecanethiol stabilized cluster to be more stable in solution, and thus it may provide a more robust platform for room-temperature Coulomb-blockade structures [6].

In this paper we report the transport characteristics of $Au_{55}[P(C_6H_5)_3]_{12}Cl_6$ and octadecanethiol-stabilized Au_{55}. We present data on two different sample configurations. In the first, electron-beam lithography is used to delineate the samples [7], while in the second a solution containing the clusters is drop cast on a prefabricated electrode array. In both methods we observe nonlinear current–voltage (I–V) characteristics near room temperature that are consistent with Coulomb-blockade-dominated transport. We also present data showing the response of the direct-current (DC) I–V characteristic to the application of an external RF signal.

2. Experiment

The $Au_{55}[P(C_6H_5)_3]_{12}Cl_6$ material was synthesized using the Schmid procedure [8] and purified to remove monomer material. Thin films for the first type of sample were

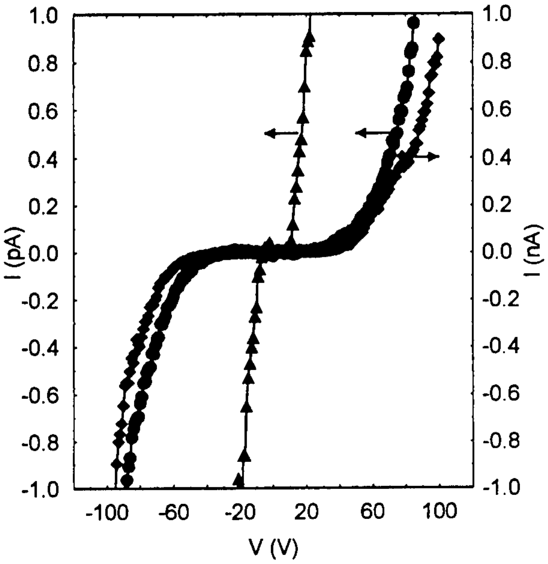


Figure 1. Current–voltage characteristics of patterned ▲ and non-patterned ● $Au_{55}[P(C_6H_5)_3]_{12}Cl_6$ clusters and octadecanethiol-stabilized Au_{55} clusters (♦). The patterned sample was measured at 195 K; the other two were measured at 295 K.

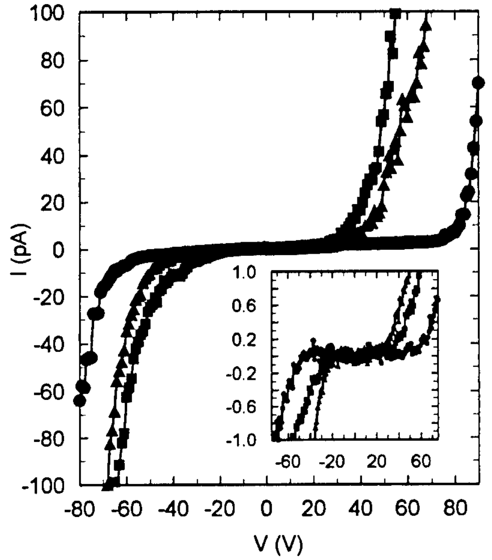


Figure 2. Threshold voltage shift as a function of voltage sweep for the non-patterned octadecanethiol-stabilized Au_{55} clusters. The inset shows the behaviour for the $Au_{55}[P(C_6H_5)_3]_{12}Cl_6$ clusters. The current scale for the inset is $\times 10^{-13}$ A. Both samples were measured at 295 K.

produced on Si_3N_4 as described elsewhere [4]. Exposure to a 40 kV electron beam with a line dosage of 100 nC cm^{-1} followed by development in dichloromethane produced well-defined structures with dimensions as small as 0.1 μm. Atomic force microscopy measurements determined the thickness of the structures to be 50 nm. Gold contacts to the patterned samples were fabricated using conventional electron beam lithography, lift-off and thermal evaporation. Non-patterned $Au_{55}[P(C_6H_5)_3]_{12}Cl_6$ samples were fabricated by drop casting onto interdigitated gold electrodes on glass. The individual electrode width and separation were 15 μm. Octadecanethiol-stabilized clusters were synthesized as described elsewhere [6]. Non-patterned samples suitable for electrical measurement were made by drop casting this material onto interdigitated electrodes on glass.

Electrical measurements were made under vacuum in a shielded vessel that was temperature regulated from 195 to 350 K. The samples were mounted on a clean Teflon stage and connected to a DC voltage source and electrometer with rigid triaxial lines. The background leakage current of the apparatus set the minimum resolvable conductance at about 10^{-15} Ω^{-1}. For each measurement, a control experiment was performed to determine the leakage current of the apparatus and sample holder. The I–V characteristics were obtained by correcting for the intrinsic leakage. Constant-amplitude RF signals with frequencies in the range 0.1–5 MHz were applied to some samples via a dipole antenna. The RF coupling between antenna and the sample was not optimized.

3. Results and discussion

Both patterned and non-patterned triphenylphosphine-stabilized samples exhibited highly nonlinear I–V behaviour, as shown in figure 1. For the patterned sample, clear blockade was seen at 195 K with current suppression up to a threshold voltage of magnitude 6.7 ± 0.6 V. For the non-patterned sample, the blockade behaviour was seen at 295 K. The threshold voltage in the non-patterned samples decreased as a function of the time under bias, as illustrated in the inset to figure 2. Threshold behaviour was also seen in the octadecanethiol sample at 295 K, as

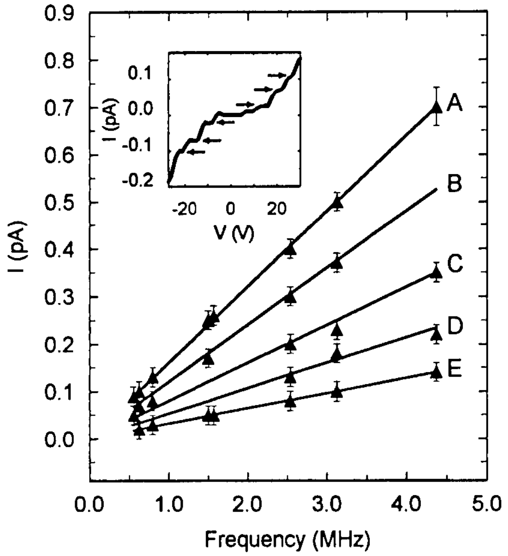


Figure 3. Current values of the observed plateaux as a function of applied radio frequency measured in the patterned $Au_{55}[P(C_6H_5)_3]_{12}Cl_6$ clusters at 195 K. The full lines have slopes (A) e, (B) $3e/4$, (C) $e/2$, (D) $e/3$ and (E) $e/5$, where $e = 1.6 \times 10^{-19}$ C. The inset is the plateau structure at $f = 0.626$ MHz. The arrows indicate the plateau positions.

shown in figure 1. The currents above threshold were much larger than observed in the patterned and non-patterned triphenylphosphine samples. The size of the threshold voltage in the octadecanethiol systems also decreased as a function of time bias was applied, as seen in figure 2. RF signals introduced constant-current plateaux in the I–V characteristics of all samples; however, the threshold instability of the drop-cast samples prevented a detailed study of this effect. A characteristic plateau structure in a patterned $Au_{55}[P(C_6H_5)_3]_{12}Cl_6$ sample is shown in the inset to figure 3. The current at which the plateaux occur is proportional to the applied signal frequency. Several constants of proportionality were observed with the largest being $1.59 \pm 0.04 \times 10^{-19}$ C. A plot of observed current position of the plateau versus applied frequency is shown in figure 3.

Several groups have modelled the transport behaviour of ordered and disordered one- and two-dimensional arrays of metal dots [9–12]. As discussed previously [4], the Middleton and Wingreen (MW) model of a disordered array of normal metal dots best describes our data. In this model the threshold voltage, V_T, scales with the number of junctions between source and drain, and above threshold the current is predicted to scale as $I \sim (V/V_T - 1)^\gamma$. Analytically, γ is 1 for one-dimensional arrays and 5/3 for infinite two-dimensional arrays. Numerical simulations find $\gamma = 2.0 \pm 0.2$ for finite two-dimensional arrays. In all three types of samples, the I–V characteristics above threshold scaled as predicted by MW. For the triphenylphosphine-stabilized clusters, $\gamma = 1.6 \pm 0.2$ and 2.1 ± 0.3, for the patterned and non-patterned cases respectively. We interpret these values to indicate that the non-patterned samples are two dimensional while the patterned samples fall somewhere between the one- and two-dimensional regimes. For octadecanethiol-stabilized clusters $\gamma = 2.6 \pm 0.3$. This value falls outside the range predicted by MW and may indicate that the samples have a dimensionality greater than 2. For both types of cluster, the threshold voltages obtained from scaling were consistent with those estimated directly from the data.

The RF-induced response in the triphenylphosphine-patterned samples is similar to that reported in one-dimensional systems [13, 14]. The effect is caused by phase locking of single-electron tunnelling events and the applied RF signal [15]. In principle, phase locking is possible in two-dimensional arrays but could be quite sensitive to inhomogeneity. However, in the case of a very inhomogeneous two-dimensional array with only a single or few percolation paths the one-dimensional description should apply. Phase locking occurs when the nth harmonic of the applied frequency, f, corresponds to the mth harmonic of the frequency of tunnelling in the system. Then the current becomes locked and plateaux are seen at $I = (n/m)ef$. Fits to the data occur for rational fractions, n/m, of 1/5, 1/3, 1/2, 3/4 and 1, as seen in figure 3. Thus, the RF response supports the hypothesis that correlated tunnelling occurs in the samples. By studying the activated behaviour of the current in the Coulomb gap [4], we have argued that transport is dominated by the charging of single Au_{55} cores.

Two differences between the patterned and non-patterned triphenylphosphine samples are the long-term stability of the transport characteristics and the temperature at which clear blockade is seen. The increased stability of the patterned sample is probably the result of a more rigid structure created by the electron-beam irradiation. Such irradiation may crosslink the ligand spheres, thereby locking the metal cores in place. The irradiation seems to lower the characteristic charging energy of the system as evidenced by measurable conduction below the threshold at 295 K. In the case of the non-patterned samples, both triphenylphosphine- and octadecanethiol-stabilized clusters show clear blockade behaviour at 295 K. Thus, another consequence of the crosslinking appears to be an increase in the capacitance of the clusters within the array.

The major difference between the non-patterned triphenylphosphine- and octadecanethiol-stabilized cluster arrays is the magnitude of the current above threshold, as seen in figure 1. Currents observed in the octadecanethiol samples are about 3 orders of magnitude greater than in either type of triphenylphosphine sample although the threshold voltages and the decrease in threshold voltage with measurement are similar in the two materials. The current magnitude is expected to be inversely proportional to the tunnel resistance and proportional to the number of parallel current paths [12]. Therefore, the increased current could be the result of either a lower tunnel resistance or a greater number of paths, or a combination of both effects. The larger γ in the octadecanethiol material may result from higher dimensionality, which is consistent with additional current paths, and thus higher current magnitudes. An effect that could influence both the tunnelling resistance and the dimensionality of the system is the interdigitation of the octadecanethiol chains to form small, three-dimensional aggregates. Some evidence for this mechanism comes from the fact that solutions can only be formed with the addition of heat. Also, unlike the triphenylphosphine samples, when attempts are made redissolve the octadecanethiol samples aggregation is observed. Further work is in progress to determine whether interdigitation is responsible for the enhanced current magnitude in the octadecanethiol-stabilized material.

4. Conclusion

By investigating the current–voltage and RF response of triphenylphosphine and octadecanethiol ligand stabilized gold clusters, we have shown that single-electron effects dominate the near-room-temperature transport. Samples patterned by electron-beam lithography have increased stability over non-patterned samples. We believe that this increased stability is the result of crosslinking between ligand shells which tends to lock the gold clusters in place. We have also found the different ligand shells studied show significantly different currents above threshold. One suggested mechanism for this increase is that the long-chain ligands allow interdigitation which may increase the dimensionality of the sample.

Acknowledgment

This work was supported in part by the Office of Naval research under Contracts N00014-93-0618 and N00014-93-1-1120.

References

[1] Yano K, Ishii T, Hashimoto T I, Kobayashi T, Murai F and Seki K 1994 *IEEE Trans. Electron Devices* **41** 1628

[2] Schon G and Simon U 1995 *Colloid Polym. Sci.* **273** 101

[3] Houbertz R, Feigenspan T, Mielke F, Memert U, Hartmann U, Simon U, Schon G and Schmid G 1994 *Europhys. Lett.* **28** 641

[4] Clarke L, Wybourne M N, Yan M, Cai S X and Keana J F W 1997 *Appl. Phys. Lett.* **71** 617

[5] Andres R P, Bein T, Dorogi M, Feng S, Henderson J I, Kubiak C P, Mahoney W, Osifchin R G and Reifenberger R 1996 *Science* **272** 1323

[6] Brown L O and Hutchison J E 1997 *J. Am. Chem. Soc.* **119** 12384

[7] Yan M, Cai S X, Wu J C, Duchi C A, Kanskar M, Wybourne M N and Keana J F W 1994 *Polym. Mater. Sci. Eng.* **70** 36

[8] Schmid G 1990 *Inorg. Synth.* **27** 214

[9] Bakhvalov N S, Kazacha G S, Likharev K K and Serdyukov S I 1989 *Sov. Phys.–JETP* **68** 581

[10] Bakhvalov N S, Kazacha G S, Likharev K K and Serdyukov S I 1991 *Physica* B **173** 319

[11] Geigenmuller U and Schon G 1989 *Europhys. Lett.* **10** 765

[12] Middleton A A and Wingreen N S 1993 *Phys. Rev. Lett.* **71** 3198

[13] Geerligs L J, Anderegg V F, Holweg P A M, Mooij J E, Pothier H, Esteve D, Urbina C and Deverot M H 1990 *Phys. Rev. Lett.* **64** 2691

[14] Delsing P, Likharev K K, Kuzmin L S and Claeson T 1989 *Phys. Rev. Lett.* **63** 1861

[15] Averin D V and Likharev K K 1991 *Mesoscopic Phenomena in Solids* ed B Al'tshuler, P Lee and R A Webb (Amsterdam: Elsevier)

Semicond. Sci. Technol. **13** (1998) A115–A118. Printed in the UK		PII: S0268-1242(98)91875-2

Single-electron tunnelling transistor in SiGe/Si double-barrier structures

Chomsik Lee

Department of Electronics Engineering, Honam University, Kwangju 506-714, Korea

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. A single-electron transistor (SET) on an SiGe/Si double-barrier structure has been studied with side contact. In Si/SiGe vertical structure transistors we can more easily control the depletion region by gate voltage than in any other three-terminal device. Drain and source regions are one-dimensional and the well region (gate region) of the device is zero-dimensional with reverse gate bias. The resonant tunnelling (one dimension to zero dimension, zero dimension to one dimension) and current–voltage curves are investigated in a two-dimensional quantum model. The effect of delta doping results in shifting and reduction of the peak value of transmission probabilities. We found current enhancement in delta-doped heterostructures at low temperatures. We proved the possibility of an Si single-electron transistor in a two-dimensional simulation of the resonant tunnelling transistor.

1. Introduction

Molecular beam epitaxy of silicon allows the growth of a wide variety of doping profiles with high quality and reproducibility. This technology offers a very appropriate tool for the fabrication of vertical devices with very short channel lengths. The tunnelling transistor results in better transistor performance without a decrease in metal oxide silicon field effect transistor (MOSFET) channel length, allowing an increase in circuit density for microelectronics applications by three-dimensional device integration. The motivation for the development of SiGe devices was the potential to improve device performance by utilizing heterostructures of a single-electron transistor. The recent progress in Si/SiGe heterostructures indicates that a new concept of band engineering can be introduced into the Si ultralarge-scale integrated circuits field. These topics have been reported in many review papers [1, 2]. The first detailed investigation of modulation Si/SiGe heterostructures was done by People *et al* [3]. Two-dimensional electron gases were realized in tensile strained silicon channels between strain-relaxed silicon–germanium barriers grown on Si(100) substrates by molecular beam epitaxy (MBE). Kasper *et al* increased the transit frequency from 20 GHz to 100 GHz using the vertical structure of the SiGe heterostructure bipolar transistor (HBT) [4]. SiGe HBTs have the potential for outstanding analogue and digital or mixed-signal high-frequency circuits widely based on standard Si technology [5]. A minimum noise figure of 0.9 dB at 10 GHz demonstrates the advantage of using MBE samples with steep and high base doping and high germanium contents.

Band engineering of Si/Ge heterostructures can be employed to produce conduction and valence band effects for enhanced carrier mobilities in both bands. Lattice mismatch between Si and Ge is utilized to grow the Si channels under tensile strain on strain-relaxed SiGe alloys. Strained SiGe has a valence band offset which can be used for improved transport of holes, while using strained Si on relaxed SiGe yields a conduction band offset for enhanced electron mobility. This strain is responsible for lifting the six-fold degeneracy of the Si conduction band, resulting in a type II alignment at the Si/SiGe heterostructure. The representative multilayer structure of a resonant tunnelling transistor (RTT) is presented in figure 1. The RTT consists of the following: 50 Å undoped Si buffer layer on n^+-Si substrate (0.5 μm), 50 Å undoped $Si_{0.4}Ge_{0.6}$ barrier, 50 or 60 Å undoped Si well layer, 50 Å undoped $Si_{0.4}Ge_{0.6}$ barrier, 50 Å undoped Si buffer layer followed by a 0.5 μm Si surface layer (n^+). The ohmic contact of the source was used as a shadow mask for non-conformal deposition of chromium on the etched MBE surface. The gate was side contacted over the low-doping concentration region (lower than 10^{10} cm^{-2}) to ensure a large depletion region. Details of the fabrication of such a device have been previously reported [6, 7].

2. Theoretical technology

Silicon and germanium have different lattice constants at room temperature, a 4.17% mismatch. Due to the relatively large lattice mismatch between SiGe and Si, commensurate (defect free) SiGe alloy films cannot be grown on silicon substrates without introducing large amounts of strain. Lattice constants of the strained layer parallel to the interface adjust such that the two materials have perfectly matching lattice constants. Note that a type II lineup sets in for the mole fraction, $x > 0.6$, quantum confinement shift.

0268-1242/98/SA0115+04$19.50

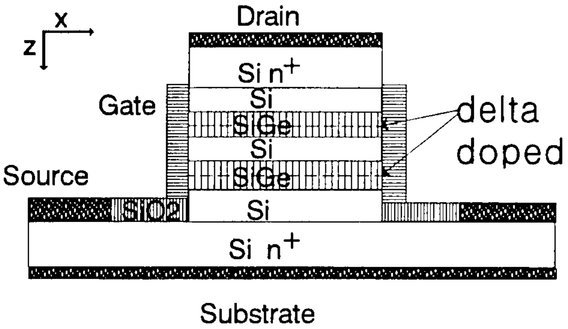


Figure 1. Double-barrier heterostructure of a resonant tunnelling transistor: type II alignment, strained (Si) and relaxed (SiGe) on Si substrate.

In figure 1, numerical solutions to the two-dimensional Poisson equation and the continuity equation are used to calculate the conduction band profile by the finite difference method (FDM). Electron transport in the double-barrier structure is calculated by a self-consistent method [8]. Scattering processes are also included to calculate the resonant tunnelling transmission coefficient. The finite difference method was used to calculate the electrostatic potential (Ψ), the quasi-Fermi level for electrons (Φ), the doping concentration (N_d) and the electron concentration (n). In the numerical analysis, source and drain are described as ohmic contacts. It is assumed that average electron velocity is dependent on the local electric field. We have solved the two-dimensional Poisson equation using the symmetric structure of x and y coordinates [9]

$$\nabla \cdot (\epsilon \nabla \Psi(x, z)) = e(N_d - N_a + p - n - N_\delta) \qquad (1)$$

where ϵ is the dielectric constant, p the hole concentration, N_a the ionized acceptor concentration and N_δ the delta doping concentration.

To calculate the tunnelling current density J_t in [8], continuous variations of potential energies have been split into as many segments as possible. The transmission coefficient can then be found from the incident and transmitted wavevectors and transfer matrix elements. The complete set of eigenfunctions is given by

$$\psi(x, z) = \sum_{n=1} \phi_n(x)\phi_n(z). \qquad (2)$$

Continuity of the total wavefunction $\psi(x, z)$ requires that each propagating wavefunction $\phi_n(z)$ be continuous; $\phi_n(x)$ are uniform plane waves. We assume there is no electron translation in the x direction. The general solution of the Schrödinger equation is a linear combination of a reflected and incident plane wave, in which the incident coefficient is normalized by unity

$$\phi_n = e^{ik_n z} + R\, e^{-ik_n z} \qquad (3)$$

where k_n is a wavevector, R is the reflection coefficient. With the high bias of source–drain or the rapid slope of band bending in the well region, the calculation of the Schrödinger equation is performed using Airy functions. When an electrical field is present in the structure, exact analytical solutions are not available for this problem. Therefore, we adopted a numerical technique based on the Airy function approach [10]. Then, the Schrödinger equation can be written as

$$\phi_n''(\eta) - \eta\phi_n(\eta) = 0 \qquad (4)$$

where $\eta = (2m^* eV_{\mathrm{ap}}/L\hbar)^{1/3}(z-\xi)$, $\xi = (L/eV_{\mathrm{ap}})(V_0 - E_z)$. Here V_0 is the barrier height, eV_{ap}/L is the applied electric field, E_z is the z-direction energy, and L is the total width of the wells and barriers. The solution of equation (4) is readily expressed in terms of Airy functions and its complimentary functions

$$\phi_n = C_n Ai(-\eta) + D_n Bi(-\eta) \qquad (5)$$

where C_n and D_n are constant. The wave equation with the potential discontinuity is given by

$$\phi_n''(\eta) - \eta\phi_n(\eta) - \sum_{n=1}^{M} \Omega_{nm}\phi_m(\eta) = 0 \qquad (6)$$

where M is the total number of modes and

$$\Omega_{nm} = \frac{2m^*}{\hbar^2} \int \phi_n^*(x) V_{\mathrm{sc}} \phi_m(x)\, \mathrm{d}x \qquad (7)$$

where the scattering potential is given by $V_{\mathrm{sc}} = \beta\delta(z)\delta \times (x - x_0)$, Ω_{nm} are mode coupling constants, x_0 is random position and β is the coupling strength. The wavefunction and its first derivative are matched at each interface throughout the structure. By a similar process to that used in [8], the scattering matrix is obtained: $S_j = M(j)M(j)_{\mathrm{sc}}^{-1}$. $M(j)$ is the transfer matrix in region j and $M_{\mathrm{sc}}(j)$ is the transfer matrix in the region of V_{sc}

$$M_{\mathrm{sc}} = \begin{pmatrix} Ai_j(-\eta) & Bi_j(-\eta) \\ Ai_j'(-\eta)/m_w^* + \gamma Ai_j & Bi_j'(-\eta)/m_w^* + \gamma Bi_j \end{pmatrix} \qquad (8)$$

where $\gamma = \sum \Omega_{nm}$. In the case of one propagating mode and one evanescent mode ($E_{x1} < E < E_{x2}$), $E_{x1,2}$ are eigenvalues of the x direction. One subband is occupied in the source region, then the scattering matrix S_j (2×2) is

$$S_j = \begin{pmatrix} 1 & 0 \\ \gamma & 1 \end{pmatrix}. \qquad (9)$$

If $V_{\mathrm{sc}} = 0$, then S_j is a unit matrix.

3. Discussion

We employ a transfer matrix formalism to calculate the transmission probabilities and current densities through the double barrier. We also describe the scattering matrices using the presence of evanescent modes in various lateral confinement configurations. We calculate the effects of delta doping on the current–voltage curves at finite temperature. Elastic scattering mechanisms and evanescent modes are included to calculate the tunnelling transmission coefficient [8]. The band offset of Si/SiGe is $\Delta E(\mathrm{eV}) = 0.5\ [0.31 + 0.53(1 - x)]$, and the dielectric constant is

$$\epsilon = \frac{1 + 2[x\epsilon_1 + (1 - x)\epsilon_2]}{1 - x\epsilon_1 - (1 - x)\epsilon_2} \qquad (10)$$

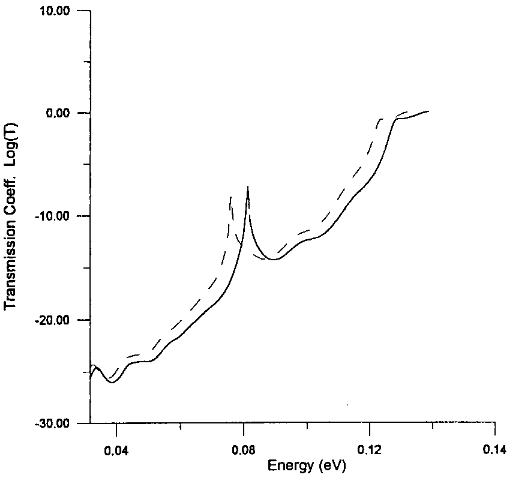


Figure 2. Logarithmic plots of transmission probabilities as a function of electron energy at 77 K, without delta doping (full curve) and with delta doping (concentration 2×10^{11} cm^{-2}) (broken curve). The well width is 42 Å and the barrier widths 51 Å.

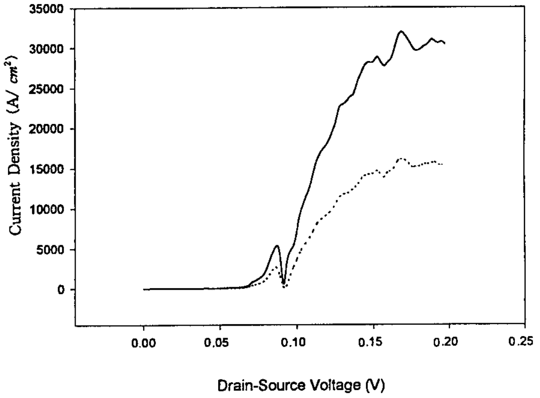


Figure 3. Current–voltage characteristics at 77 K: gate bias −1.5 V (full curve), −2.5 V (dotted curve). The well width is 51 Å and the barrier widths 51 Å; coupling strength $\beta = 0.05$ feV cm^2.

where $\epsilon_1 = \epsilon_{\mathrm{Ge}} - 1/(\epsilon_{\mathrm{Ge}} + 2)$, $\epsilon_2 = \epsilon_{\mathrm{Si}} - 1/(\epsilon_{\mathrm{Si}} + 2)$, ϵ_{Ge} being the dielectric constant of Ge. We have studied a gate-to-gate space of 0.4 μm with delta doping and without delta doping in barriers. In this calculation, the delta doping concentration was 2×10^{11} cm^{-2}, the coupling strength was $\beta = 0.05$ feV cm^2, the barrier height was $V = 0.13$ eV, the mole fraction of Ge was $x = 0.6$, the effective mass of Si electrons at temperature T was $m^*/m_0 = (1.045 + 4.5 \times 10^{-4} T)$. We calculated eigenstates in the quantum well with strong lateral confinement and verified negative differential conductances.

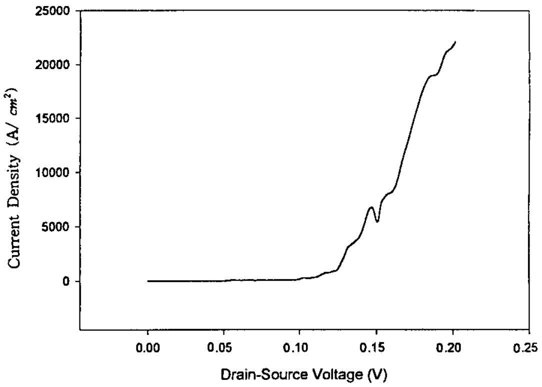


Figure 4. Current–voltage characteristics at 77 K: gate bias −2.1 V (full curve), well width 42 Å, barrier widths 51 Å, coupling strength $\beta = 0.05$ feV cm^2. One quasi state is occupied in the well region.

A plot of transmission probabilities as a function of electron energy at 77 K is shown in figure 2. The full width at half maximum (FWHM) of the resonant transmission is increased with delta doping in figure 2. It was shown that modulation doping causes a strong enhancement of tunnelling probabilities compared with undoped samples. The quasi-bound state in the well region is shifted to lower energy levels. This might be explained by the additional confinement of the injected electrons. The peak-to-valley current ratio will be reduced with delta doping. It is important to notice that the resonance peak shifts monotonically from 0.085 to 0.075 V with delta doping. Figure 3 shows the evidence for subband mixing in the case in which one or more subbands are occupied in the source with barrier widths 51 Å and well width 51 Å. At low bias, the subband is higher than the electron energy, and there is no current flow. As the applied bias voltage is increased, the current begins to flow at a source–drain voltage of $V_{\mathrm{sd}} = 0.06$ V. There is always a continuum of states which can be mixed by other interactions. As the gate bias increases, the source region shows strong quantization. If well width is 60 Å, the resonance peak is shifted from 0.085 to 0.087 V by increasing reverse gate bias $V_g = 1$ V. As reverse gate bias increases from 1.5 V to 2.5 V in figure 3, the current density is reduced by about 50% and quasi-bound states shift to a higher voltage. A single-electron state is observed in figure 4. We used barrier widths of 51 Å, a well width of 42 Å and a gate-to-gate space of 0.4 μm in figure 4. However, the peak-to-valley current ratio is very small. The scattering and delta doping destroy the coherence of the wavefunction so that the peak current from resonances is reduced.

4. Conclusion

In the Si/SiGe RTT, we can easily control the dimensionality of the well region using the field effect. We also present the possibility of an Si single-electron transistor in a two-dimensional simulation of the current–voltage characteristics of the RTT. The presence of delta doping leads to changes in the scattering matrices calculated using Airy functions. Thus, the transmission coefficient is changed dramatically in the presence of delta doping and a strong scattering potential. We find that the resonant peak is substantially shifted to lower bias by delta doping. The full width at half maximum is increased with delta doping.

Acknowledgments

This work was supported by grant no KOSEF 971-0914-083-1 from the Korea Science and Engineering Foundation.

References

[1] Nakagawa K and Miyao M 1991 *J. Appl. Phys.* **69** 3085

[2] Pearsall T 1989 *CRC Solid State Mater. Sci.* **15** 551

[3] People R, Bean J C and Lang D V 1984 *J. Vac. Sci. Technol.* A **3** 846

[4] Kasper E, Kibbel H, Hergoz H and Gruhle A 1994 *Japan. J. Appl. Phys.* **33** 2415

[5] Krishna G, Aditya A, Chakrabarti N and Banerjee S 1995 *IEEE Trans. Comput. Aided* **14** 803

[6] Lee C 1996 *Japan. J. Appl. Phys.* **35** 2583

[7] Austing D, Honda T and Tarucha S 1996 *Solid-State Electron.* **40** 237

[8] Lee C and Weichold M 1997 *J. Appl. Phys.* **81** 8064

[9] Bozler C A and Alley G D 1980 *IEEE Trans. Electron Devices* **27** 1128

[10] Cruz H, Hernandez-Cabrera A and Munoz A 1990 *Semicond. Sci. Technol.* **6** 218

Semicond. Sci. Technol. **13** (1998) A119–A123. Printed in the UK PII: S0268-1242(98)91879-X

Ångström-level, real-time control of the formation of quantum devices

S-B Carlsson, K Deppert, T Junno, M H Magnusson, L Montelius and L Samuelson

Lund University, Solid State Physics/Nanometer Structure Consortium, Box 118, S-221 00 Lund, Sweden

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We report a novel approach for the realization of quantum devices which require device structures of sub-10 nm dimensions and position control better than 1 nm. In this approach we combine three methods from nano-technology: (i) an aerosol technique for the fabrication of metallic and semiconducting nano-crystals or nano-particles with diameters in the range 5–50 nm, (ii) extreme electron beam lithography to define contact gap geometries with dimensions of 10–50 nm and (iii) a manipulation technique based on atomic-force microscopy, combined with *in situ* electrical measurements of the device characteristics, by which pre-fabricated nano-structures can be positioned with high accuracy. We present details of room-temperature measurements on quantized conductance devices, formed in the neck structures between neighbouring gold particles. These necks or wires have a cross-section of only one or a few gold atoms, leading to quantized conductance of $G = n(2e^2/h)$, with values of n between 1 and 10 having been observed. Such lateral quantum resistor devices are found to be remarkably stable, frequently maintaining the conductance levels on the time scale of hours. We also discuss the prospect for novel devices in which a single nano-particle or a single molecule is controllably positioned with high accuracy, with tunnel gaps surrounding the island.

1. Introduction

In the endeavour to realize quantum devices which can operate at room temperature, one crucial aspect is to make the relevant energy quantization, due to either single electron charging or due to quantum confinement, large compared to the thermal energy, kT. In order for the energy quantization in semiconducting quantum dots to exceed kT, this typically requires confining structures to be smaller than about 10 nm. For devices relying on Coulomb blockade, conducting nodes should be smaller than 5 nm in order for $E_C = e^2/2C$ to be much larger than kT. In order for quantization effects in metallic conductance structures to be operational at room temperatures, metallic wires of dimensions of about or less than 1 nm are required. It is clear that straightforward lithography will not easily solve this problem and, therefore, novel approaches for nano-structure fabrication are needed.

In this paper we describe methods to make quantum devices controllably, based on:

- fabrication of mono-disperse nano-crystals and nano-structures [1],
- pre-defined electrodes and gates, made by electron beam lithography, metallization and lift-off, resulting in contact distances controlled to about 10–50 nm,
- atomic-force microscopy to image and manipulate nano-structures [2] and
- *in situ* measurements of electrical characteristics of devices during the manipulation of particles [3].

In section 2, nano-fabrication techniques are summarized, with a fabrication method for making mono-disperse nano-particles in metallic as well as semiconducting material. Results will be presented for different metallic nano-particles, such as silver and indium, as well as for compound semiconductor nano-crystals, for instance gallium arsenide and indium phosphide, all of these particles being in the size-range 5–30 nm in diameter. This is followed by a description of electron beam lithography with metallization and lift-off to produce small electrode gaps, down to 10 nm in width. In this section, finally, the technique we have developed for combined imaging and manipulation of nano-particles as small as 5 nm, using an atomic force microscope (AFM), including *in situ* electrical measurements is described.

In section 3 we first demonstrate a nano-mechanical switch based on the short distance displacement of one nano-particle in a chain. We then present device data for metallic quantum point contact (QPC) devices which can be built with ångström-level control, by which room-temperature operating QPCs with stable conductance levels of $G = n\,2e^2/h$ can be made. Finally in this section

0268-1242/98/SA0119+05$19.50

Figure 1. Transmission electron micrograph of 40 nm Ag aerosol particles.

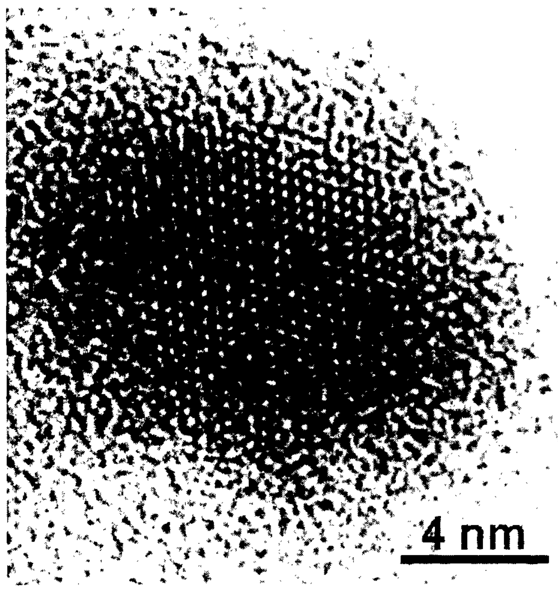


Figure 2. Transmission electron micrograph of a single 10 nm InP aerosol particle.

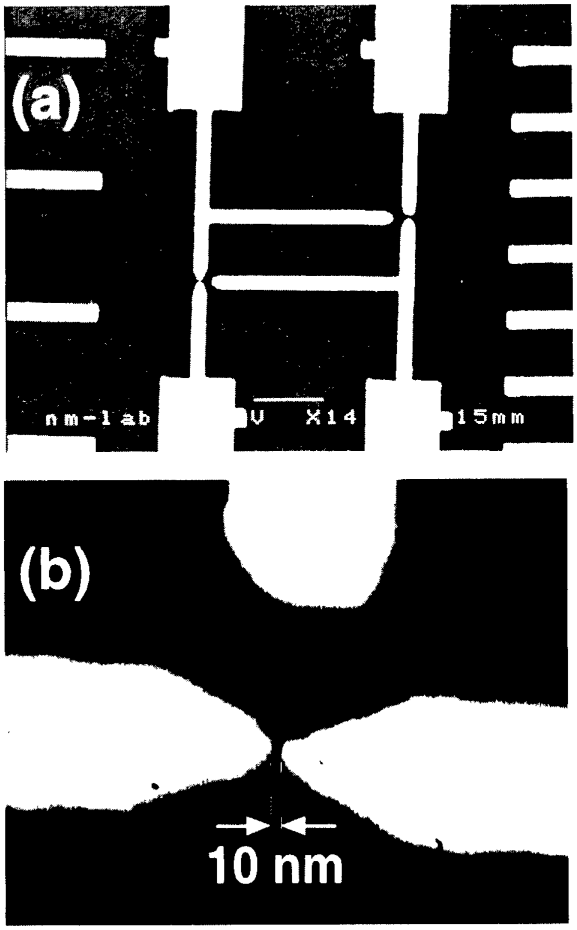


Figure 3. Scanning electron micrographs of e-beam-defined electrodes. (a) Two gaps with two gate electrodes, (b) a 10 nm gap separating two electrodes.

we discuss future possible applications of these techniques for the controlled formation of single-electron tunnelling devices, with manipulation of tunnelling gaps on a sub-nm level, for which preliminary measurements of Coulomb gaps have been made. Similarly we predict that contacting of single molecules will be feasible using our approach.

2. Nano-technology

In order that operation of Coulomb blockade devices at room temperature can be made possible, metallic islands with a diameter of less than 5 nm must be used. We have chosen a nano-particle fabrication approach [1] based on aerosol technology to create such objects. Ideally, a Coulomb blockade device can be made by positioning such a nano-particle between two electrodes, separated by a gap of 5–10 nm. We show here that electrode gaps down to 10 nm can be fabricated by electron beam lithography. Different proposals for the localization of nano-particles relative to such ultra-fine electrodes have been suggested, for example, direct positioning of particles by the electric field from these electrodes [4]. Our approach which we have previously published is the technique for manipulating nano-particles by an AFM in which the AFM acts both as a microscope and a tool for nanometre-scale manipulation [2].

2.1. Aerosol nano-crystal fabrication

For the fabrication of small metal and semiconductor particles we use a technology which is naturally suited for the large scale production of metal droplets of uniform size, namely an aerosol technique. This approach is very good in terms of size distribution and controllability of the process, but also offers the possibility of using a large variety of metals and non-metals that can be deposited on a substrate. The production of the aerosol takes place in a special apparatus. The metal, for instance gallium, indium or silver, is evaporated in a tube furnace and mixed with a clean carrier gas flow. An ultra-fine metal aerosol consisting of particles of nm size is formed by homogeneous nucleation in a cooling section. After a charging process, a monodisperse aerosol fraction, with a particle diameter of 20 ± 1 nm, for instance, is selected using a differential mobility analyser (DMA). A DMA apparatus exploits the fact that the electrical

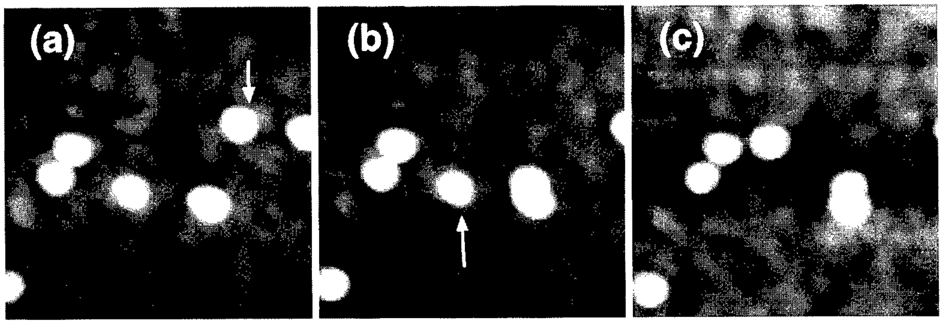


Figure 4. AFM manipulation of 8–10 nm In particles. The arrows indicate the direction of the movement.

