NanoScience and Technology

Springer
Berlin
Heidelberg
New York
Barcelona
Budapest
Hong Kong
London
Milan
Paris
Santa Clara
Singapur
Tokyo

NanoScience and Technology

Series Editors: K. von Klitzing R. Wiesendanger

Sliding Friction
Physical Principles and Applications
By B. N. J. Persson

Scanning Probe Microscopy
Analytical Methods
Editor: R. Wiesendanger

Mesoscopic Physics and Electronics
Editors: T. Ando, Y. Arakawa, K. Furuya, S. Komiyama, and H. Nakashima

T. Ando · Y. Arakawa · K. Furuya
S. Komiyama · H. Nakashima (Eds.)

Mesoscopic Physics and Electronics

With 203 Figures

Springer

Professor T. Ando

Institute for Solid State Physics
University of Tokyo
7-22-1 Roppongi, Minato-ku, Tokyo 106
Japan

Professor Y. Arakawa

Institute of Industrial Science
University of Tokyo
7-22-1 Roppongi, Minato-ku, Tokyo 106
Japan

Professor K. Furuya

Department of Electrical and Electronic Engineering
Tokyo Institute of Technology
2-1-12 O-okayama, Meguro-ku, Tokyo 152
Japan

Professor S. Komiyama

Department of Basic Science
University of Tokyo
3-8-1 Komaba, Meguro-ku, Tokyo 153
Japan

Professor H. Nakashima

Institute of Scientific and Industrial Research
Osaka University
8-1 Mihogaoka, Ibaraki, Osaka 567
Japan

Series Editors:

Professor Dr., Dres. h. c. Klaus von Klitzing

Max-Planck-Institut für Festkörperforschung, Heisenbergstrasse 1
D-70569 Stuttgart, Germany

Professor Dr. Roland Wiesendanger

Institut für Angewandte Physik, Universität Hamburg, Jungiusstrasse 11
D-20355 Hamburg, Germany

ISSN 1434-4904
ISBN-13:978-3-642-71978-3

Library of Congress Cataloging-in-Publication Data applied for

Die Deutsche Bibliothek – CIP-Einheitsaufnahme

Mesoscopic physics and electronics / T. Ando ... (ed.). – Berlin ; Heidelberg ; New York ; Barcelona ; Budapest ; Hong Kong ; London ; Milan ; Paris ; Santa Clara ; Singapore ; Tokyo : Springer, 1998
(NanoScience and technology)
ISBN-13:978-3-642-71978-3 e-ISBN-13:978-3-642-71976-9
DOI: 10.1007/978-3-642-71976-9

This work is subject to copyright. All rights are reserved, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilm or in any other way, and storage in data banks. Duplication of this publication or parts thereof is permitted only under the provisions of the German Copyright Law of September 9, 1965, in its current version, and permission for use must always be obtained from Springer-Verlag. Violations are liable for prosecution under the German Copyright Law.

© Springer-Verlag Berlin Heidelberg 1998
Softcover reprint of the hardcover 1st edition 1998

The use of general descriptive names, registered names, trademarks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

Typesetting: Camera ready by authors/editors

SPIN: 10648070 54/3144 - 5 4 3 2 1 0 – Printed on acid-free paper

Preface

Semiconductor technology has developed considerably during the past several decades. The exponential growth in microelectronic processing power has been achieved by a constant scaling down of integrated circuits. Smaller feature sizes result in increased functional density, faster speed, and lower costs. One key ingredient of the LSI technology is the development of the lithography and microfabrication. The current minimum feature size is already as small as 0.2 μm, beyond the limit imposed by the wavelength of visible light and rapidly approaching fundamental limits. The next generation of devices is highly likely to show unexpected properties due to quantum effects and fluctuations.

The device which plays an important role in LSIs is MOSFETs (metal-oxide-semiconductor field-effect transistors). In MOSFETs an inversion layer is formed at the interface of silicon and its insulating oxide. The inversion layer provides a unique two-dimensional (2D) system in which the electron concentration is controlled almost freely over a very wide range. Physics of such 2D systems was born in the mid-1960s together with the development of MOSFETs. The integer quantum Hall effect was first discovered in this system.

The development of crystal growth technology such as MBE (molecular beam epitaxy) and MOCVD (metalorganic chemical vapor deposition) made it possible to fabricate semiconductor heterostructures, quantum wells, superlattices, etc. Using supreme quality heterointerfaces a 2D electron gas having a very high mobility and a long mean free path has been realized. Thus 2D physics has developed considerably and entered a new era. The fractional quantum Hall effect observed in heterostructures for the first time is known to be a manifestation of a new quantum liquid possible only in ideal 2D systems without disorder. Semiconductor heterostructures are already used as real devices such as the HEMT (high electron mobility transistor) and quantum well lasers.

The combination of lithography technology and semiconductor heterostructures has produced quantum wires, quantum dots, antidots, quantum point contacts, etc. In such new quantum systems or structures, the feature size is smaller than the electron mean free path and various new phenomena are observed in the electron transport at low temperatures. They include conductance quantization due to ballistic electron motion, universal conductance fluctuations due to quantum interference effects, the Coulomb blockade or single-electron tunneling due to the discreteness of the electronic charge, etc. These new phenomena have provided challenging subjects important from

the point of view of fundamental physics, and various quantum-effect devices utilizing quantum structures are actively being pursued.

Quantum wires and dots are expected to be able to significantly improve optical devices such as lasers. A full confinement of photons as well as electrons in microcavities is now under discussion for future realization of zero-threshold lasers in which electron-hole transitions are coupled with only a single spontaneous emission mode.

In April 1994, a three-year inter-university research program "Quantum Coherent Electronics: Physics and Technology (QCEPT)" was launched supported by the Grant-in-Aid for Scientific Research on Priority Area from the Ministry of Education, Science, and Culture, Japan (T. Ando, project head). The program consisted of five groups: [a] theory of quantum transport (T. Ando, group leader), [b] quantum coherent transport (S. Komiyama, group leader), [c] coherent electron-photon interaction (Y. Arakawa, group leader), [d] quantum coherent devices (K. Furuya, group leader), and [e] quantum process technologies (H. Nakashima, group leader). The total number of researchers involved in this program is about 60. The aim of the program is to explore a new research field to be called QCEPT through understanding basic physics of quantum structures, pursuing their controllability, proposing new devices, and also developing new technologies to fabricate atomically-controlled structures.

This book was written for the purpose of giving general reviews on various subjects in this new research field and discussing their future prospects, being supported in part by Grant-in-Aid for Publication of Scientific Research Result from the Ministry of Education, Science, Sports and Culture, Japan. It was edited in such a way that it may be used as an introductory textbook for graduate students who have just entered this field. Because most of the authors of this book are active members of the QCEPT program, it inevitably contains some of the scientific and technological achievements obtained in the course of the program.

Tokyo and Osaka
September 1997

Tsuneya Ando
Yasuhiko Arakawa
Kazuhito Furuya
Susumu Komiyama
Hisao Nakashima

Contents

1. Introduction – Mesoscopic Systems 1

1.1 Introduction 1
T. Ando

1.2 Length Scales Characterizing Mesoscopic Systems 3
T. Ando

1.2.1 Fermi Wavelength 3
1.2.2 Mean Free Path 3
1.2.3 System Size 3
1.2.4 Thermal Diffusion Length and Thouless Energy 4
1.2.5 Phase Coherence Length 5
1.2.6 Diffusive Regime and Ballistic Regime 5
1.2.7 Quantum Wires, Dots, and Antidots 6
1.2.8 Anderson Localization 8
References 10

1.3 Landauer's Formula 11
T. Ando

1.3.1 Conductance and Transmission Probability 11
1.3.2 Some Applications 12
a. Universal Conductance Fluctuations 12
b. Conductance Quantization 13
References 14

1.4 Fluctuations and Aharonov–Bohm Effect 15
H. Fukuyama

1.4.1 Aharonov–Bohm Effect 15
1.4.2 Universal Conductance Fluctuations 16
1.4.3 Persistent Current 17
1.4.4 Fluctuations of Orbital Susceptibility 19
References 20

1.5 Ballistic Electron Transport 22
A. Kawabata

1.5.1 Quantization of Conductance 22
1.5.2 Interaction Effects on Conductance Quantization 25
1.5.3 Magnetic Focusing 27
1.5.4 Bend Resistance and Transfer Resistance 27
1.5.5 Anomaly in Weak-Field Hall Effect 28
References 30

1.6 Coulomb Blockade 31
A. Kawabata
1.6.1 Introduction 31
1.6.2 Single Electron Tunneling 31
1.6.3 SET Oscillation 32
1.6.4 Tunneling in Superconducting Junctions 33
1.6.5 Coulomb Blockade in Quantum Dots 36
1.6.6 Resonant Transmission and Kondo Effect 37
1.6.7 KTB Transition in Junction Network 40
References 43

2. Transport in Quantum Structures 45

2.1 Tomonaga–Luttinger Liquid in Quantum Wires 45
M. Ogata and H. Fukuyama
2.1.1 Introduction 45
2.1.2 Tomonaga–Luttinger Liquid 45
2.1.3 Conductance of Finite-Length Quantum Wire 49
2.1.4 Quantized Value of Conductance 50
2.1.5 Mott–Hubbard Insulator 52
References 52