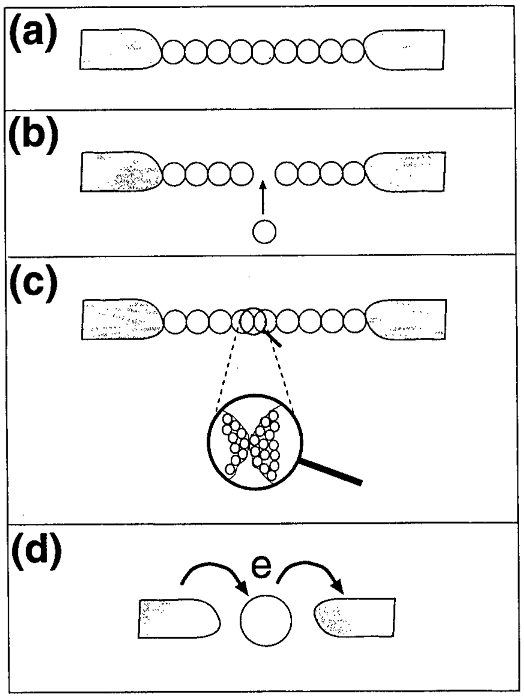


Figure 5. Schematic picture of possible quantum devices. (a) Ohmic nano-bridge and (b) bi-stable nano-mechanical switch. (c) Stable quantized conductance device. (d) Single electron tunnelling device.

mobility of singly charged, nanometre-sized particles is a monotonically decreasing function of particle size. Size selection takes place by balancing the electrical mobility with the force of the gas used to flush out unwanted particles. The selected metal aerosol particles, charged and uniform in size, are carried through a second furnace, which can be used for different purposes, either for reshaping or chemical modification of the particles. During reshaping the metal particles undergo a sintering process, and leave the furnace as perfectly spherical particles as illustrated in figure 1 [5]. For reaction, a gaseous reactant, for example arsine or phosphine, is fed into the furnace where the metal particles react with it to form the desired material. By this process, compound semiconductor nano-crystals have been fabricated by the reaction of size-selected metal particles with hydrides [1]. The stoichiometry of these particles has been verified by electron diffraction and x-ray analysis [6]. Figure 2 shows an example of an InP particle created by this process. The modified particles can be deposited on a substrate by means of an electric field.

2.2. e-beam lithography

Nanometre-scale electrodes were fabricated on a 300 nm thick SiO_2 layer on a Si substrate by e-beam lithography, followed by evaporation of 3 nm Ti (for adhesion) and 25 nm Au. The electrodes were separated by a gap of typically 10–50 nm (figure 3). For the quantized conductance experiments a grid of disc-shaped Au particles, 30–100 nm in diameter and 30 nm thick, were formed in a second e-beam lithography and lift-off step.

2.3. Manipulation of nano-particles

The AFM manipulation of nano-particles is done using the tapping mode for imaging and the contact mode, without feedback control, for repositioning. First, the particles are imaged in tapping mode, a particle is selected and the AFM tip is positioned 10–50 nm behind it with respect to the intended direction of relocation, the feedback is broken and the tip is moved in the vertical direction into contact with the sample surface. By increasing the normal force between the tip and the sample we can vary the amount of lateral force that can be exerted on the particle. The tip is then moved along a line through the centre of the particle. Because the feedback is turned off, the tip will not follow the contours of the particle but will force the particle to move in front of it. The tip is then retracted from the surface, the feedback loop turned on again and an image is scanned to observe the new position of the particle. Aerosol particles as well as e-beam-defined metal discs can be manipulated in this way. The discs appear to stick more strongly to the surface than the aerosol particles. However, in the present setup we cannot measure the lateral forces required to move an object. In figure 4 a sequence of AFM images taken during the manipulation of 8–10 nm aerosol-produced In particles is shown. The particles can be moved in steps as small as 1 Å, which, in combination with real-time monitoring of the electrical device characteristics, can be used to build devices with highly controlled properties.

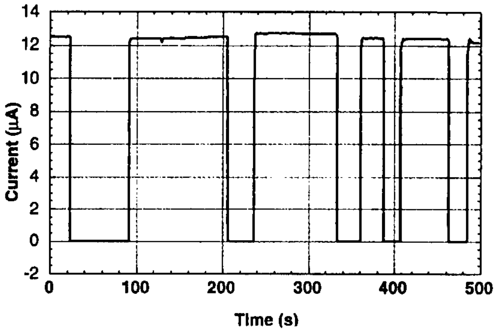


Figure 6. Current as a function of time as one Au particle in a chain between the electrodes is pushed in and out repeatedly. Bias voltage = 5 mV.

2.4. On-line electrical measurements

To perform on-line electrical measurement the sample is placed on a standard DIL chip carrier, bonded and placed in a specially designed AFM sample holder allowing on-line electrical connections. After the manipulation the sample can be transferred to a cryostat or any other measuring setup.

3. Applications

By combining these nano-technologies we have fabricated, in principle, four types of device, as illustrated in figure 5. First, by using Au particles in a chain, it is possible to form an ohmic ultra-narrow wire. Then, by moving one of these particles in and out of the chain, a bi-stable, nano-mechanical switch is formed. Furthermore, at the moment of making or breaking contact a narrow constriction is formed which shows quantum point contact behaviour. This effect could be used for stable quantized conductance devices [7]. By using particles covered with an insulating layer, for instance, an oxide layer or some other coating, one can fabricate single electron tunnelling devices based on the Coulomb-blockade effect. Finally, by starting with e-beam-defined electrodes and then constructing nano-scale contacts of metal particles, it will be possible to make electrical contact with very small objects on the surface, including molecules.

3.1. Nano-bridge and nano-mechanical switch

When an e-beam-defined Au particle is pushed into full contact with the electrodes an ohmic bridge forms between them. It is possible to make ohmic contact with both macroscopic objects as well as with other particles and one can thereby build a smaller gap than originally defined by the electrodes. The same particle can be pushed several times into and out of the gap on a time scale of minutes, as can be seen in figure 6. The resistance in the off-state is in the $T\Omega$ regime and in the on-state typically 120 Ω. Current–voltage measurement in the on-state shows perfect ohmic behaviour for all bias voltages below 2 V.

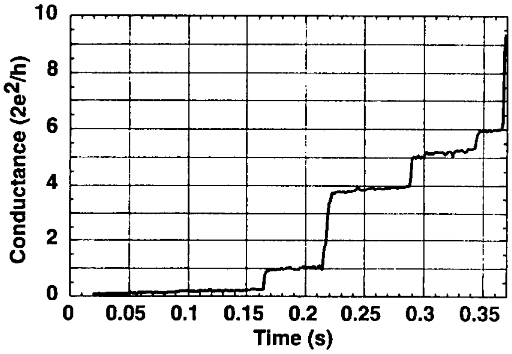


Figure 7. Conductance as a function of time as a Au particle is pushed into contact with the electrodes. Distinct steps corresponding to integer values of the conductance quantum are observed. Bias voltage = 10 mV, $T = 300$ K.

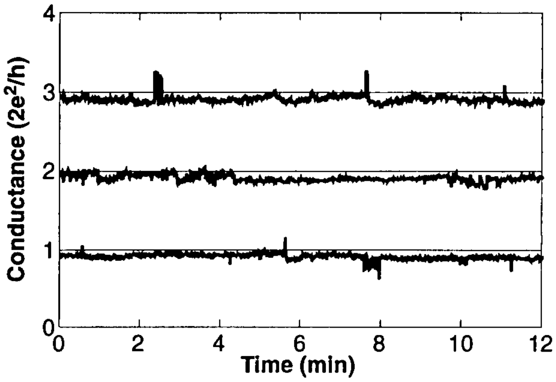


Figure 8. Stable quantum point contacts at quantum levels 1, 2 and 3.

3.2. Quantized conductance

In a constriction of dimensions of the order of the Fermi wavelength for the electrons in the material, quantized conductance can be observed [8–10]. The steps are in integer units of the conductance quantum $G_0 = 2e^2/h$ [11]. In a time-resolved measurement, during which a Au particle is pushed by the AFM tip into contact, these steps are clearly observed (figure 7). By selecting the pushing rate of the AFM tip it is possible to record this behaviour on a time scale ranging from ms to minutes. This effect can be observed both when making or breaking the contact. The bias voltage can be varied below 100 mV without any noticeable change in the measurement. If the particle is pushed slowly towards contact and the conductance is recorded continuously, it is possible to stop the manipulation at any predefined value of the conductance. The tip can then be retracted and the point contact is allowed to self-develop. Usually the conductance of the junction decreases over a period of a few minutes and then stabilizes at an integer value of G_0. Figure 8 shows conductance plateaus 1, 2 and 3, stable for more than 10 minutes. Plateaus stable for more than one hour have been recorded. On these plateaus the junction is quite insensitive to mechanical disturbances in the surroundings, such that it is possible to tap the measurement setup gently without any change in the electrical characteristics. Eventually, the junction normally breaks. It is unclear whether this is due to electromigration in the wire or nano-particle dynamics. Even though the QPC region is ballistic and should not have losses, we expect electromigration to occur very close to this region, when the current is still confined to a cross-section diameter of a few nanometres. In our samples this corresponds to current densities of 10^{10} A m^{-2} which is known to cause electromigration.

From the observation of random-telegraph-like noise on the stable plateaus we speculate that a single atom is flicking between two alternative positions in the QPC region.

3.3. Future applications

So far we have reported only electrical measurements on e-beam-defined particles. However, initial experiments with aerosol particles of In have been performed, showing clear signs of Coulomb blockade at low temperatures. These In particles are expected to have a thin layer of oxide acting as the tunnelling barrier. As this technique allows manipulation of very small particles, it is one candidate for the fabrication of future, room-temperature, single electron transistors [12, 13]. By using Au aerosol particles we expect that quantum point contact devices with better defined constrictions can be made. As an alternative to the aerosol fabrication technique, both for single electronics and quantized conductance devices, one can use colloidal particles which can also be fabricated with a narrow size distribution and with different layers on the surface. Finally, we expect the fields of both molecular and bio-physics to take advantage of this technique, both for basic characterization and device applications, that is, building an ohmic wire that connects to a single molecule.

Acknowledgments

We want to thank Jan-Olle Malm and Chatrin Svensson at Lund University for the transmission micrograph pictures and Hongqi Xu for discussions on the theory for quantum transport. The work was performed within the Nanometer Structure Consortium and was financed by NFR, TFR, NUTEK, SSF and the ESPRIT Charge programme project 22953.

References

[1] Deppert K and Samuelson L 1996 *Appl. Phys. Lett.* **68** 1409
[2] Junno T, Deppert K, Montelius L and Samuelson L 1995 *Appl. Phys. Lett.* **66** 3627
[3] Samuelson L, Carlsson S-B, Junno T, Xu Hongqi and Montelius L 1998 *Proc. NATO Forum—Nano-scale Science and Technology* ed N Garcia *et al* (Netherlands: Kluwer) at press
[4] Bezryadin A, Dekker C and Schmid G 1997 *Appl. Phys. Lett.* **71** 1273
[5] Deppert K, Maximov I, Samuelson L, Hansson H-C and Wiedensohler A 1994 *Appl. Phys. Lett.* **64** 3293
[6] Deppert K, Magnusson M H, Samuelsson L, Malm J-O, Svensson C and Bovin J-O 1998 *J. Aerosol Sci.* **29** 737
[7] Junno T, Carlsson S-B, Xu Hongqi, Montelius L and Samuelson L 1998 *Appl. Phys. Lett.* **72** 548
[8] van Wees B J, van Houten H, Beenakker C W J, Williamson J G, Kouwenhoven L P, van der Marel D and Foxon C T 1988 *Phys. Rev. Lett.* **60** 848
[9] Muller C J, van Ruitenbeek J M and de Jongh L J 1992 *Phys. Rev. Lett.* **69** 140
[10] Pascual J I, Méndez J, Gómez-Herrero J, Baró A M, Garcia N and Vu Thien Binh 1993 *Phys. Rev. Lett.* **71** 1852
[11] Landauer R 1989 *J. Phys.: Condens. Matter* **1** 8099
[12] Averin D V and Likharev K K 1986 *J. Low Temp. Phys.* **62** 345
[13] Fulton T A and Dolan D J 1987 *Phys. Rev. Lett.* **59** 109

Semicond. Sci. Technol. **13** (1998) A124–A126. Printed in the UK PII: S0268-1242(98)91880-6

Coulomb blockade oscillation in a single atomic junction

F Yamaguchi†, D H Huang‡ and Y Yamamoto†§

† ERATO Yamamoto Quantum Fluctuation Project, Edward L Ginzton Laboratory, Stanford University, Stanford, CA 94305, USA
‡ ERATO Yamamoto Quantum Fluctuation Project, NTT Musashino R&D Centre, Musashino, Tokyo 180, Japan
§ NTT Basic Research Laboratories, 3-1 Morinosato-Wakamiya Atsugi, Kanagawa 243-01, Japan

Received 7 December 1997, in final form 11 March 1998

Abstract. The first observation of the Coulomb blockade effect in the smallest possible system with a single atom as the central island of a double-barrier tunnel junction is reported. Our system consists of a single tungsten atom as the central island and a tungsten STM tip and a silicon (100)2 × 1 reconstructed surface as the two electrodes. The use of a single atom as the central island makes the change in the electrostatic potential due to a variation of number of electrons in the island of the order of 1 eV and thus the Coulomb blockade effect is made more controllable and stable even at room temperature. A specific shape of the tip apex forms a tunnel junction between an apex atom and the rest of the tip with the energy-level broadening of the apex atom smaller than the change in the charging energy due to the change in the number of electrons in the single tungsten atom. This theoretical prediction was confirmed by the experimental results of I–V measurements with an STM tip made from a W(111) single-crystal wire where the change in the charging energy is 1.1 eV.

The Coulomb blockade effect can be seen in a system consisting of two electrodes separated by an insulating gap and a third electrode ('central island') in the middle of the gap [1]. If the central island is small, a change in the electrostatic potential due to the change in the number of electrons in the island can be large compared with the thermal energy and the energy-level broadening due to the finite tunnelling rate between the central island and the electrodes. The electron tunnelling events are blocked unless excess energy is supplied by an external voltage source to compensate for this increased potential change. This phenomenon has been extensively studied for the past two decades due to fundamental interest and potential applications to future nanoscale electronic/photonic devices [2–5]. We report here the first observation of the Coulomb blockade effect in a single-atomic-junction (single-tungsten-atom STM tip). The use of a single atom as the central island makes the change in the electrostatic potential due to the variation of the number of electrons in the island greater than 1 eV and thus the Coulomb blockade effect is made more controllable and stable even at room temperature.

Coulomb blockade oscillation shown in figure 1(a) was observed using an ultra-high vacuum (UHV) STM (JSTM-4500XT) with an operating pressure of 2 × 10^{-8} Pa at room temperature. An STM tip which has a pyramidal protrusion structure terminated with a single tungsten atom ('single-atom tip') was prepared by field evaporation from a single-crystal tungsten wire with (111) orientation, following Binh and Purcell [6]. After this critical step, in order to remove oxide layers and other foreign atoms adsorbed on the tip apex during the field evaporation process, the tip was cleaned again with electron bombardment.

A schematic configuration of a single-atom tip with (111) orientation, which is self-organized to minimize the surface energy during the field evaporation process, is shown in figure 2(a). A tight-binding (TB) calculation for the electronic states of a single-atom tip, consisting of a single atom in the first layer, three atoms in the second layer, six atoms in the third layer and a bulk reservoir, shows the presence of electronic states where electrons are localized into the apex atom and depleted from the second and third layers, as shown in figure 3. The field-induced atomic energy shift in the single-atom tip effectively decouples the apex atom from the rest of the tip. The localized electronic state into the apex atom is thus regarded as a central island which capacitively couples to the two electrodes, the rest of the tip, and the silicon surface, with capacitances C_T and C_S, respectively, as shown in figure 2(b), which explains the oscillatory behaviour of the conductance shown

0268-1242/98/SA0124+03$19.50

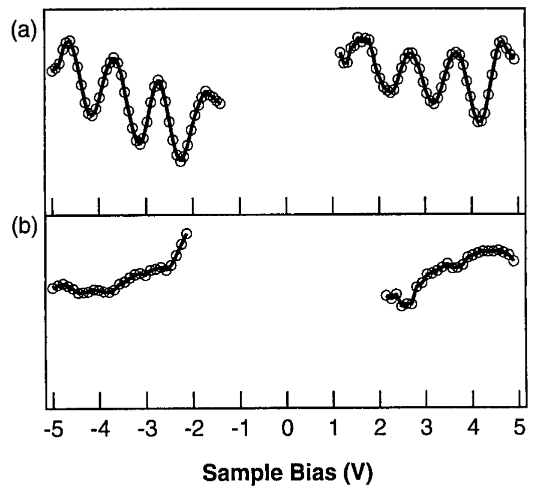


Figure 1. Normalized differential conductance $(\mathrm{d}I/\mathrm{d}V)/(I/V)$ measured by a single-atom tip (a) and by a truncated tip (b) against applied voltage V. Experimental data (a) were obtained with a tip–sample separation corresponding to a tunnel current of 6 nA at an applied voltage of $V = -2$ V. The I–V measurements were performed with a constant tip–sample separation. The observed peak separation of $(e^2/C_T)_{\mathrm{exp}} \simeq 1.1$ eV in the Coulomb blockade oscillation is compared to the theoretical predictions $(e^2/C_T)_{\mathrm{th}} \simeq 1.1$ eV.

in figure 1(a). The charging energy for the 5d orbital of an isolated tungsten atom can be estimated from the linearly increasing third to sixth photoionization energies, about 12 eV [7, 8]. The pyramidal structure of the single-atom tip shown in figure 2(a) reduces the charging energy of an apex atom due to the distributed capacitive coupling to the second, third and fourth layers of the tip. This charging energy is calculated to be $(e^2/C_T)_{\mathrm{th}} \sim 1.1$ eV [9].

When the tunnel rate between the apex atom and the sample (Si surface) is much smaller than that between the apex atom and the rest of the tip, which is the case for the experimental data of figure 1(a), the tunnel current is governed solely by the (slower) apex atom–sample tunnelling event. As a positive sample bias voltage is applied, the Fermi level of the sample, μ_{sample}, moves downward relative to the discrete energy level of the localized state. When μ_{sample} goes across each one of the resonances at voltages $V(n) = e(2n+1)/2C_T + \Delta(n = 0, 1, 2, \ldots)$, the sample starts extracting an electron from the localized state and new channels for current flow between the tip and the sample open up one by one, which results in peaks in the differential conductance, as shown in figure 1(a). Here, Δ is an offset voltage determined by the peak energy shift of the localized state relative to the Fermi level. Peaks in the differential conductance are also seen when a negative sample bias voltage is applied. When μ_{sample} moves upward to cross each one of the resonances at voltages $V(n) = -e(2n+1)/2C_T + \Delta(n = 0, 1, 2, \ldots)$, the sample starts supplying the localized state with an additional electron, which results in peaks in the differential conductance. We could identify the eight differential conductance oscillation peaks for $n = 1$ to $n = 4$. We could not identify the two peaks for $n = 0$ due to the poor signal-to-noise ratio in the I–V measurement in the small bias voltage ($V < 0.5$ V) region. The observed single charging energy $(e^2/C_T)_{\mathrm{exp}} \simeq 1.1$ eV is in fairly good agreement with the theoretical values mentioned above.

The result reported here is the first proof of the existence of Coulomb blockade oscillation in a single atomic tunnel junction. On the fundamental side, such a system provides a unique experimental method to characterize the highest occupied energy level (energy-level broadening $\hbar\Gamma$

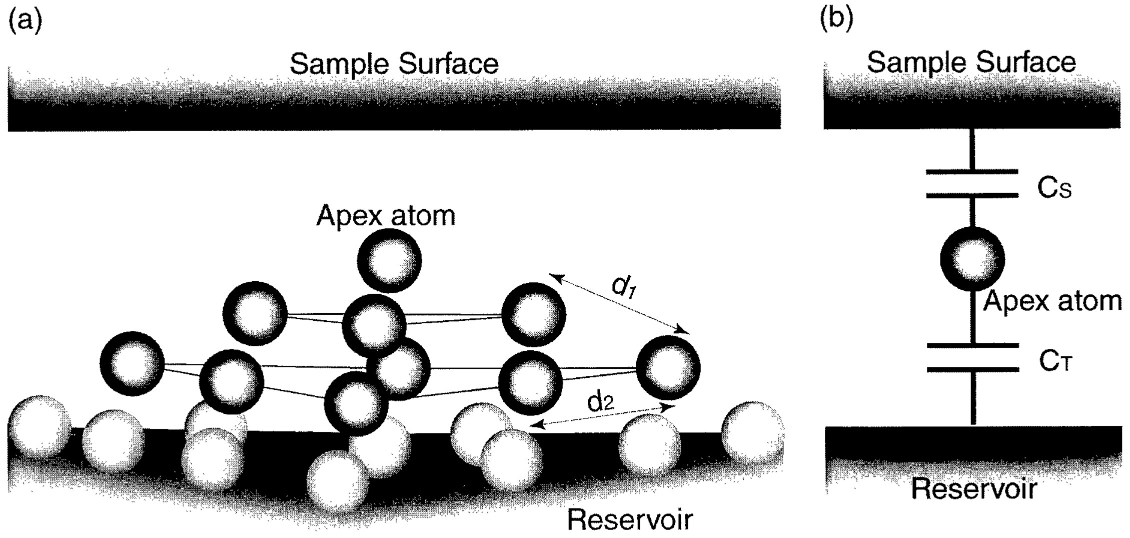


Figure 2. (a) A pyramidal structure of a single-atom tip which has a single atom, three atoms (second layer), six atoms (third layer) and ten atoms (fourth layer). $d_1 = 2.73$ Å and $d_2 = 4.46$ Å. (b) The double-barrier tunnel junction model of a single-atom tip.

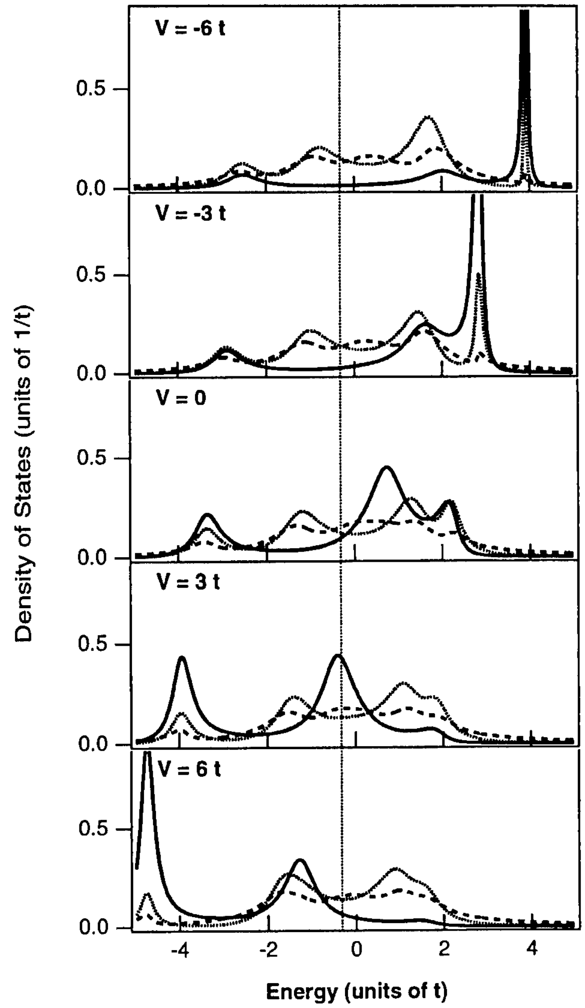


Figure 3. The density of states per atom of a pyramidal tungsten(111) tip in the first layer (solid line), second layer (dashed line) and third layer (broken line) calculated by the tight-binding calculation, in which the field-induced atomic energy shift of each constituent atom is taken into account. The electron energy and applied bias are scaled by the tunnel matrix *t* between the nearest neighbour atoms and the tunnelling rates up to the third nearest neighbour atoms are taken into account.

and single-electron charging energy e^2/C_T) for the single atom adsorbed on the electrode. On the application side, various mesoscopic devices based on the Coulomb blockade effect such as single-electron transistors and logics [2], single-electron turnstile devices [3], single-electron pumps [4], single-photon turnstile devices [5], etc, which require low-temperature operation due to the relatively small charging energy, should be realized rigidly at room temperature on an atomic scale.

References

[1] Devoret M H, Esteve D and Urbina C 1992 Single-electron transfer in metallic nanostructures *Nature* **360** 547–53

[2] Averin D V and Likharev K K J 1986 Coulomb blockade of single-electron tunnelling, and coherent oscillations in small tunnel junctions *Low Temp. Phys.* **62** 345–73

[3] Geerligs L J, Anderegg V F, Holweg P A M, Mooij J E, Pothier H, Esteve D, Urbina C and Devoret M H 1990 Frequency-locked turnstile device for single electrons *Phys. Rev. Lett.* **64** 2691–94

[4] Pothier H, Lafarge P, Orfila P F, Urbina C, Esteve D and Devoret M H 1991 Single electron pump fabrication with ultrasmall normal tunnel junctions *Physica* B **169** 573–4

[5] Imamoglu A and Yamamoto Y 1994 Turnstile device for heralded single photons: coulomb blockade of electron and hole tunnelling in quantum confined p–i–n heterostructures *Phys. Rev. Lett.* **72** 210–13

[6] Binh Vu T and Purcell S T 1997 Field emission from nanotips *Appl. Surf. Sci.* **111** 157–64

[7] Iafrate G J, Hess K, Krieger J Band Macucci M 1998 Capacitive nature of atomic-sized structures *Phys. Rev.* B **52** 10 737–9

[8] Emsley J 1989 *The Elements* (Oxford: Clarendon) p 201

[9] The charging energy of an isolated tungsten atom is estimated as 12.3 eV from the photoionization energy data [8]. This corresponds to the atomic capacitance $C_{5d} = 1.3 \times 10^{-20}$ F, which is equivalent to the self-capacitance of a metal sphere of a radius $a_{5d} = 1.12$ Å. We have calculated the effective charging energy e^2/C of the apex atom based on the distributed capacitively coupled metal spheres aligned in the crystallographic configuration shown in figure 2(a), in which the ten atoms in the fourth layer are assumed to form a grounded reservoir electrode.

Semicond. Sci. Technol. **13** (1998) A127–A129. Printed in the UK

PII: S0268-1242(98)91882-X

Single-photon turnstile device: simultaneous Coulomb blockade for electrons and holes

J Kim†, O Benson†, H Kan‡ and Y Yamamoto†§

† ERATO Quantum Fluctuation Project, E L Ginzton Laboratory,
Stanford University, Stanford, CA 94305, USA
‡ ERATO Quantum Fluctuation Project,
Hamamatsu Photonics Central Research Laboratory, Hamakita, Shizuoka, Japan

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. Utilizing simultaneous Coulomb blockade for electrons and holes in a p–n junction, we can realize a device where a single electron and a single hole are injected into the active region to produce a single photon with well-regulated time interval. The photons emitted from such a device can be studied with single photon counting detectors using a Si solid-state photomultiplier. We report locking of the photon emission with external driving pulse, in the regime where one electron and one hole are injected into the active region on the average.

Semiconductor light-emitting devices convert injected carriers (electrons and holes) to photons by radiative recombination processes at a p–n junction. The statistics of the photons thus generated are mainly governed by the statistics of injected carriers, provided that the conversion efficiency is high [1]. In semiconductor lasers and light-emitting diodes (LEDs) where the quantum efficiency is high enough (30–70%), regulated carrier injection results in the generation of a squeezed state of light [2]. It was recently understood that the charging effect at the junction regulates the carrier injection and produces light with intensity noise less than the standard quantum limit [3].

It was also proposed that such an electron-to-photon conversion process can be manipulated at the single-particle level, where a single electron and a single hole are injected to the active region of an LED to produce a single photon [4]. Such a process can generate a regulated single-photon stream, which provides a fundamental technique for experimental tests of quantum mechanics [5] and future advancement in quantum cryptography and quantum computation. In this paper we present our recent experimental progress in such a device, called a single-photon turnstile device.

A single-photon turnstile device is based on simultaneous Coulomb blockade for electrons and holes in a mesoscopic double-barrier p–n junction. A single electron resonantly tunnels into an electron subband in a central island at a certain bias voltage (figure 1(*a*)). When one electron tunnels, the Coulomb blockade effect inhibits subsequent electron tunnelling. When the bias is increased to satisfy the hole resonant tunnelling condition (figure 1(*b*)), only a single hole is allowed to tunnel into the hole subband of the central island because of the Coulomb blockade effect for the holes. By modulating the bias voltage between the electron and the hole resonant tunnelling conditions periodically, we can inject a single electron and a single hole into the central island periodically, which will be followed by a single photon emission. Figure 2 shows the numerical simulation of such a single-photon turnstile device operation. A single electron and a single hole are

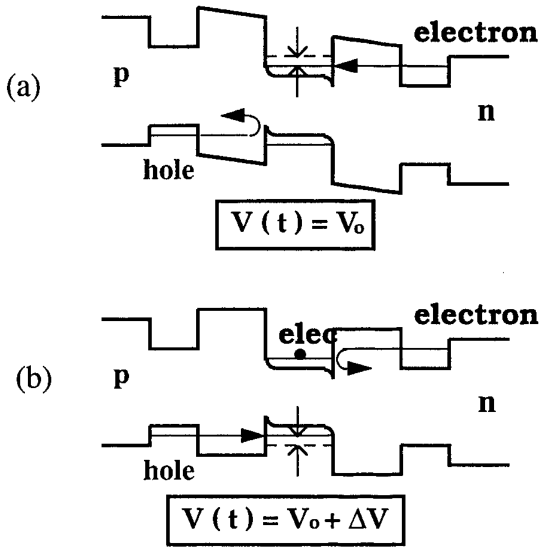


Figure 1. Operation principle of single-photon turnstile device.

---

§ NTT Basic Research Laboratories, Atsugi, Kanagawa, Japan.

0268-1242/98/SA0127+03$19.50 © 1998 IOP Publishing Ltd

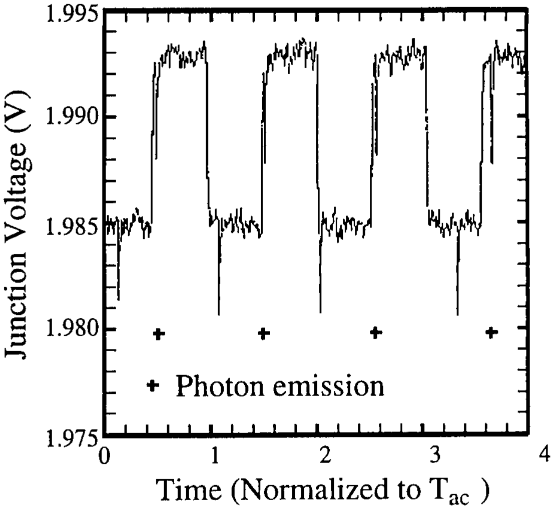


Figure 2. Numerical simulation for the operation of a single-photon turnstile device. The junction voltage is modulated between two values, and a single electron and a single hole are injected per driving cycle, resulting in single-photon emission per driving pulse.

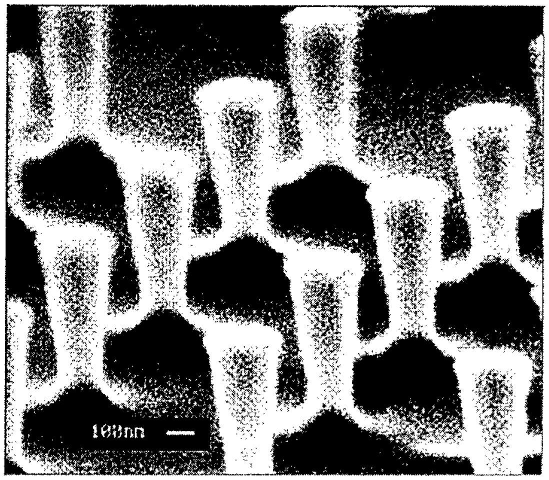


Figure 3. A typical p–n junction post structure. This particular structure is 0.2 μm in diameter.

injected per cycle and a single photon is generated in each cycle.

The device is realized in a GaAs/AlGaAs semiconductor system. A double-barrier p–n junction is carefully designed and grown on a GaAs wafer by the MBE growth technique. Small devices with diameters in the range 0.2–1.0 μm were defined by evaporating 300 Å of Cr/Au by electron beam lithography and liftoff. The Cr/Au was used as the mask to etch posts with a height of ~0.8 μm with a BCl_3/Cl_2 ECR plasma. Typical etched posts are shown in figure 3. The etched surface of the structure was passivated by sulphur and encapsulated by 500 Å of silicon nitride film. The structure was planarized with hardbaked photoresist, and p-type contact pads were evaporated. An n-type contact was formed in the substrate.

The measured current and photocurrent as a function of the bias voltage at 4 K for a device with 0.6 μm nominal size are shown in figure 4. The photocurrent shown in this figure was measured by a large-area avalanche photodiode, normalized by the avalanche gain of ~60. We see the resonant tunnel current and the corresponding photon emission, with negligible background leakage current. The detected photocurrent I_p is given by

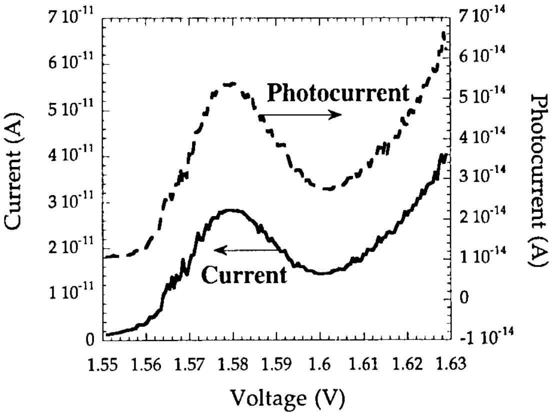


Figure 4. Current and photocurrent as a function of bias voltage for a 0.6 μm device at 4 K. The photocurrent is measured using a large-area avalanche photodiode.

$$I_p = \eta_i \eta_e I_e \tag{1}$$

where η_i is the internal quantum efficiency for conversion of an electron–hole pair to a photon, η_e is the external coupling efficiency for the generated photon to reach the detector and I_e is the electrical current flowing through the single-photon turnstile device. From this measurement, we can conclude that the overall current-to-photocurrent coupling efficiency ($\eta_i \eta_e$) is about 2×10^{-3}. From simple geometric considerations, the external coupling efficiency (η_e) for a photon to escape the single-photon turnstile device and to reach the detector is expected to be less than 2×10^{-2}. This puts the lower bound on the intrinsic electron-to-photon conversion efficiency (η_i) of the device to be $\geq 10\%$.

To confirm the turnstile device operation, we need to measure the emitted photons with high quantum efficiency and good time resolution. We used a Si solid-state photomultiplier (SSPM) as our single photon counting detector. This detector has a quantum efficiency of more than 70%, an electron multiplication factor of ~20 000, a response time of ~2 ns and virtually no excess noise in the multiplication [6]. In our setup the detector is optimized to have a low dark count rate to improve the signal-to-noise ratio at our low signal level, and the time resolution was limited to 40 ns, corresponding to the bandwidth of the low-temperature amplifier used in the experiment.

The single-photon turnstile device and the SSPM single photon counting detector were mounted to a 4 K cryostat in a face-to-face coupling configuration. Since the area of the single photon counting detector was smaller than that of the avalanche photodiode used in figure 4, the photon coupling efficiency from the single-photon turnstile device to the detector was about 10^{-4} in this experiment.

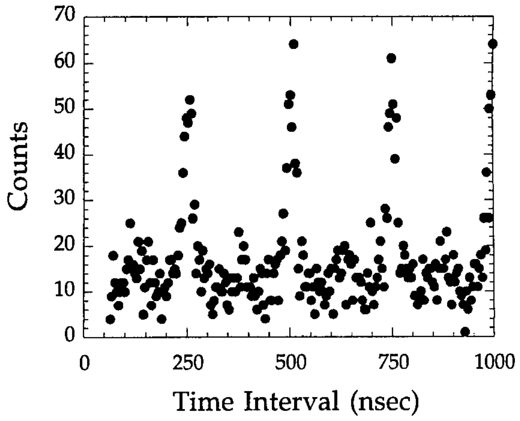


Figure 5. Time interval histogram between two photons.

To operate the device, we modulated the single-photon turnstile device with 50 ns pulses separated by 200 ns. The DC bias voltage was adjusted so that the average current flowing through the device was about 0.8 pA. This current value corresponds to injection of 1.2 electron–hole pairs into the active region per driving pulse on the average. To find out whether the photon emission event is locked to the driving signal, we performed time interval measurement. The signal output from the single photon counting detector was split into two and was used to provide 'start' and 'stop' triggers for a time interval counter. By accumulating the data, we can plot the photon detection counts as a function of the time delay after a single photon is detected. A typical result is shown in figure 5. We find that the time separation between two photons is strongly modulated at the driving frequency. Since we are injecting only about 1 electron–hole pair per driving cycle, only one photon is emitted per driving pulse on the average. This measurement shows that the single-photon emission is locked to the driving frequency.

We would like to point out that this measurement does not confirm any non-classical photon statistics. Similar results are expected if we start with a modulated macroscopic laser or LED and simply attenuate the light until one photon is detected per pulse on the average. However, a mesoscopic light-emitting device as used in our experiment has the potential to generate one and only one photon per driving pulse by utilizing the Coulomb blockade effect, approaching the ideal number state of photons that is otherwise difficult to generate. When such a device is realized, we expect to see non-classical photon statistics such as antibunching and sub-Poissonian distribution.

To summarize, we have fabricated a mesoscopic p–n junction where the electron and hole injection can be regulated to generate a single photon with well-defined time interval. The photons generated by the turnstile device are measured by an SSPM single photon counting detector. We were able to reach the operating regime where single electrons are injected per driving pulse on the average, to generate single photons that are locked to the driving frequency.

References

[1] Yamamoto Y, Machida S and Nilsson O 1986 *Phys. Rev.* A **34** 4025
Yamamoto Y and Machida S 1987 *Phys. Rev.* A **35** 5114
[2] Machida S, Yamamoto Y and Itaya Y 1987 *Phys. Rev. Lett.* **58** 1000
Machida S and Yamamoto Y 1988 *Phys. Rev. Lett.* **60** 792
[3] Imamoḡlu A and Yamamoto Y 1993 *Phys. Rev. Lett.* **70** 3327
Kim J, Kan H and Yamamoto Y 1995 *Phys. Rev.* B **52** 2008
[4] Imamoḡlu A and Yamamoto Y 1994 *Phys. Rev. Lett.* **72** 210
[5] Eberhard P H 1993 *Phys. Rev.* A **47** R747
[6] Kwiat P G, Steinberg A M, Chiao R Y, Eberhard P H and Petroff M D 1993 *Phys. Rev.* A **48** R867
Kim J, Yamamoto Y and Hogue H H 1997 *Appl. Phys. Lett.* **70** 2852

Semicond. Sci. Technol. **13** (1998) A130–A134. Printed in the UK

PII: S0268-1242(98)91835-1

Experimental demonstration of quantum-dot cellular automata

G L Snider, A O Orlov, I Amlani, G H Bernstein, C S Lent, J L Merz and W Porod

Department of Electrical Engineering, University of Notre Dame, Notre Dame, IN 46556, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We present the experimental demonstration of a basic cell of quantum-dot cellular automata (QCA), a transistorless computation paradigm which addresses the issues of device density and interconnection. The device presented is a six-dot quantum-dot cellular system consisting of a four-dot QCA cell and two electrometer dots. The system is fabricated using metal dots which are connected by capacitors and tunnel junctions. The operation of a basic cell is confirmed by the externally controlled polarization change of the cell. The cell exhibits a bistable response, with more than 80% polarization of the charge within a cell.

1. Introduction

For more than 30 years the microelectronics industry has enjoyed dramatic improvements in the speed and size of electronic devices, and continued growth in the industry requires a further increase in the number of devices fabricated on a chip. This trend has long obeyed Moore's law, which predicts that the number of devices integrated on a chip will double every 18–24 months. Adherence to this exponential growth curve has been a monumental task requiring rapid improvements in all aspects of integrated circuit fabrication, permitting manufacturers both to shrink the size of devices and to increase chip size while maintaining acceptable yields. Since the early 1970s the device of choice for high levels of integration has been the field effect transistor (FET), and while the FET of today is a vast improvement over that of 1970, it is still a current switch like the mechanical relays first used to encode binary information. At gate lengths below 0.1 μm, FETs will begin to encounter fundamental effects which make further scaling difficult. A possible method for the microelectronics industry to maintain continued growth in device density is to change from the FET-based paradigm to one based on nanostructures. Here, instead of fighting the effects that come with feature size reduction, these effects are used to advantage. One nanostructure paradigm, proposed by Lent *et al* [1,2], is quantum-dot cellular automata (QCA), which employs arrays of coupled quantum dots to implement Boolean logic functions [3,4]. The advantage of QCA lies in the extremely high packing densities possible due to the small size of the dots, the simplified interconnection and the extremely low power–delay product which can be arbitrarily reduced by adiabatic switching [5]. Using QCA cells with dots of 20 nm diameter, an entire full adder can be placed within 1 μm^2, approximately the area of a single 0.07 μm gate length FET.

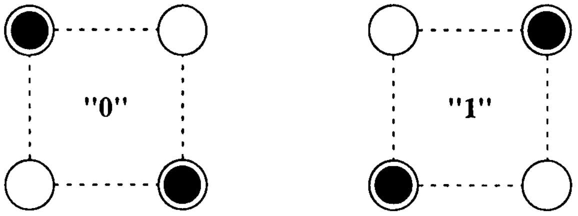


Figure 1. Basic four-dot QCA cell showing the two possible polarizations.

A basic QCA cell consists of four quantum dots located at the corners of a square, coupled by tunnel barriers. If the cell is biased so that there are two excess electrons among the four dots, Coulomb repulsion will force the electrons to opposite corners. There are thus two energetically equivalent polarizations, as shown in figure 1. These two polarizations can be labelled logic '0' and '1', and, by properly arranging cells so that the polarization of one cell sets the polarization of a nearby cell, it is possible to implement all combinational logic functions. A tremendous advantage of QCA devices is the simplified interconnect which is possible with this paradigm. Since the cells communicate only with their nearest neighbours, there is no need for long interconnect lines. The inputs are applied to the cells at the edge of the system and the computation proceeds until the output appears at cells at the edge of the QCA array.

2. Experiment

The experimental work presented is based on a QCA cell using aluminium islands with aluminium oxide tunnel junctions, fabricated on an oxidized silicon wafer. The fabrication used standard electron beam lithography and shadow evaporation to form the islands and tunnel

0268-1242/98/SA0130+05$19.50

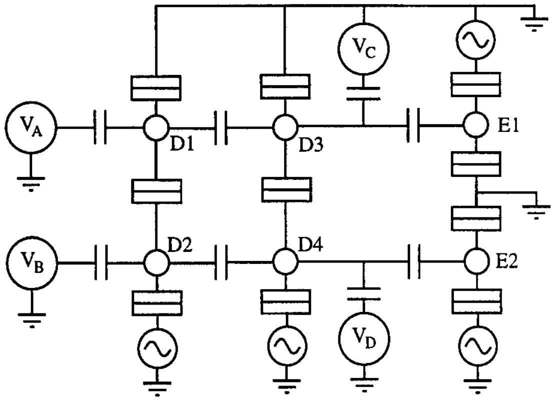


Figure 2. Schematic diagram of the four-dot QCA cell with two electrometers.

junctions [6]. The area of the tunnel junctions is an important quantity since this dominates island capacitance, determining the charging energy of the island and hence the operating temperature of the device. For our devices the area is approximately 50 nm by 50 nm, giving a junction capacitance of 240 aF.

We recently demonstrated the first step in the development of QCA systems, i.e. a functional QCA cell where we can switch the polarization of a cell, confirming that the switching of a single electron between coupled quantum dots can control the position of a single electron in a second set of dots [7]. A schematic diagram of our four-dot QCA cell with additional electrometers is shown in figure 2. The four dots of the QCA cell consist of two double dots, where the dots are joined by a tunnel junction. This breaks the QCA cell into two half-cells, where electrons are allowed to tunnel 'vertically' between dots but not 'horizontally'. The input voltages V_A and V_B are connected to double dot D1 and D2, which form the input half-cell. These are connected capacitively to the output half-cell D3 and D4, which are in turn capacitively coupled to the electrometers E1 and E2. The circuit was mounted on the cold finger of a dilution refrigerator whose base temperature is 10 mK. The conductance through each double-dot half-cell and each individual electrometer can be measured simultaneously using standard ac lock-in techniques. An excitation voltage of 4 μV was typically used, with a frequency between 16 and 40 Hz. A magnetic field of 1 T was applied to suppress the superconductivity of the aluminium metal. The capacitances between the gates and islands were extracted from the period of the Coulomb blockade oscillations [8].

The operation of a QCA cell is best understood by examining the conductance through the output half-cell as a function of the two gate voltages V_C and V_D, as shown in the contour plot of figure 3. A peak in the conductance is observed each time the Coulomb blockade is lifted for the double-dot system, and because of the capacitive coupling between the dots each peak splits into a double peak. These peaks form the vertices of a hexagonal structure which we refer to as the 'honeycomb', delineated by the broken lines in figure 3 [9]. The electron population of the dots is stable within each hexagon of the honeycomb and changes when a border between cells is crossed. The excess electron population within each hexagon can thus be labelled, with the (0, 0) hexagon centred at $V_C = V_D = 0$ V. A point in the honeycomb defined by a single setting of V_C and V_D is called the working point, which defines a particular configuration of electrons. If V_C is swept in the positive direction with V_D fixed, electrons are added one by one to the top dot as the working point moves horizontally through the hexagons (1, 0) then (2, 0) and so on. If V_D is

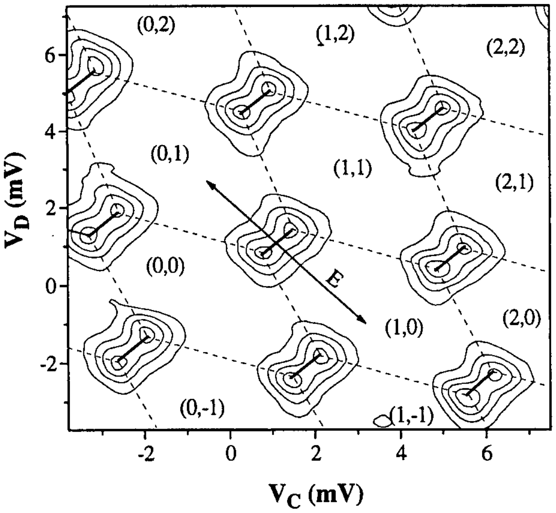


Figure 3. Contour plot of the measured conductance through the output half-cell as a function of V_C and V_D. The excess electron population is noted as (n_3, n_4).

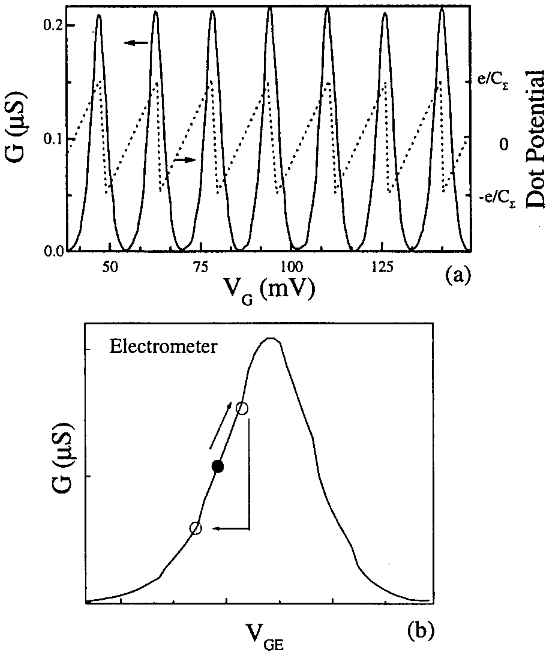


Figure 4. (*a*) The measured conductance and calculated potential of a single dot as a function of its gate bias. (*b*) A single conductance peak of the electrometer dot, as a function of its gate bias. The nominal working point is set to the left side of the peak, and the potential variations on the coupled dot cause excursions away from this point, causing large changes in the electrometer conductance.

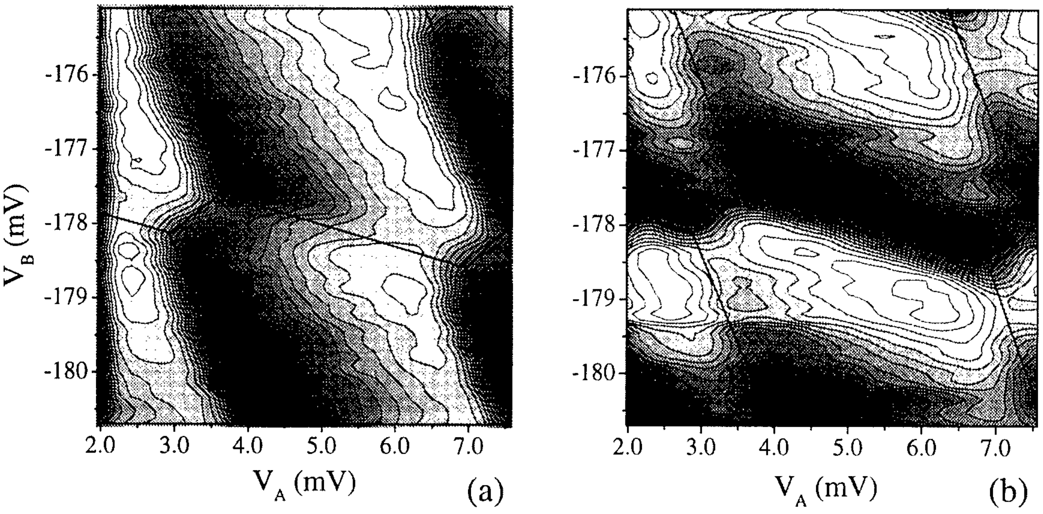


Figure 5. (*a*) Contour plot of the conductance through electrometer E1 as a function of V_C and V_D. Light regions represent areas of high conductance and dark regions those of low conductance. (*b*) Contour plot of the conductance through electrometer E2.

swept positive, electrons are added to the bottom dot as the working point moves vertically through the hexagons (0, 1) and (0, 2). Most important for QCA operation is motion of the working point in the direction shown by arrow E in figure 3. This movement between the (0, 1) and (1, 0) hexagon represents the switching of an electron between the bottom dot and top dot. The goal of our experiment is to demonstrate QCA operation by using electrodes to force this transition in the input half-cell and then to have the potential changes on the input half-cell force an opposite transition in the output half-cell.

Since the operation of a QCA cell depends on the position of a single electron, it is necessary to track the positions of electrons within the cell. One way to do this is to measure the conductance through each half-cell. A peak in the conductance as the gate voltages are changed indicates that the Coulomb blockade has been lifted for both dots simultaneously, and a change in the dot population has occurred. However, as seen in figure 3, not all electron transitions can be detected in this manner. If V_C alone is swept, electrons are added to only the top dot, and no change in the conductance through the dots is seen. To characterize fully the QCA cell it is therefore necessary to detect externally the charge state of each dot individually. This is done using additional dots as electrometers [10], capacitively coupled to the output half-cell, as shown in the schematic of figure 2. The electrometer operates by detecting small potential changes in the dot being measured. Figure 4(*a*) plots the measured dot conductance and theoretical potential as a function of an applied gate voltage. The potential on a dot exhibits a sawtooth behaviour, increasing with positive gate voltage, followed by an abrupt reset when an additional electron is added to the dot. For the dots used in our experiments, the amplitude of this sawtooth oscillation is approximately 110 μV. The operation of the electrometer is shown in figure 4(*b*), which expands a single conductance peak. The gate voltage of the electrometer is adjusted so that the working point lies on the side of one of the conductance peaks and is then held constant through the rest of the experiment. Any potential variations coupled to the electrometer, such as those on a nearby dot, will act as an additional gate voltage to the electrometer dot which causes a shift of the working point shown in figure 4(*b*). Since the conductance peaks are quite sharp, a small change in the gate voltage gives a large change in the conductance through the electrometer dot. Knowing the capacitance coupling the electrometer to the dot and the shape of the electrometer conductance peak, it is possible to calculate the potential on the measured dot. In these measurements it is important to know which side of the conductance peak is chosen for the electrometer working point since the sign of the transconductance changes from side to side. In our experiments the electrometers are always biased on the left side of the peak, giving a positive transconductance.

Figure 5, which shows a contour plot of the conductance through each of the electrometers, demonstrates that the electrometers can externally detect the movement of an electron in a double-dot half-cell. Figure 5(*a*) shows the conductance of the top electrometer as a function of V_C and V_D. The overlying lines show the honeycomb borders as determined by a simultaneous measurement of the conductance through the double dot. It is clearly seen that the largest changes in the electrometer conductance occur at the honeycomb borders where electrons are added to D3. Only a small change is seen when electrons are added to D4, because of the unavoidable, but small, parasitic capacitance coupling the upper electrometer to D4. Similarly, the conductance through the lower electrometer shows a large change when electrons are added to D4, as shown in figure 5(*b*). In each case the conductance through the electrometer is proportional to the potential on the corresponding dot. In our experiments we typically use both the conductance through each half-cell and the electrometer signals to characterize completely the operation of the QCA cell.

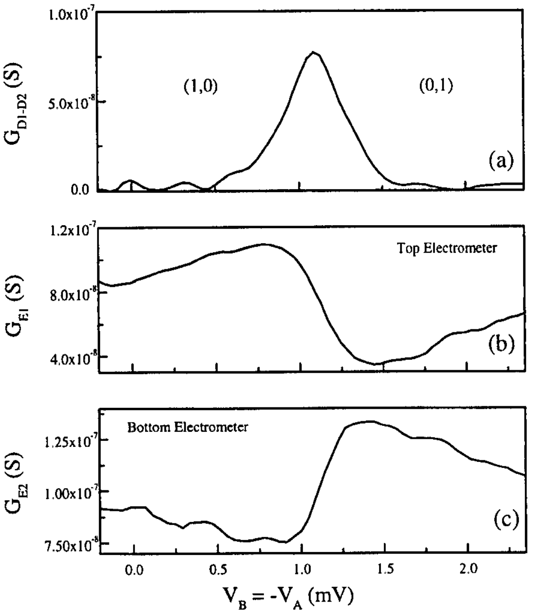


Figure 6. (*a*) Conductance through the input half-cell, where the peak indicates the switch of an electron from D1 to D2. (*b*) Conductance through the electrometer E1 indicating the addition of an electron to D3. (*c*) Conductance through electrometer E2 indicating the removal of an electron from D4.

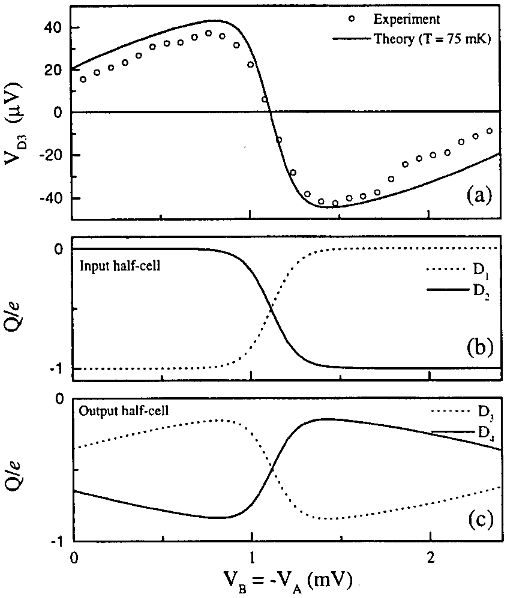


Figure 7. (*a*) Measured potential on dot D3 as a function of the differential input $V_B = -V_A$ together with theory at 75 mK. Calculated excess electron populations for (*b*) the input half-cell D1 and D2 and (*c*) the output half-cell D3 and D4 as a function of the differential input are also shown.

QCA operation is demonstrated by applying a differential voltage to the input half-cell, a positive bias to V_B and a negative bias to V_A. As this differential voltage is swept, electrons are moved one by one from D1 to D2, and the electrostatic potential on D1 and D2 changes in response to the applied gate voltages and the position of electrons. The potential on D2 increases with the positive voltage V_B, until an abrupt reset which occurs when an electron enters the dot, corresponding to the sawtooth pattern seen in figure 4(*a*). Similarly, the potential on dot D1 is also a sawtooth as a function of V_A, but with the opposite phase. The amplitude of each of these sawtooth oscillations is calculated to be 100 μV. It is this voltage which must force an electron in the output half-cell to move from D4 to D3. Since the potentials on D1 and D2 act as additional gate voltages for D3 and D4, the honeycomb of the output half-cell will shift in response to potential changes in the input half-cell. The working point of the output half-cell (V_C, V_D) is set close to the border separating the (0, 1) and (1, 0) hexagons. For QCA operation we must have a shift that is sufficient to move the honeycomb border from one side of this working point to the other, representing a switch of an electron in the output half-cell. This switching in the output half-cell will be detected by the two electrometers, where the conductance of one electrometer will increase as an electron leaves its adjacent dot, while the conductance of the other electrometer will decrease as an electron enters its adjacent dot. The experimental measurements confirm this behaviour, as shown in figure 6, which plots the conductance through the input half-cell, together with the conductance through each electrometer as a function of the input voltages $V_B = -V_A$. The peak in the conductance through the input half-cell, seen in figure 6(*a*) as V_B increases, indicates that an electron has moved from D1 to D2. As the electron switches in the input half-cell the electrometer conductances shift in opposite directions, a shift down for the top electrometer of figure 6(*b*) and a shift up for the bottom electrometer of figure 6(*c*). This indicates that an electron has moved from D4 to D3 as expected because of the electron switch in the input half-cell. This confirms the polarization change required for QCA operation.

Using the electrometer signal of figure 6(*b*) we can calculate the potential on D3 as a function of the input differential voltage. This is plotted in figure 7(*a*) together with the theoretically calculated potential at a temperature of 75 mK. Although at a temperature of 0 K the potential changes are abrupt, the observed potential shows the effects of thermal smearing, and theory at 75 mK shows good agreement with experiment. The heating of the electron system to temperatures above that of the dilution refrigerator is probably due to the applied excitation voltage and noise voltages coupled into the sample by the leads. This effect is commonly seen in measurements of this type [11]. Figures 7(*b*) and 7(*c*) plot the theoretical excess charge on each of the dots in the input and output half-cells at 75 mK. This shows an 80% polarization switch of the QCA cell, and the polarization change can be further improved with an increase in the capacitances coupling input and output half-cells.

3. Summary

A device paradigm based on QCA cells offers the opportunity to break away from FET-based logic and to exploit the quantum effects that come with small size. In this new paradigm, logic levels are no longer encoded as voltages but as the position of electrons within a quantum dot cell. QCA cells are scalable to molecular dimensions, and the performance improves as the size shrinks. A QCA cell with molecular dimensions should operate at room temperature since the energy spacings of the dot states will be larger than kT, even at 300 K. Using aluminium island dots, with aluminium oxide tunnel junctions, we have demonstrated the operation of a QCA cell. The cell exhibits a bistable distribution of electrons, and the polarization of the cell can be switched by externally applied bias voltages.

Acknowledgments

This work was supported in part by the Defense Advanced Projects Agency, Office of Naval Research (Contract N00014-95-1-1166) and the National Science Foundation.

References

[1] Lent C S, Tougaw P D, Porod W and Bernstein G H 1993 *Nanotechnology* **4** 49

[2] Lent C S and Tougaw P D 1997 *Proc. IEEE* **85** 541

[3] Lent C S and Tougaw P D 1993 *J. Appl. Phys.* **74** 6227

[4] Tougaw P D and Lent C S 1994 *J. Appl. Phys.* **75** 1818

[5] Tougaw P D and Lent C S 1996 *J. Appl. Phys.* **80** 4722

[6] Fulton T A and Dolan G H 1987 *Phys. Rev. Lett.* **59** 109

[7] Orlov A O, Amlani I, Bernstein G H, Lent C S and Snider G L 1997 *Science* **277** 928

[8] Likharev K K 1988 *IBM J. Res. Dev.* **32** 144

[9] Pothier H, Lafarge P, Orfila P F, Urbina C, Esteve D and Devoret M H 1991 *Physica* B **169** 573

[10] Amlani I, Orlov A O, Snider G L, Lent C S and Bernstein G H 1997 *Appl. Phys. Lett.* **71** 1730

[11] Lafarge P, Pothier H, Williams E R, Esteve D, Urbina C and Devoret M H 1991 *Z. Phys.* B **85** 327

Semicond. Sci. Technol. **13** (1998) A135–A139. Printed in the UK PII: S0268-1242(98)91885-5

On the validity of quantum hydrodynamics for describing antidot array devices

J R Barker† and D K Ferry‡

† Department of Electronics and Electrical Engineering, University of Glasgow, Glasgow G12 8LT, UK
‡ Center for Solid State Electronics Research, Arizona State University, Tempe, AZ 85287-6206, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. Quantum hydrodynamic models are becoming increasingly used for modelling both conventional and novel semiconductor devices. However, there are several controversial problems which are unresolved including the form of the quantum potential in different formulations. It is shown here that present quantum hydrodynamics is not consistent with mixed quantum states in which spatial fluctuations are significant and that where transient bound state formation is possible it becomes important to add vorticity terms to account for multiply connected regions. These problems are of importance to transport studies of dense antidot systems.

1. Introduction

The present trend in serious semiconductor device modelling for both silicon MOSFET and III–V heterostructures is towards full-band, ensemble Monte Carlo simulation coupled to detailed 2D and 3D device models based on finite element models [1]. The model is essentially classical and is built on the Boltzmann–Bloch high-field transport equation (BTE) for the carrier distribution function in phase space. This approach becomes more important than traditional hydrodynamic (HD) and drift–diffusion (DD) models [2] as one considers hot-electron and transient effects arising from deep submicron structures and/or complex material geometries such as Si–Ge heterostructures or III–V heterostructures. At the same time many mesoscopic devices have been proposed which rely on strictly quantum effects ranging from tunnelling to interference and diffraction processes associated with electron-waveguide-like structures [3]. Quantum devices have only recently been considered from a full device modelling standpoint, particularly resonant tunnelling devices [4]. The underpinning theoretical framework for quantum functional devices is less clear than for classical devices, but there is now a strong interest in using quantum versions of the HD approach to yield rapid turn-round simulations. However, even for classical devices there are significant controversies concerning the correctness of the HD formulation [5].

A number of groups [6] have reported device simulations based on quantum hydrodynamic (QHD) models which have the advantage of fitting cleanly into the conventional device modelling framework with the promise of building in quantum transport in a very simple fashion. At a later stage one might imagine progressing to the underlying quantum kinetic picture, corresponding to the Boltzmann model in classical transport, and it would be useful if that regime were amenable to Monte Carlo simulation. In the present paper we examine whether or not this scheme is really feasible and in particular whether present QHD models really describe quantum effects correctly. Our objective is to find suitable HD models that will describe transport through dense antidot arrays imposed on 2DEG systems. These structures have interesting quantum transport properties as evidenced both experimentally and by recursive Green function techniques [6]. It would be advantageous to make more predictive studies by solving suitable HD equations to cut down the very large computation time for more exact methods particularly in self-consistent fields.

2. Classical hydrodynamic models

The classical HD model for semiconductor transport (hot-electron mobility and diffusion) was introduced by Stratton [7] within the relaxation time approximation. The HD equations for a two-valley semiconductor were derived by Blotekjaer [8] by taking the first three moments of the Boltzmann transport equation again using the relaxation time approximation. Some 10 years later the formalism was used for modelling GaAs MESFETs [10]. Since then there have been many modifications of the HD model, notably to deal with transient and non-local effects found in Monte Carlo simulation. More recently there have been attempts to deal more satisfactorily with the collision moments [11]. For our purposes let us focus on the derivation of the classical HD model from the Boltzmann transport equation. The latter may be taken without loss of generality to have the simple form

$$(\partial/\partial t + \boldsymbol{v}\cdot\nabla + \boldsymbol{F}\cdot\nabla_p) f(r, p, t) = (\partial f/\partial t)_{coll} \qquad (1)$$

where f is the distribution function, $\boldsymbol{F} = e\boldsymbol{E}$, and $\boldsymbol{p} = h\boldsymbol{k}$. Multiplying equation (1) by the functions 1, $\boldsymbol{p}$, $\varepsilon(\boldsymbol{p})$ and $\boldsymbol{p}\varepsilon$, and integrating over momentum space, we obtain the coupled macroscopic conservation equations:

$$\partial n/\partial t + \nabla\cdot\{n\boldsymbol{V}\} = nC_n \qquad (2)$$

$$\partial(n\boldsymbol{P})/\partial t + \nabla\cdot\{n\boldsymbol{U}\} - n\boldsymbol{F} = nC_p \qquad (3)$$

$$\partial(nW)/\partial t + \nabla\cdot\{n\boldsymbol{S}\} - n\boldsymbol{V}\cdot\boldsymbol{F} = nC_\varepsilon \qquad (4)$$

$$\partial(n\boldsymbol{Q})/\partial t + \nabla\cdot\{n\boldsymbol{R}\} - n(W\boldsymbol{1} + \boldsymbol{U})\cdot\boldsymbol{F} = nC_{\varepsilon p}. \qquad (5)$$

Here the physical averages are given by $\boldsymbol{V} = \langle\boldsymbol{v}\rangle$ is the average carrier velocity, $\boldsymbol{P} = \langle\boldsymbol{p}\rangle$, $\boldsymbol{U} = \langle\boldsymbol{vp}\rangle$ is the momentum flux tensor, $W = \langle\varepsilon\rangle$ is the average carrier energy, $\boldsymbol{S} = \langle\varepsilon\boldsymbol{v}\rangle$ is the average energy flux, $\boldsymbol{Q} = \langle\boldsymbol{p}\varepsilon\rangle$ and $\boldsymbol{R} = \langle\boldsymbol{vp}\varepsilon\rangle$ is a fourth-order moment. n is the carrier concentration. Finally, the terms labelled C on the right-hand sides of equations (2)–(5) are the first four moments of the collision integral. The equations (2)–(5) form a set of eight coupled equations for rather more unknowns: n, $\boldsymbol{V}$, $\boldsymbol{P}$, $\boldsymbol{U}$, W, $\boldsymbol{S}$, $\boldsymbol{Q}$, $\boldsymbol{R}$. In the present paper we shall ignore the collisional moments which are themselves a source of controversy and concentrate on the dynamical parameters.

The number of unknowns is reduced by setting an additional group of closure relations which relate some of the unknowns. The three non-controversial closure relations relate average velocity to average momentum, the momentum flux tensor to the velocity flux tensor and the electron momentum temperature tensor $\boldsymbol{T}_p$, and the average total energy to the average velocity and the electron momentum temperature tensor. For a simple parabolic band, if we define

$$k_B\boldsymbol{T}_p = m\langle(\boldsymbol{v} - \boldsymbol{V})(\boldsymbol{v} - \boldsymbol{V})\rangle \qquad (6)$$

the first set of closure relations are

$$\boldsymbol{P} = m\boldsymbol{V} \qquad (7)$$

$$\boldsymbol{U} = m\boldsymbol{VV} + k_B\boldsymbol{T}_p \qquad (8)$$

$$W = \tfrac{1}{2}mV^2 + \tfrac{1}{2}\,\mathrm{Trace}(k_B\boldsymbol{T}_p). \qquad (9)$$

Here m is the conductivity effective mass.

The remaining closure relations are controversial. For example, if equipartition is assumed, it follows that we may define an electron temperature T such that

$$k_B\boldsymbol{T}_p = k_BT\boldsymbol{1}. \qquad (10)$$

From equation (10) we may derive a closure relation between $\boldsymbol{U}$ and $\boldsymbol{V}$ and W:

$$\boldsymbol{U} = \boldsymbol{VP} + [2(W - \tfrac{1}{2}mV^2)/3]\boldsymbol{1}. \qquad (11)$$

Equipartition is not generally valid for non-equilibrium transport and worse still a closure relation for $\boldsymbol{U}$, W and $\boldsymbol{V}$ is inaccessible for general non-parabolic bands. In practice, it is commonly assumed that a closure relation of the following form exists:

$$\boldsymbol{U} = \boldsymbol{VP} + u(W)\boldsymbol{1} \qquad (12)$$

where $u(W)$ is a non-linear function of W.

3. Quantum hydrodynamic models

QHD models have been derived within various approximations by similar techniques to the classical route but starting from the Bloch equation or the equation of motion for the Wigner function [12]. Collisional moments are dealt with phenomenologically. Many device simulators and indeed commercial simulator packages have been produced for studying quantum devices such as resonant tunnelling diodes and for providing some feel for the softening effects of sharp boundaries on carrier flow. At first sight there should be no difficulty in describing quantum ballistic transport in a self-consistent smoothly varying electric field with simple boundaries. Indeed quite simple modifications appear in current QHD models.

The major change may be seen for example in the momentum conservation equation corresponding to equation (3) where quantum effects are manifest through the appearance of a quantum force $\boldsymbol{F}_Q$ derived from the Bohm quantum potential ϕ_Q, which is determined by the carrier concentration:

$$\phi Q = -(\hbar^2/2mn^{1/2})\nabla^2 n^{1/2}. \qquad (13)$$

The most common form of QHD, in one dimension for simplicity, gives in place of equations (2) and (3)

$$\partial n/\partial t + (\partial/\partial x)\{nV\} = 0 \qquad (14)$$

$$\partial(nP)/\partial t + (\partial/\partial x)\{nU\} + (n\,\partial/\partial x)(\phi + \phi_Q/3) = 0 \qquad (15)$$

where ϕ is the classical potential energy and

$$U = mV^2 + k_BT. \qquad (16)$$

Similarly the energy conservation equation (4) is modified from classical HD: the Bohm potential is added to the classical potential energy and a further quantum term appears, related to the so-called Wigner potential ϕ_W which satisfies the identity

$$2\nabla\{n\phi_W\} = n\nabla\phi_Q. \qquad (17)$$

4. The single-carrier quantum hydrodynamic model

In derivations of QHD based on the moment method, the quantum potential terms arise through the closure relations and are subject therefore to similar doubts to those in the classical case. However, more serious problems are hidden in QHD as we see if we consider a single carrier described by a pure-state wavefunction $\psi = R\exp(\mathrm{i}S/\hbar)$ where we identify the carrier concentration as $n = R^2$, and $S/\hbar$ is the phase. Choosing a simple Hamiltonian $H = p^2/2m + \phi$ we may insert the polar form for the wavefunction into the Schrödinger equation and separating real and imaginary parts we obtain

$$\partial n/\partial t + \nabla \cdot \{nV\} = 0 \qquad (18)$$

where we have used the exact relation $mV = \langle p \rangle_Q = \langle \psi | p | \psi \rangle$ evaluated in the position representation as $P = \langle p \rangle_Q = \nabla S$ and where $\langle \ldots \rangle_Q$ is the quantum average. We also obtain the quantum Hamilton–Jacobi equation:

$$-\partial S/\partial t = (\nabla S)^2/2m + \phi + \phi_Q. \qquad (19)$$

Equation (18) is just the continuity equation and we recover the equivalent of equation (3) or (15) by taking the gradient of equation (19):

$$\partial \boldsymbol{P}/\partial t + \nabla(\phi + \phi_Q) = \boldsymbol{0}. \qquad (20)$$

Finally, using the continuity equation we find equation (3) again:

$$\partial(n\boldsymbol{P})/\partial t + \nabla \cdot \{n\boldsymbol{U}\} + n\,\nabla\phi = \boldsymbol{0} \qquad (21)$$

where the momentum current density (stress tensor) contains quantum contributions

$$\boldsymbol{U} = m\boldsymbol{V}\boldsymbol{V} + m\boldsymbol{V}'\boldsymbol{V}' - (h^2/mn)\,\nabla^2 n\,\boldsymbol{1}. \qquad (22)$$

Here the pseudo-velocity $\boldsymbol{V}'$ and pseudo-momentum $\boldsymbol{P}'$ are defined by

$$\boldsymbol{P}' = m\boldsymbol{V}' = (h/2n)\,\nabla n. \qquad (23)$$

They are related to the quantum potential by [13]

$$\phi_Q = \tfrac{1}{2}mV'^2 - (\hbar^2/2mn)\,\nabla^2 n. \qquad (24)$$

To compare with the result (15) in section 3 we project into one dimension to find

$$\partial(nP)/\partial t + (\partial/\partial x)\{nU_0\} + (n\partial/\partial x)(\phi + \phi_Q) = 0 \quad (25)$$

where

$$U_0 = mV^2. \qquad (26)$$

These results are exact. The similarity to the QHD result (15) is striking. The non-appearance of temperature is not surprising. What is surprising is that the quantum potential appears with a coefficient of unity and not the 1/3 factor seen in equation (15). This problem has been recognized before and Ancona and Iafrate [15] for example have suggested that the prefactor to the quantum potential should be regarded as an adjustable parameter. Grubin *et al* [16] have examined the effect of different prefactors on device simulations. A detailed discussion of the different formulations and comparison with direct quantum calculations is given by Grubin [4].

5. Interpretation

We may shed some light on the results of section 4 by noting that equations (21) and (22) follow directly from the conservation equation

$$\partial(n\boldsymbol{P})/\partial t + \nabla \cdot \{n\boldsymbol{U}\} - n\boldsymbol{F} = \boldsymbol{0}$$

with $\boldsymbol{P}$ and $\boldsymbol{U}$ now defined as originally by equations (7) and (8) but with the quantum average $\langle\rangle_Q$:

$$\boldsymbol{U} = \langle \boldsymbol{pp}/m \rangle_Q = m\boldsymbol{V}\boldsymbol{V} + m\langle(\boldsymbol{v} - \boldsymbol{V})(\boldsymbol{v} - \boldsymbol{V})\rangle_Q. \quad (27)$$

This equation follows directly from moment decomposition of the Wigner equation or the density matrix equations and has an identical form to the classical result (3) except that we use the quantum expectation value.

As noted elsewhere [13] it is clear that the quantum potential is actually a dynamical quantity describing fluctuations in the electron velocity through the velocity autocorrelation function: for a pure state we find by direct construction using the Schrödinger equation

$$k_B\boldsymbol{T}_p = m\langle(\boldsymbol{v} - \boldsymbol{V})(\boldsymbol{v} - \boldsymbol{V})\rangle_Q = m\langle \boldsymbol{vv} \rangle_Q - \langle \boldsymbol{V} \rangle_Q \langle \boldsymbol{V} \rangle_Q. \qquad (28)$$

Equation (28) provides a closure equation which is exact in this simple pure state case. If the velocity is regarded as a stochastic dynamical variable equation (28) may be interpreted as the covariance matrix of the velocity suggesting fluctuations with an effective internal 'temperature' $T_Q \sim \phi_Q/kB$. This picture is consistent with the proposed stochastic extension to Bohm's deterministic picture of quantum mechanics [13, 14]. Following this line we may use the average energy expression with quantum averages to obtain

$$\begin{aligned} W &= \tfrac{1}{2}mV^2 + \tfrac{1}{2}\,\mathrm{Trace}(k_B\boldsymbol{T}_p) \\ &= \tfrac{1}{2}(mV^2 + \phi_Q) \qquad \text{1D result} \\ &= \tfrac{1}{2}(mV^2 + 3\phi_Q) \qquad \text{3D result.} \end{aligned} \qquad (29)$$

The prefactor to the quantum potential is here seen to be a property of the dimensionality of the system. The average energy comprises kinetic energy associated with the mean motion plus 'thermal' energy derived from the velocity fluctuations. An interpretation of the quantum potential in terms of inherent fluctuations and its reconciliation with the Bohm picture are described in [13].

6. The mixed-state quantum hydrodynamic model

The simple picture developed exactly in section 5 suggests that by using its extension to mixed states we should easily recover one of the proposed QHD models. This is not the case. To see this let us define a pure-state density matrix ρ_j for each the pure state j. Assigning a probability P_j for state j we may define the mixed-state density matrix, the average particle density n and average velocity $\boldsymbol{V}$ via

$$\rho = \sum_j P_j \rho_j \qquad \boldsymbol{V}_j = \nabla S_j/m$$

$$\phi_{Qj} = -(h^2/2mn_j^{1/2})\,\nabla^2 n_j^{1/2} \qquad \sum_j P_j = 1 \qquad (30)$$

$$nV = \sum P_j n_j V_j \qquad n = \sum P_j n_j. \tag{31}$$

The equation of motion of $P = mV$ follows as

$$\partial(nP)/\partial t - \sum P_j n_j \nabla\cdot\{n_j V_j\} + n\nabla(\phi + \phi_{QQ}) = \mathbf{0} \tag{32}$$

where ϕ_{QQ} is an effective quantum potential:

$$\phi_{QQ} = \sum P_j n_j \phi_j / n. \tag{33}$$

The effective quantum potential is no longer a simple function of the average density and depends on the states and state distributions in a complex manner.

7. Fluctuation analysis

We can elucidate some of the physics by the transformation to fluctuation coordinates:

$$n_j = n + \Delta n_j \qquad \langle\ldots\rangle = \sum P_j \ldots \qquad n = \langle n_j \rangle. \tag{34}$$

The effective quantum potential becomes to lowest orders in Δn_j

$$\phi_{QQ} \sim \phi_Q(1 - \alpha\langle\Delta n_j\,\Delta n_j\rangle) \tag{35}$$

$$\phi_Q = -(\hbar^2/2mn^{1/2})\,\nabla^2 n^{1/2}. \tag{36}$$

The second term in equation (35) represents a significant contribution to the overall density–density correlation function, which is known to produce weak localization and universal conductance fluctuations in mesoscopic devices. For small fluctuations, we recover the familiar form for the quantum potential. However, any significant spatial variation of the substates j will lead to spurious results if the simplified QHD equations are used. This will particularly apply to confined systems where size quantization is important. For resonant tunnelling devices with a single resonance, that state dominates and the approximation may work quite well. This is not likely to be the case for multiple resonances in multistate devices. These remarks will also apply to flows through 2DEG structures which have any residual potential field varying on the scale of the carrier wavelength. Thus dense antidot potentials [17] are not likely to be modellable by simplistic versions of QHD.

8. Is quantum hydrodynamics complete?

There is a further challenge that faces QHD models which concerns the requirement for a pure state that the action S must satisfy $S = S \bmod h$. Since $V = \nabla S/m$, this non-local boundary condition on S leads to the quantization condition (N an integer)

$$\Gamma = \int V \cdot \mathrm{d}r = Nh/m \tag{37}$$

which in HD terms corresponds to the quantization of velocity circulation which is related to vorticity via Stoke's theorem. Γ vanishes for irrotational flow (Kelvin's theorem) which makes the vorticity constant, except at boundaries, or in shocks or scattering processes. All three of these conditions may occur in ultrasmall devices.

Kelvin's theorem tells us that generally, in singly connected systems, the effects of size quantization will not evolve from a classical distribution, but must have propagated from initial conditions. Moreover, systems which exhibit quantization cannot be treated as irrotational hydrodynamic flows. The irrotational regions correspond to nodal lines in the wavefunction field. Closed orbital flows can be set up between reflecting barriers as in resonant tunnelling, but also by multiple scattering from point defects or along surfaces and interfaces (all potential sources of weak localization). Whereas condition (37) can be met by appropriate choice of boundary state in non-dissipative hydrodynamics, there are severe problems for mixed states where the corresponding condition leads to

$$\Gamma = \int V \cdot \mathrm{d}r = \int \mathrm{d}r \cdot \sum P_j n_j V_j / n \tag{38}$$

subject to a microscopic boundary condition of the form (37) for each component state in the mixture:

$$\Gamma_j = \int V_j \cdot \mathrm{d}r = N_j h/m. \tag{39}$$

This renders closure of the QHD moment hierarchy extremely difficult. One way in which Γ could vanish is if phase mixing to zero occurs within the average over individual circulations which would restore relative simplicity to the QHD equations. This remains a topic for future exploration.

9. Conclusions

Confidence in QHD models largely derives from the sensible conclusions which are reached when a solution is known to the Schrödinger equation and the result used in the subsequently derived HD models. The reverse procedure is, however, seriously flawed: solutions to the HD equations under conventional boundary conditions are not sufficient to reconstruct the true quantum solutions unless certain non-local conditions (37) and (38) are imposed which severely alter the nature of the solutions and indeed the simple kinetic or HD interpretation. Multiply connected device geometries such antidot array devices provide just such a scenario. In HD models, quantum effects are induced into the solution via restrictions on the initial-condition boundary relations, the dynamical evolution of the system and non-local boundary conditions which act to preserve the continuity and single valuedness of the underlying wavefunctions.