2.2 Quantum Wires 54
A. Kawabata
2.2.1 Magnetoresistance and Boundary-Roughness Scattering 54
2.2.2 One-Dimensional Electron in Slowly Varying Potential 56
2.2.3 Interaction Effects in Quantum Wires 58
References 60

2.3 Magnetophonon Resonance in Quantum Wires 61
N. Mori and C. Hamaguchi
2.3.1 Introduction 61
2.3.2 Theory 62
2.3.3 Experiments 64
References 65

2.4 Quantum Dots and Artificial Atoms 66
S. Tarucha
2.4.1 Quantum Dots Containing a Few Electrons 66
2.4.2 Atom-like Properties – Shell Filling 67
2.4.3 Atom-like Properties – Spin Effects 68
References 71

2.5 Antidot Lattices – Classical and Quantum Chaos 72
T. Ando
2.5.1 Antidot Lattices 72
2.5.2 Commensurability Peaks 72
2.5.3 Aharonov–Bohm Type Oscillation 76
2.5.4 Altshuler–Aronov–Spivak Oscillation 80
2.5.5 Scattering Matrix Formalism 82

2.5.6 Anderson Localization 87
References 88

2.6 Electric and Magnetic Lateral Superlattices 90
Y. Iye
2.6.1 Lateral Modulation 90
2.6.2 Weiss Oscillation 90
2.6.3 Magnetic Weiss Oscillation 91
References 95

2.7 Terahertz Spectroscopy of Nanostructures 96
K. Hirakawa
2.7.1 Introduction 96
2.7.2 Swept-Frequency THz Spectroscopy 96
2.7.3 Electronic States in Single Quantum Wire Structure 96
2.7.4 Blackbody Radiation from Hot Carriers 99
2.7.5 Summary 102
References 102

2.8 Wannier–Stark Effect in Transport 104
M. Morifuji and C. Hamaguchi
2.8.1 Wannier–Stark Effect 104
2.8.2 Zener Tunneling and Wannier–Stark States 105
2.8.3 Measurements of Zener Current through a p-i-n Diode 106
References 107

3. Quantum Hall Effect 109

3.1 Crossover from Quantum to Classical Regime 109
T. Ando
3.1.1 Bulk Versus Edge Current Picture 109
3.1.2 Edge Transport and Bulk States 112
3.1.3 Voltage Distribution 115
3.1.4 Summary 118
References 119

3.2 Edge States and Nonlocal Effects 120
S. Komiyama
3.2.1 What Is Edge Current? 120
3.2.2 Halperin's Edge Current 120
3.2.3 Local Current Distribution 122
3.2.4 Büttiker's Edge Current 123
3.2.5 Nonlocal Resistance 125
References 130

3.3 Magnetocapacitance and Edge States 132
K. Murase, S. Takaoka, and K. Oto
3.3.1 Spatial Dispersion of Edge States 132
3.3.2 Edge States Width and Magnetocapacitance 132
References 137

4. Electron-Photon Interaction in Nanostructures 139

4.1 Introduction 139
Y. Arakawa
References 139

4.2 Theory of Electron-Photon Interaction 140
A. Shimizu
4.2.1 Electron and Hole Operators in Insulating Solids 140
4.2.2 Effective-Mass Approximation 141
4.2.3 Optical Matrix Elements 146
4.2.4 Quantum States in Nanostructures 151
4.2.5 Quantum Optical Phenomena in Nanostructures 154
References 154

4.3 Electron-Photon Interaction in Microcavities 156
K. Ujihara
4.3.1 Concept of Spontaneous Emission Control 156
a. Outline 156
b. Spontaneous Emission in a Single Mode 157
c. Spontaneous Emission in Free Space 159
d. Controlled Spontaneous Emission 160
4.3.2 Experimental Results of Spontaneous Emission Control 163
4.3.3 Cavity-Polariton Effects 164
References 166

4.4 Photonic Crystals 167
T. Baba
4.4.1 Photonic Crystals and Spontaneous Emission Control 167
4.4.2 Band Structure of Photonic Crystals 169
4.4.3 Technologies of Photonic Crystals 171
References 175

4.5 Microcavity Surface Emitting Lasers 176
K. Iga
4.5.1 Overview 176
4.5.2 Technology for Low Threshold Surface Emitting Lasers 177
4.5.3 New Materials for Surface Emitting Lasers 178
References 180

4.6 Toward Lasers of the Next Generation 181
Y. Arakawa
4.6.1 Quantum Dot Lasers 181
4.6.2 Microcavity Quantum Dot Lasers 183
References 185

5. Quantum-Effect Devices 187

5.1 Introduction 187
K. Furuya
References 187

5.2 Electron-Wave Reflection and Resonance Devices ····· 188
M. Asada
5.2.1 Introduction ····· 188
5.2.2 Epitaxial Growth of $CoSi_2/CaF_2$ on Si ····· 188
5.2.3 Resonant Tunneling Transistor ····· 190
5.2.4 Observation of Hot Electron Interference ····· 191
5.2.5 Field-Effect Quantum Device ····· 192
References ····· 193

5.3 Electron-Wave Coherent Coupling Devices ····· 195
N. Tsukada
5.3.1 Coherent Coupling in Double Quantum Well ····· 195
5.3.2 Electron Directional Coupler ····· 196
5.3.3 Coherent Oscillation Devices ····· 198
5.3.4 Bloch Oscillation Devices ····· 199
5.3.5 Coherent Oscillations in ac-Field ····· 200
References ····· 202

5.4 Electron-Wave Diffraction Devices ····· 204
K. Furuya
5.4.1 Electron Wavefront and Its Manipulation ····· 204
5.4.2 Coherence of Electron Wave ····· 205
a. Phase Breaking Time Required for Interference ····· 205
b. Energy Sharpness Required for Interference ····· 206
c. Phase Breaking Time Estimated ····· 206
d. Coherence of Electron Wave ····· 207
5.4.3 Diffraction of Hot Electron Wave ····· 210
References ····· 212

5.5 Devices Using Ultimate Silicon Technology ····· 213
T. Hiramoto
5.5.1 Future of VLSI Device Technology ····· 213
5.5.2 Silicon Single-Electron Devices ····· 214
5.5.3 Integration of MOS and Single-Electron Devices ····· 218
References ····· 219

5.6 Circuit Systems Using Quantum-Effect Devices ····· 220
Y. Amemiya
5.6.1 Information Processing Architectures ····· 220
5.6.2 Binary-Decision-Diagram Circuits ····· 220
5.6.3 Local-Interaction Logic Circuits ····· 221
5.6.4 Analog Computation Systems ····· 222
5.6.5 MOBILE Circuit Systems ····· 223
5.6.6 RHET Circuit Systems ····· 224
References ····· 225

6. Formation and Characterization of Quantum Structures ···· 227

6.1 Introduction ····· 227
H. Nakashima
References ····· 227

6.2 Quantum Wires and Dots by MOCVD (I) ····· 228
T. Fukui
6.2.1 Quantum Wires on Vicinal Surfaces ····· 228
6.2.2 Quantum Dot Formation on Masked Substrates ····· 233
References ····· 236

6.3 Quantum Wires and Dots by MOCVD (II) ····· 238
Y. Arakawa
6.3.1 Quantum Wires by Selective MOCVD ····· 238
6.3.2 Quantum Dots by Selective MOCVD ····· 240
6.3.3 Quantum Dots in 2D V-Grooves ····· 241
6.3.4 Self-Assembled InGaAs Quantum Dots ····· 243
6.3.5 Use of Spinodal Phase Separation ····· 244
References ····· 245

6.4 Quantum Wires on Vicinal GaAs (110) Surfaces ····· 247
H. Nakashima
6.4.1 Introduction ····· 247
6.4.2 Step Structures ····· 247
6.4.3 AlGaAs Quantum Wires ····· 249
6.4.4 GaAs Quantum Wires ····· 250
References ····· 254

6.5 Tilted T-Shaped and (775)B Quantum Wires ····· 255
S. Hiyamizu
6.5.1 Introduction ····· 255
6.5.2 $GaAs/Al_{0.3}Ga_{0.7}As$ Tilted T-shaped QWRs ····· 255
6.5.3 Fabrication of $GaAs/Al_{0.3}Ga_{0.7}As$ Tilted T-QWRs ····· 256
a. Cathodoluminescence Measurements ····· 257
b. Calculation of Electron and Hole States in T-QWRs ····· 257
6.5.4 Naturally Formed QWRs on (775)B GaAs Substrates ····· 259
a. MBE Growth of $GaAs/(GaAs)_m(AlAs)_n$ QWRs ····· 259
b. Photoluminescence Measurements ····· 261
References ····· 262

6.6 SiGe Quantum Structures ····· 264
N. Usami and Y. Shiraki
6.6.1 Band Modification by SiGe/Si Heterostructures ····· 264
6.6.2 SiGe Quantum Wells ····· 265
6.6.3 SiGe Quantum Wires and Dots ····· 268
References ····· 271