References

[1] Asenov A, Brown A, Roy S, Arokianathan C, Davies J H and Barker J R 1997 *Proc. 2nd NASA Device Modelling Workshop* ed S Saini (NASA) pp 86–99

[2] Snowden C (ed) 1989 *Semiconductor Device Modelling* (London: Springer)

[3] Ferry D K and Barker J R *VLSI Des.* to be published

[4] Grubin H L 1995 *Quantum Transport in Ultra-Small Devices* ed D K Ferry, H L Grubin, C Jacoboni and A P Jauho (New York: Plenum) pp 241–80

[5] Tang T and Ieong M 1997 *Proc. 2nd NASA Device Modelling Workshop* ed S Saini (NASA) pp 127–36

[6] Ferry D K, Akis R, Vasileska D, Bird J P and Barker J R 1997 *Proc. 2nd NASA Device Modelling Workshop* ed S Saini (NASA) pp 20–31

[7] Stratton R 1962 *Phys. Rev.* B **126** 2002–14

[8] Blotekjaer K 1970 *IEEE Trans. Electron Devices* **17** 38–47

[9] Curtice W R and Yun Y H 1981 *IEEE Trans. Electron Devices* **28** 954–62

[10] Hansch W and Miura-Mattausch M 1986 *J. Appl. Phys.* **60** 650–6

[11] Gardener C L 1994 *SIAM J. Appl. Math.* **54** 409–27

[12] Ferry D K and Zhou J-R 1993 *Phys. Rev.* B **48** 7944–50

[13] Barker J R 1994 *Semicond. Sci. Technol.* **9** 911–7

[14] Barker J R 1995 *Quantum Transport in Ultra-Small Devices* ed D K Ferry, H L Grubin, C Jacoboni and A P Jauho (New York: Plenum) pp 171–80

[15] Ancona M A and Iafrate G J 1989 *Phys. Rev.* B **39** 9536

[16] Grubin H L, Kreskovsky J P, Govindan T R and Ferry D K 1994 *Semicond. Sci. Technol.* **9** 855

[17] Zozoulenko I V, Maao F A and Hauge E H 1996 *Phys. Rev.* B **53** 7975

Semicond. Sci. Technol. **13** (1998) A140–A142. Printed in the UK

PII: S0268-1242(98)91892-2

High electron mobility in strained Si channel of $Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x$ heterostructure with abrupt interface

N Sugii, K Nakagawa, Y Kimura, S Yamaguchi and M Miyao

Central Research Laboratory, Hitachi Ltd, Kokubunji-shi, Tokyo 185, Japan

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. $Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x$ heterostructures with high electron mobility were grown by molecular beam epitaxy (MBE) combined with solid-phase epitaxy (SPE). The SPE growth completely suppressed Ge segregation in the $Si_{1-x}Ge_x$ layer on top of the Si channel layer. High-resolution cross-sectional transmission electron microscopy observation revealed that the heterointerfaces formed by the SPE growth technique were very flat compared with those for the MBE-grown samples. The electron mobility for an SPE-grown sample with a graded buffer layer was nearly one order of magnitude higher than without such a layer. A high electron mobility of 8.0×10^5 cm^2 V^{-1} s^{-1} was obtained at 15 K for the sample with a graded buffer layer ($x = 0.2$).

The application of $Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x$ heterostructures in Si devices is highly advantageous because of their higher carrier mobilities [1–8] compared with those in conventional Si devices. However, the electron mobilities reported to date have not come close to the ideal value [9]. To realize ultrahigh electron mobilities, an abrupt interface must be formed on top of the channel through which the confined electrons pass, while also reducing the defect density within the channel.

During a molecular beam epitaxial (MBE) growth of Ge on Si, the Ge atoms are segregated [10]. The segregation lengths are reported to be longer than 10 nm at temperatures from 400 to 600 °C, which are frequently used for MBE growth [11]. Consequently, it is impossible to obtain an abrupt interface through conventional MBE techniques. When an SiGe layer is grown at temperatures below 100 °C, an abrupt interface can be obtained because the segregation length is shorter than 3 nm at such temperatures [11]. However, low-temperature growth causes deterioration of the crystallinity in this layer. These problems must be solved to realize SiGe devices with ultrahigh mobility.

We have developed a growth technique that combines MBE with solid-phase epitaxy (SPE) to solve these problems. This technique enables $Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x$ heterostructures to be fabricated with an atomically flat interface between the channel and the doped region. As a result, a high electron mobility of 8.0×10^5 cm^2 V^{-1} s^{-1} can be obtained at 15 K.

A schematic cross-section of the heterostructure we used in our experiments is shown in figure 1. The p-Si (100) substrate was cleaned by chemical etching followed by thermal cleaning in a UHV chamber. An $Si_{1-x}Ge_x$ ($0.1 \le x \le 0.4$) buffer layer (1.0–2.0 μm thick) was then grown at 600 °C to minimize the strain caused by the lattice mismatch between the buffer layer and the substrate. Two types of buffer layers were used: (1) a single layer with a constant Ge content (x) or (2) a stacked layer combining the constant Ge content (x) layer with a graded Ge content ($0 \le y \le x$) layer below it. Next, an Si channel layer (20 nm thick) was grown at 400 °C. This layer was made as thick as possible, within the critical range for pseudomorphic growth of Si on $Si_{1-x}Ge_x$ ($0.2 \le x \le 0.3$), to eliminate

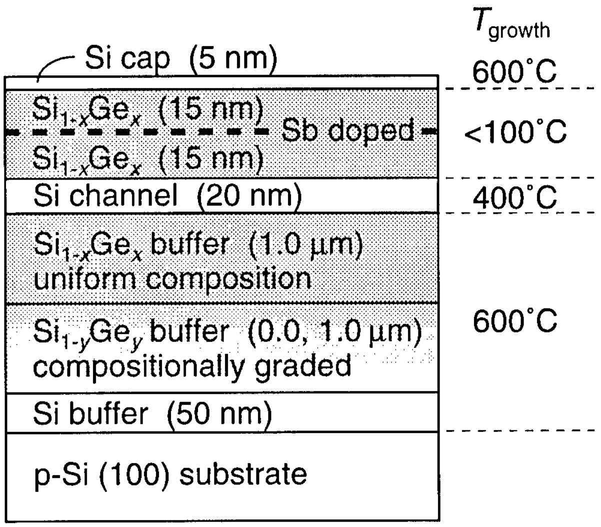


Figure 1. Cross-section of SiGe/Si/SiGe heterostructure.

0268-1242/98/SA0140+03$19.50 © 1998 IOP Publishing Ltd

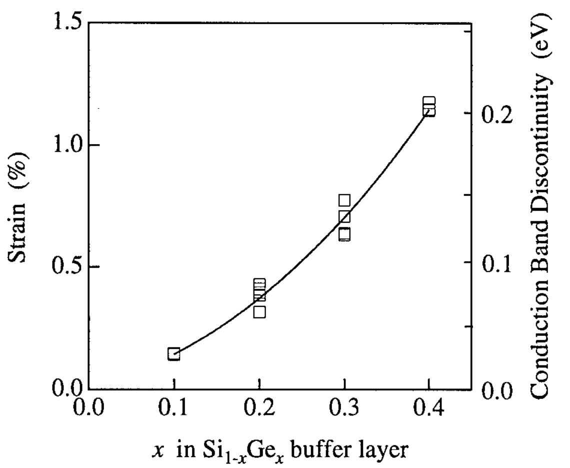


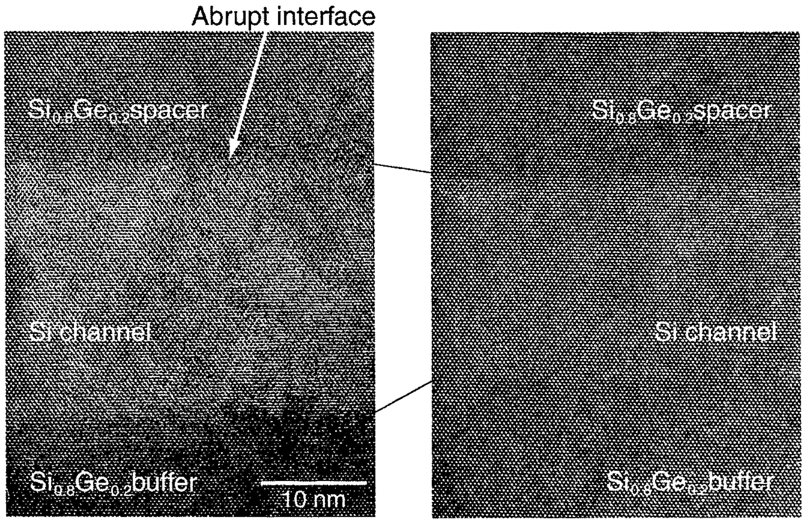
Figure 2. Strain of the Si channel for heterostructures as a function of Ge content of the buffer layers of various thicknesses: ○, uniform 1.0 μm thick buffer layer; □, combination of 1.0 μm thick uniform buffer layer and 1.0 μm thick graded buffer layer. The conduction band discontinuity calculated from the strain is indicated on the right-hand axis.

the surface roughness of the buffer layer. Above the channel layer, an $Si_{1-x}Ge_x$ spacer layer, an Sb delta-doped layer, and an $Si_{1-x}Ge_x$ cap layer were formed by SPE to prevent any surface segregation of Ge and Sb, i.e. each layer was deposited at below 100 °C then crystallized at 600 °C. Finally, an Si cap layer was grown at 600 °C. Other samples having a spacer layer grown by MBE at 450 °C were prepared to compare with the SPE-grown samples. The strain and microstructures in the heterostructures were characterized by Raman spectroscopy and cross-sectional transmission electron microscopy. The transport properties of the channel were evaluated by Hall effect measurement with bar-shaped samples.

The strains in the Si channel layer of the heterostructures are summarized in figure 2 as a function of Ge content (x) of the buffer layer of uniform composition (1 μm thick) or compositionally graded (2 μm thick). The values were evaluated assuming Vegard's law from the Raman shift related to Si–Si vibration. The conduction band discontinuity was calculated from the strain in the channel by using the procedure proposed by van de Walle and Martin [12] and People and Bean [13]. This value is indicated on the right-hand axis in the same figure. The strain relaxation of the samples seems to have been almost complete because there was little difference in the strains for the samples with different buffer layer thicknesses. For the samples with Ge content more than 0.2, the conduction band discontinuity was more than 0.08 eV. This is sufficient for electron confinement in a channel at room temperature.

High-resolution cross-sectional transmission electron microscopy images at the channel region of heterostructures formed with and without an SPE growth technique are shown in figure 3. The heterointerface on top of the channel for the sample grown with SPE seemed to be smoother than for the sample grown by MBE. In addition, low-magnification images (not shown) revealed that misfit dislocations were limited to the bottom part of the buffer layer and to the substrate and that no threading dislocations were observed in or around the channel layer.

The electron mobilities (μ) at 4.2 K for samples with and without a graded buffer layer are summarized in figure 4 as a function of the Ge content (x) and the carrier concentration (n), which is controlled by the Sb doping level. All the mobilities increased proportionally with n owing to the screening of the ionized-impurity scattering. In addition, the mobility for the sample with a graded buffer layer was nearly one order of magnitude higher than the



sample with an SPE-grown SiGe spacer layer

MBE-grown sample

Figure 3. High-resolution images at the channel region of heterostructures formed with and without an SPE growth technique.

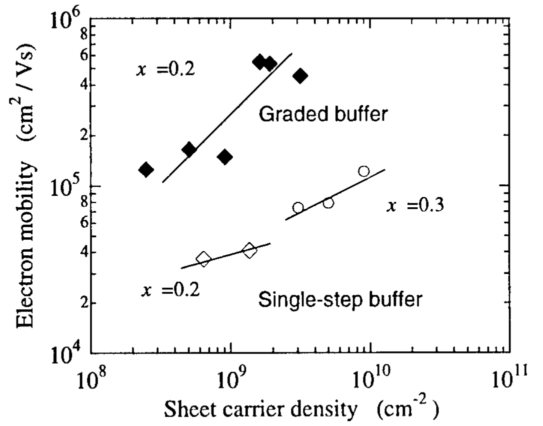


Figure 4. Electron mobilities as a function of carrier concentration for heterostructures with various types of buffer layers.

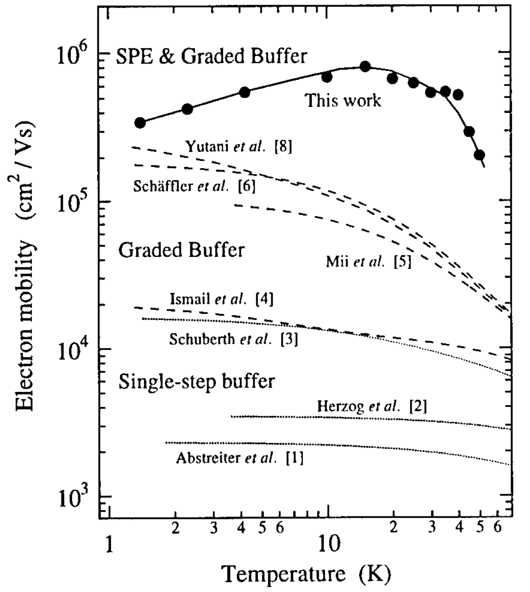


Figure 5. Temperature dependences of electron mobility for a heterostructure with a graded buffer layer ($x = 0.2, 0 \leq y \leq x$). Data from the literature [1–6, 8] are shown: ······, samples with a single-step buffer layer; - - - -, samples with a graded buffer layer.

mobility of a sample with the same n but without a graded buffer layer. Figure 5 shows the temperature dependences of electron mobility for our best sample (circles) with an SPE-grown spacer layer and a graded buffer layer ($x = 0.2$). The electron mobility had a peak at 15 K. A monotonic decrease in mobility below this temperature was observed, possibly due to ionized-impurity scattering because the carrier density (2.0×10^9 cm^{-2}) was too low to screen the scattering sufficiently. If the carrier density is increased more than one order of magnitude, the mobility at low temperatures should go up. Nevertheless, we have obtained higher mobilities compared with the data in the literatures [1–6, 8], suggesting that the interface mixing between Si and Ge was drastically suppressed by the SPE process we used. This method combining MBE and SPE promises to be a powerful tool for realizing sophisticated high-speed SiGe devices.

In summary, we have fabricated $Si_{1-x}Ge_x/Si/Si_{1-x}Ge_x$ heterostructures with an atomically flat channel. The SPE growth completely suppressed Ge segregation in the upper $Si_{1-x}Ge_x$ layer and made the interface on top of the channel flat. No threading dislocations were observed in the channel layer. By using SPE growth and a graded buffer layer, high electron mobility (8.0×10^5 cm^2 V^{-1} s^{-1} at 15 K) was obtained.

Acknowledgments

We thank Mr Takeshi Ikezu of the Central Research Laboratory, Hitachi Ltd (HCRL) for the MBE growth of the heterostructure and Dr Masahiko Hiratani of HCRL for the Hall effect measurement.

References

[1] Abstreiter G, Brugge H and Wolf T 1985 *Phys. Rev. Lett.* **54** 2441
[2] Herzog H-J, Jorge H and Schäffler F 1990 *Thin Solid Films* **184** 237
[3] Schuberth G, Schäffler F, Besson M, Abstreiter G and Gornik E 1991 *Appl. Phys. Lett.* **59** 3318
[4] Ismail K, Meyerson B S and Wang P J 1991 *Appl. Phys. Lett.* **58** 2117
[5] Mii Y J, Xie Y H, Fitzgerald E A, Monroe D, Thiel F A, Weir B E and Feldman L C 1991 *Appl. Phys. Lett.* **59** 1611
[6] Schäffler F, Többen D, Herzog H-J, Abstreiter G and Holländer B 1992 *Semicond. Sci. Technol.* **7** 260
[7] Ismail K, Arafa M, Saenger K L, Chu J O and Meyerson B S 1995 *Appl. Phys. Lett.* **66** 1077
[8] Yutani A and Shiraki Y 1997 *J. Crystal Growth* **175–176** 504
[9] Sze S M 1981 *Physics of Semiconductor Devices* 2nd edn (New York: Wiley) p 30
[10] Iyer S S, Tsang J C, Copel M W, Pukite P R and Tromp R M 1989 *Appl. Phys. Lett.* **54** 219
[11] Nakagawa K and Miyao M 1991 *J. Appl. Phys.* **69** 3058
[12] van de Walle C G and Martin R M 1986 *Phys. Rev.* B **34** 5621
[13] People R and Bean J C 1986 *Appl. Phys. Lett.* **48** 538

Semicond. Sci. Technol. **13** (1998) A143–A146. Printed in the UK

PII: S0268-1242(98)91883-1

Interface effects on intersubband carrier relaxation in GaAs/AlGaAs quantum wells

M Dür and S M Goodnick

Department of Electrical Engineering, Arizona State University, Tempe, AZ 85287-5706, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. Using an ensemble Monte Carlo simulation, we investigate intersubband relaxation of photoexcited electrons in GaAs/AlGaAs quantum wells having a subband separation smaller than the polar optical phonon energy. In this paper, we report on simulation of modulation-doped structures similar to experimental samples recently studied with intersubband pump and probe spectroscopy using a free-electron laser. For modulation-doped samples, intersubband relaxation due to intercarrier scattering is an important effect at short times, whereas at longer times polar optical and acoustic phonons dominate the intersubband decay. For low excitation density conditions, interface scattering is found to contribute greatly to the intersubband decay of the carriers during the initial pulse for even moderate interface charge densities.

The decay of photoexcited carriers in nanostructured systems from higher-energy subbands to the ground subband is a topic of potential technological interest in, for example, proposed optoelectronic devices such as infrared lasers and photodetectors in low-dimensional systems. During the past decade, intersubband relaxation of electrons from the second to the first subband in GaAs/Al$_x$Ga$_{1-x}$As quantum well (QW) structures has been the subject of extensive experimental study via optical techniques [1]. It is generally accepted that the intersubband relaxation in narrow wells, where the subband separation is larger than the polar optical phonon energy (36 meV in GaAs), is controlled primarily by polar optical phonon emission. Measured values for the electron lifetime are very short, on the order of half a picosecond, and are in good agreement with theoretical estimates of the electron–phonon scattering rate.

In wide wells, in which the subband separation is smaller than the polar optical phonon energy, electrons at the bottom of the first excited subband are forbidden by energy conservation from emitting phonons, and therefore the intersubband relaxation is expected to be slower. Long relaxation times have indeed been observed, on the order of several hundred picoseconds [2]. However, much shorter time constants, between 15 and 40 ps, have also been measured. Recently, Murdin *et al* have performed time-resolved pump and probe experiments on both uniformly doped [3] and modulation-doped [4] QWs with a free-electron laser. In these experiments, electrons were excited resonantly from the first to the second subband by a pump pulse, and the subsequent relaxation of electrons back to equilibrium was monitored by the transmission of a time-delayed probe pulse. Employing an infrared bleaching technique, relaxation times of approximately 40 ps were observed. These time constants are shorter than can be explained by acoustic phonon scattering [5]. Therefore, Murdin *et al* have suggested that, even at wide well widths, electrons in the second subband can relax by polar optical phonon emission through the high-energy tail of the thermal distribution as a result of intercarrier scattering, as was demonstrated in earlier Monte Carlo simulations [6]. Monte Carlo simulation of the relaxation in uniformly doped wide wells corresponding to [3] showed good agreement between the measured decay of the differential transmission spectrum (proportional to the excess carrier population in the second subband) and the simulated decay from the upper to lower subband [7, 8].

The purpose of the present work is to interpret more fully the Murdin experiments by means of ensemble Monte Carlo (EMC) simulations, comparing with more recent results in modulation-doped samples with variable intensity excitation [4]. The basic EMC model used here to simulate carrier relaxation in QWs has been described in detail in [1]. The details related to modelling intersubband relaxation were reported in our previous paper on intersubband relaxation in wide, uniformly doped QWs [7]. Briefly, single-band self-consistent solutions for the coupled Schrödinger and Poisson equations for the conduction band states in a GaAs/Al$_x$Ga$_{1-x}$As heterostructure are utilized to calculate the intra- and intersubband scattering rates of the quantum confined system. Scattering due to bulk-like polar optical and

0268-1242/98/SA0143+04$19.50

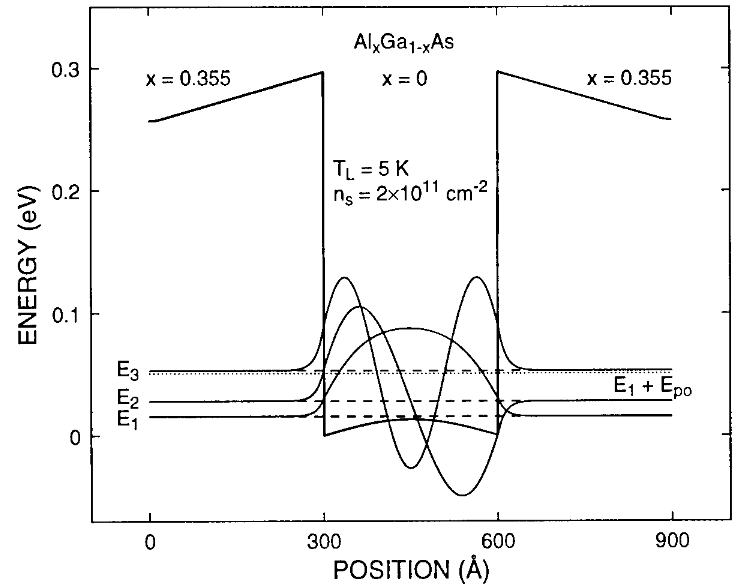


Figure 1. Self-consistent calculation of the first three energy eigenvalues and envelope functions for a modulation-doped QW structure. The separation between the first and second subband is $\Delta E_{12} = 12.1$ meV.

acoustic phonons (including its full inelastic form), ionized impurities and intercarrier interaction are considered in this paper. In order to simulate the effect of scattering from the heterojunction interfaces, we have included interface impurities (arising from interface states or from trapped impurities at the interface). These impurities are modelled as a sheet of charge located at both heterolayer interfaces with a given sheet density. Interface impurities are assumed to exist with equal density at both interfaces and are uncorrelated both within the plane and between interfaces. The Coulomb interaction is treated in the usual fashion, with a time-varying, statically screened interaction. In order to account for the change in distribution function during relaxation, the dielectric matrix is continually updated during the simulation. With the aim of more properly treating electron–electron scattering, we have improved the static screening model of [7] and included dynamic screening within the random phase approximation for intersubband intercarrier scattering [9]. The photoexcitation model in [7] has been modified to account for saturation of the intersubband absorption at high laser intensities as well.

In our previous simulation of the Murdin experiments in uniformly doped samples, results indicate that, under bleaching conditions, the intersubband decay of electrons during the photoexcitation process is limited mainly by bulk ionized impurity scattering. Polar optical phonon emission also contributes considerably to the electron decay and is associated with the thermal tail of the heated distribution function in both subbands. In more recent experiments, modulation-doped samples were employed which avoid complications due to impurity scattering. Figure 1 shows the calculated potential profile and electron energy levels in the system of Murdin *et al* [4]. The AlGaAs barriers are delta doped at a distance of 30 nm from the interface on either side, and the well width is 30 nm. The calculated subband spacing is 12.1 meV, which is smaller than the resonant absorption energy due to the depolarization shift.

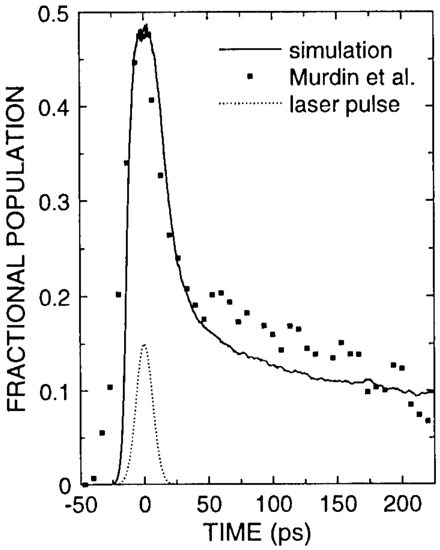


Figure 2. Population of the second subband (as a fraction of the total population) versus time for a modulation-doped sample under bleaching conditions. The simulated occupancy is compared with the experimental data of Murdin *et al.* The temporal evolution of the excitation rate due to the laser pulse is also shown (a.u.).

The simulated fractional population of the upper subband versus time for the modulation-doped sample is shown in figure 2 at 5 K for an electron sheet density of 2×10^{11} cm^{-2}. A Gaussian pulse is assumed in the simulation, and the initial and final state occupancies are included to account properly for bleaching of the absorption. The

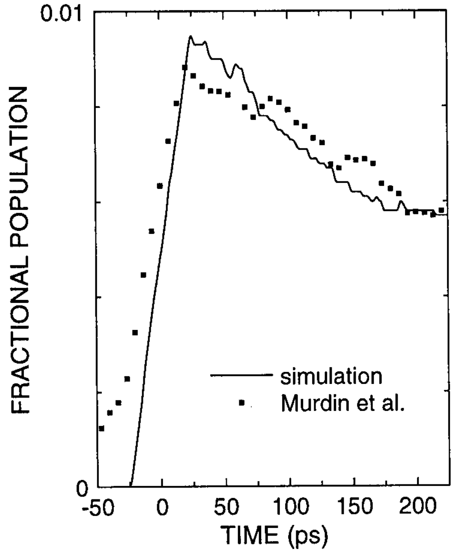


Figure 3. Population of the second subband (as a fraction of the total population) versus time for a modulation-doped sample under low excitation density conditions. The simulated occupancy is compared with the experimental data of Murdin *et al* for low-intensity excitation.

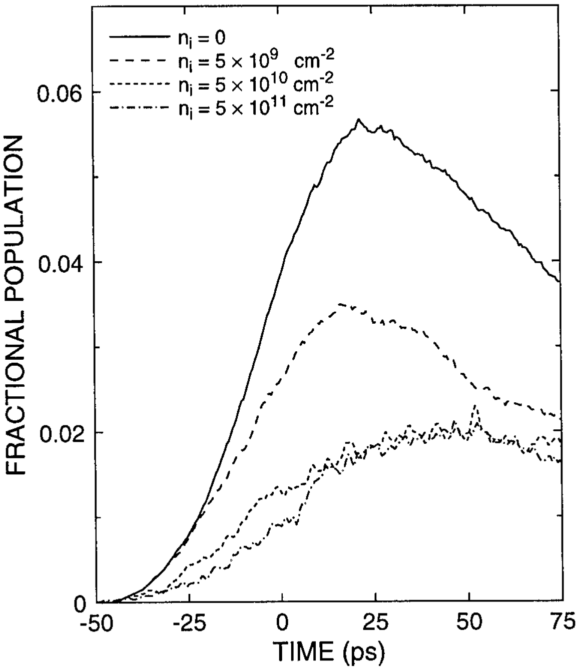


Figure 4. Simulated fractional population in the second subband versus time for several different interface impurity concentrations under low excitation density conditions.

experimental data for the measured transmission change of the sample, given in arbitrary units, are included in the figure for comparison. The population in the upper subband decays within 20 ps to half of its peak value. For longer times, the decay slows down markedly for both the experimental and the simulated results. As we previously found for uniformly doped samples, after the end of the pulse (approximately 15 ps), the two subbands quickly reach a common temperature and chemical potential as a result of intersubband scattering, and, in the long time constant regime, the decay from the second to the first subband is governed by the cooling of the distribution rather than the intersubband scattering time.

For lower excitation densities, the population in the second subband cannot build up as rapidly as in the case of intense pumping conditions, and the carrier decay shows only one longer time constant as shown in figure 3. Here the agreement between the Monte Carlo result and the experimental decay is quite good, and does not evidence the initial rapid decay. Under low excitation conditions, the two subbands thermalize (i.e. reach a common temperature and chemical potential) almost immediately, and the decay is characteristic of the energy loss rate of the system due to polar optical and acoustic phonons. The electron temperature itself is much higher than the lattice temperature because of the kinetic energy imparted to the electrons in the ground subband by the electrons in the upper subband as they relax.

Finally, figure 4 illustrates the effect of varying interface charge densities on the rise and fall of the second subband population during the laser pulse for low excitation. Interface scattering introduces an additional intersubband channel and, depending on the charge density, this may be larger than the electron–electron and electron–phonon rates. Interface scattering causes carriers to transfer back to the first subband during photoexcitation, hence reducing the peak population in the second subband. Therefore, the effect of interface scattering is to suppress the short time constant component of the relaxation resulting in a single time constant behaviour.

In summary, we have simulated the intersubband relaxation of carriers from the second subband to the first in wide QWs in which phonon bottleneck effects are expected. The simulated results show good agreement at both low and high intensity with the measured decay in modulation-doped QW structures determined using far infrared intersubband absorption spectroscopy. The electrons quickly thermalize between both bands after the pulse because of electron–electron and optical phonon scattering, resulting in a decay governed by the cooling of the combined distributions of the two subbands. The role of interface scattering is to quicken this thermalization process by providing more transfer of carriers from the upper subband to the lower one during photoexcitation.

Acknowledgments

This work was partially supported by the Office of Naval Research, Grant N00014-93-1-0618. We would also like to acknowledge the support of the Alexander von Humboldt Foundation (SMG) and the Austrian Science Foundation through the Schrödinger Fellowship (MD) in this work.

References

[1] Goodnick S M and Lugli P 1992 *Hot Carriers in Semiconductor Nanostructures* ed J Shah (New York: Academic) ch III.1, p 191

[2] Oberli D Y, Wake D R, Klein M V, Klem J, Henderson T and Morkoç H 1987 *Phys. Rev. Lett.* **59** 696
[3] Murdin B N *et al* 1994 *Semicond. Sci. Technol.* **9** 1554
[4] Murdin B N, Heiss W, Langerak C J G M, Lee S-C, Galbraith I, Strasser G, Gornik E, Helm M and Pidgeon C R 1997 *Phys. Rev.* B **55** 5171
[5] Ferreira R and Bastard G 1989 *Phys. Rev.* B **40** 1074
[6] Goodnick S M and Lugli P 1989 *Superlattices Microstruct.* **5** 51
[7] Dür M, Goodnick S M and Lugli P 1996 *Phys. Rev.* B **54** 17 794
[8] Dür M, Goodnick S M and Lugli P 1997 *Phys. Status Solidi* b **204** 170
[9] Jain J K and Das Sarma S 1987 *Phys. Rev.* B **36** 5949

Semicond. Sci. Technol. **13** (1998) A147–A151. Printed in the UK PII: S0268-1242(98)91884-3

Electron–interface phonon interaction in multiple quantum well structures

J P Sun†, H B Teng†, G I Haddad† and M A Stroscio‡

† Center for High Frequency Microelectronics, Solid State Electronics Laboratory,
Department of Electrical Engineering and Computer Science,
The University of Michigan, Ann Arbor, MI 48109-2122, USA
‡ US Army Research Office, PO Box 12211, Research Triangle Park,
NC 27709-2211, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. Intersubband relaxation rates due to electron interactions with the interface phonons are evaluated for multiple quantum well structures designed for step quantum well lasers operating at mid-infrared to submillimetre wavelengths. The interface phonon modes and electron–phonon interaction Hamiltonians for the structures are derived using the transfer matrix method, based on the macroscopic dielectric continuum model, whereas the electron wavefunctions are obtained by solving the Schrödinger equation. Fermi's golden rule is employed to calculate the electron relaxation rates between the subbands in these structures. The relaxation rates for two different structures are examined and compared with those calculated using the bulk phonon modes and the Fröhlich interaction Hamiltonian. The sum rule for the relationship between the form factors of the various localized phonon modes and the bulk phonon modes is verified. The results obtained in this work illustrate that the transfer matrix method provides a convenient way for deriving the properties of the interface phonon modes in different structures of current interest and that, for preferential electron relaxation in intersubband laser structures, the effects of the interface phonon modes are significant and should be considered for optimal design of these laser structures.

1. Introduction

In recent years, several quantum well structures have been proposed and demonstrated for achieving lasing wavelengths from the mid-infrared to the submillimetre wave region [1–4]. These lasers incorporate narrow quantum wells that must have thicknesses as small as 30–50 Å and have transition level separations engineered to one unit or a multiple of the longitudinal-optical (LO) phonon energy. The level separations are critical design parameters for these laser structures, because preferential relaxation rates between these levels are required to maintain adequate carrier population inversion between the active laser states and successful lasing operation at the designed wavelength. It is well known that electron–phonon interactions are the major physical processes for electron relaxation from the upper energy levels to the lower levels. However, similar to the modification of electron behaviour in confined structures, it is now well recognized that the shapes and energies of phonon modes are modified by quantum wells. Specifically, LO phonons in quantum wells may be described by the dielectric continuum model in terms of confined and interface (IF) modes that have properties different from those of bulk phonons [5–8]. There is clearly a need to formulate the phonon modes localized in these quantum well structures and to determine how these localized phonon modes affect electron transitions or relaxation and how they influence laser performance.

Two laser structures shown in figures 1 and 2 are studied: a symmetric quantum well structure and an asymmetric quantum well structure, consisting of $Al_{0.6}Ga_{0.4}As$ as the outermost layers, $Al_{0.25}Ga_{0.75}As$ as the step layers and the GaAs quantum well. These two structures represent basic structural cells used in novel laser structures such as the step quantum well lasers [4]. The quantities of interest in this study include the relevant confined, IF and bulk-like modes, the bound-state electronic energies and wavefunctions and the significant phonon-assisted transition rates, or the relaxation rates, for these structures. These relaxation rates are also important in quantum cascade lasers [1] and tunnelling injection lasers [2].

2. Solution method

We outline in this section the solution procedures of the relaxation rate calculations based on the theory of localized phonons in quantum well systems. The main formulations employed and developed in this work are also presented.

2.1. Dielectric continuum model

Mori and Ando [8], in their study of electron–optical phonon interaction in heterostructures, derived the interaction Hamiltonian using orthonormal eigenmodes of phonons calculated in the dielectric continuum model.

0268-1242/98/SA0147+05$19.50

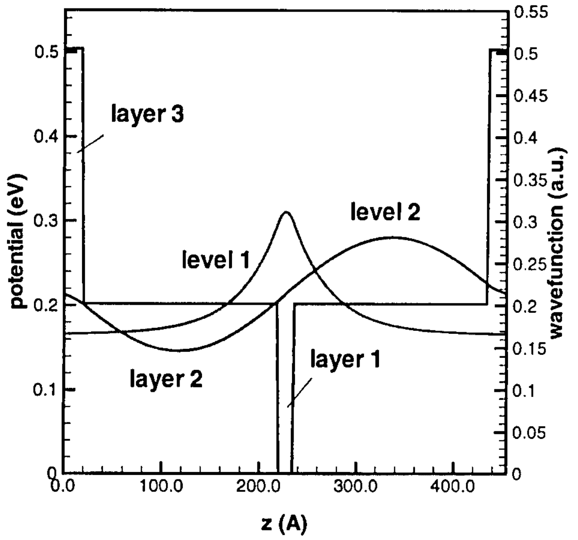


Figure 1. Symmetric quantum well structure: electron potential and calculated wavefunctions of the first two levels.

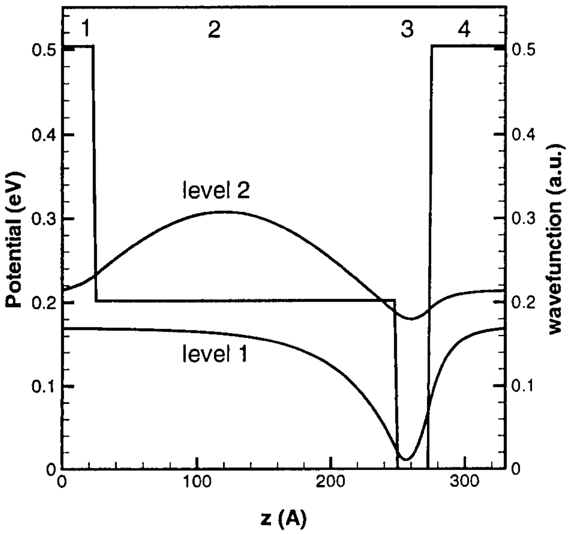


Figure 2. Asymmetric quantum well structure: electron potential and calculated wavefunctions of the first two levels. The numbers 1, 2, 3, 4 denote the different regions.

The dielectric continuum model determines optical phonon modes using classical electrostatics by solving the following equations in each layer of the heterostructure

$$\nabla^2\Phi(\boldsymbol{r}) = \frac{1}{\epsilon_0}\nabla\cdot\boldsymbol{P}(\boldsymbol{r}) \tag{1}$$

$$\boldsymbol{P}(\boldsymbol{r}) = \epsilon_0\chi_n(\omega)\boldsymbol{E}(\boldsymbol{r}) \tag{2}$$

$$\boldsymbol{E}(\boldsymbol{r}) = -\nabla\Phi(\boldsymbol{r}) \tag{3}$$

with the boundary conditions at the heterointerfaces. In equations (1)–(3), $\Phi(\boldsymbol{r})$ is the electrostatic potential as a function of the three-dimensional position vector $\boldsymbol{r}$ due to the optical phonon modes, $\boldsymbol{P}(\boldsymbol{r})$ the polarization field, $\boldsymbol{E}(\boldsymbol{r})$ the electric field, ϵ_0 the permittivity of vacuum and $\chi_n(\omega) = \epsilon_n(\omega) - 1$ the dielectric susceptibility in layer n.

The dielectric function $\epsilon_n(\omega)$ is given by [9]

$$\epsilon_n(\omega) = \epsilon_n(\infty)\frac{\omega^2 - \omega_L^2}{\omega^2 - \omega_T^2} \tag{4}$$

for a binary compound semiconductor, or

$$\epsilon_n(\omega) = \epsilon_n(\infty)\frac{(\omega^2 - \omega_{LA}^2)(\omega^2 - \omega_{LB}^2)}{(\omega^2 - \omega_{TA}^2)(\omega^2 - \omega_{TB}^2)} \tag{5}$$

for a ternary compound semiconductor, following the generalized Lyddane–Sachs–Teller relation, where the subscript L denotes the LO phonon modes and T the transverse-optical (TO) phonon modes, respectively, and $\epsilon_n(\infty)$ is the high-frequency dielectric constant.

For a given heterojunction system the normal LO phonon modes (mode frequencies and dispersion relationships) and the corresponding $\Phi(\boldsymbol{r})$, $\boldsymbol{P}(\boldsymbol{r})$ and $\boldsymbol{E}(\boldsymbol{r})$ can be obtained by solving equations (1)–(3) simultaneously with the boundary conditions at the heterointerfaces. More details of the solution procedures can be found in [8]. In their study of single heterostructures, three types of normal phonon modes have been identified: (1) IF modes, (2) half-space LO modes and (3) half-space TO modes. For double heterostructures, there are six types of modes: (1) symmetric IF modes, (2) antisymmetric IF modes, (3) confined LO modes, (4) confined TO modes, (5) half-space LO modes and (6) half-space TO modes. We will investigate properties of these modes and their influence on the electron relaxation process in the laser structures.

2.2. Relaxation rates and interaction Hamiltonian

For evaluation of electron relaxation rates between the subbands, Fermi's golden rule is employed. The relaxation rate or transition probability per unit time by the electron–LO phonon interaction is given by

$$W_{if} = \frac{2\pi}{\hbar}|\langle f|\mathcal{H}_{int}|i\rangle|^2\delta(E_f - E_i \pm \hbar\omega) \tag{6}$$

where $\mathcal{H}_{int} = -e\Phi$ is the interaction Hamiltonian and e the magnitude of electronic charge. i and f denote the electron initial and final state, respectively. ω is the phonon frequency. Once the Hamiltonians and wavefunctions are known, the relaxation rates can be calculated by numerical integration of equation (6). For the bulk cases, $\mathcal{H}_{int}$ is in the well-known Fröhlich form, whereas, in the heterostructures, changes of the Fröhlich Hamiltonian due to phonon confinement and localization must be evaluated. For simple heterostructures, the interaction Hamiltonians can be readily derived in analytical form from the localized phonon modes based on the dielectric continuum model. However, for more complicated structures such as those involving multiple quantum wells or mixed binary–ternary material systems, dozens of localized phonon modes exist and the interaction Hamiltonians become very complicated and have to be treated numerically, as performed in this work.

2.3. Form factors and sum rule

The interaction Hamiltonians are derived from the normal phonon modes and are therefore dependent on the device structure. However, the terms in the interaction Hamiltonian which are independent of material parameters such as the phonon frequencies and dielectric constants can be grouped and defined to be the *form factors* [8]. Because of the complete orthonormality of the eigenmodes, a *sum rule* exists. For example, for the double heterostructures,

$$F_B(q) = F_S(q) + F_A(q) + F_C(q) + F_{H^+}(q) + F_{H^-}(q) \quad (7)$$

where $F_B(q)$ is the form factor in the case where the electrons are assumed to interact with bulk LO phonons and q is the in-plane wavevector, F_A is for the asymmetric IF phonon modes, F_S for the symmetric IF modes, F_C for the confined modes and F_{H^+} and F_{H^-} are for the half-space phonons in the right and left half-space, respectively. A major advantage of the sum rule is that one can justify the calculations of the various phonon modes using equation (7). We made use of the sum rule in this work to verify our calculations.