Index ····· 273

List of Contributors

Yoshihito AMEMIYA
Department of Electronics, Hokkaido University, *North 13, West 8, Sapporo 060, Japan*

Tsuneya ANDO*
Institute for Solid State Physics, University of Tokyo, *7–22–1 Roppongi, Minato-ku, Tokyo 106, Japan*

Yasuhiko ARAKAWA*
Institute of Industrial Science, University of Tokyo, *7–22–1 Roppongi, Minato-ku, Tokyo 106, Japan*

Masahiro ASADA
Department of Electrical and Electronic Engineering, Tokyo Institute of Technology, *2–1–12 O-okayama, Meguro-ku, Tokyo 152, Japan*

Toshihiko BABA
Division of Electrical and Computer Engineering, Yokohama National University, *156 Tokiwadai, Hodogaya-ku, Yokohama 240, Japan*

Takashi FUKUI
Research Center for Interface Quantum Electronics, Hokkaido University, *North 13 West 8, Sapporo 060, Japan*

Hidetoshi FUKUYAMA
Department of Physics, University of Tokyo, *7–3–1 Hongo, Bunkyo-ku, Tokyo 113, Japan*

Kazuhito FURUYA*
Department of Electrical and Electronic Engineering, Tokyo Institute of Technology, *2–1–12 O-okayama, Meguro-ku, Tokyo 152, Japan*

Chihiro HAMAGUCHI
Department of Electronic Engineering, Osaka University, *Suita, Osaka 565, Japan*

Kazuhiko HIRAKAWA
Institute of Industrial Science, University of Tokyo, *7–22–1 Roppongi, Minato-ku, Tokyo 106, Japan*

Toshiro HIRAMOTO
Institute of Industrial Science, University of Tokyo, *7–22–1 Roppongi, Minato-ku, Tokyo 106, Japan*

Satoshi HIYAMIZU
Faculty of Engineering Science, Osaka University, *1-1 Machikaneyama, Toyonaka, Osaka 560, Japan*

Kenichi IGA
Precision and Intelligence Laboratory, Tokyo Institute of Technology, *4259 Nagatsuda, Midori-ku, Yokohama 226, Japan*

Yasuhiro IYE
Institute for Solid State Physics, University of Tokyo, *7–22–1 Roppongi, Minato-ku, Tokyo 106, Japan*
Arisato KAWABATA
Department of Physics, Gakushuin University, *1–5–1 Mejiro, Toshima-ku, Tokyo 171, Japan*
Susumu KOMIYAMA
Department of Basic Science, University of Tokyo, *3–8–1 Komaba, Meguro-ku, Tokyo 153, Japan*
Nobuya MORI
Department of Electronic Engineering, Osaka University, *Suita, Osaka 565, Japan*
Masato MORIFUJI
Department of Electronic Engineering, Osaka University, *Suita, Osaka 565, Japan*
Kazuo MURASE
Department of Physics, Osaka University, *1-1 Machikaneyama, Toyonaka 560, Japan*
Hisao NAKASHIMA*
Institute of Scientific and Industrial Research, Osaka University, *8-1 Mihogaoka, Ibaraki, Osaka 567, Japan*
Masao OGATA
Department of Basic Science, University of Tokyo, *3–8–1 Komaba, Meguro-ku, Tokyo 153, Japan*
Kenichi OTO
Department of Physics, Osaka University, *1-1 Machikaneyama, Toyonaka 560, Japan*
Akira SHIMIZU
Institute of Physics, University of Tokyo, *3–8–1 Komaba, Meguro-ku, Tokyo 153, Japan*
Yasuhiro SHIRAKI
Research Center for Advanced Science and Technology, University of Tokyo, *4–6–1 Komaba, Meguro-ku, Tokyo 153, Japan*
Sadao TAKAOKA
Department of Physics, Osaka University, *1-1 Machikaneyama, Toyonaka 560, Japan*
Seigo TARUCHA
NTT Basic Research Laboratories, *3–1 Morinosata Wakamiya, Atsugi 243, Japan*
Noriaki TSUKADA
Department of Electronics and Information Engineering, Aomori University, *2–3–1 Kobata, Aomori 030, Japan*
Kikuo UJIHARA
University of Electro-Communications, *1–5–1 Chofugaoka, Chofu, Tokyo 182, Japan*
Noritaka USAMI
Research Center for Advanced Science and Technology, University of Tokyo, *4–6–1 Komaba, Meguro-ku, Tokyo 153, Japan*

---

* *Editors*

1. Introduction – Mesoscopic Systems

1.1 Introduction

Structures of integrated circuits are approaching fundamental limits and the next generation of devices might show unexpected properties due to quantum effects and fluctuations. A long history of research has shown that when features in electronic materials approach the size of electronic wavelengths, quantum effects manifest themselves in their properties. A new research field has been developed in which we pursue understanding of basic physics associated with such quantum structures, explore their controlability, and propose new devices. The purpose of this book is to give a comprehensive description of developments in this new field of mesoscopic physics and electronics.

From the point of view of fundamental physics, mesoscopic systems have provided various basic problems, such as universal conductance fluctuations, the quantum Hall effect, and the crossover between microscopic and macroscopic regimes of quantum transport. Further, new quantum structures including quantum wires have shown fascinating new phenomena associated with ballistic transport in which electrons can move without being scattered by impurities or random potential fluctuations, such as conductance quantization, the Tomonaga–Luttinger liquid, and classical and quantum chaos. Quantum dots have provided artificial atoms in which the Coulomb interaction plays important roles, leading to various phenomena, including single-electron tunneling.

Quantum structures, particularly quantum wires and dots, are expected to be able to significantly improve optical devices such as lasers. A full confinement of photons is currently being discussed in microcavities for future realization of zero-threshold lasers in which electron-hole transitions are coupled with only a single spontaneous emission mode. The full-quantization of both electrons and photons has now become a quite important issue. Various new quantum-effect devices, which use the coherent electron wave, i.e., interference, multiple reflection, Bragg reflection, diffraction, etc., and the Coulomb blockade and charging effect, are also being pursued. The realization of these devices requires extreme control of fabrication tolerances, such as uniformity, defects, and damage-free structures. A variety of techniques for the formation of quantum structures are being developed.

In Chap. 1 a short introduction is given on quantum transport in such mesoscopic systems, including various characteristic length scales, Landauer's formula, new phenomena such as conductance fluctuations and Aharonov–Bohm effects, ballistic transport, and the Coulomb blockade and single-electron tunneling. In Chap. 2 various investigations on quantum transport in

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

mesoscopic systems are discussed. Some topics on the quantum Hall effect are described in Chap. 3. Chapter 4 explores possibilities of applications of such mesoscopic systems as new optical devices. Chapter 5 contains similar discussions in the field of electronics devices. Some examples of new developments in the formation and characterization of semiconductor structures are presented in Chap. 6.

1.2 Length Scales Characterizing Mesoscopic Systems

1.2.1 Fermi Wavelength

One important length scale characterizing mesoscopic systems is the Fermi wavelength $\lambda_F = 2\pi/k_F$, where k_F is the Fermi wave number. At zero temperature, electrons occupy states specified by the wave vector $\mathbf{k}$ with $|\mathbf{k}| \leq k_F$. The Fermi wave vector is related to the electron density n through

$$n = \begin{cases} \dfrac{2}{(2\pi)^3}\dfrac{4\pi}{3}k_F^3 & (d=3), \\ \dfrac{2}{(2\pi)^2}\pi k_F^2 & (d=2), \\ \dfrac{2}{2\pi}k_F & (d=1), \end{cases} \tag{1.2.1}$$

where d is the system dimension (0D for $d=0$, 1D for $d=1$, 2D for $d=2$, and 3D for $d=3$) and the factor 2 comes from the electron spin. In typical metals such as Cu and Ag, the Fermi wavelength is of the order of a few angstrom and in semiconductors such as 2D systems realized in GaAs/AlGaAs heterostructures we have $\lambda_F \sim 400$ Å for the electron concentration $n \sim 3\times10^{11}$ cm^{-2}.

1.2.2 Mean Free Path

In external fields the time development of the electron system is governed by the Boltzmann transport equation [1], which introduces the important length scale called the mean free path Λ. The mean free path is the average orbit-length covered by an electron before being scattered into a different wave-vector direction. At sufficiently low temperatures, the transport is determined by electrons in the vicinity of the Fermi energy E_F and therefore we have $\Lambda = v_F\tau$ where v_F is the Fermi velocity and τ is the relaxation time. The conductivity is usually written as

$$\sigma = \frac{ne^2\tau}{m}, \tag{1.2.2}$$

where m is the electron mass.

The relaxation time is different from the conventional scattering time defined as the inverse of the scattering probability for an electron to be scattered into different states. In fact, scattering that does not change the direction of the electron motion does not give rise to reduction in the current. For a scatterer having a slowly varying potential, the probability of small-angle scattering becomes much larger than that of large-angle scattering, leading to a large difference between the transport relaxation time $\tau_{\rm tr}$ and τ. The relaxation time appearing in the above equation is actually $\tau_{\rm tr}$ instead of τ.