2.4. Transfer matrix method

The various localized phonon modes, the related interaction Hamiltonians and the relaxation rates for simple heterostructures such as single or double quantum wells composed of binary semiconductors may be derived in analytical forms for further computation. This has been done by Stroscio [10] for a symmetric quantum well structure. However, for heterosystems consisting of multiple heterointerfaces or asymmetric quantum wells, application of even a simple macroscopic dielectric continuum model becomes highly complicated owing to the coupling of adjacent interfaces. In addition, ternary or quaternary composition will add further complexity to the phonon modes. Yu *et al* [11] developed a general transfer matrix formalism to determine the electrostatic potential and dispersion relations of the IF phonons in a multiple-IF heterostructure within the framework of the dielectric continuum model. In this work, the transfer matrix method is employed to generate the IF phonon modes and numerical Hamiltonians for the asymmetric structure. The analysis follows the procedures described in [11]. For a given IF phonon mode, the electrostatic potential $\Phi(\boldsymbol{r})$ in the region $[z_i, z_{i+1}]$ and its two-dimensional Fourier transform $\Phi_i(\boldsymbol{q}, z)$ are defined by

$$\Phi_i(\boldsymbol{r}) = \sum_{\boldsymbol{q}} \exp(-\mathrm{i}\boldsymbol{q} \cdot \boldsymbol{\rho}) \Phi_i(\boldsymbol{q}, z) \quad (8)$$

$$\Phi_i(\boldsymbol{q}, z) = c_{i-}\,\mathrm{e}^{-qz} + c_{i+}\,\mathrm{e}^{+qz} \quad (9)$$

where $\boldsymbol{\rho} = (x, y)$ and $\boldsymbol{q}$ denote the positions and wavevectors in the two-dimensional plane of the interface. The z-axis is chosen as the direction of growth. Specifically, for the asymmetric structure,

$$\Phi = \begin{cases} \Phi_1 & -\infty < z < 0 \\ \Phi_2 & 0 \le z < a \\ \Phi_3 & a \le z < b \\ \Phi_4 & b \le z < +\infty \end{cases} \quad (10)$$

where for the regions $i = 1, 2, 3, 4$ (see figures 1 and 2), Φ_i is given by

$$\Phi_i = A(C_{i-}\,\mathrm{e}^{-qz} + C_{i+}\,\mathrm{e}^{qz}) \quad (11)$$

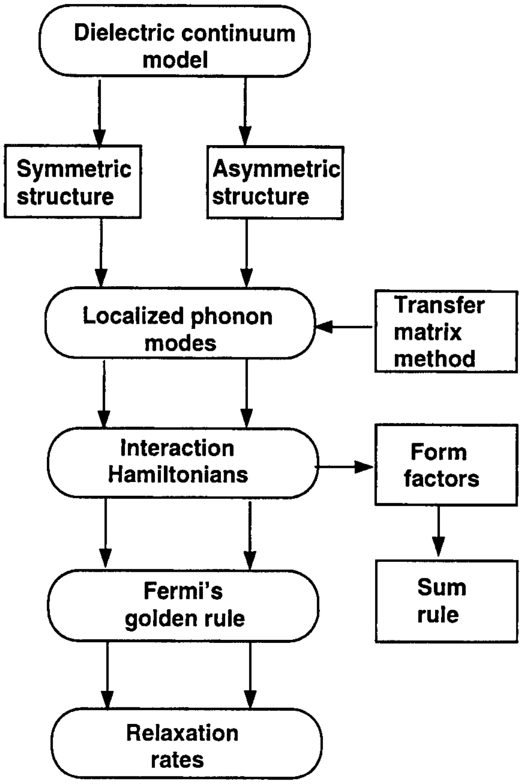


Figure 3. Outline of solution procedures employed and developed in this work.

where the normalization factor A is derived to be

$$A = \left[\sum_i \frac{\epsilon_0}{\hbar} \frac{\partial \epsilon_i(\omega)}{\partial \omega} q (C_{i+}^2\,\mathrm{e}^{2qz} - C_{i-}^2\,\mathrm{e}^{-2qz})|_{z_i}^{z_{i+1}} \right]^{-1/2} \quad (12)$$

and the coefficients $C_{i\pm}$ are functions of structural parameters and phonon in-plane wavevectors and are obtained by the transfer matrices set by the boundary conditions at the interfaces. The dispersion relations of the phonon modes are given by the solutions of

$$\begin{aligned} &(\epsilon_4 - \epsilon_3)(\epsilon_3 + \epsilon_2)(\epsilon_2 - \epsilon_1)\,\mathrm{e}^{-2qb} \\ &\quad + (\epsilon_4 - \epsilon_3)(\epsilon_3 - \epsilon_2)(\epsilon_2 + \epsilon_1)\,\mathrm{e}^{-2q(b-a)} \\ &\quad + (\epsilon_4 + \epsilon_3)(\epsilon_3 - \epsilon_2)(\epsilon_2 - \epsilon_1)\,\mathrm{e}^{-2qa} \\ &\quad + (\epsilon_4 + \epsilon_3)(\epsilon_3 + \epsilon_2)(\epsilon_2 + \epsilon_1) = 0. \end{aligned} \quad (13)$$

The solution procedures employed and developed in this work are outlined in figure 3.

3. Results and summary

For the symmetric laser structure, evaluation of the various localized phonon modes and their effects on the relaxation rates has been carried out based on the formulations of the phonon modes and interaction Hamiltonians developed by Stroscio [10]. However, the analytical approximations for electron wavefunctions at the different subbands used in [10] are replaced with more accurate wavefunctions obtained by numerically solving the Schrödinger equation for this structure. The relaxation rates between the

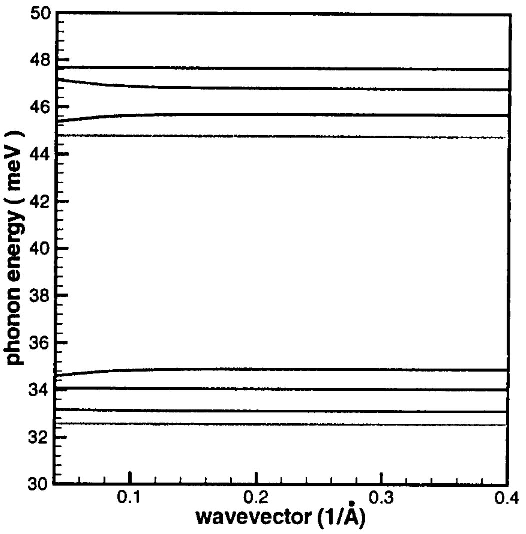


Figure 4. Calculated IF phonon modes in the asymmetric structure based on the transfer matrix method. Note that two of the modes are merged together.

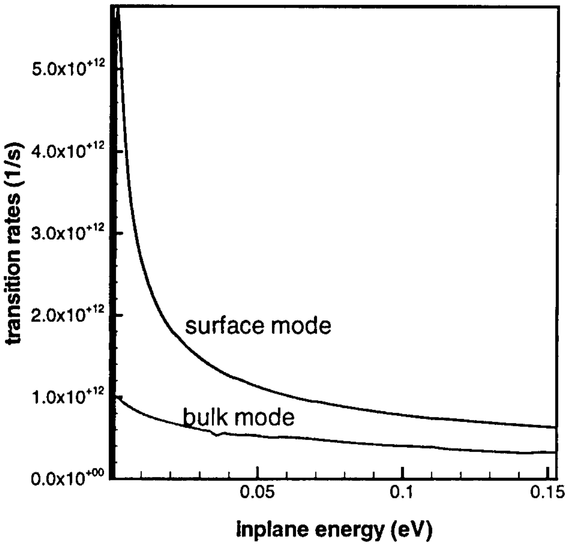


Figure 5. Transition rates as functions of electron in-plane energy in the asymmetric structure.

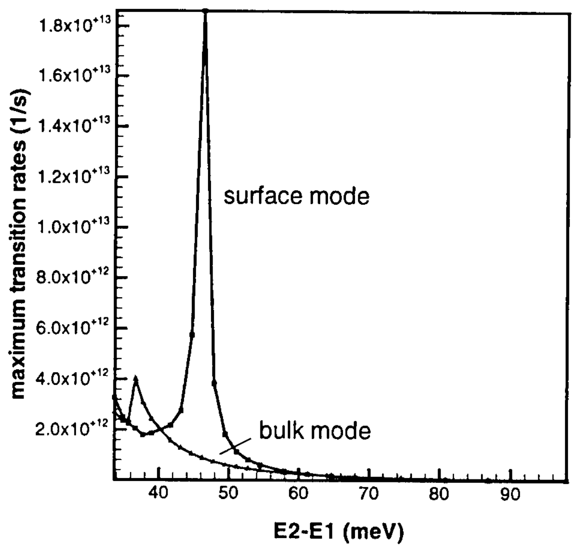


Figure 6. Resonant transition rates as functions of energy level separation in the asymmetric structure.

subbands are calculated using Fermi's golden rule with numerical integration. The electron relaxation rates are found to be dominated by the IF phonon scattering. In particular, the antisymmetric IF-phonon-assisted transition dominates at a level separation close to the IF phonon frequency ($E_2 - E_1 = 47$ meV). A maximum or resonant relaxation rate was found when the electron in-plane energy compensates the difference between the mode frequency and energy level separation. A device structure designed to have an energy level separation of the antisymmetric IF mode frequency would have a relaxation rate several times larger than that designed with the bulk phonon frequency. In order to justify our calculations, the sum rule of the form factors is verified numerically for various energy levels. The sum of the form factors of the various phonon modes added together is equal to that of the bulk mode. The agreement is excellent. More details of the findings on the symmetric structure are described elsewhere [12].

Effects of the IF phonon on electron relaxation for the asymmetric structure have also been calculated. There are 10 branches of the IF phonon modes in their dispersion relations, as shown in figure 4. Unlike the symmetric structure, the asymmetric structure does not produce distinctive symmetric and antisymmetric IF phonon branches in the dispersion spectrum. We have therefore referred to these phonon modes as the surface modes. In figure 5, the IF-phonon-assisted relaxation rates are plotted versus the electron in-plane energy. It is seen that the relaxation rate assisted by the surface phonons is approximately 6 times as large as that of the bulk phonon. Finally, in figure 6, the resonant relaxation is shown to have a sharp increase at the IF phonon energy and to have a higher relaxation rate as compared with the case of the symmetric structure.

In summary, evaluations of localized phonon effects on the electron relaxation rates in the novel quantum well laser structures performed in this work have shown that, to facilitate preferential electron relaxation for optimal laser performance, it is important to include design considerations of electron transitions induced by the IF phonons and to have electron subbands engineered with their separation close to the interface phonon energy. Further work in this direction may lead to more optimized design modifications for novel quantum well laser devices.

Acknowledgment

This work was supported by the US Army Research Office Contract DAAL03-91-C-0034.

References

[1] Faist J, Capasso F, Sivco D L, Sirtori C, Hutchinson A L and Cho A Y 1994 Quantum cascade laser *Science* **264** 553

[2] Sun H C, Davis L, Sethi S, Singh J and Bhattacharya P 1993 Properties of a tunneling injection quantum-well laser: recipe for a 'cold' device with a large modulation bandwidth *IEEE Photon. Tech. Lett.* **5** 870

[3] Afzali-Kushaa A, Haddad G I and Norris T B 1993 Optically pumped intersubband lasers *Int. Semiconductor Device Research Symp. (Charlottesville, VA, 1–3 December 1993)*

[4] Zhang X, Haddad G I, Sun J P, Afzali-Kushaa A and Sung C Y 1995 Population inversion in step quantum wells at 10 μm wavelength *53rd IEEE Device Research Conf. (Charlottesville, VA, 19–21 June 1995)*

[5] Fuchs R and Kliewer K L 1965 Optical modes of vibration in an ionic crystal slab *Phys. Rev.* A **140** 2076

[6] Lucas A A, Kartheuser E and Badro R G 1970 Electron–phonon interaction in dielectric films: application to electron energy loss and gain spectra *Phys. Rev.* B **2** 2488

[7] Licari J J and Evrard R 1977 Electron–phonon interaction in a dielectric slab: effect of the electronic polarizability *Phys. Rev.* B **15** 2254

[8] Mori N and Ando T 1989 Electron–optical-phonon interaction in single and double heterostructures *Phys. Rev.* B **40** 6175

[9] Kim K W and Stroscio M A 1990 Electron–optical-phonon interaction in binary/ternary heterostructures *J. Appl. Phys.* **68** 6289

[10] Stroscio M A 1996 Interface-phonon-assisted transitions in quantum-well lasers *J. Appl. Phys.* **80** 6864

[11] Yu S, Kim K W, Stroscio M A, Iafrate G J, Sun J P and Haddad G I 1997 Transfer matrix method for interface optical phonon modes in multiple-interface heterostructure system *J. Appl. Phys.* **82** 3363

[12] Sun J P, Teng H B, Haddad G I, Stroscio M A and Iafrate G J Intersubband relaxation in step quantum well structures *International Workshop on Computational Electronics (Notre Dame, IN, May 1997), VSLI Des.* to be published

Semicond. Sci. Technol. **13** (1998) A152–A154. Printed in the UK PII: S0268-1242(98)91836-3

High dopant and carrier concentration effects in gallium aluminium arsenide: band structure, effective carrier concentrations and mobilities

H S Bennett

Semiconductor Electronics Division, National Institute of Standards and
Technology, Bldg 225, Room B344, Gaithersburg, MD 20899, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. The changes in carrier densities of states, band edges, effective carrier concentrations and carrier mobilities due to high dopant and carrier concentration effects in $Ga_{1-y}Al_yAs$ are calculated at 300 K for donor densities, N_D, between 10^{16} cm^{-3} and 10^{19} cm^{-3} and for acceptor densities, N_A, between 10^{16} cm^{-3} and 10^{20} cm^{-3}. The mole fraction of AlAs, y, varies between 0.0 and 0.3 in these calculations. Only one quantum mechanical theory is used to treat both sides of the Mott transition in these calculations that give an internally self-consistent description of carrier transport for $Ga_{1-y}Al_yAs$ heterostructures.

1. Introduction

A common critical need identified in the technology roadmaps from the Optoelectronics Industry Development Association [1] and the National Electronics Manufacturing Initiative [2] is the need for predictive computer simulations of processes, devices and circuits. The goal of this paper is to respond to this need by calculating self-consistently from one quantum mechanical theory bandgap changes, distorted densities of states for the carriers and effective carrier concentrations for $Ga_{1-y}Al_yAs$. These calculations and the recently reported majority and minority electron and hole mobilities [3] now give together an internally self-consistent description of carrier transport in $GaAs/Ga_{1-y}Al_yAs$ heterostructures based on first principles with no fitting parameters to experimental measurements. This new self-consistent description of carrier transport across heterostructure interfaces will reduce the number of unknown or variational parameters in simulators for heterostructures and should lead to improved predictive capabilities for III–V device simulators. Even though this paper reports on the $GaAs/Ga_{1-y}Al_yAs$ system at 300 K, other ternary or elemental semiconductors may be treated by the same theory. Also, the theory and models summarized here are valid for temperatures other than 300 K.

2. Approach

This paper is the first reported use of one theory to calculate the distorted carrier densities of states for dopant densities on both sides of and at the Mott transition. The effects of carrier–dopant ion interactions on the densities of states for carriers and on the valence and conduction bands are calculated by using the Klauder self-energy method (fifth level of approximation) [4]. This approximation uses a renormalized propagator and is valid for cases both of low densities for which the scattering potentials from the dopant ions do not overlap appreciably and of high densities for which the scattering potentials from the dopant ions overlap. Solutions to the complex, very non-linear and inhomogeneous integral equations that result from the Klauder fifth-level approximation are obtained for low dopant densities that involve both bound and continuum states and for high dopant densities that involve primarily continuum states. The effects of carrier–carrier interactions (majority-carrier exchange and minority-carrier correlation) are calculated by extending the work of Abram *et al* [5] to include both finite wavenumbers and finite temperatures. The results show that quantitatively significant changes occur in the values of densities of states, locations of band edges and effective carrier concentrations when they are compared with the values usually used by researchers, particularly with those values found in most discussions of III–V heterostructures.

3. Results

The changes in carrier densities of states, band edges, effective carrier concentrations and minority electron mobilities due to high dopant and carrier concentration effects in $Ga_{1-y}Al_yAs$ are calculated at 300 K for donor

0268-1242/98/SA0152+03$19.50

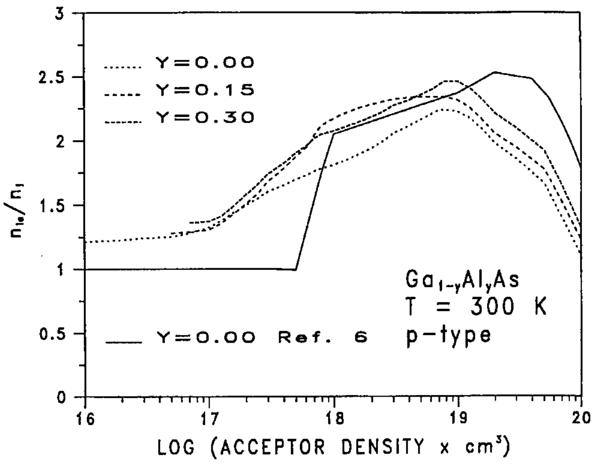


Figure 1. Effective intrinsic carrier concentration ratios n_{ie}/n_i for p-type $Ga_{1-y}Al_yAs$ as functions of acceptor density for three representative values of mole fraction $y = 0.0$, 0.15 and 0.30.

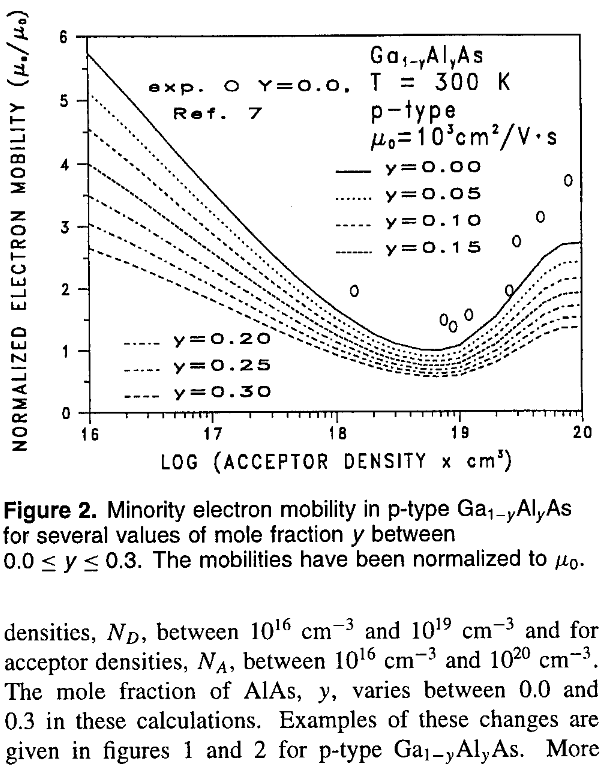


Figure 2. Minority electron mobility in p-type $Ga_{1-y}Al_yAs$ for several values of mole fraction y between $0.0 \le y \le 0.3$. The mobilities have been normalized to μ_0.

densities, N_D, between 10^{16} cm^{-3} and 10^{19} cm^{-3} and for acceptor densities, N_A, between 10^{16} cm^{-3} and 10^{20} cm^{-3}. The mole fraction of AlAs, y, varies between 0.0 and 0.3 in these calculations. Examples of these changes are given in figures 1 and 2 for p-type $Ga_{1-y}Al_yAs$. More numerical analyses are needed to determine whether the crossings of the broken curves in figure 1 for different y are physically real. The full curve in figure 1 gives past results [6], based on Klauder's third level of approximation for high densities and on the screened bound state for low densities. An interpolation method based on Klauder's fifth level evaluated at one acceptor density in the region of the Mott transition is used to go between the low- and high-density regions. The three broken curves give present results from the Klauder fifth approximation for $Ga_{1-y}Al_yAs$ with $y = 0.0$, 0.15 and 0.30. The carrier mobilities are calculated by solving the Boltzmann transport equation with a variational procedure that avoids the use of the relaxation-time approximation when the latter is not valid. Scattering rates are summed prior to the variational solutions so that Matthiesson's rule is not used. The ionized impurity and carrier–carrier scattering rates are calculated in terms of quantum mechanical phase-shift analyses so that the Born approximation is not used for these long-ranged or screened potentials. The minority electron mobility curves in figure 2 include the effects of the electrons scattering from hole plasmons and of deducting the density of holes with energies below the Fermi energy from the electron–hole scattering process. The experimental data points for the case $y = 0.0$ are from [7] and are indicated by open circles. The foregoing changes as functions of carrier density, dopant density and mole fraction of AlAs may be incorporated into simulators for heterostructures through the use of 'look-up tables' or non-linear analytic fits to the theoretical data. Interested readers may obtain by writing to the author the data shown in figures 1 and 2 and the data for band edge, bandgap and densities of states changes that are not given here.

4. Conclusions

The results of this work are important for predictive simulations of optoelectronic, microwave and electronic devices because such computer simulations require physically reasonable values of the changes in the locations of band edges, densities of states and effective carrier concentrations. For example, figure 2 suggests a different design from the one at present used for linear amplifiers in wireless networks based on air-interface standards such as code division multiple access and time division multiple access. Such amplifiers are often $GaAs/Ga_{1-y}Al_yAs$ heterojunction bipolar transistors (HBTs). To increase their operating frequencies, the product of the base resistance, R_b, and the input capacitance, C, i.e. the R_bC time constant, should be as small as possible and the minority carrier (electron) mobility in the p-type base, $\mu_e(p;\, N_A)$, should be as large as possible. Most mobility models in current HBT simulators predict that $\mu_e(p;\, N_A)$ decreases monotonically with increasing N_A, whereas figure 2 shows that $\mu_e(p;\, N_A)$ has a relative minimum. If all other parameters remain essentially the same, then figure 2 suggests that increasing the base doping N_A beyond values used currently in linear HBT amplifiers should increase their operating frequencies because a small range of acceptor densities exists for which the minority mobility increases with increasing acceptor density.

The distorted densities of states, band edge changes and effective intrinsic carrier concentrations reported here and the mobilities reported in [3] are a consistent set of input models for device simulations of heterostructures. Combining portions of the results given here, however, with other models for these quantities, which are derived from the interpretation of electrical measurements on devices themselves, may lead to incorrect descriptions of optical and electronic behaviour of heterostructures, unless great care occurs in making and in interpreting the electrical measurements.

References

[1] Pilitsis J, Pankov P and Bergh A (eds) 1996 *Optoelectronic Technology Roadmap* (Washington, DC: Optoelectronics Industry Development Association) pp 15, 16 and 32

[2] 1996 *National Electronics Manufacturing Technology Roadmaps* (Herndon, VA: National Electronics Manufacturing Initiative) pp 143, 205 and 208

[3] Bennett H S 1996 *J. Appl. Phys.* **80** 3844

[4] Klauder J R 1961 *Ann. Phys.* **14** 43

[5] Abram R A, Childs G N and Sauderson P A 1984 *J. Phys. C: Solid State Phys.* **17** 6105

[6] Bennett H S and Lowney J R 1990 *Solid-State Electron.* **33** 675

[7] Harmon E S, Lovejoy M L, Melloch M R and Lundstrom M S 1993 *J. Appl. Phys.* **63** 536

Semicond. Sci. Technol. **13** (1998) A155–A159. Printed in the UK PII: S0268-1242(98)91837-5

Experimental and theoretical study of ultra-thin oxides

E S Daniel†, **D Z-Y Ting**†‡ **and T C McGill**†§

† T J Watson Sr Laboratory of Applied Physics, California Institute of Technology, Pasadena, CA 91125, USA
‡ Department of Physics, National Tsing Hua University, Hsinchu 30043, Taiwan

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We report on an experimental and theoretical study of transport through thin oxides. The experimental study was carried out on the tunnel switch diode (TSD) which consists of an MOS junction on top of a pn junction. The properties of the TSD depends critically on the properties of the tunnel oxide layer. Our results indicate that these devices can exhibit two different modes of behaviour depending on the stress history of the oxide. An unstressed device exhibits a thyristor-like *I*–*V* characteristic with fairly low current density. As the oxide is stressed, however, the *I*–*V* characteristic discontinuously shifts into a higher-current thyristor-like mode in which current transport appears to be highly non-uniform and depends strongly on stress history. This suggests a possible structural change in the oxide layer which is not completely destructive in that the device continues to function. We present a possible theoretical model of such a structural change in which microscopic filaments are generated in the oxide. Calculations of *J*–*V* curves for such structures with varying filament heights qualitatively match stressed MOS *I*–*V* curves found in the literature and qualitatively explain the dual-mode behaviour of the TSD.

1. Introduction

Advanced metal–oxide–semiconductor (MOS) field-effect transistors and novel silicon-based devices require the use of ultrathin gate or tunnel oxides. As a result, transport and breakdown mechanisms in such oxides have been studied extensively in MOS structures, although these properties are still not completely well understood. It has been shown that the tunnel switch diode (TSD) device, consisting of an MOS junction on top of a pn junction, is quite sensitive to the transport properties of its thin oxide layer [1–3]. Therefore, studies of this device may reveal information concerning the oxide properties supplemental to that obtained in MOS research. Here, we present the stress-dependent behaviour of a set of TSD devices and we discuss the role of the oxide in determining this behaviour. The variation of these *I*–*V* characteristics has motivated the study of transport through non-uniform thin oxide layers. We present 3D quantum mechanical scattering calculations which allow theoretical investigation of the tunnelling properties of non-uniform oxides. In particular, we examine n-poly Si/oxide/p-Si devices with conducting filaments in an attempt to model stressed oxide behaviour.

2. Tunnel switch diode results

2.1. Background

The TSD, first discovered by Yamamoto in 1972 [4], consists of a stack of layers as shown in figure 1(*a*). If current is driven through these layers with negative bias on the gate, a thyristor-like *I*–*V* curve results, as shown in figure 1(*b*). For applied voltages between the peak and valley voltages, two stable current states are allowed. It has been demonstrated theoretically [2] and experimentally [1, 3] that this device relies critically on the oxide as a tunnel barrier. If the oxide is either too thin or too thick, the peak in the *I*–*V* curve disappears, and there is no range of voltage over which bistable current states may be supported.

Given that the TSD device is so sensitive to the oxide transmission properties, one would expect that the effects of electrical stress on the device would be primarily determined by the stress modification of the oxide layer. Typical MOS electrical stress studies, in which an MOS device is subjected to a constant current while the voltage is monitored over time, indicate that the oxide becomes more transmissive with time, as the voltage across the device drops [5, 6]. Initially, this increase in conductivity is gradual, and is often termed stress-induced leakage current (SILC). After some time, however, the conductivity increases sharply, indicating a 'breakdown', characterized by the total integrated charge per unit area which has

---

§ Author to whom correspondence should be addressed.

0268-1242/98/SA0155+05$19.50

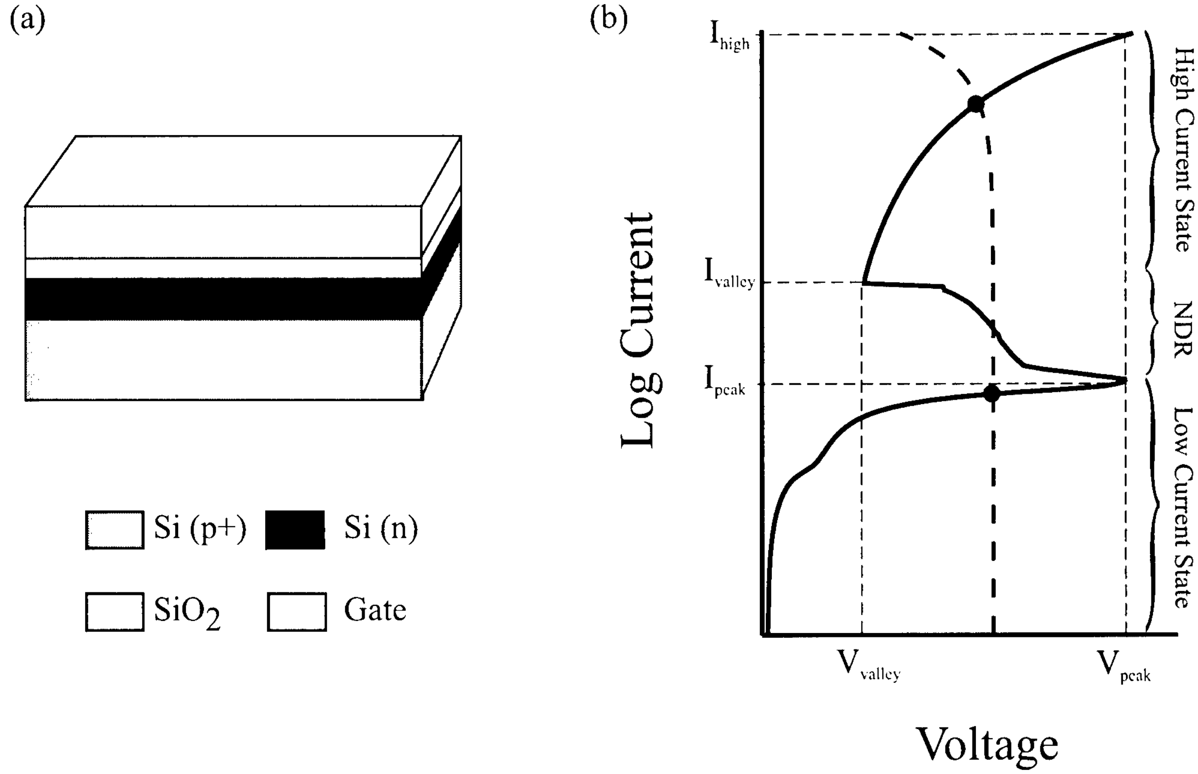


Figure 1. TSD. (*a*) The TSD device consists of a p$^+$-Si substrate with an n-Si epilayer, typically a few microns thick, a tunnel oxide (1.0–4.5 nm) and a conductive gate on top. (*b*) A typical measured *I–V* curve for a 2.5×10^{-5} cm^2 device plotted together with a 2.5 V, 10 Ω load line, illustrating the two stable current states. For this particular device, I_{high} = 100 mA, I_{valley} = 100 μA, I_{peak} = 9 μA, V_{valley} = 1 V and V_{peak} = 4 V.

by the total integrated charge per unit area which has permeated the device at this point, the charge to breakdown (Q_{bd}). Even though typical operation of the TSD device exposes the oxide to charge fluences much greater than typical Q_{bd} values, most reported studies of the TSD do not discuss any stress-related effects, except for an occasional mention of an initial 'forming' behaviour which occurs the first time the device is ever switched into the high-current state [3]. In this section, we present stress-dependent measurements of a set of TSD devices in an attempt to understand this behaviour.

2.2. Fabrication and measurement procedures

Wafers of size 6 in, consisting of p-type 2×10^{18} cm^{-3} substrates with n-type 2×10^{15} cm^{-3} epilayers, either 1.5 μm or 2.5 μ thick, were used. A field oxide was defined using a local oxidation of silicon procedure [7], and the thin tunnel oxide grown within the field oxide window. Samples with 2.0 nm, 2.5 nm, 3.5 nm and 4.5 nm oxides were prepared. Polysilicon gates (n type) were defined, covering the thin oxide region and extending onto the field oxide. Square devices with edges 85, 200, 500, 1320 μm were defined in this way. Roughly 1 μm of Al was deposited on the back surface of the wafers to provide the back contact.

Measurements of the devices were performed using an HP4145 parameter analyser connected to an Alessi probe station. Two probes were used for each measurement: one to source the current to the gate contact and the other to sense the voltage. This allowed elimination of probe contact resistance from the measurement. The contact resistance between the conducting probe station chuck and the wafer back contact was assumed to be negligible. As the HP4145 was operated as a current source rather than a voltage source, the entire *I–V* curve could be measured in one sweep without any hysteresis. Further processing and measurement details can be found elsewhere [8].

It was found that the act of measuring a single *I–V* curve over a current range such as that shown in figure 1(*b*) could significantly change the TSD behaviour, as evidenced by subsequent *I–V* curve measurements. As mentioned above, this 'forming' behaviour has been observed by others as well. In order to characterize better this forming process, a simple method was used, which proceeded as follows. A device which had never been subjected to electrical stress would be connected to the parameter analyser, and an *I–V* curve would be scanned from zero current to some small maximum current. This process yielded the 'virgin' scan. While this measurement proceeded, the stress induced by the measured currents may have modified the device. A subsequent scan to a slightly higher maximum current would then reveal a modified *I–V* curve, while further stressing the device. Another scan to a still higher maximum current could then be taken, and so on. This method, which we will refer to as 'incremental current stressing', was used to obtain sets of *I–V* curves for each device studied.

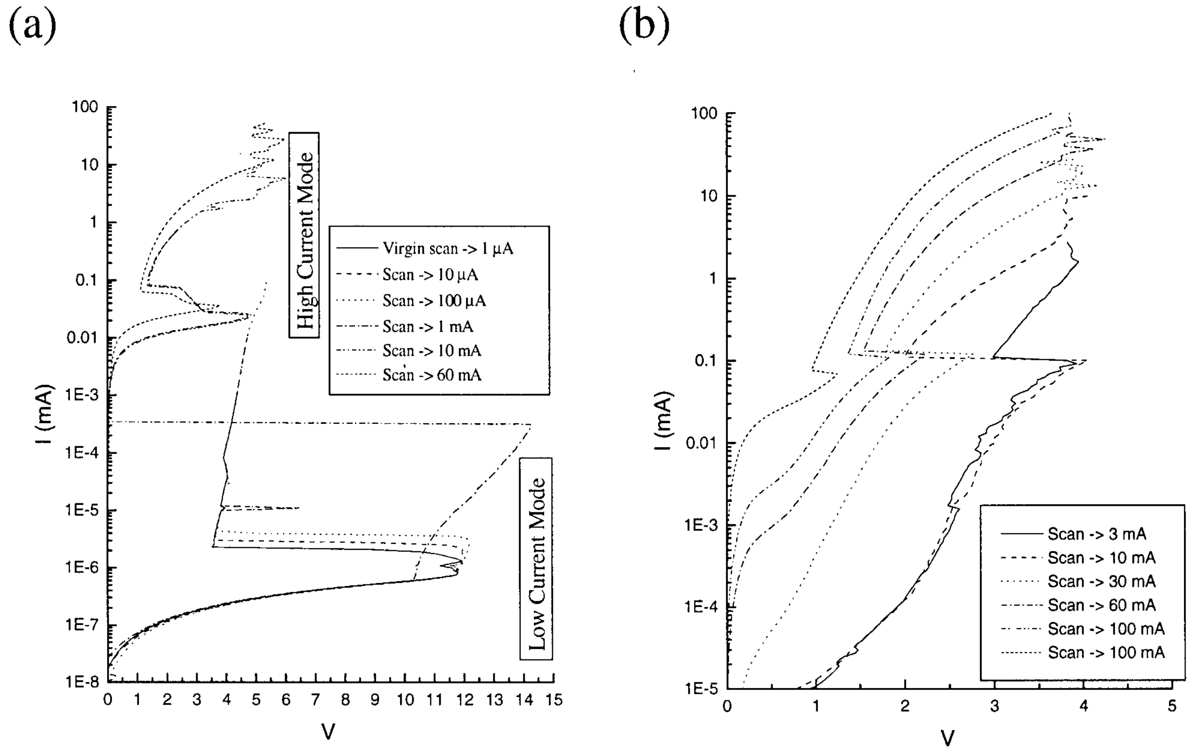


Figure 2. (*a*) Incremental current stress *I–V* curves for a 200 × 200 μm^2 TSD device with a 3.5 nm oxide layer and a 2.5 μm epilayer. The device exhibits two distinct 'modes' of operation, each of which can support low- and high-current states characteristic of the TSD. (*b*) Incremental current stress *I–V* curves for an 85 × 85 μm^2 device with a 2.0 nm oxide layer and a 1.5 μm epilayer, illustrating a voltage-limiting behaviour.

2.3. Results and discussion

Consider the incremental current stress curves shown in figure 2(*a*). The three lowest stress scans each sweep out a very low current bistable TSD *I–V* curve and are very nearly identical except for a slight increase in the switching current with stress. After a maximum current stress of 0.1 mA (corresponding to a voltage of ~ 5.4 V), however, the *I–V* curve begins to change significantly. During the subsequent sweep, the voltage drops suddenly to zero in what appears to be some kind of breakdown but, as the sweep continues, a higher-current TSD *I–V* curve results, and remains (albeit modified) in subsequent scans. Thus, a 'low-current TSD mode' is observed when the device has been subjected to little or no stress, and a 'high-current TSD mode' abruptly takes over once some critical stress has been reached. Both 'modes' exhibit bistable current states for a range of voltage. This behaviour was observed in nearly all devices measured.

Note that there are several major differences between these two TSD modes. For example, the low-current mode does not change appreciably with incremental current stress until the device is forced into the high-current mode. The high-current mode, however, changes substantially with each incremental scan, and the oxide seems to become just conductive enough to limit the voltage to a fixed maximum of a few volts (see figure 2(*b*)). This limit voltage was found to decrease with decreasing oxide thickness, but the limiting electric field was not constant, as might be expected. Another difference is found in the shape of the high-current state in each of the two modes. The *I–V* curve in the high-current state of the low-current mode is exponential in shape (linear on the log–linear plot), and extremely steep, spanning nearly 5 orders of magnitude in current over a range of roughly 2 V. The high-current state of the high-current mode, however, does not have an exponential shape, and increases only 2 orders of magnitude over a range of a few volts, suggesting a somewhat different transport mechanism. Many more differences between the two TSD modes were found, as discussed elsewhere [8].

The relatively weak stress dependence and the exponential high-current state of the low-current mode of the TSD suggest that the oxide has not broken down and is in the SILC regime. Similarly, the high stress levels and the voltage-limiting behaviour of the high-current mode suggest that this mode results after the oxide has reached breakdown. If this is true, however, the fact that the device can support two current states in the high-current mode suggests that the broken-down oxide still acts as a partial barrier. In the next section, we explore one possible physical model which can reflect this behaviour.

3. Simulations of stressed MOS structures

3.1. Background

Non-uniformities in oxides are typically not treated theoretically because of the dramatically increased complexity and computational demands. However, they can have dramatic effects on the current–voltage characteristics of

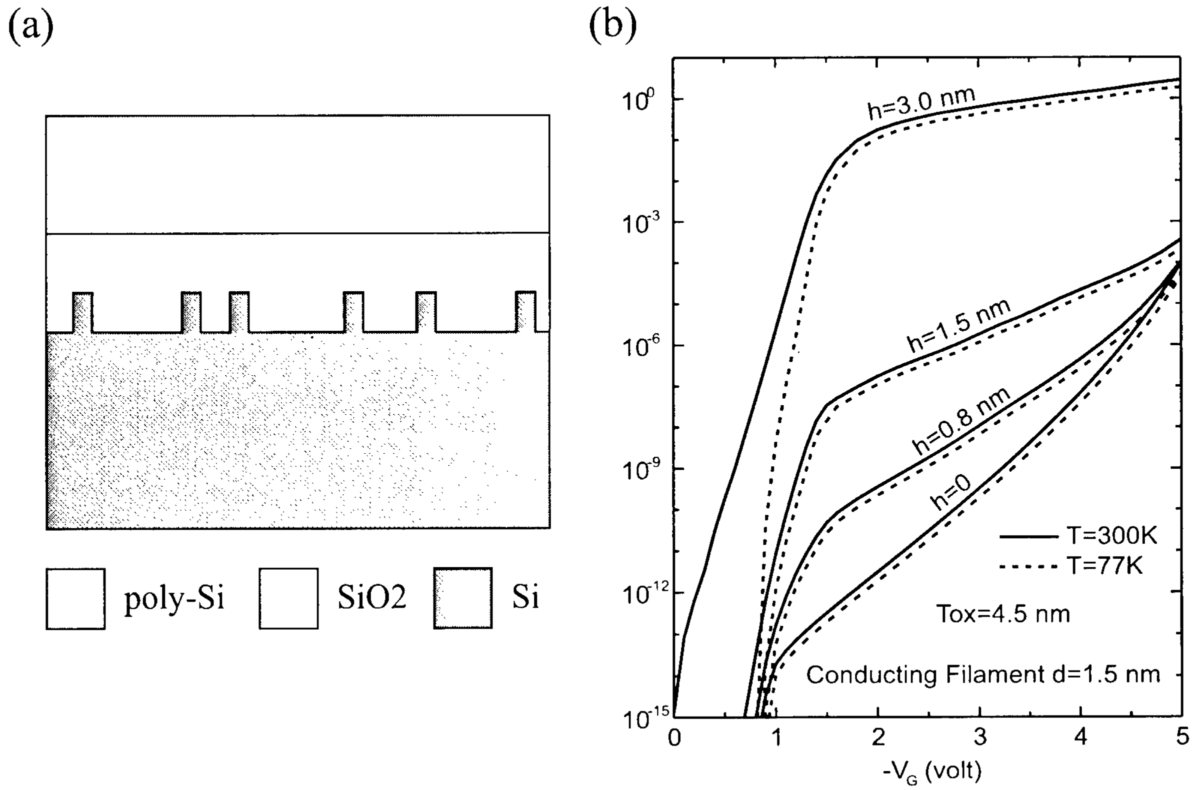


Figure 3. (*a*) Modelled MOS structure with Si filaments embedded within the oxide. (*b*) Calculated current density–voltage curves for a set of n^+ poly-Si/SiO$_2$/p-Si tunnel structures with oxide-embedded cylindrical conducting filaments with various cylinder heights. The oxide thickness is 4.5 nm. Conducting filaments have a diameter of 1.5 nm, and cover approximately 10% of the cross-sectional area.