1.2.3 System Size

Consider a system confined into a box with size $L_x \times L_y \times L_z$ $(L_x < L_y < L_z)$

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

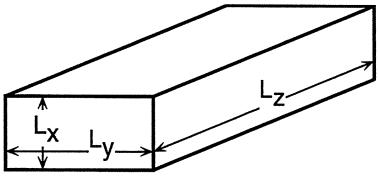


Fig. 1.2.1 A system confined into a box with size L_x, L_y, and L_z. The system varies among 3D, 2D, 1D, and 0D depending on the relative magnitude of the sizes and the Fermi wavelength

as shown in Fig. 1.2.1. The energy of this system is given by

$$\varepsilon(n_x, n_y, n_z) = \frac{\hbar}{2m}\left[\left(\frac{\pi n_x}{L_x}\right)^2 + \left(\frac{\pi n_y}{L_y}\right)^2 + \left(\frac{\pi n_z}{L_z}\right)^2\right], \tag{1.2.3}$$

where n_x, n_y, and n_z are integers. The system can roughly be categorized as follows depending on the relative magnitude of L_x, L_y, L_z, and λ_F.

(1) $\lambda_F \ll L_x < L_y < L_z$: 3D (bulk)
(2) $\lambda_F \sim L_x < L_y < L_z$: quasi 2D (thin films)
(3) $L_x < \lambda_F \ll L_y < L_z$: 2D (MOSFETs, heterostructures)
(4) $L_x < L_y \sim \lambda_F \ll L_z$: quasi 1D (quantum wires)
(5) $L_x < L_y < \lambda_F \ll L_z$: 1D
(6) $L_x < L_y < L_z < \lambda_F$: 0D (quantum dots)

When many electrons are confined into a box with size $L \sim L_x \sim L_y \sim L_z$, we have to consider the charging energy E_c due to the Coulomb interaction among electrons in addition to the quantum mechanical quantization energy mentioned above. The typical charging energy required for adding or removing an electron is given by $E_c = e^2/L$. This charging energy is more important than the quantization energy for systems with a large size and becomes less important with the decrease in the size. The crossover occurs around $L \sim \sqrt{\lambda_F a_B}$, where a_B is the Bohr radius. The charging energy plays an important role in the single-electron tunneling through an isolated metallic island or a quantum dot.

1.2.4 Thermal Diffusion Length and Thouless Energy

An electron makes a ballistic motion without being scattered on a scale smaller than the mean free path Λ, but it makes a diffusive motion like a random walk on a scale larger than Λ. This diffusive motion is characterized by the diffusion coefficient D, given by

$$D \sim v_F^2 \tau = \frac{\Lambda^2}{\tau} = \Lambda v_F. \tag{1.2.4}$$

Because of scattering an electron specified by wave vector $\mathbf{k}$ has a broadening in the $\mathbf{k}$ space corresponding to the energy $\hbar/\tau$. This broadening is given by $\Delta k \sim 1/\Lambda$ through $(\partial\varepsilon(\mathbf{k})/\partial k)\Delta k \sim \hbar/\tau$. Note, however, that an electron is in an eigenstate having a definite energy and wave function. This

wave function is no longer given by a plane wave but shows a complicated spatial variation corresponding to the broadening in the **k** space. The phase coherence of each electron is not destroyed by scattering from impurities.

At nonzero temperatures, transport quantities are determined by a thermal average for states having an energy lying in the region of width of the order of $k_B T$ around the Fermi energy, where T is the temperature and k_B is the Boltzmann constant. This means that the electron loses its phase memory after a time interval of the order of $\hbar/k_B T$. The size corresponding to this time interval is given by $L_T = \sqrt{D\hbar/k_B T}$ when the electron makes a random walk characterized by the diffusion constant D. This length is called the thermal diffusion length.

For a system with finite size L, we can determine an energy scale given by $E_T = D\hbar/L^2$ which is called the Thouless energy [2,3]. Two states having energies close to each other are correlated to each other, but those having energies which are separated from each other by more than E_T can be regarded as almost independent of each other. The Thouless energy plays an important role also in Anderson localization.

1.2.5 Phase Coherence Length

So far we have considered the electron system in which the electron motion is described by a single Schrödinger equation. In actual systems, however, electrons interact with many other degrees of freedom and exchange their energies. For example, the electron moves from one eigenstate to another due to scattering from other electrons or lattice vibrations. The phase of the electron is destroyed by such dynamical interactions and the distance over which an electron maintains its phase memory is called the phase coherence length. For the phase relaxation time τ_ϕ determined by inelastic scattering, the phase coherence length is given by $L_\phi = \sqrt{D\tau_\phi} = \Lambda\sqrt{\tau_\phi/\tau}$ in the diffusive regime.

The electron-electron and electron-phonon scattering decrease at low temperatures. Usually $\tau_\phi \propto T^{-p}$ [4], where $p = 2$ in the case of electron-electron scattering according to the Fermi liquid theory. In the presence of strong disorder, p becomes smaller. For scattering from phonons, the power p is larger. The phase coherence increases with decreasing temperature in proportion to $L_\phi \propto T^{-p/2}$. In the ballistic regime (see below), we have $L_\phi \sim v_F \tau_\phi$.

1.2.6 Diffusive Regime and Ballistic Regime

In mesoscopic systems, sizes can be comparable to or smaller than the phase coherence length. Usually observed quantities in macroscopic systems are the average over ensembles of various systems, since each system consists of many subsystems described by independent Schrödinger equations. When the phase coherence becomes comparable to the system size, however, the whole system is governed by a single Schrödinger equation and observed quantities are not averaged out and can be different from system to system.

A mesoscopic system is divided into the diffusive and ballistic regimes depending on the relative magnitude of the mean free path Λ and the system

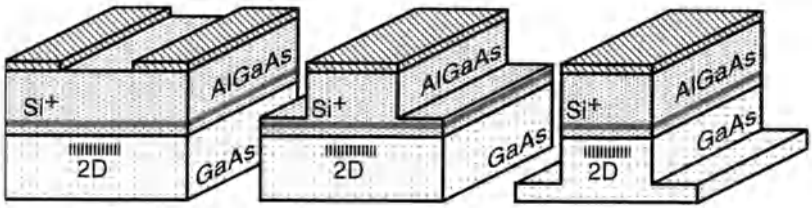


Fig. 1.2.2 A schematic illustration of quantum wires fabricated by (left to right) split-gate, shallow mesa etching, and deep mesa etching

size L. In the diffusive regime where $L \gg \Lambda$, the transport can be understood as a diffusive process and is essentially independent of the form of the system. In the ballistic regime where $L \ll \Lambda$, on the other hand, the electron makes a ballistic motion inside the system and the system boundary plays the role of scatterer instead of impurities.

Metallic wires or dots are in the diffusive regime, as the mean free path is of the order of ~ 100 Å and the system size is larger. Further, the Fermi wavelength (1∼2 Å) is much smaller than the system size and therefore the quantization of energy levels is not important except at very low temperatures where the separation between nearest neighbor energy levels becomes comparable to the temperature. In metallic dots, therefore, the most important energy scale is the charging energy due to the Coulomb interaction.

In semiconductor heterostructures having a supreme quality, the mean free path can be as large as 50 μm and ballistic transport occurs. Further, the Fermi wavelength is large (300∼500 Å) and can be comparable to the system size. Therefore, the quantization due to the system confinement usually plays a very important role.

1.2.7 Quantum Wires, Dots, and Antidots

There are various ways to fabricate quantum wires and dots in semiconductors. Figure 1.2.2 is a schematic illustration of wires made by the split gate and shallow and deep mesa etching techniques. In such wires, 2D electrons are confined in a space far from the structure made by lithography by an electrostatic potential. Consequently, the confinement is not so strong and at the same time the potential fluctuations arising from those of the lithographic structure are less important.

It is necessary to perform a self-consistent calculation to determine the confinement potential and energy level structure of such quantum wires. Several calculations have been reported [5–8], which seem to show that the potential is nearly parabolic like a harmonic potential in sufficiently narrow wires but has a flat central part and a rapidly varying barrier part at the edge of the wire. The potential can be regarded as almost parabolic, if the part coming from electrons confined in the wire is completely neglected.

Quantum dots can be fabricated using 2D systems similarly as for 2D lithographic structures. The number of electrons confined in a single dot ranges from 20 to 300 in the case of dots fabricated by split gates or in-plane gates. The number of electrons can be much smaller down to a few

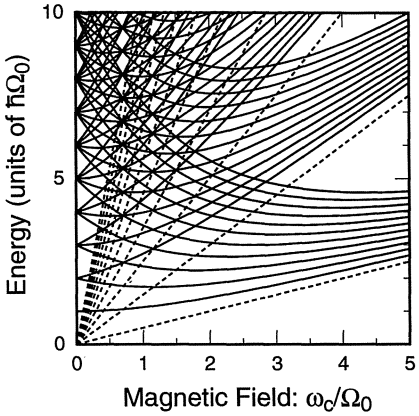


Fig. 1.2.3 Energy levels of a quantum dot in magnetic fields. The dotted lines denote the energy of the Landau levels given by $(N+1/2)\hbar\omega_c$ with $N=0, 1, 2, \ldots$

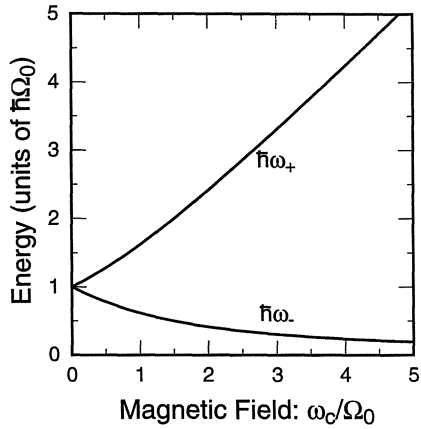


Fig. 1.2.4 Resonance energies versus the magnetic field strength in a quantum dot confined by a parabolic confining potential. In the limit of high magnetic fields, E_+ approaches $\hbar\omega_c$ (cyclotron energy) while E_- which corresponds to an edge mode vanishes

for dots fabricated more recently as discussed in Sect. 2.4. In addition to the techniques based on structured gates there has been rapid development in various new methods of creating quantum wires and dots, which will be discussed extensively in Chap. 6.

The Schrödinger equation for an electron confined by a harmonic potential in a magnetic field has been solved [9]. The energy eigenvalues are given by

$$E_{nm} = (2n+|m|+1)\sqrt{(\hbar\Omega_0)^2+\left(\frac{\hbar\omega_c}{2}\right)^2} + \frac{1}{2}m\hbar\omega_c, \tag{1.2.5}$$

where Ω_0 is the frequency characterizing the confining potential, $n=0,1,\ldots$ is the radial quantum number, and $m=0,\pm1,\pm2,\ldots$ is the azimuthal quantum number (Fock–Darwin states). Dipole allowed transitions have energies

$$E_\pm = \sqrt{(\hbar\Omega_0)^2+\left(\frac{\hbar\omega_c}{2}\right)^2} \pm \frac{1}{2}\hbar\omega_c. \tag{1.2.6}$$

These are shown in Figs. 1.2.3 and 1.2.4.

The energy level structure containing many electrons becomes quite complicated because of an interplay of the confining potential and the mutual Coulomb interaction. In spite of such complications, however, experimentally observed FIR spectra resemble that shown in Fig. 1.2.4. This can be understood in terms of the so-called generalized Kohn's theorem stating that optically active collective modes of many interacting electron systems in a parabolic confinement are those of a single particle with mass and charge given by the total number of electrons [10,11]. The spectra deviate from

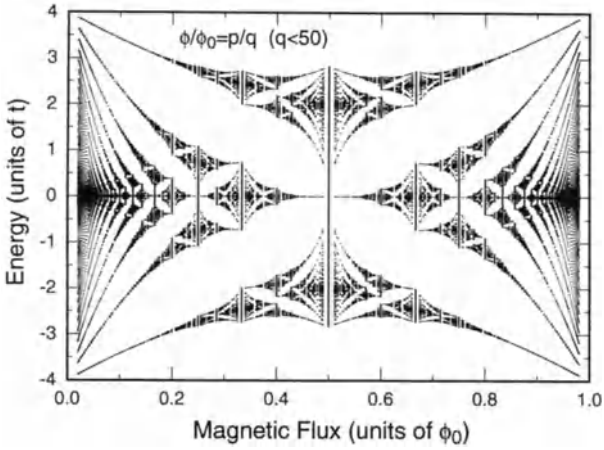


Fig. 1.2.5 Energy levels of a square lattice in magnetic fields. The spectrum exhibits a characteristic self-similar structure

that shown in Fig. 1.2.4 only when the confinement becomes different from a harmonic potential, for example, due to couplings between neighboring dots.

The lateral superlattice can be realized by a periodic array of quantum-wire structures. Usually, 2D electrons are affected by periodic electric potential, but they are also influenced by periodic magnetic fields. This will be discussed in Sect. 2.6. When periodic potentials are introduced in both x and y directions, we have a simple form of man-made crystal. One typical system is known as an antidot lattice, in which periodic circular barrier potentials are introduced into a 2D system. The antidot lattice is particularly interesting in magnetic fields, because the energy spectrum is known to become self-similar and fractal.

Figure 1.2.5 shows an example of Hofstadter's butterfly spectrum calculated in a square lattice in the presence of magnetic flux ϕ [12]. The spectrum is periodic as a function of $\nu = \phi/\phi_0$ where $\phi_0 = ch/e$ is the magnetic flux quantum. For $0 < \nu < 1$, the spectrum is self-similar. It is difficult to observe such a butterfly spectrum in a real crystal because the required magnetic field is extremely large. Further, the classical electron motion in this system is known to be chaotic. Interesting magnetotransport phenomena in antidot lattices are discussed also in Sect. 2.5.

1.2.8 Anderson Localization

When the mean free path becomes short and comparable to the Fermi wavelength, multiple scattering from scatterers becomes important, the electron wave function is spatially localized, and the conductivity vanishes at zero temperature. Such a metal-insulator transition due to disorder is called the Anderson localization [13]. In the localized states, the wave function decays exponentially away from a certain localization center, i.e., $\psi(r) \sim \exp(-r/\xi)$ or $\sim \exp(-\alpha r)$, where r is the distance from the localization center, ξ is called localization length, and $\alpha = \xi^{-1}$ is the inverse localization length.

Anderson localization depends strongly on the dimensionality [13–17]. This can most easily be understood in terms of a scaling argument [4]. Consider a d-dimensional cube with size L. By combining different cubes of the same size, we can construct a cube with size $2L$. By repeating such a procedure starting from a small system, we can obtain energy levels and wave functions of a system having arbitrary size. The combination of two cubes causes mixing of energy levels. Couplings of levels closest to each other in energy are most important. Let $V(L)$ be the resonance energy of such energy levels and $W(L)$ their energy difference. We can safely assume $W(L)$ to be the average energy separation, i.e., $V(L) \sim 1/L^d D(E)$ with $D(E)$ being the density of states.

The ratio $g(L) \equiv V(L)/W(L)$ is called the Thouless number [2,3]. If states are localized, $g(L)$ decreases exponentially with the increase of the system size L, i.e., $g(L) \propto \exp(-\alpha L)$. When states are extended, it approaches a positive constant. Let $\tau(L)$ be the time during which an electron stays in a cube with size L. Then, the diffusion coefficient is given by $D^*(L) \sim L^2/\tau(L)$. The uncertainty relation $\tau(L)V(L) \sim \hbar$ and the Einstein relation lead to

$$\sigma = e^2 D^*(L) D(E_F) \sim \frac{e^2}{\hbar}\frac{V(L)}{W(L)} \sim \frac{e^2}{\hbar} g(L)\, L^{2-d}, \tag{1.2.7}$$

where σ is the conductivity and $D(E_F)$ is the density of states at the Fermi level. This means that $(e^2/\hbar)g(L)$ is the conductance of the cube with size L.

In the single-parameter scaling theory, we assume that the Thouless number of a larger system is determined by that of a smaller system alone, i.e.,

$$g(bL) = f[g(L), b], \tag{1.2.8}$$

with $b > 1$. By differentiating this equation with respect to b and taking the limit $b \to 1$, we get the scaling relation:

$$\frac{d \ln g(L)}{d \ln L} = \frac{1}{g(L)} \frac{\partial f}{\partial b}\bigg|_{b=1} \equiv \beta[g(L)]. \tag{1.2.9}$$

The function $\beta(g)$ has the following properties: $\beta \sim g$ for $g \ll 1$ and $\beta \sim d-2$ for $g \gg 1$. It is a reasonable assumption that $beta(g)$ increases with g and has no singularity. This immediately leads to the conclusion that $g < 0$ for $d \leq 2$, meaning that states are all localized in 1D and 2D. In fact, in 1D systems, states are all known to be localized exponentially, however small the randomness may be [18]. In 3D systems, a metal-insulator transition occurs at a certain critical randomness when the degree of the disorder is varied. 2D systems are marginal and the situation varies dependent on the symmetry.

In the absence of a magnetic field, where the time reversal symmetry is present, the Hamiltonian is always chosen as a real symmetric matrix and the corresponding wave functions are real. In this orthogonal case, the states are believed to be localized. There are two regimes. In the weak localization regime, the wave function varies only logarithmically and effects of disorder can be considered as a small perturbation. In the strong localization regime,

the wave function exhibits an exponential decay.

In the presence of a magnetic field, the time reversal symmetry is broken, the Hamiltonian is given by a complex unitary matrix, and the wave function becomes complex. In this unitary case, the localization effect becomes weaker in weak magnetic fields, leading to a negative magnetoresistance. In high magnetic fields, where Landau levels are well separated from each other, delocalized states appear in the vicinity of each broadened Landau level. The presence of current-carrying extended states is a key ingredient in the integral quantum Hall effect, as will be discussed in Sect. 3.1.

In the presence of a strong spin-orbit interaction, the Hamiltonian is given by a quarternion matrix and the states are always doubly degenerate in the absence of a magnetic field (the Kramers degeneracy). In this symplectic case, the localization effect becomes further weakened and there seems to be a metal-insulator transition as a function of the degree of disorder.