MOS tunnel structures with ultrathin oxide barriers. One possible type of non-uniformity results when conducting filamentary paths are created in an oxide layer. Several researchers have postulated the stress-induced formation of localized conducting filaments extending from the Si–SiO$_2$ interface into the oxide as a model for breakdown or quasi-breakdown [5, 9, 10]. Here, we use a 3D model which allows calculation of the J–V characteristics of MOS tunnel structures containing such non-uniform oxide layers.

3.2. Method

A standard treatment of oxide tunnelling uses a 1D potential to describe the oxide barrier. In order to treat interfacial non-uniformity, a 3D description is necessary. In principle, variations in the non-uniform potential extend indefinitely in the directions along the interface. In practice, we do not perform computation on an infinite domain, but use instead a quasi-3D supercell geometry, in which a finite 3D cell is replicated infinitely in the two dimensions parallel to the layer interfaces, to approximate the physical structure. We treat the problem of tunnelling through this system using the open boundary planar supercell stack method [11].

Our model of the MOS tunnel structure consists of an n^+ poly-Si electrode, followed by the non-uniform oxide layer, and finally a p-type silicon region (see figure 3(*a*)). Because the nature of the filamentary conducting material is not well known, we choose to fill the cylinders with silicon for simplicity. In order to obtain J–V curves, transmission coefficients are calculated for a range of energies at each bias voltage and integrated to obtain a current density. The values of the parameters used and the details of the calculations are presented in [12].

3.3. Results and discussion

We consider a set of structures with 4.5 nm thick oxides, embedded with cylindrical conducting filaments 1.5 nm in diameter. The filaments account for approximately 10% of our computational domain in cross-sectional area, and extend from the SiO$_2$–Si interface into the oxide layer with cylinder heights of h = 0.8, 1.5 and 3.0 nm. A fourth, 'undamaged' (h = 0) structure is also included for comparison.

Figure 3(*b*) shows the J–V curves for these structures calculated at 300 K and 77 K. In general, the current densities increase dramatically with the filament length, h. This is due to lateral localization of transmitting state wavefunctions in the more highly conductive filaments. In order to estimate quantitatively the fraction of current flowing through the filaments, we define the 'filament transmission fraction' for a transmitting state as the sum of probability densities over the filament sites, divided by the total probability densities in the layers which contain the filaments.

Filament transmission fractions for tunnelling states with incoming energy equal to E_F^M at forward gate biases of 2 V (direct tunnelling) and 5 V (Fowler–Nordheim tunnelling) were calculated in the case of each cylinder height. In all cases, the filament transmission fraction was found to exceed greatly 10%, the fraction of cross-sectional area occupied by the filaments. This clearly indicates that conduction is strongly localized to the filaments. In the direct tunnelling regime, the filament transmission fraction was found to be roughly 80–90% for all cylinder heights. The localization was found to be weaker in the Fowler–Nordheim regime, ranging from 26% for $h = 0.8$ nm to 46% for $h = 1.5$ nm to 87% for $h = 3.0$ nm. This is because the rough edge of the oxide is biased below the incoming electron energy in the Fowler–Nordheim regime, so the transmitted electron wavefunction is evanescent only through part of the filament. Therefore, the electron only tunnels through some fraction of the oxide layer adjacent to the gate. Since tunnelling properties in this regime are primarily determined by this leading edge of the oxide barrier, the $h = 0.8$ and 1.5 nm structures show current densities which converge with that of the undamaged ($h = 0$) structure as high bias. The filament in the $h = 3.0$ nm structure, however, extends sufficiently close to the leading edge of the tunnel barrier so that the large current density increases persist even at higher biases.

Our J–V curves of increasing filament heights bear a very strong qualitative resemblance to experimental I–V curves of MOS devices under increasing levels of stress [5, 6, 9]. It is worth noting that within our model we have reproduced levels of stress all the way from 'unstressed' to 'breakdown' using the same model, with the only difference being the height of the filaments. We note that in particular we can reproduce the breakdown behaviour without using filaments which completely permeate the oxide; we have kept the filament height to under 3 nm for the 4.5 nm oxide, as suggested by Hirose *et al* [9]. This may explain why the TSD device may operate in the high-current mode even after breakdown has been reached, as the oxide still is a partial contiguous barrier.

Another issue of note is the dependence on the fraction of gate area occupied by the filaments. We have performed a few calculations where we kept the filament size the same, but enlarged the supercell size by a factor of 4, effectively reducing filament density by a factor of 4. We find that the corresponding current density reduction is approximately fourfold in the direct tunnelling regime, but less in the Fowler–Nordheim regime. In fact, in the case of short filaments, the high bias current densities are essentially independent of filament density, since, again, high-bias currents are primarily controlled by the leading edge of the barrier, which is out of reach for the short filaments. These variations in scaling as a function of applied bias are consistent with the above discussion of filament transmission fraction.

The $h = 3.0$ nm curve differs from the others in its low-bias temperature dependence. In the $h = 0$, 0.8, and 1.5 nm cases, the 77 K and 300 K results appear essentially the same. In the $h = 3.0$ nm curve, however, current densities at low biases increase significantly with temperature. This is due to resonant tunnelling through quantum-confined states in the filaments. This behaviour is described in detail elsewhere [12]. It is therefore possible that temperature-dependent measurements of stressed-oxide I–V curves may provide further insight into the applicability of this filament model.

4. Conclusions

We have found that n^+ poly-Si/SiO$_2$/n-Si/p-Si TSD devices can exhibit two modes of operation depending on the level of stress to which they have been subjected. These modes seem to correspond to the SILC and breakdown phases found in MOS device stress experiments. Even though the oxide in the TSD devices has reached breakdown in the high-current mode of operation, the device can support bistable current states, suggesting that the oxide still acts as somewhat of a tunnel barrier. These results, as well as MOS stress results, can be understood in terms of a model in which increased stress creates larger and larger conducting filaments extending from the Si–SiO$_2$ interface. We performed 3D quantum mechanical calculations to analyse the current–voltage characteristics of n^+ poly-Si/SiO$_2$/p-Si tunnel structures containing such oxides with embedded filaments. The filaments are found to act as localized conduction paths, leading to dramatic increases in current densities. Depending on the filament length, this model can produce current–voltage characteristics reminiscent of those observed experimentally for ultrathin oxides subjected to wide ranges of stress.

Acknowledgments

This research was supported in part by the Defense Advanced Research Projects Agency and monitored by the Office of Naval Research under contract number N0014-93-1-0710 and by the Office of Naval Research under contract number N00014-89-J-1141.

References

[1] Pettersson P O, Zur A, Daniel E S, Levy H J, Marsh O J and McGill T C 1998 *IEEE Trans. Electron Devices* submitted

[2] Simmons J G and El-Badry A A 1978 *Radio Electron. Eng.* **48** 215

[3] Kroger H and Wegener H A R 1978 *Solid-State Electron.* **21** 643

[4] Yamamoto T and Morimoto M 1972 *Appl. Phys. Lett.* **20** 269

[5] Halimaoui A, Briére O and Ghibaudo G 1997 *Microelectron. Eng.* **36** 157

[6] Depas M, Nigam T and Heyns M M 1996 *IEEE Trans. Electron. Devices* **43** 1499

[7] Weste N H E and Eshraghian K 1993 *Principles of CMOS VLSI Design* (Reading, MA: Addison-Wesley)

[8] Daniel E S 1997 *PhD Thesis* California Institute of Technology

[9] Hirose M, Alay J L, Yoshida T and Miyazaki S 1996 *Proc. Electrochem. Soc.* **96-1** 485

[10] Apte P P and Saraswat K C 1994 *IEEE Trans. Electron Devices* **9** 1595

[11] Ting D Z-Y, Kirby S K and McGill T C 1993 *J. Vac. Sci. Technol.* **11** 1738

[12] Ting D Z-Y and McGill T C 1998 *Appl. Phys. Lett.* submitted

Semicond. Sci. Technol. **13** (1998) A160–A162. Printed in the UK

PII: S0268-1242(98)91838-7

Electrically programmable resistor on semi-insulating GaAs

Jun Shen†

Department of Electrical Engineering and Center for Solid State Electronics
Research, Arizona State University, Tempe, AZ 85287, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We report the observation and results of subsequent studies of an electrically programmable resistor on semi-insulating GaAs. Several possible explanations of the phenomenon (heating and surface state charging) are proposed.

Non-volatile memory devices are readily available in Si technology in the form of electrically programmable read-only memories (EPROMs) for code storage in an integrated circuit (IC). The EPROMs are based on the principle that threshold voltage shifts can be induced by charging and discharging the floating gate in a field effect transistor. In GaAs technology, however, EPROMs, and, in general, non-volatile memories, are not available because of the lack of a good gate insulator with low interface state densities. Recently, the growing wireless communication industry has revitalized GaAs technology and at the same time pointed out the need for non-volatile memories on GaAs ICs. Here we report the observation and results of subsequent studies of an electrically programmable resistor on semi-insulating GaAs which can potentially be used for such applications.

The basic structure of the resistor consists of a pair of metal contacts evaporated on the surface of a semi-insulating GaAs wafer, with a spacing of about several microns. A filament ('bridge', 'tunnel', with a cross-sectional width of ~1 μm) can be formed when several volts of bias (electrical field strength ~10^6 V cm^{-1}) are applied across the contacts (figure 1). Then, the structure becomes conductive (5–50 Ω) and its resistance can be electrically programmed (figure 2).

A typical programming procedure is as follows: (i) the low-resistance state (~5 Ω) can be obtained by applying a relatively high voltage (say ~1.5 V) across the contacts, and (ii) the high-resistance state (~50 Ω) can be achieved by applying a voltage of ~0.8 V. An apparent negative differential resistance (NDR) region usually accompanies the switching in case (ii). Resistance of the states can be 'read' by setting a voltage (say 0.2 V) smaller than the 'writing' voltage signals so that the 'read' is non-destructive. These effects can be found in many similar structures and the sequence is repeatable (figure 3). The phenomenon does not seem to be sensitive to substrate choice (as long as it is semi-insulating GaAs) or processing. Also, once the states (low or high) are set, their resistance values are basically unchanged over time when the voltage goes to zero.

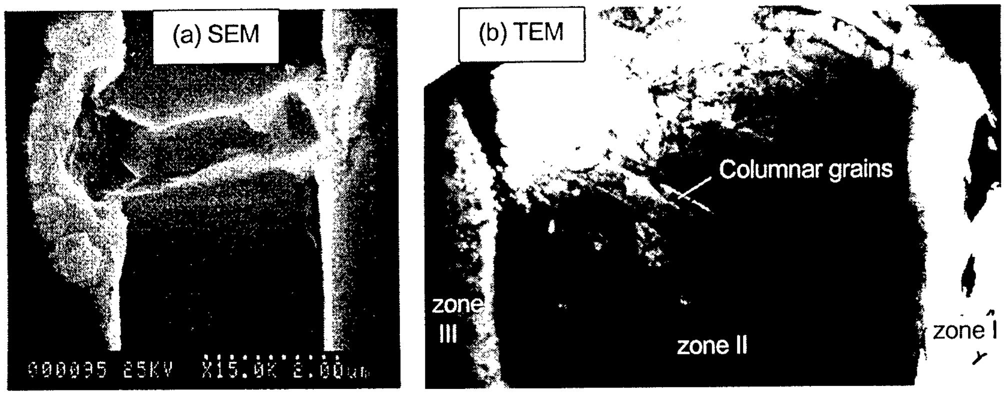


Figure 1. SEM and TEM pictures of a filament. Polycrystalline regions can be seen in the structure. Zone I corresponds to the void. Zone III is closest to the unperturbed GaAs single crystal substrate. Zone II is the columnar polycrystalline region in between I and II.

---

† E-mail address: jshen@asu.edu

0268-1242/98/SA0160+03$19.50 © 1998 IOP Publishing Ltd

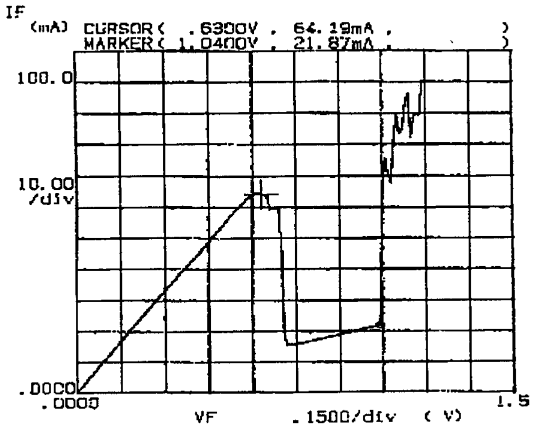


Figure 2. Typical current–voltage characteristics of the filament. An apparent negative differential resistance region and two distinctive resistance states can be seen from the curve.

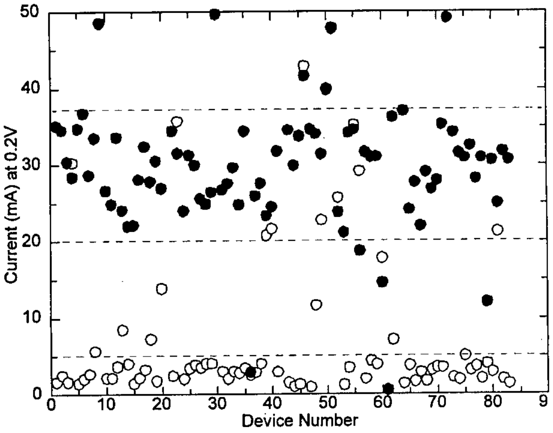


Figure 3. Random sampling of a matrix of devices on an as-fabricated wafer according to the standard 'write' and 'read' cycles. Full and open circles represent the nominal low- and high-resistance states, respectively. Some 'wrong' bits are present, as can be seen by their wrong places (open circles in the upper band and vice versa).

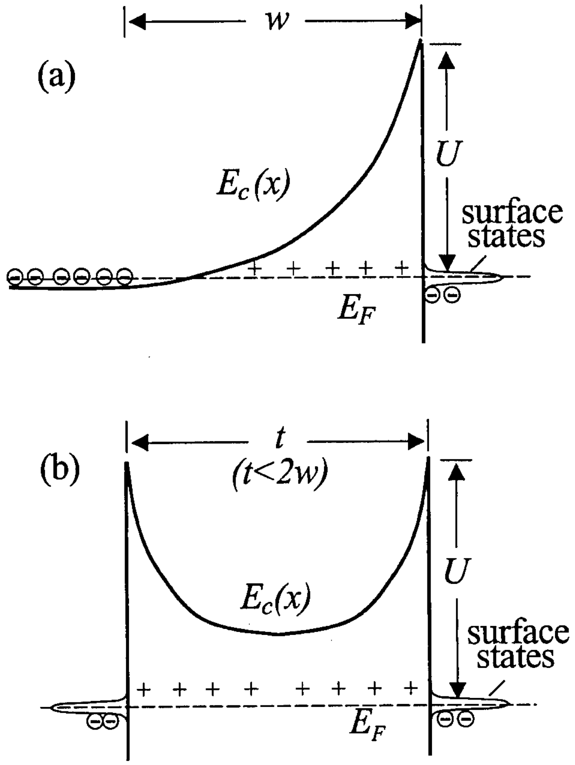


Figure 4. Illustration of conduction band profiles near a semiconductor surface. (*a*) Thick semiconductor (bulk). (*b*) Thin nanostructure. The nanostructure can be either insulating (depleted) or conductive depending on the surface charge states.

In order to understand the mechanism leading to the phenomenon, several analyses (SEM, TEM, EDS, etc) were performed (figure 1). Cross-sectional TEM pictures showed the formation of polycrystalline regions with different grain sizes. Traces of contact metal (Au in this case) were also found in the filament.

The formation of the filament can be understood as follows. A localized high current (due to localized high electric fields arising from local asperities in the metal pads) causes heating of a strip of GaAs. The heating power per unit volume is approximately 10^{12} W cm^{-3} ($V = 10$ V, $I = 100$ mA, volume $v = 1 \times 1 \times 1$ μm^3 = 10^{-12} cm^3). Assuming adiabatic joule heating (short time-scale), the time that takes the filament temperature to increase by ΔT is $\tau \sim (c\rho v/VI)\Delta T$, where c is the filament specific heat and ρ is the specific mass. For GaAs, $c = 0.35$ (J g^{-1} K^{-1}) and $\rho = 5.32$ (g cm^{-3}). With the above heating power and volume, the time to cause local melting (GaAs melting temperature $T = 1238$ °C) is about 2 ns. During the process, part of the material evaporates, forming a trench, and the rest re-crystallizes when the current flow stops, forming the polycrystalline conductive channel.

The mechanisms that determine the resistance of the channel are not well understood. One possibility may be related to the conductive volume changes associated with the subsequent heating and cooling effects during programming, as proposed for the Si antifuses [1]. Another possibility may be related to the charging and discharging of surface states in the grain boundaries, to be explained as follows.

Assuming that the conductivity of the filament is controlled by a few large single-crystal grains of the size of several hundred angstroms, then the surface charge states of the grains can completely determine the conductivity. For most semiconductors, it is well known that the surface Fermi level is pinned near the middle of the bandgap. Then, depending on the type and concentration of extrinsic doping, different degrees of band bending occur near the surface. The band profile can be obtained by solving Poisson's equation. At equilibrium, the rules of thumb are: (i) the total charge is neutral; (ii) the Fermi level lines up everywhere. For a uniform doping density of N, one can obtain the conduction band profile $E_c(x)$ to be parabolic in x, and a depletion length $w = \sqrt{2\varepsilon\varepsilon_0 U/qN}$ where the band becomes flat (or approaches the bulk) (figure 4(*a*)). For $U = 0.7$ eV, $N = 10^{18}$ cm^{-3} (typical GaAs parameters), $w = 305$ Å. Now, if the crystal is so small such that its radius (t) is shorter than $2w$ (figure 4(*b*)), then the surface depletion region extends to cover the whole bulk of the crystal, or in other words, the whole crystal becomes completely depleted. The bulk charge may not be able to balance the surface charge. This effect can have significant consequences to the nanostructures:

(i) A heavily doped nanostructure can be an insulator. This is quite obvious because the conduction of the nanostructure is no longer controlled by the bulk, but rather, by the surface depletion region, which has few mobile carriers.

(ii) An undoped nanostructure can be conductive. This happens when the surface is in a metastable charged state that induces a conductive surface 'inversion' layer.

(iii) The conductive and insulating states can have very long lifetimes (metastable). This is due to the localized nature of the surface states and the non-equilibrium band bending which prohibits efficient exchange of the charges between the bulk and surface. By charging and discharging these surface states, one can change the conductivity of the crystal grains, leading to the observed memory effect.

Other non-volatile memory phenomena in intentionally fabricated small (diameter ~0.1 μm) resonant tunnelling diodes have also been reported and may be related to our observations and explanations [2, 3].

In summary, we have reported the results of an electrically programmable resistor on semi-insulating GaAs. Several possible explanations are proposed. Further study on the origin of the phenomenon and improvements on its reliability can potentially render the device for high-density non-volatile memory applications.

Acknowledgments

The results were obtained when the author was with Motorola, Inc. TEM analysis was performed by Dr David Theodore of Motorola, Inc.

References

[1] Shacham-Diamand Y 1993 *IEEE Trans. Electron Devices* **40** 1780

[2] Gullapalli K K, Tsao A J and Neikirk D P 1993 *Appl. Phys. Lett.* **62** 2856

[3] Alphenaar B W, Durrani Z A K, Heberle A P and Wagner M 1995 *Appl. Phys. Lett.* **66** 1234

Semicond. Sci. Technol. **13** (1998) A163–A164. Printed in the UK PII: S0268-1242(98)91839-9

Interface roughness and polar optical phonon scattering in $In_{0.53}Ga_{0.47}As/AlAs/InAs$ RTDs

R Lake, G Klimeck and D Blanks

Applied Research Laboratory, Raytheon TI Systems, PO Box 655936, MS 134,
Dallas, TX 75265, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. The contributions of interface roughness scattering and polar optical phonon scattering to the valley current of $In_{0.53}Ga_{0.47}As/AlAs/InAs$ resonant tunnelling diodes (RTDs) are theoretically found to be comparable. An $In_{0.53}Ga_{0.47}As/AlAs/InAs$ RTD design is suggested to experimentally observe the phonon peak which has never been observed in this material system. Such a device will provide a calibration point for the theoretical calculations.

For most resonant tunnelling diodes (RTDs), sp^3s^* calculations indicate that the valley current at room temperature is largely the result of coherent transport through the second resonant state in the well [1,2]. State-of-the-art RTDs use an InAs notch in the well to increase the separation of the first and second resonant states [3]. We are aware of only one theoretical investigation of incoherent scattering in these devices [4]; interface roughness was found to contribute far more to the valley current than polar optical phonon scattering. We find the relative contributions to be comparable. An $In_{0.53}Ga_{0.47}As/AlAs/InAs$ RTD design is presented to observe the phonon peak [5–7] which has never been observed in this material system. Such a device will provide a calibration point for the theory.

Our calculations use a single-band tight-binding model and the scattering is included with the models and self-energies described in appendix A of [8]. Interface roughness is modelled with 5 nm exponential correlation [8,9]. The bulk $In_{0.53}Ga_{0.47}As$ polar optical phonon model is used. Self-energies are included using the multiple sequential scattering algorithm illustrated by figure 4 of [8] with $N = 4$. The electrostatic potential is calculated self-consistently in the absence of incoherent scattering and it is then used for the scattering calculations [9].

We are not aware of any studies of the confined and interface phonon modes in $In_{0.53}Ga_{0.47}As/AlAs/InAs$ RTDs, while numerous studies exist for the GaAs/AlAs devices [7,10]. For a GaAs/AlAs RTD with 2.5 nm AlAs barriers and a 5 nm GaAs well, the phonon peak resulting from bulk phonons and the phonon peak resulting from confined and interface modes are essentially identical [11]. Therefore, the bulk phonon model should provide an accurate prediction of the shape, magnitude and voltage position of the phonon peak.

Figure 1 shows both the experimental and calculated current–voltage (I–V) curves for an $In_{0.53}Ga_{0.47}As/AlAs/InAs$ RTD at room temperature. Interface roughness is included at all interfaces. The experimental I–V is asymmetric, although the structure is nominally symmetric. We have flipped the third quadrant onto the first to show the asymmetry. The simulation with the lowest valley current is the coherent tunnelling calculation labelled 'NS'. The inclusion of acoustic phonon and alloy scattering raises the valley current, but it is barely noticeable on a linear scale. The curve labelled 'AP Alloy IR' includes scattering from acoustic phonons, alloy disorder and interface roughness.

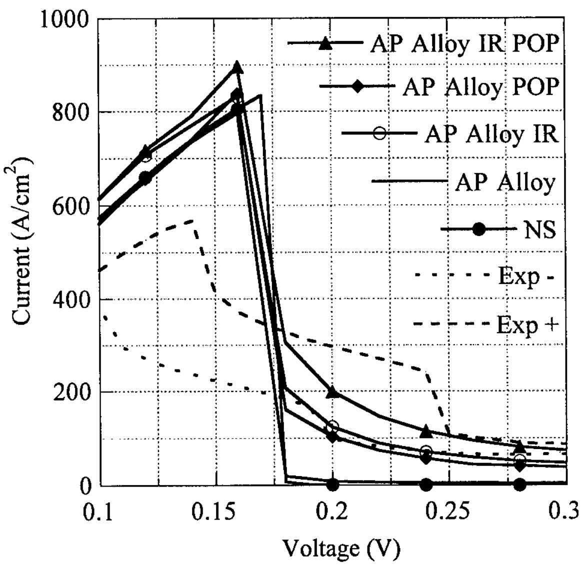


Figure 1. Comparison of experimental and calculated I–Vs at 300 K including various incoherent scattering mechanisms. The RTD consists of 10^{18} cm^{-3} n^+ $In_{0.53}Ga_{0.47}As$ leads, 2 nm intrinsic $In_{0.53}Ga_{0.47}As$ spacer layers, 2.9 nm AlAs barriers, and a 1.5 nm/2.0 nm/1.5 nm $In_{0.53}Ga_{0.47}As/InAs/In_{0.53}Ga_{0.47}As$ well.

0268-1242/98/SA0163+02$19.50 © 1998 IOP Publishing Ltd

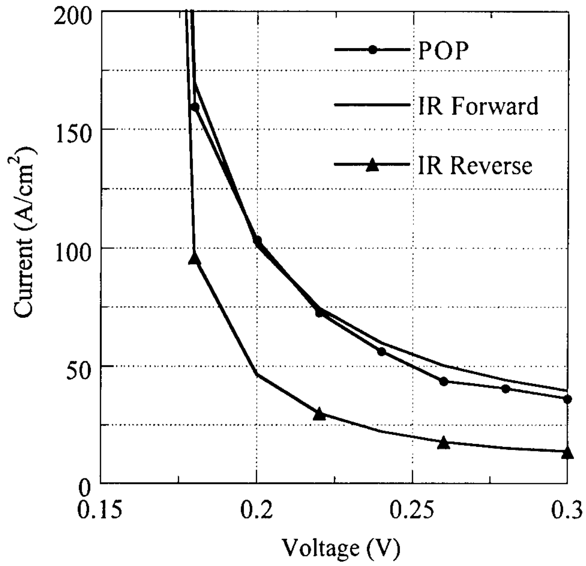


Figure 2. I–V characteristic of the same device as in figure 1 but with interface roughness on alternate interfaces.

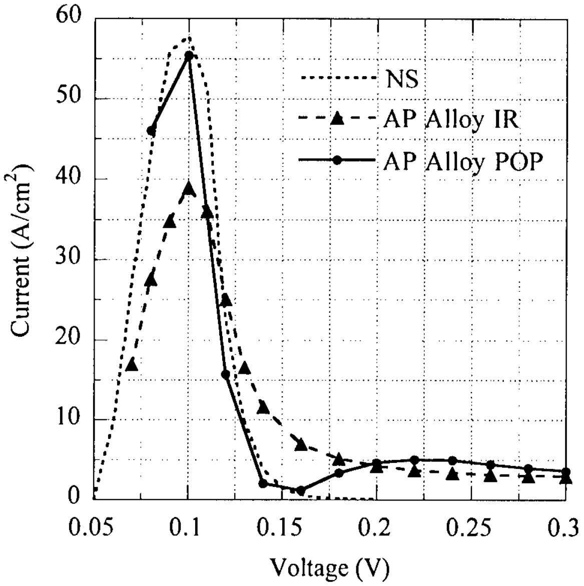


Figure 3. I–V characteristic at 4.2 K which exhibits a phonon peak. The RTD has 10^{18} cm^{-3} n$^+$ $In_{0.53}Ga_{0.47}As$ leads, 50 nm intrinsic $In_{0.53}Ga_{0.47}As$ spacer layers, a 2.9 nm AlAs emitter barrier, a 2 nm/2 nm/2 nm $In_{0.53}Ga_{0.47}As/InAs/In_{0.53}Ga_{0.47}As$ well and a 2.3 nm collector barrier.

The curve labelled 'AP Alloy POP' includes scattering from acoustic phonons, alloy disorder and polar optical phonons. These calculations indicate that the contribution from interface roughness scattering and polar optical scattering are comparable. The curve labelled 'AP Alloy IR POP' includes scattering from all four scattering mechanisms.

In figure 2 we investigate the effects of asymmetric roughness [9]. The curve labelled 'POP' is replotted from figure 1 where it is labelled 'AP Alloy POP'. For the curve labelled 'IR Forward', the first, third and fifth interfaces counting from the emitter are rough and the other interfaces are smooth. For the curve labelled 'IR Reverse', the rough and smooth interfaces are interchanged. For the 'IR Reverse' curve, the excess current is scattered from the rough InAs interface in the well since scattering from the second and sixth interfaces is negligible [9]. A comparison of the 'IR Reverse' and 'IR Forward' curves indicates that the scattering from the first AlAs interface is roughly twice as large as the scattering from the InAs interface. Knowing the relative contributions from the InAs and AlAs interfaces and taking into account the difference in material parameters, we can assert that our result for the $In_{0.53}Ga_{0.47}As/AlAs/InAs$ RTD is consistent with our result for a GaAs/AlAs RTD [9].

A device that displayed the phonon peak would be useful for assessing the theory. The calculated I–V of such a device at 4.2 K is shown in figure 3. The three curves show the coherent tunnelling calculation labelled 'NS', the calculation including acoustic phonons, alloy disorder and interface roughness labelled 'AP Alloy IR', and the calculation including acoustic phonons, alloy disorder and polar optical phonons labelled 'AP Alloy POP'. We note that (i) the reduction of the peak current in the presence of interface roughness scattering is a numerical artefact of the finite truncation of the scattering series and the lack of full charge self-consistency, and (ii) these approximations have little impact on the calculation of the valley current which is our concern [8, 12]. The phonon peak should be clearly visible in such a structure.

In summary we find the relative contributions of interface roughness scattering and polar optical phonon scattering to the valley current of $In_{0.53}Ga_{0.47}As/AlAs/InAs$ RTDs to be comparable, which is consistent with results for a GaAs/AlAs RTD [9]. For the GaAs/AlAs RTD, the phonon peak is clearly visible. Figure 3 shows that the phonon peak should be observable for $In_{0.53}Ga_{0.47}As/AlAs/InAs$ RTDs as well. Experimental data [5–7] are required to assess the theory.

References

[1] Bowen R C, Klimeck G, Lake R, Frensley W and Moise T S 1997 *J. Appl. Phys.* **81** 3207

[2] Klimeck G, Boykin T, Bowen R C, Lake R, Blanks D, Moise T S, Kao Y-C and Frensley W 1997 *55th Annual Device Research Conference Digest* (Piscataway, NJ: IEEE) p 92

[3] Moise T S, Kao Y-C, Katz A J, Broekaert T P E and Celli F G 1995 *J. Appl. Phys.* **78** 6305

[4] Roblin P, Potter R C and Fathimulla A 1996 *J. Appl. Phys.* **79** 2502

[5] Zaslavsky A, Goldman J V, Tsui D C and Cunningham J E 1988 *Appl. Phys. Lett.* **53** 1408

[6] Alves E S, Eaves L, Henini M, Hughes O H, Leadbeater M L, Sheard F W, Toombs G A, Hill G and Pate M A 1988 *Electron. Lett.* **24** 1190

[7] Turley P J and Teitsworth S W 1994 *Phys. Rev.* B **50** 8423

[8] Lake R, Klimeck G, Bowen R C and Jovanovic D 1997 *J. Appl. Phys.* **81** 7845

[9] Lake R, Klimeck G, Bowen R C, Fernando C, Leng M, Moise T S and Kao Y-C 1996 *Superlatt. Microstruct.* **20** 279

[10] Tsen K T 1993 *Int. J. Mod. Phys.* B **7** 4165

[11] Mori N, Taniguchi K and Hamaguchi C 1992 *Semicond. Sci. Technol.* **7** B83

[12] Johansson P 1993 *Phys. Rev.* B **48** 8938

Semicond. Sci. Technol. **13** (1998) A165–A168. Printed in the UK PII: S0268-1242(98)91844-2

Numerical approximations to the treatment of interface roughness scattering in resonant tunnelling diodes

G Klimeck†, R Lake and D K Blanks

Raytheon TI Systems, Dallas, TX 75243, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. Current calculations for a resonant tunnelling diode which include a self-consistent Born treatment of interface roughness combined with full charge self-consistency are presented. Approximate but more computationally tractable solutions which either ignore the scattering-assisted charge or the real part of the self-energy are compared with the full solution. The coherent tunnelling calculation combined with quantum charge self-consistency provides an accurate estimate of the resonant current. To obtain the same accuracy while including interface roughness scattering for our particular device requires a treatment of both the real and the imaginary parts of the self-energy combined with a fully self-consistent treatment of the scattering-assisted charge.

1. Introduction

Electron transport through a resonant tunnelling diode (RTD) is a function of the electron charging, bandstructure, and scattering effects. The relative importance of these physical effects depends on the device design, device composition and the temperature. To model these effects for layered heterostructures, we developed a comprehensive device simulator [1–7] called NEMO (nanoelectronic modelling), which addresses the design and analysis needs of device engineers and physicists. The underlying theory is documented in detail in [5].

In this paper, we use NEMO to study numerical approximations to the treatment of incoherent elastic scattering resulting from interface roughness (IR) with varying average island sizes of $\Lambda = 0, 5, 10, 40, 100$ nm. An increasing IR average island size implies an increasing strength of the scattering self-energies for small transverse momenta [5, 6]. An exponential correlation model is used [5, 6]. Incoherent averaging is appropriate when the dephasing length exceeds the average island size. The 100 nm island size is pushing the validity of this approach and we are using it to test the numerical convergence properties of our algorithms. Scattering is treated with self-energies calculated in the self-consistent Born approximation using the non-equilibrium Green's function formalism. The self-energies have two effects: they both shift and broaden the resonance spectrum in the RTD and they inject electrons into unoccupied states. Combined, both effects change the electron density in the well of the RTD, which in turn modifies the electrostatic and exchange-correlation [8] potentials. A charge self-consistent treatment is required if the scattering is strong enough. Previously, we presented a finite-order treatment of IR and polar-optical phonon scattering with a fully self-consistent treatment of the charge [2]. Here we present the first self-consistent Born treatment of IR scattering with a fully self-consistent treatment of the charge. We refer to simulations which include scattering-assisted charge in the potential as 'fully charge self-consistent'.

A fully self-consistent treatment is numerically expensive since the scattering must be calculated multiple times in the iterative potential scheme. To reduce this computational load, we can calculate the potential without scattering and then perform a single-pass scattering calculation using the no-scattering-based potential. We refer to such a model as 'partially charge self-consistent'.

Other approaches [9, 10] have modelled scattering with an optical potential of the type $i\hbar/2\tau$ on the diagonal of the Hamiltonian characterized by a relaxation time τ. The physical picture to keep in mind here is that of a damped oscillator. Friction in this oscillator is not only going to widen the resonance spectrum ($\mathrm{Im}\,(\Sigma^R) \approx \hbar/2\tau$) but also shift the centre frequency ($\mathrm{Re}\,(\Sigma^R)$). Ignoring the real part of the self-energy is a numerically appealing since it reduces the shift in the resonance and therefore increases the rate of convergence of the self-consistent Born calculation. In this paper we examine the effects of $\mathrm{Re}\,(\Sigma^R)$ on the I–V characteristic of an RTD with both partial and full charge self-consistency.

---

† Present address: Jet Propulsion Laboratory, Pasadena CA, 911809 8099, USA. E-mail address: gekco@jpl.nasa.gov

We find that the fully charge self-consistent treatment, including the altered resonance spectrum and scattering-assisted charge, closely reproduces the resonant current calculated from a coherent tunnelling model. Since the full calculation requires approximately 2000 times the CPU time of the coherent tunnelling calculation, it is tempting to use approximations such as partial charge self-consistency or to ignore the real part of the self-energy. However, such approximations can lead to spurious shapes of the resonant current.

2. Partial charge self-consistent simulations

The simulated device is a nominally symmetric GaAs/AlAs RTD consisting of 3.1 nm barriers and a 6.2 nm well. The RTD is clad with 20 nm undoped spacers and highly doped (1×10^{18} cm^{-3}) contact layers. The first and the third heterostructure interfaces are assigned to have IR [6] (see the inset of figure 1(*a*)). All simulations are performed at 4.2 K to avoid Γ–X–Γ tunnelling in the AlAs barriers.

First, we compare the inclusion and exclusion of the real part of the scattering self-energies for partial charge self-consistent simulations. If the real part of the scattering self-energy is included (see figure 1(*a*)), the I–V peak becomes more distorted for increasing scattering strengths. Such distortions of the main I–V peak have been observed by other authors as well [11, 12]. We will show in section 3 that these distortions disappear when a full charge self-consistent calculation is performed.

One numerical approximation applied in earlier work [7, 9, 10, 13] is the neglect of the real part of the self-energies. This approximation improves numerical convergence since the resonances do not move as strongly in energy. For a single polar optical phonon scattering event with partial self-consistency, this approximation actually improves the agreement between simulations and experiment [7] by removing a large distortion of the peak simular to the one shown in figure 1. For a self-consistent Born treatment of IR, neglect of the real part of the self-energy always degrades the results. Figure 1(*b*) shows that the distortion of the shape of the current peak is altered and the peak amplitude is reduced if the real part of the scattering self-energy is excluded from the simulation. The distortion of the peak shape is shown for a single IR correlation length of Λ = 40 nm in figure 2 in comparison with experimental data.

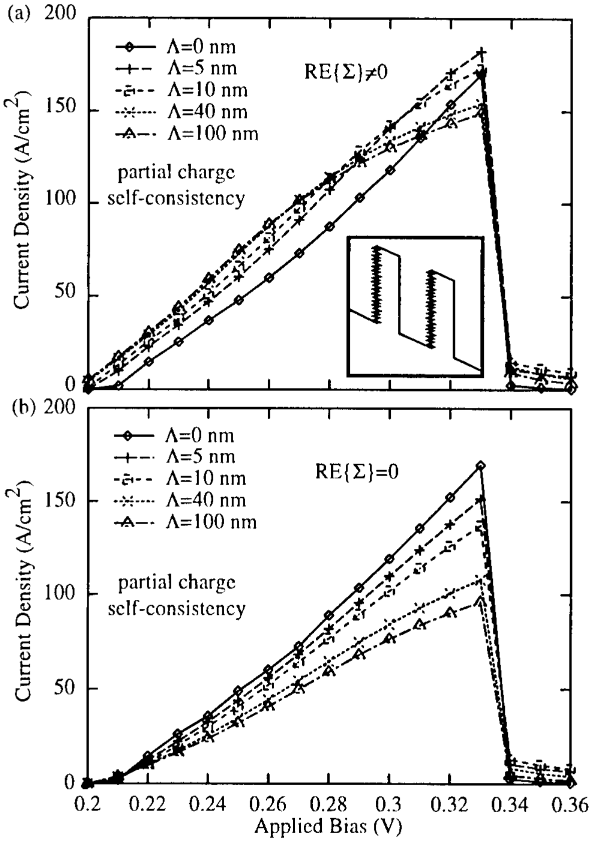


Figure 1. I–V characteristic for four different IR correlation lengths (Λ) computed on a partial charge self-consistent potential (*a*) including and (*b*) excluding the real part of the scattering self-energies. The inset shows the conduction band edge under bias indicating the rough interfaces.

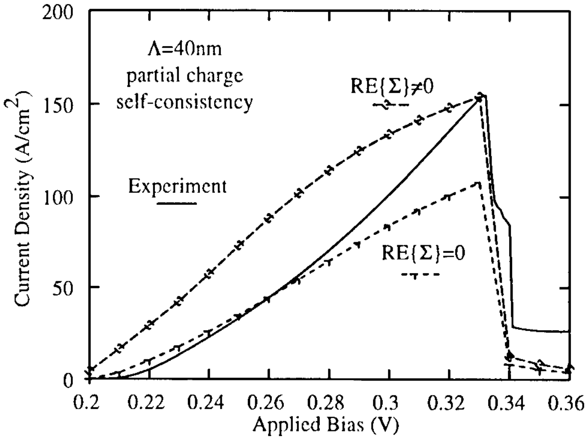


Figure 2. I–V characteristics for a single Λ = 40 nm comparing the effect of the inclusion and exclusion of the real part of the scattering self-energies explicitly. The potential is based on partial charge self-consistency. The I–V peak is significantly distorted if the real part of the scattering self-energies is included. Experimental data in the full line.