References

1 See, for example, *Solid State Physics*, N.W. Ashcroft and N.D. Mermin (HRW International Editions, 1976), Chapt. 16
2 J.T. Edwards and D.J. Thouless, J. Phys. C **5**, 807 (1972)
3 D.C. Licciardello and D.J. Thouless, J. Phys. C **8**, 4157 (1975); Phys. Rev. B **35**, 1475 (1975); J. Phys. C **11**, 925 (1978)
4 E. Abrahams, P.W. Anderson, D.C. Licciardello, and T.V. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979)
5 S.E. Laux and F. Stern, Appl. Phys. Lett. **59**, 91 (1986)
6 A. Kumar, S.E. Laux, and F. Stern, Appl. Phys. Lett. **54**, 1270 (1989); Phys. Rev. B **42**, 5166 (1990)
7 A. Kumar, S.E. Laux, and F. Stern, Phys. Rev. B **42**, 5166 (1990)
8 T. Suzuki and T. Ando, J. Phys. Soc. Jpn. **62**, 2986 (1993); Physica B **201**, 345 (1994); **227**, 46 (1996)
9 V. Fock, Z. Phys. **47**, 446 (1928)
10 L. Brey, N. Johnson, and B. Halperin, Phys. Rev. B **40**, 10647 (1989)
11 P. Maksym and T. Chakraborty, Phys. Rev. Lett. **65**, 108 (1990)
12 D.R. Hofstadter, Phys. Rev. B **14**, 2239 (1976)
13 P.W. Anderson, Phys. Rev. **109**, 1492 (1958)
14 *Anderson Localization*, ed. Y. Nagaoka and H. Fukuyama (Springer, Berlin, 1982)
15 G. Bergmann, Phys. Rev. **107**, 1 (1984)
16 P.A. Lee and T.V. Ramakrishnan, Rev. Mod. Phys. **57**, 287 (1985)
17 *Anderson Localization*, ed. T. Ando and H. Fukuyama (Springer, Berlin, 1988)
18 V.L. Berezinskii, Sov. Phys. JETP **38**, 620 (1974)

1.3 Landauer's Formula

1.3.1 Conductance and Transmission Probability

Landauer derived the relation between the conductance of a one-dimensional (1D) wire and the transmission and reflection probabilities at the Fermi level [1]. Let us consider the system schematically illustrated in Fig. 1.3.1. A wire is connected at both ends to an ideal wire which is infinitely long and eventually connected to an electron reservoir. The ideal reservoir satisfies the following two conditions:

(1) All incident electrons are absorbed by the reservoir irrespective of their energy and phase.
(2) It constantly provides electrons with energy below chemical potential μ. The energy and phase of these electrons are independent of those of absorbed electrons.

Let μ_1 and μ_2 represent the chemical potential of the reservoir 1 and 2, respectively, and T and R be the transmission and reflection probability, respectively, of the wire. The conservation of the number of particles requires $T+R=1$. Further, let v be the electron velocity in the ideal lead and $\partial n_+/\partial E = 1/\pi\hbar v$ be the density of states for electrons moving from left to right. The total current flowing across the system is given by

$$I = (-e)v\frac{\partial n_+}{\partial E}T(\mu_1-\mu_2). \tag{1.3.1}$$

Because the voltage difference between 1 and 2 is $(-e)V_{21}=\mu_1-\mu_2$, the *two terminal* conductance measured between 1 and 2 is given by

$$G = \frac{I}{V_{21}} = \frac{e^2}{\pi\hbar}T. \tag{1.3.2}$$

This relation between the conductance and the transmission probability is frequently used as the basis of understanding the transport in mesoscopic systems.

The original version of Landauer's formula is slightly more complicated, because the above conductance is not that of the wire but that of the system containing the ideal leads. In fact, this conductance does not become infinite even for a perfect conducter corresponding to vanishing reflection probability, $R=0$ or $T=1$. Usually, the conductance represents the ratio of the current to the voltage applied across the system. This voltage is the difference of the effective Fermi level of the ideal leads. When $T=0$, each ideal lead is in equilibrium with the connected reservoir and $(-e)V=\mu_1-\mu_2$. When $T=1$, on the other hand, we have $(-e)V=0$. For general T and R

$$(-e)V = R(\mu_1-\mu_2). \tag{1.3.3}$$

Therefore, the final expression of the conductance is

$$G = \frac{e^2}{\pi\hbar}\frac{T}{1-T}. \tag{1.3.4}$$

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

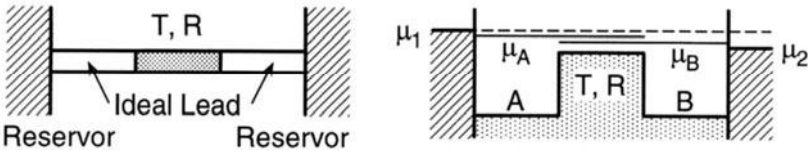


Fig. 1.3.1 A one-dimensional wire connected at both ends by an ideal lead to reservoirs and its schematic energy diagram

In analyzing the transport in mesoscopic systems, the formula (1.3.2) and its multi-channel and multi-terminal versions are more appropriate than the original formula (1.3.4).

The effective chemical potential μ_A of the left ideal lead is determined by the condition that the number of electrons having an energy above μ_A and that of holes below μ_A should be the same. They are given by

$$(1+R)\frac{\partial n_+}{\partial E}(\mu_1-\mu_A) \quad \text{and} \quad T\frac{\partial n_+}{\partial E}(\mu_A-\mu_2), \tag{1.3.5}$$

respectively. Equating these, we have

$$\mu_A = \frac{1}{2}\,[(1+R)\mu_1+(1-R)\mu_2]. \tag{1.3.6}$$

Similarly, the number of electrons having an energy above μ_B and that of holes below μ_B in the right lead are given by

$$T\frac{\partial n_+}{\partial E}(\mu_1-\mu_B) \quad \text{and} \quad 2\frac{\partial n_+}{\partial E}(\mu_B-\mu_2) - T\frac{\partial n_+}{\partial E}(\mu_B-\mu_2), \tag{1.3.7}$$

respectively. This gives

$$\mu_B = \frac{1}{2}\,[(1-R)\mu_1+(1+R)\mu_2], \tag{1.3.8}$$

Equations (1.3.6) and (1.3.8) give the relation $\mu_A-\mu_B=R(\mu_1-\mu_2)$ leading to (1.3.3).

Landauer's formula can be derived in various other ways and has been extended to systems with higher dimensions having many current carrying channels [2–5]. It has been derived based on the Kubo formula and there have been some discussions on the relation between (1.3.2) and (1.3.4) [6–9].

1.3.2 Some Applications

a. Universal Conductance Fluctuations

The universality of the conductance fluctuations has been shown first by perturbational calculation based on the Kubo formula, as discussed later in this chapter. This can be understood as follows in terms of Landauer's formula, as has been shown by Imry [10].

Let us consider a hypercube with size L in dimension d. The conductance

is written as

$$G = \frac{e^2}{\pi\hbar}\sum_{i,j}^{N}|t_{ij}|^2, \tag{1.3.9}$$

where $N \sim (k_F L)^{d-1}$ is the total number of the incoming and outgoing channels. It is a reasonable assumption that $|t_{ij}|$ itself does not vary so much from channel to channel in the diffusive regime. The most important point is that the $|t_{ij}|$ for different channels are correlated to each other because the system is described by a single Schrödinger equation. Let N_c be the effective number of channels independent of each other and t_0 be the typical value of $|t_{ij}|$. Then, the conductance and its fluctuation are given by

$$G \sim \frac{e^2}{\pi\hbar}N_c^2 t_0^2, \quad \Delta G \sim \frac{e^2}{\pi\hbar}N_c t_0^2. \tag{1.3.10}$$

Using the usual expression for the conductivity,

$$\sigma = \frac{ne^2\tau}{m} \sim \frac{e^2}{\pi\hbar}k_F^{d-1}\Lambda, \tag{1.3.11}$$

with Λ being the mean free path, we have

$$G = \sigma L^{d-2} = \frac{e^2}{\pi\hbar}(k_F L)^{d-1}\frac{\Lambda}{L}, \tag{1.3.12}$$

and

$$\Delta G \sim \frac{G}{N_c} \sim \frac{e^2}{\pi\hbar}\frac{N}{N_c}\frac{\Lambda}{L}. \tag{1.3.13}$$

Because $L \gg \Lambda$, the fluctuation would be much smaller than the average if the effective channel number N_c is of the same order as the total channel number N.