3. Full charge self-consistent simulations

The distortions in the shape and amplitude of the I–V characteristics indicate that the resonance spectrum in the RTD is significantly changed. This implies that the charge distribution in the device may be changed significantly as well. We now include the scattering-assisted charge in the charge self-consistent loop and again examine the effects of the exclusion and inclusion of the real part of the scattering self-energies.

If the real part of the scattering self-energies is excluded (see figure 3(*b*)), the peak amplitudes remain (see figure 1(*b*)) strongly reduced with increasing scattering strength. In addition, the I–V turn-offs are shifted to lower voltages.

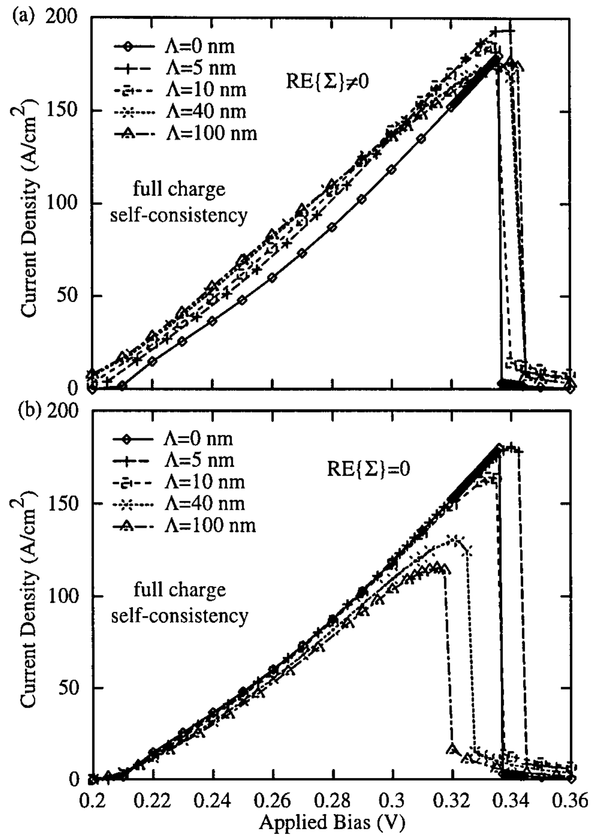


Figure 3. I–V characteristic for four different IR correlation lengths (Λ) computed on a full charge self-consistent potential including the effects of the scattering-assisted charge (*a*) including and (*b*) excluding the real part of the scattering self-energies.

If no approximations are made and the scattering-assisted charge and the real parts of the scattering self-energies are included (see figure 3(*a*)), the I–V curves are no longer distorted and they compare closely with the resonant current calculated with the coherent tunnelling model. Figure 4 compares the $\Lambda = 40$ nm curves from figure 3 with the experimental data.

4. The valley current

Finally, the effects of the real part of the self-energies on the RTD valley current are compared for various IR correlation lengths ($\Lambda = 5, 10, 40, 100$ nm) in the full charge self-consistent simulation. The coherent valley current ($\Lambda = 0$ nm) is for this device 2–3 orders of magnitude lower than the scattering-enhanced valley current and therefore not shown. In the valley current region, the current and the occupation of the well are small. Therefore, the effects of charging are expected to be small. We have indeed verified this and defer this discussion to a separate publication [14].

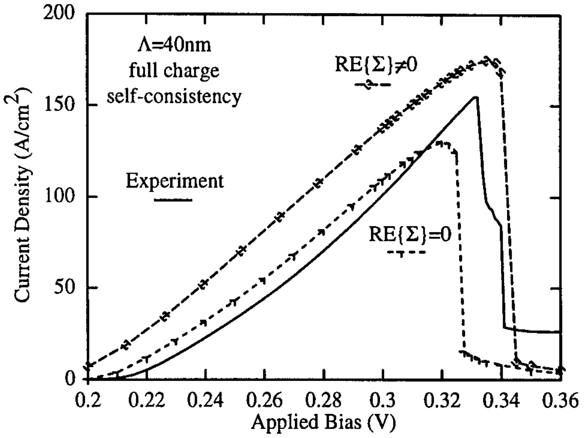


Figure 4. I–V characteristics for a single $\Lambda = 40$ nm comparing the effect of the inclusion or exclusion of the real part of the scattering self-energies. The potential is based on full charge self-consistency, which includes the effects of the scattering-assisted charge. The I–V peak is reduced and the peak voltage is shifted to lower values if the real part of the scattering self-energies is excluded from the simulation. Experimental data in the full line.

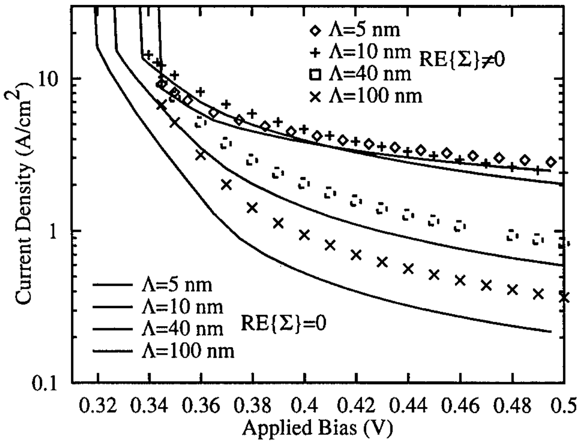


Figure 5. Valley current calculated in full charge self-consistency for four different IR correlation lengths (Λ) with and without the real parts of the scattering self-energies.

However, we have seen in section 3 that the real part of the scattering self-energies has a strong effect on the peak amplitude and peak voltage of the RTD. Figure 5 compares the valley current computed with and without the real part of the scattering self-energies. The simulations without the real part appear shifted to lower voltages from the corresponding simulations with the real part. This shift to lower voltages was already evident in figure 3(*b*). In this presentation we are not intending to simulate fully the valley current. It is shown in [6, 7] that the valley current can be properly modelled with the inclusion of polar optical phonon and IR scattering.

5. Summary

The effects of IR scattering have been computed in the self-consistent Born algorithm for various IR correlation lengths, charging models and self-energy approximations. For small average island sizes, $\Lambda \leq 5$ nm, the peak current is relatively unaffected by the approximation used. For larger average island sizes, a fully charge self-consistent solution combined with both the real and the imaginary parts of the scattering self-energy are required for our device to reproduce the coherent tunnelling calculation of the resonant current. Any approximations lead to spurious distortion of the resonant current, while the off-resonant current is relatively unaffected.

Acknowledgments

We would like to thank R Chris Bowen, Chenjing L Fernando, Manhua Leng and William R Frensley for their work during the NEMO software project. We would like to thank Ted Moise and Yung-Chung Kao for the experimental data.

References

[1] Klimeck G *et al* 1995 *53rd Ann. Device Research Conf. Digest* (New York: IEEE) p 52

[2] Lake R *et al* 1996 *54th Ann. Device Research Conf. Digest* (New York: IEEE) p 174

[3] Klimeck G *et al* 1997 *55th Ann. Device Research Conf. Digest* (New York: IEEE) p 92

[4] Klimeck G *et al* 1995 *Appl. Phys. Lett.* **67** 2539

[5] Lake R, Klimeck G, Bowen R C and Jovanovic D 1997 *J. Appl. Phys.* **81** 7845

[6] Lake R *et al* 1996 *Superlattices Microstruct.* **20** 279

[7] Klimeck G *et al* 1997 *Quantum Devices and Circuits* ed K Ismail, S Bandyopadhyay and J P Leburton (London: Imperial College Press) pp 154–9

[8] Gawlinski E, Dzurak T and Tahir-Kheli R A 1992 *J. Appl. Phys.* **72** 3562

[9] Zohta Y and Tanamoto T 1983 *J. Appl. Phys.* **74** 6996

[10] Sun J P and Haddad G I 1998 *Int. Workshop for Computational Electronics (Tempe, AZ, 1995) VLSI Des.* to be published

[11] Roblin P, Potter R C and Fathimulla A 1996 *J. Appl. Phys.* **79** 2502

[12] Wang S-J *et al* 1996 *Japan. J. Appl. Phys.* **35** 3858

[13] Lake R and Datta S 1992 *Phys. Rev.* B **45** 6670

[14] Klimeck G, Lake R and Blanks D K *Phys. Rev.* B, submitted

Semicond. Sci. Technol. **13** (1998) A169–A172. Printed in the UK

PII: S0268-1242(98)91840-5

Butt-coupling loss of 0.1 dB/interface in InP/InGaAs multiple-quantum-well waveguide–waveguide structures grown by selective area chemical beam epitaxy

C A Verschuren†‡, P J Harmsma†§, Y S Oei†§, M R Leys†‡, H Vonk†‡ and J H Wolter†‡

† COBRA Inter-University Research Institute, PO Box 513, 5600 MB Eindhoven, The Netherlands
‡ Eindhoven University of Technology, Department of Physics, Eindhoven, The Netherlands
§ Delft University of Technology, Department of Electrical Engineering, Delft, The Netherlands

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. The lateral coupling of waveguiding structures in both [011] and $[0\bar{1}1]$ directions is studied using embedded selective area epitaxy by chemical beam epitaxy. All growth steps are carried out under the same growth conditions on (100) InP substrates misoriented by 0.5° towards (111)B. Both planar and selectively grown material exhibits bright luminescence and narrow PL linewidths (8 meV FWHM at 4 K), up to the lateral junction. Moreover, no degradation of the original material properties is observed after regrowth. SEM images show very flat layers and excellent lateral coupling for all four types of junctions. After reactive ion etching of waveguide ridges, the optical coupling losses have been determined using a Fabry–Pérot setup at 1530 nm (TE polarization). Values of 0.1 dB/interface with excellent uniformity are presented. From our results we conclude that, by optimization of the sample preparation prior to regrowth, values of 0.1 dB/interface can be obtained reproducibly for both perpendicular coupling directions.

1. Introduction

Monolithic integration of InP-based opto-electronic devices operating at 1.55 μm is increasingly important for applications such as WDM networks. For the best overall performance the different devices must be optimized individually and the coupling of the optical signal from one device to another must be as efficient as possible. An attractive way for the production of such photonic integrated circuits is embedded selective area epitaxy. By selective area growth in etched recesses, different vertically aligned structures with lateral (butt) couplings can be realized, combining freedom in device design with an optimal overlap of the optical modes. In practice, the growth inside a recess is quite complex, especially near the side walls, where different crystallographic planes are exposed simultaneously and diffusion and shadowing effects influence the growth behaviour. This may lead to air gaps, local composition changes and severe growth rate changes, which can degrade the coupling efficiency.

In the last few years some groups obtained promising results using low-pressure metal–organic vapour-phase epitaxy [1] and chemical beam epitaxy (CBE) [2, 3]. The latter CBE results are obtained by using a perpendicular group III injector instead of the usual tilted geometry. In this way shadowing effects are nearly eliminated, which is advantageous for the realization of good lateral couplings. Unfortunately, in most CBE systems, as in our Riber 32P, the perpendicular geometry is impossible. However, we find that the growth behaviour can be much improved by using a low substrate misorientation with only B type steps.

Here, we present our most recent results in InP/InGaAs multiple-quantum-well (MQW) waveguide–waveguide structures grown on substrates with a misorientation of 0.5° → (111)B. We have achieved very good coupling results in both perpendicular crystallographic directions.

0268-1242/98/SA0169+04$19.50

2. Experimental procedure

In the first epitaxy step MQW waveguide structures are grown by CBE in a Riber 32P system on 2 in (100) InP substrates, 0.5° misoriented towards (111)B. The waveguiding (WG) core typically consists of 60 periods of 2–3 nm lattice-matched InGaAs wells and 6 nm InP barriers, resulting in a room-temperature photoluminescence at 1350–1400 nm. The top cladding layer of InP is 300 nm. Details of the growth can be found in [4]. After deposition of 100 nm SiN_x, series of windows are defined by lithography on quarter wafers. The mask pattern, illustrated in figure 1, is aligned in either the [011] or the $[0\bar{1}1]$ direction. It contains three groups of windows, each with five series of three identical windows, with lengths ranging from 100 to 900 μm, in steps of 200 μm. With this mask, recesses are etched by reactive ion etching (RIE), through the WG core and about 200 nm into the InP buffer. The final step in sample preparation is a brief etch in diluted $\mathrm{Br_2}$–$\mathrm{CH_3OH}$. This gives a slight underetch below the SiN_x mask which, if properly applied, prevents the formation of so-called 'ears'. In the second growth step, the recesses are selectively filled with a similar WG structure using identical CBE growth conditions as for the original structure. Following the regrowth, 10 mm long waveguide stripes of 2, 3 and 4 μm width are defined in SiN_x, each running through one regrown recess, as indicated by the broken lines in figure 1. Near every series of regrown windows stripes are defined which only run through the original WG material, to serve as reference. The stripes are then used to etch shallow WG ridges, 100 nm into the core.

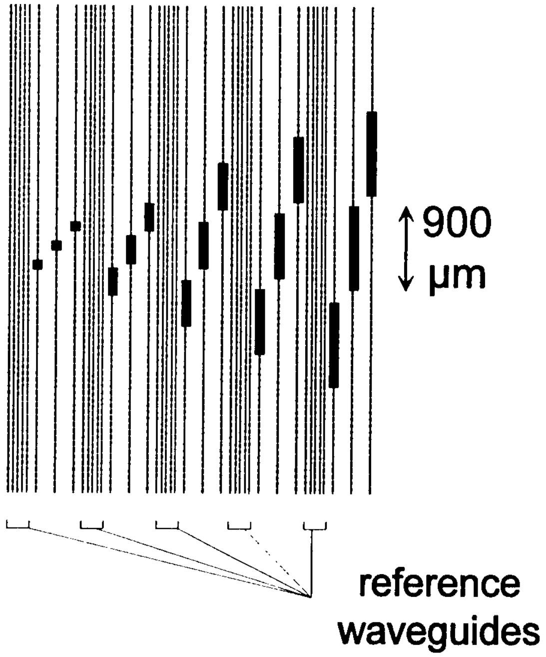


Figure 1. Schematic mask layout. The rectangles represent the windows–recesses for the selective regrowth and the broken lines indicate the waveguide stripes for transmission measurements. This layout is repeated three times, with waveguide widths of 2, 3 and 4 μm, respectively.

3. Results and discussion

3.1. Waveguide substrates

Our 60 period MQW waveguide structures show bright luminescence and a high material quality, as indicated by for example the 4 K PL linewidth of 7–9 meV and mirror-like morphology with low defect density. For the (original) waveguide substrates we use for the integration experiments, the PL peak positions at room temperature, λ_{RT}, and the 4 K linewidths are listed in table 1.

3.2. Integration results

Characterization of both regrown material and coupling quality is important, since it provides information for the optimization of growth conditions and sample processing. The first method we use is Nomarski or phase contrast microscopy for the visual inspection of the surface. The morphology and the number of defects indicate mainly the success of the sample preparation and are a first check of the growth process.We find that the morphology in regrown areas is as smooth as on planar substrates, while the defect density is only slightly higher.

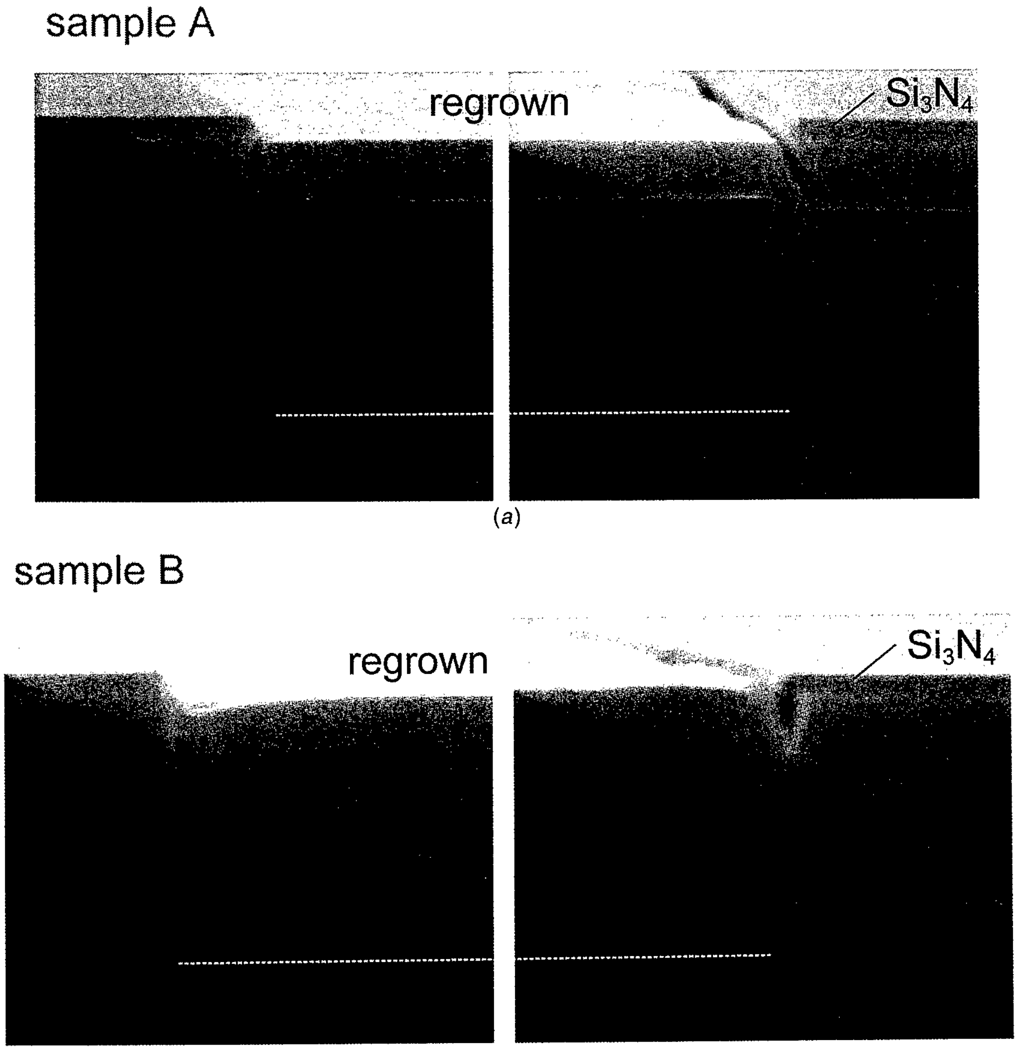
Information on the regrowth and coupling quality is obtained by examining cross-sections using scanning electron microscopy (SEM). After cleavage, the samples are stain etched using $\mathrm{K_3Fe(CN)_6}$–KOH to enhance the material contrast. Figure 2 shows the cross-sections (see table 1 for sample details). It is clear that in all cases the vertical alignment and the flatness of the layers is excellent. The absence of air gaps and a transition over only about 0.1 μm suggest low coupling loss at the lateral interfaces. The coupling in figure 2(*a*) (along $[0\bar{1}1]$) looks somewhat better than the coupling in figure 2(*b*) (along [011]). In our opinion this is because the combination of fast surface diffusion and chemically active B steps in the $[0\bar{1}1]$ direction helps to compensate shadowing effects at the recess side walls [5]. Notwithstanding the anisotropic diffusion, the results in figure 2(*b*) look surprisingly good.

In order to check both original and regrown waveguide material properties after regrowth, we have used a spatially resolved PL (SRPL) setup with an Ar^+ laser focused to a spot diameter of about 5 μm and a He flow cryostat which can be moved by two stepper motors with 0.1 μm step resolution. For the original waveguide we find no degradation or shift of the optical properties after regrowth. Both PL position and linewidth are the same within the experimental error of about 1 nm. The regrown material has PL intensities and linewidths which are uniform over the wafer and are comparable with those for material grown on 2 in epiready substrates (see table 1).

The same SRPL setup is also used to quantify possible optical non-uniformities near the recess edges. With a scan along edges in the B type step propagation direction ($[0\bar{1}1]$), we find a red shift of about 12 nm, over a distance which extends more than 25–30 μm from the edge onto the regrown area. This distance corresponds to the surface corrugation that is always observed at this edge [5, 6]. The

Table 1. Sample description.

Sample	InP top (nm)	Original waveguide λ_{RT} (nm)	Original waveguide 4 K FWHM (meV)	Regrown waveguide λ_{RT} (nm)	Regrown waveguide 4 K FWHM (meV)	Ridge direction	Coupling loss (dB/interface)
A	300	1435	8	1400	9	$[0\bar{1}1]$	0.3 ± 0.1
B	300	1415	9	1400	9	[011]	$< 0.1 \pm 0.04$



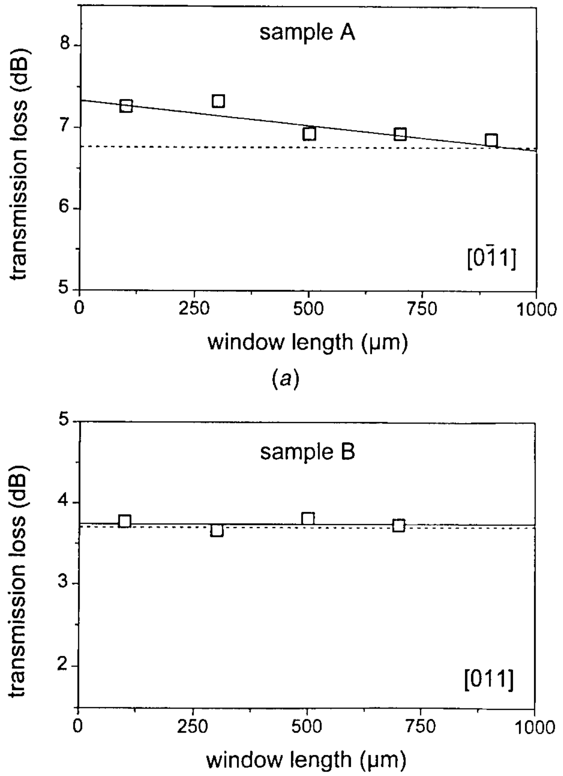
(*a*)

(*b*)

Figure 2. SEM cross-sections showing the lateral coupling for (*a*) sample A and (*b*) sample B. The broken lines show the bottom of the etched recess; the scale bars indicate 1 μm.

red shift may be attributed to a locally enhanced growth rate at the macroscopic surface steps. The spectra for the three remaining edges, $[01\bar{1}]$, [011] and $[0\bar{1}\bar{1}]$, are nearly identical. Analysis reveals a red shift of about 12 nm, but here it occurs within less than 1 μm from the edge. As indicated by high-magnification SEM images, tiny (111) facets are formed near the recess side walls, which are known to give rise to a red shift [7, 8]. We conclude that the actual extent of the red shift at these edges is about 0.1 μm or less.

We have determined the transmission loss for all waveguide ridges using a Fabry–Pérot setup at 1530 nm (TE polarization). Extrapolation of the measured transmission losses to a regrown length of zero and substraction of the reference's transmission loss yields the loss due to the lateral interfaces only. The results for all samples are plotted in figure 3. The small scatter of the data around the curves illustrates the excellent uniformity of the lateral couplings. For sample A the transmission loss is reduced as the regrown length increases, which indicates



(*a*)

(*b*)

Figure 3. Fabry–Pérot transmission measurements versus the length of a regrown window for 3 μm WG ridges. The broken lines represent the reference waveguide data.

that the regrown material has a loss smaller than that of the original material. This is caused by the different PL positions: the original material with $\lambda_{RT} = 1430$ nm gives rise to significant absorption at 1530 nm, while the regrown material with $\lambda_{RT} = 1390$ nm is more transparent. It is clear that, for negligible absorption effects, λ_{RT} must be well below 1400 nm.

From the data in figure 3(*a*) we obtain a coupling loss for sample A of about 0.6 dB for two lateral interfaces ($=2\epsilon_i$) along $[0\bar{1}1]$. Since no irregularities are observed in the SEM cross-sections (figure 2(*a*)) part of this loss is attributed to the observed corrugation-related red shift over a distance of more than 25 μm. Thus, it should be possible to reduce this coupling loss by growing material with a lower λ_{RT}. Figure 3(*b*) shows equal values for both reference and partly regrown waveguides. Since no dependence on the regrown length is observed, the coupling loss is in fact determined by the standard deviation of the measurements, resulting in a *maximum* value as low as $\epsilon_i = 0.1 \pm 0.04$ dB/interface for sample B. These results are clearly not optimal. However, further experiments [4] indicate that coupling losses around 0.1 dB/interface can be reproducibly achieved in both crystallographic directions.

4. Conclusions

We have investigated the lateral integration of waveguiding structures by embedded selective area CBE using (100) InP 0.5° → (111)B substrates. The second growth step is carried out under identical growth conditions to those of the first, planar growth step. PL characterization shows no degradation or shift for the original material after regrowth. Furthermore, the morphological and PL quality of the regrown material is as good as on normal substrates, up to the lateral junction. SEM and Fabry–Pérot measurements reveal very good coupling results with excellent uniformity in both perpendicular crystallographic directions. At the $[0\bar{1}1]$ junction a small PL red shift is observed for the regrown structures. To eliminate resulting absorption effects in regrown passive structures, λ_{RT} must be below ~1380 nm. For planar regrowth surfaces, growth irregularities or 'ears' at the recess edge must be prevented. This can be done by carefully underetching the RIE recess prior to the regrowth. Loss values around 0.1 dB/interface can be realized for lateral coupling along both [011] and $[0\bar{1}1]$, which is extremely attractive for applications in the integration of photonic devices.

Acknowledgments

This work was financially supported by the Dutch Ministry of Economic Affairs (IOP) and the ACTS programme of the European Union. Special thanks are due to Geraldine O'Donovan for carrying out the SRPL measurements.

References

[1] Ahn J-H, Oh K R, Kim J S, Lee S W, Kim H M, Pyun K E and Park H M 1996 *IEEE Photon. Technol. Lett.* **8** 200
[2] Legay P, Alexandre F, Benchimol J L, Allovon M, Laune F and Fouchet S 1996 *J. Cryst. Growth* **164** 314
[3] Wachter M, Schöffel U, Schier M and Heinecke H 1997 *J. Cryst. Growth* **175–176** 1186
[4] Verschuren C A, Harmsma P J, Oei Y S, Leys M R, Vonk H and Wolter J H *J. Cryst. Growth* accepted for publication
[5] Verschuren C A, Leys M R, Oei Y S, Vreeburg C G M, Vonk H, Rongen R T H and Wolter J H 1997 *J. Cryst. Growth* **170** 650
[6] Matz R, Heinecke H, Baur B, Primig R and Cremer C 1993 *J. Cryst. Growth* **127** 230
[7] Heinecke H, Milde A, Matz R, Baur B and Primig R 1994 *Phys. Scr.* T **55** 14
[8] Chiu T H, Chen Y, Zucker J E, Marshall J L and Shunk S 1993 *J. Cryst. Growth* **127** 169

Semicond. Sci. Technol. **13** (1998) A173–A176. Printed in the UK PII: S0268-1242(98)91841-7

Mesh-based particle simulation of sub-0.1 micron FETs

C Arokianathan†, A Asenov‡ and J Davies

Department of Electronics and Electrical Engineering, Glasgow University,
Glasgow G12 8LT, UK

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. When the devices in modern integrated circuits shrink to dimensions below 100 nm the discrete nature of the charge of impurities and carriers starts to exert its influence. In this paper we describe a parallel mesh-particle approach to the simulation of sub-100 nm FETs when the impurities and the carriers are treated as individual charges. The mesh-based solution of the Poisson equation has advantages when devices with complex shapes and boundaries have to be simulated, but introduces problems related to the convergence of the parallel solvers and in resolving the short-range interparticle forces. We found that the parallel multigrid solvers have superior convergence properties for discrete charge simulations, uniformly damping the short- and long-range components of the residuals. We developed an economical approach to analytically correct the erroneous short-range component of the interparticle forces derived from the mesh-based solution. The mesh-particle approach is illustrated and verified in the simulation of a dual-gate MESFET with a channel length of 80 nm.

1. Introduction

At the end of the revised Silicon Roadmap in the year 2012 the feature size of integrated MOSFETs in commercial integrated circuits will be below 50 nm [1]. However, devices with a channel length below 50 nm have already been fabricated and investigated in leading research laboratories [2]. It is widely accepted [3, 4] that the discrete nature of the charge on electrons and impurities will have a significant impact on the characteristics of these devices and will become important limiting factors to device scaling and integration. The random fluctuations of the relatively small number of dopants and their discrete microscopic arrangement in the channel and in the source/drain regions of a sub-100 nm MOSFET lead to an average decrease and spread in the threshold voltage, and degradation and fluctuations in the subthreshold slope [5, 6]. The particulate nature of the small number of carriers is also expected to influence the behaviour of such mesoscopic devices. Trapping and detrapping of individual carriers can produce appreciable random telegraph noise [7], the importance of which increases as the area of a device falls. Statistical fluctuations in the number of electrons will enhance the shot noise.

There is no general agreement upon the techniques that should be employed to simulate and optimize devices in the transition from semiclassical (100 nm) to pure quantum mechanical (10 nm) limits, and to account for all the above types of fluctuations. In principle, quantum statistical mechanics is the only method to embrace this range, but it is not clear that satisfactory techniques for solving the equations yet exist, and the scale of the simulations is wholly impractical. There are, fortunately, indications that the well-established approaches to simulation based on partial differential equations (drift–diffusion, hydrodynamic) [6], and stochastic techniques based on classical particle trajectories, can provide useful insight and information [8].

In this paper we discuss different aspects of mesh-based 'atomistic' particle simulations of sub-0.1 μm FETs. To account for fluctuations in doping between devices and the small number of carriers, both the ionized impurities and the carriers are treated as discrete charges. This imposes severe demands on the Poisson solver when combined with the complicated geometries of a real device. The mesh-based solution of the Poisson equation has advantages, compared with the direct analytical calculation of the interparticle Coulomb forces [9], when devices with complex geometry and contacts must be simulated. At the same time it introduces problems associated with the convergence of most iterative solvers and with resolving the short-range interparticle forces. The performance and efficiency of parallel coloured SOR, BiCGSTAB(2) and multigrid solvers are compared for complex potential landscapes resulting from point-like charges. Efficient analytical corrections are introduced for the short-range part of the Coulomb interaction estimated from the mesh solution. The mesh-based approach is illustrated and validated in a Brownian type [10] simulation of a 0.08 μm dual-gate MESFET.

---

† Present address: Roke Manor Research Ltd, Roke Manor, Romsey, Hampshire SO51 0ZN, UK.

‡ Author to whom correspondence should be addressed. E-mail address: a.asenov@elec.gla.ac.uk

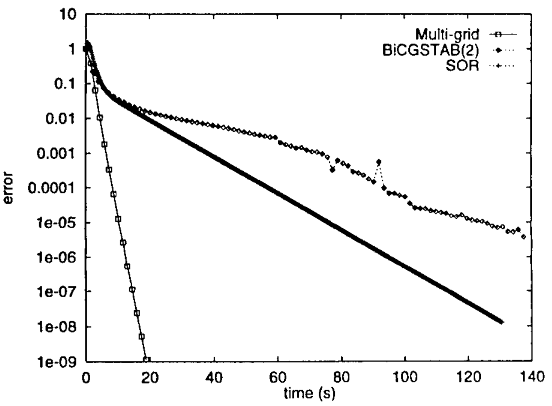


Figure 1. Error as a function of the CPU time for a parallel coloured SOR, BiCGSTAB(2) and multigrid solvers for a cubic 65^3 grid with uniform dielectric constant and randomly distributed impurities.

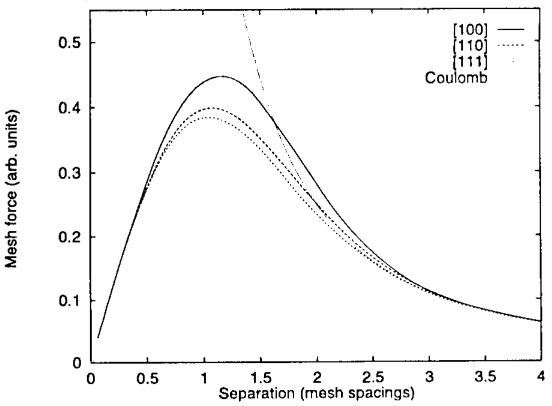


Figure 2. The force calculated from the mesh solution of the Poisson equation when the line joining the two charges lies on the [100], [110] and [111] directions.

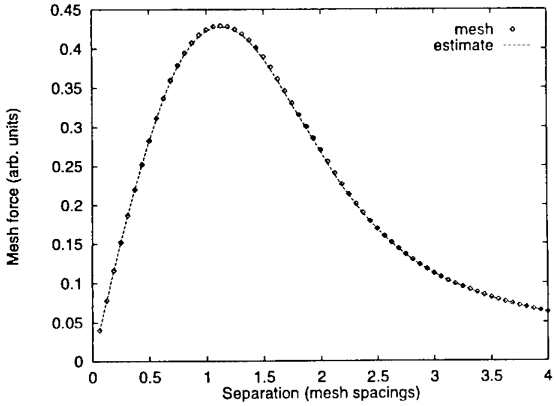


Figure 3. Test for the analytical correction of the mesh force for the [411] direction.

2. Computational aspects

The correct treatment of discrete charges and their interaction in 'atomistic' mesh-based particle device simulations requires a 3D solution of the Poisson equation with fine-grain discretization to properly resolve the short-range interparticle forces. This is important because the ionized impurity scattering has to be included *ab initio* in such simulations. We use spatial domain decomposition techniques to uniformly distribute the solution domain among the mesh connected processors of an eight-node Parsytec Power PC X-Plorer [11] and to carry out the solution in parallel. The sharp spikes in the solution resulting from the discrete nature of the charges on short length scales have an adverse effect on the convergence of most parallel iterative solvers. In figure 1 we compare the performance and efficiency of parallel coloured SOR, BiCGSTAB(2) and multigrid solvers for a cubic 65^3 grid with uniform dielectric constant and randomly distributed impurities. The methods for calculating the cross products were kept as similar as possible to ensure fairness of the comparison. Although converging for fewer iterations than the coloured SOR solver, the BiCGSTAB(2) method stalls for the complex potential landscape resulting from the point-like charges. Reduction of the highly localized errors seems to be problematic since only a small fraction of the nodes are affected, introducing small changes in the vector norms. Due to the much larger computational work per iteration the parallel BiCGSTAB(2) usually consumes more time than the parallel SOR solver. The multigrid method reducing, on different levels, both the high- and low-frequency residual components performs excellently and is the obvious choice.

A four-node differencing scheme was used to calculate the electric field. It produces good agreement with the true Coulomb force for distances larger than three mesh spacings. However, the mesh solution of Poisson's equation, even with a spacing below 1 nm, is band-limited and introduces significant errors in the estimated interparticle forces when the distance between the charges becomes smaller than two to three mesh spacings. The remedy to this problem is to use the mesh solution to obtain the forces introduced by the boundary conditions and the charges more than three mesh spaces away, and to introduce analytical corrections for the charges on shorter distances. This is a complicated task because the analytical correction varies with the relative position of the charges with respect to the grid nodes and is determined by the range and the relative orientation of the charges with space in respect to each other (figure 2). We have developed an inexpensive approach to analytically correct the short-range part of the interaction estimated from the mesh solution. For the correction we use an efficient interpolation scheme which reflects the symmetry of the cubic mesh cell. The electric field is tabulated as a function of the relative displacement r_{ij} between the two charges along each of the three directions of symmetry of the cubic cell and simple polynomial functions are used to interpolate between them to get the mesh force. The first step is to sort the components of r_{ij} by decreasing order of magnitude and label them l_1, l_2 and l_3. The quantity

$$\cos^2\theta = \frac{l_1^2}{r_{ij}^2} \qquad (1)$$

is used both to interpolate between the forces along the [100] and [111] directions and along [100] and [110]

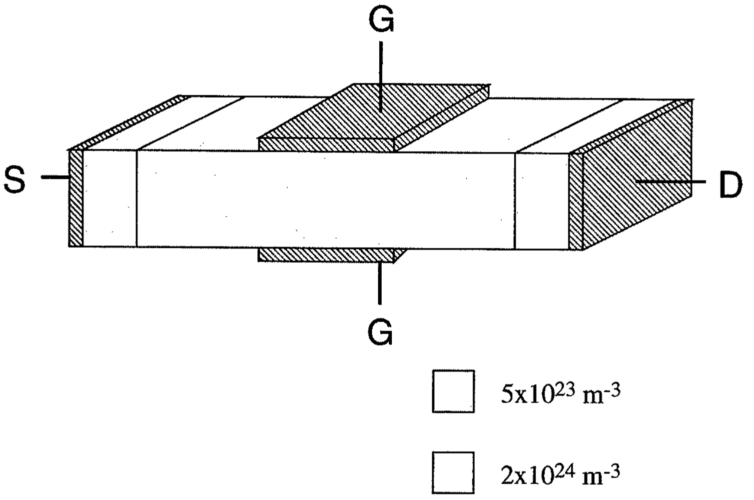


Figure 4. Schematic view of a dual-gate MESFET.

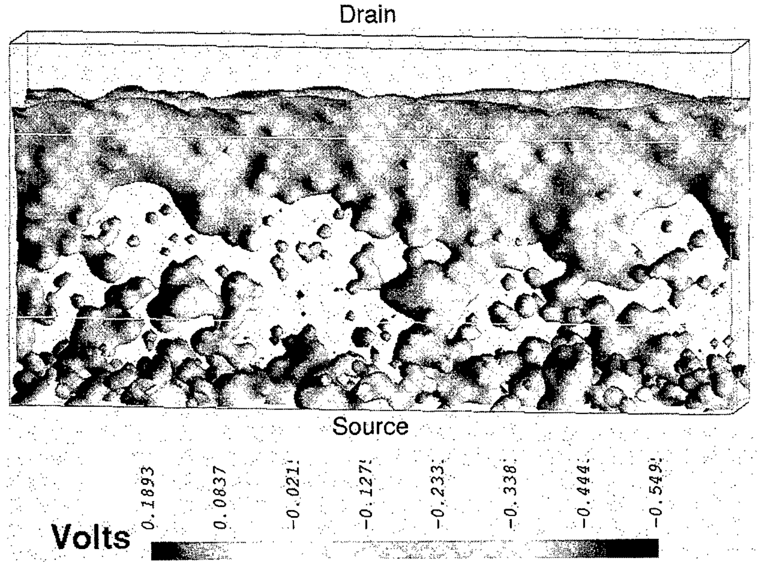


Figure 5. Equipotential surfaces in the channel of 80 nm channel dual-gate GaAs MESFET at threshold.

directions. Finally

$$\sin^2\phi = \frac{l_3^2}{l_2^2 + l_3^2} \tag{2}$$

is used to interpolate the net force from these two quantities. The quality of this correction is illustrated in figure 3.

Accurate integration of the equations of motion is another problem which has to be carefully considered because the magnitude of the Coulomb force varies rapidly near the impurities and is addressed elsewhere [12].

3. Simulation example

The mesh-based solution of the Poisson equation was tested in combination with the Brownian particle method [10] in the simulation of a 80 nm dual-gate MESFET. The device has a channel doping concentration of $N_D = 5 \times 10^{23}$ m^{-3}. The schematic view of the MESFET is given in figure 4. A $129 \times 129 \times 33$ grid was used with a grid spacing of 2.5 nm. The carriers were placed at random in the solution domain but their numbers were in accordance with the doping level in the different device regions. Figure 5 illustrates the potential in the channel near threshold; fluctuations due to the random donors are evident. Figure 6 shows that the results of this atomistic simulation agree well with drift-diffusion results from MEDICI, at low source–drain voltage. The slope of the $I_D(V_G)$ characteristic is proportional to the mobility, and shows that impurity scattering is properly included in the Brownian simulation.

4. Conclusions

In this paper we have demonstrated that particle type simulation of sub-100 nm semiconductor devices when both the impurities and carriers are treated as individual charges can be carried out using a mesh-based solution of the Poisson equation. The impurity scattering is introduced

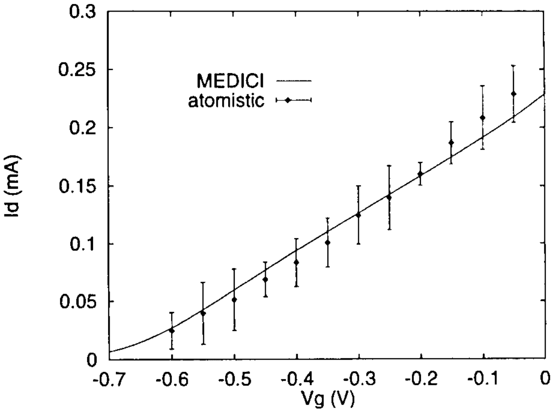


Figure 6. Transfer characteristics of an 80 nm channel dual-gate GaAs MESFET, comparing the 'atomistic' Brownian simulations and drift–diffusion MEDICI simulations at low drain voltage (V_D = 0.05 V).

ab initio in the simulations through the electrostatic potential. We have developed an economical approach for analytical correction of the Coulomb interparticle forces which surmounts the difficulties associated with the band limitations of the mesh-based solution.

References

[1] *The National Technology Roadmap for Semiconductors* revised edn 1997 (San Jose, CA: Semiconductor Industry Association)

[2] Hori A, Nakaoka H, Umimoto H, Yamashita K, Tkese M, Shimitzu N, Mizuno B and Odanka S 1994 *Proc. 1994 IEDM Tech. Dig.* 485

[3] Taur Y *et al* 1997 *Proc. IEEE* **85** 486

[4] Zhou J-R and Ferry D K 1994 *Proc. 3rd Int. Workshop on Computational Electronics* (New York: Plenum) p 74

[5] Mizuno T, Iwase M, Niiyama H, Shibata T, Fujisaki K, Nakasugi T, Toriumi A and Ushiku U 1996 *1996 IEDM Tech. Dig.* 13

[6] Wong H-S and Taur Y 1993 *1993 IEDM Tech. Dig.* 705

[7] Hung K K, Ho P, Hu C and Cheng Y C 1990 *IEEE Electron Device Lett.* **11** 90

[8] Ferry D K, Akis R, Vasileska D, Bird J P and Barker J R 1997 *Proc. 2nd NASA Workshop on Device Modelling, (Moffet Field, NASA)*

[9] Ferry D K, Kriman A M, Kann M J and Joshi R P 1991 *Comput. Phys. Commun.* **67** 119

[10] Arokianathan C R, Asenov A and Davies J H 1996 *J. Appl. Phys.* **80** 1

[11] Asenov A, Barker J R, Brown A R and Lee G L 1996 *J. Simulation Practice Theory* **4** 155

[12] Arokianathan C R, Davies J H and Asenov A 1997 *VLSI Design* at press

Semicond. Sci. Technol. **13** (1998) A177–A179. Printed in the UK
PII: S0268-1242(98)91888-0

Cellular automata simulation of nanometre-scale MOSFETs

M Saraniti†‖, G Zandler‡, G Formicone§, S Wigger† and S Goodnick†

† Electrical Engineering Department, Arizona State University, Tempe,
AZ 85287-5706, USA
‡ Walter Schottky Institut, TU-München, D-85748 Garching, Germany
§ Digital Device Engineering Group, Motorola SPS, Tempe, AZ 85284, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We present systematic theoretical cellular automata studies of vertically grown, nanometre-scale, MOSFETs. The predicted drain characteristics and output conductance are in excellent agreement with experimental data from fabricated devices. The inclusion of an inhomogeneous *p*-doping profiles along the channel is investigated, which is shown to improve current saturation and therefore allows the reduction of the device dimensions.

1. Introduction

As a numerically efficient discrete variant of the Monte Carlo technique [1], the cellular automaton (CA) method [2–4] has been shown to reduce significantly the computational burden required by particle-based simulation approaches. Recently, chemical vapour deposition epitaxy was utilized to grow vertical MOSFETs with channel lengths ranging from 65 to 170 nm [5, 6]. A systematic set of CA calculations is performed, to investigate the influence of nonlinear transport and of doping profile on the electrical behaviour of such ultrashort devices.

2. Device characterization

2.1. Current–voltage curves

Based on the geometry sketched in figure 1, we investigated the drain current saturation behaviour of the vertical MOSFET for different channel lengths. Saturation is found for gate lengths down to 65 nm as can be seen in figure 2. The simulation results are in excellent agreement with the experiment; no tuning of geometrical or physical parameters was needed.

The output current and output conductance increase as the channel length is reduced, reflecting a strong barrier reduction at the source junction. These effects are further enhanced by velocity overshoot in the shorter channels.

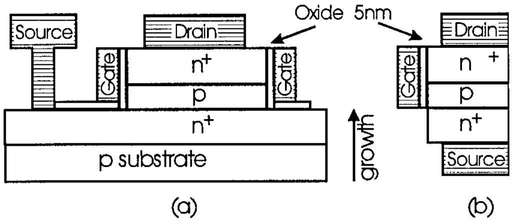


Figure 1. (*a*) Layout of the vertical CVD–EPI MOS [6]. The growth direction is shown. (*b*) The equivalent simulated region.

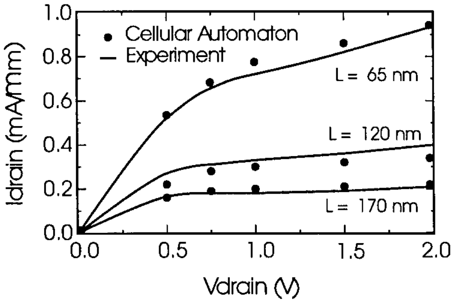


Figure 2. Computed (●) and experimental (——) [5] current–voltage characteristics with decreasing channel length at $V_G = 2$ V gate bias.

For the 65 nm device, simulations were performed up to a bias value of $V_D = 2.5$ V, to investigate the occurrence of impact ionization at higher bias. Up to the highest bias point shown, negligible contributions due to impact ionization were found. However, this process becomes more and more important at higher drain bias. Simulation of the same device performed with commercial software based on the momenta method was unable to reproduce this result when impact ionization was turned on.

2.2. Effects of the doping profile, shorter devices

The saturation behaviour of the simulated devices is clearly affected by bulk parasitic conduction which occurs in the source–drain direction, away from the inversion layer. This parasitic *n*–*p*–*n* structure acts as a barrier for electron current until its spatial extension is larger than the depleted

---

‖ To whom all correspondence should be addressed.

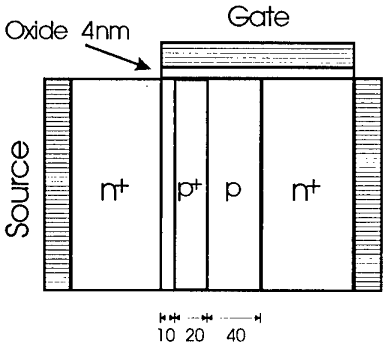


Figure 3. Cross-section of the saturation optimizing geometry. A 3 × 10^{18} cm^{-3} p^+ buffer is included within the central *p* region.

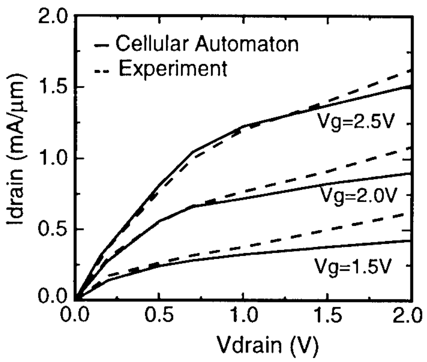


Figure 4. Improved saturation behaviour of the simulated MOSFET after adding an additional 20 nm p^+ buffer.

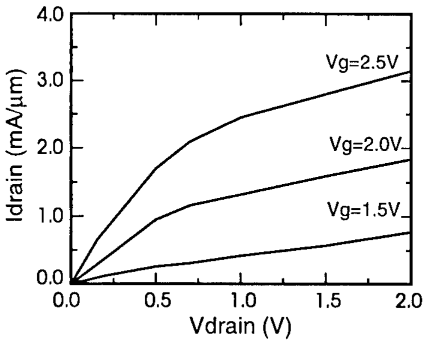


Figure 5. Current–voltage curves for a 50 nm vertical MOS with inhomogeneously doped channel. The doping concentration in the channel is 4 × 10^{18} cm^{-3}, whereas a density of 5 × 10^{18} cm^{-3} acceptors is set in the p^+ buffer.

regions at the *n–p* and *p–n* sides, i.e. when the central *p* region is not depleted.

When the dimension of the *p* region is reduced or, analogously, when the source–drain bias extends the above-mentioned depletion regions, electrons can flow through the *n–p–n* region, and the influence of the gate bias on the output current is negligible.

To improve the saturation of the simulated vertical MOSFET and possibly to reduce its channel length, an additional 20 nm p^+ buffer was introduced in the central *p* region. The chosen doping concentration in this p^+ buffer is 3 × 10^{18} cm^{-3}. The buffer was placed 10 nm from the *n–p* source junction, as shown in figure 3.

The buffer separation from the *n–p* source junction keeps the value of the source–junction field low enough to avoid impact ionization. Additionally, punch-through is reduced because of the higher doping. The drain current reduction due to the buffer was compensated by reducing the oxide thickness to 4 nm.

Simulation results for the device with the p^+ buffer are compared in figure 4 with measurements made on a transistor with a homogeneously doped channel. The simulated saturation is improved compared with the measured characteristics shown with broken lines. The reduction of the punch-through due to the p^+ buffer allows a further reduction of the channel length to 50 nm. The simulated drain current curves of a 50 nm transistor are shown in figure 5.

A slight increase in the doping concentration in the channel was used to improve the output conductance, where values of 4×10^{18} cm^{-3} and 5×10^{18} cm^{-3} were used within the channel and in the p^+ buffer, respectively.

3. Future work

The CA approach for semiconductor device simulation allows the implementation of the basic algorithms in fast and accurate programs, reducing the trade-off between modelling capabilities and computational efficiency which is typical of particle-based simulation techniques. On the other hand, the method can nowadays be considered mature enough to start working on its improvements.

A full-band approach [7] is required to improve the band structure model used in the simulation programs and to allow better modelling of high-energy, hot-electron effects. The full-band phonon dispersion has to be included as well in the scattering rate computation, and all phonon modes have to be considered, such that the asymmetric nature of scattering rates in momentum space is accounted for [8].

Three-dimensional (3D) simulation capabilities in real space must be included to model 3D effects in small devices and exotic geometries [9]. This aspect does not require additional theoretical effort, but the frequent solution of Poisson's equation in 3D is a formidable task from the computational viewpoint, for both the time *and* the memory required. This requires the use of state-of-the-art approaches, such as multigrid and/or conjugate gradient, which we are currently implementing in 3D in scalar and parallel versions.

The inclusion of anisotropic scattering within a full-band framework requires the full transition table for every initial and final state across the full Brillouin zone. The very intensive memory usage has to be optimized and a given indetermination in energy has to be accounted for in a full-band approach when compared with the simpler, symmetric band models used in the current programs. Although the problems due to memory requirements appear not to have a trivial solution, the possibility to have a complete (and efficient) modelling of the carrier dynamics in *k* space appears feasible.

We are currently working to improve the original algorithm proposed by Kometer *et al* [2], and a demonstration of this new version is expected, including the aforementioned issues.

4. Conclusions

The capability of the CA approach to predict accurately the highly nonlinear transport behaviour of real nanostructured semiconductor devices was demonstrated. A novel family of vertical MOSFETs was investigated, giving results in excellent agreement with the experiment. Simulation results were shown of an inhomogeneous doping profile that make it possible to scale the device down to 50 nm. Improvements of the CA method to include full-band structure information and full anisotropic phonon scattering were discussed.

Acknowledgments

This work was partially supported by the NSF Grant ECS–9312240 and by the Siemens project SFE.

References

[1] Jacoboni C and Lugli P 1989 *The Monte Carlo Method for Semiconductor Device Equations* (New York: Springer)

[2] Kometer K, Zandler G and Vogl P 1992 Lattice-gas cellular-automaton method for semiclassical transport in semiconductors *Phys. Rev.* B **46** (3) 1382–94

[3] Kometer K 1992 Lösung der Boltzmann–Gleichung mit der Methode zellulärer Automaten *PhD Thesis* Technische Universität München, Walter Schottky Institut

[4] Rein A 1995 Zelluläre Automaten in der Transporttheorie: Konzepte und Anwendungen *PhD Thesis* Technische Universität München, Walter Schottky Institut

[5] Hofmann F, Krautschneider W H, Risch L and Schaefer H 1995 CVD–EPI MOS transistor with a 65 nm vertical channel *Extended Abstracts of the 1995 Int. Conf. on Solid State Devices and Materials (Osaka)* pp 46–8

[6] Risch L, Krautschneider W H, Hofmann F, Schäfer H, Aeugle T and Rösner W 1996 Vertical MOS transistors with 70 nm channel length *IEEE Trans. Electron Devices* **43** (9) 1495–8

[7] Fischetti M V and Laux S E 1988 Monte Carlo analysis of electron transport in small semiconductor devices including band-structure and space-charge effects *Phys. Rev.* B **38** (14) 9721–45

[8] Kunikiyo T, Takenaka M, Kamakura Y, Yamaji M, Mizuno H, Morifuji M, Taniguchi K and Hamaguchi C 1994 A Monte Carlo simulation of anisotropic electron transport in silicon including full band structure and anisotropic impact-ionization model *J. Appl. Phys.* **75** (1) 297–312

[9] Auth C P and Plummer J D 1997 Scaling theory for cylindrical, fully-depleted, surrounding-gate MOSFET's *IEEE Electron Device Letters* **18** (2) 74–6

Semicond. Sci. Technol. **13** (1998) A180–A183. Printed in the UK PII: S0268-1242(98)91842-9

Double electron layer tunnelling transistor (DELTT)

M A Blount[†][‡], J A Simmons[†], J S Moon[†][§], W E Baca[†], J L Reno[†] and M J Hafich[†]

[†] Sandia National Laboratories, Albuquerque, NM 87185-1415, USA
[‡] Department of Physics and Astronomy, University of New Mexico, NM, USA
[§] Department of Physics, Michigan State University, MI, USA

Received 7 December 1997, accepted for publication 11 March 1998

Abstract. We demonstrate the double electron layer tunnelling transistor (DELTT), based on the gate control of two-dimensional–two-dimensional tunnelling in a double quantum well. Unlike previously proposed resonant tunnelling transistors, the DELTT is entirely planar and can be easily fabricated in large numbers. At 1.5 K we demonstrate peak-to-background ratios of ~ 50:1 in source–drain conductance versus gate voltage and peak-to-valley ratios of ~ 20:1 in the source–drain current versus source–drain voltage. Using a single DELTT in series with a load resistor, we demonstrate low-power bistable memories at 1.5 K. We also demonstrate a unipolar complementary static RAM by connecting two DELTTs in series.

Quantum devices based on resonant tunnelling have long been of great interest, because of two highly desirable attributes. First, coherent tunnelling is inherently extremely high speed: double barrier resonant tunnelling diodes (DBRTDs) have been demonstrated at terahertz frequencies [1]. Second, the negative differential resistance (NDR) inherent in resonant tunnelling makes possible multifunctionality, enabling the same circuit functions to be achieved with far fewer devices [2]. However, DBRTDs suffer from the lack of a third control terminal, precluding electrical isolation between the input and output terminals. While hybrid devices formed by combining DBRTD structures with conventional transistors circumvent this difficulty [3], they do so by sacrificing much of the speed advantage of resonant tunnelling. Schemes to add a third terminal to the DBRTD so as to achieve a true quantum tunnelling transistor have been of two types. The first type proposes to make electrical contact to the resonant tunnelling layer [4], which has largely proven problematic. The second type proposes lateral gate control of the potential of the resonant tunnelling layer [5], or lateral quantum confinement. [6] However, such devices are extremely sensitive to slight variations in lateral dimension, making them nearly impossible to fabricate in the large numbers necessary for real-world applications, a difficulty that is shared by single-electron transistors.

In this work we demonstrate a novel quantum effect FET, the double electron layer tunnelling transistor (DELTT), which does not require lateral confinement but rather is entirely planar in configuration. The DELTT operates through tunnelling between the two parallel two-dimensional (2D) electron layers in a double quantum well (DQW) heterostructure. The critical dimensions all reside in the growth direction, and thus take full advantage of the monolayer resolution of molecular beam epitaxy. Further, because both layers are 2D, the tunnelling is easily modulated by surface Schottky gates. The DELTT shows excellent resonant NDR characteristics, with I–V peak-to-valley ratios of ~ 20:1 at 1.5 K. More importantly, the height and position of the resonant tunnelling I–V peak are controllable by gate voltage over a wide range. A single DELTT exhibits large regions of both negative and positive transconductance, readily enabling multifunctional applications. Using these DELTTs, we have demonstrated both single- and two-DELTT bistable static memory cells at 1.5 K, which use half as many transistors as their conventional counterparts.

Devices were processed from a modulation-doped DQW GaAs/$Al_{0.3}Ga_{0.7}As$ heterostructure grown by molecular beam epitaxy. The DELTTs had 150 Å wide top and bottom QWs, with electron densities of 2.0×10^{11} cm^{-2} and 1.4×10^{11} cm^{-2}, respectively, and a 125 Å $Al_{0.3}Ga_{0.7}As$ barrier between the QWs. The source terminal makes electrical contact to the top QW only, while the drain terminal contacts the bottom QW only. The contacts are formed by diffusing Au/Ge/Ni to both QWs, and then using a fixed DC bias on adjacent 10 μm wide Ti/Au depletion gates to deplete electrons from the QW one does not wish to contact [7]. An additional top surface Ti/Au control gate acts as a third terminal. The control gate was 40 mm long and 500 mm wide. We used our previously developed flip-chip processing scheme, the epoxy-bond-and-stop-etch technique [8], to place the backgates at a distance of ~ 2 μm from the bottom QW.

We now discuss the DELTT's principle of operation. Resonant tunnelling can occur only when there exist states

0268-1242/98/SA0180+04$19.50

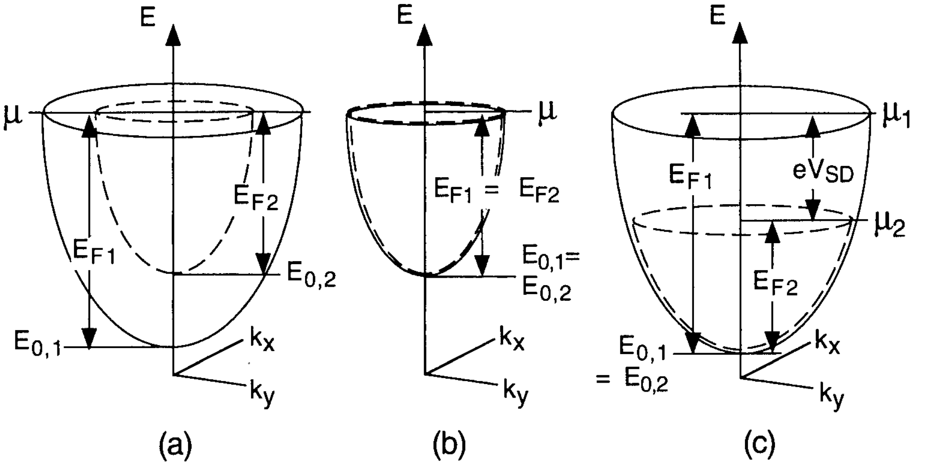


Figure 1. Sketch of the dispersion of a DQW. Tunnelling occurs when states in the two paraboloids coincide. In (a) $V_{SD} = 0$ and the densities are unequal at $V_{TC} = 0$. In (b) the densities have been made equal by applying a finite V_{TC} at $V_{SD} = 0$. In (c) the densities remain unequal and $V_{TC} = 0$, but the paraboloids are made to coincide by applying a source–drain voltage $V_{SD} = (E_{F1} - E_{F2})/e$.

in both QWs with identical energy and in-plane momentum, since these quantities must be conserved in a tunnelling event [9]. One can determine when this condition is met by sketching the allowed states of the DQW in energy–momentum space, i.e. the dispersion curve. Figure 1(a) shows the dispersion of a DQW with a higher electron density in the top QW than the bottom, at a top control gate voltage $V_{TC} = 0$, and an applied source–drain bias $V_{SD} \approx 0$. Because both layers are 2D, their allowed states each form a paraboloid having states only on the surface and none in the interior. While the chemical potentials $\mu_{1,2}$ of the two QWs coincide ($V_{SD} = 0$), their Fermi energies $E_{F1,2}$ differ. Thus the paraboloids are offset in energy, no pairs of states of identical momentum and energy exist and tunnelling does not occur. Tunnelling can be switched on by (1) varying the densities of either 2D layer with surface gates, (2) changing the chemical potential difference between the QWs by applying a source–drain bias or (3) both. Figure 1(b) shows the dispersion when $V_{SD} = 0$ and the density of the top QW is decreased via a negative V_{TC} until it equals that of the bottom QW. The two paraboloids then coincide, allowing tunnelling to occur. This situation corresponds to a peak in the small-signal source–drain conductance G_{SD} as a function of V_{TC}. Figure 1(c) shows the dispersion when $V_{TC} = 0$ and V_{SD} is increased to equal $(E_{F1} - E_{F2})/e$. In that case the two paraboloids again coincide. Occupied states in the top QW are now paired with unoccupied states in the bottom QW, and tunnelling can again occur. This situation corresponds to a peak in the source–drain current–voltage (I–V) curve. We note that while the physics of 2D–2D tunnelling has been previously explored [9], the use of such structures for device applications, particularly their source–drain I–V curves, has received little attention [9–11].

Measurements of the DELTT confirm the above description of its operating principles. Figure 2(a) shows G_{SD} versus V_{TC} at 0.3 K. A strong tunnelling resonance appears and is 50 times larger than the background conductance. This happens at $V_{TC} = -0.12$ V, when the top QW has been partially depleted so as to have the same density as the bottom QW, 1.4×10^{11} cm^{-2}. The device thus exhibits a strong negative transconductance for $V_{TC} > -0.12$ V, which actually exceeds the positive

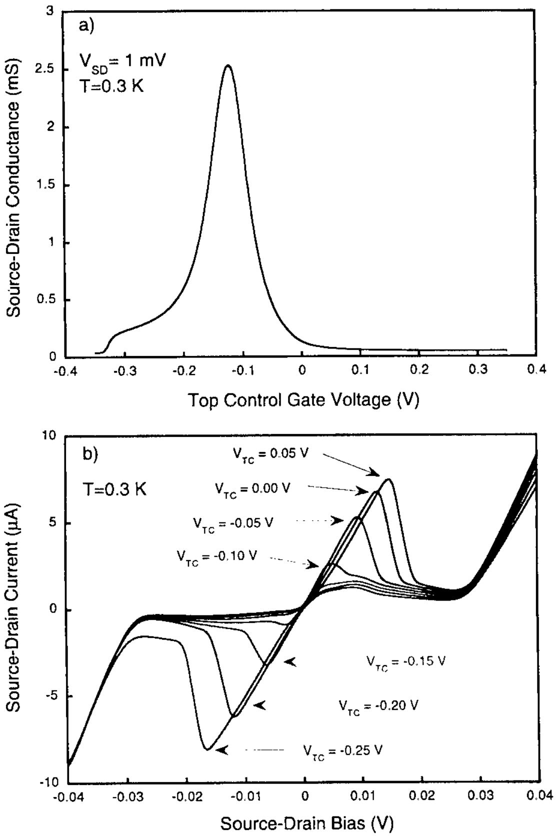


Figure 2. DELTT characteristics at 0.3 K. (a) Source–drain conductance G_{SD} versus top control gate voltage V_{TC}. (b) Source–drain current I_{SD} versus source–drain voltage V_{SD} for several values of V_{TC}.

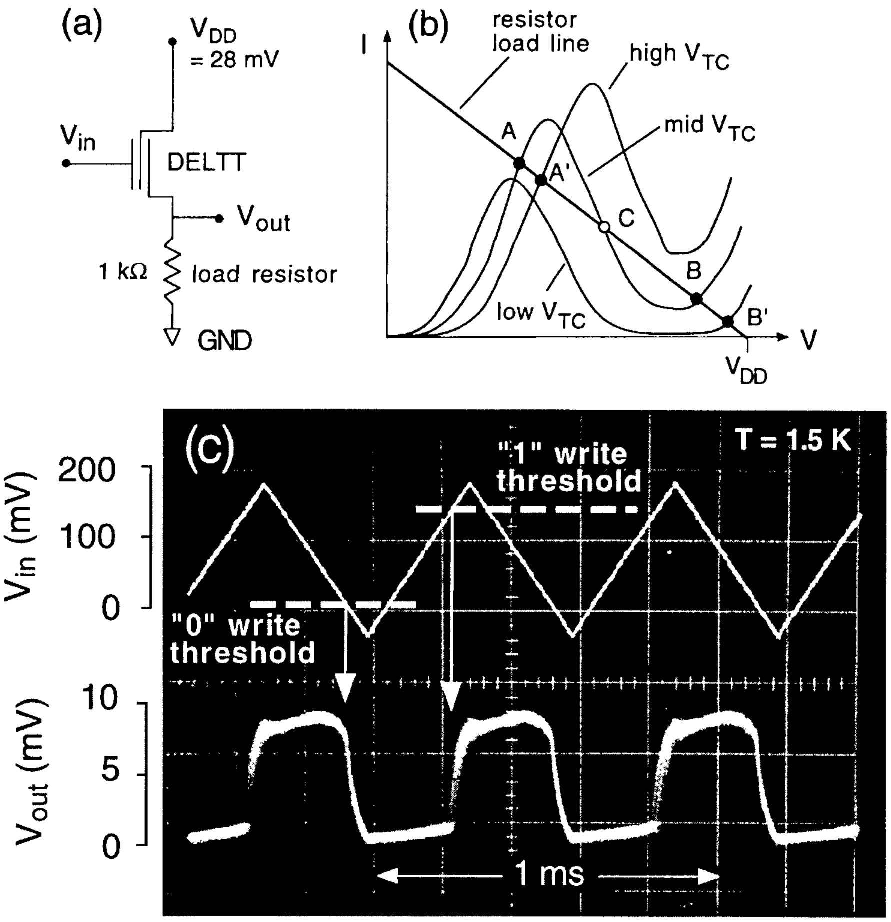


Figure 3. (a) Bistable memory circuit using a single DELTT. (b) Sketch of DELTT I–V curves at three characteristic values of V_{TC} and the resistor load line. (c) Oscilloscope data demonstrating that the circuit behaves as a memory.

transconductance occuring for $V_{TC} < -0.12$ V. As will be shown below, this property is extremely useful for circuit applications. Figure 2(b) shows the source–drain I–V for several fixed values of V_{TC}, all at 0.3 K. A resonant tunnelling peak occurs when the offset in chemical potential μ is equal to the difference in Fermi energies between the two QWs. (V_{SD} required to achieve this condition is actually larger than $(E_{F1} - E_{F2})/e$ because of a significant resistive voltage drop under the depletion gates.) At higher V_{SD} the current drops dramatically, resulting in a strong NDR. Peak-to-valley ratios approaching 20:1 are observed. More importantly, the position of the resonance shifts strongly with V_{TC}, clearly demonstrating the gate control of tunnelling. Note that the position of the resonance in V_{SD} can be moved through $V_{SD} = 0$, as the density difference between the QWs changes sign with variation of V_{TC}. For a fixed V_{TC} the resonance occurs only for one polarity of V_{SD}, in contrast to DBRTDs. These characteristics change little until ~ 10 K.

Because of the multifunctionality enabled by the DELTTs NDR characteristics, highly compact circuits can be constructed. Two different bistable static memory cells were demonstrated. The first was a circuit consisting of a DELTT in series with a 1 kΩ load resistor, as shown in figure 3(a). Figure 3(b) shows a sketch of the I–V curves of the DELTT at three different V_{TC}, together with the resistor's load line. For 10 mV < V_{TC} < 140 mV, a V_{DD} of 28 mV causes the resistor load line to intersect the I–V curve at two stable points A and B and one unstable point C. By sweeping the control gate above an upper threshold (140 mV) or below a lower threshold (10 mV), the circuit is forced to contain only one stable point, A′ for high V_{TC} and B′ for low V_{TC} . The circuit will then latch in that particular stable point and will remain there once V_{TC} is returned to an intermediate value. Thus at intermediate V_{TC} the state of the circuit depends on its past history, meaning that it behaves as a bistable static memory cell. Figure 3(c) shows an oscilloscope trace of the input and output voltages of this circut, at 1.5 K. A triangular wave is fed to the input, causing the output to switch to the low or high state whenever the write thresholds are reached, clearly demonstrating the memory effect. The on–off output voltage ratio is ~ 10 : 1, allowing simple biasing and good state discrimination. We note that because of the DELTTs NDR, this memory requires only one transistor and one load resistor, in contrast to the two transistors and two load resistors of a conventional NMOS SRAM, which behaves similarly.

The second bistable memory cell utilized two DELTTs in series, as shown in the inset to figure 4(a). In a manner similar to the single-DELTT memory element, this CDELTT memory cell utilizes a pair of stable intersection

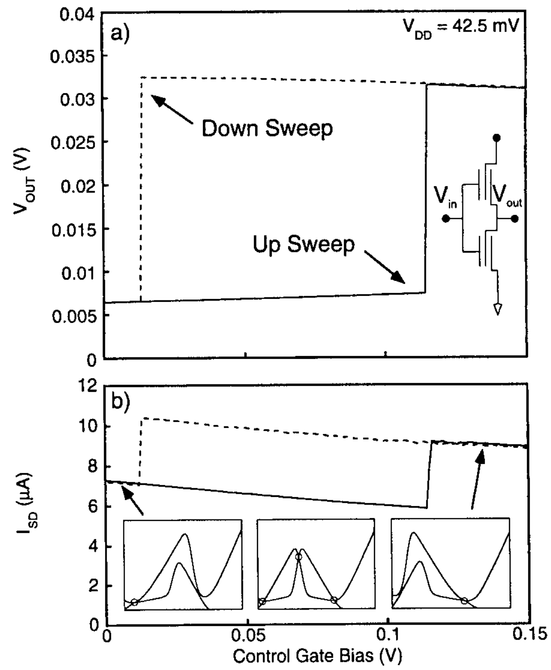


Figure 4. Bistable complementary DELTT (CDELTT) memory. (a) Output voltage and (b) circuit current as a function of control gate bias, demonstrating memory function. Inset to (a) shows the complementary DELTT circuit, while the insets to (b) show the *I–V* and load line diagrams for various ranges of input voltage.

points and a region of bistability. In this case, however, one DELTT provides the load line seen by the other. By raising or lowering the *I–V* characteristic of one DELTT relative to the other with V_{IN}, the circuit can be latched into a high or low state. This configuration has a tremendous advantage over the single-DELTT memory element in that the static current is determined by the valley current of the DELTTs used. For large peak-to-valley ratios, the current will always be low regardless of which state the circuit is in (see inset to figure 4(b)). Figure 4 shows the voltage output and source–drain current of the CDELTT memory cell operated at $V_{DD} = 42.5$ mV as a function of V_{IN}. While the output voltage states latch at ~ 7 mV and ~ 32 mV, a factor of ~ 5 difference, the current switches between 7 μA and 10 μA, therefore changing by only $\sim 30\%$. Thus, although the circuit is unipolar, because the DELTT exhibits both positive and negative transconductance, the circuit can act in a complementary fashion similar to CMOS SRAMs. Because of the DELTT's NDR, however, only two transistors are required instead of four.

In summary, we have demonstrated a novel quantum tunnelling transistor based on the gate control of 2D–2D tunnelling, whose structure is entirely planar. Using a single DELTT and a load resistor, we demonstrated bistable memories at 1.5 K. By placing two DELTTs in series, we also demonstrated a unipolar complementary memory with relatively low current in both memory states. Elsewhere we report on high electron density single-DELTT memories operating at 77 K [12] and the use of the DELTT to construct digital logic gates and gate-controlled oscillators [13].

Acknowledgment

Sandia is a multiprogramme laboratory operated by Sandia Corporation, a Lockheed Martin company, for the US Department of Energy under Contract DE-AC04-94AL85000.

References

[1] Sollner T C L G, Goodhue W D, Tannenwald P E, Parker C D and Peck D D 1983 *Appl. Phys. Lett.* **43** 588
Brown E R, Soderstrom J R, Parker C D, Mahoney L J, Molvar K M and McGill T C 1991 *Appl. Phys. Lett.* **59** 2291

[2] Capasso F, Sen S, Beltram F, Lunardi L M, Vengurlekar A S, Smith P R, Shah N J and Malik R J 1989 *IEEE Trans. Electron Dev.* **36** 2065
Shen J, Kramer G, Tehran S, Goronkin H and Tsui R 1995 *IEEE Electron Device Lett.* **16** 178

[3] Capasso F and Kiehl R A 1985 *J. Appl. Phys.* **58** 1366
Woodward T K, McGill T C, Chung H F and Burnham R D 1987 *Appl. Phys. Lett.* **51** 1542

[4] Schulman J N and Waldner M 1988 *J. Appl. Phys.* **63** 2859
Reed M A, Frensley W R, Matyi R J, Randall J N and Seabaugh A C 1989 *Appl. Phys. Lett.* **54** 1034

[5] Dellow M W, Beton P H, Henini M, Main P C, Eaves L, Beaumont S P and Wilkinson C D W 1991 *Electron. Lett.* **27** 134
Peatman W C B, Brown E R, Rooks M J, Maki P, Grimm W J and Shur M 1994 *IEEE Electron Dev. Lett.* **15** 236
Kolagunta V R, Janes D B, Chen G L, Webb K J, Melloch M R and Youtse C 1996 *Appl. Phys. Lett.* **69** 374

[6] Chou S Y, Allee D R, Pease R F W and Harris J S 1989 *Appl. Phys. Lett.* **55** 1762
Ismail K, Antoniadis D A and Smith H I 1989 *Appl. Phys. Lett.* **55** 589

[7] Eisenstein J P, Pfeiffer L N and West K W 1990 *Appl. Phys. Lett.* **57** 2324

[8] Weckwerth M V, Simmons J A, Harff N E, Sherwin M E, Blount M A, Baca W E and Chui H C 1996 *Superlattices Microstruct.* **20** 561

[9] Eisenstein J P, Gramila T J, Pfeiffer L N and West K W 1991 *Phys. Rev.* B **44** 6511
Simmons J A, Lyo S K, Klem J F, Sherwin M E and Wendt J R 1993 *Phys. Rev.* B **47** 15741
Brown K M, Linfield E H, Ritchie D A, Jones G A C, Grimshaw M P and Pepper M 1994 *Appl. Phys. Lett.* **64** 1827

[10] Katayama Y and Tsui D C 1993 *Appl Phys. Lett.* **62** 2563

[11] Luryi S 1988 *Appl. Phys. Lett.* **52** 501

[12] Simmons J A, Blount M A, Moon J S, Baca W E, Reno J L and Hafich M J 1997 *IEEE Tech. Digest of the Int. Electron Devices Meeting (7–10 December 1997, Washington DC)* p 755

[13] Moon J S, Simmons J A, Blount M A, Baca W E, Reno J L and Hafich M J *Proc. Int. Semiconductor Device Research Symp. (Charlottesville, VA, 10–13 December, 1997)* p 27

Electronic journals available

Bioimaging
Classical and Quantum Gravity
Combustion Theory and Modelling
Distributed Systems Engineering
European Journal of Physics
High Performance Polymers
Inverse Problems
Journal of Micromechanics and Microengineering
Journal of Optics
Journal of Physics A: Mathematical and General
Journal of Physics B: Atomic, Molecular and Optical Physics
Journal of Physics: Condensed Matter
Journal of Physics D: Applied Physics
Journal of Physics G: Nuclear and Particle Physics
Journal of Radiological Protection
Modelling and Simulation in Materials Science and Engineering
Nanotechnology
Network: Computation in Neural Systems
Measurement Science and Technology
Nonlinearity
Physics Education
Physics in Medicine and Biology
Physiological Measurement
Plasma Physics and Controlled Fusion
Plasma Sources Science and Technology
Public Understanding of Science
Pure and Applied Optics: Journal of the European Optical Society Part A
Quantum and Semiclassical Optics: Journal of the European Optical Society Part B
Reports on Progress in Physics
Semiconductor Science and Technology
Smart Materials and Structures
Superconductor Science and Technology
Waves in Random Media

Institute *of* **Physics** PUBLISHING

▸▸▸▸ **http://www.iop.org** ◂◂◂◂

Setting the Pace in Electronic Publishing

Institute of Physics Publishing brings vital physics resources direct to your desktop with an upgraded Electronic Journals service incorporating new HyperCite™ linking technology.

All of our research journals are available on the World Wide Web at no additional cost to institutional subscribers, offering full text articles in PDF or PostScript format, and new HyperCite™ technology providing links to a range of full text and abstract archives.

Institutions can register for free electronic access to the print journals to which they subscribe.

New in 1998, electronic access to our five year full text archive (1993–1997) will also be available at no extra cost to all institutional subscribers, including new subscribers.

Registration forms are available online (http://www.iop.org/Support/EJ)

Practical, Time-saving Features

- Work where, when, and how you want – from home, office, or library, 24 hours a day, 7 days a week.
- Read print-quality full text articles with all mathematics and figures, using Adobe's free Acrobat Reader software or a PostScript printer.
- Browse tables of contents (free to all visitors) or search abstracts using familiar World Wide Web technology, such as Netscape Navigator.
- Find papers cited by and papers citing your selected article; trace author's past influence, and see the effect that an article has on future literature with new Hypercite™ technology.
- Powerful searching with Boolean, truncated, string, and nested search terms.
- An extended IoP abstract archive back to first issue (where available for subscribed titles).
- Articles available online weeks before the print edition.
- View multimedia enhancements including videos, computer simulations and data sets.
- A virtual filing cabinet in which users can store articles of interest annotated with personal notes.
- New hyperlinked e-mail alert providing advanced warning of the latest relevant papers corresponding to your chosen search items.
- Customizable menus for faster access to your areas of interest.
- Pre-set searching based on your preferences and subject area.
- No site or user ID and passwords required: access is controlled by IP address or domain name.
- Individual users may create their own personal logins to take advantage of valuable customizable features.

For further information, a full list of journals and details of how to make sure your site is registered, see http://www.iop.org/EJ/welcome or e-mail: custserv@ioppublishing.co.uk to request an information pack

Europe and Rest of World
Customer Services Department
Institute of Physics Publishing
Dirac House, Temple Back
Bristol BS1 6NX, UK

Tel: +44 (0)117 929 7481
Fax: +44 (0)117 929 4318

USA, Canada and Mexico
IOP Publishing Inc.,
The Public Ledger Building, Suite 1035
150 South Independence Mall West
Philadelphia, PA 19106, USA.

Tel: (800) 358 4677
Fax: (215) 627 0879

Institute *of* **Physics** PUBLISHING

Online Author and Referee Services

Enhance Your Communications with Institute of Physics Publishing

As part of its continuing commitment to support authors and referees throughout the publishing process, Institute of Physics Publishing provides the following services on the World Wide Web. These make the process of working with us faster and more efficient, with a range of communication options available to you right from your desktop.

Services for Referees *(http://www.iop.org/Journals/rsi)*

To simplify the peer review process, you can use your World Wide Web browser to:

- Keep your personal information up to date online, including your contact details, areas of research interest and dates of availability.
- Complete and return your reviews online (for selected journals) – no more paper forms to fill out or letters to post – send your referee forms electronically!
- Check classification codes online – use links to information on PACS, AMS, and PMB classification schemes when ever you need to check or specify the classification codes for your areas of expertise.

Register as a New Referee *(http://www.iop.org/Journals/rsi)*

To maintain the quality of our journals, we're constantly updating and expanding our list of expert referees. If you are willing to referee articles occasionally, you can submit your contact details and areas of expertise online using your World Wide Web browser (alternatively, you can email this information to refquest@ioppublishing.co.uk). This will ensure that any articles sent to you for review will be closely related to your fields of interest.

Services for Authors *(http://www.iop.org/Journals/asi)*

Our Author Enquiry Service helps you track your article through the assessment and production processes, and lets you exchange vital information with us quickly and efficiently. You can:

- Determine the status of your submitted article.
- Find out the volume and page numbers for your published article.
- Provide us with updated contact information, such as a change of address.

To use these author and referee services, visit our home page at http://www.iop.org and look for more information under Author Services or Referee Services.

Our online author and referee services are continually being updated and enhanced; your feedback is a vital part of that process, so please let us know what you think. You can provide feedback from our Web site, or send comments and questions by email to authref@ioppublishing.co.uk.

6. SINGLE ELECTRON DEVICES

7. QUANTUM DOT ARRAYS

8. QUANTUM WELLS

9. NANODEVICES

Continued

2. ORGANICS ON SEMICONDUCTORS

3. IMPURITIES IN HETEROSTRUCTURES

4. FABRICATION OF NANOSTRUCTURES

5. HETEROSTRUCTURES AND TRANSPORT

Continued

Semiconductor Science and Technology

Volume 13 Number 8A August 1998

2nd International Workshop on Surfaces and Interfaces of Mesoscopic Devices, Hawaii, 7–12 December 1997

1. ELECTRON TRANSPORT IN QUANTUM DOTS

Continued on inside back cover

Bibliographic codes CODEN: SSTEET 13 (8A) A1–A184 (1998) ISSN: 0268-1242