Let $\exp(-2\alpha L)$ denote an eigenvalue of the $N \times N$ matrix $(|t_{ij}|^2)$. Then, α distributes in the region $0 < \alpha < \Lambda^{-1}$. Because of repulsion between eigenvalues (corresponding to the strong correlation among $|t_{ij}|$ of different channels) they distribute almost uniformly. Since the eigenvalues satisfying the condition $\alpha L \lesssim 1$ contribute to the conduction, the effective channel number is given by $N_c \sim (\Lambda/L)N$. The substitution of this into the above equation immediately leads to the universality of the fluctuations, $\Delta G \sim e^2/\pi\hbar$.

b. Conductance Quantization

It is quite straightforward to show the quantization of the conductance across a quantum point contact in the adiabatic limit where the confining potential varies slowly within the Fermi wavelength. In this case, scattering between different channels can be neglected completely and the confining potential is represented by an effective one-dimensional potential for each channel. For a confinement in the y direction into a well with width $W(x)$ slowly varying

in the x direction, for example, the effective potential for the jth channel ($j=1,2,\ldots$) is given by

$$V_j(x) = \frac{\hbar^2}{2m}\left[\frac{\pi j}{W(x)}\right]^2. \tag{1.3.14}$$

When the maximum value of the effective potential is below the Fermi level ($V_j^{\max} < E_F$), an electron passes through the point contact with probability unity. When the maximum is above the Fermi level ($V_j^{\max} > E_F$), an electron is reflected back with probability unity. Therefore, the multi-channel version of Landauer's formula immediately leads to the quantization of conductance into integral multiples of $e^2/\pi\hbar$, where the integer is the number of 1D channels below the Fermi level at the point contact.

When the potential varies abruptly in comparison with the Fermi wavelength, reflection becomes appreciable and leads to an oscillation of the conductance due to interference effects. Many calculations of the conductance have been reported, some of which seem to suggest that the quantization does not necessarily require the adiabatic limit [11]. Transport quantities other than the conductance have also been studied [12–16].

References

1 R. Landauer, IBM J. Res. Dev. **1**, 223 (1957); Philos. Mag. **21**, 863 (1970).
2 P.W. Anderson, D.J. Thouless, E. Abrahams, and D.S. Fisher, Phys. Rev. B **22**, 3519 (1980)
3 P.W. Anderson, Phys. Rev. B **23**, 4828 (1981)
4 M.Y. Azbel, J. Phys. C **14**, L225 (1981)
5 M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Phys. Rev. B **31**, 6207 (1985)
6 E.N. Economou and C.M. Soukoulis, Phys. Rev. Lett. **46**, 618 (1981)
7 D.S. Fisher and P.A. Lee, Phys. Rev. B **23**, 6851 (1981)
8 D.C. Langreth and E. Abrahams, Phys. Rev. B **24**, 2978 (1981)
9 D.J. Thouless, Phys. Rev. Lett. **47**, 972 (1981)
10 Y. Imry, Europhys. Lett. **1**, 249 (1986)
11 See, for example, T. Ando, Phys. Rev. B **44**, 8017 (1991) and references cited therein
12 U. Sivan and Y. Imry, Phys. Rev. B **33**, 551 (1986)
13 P. Streda, J. Phys. Conden. Matter **1**, 1025 (1989)
14 P.N. Butcher, J. Phys. Condens. Matter **2**, 4869 (1990)
15 C. Proetto, Phys. Rev. B **44**, 9096 (1991); Solid State Commun. **80**, 909 (1991)
16 Y. Okuyama, T. Sakuma, and N. Tokuda, Surf. Sci. **263**, 258 (1992)

1.4 Fluctuations and Aharonov–Bohm Effect

1.4.1 Aharonov–Bohm Effect

The most remarkable phenomenon resulting from the wave nature of electrons is the Aharonov–Bohm (AB) effect [1,2], which not only reveals that electrons are waves but also demonstrates that the vector potential, but not the magnetic field, plays basic roles.

Consider an interference of electron waves at point Q coming through a slit P and propagating along paths 1 and 2 encircling a flux Φ as shown in Fig. 1.4.1. The electron wave function is described by the Schrödinger equation

$$\left[\frac{1}{2m}(\mathbf{p}+\frac{e}{c}\mathbf{A})^2\right]\psi(\mathbf{r}) = E\psi(\mathbf{r}) \ . \tag{1.4.1}$$

Here, m and $-e$ are the mass and charge of an electron and $\mathbf{A}(\mathbf{r})$ is the vector potential, respectively, and it is assumed that there are no potential fluctuations along these paths. It is first noted that the wave function propagating along the path i is given as follows

$$\begin{aligned} \psi_i(\mathbf{r}) &= \exp\left[-i\theta_i(\mathbf{r})\right]\psi_i^0(\mathbf{r}), \\ \theta_i(\mathbf{r}) &= \frac{2\pi}{\Phi_0}\int_P^{\mathbf{r}} \mathbf{A}(s)\cdot d\mathbf{s}_i, \end{aligned} \tag{1.4.2}$$

where $\Phi_0 = hc/e$ is the flux quantum and $\psi_i^0(\mathbf{r})$ is that in the absence of the vector potential and $d\mathbf{s}_i$ denotes the integration along the path i. Then the total amplitude of the electron wave at the point Q is given by

$$|\psi_i(Q)+\psi_2(Q)| = |\psi_1(Q)|^2 + |\psi_2(Q)|^2 + 2\mathrm{Re}\psi_1^*(Q)\psi_2(Q) \tag{1.4.3a}$$

$$\simeq\ 2\ |\psi_1^0(Q)|^2\{1+\cos[\xi_E(Q)+\theta_1-\theta_2]\}, \tag{1.4.3b}$$

with

$$\theta_1-\theta_2 = \frac{2\pi}{\Phi_0}\left[\int_P^Q \mathbf{A}\cdot d\mathbf{s}_1 - \int_P^Q \mathbf{A}\cdot d\mathbf{s}_2\right] = \frac{2\pi}{\Phi_0}\oint \mathbf{A}(s)\cdot d\mathbf{s} = 2\pi\frac{\Phi}{\Phi_0} \ . \tag{1.4.4}$$

In passing from (1.4.3a) to (1.4.3b) we assumed $|\psi_1(Q)|^2 \sim |\psi_2(Q)|^2$, which is valid for propagating plane waves, and $\xi_E(Q)$ is defined by $\psi_1^{0*}(Q)\psi_2^0(Q) = |\psi_1^0|^2 e^{i\xi_E(Q)}$. In (1.4.4) the Stokes theorem is noted, and (1.4.3b) implies that the total amplitude at point Q oscillates periodically with respect to the the magnetic flux enclosed by the two paths due to the interference effect. This is called the Aharonov–Bohm effect [1]. Note that electrons are affected by the magnetic field through the vector potential even though they are propagating in the region where there is actually no magnetic field, indicating that the vector potential has nonlocal effects. This has been demonstrated explicitly in experiments by Tonomura [2], who employed electron holography.

A beautiful example of the Aharonov–Bohm effect is the quantization of the flux Φ trapped within a superconducting ring as shown in Fig. 1.4.2, $\Phi = (hc/2e) \times$ (integer). Here $2e$ instead of e reflects the Cooper pair condensate in the superconductors.

The first observation of the AB effect in normal metal rings was reported by Webb [3] in an Au wire with 825 nm diameter and 49 nm width as shown

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

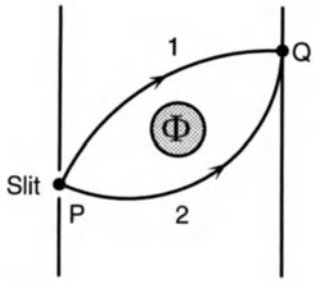


Fig. 1.4.1 Schematic representation of the Aharonov–Bohm effect

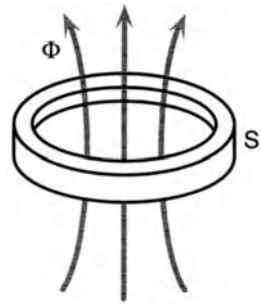


Fig. 1.4.2 Flux quantization by a superconducting ring

in Fig. 1.4.3. In this case, the magnetic field penetrates the wire region, in contrast to the theoretical discussions associated with Fig. 1.4.1, but this effect is not serious since the magnitude of the flux in the wire relative to that encircled by the wire is small [4]. Actually, in earlier experiments [5] on a wire of 320 nm diameter and 50 nm width Umbach et al. [5] obtained the data shown in Fig. 1.4.4, where there is no clear AB oscillation but only apparently irregular variation with respect to the magnetic field. This irregular variation was reported to be "reproducible but sample-specific," which are the characteristic features of fluctuations to be discussed in the following, called the magneto-fingerprint. From (1.4.3b) it is actually seen that the AB oscillation observed in Fig. 1.4.3 is a part of the fluctuations, since in this experiment there are many channels (N_c) contributing to the conductance g, and $\xi_E(Q)$ varies randomly among these channels,

$$g \sim g_0 \sum_{\alpha}^{N_c} \left[1 + \cos\left(\xi_\alpha + \frac{2\pi\Phi}{\Phi_0}\right)\right] = \bar{g}\left[1 + \frac{C_1}{\sqrt{N_c}} \cos\left(\zeta + \frac{2\pi\Phi}{\Phi_0}\right)\right] \qquad (1.4.5)$$

where $g_0 \sim e^2/h$, $\bar{g} = g_0 N_c$, C_1 is a constant of order unity, and ζ is some angle. This characteristic feature of the AB oscillation as due to the fluctuations has been explicitly demonstrated by studying the change of the amplitude of the AB oscillation of a variable number of rings arrayed in series [6].

As seen, the AB effect requires the persistence of the phase of waves propagating along the different paths. Hence, once dephasing processes are present, the amplitude of AB oscillation is reduced. A typical example of such a dephasing process is the finite temperature effect due to inelastic scattering leading to the suppression of the amplitude proportional to $\exp(-\pi R/L_\varphi)$ where L_φ is the inelastic scattering length.

Besides the AB oscillation, there exists another type of oscillation seen in resistivity, called Altshuler–Aronov–Spivak (AAS) oscillation [7], which persists even in macroscopic systems in contrast to the former. AAS oscillation is due to the modulation by the magnetic field of the interference between two waves, which are related to each other by the time-reversal symmetry. The crossover between AB oscillation and AAS oscillation as the number of ensembles is varied has been clarified by Stone and Imry [8].

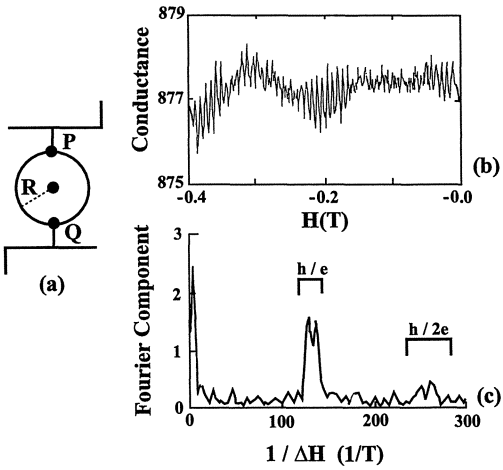


Fig. 1.4.3 Observed AB oscillation in a normal metal ring [3]

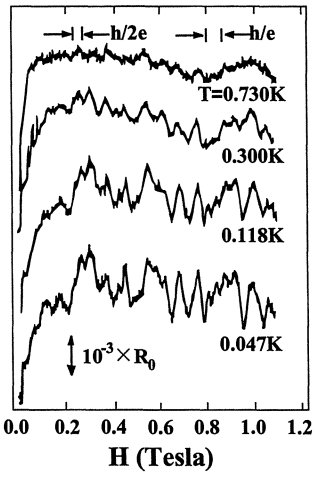


Fig. 1.4.4 Observed conductance fluctuations in a thick normal metal ring [5]

1.4.2 Universal Conductance Fluctuations

A manifestation of the remarkable feature of quantum interference in transport in solids is the universal conductance fluctuations (UCF) discovered by Altshuler [9] and by Lee and Stone [10]. The conductance g of the d-dimensional system with the size L^d is given by $g = \sigma L^{d-2}$ with σ the conductivity. If g is due to the sum of various constituents g_α, i.e., $g = \sum_\alpha g_\alpha$, the fluctuation defined by $\delta g^2 \equiv \langle (g - \bar{g})^2 \rangle$ ($\bar{g} \equiv \langle g \rangle$) will be of the order $\delta g^2/\bar{g}^2 \sim O(L^{-d})$, i.e., $\delta g^2 \propto L^{d-4}$, if there is no correlation between g_α for different α. In fact, however, δg^2 has been predicted to be independent of L [9,10] due to the quantum interference of electron waves scattered by the random distribution of impurities, $\sqrt{\delta g^2} = 0.729$ ($d=1$), 0.862 ($d=2$), 1.088 ($d=3$) in units of $e^2/\hbar$.

In reaching this important theoretical finding, the ergodic hypothesis was introduced. The characteristics of the fluctuations of the conductance of a particular sample (with a fixed configuration of impurities) arising from the variation of some external parameters (e.g., magnetic field as in Fig. 1.4.3 or gate voltage) are the same as those among different samples with various impurity configurations. This hypothesis has been numerically tested to be valid as seen in Fig. 1.4.5 [11]. Based on this hypothesis, detailed studies of the dependence of the fluctuations on system size, magnetic field, temperature, and different types of scattering, have been carried out [11].

1.4.3 Persistent Current

The preceding sections have discussed the fluctuations of transport currents in mesoscopic systems. Solids exhibit another type of electric current, which is without dissipation in contrast to the transport currents, i.e., orbital currents induced by the external magnetic field. Such orbital currents are the sources

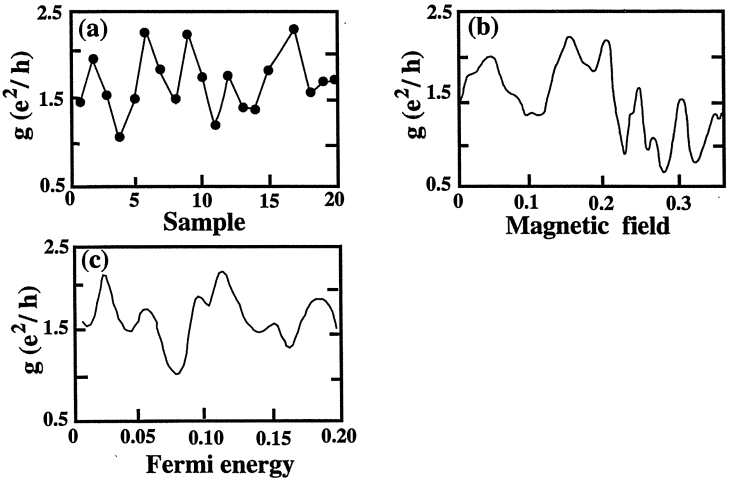


Fig. 1.4.5 The numerical confirmation of the ergodic hypothesis for the conductance fluctuations

of orbital magnetism, whose typical example is the Landau diamagnetism of free electrons, but the sign of the orbital magnetism of Bloch electrons can be positive (orbital paramagnetism) or negative (orbital diamagnetism). A simple case of this magnetic-field-induced persistent current is the one-dimensional ring encircling a flux Φ, similar to Fig. 1.4.2 but with normal metals in the present case. In the clean case, the wave function is given by (see (1.4.2)) $\psi(x) = \exp[-i\theta(x)]e^{ikx}/\sqrt{L}$ where x is the coordinate along the ring with the length L. The periodic boundary condition requires $2\pi\Phi/\Phi_0 + kL = 2\pi n$, with n an integer, resulting in

$$k = \frac{2\pi}{L}\,(n+\varphi) \equiv k_n \tag{1.4.6}$$

with $\varphi = \Phi/\Phi_0$. As φ is varied, k_n changes, and the occupation of k_n and the ground state energy E undergo periodic variations as schematically shown in Fig. 1.4.6. The electric current j along the ring is given by [12,13,14]

$$j = -c\frac{\partial E}{\partial \Phi} = -\frac{e}{h}\frac{\partial E}{\partial \varphi} \tag{1.4.7}$$

whose φ dependence is seen in Fig. 1.4.7. Here, $j_0 = ev_{\mathrm{F}}/L$. As seen, j can have both signs and is of the order of L^{-1}, which are typical features of fluctuations. In the presence of impurity potential, the fluctuations are reduced as $\sqrt{\langle j^2\rangle} \sim j_0(L/\ell)^{-1}$ where ℓ is the elastic mean free path, which is a characteristic value of a current due to one electron diffusing over length L, $e(D/L^2)$, where D is the diffusion coefficient [15].

There exist several experimental reports [16,17] of the observation of such persistent currents, which have revealed much larger magnitude than the theoretical predictions, whose cause has not been clarified yet [18].

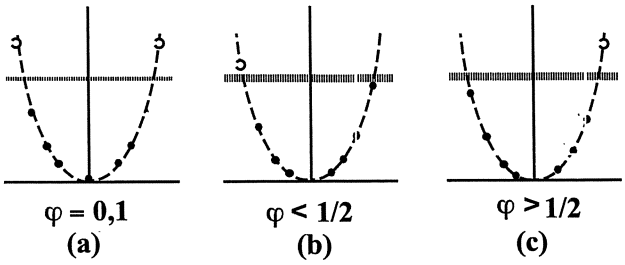


Fig. 1.4.6 Flux dependence of the energy spectra of a one-dimensional electron along a ring

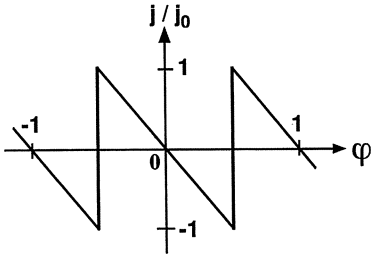


Fig. 1.4.7 Flux dependence of the current of a one-dimensional electron system in a ring

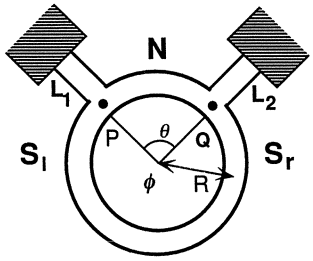


Fig. 1.4.8 The AB ring of a normal metal placed between two superconductors

As explained, the modifications of the allowed values of wave vector k_n by the flux inside the ring is crucial for the existence of the persistent currents. It is interesting and important to note that the values of flux have to be quantized, i.e., φ set equal to an integer, unlike the case where the allowed values of k_n are not changed. This latter case corresponds to the flux quantization by a superconductiong ring with a radius larger than the penetration depth of the external magnetic field. (In this case Φ_0 should be replaced by $\Phi_0/2$.) Whether φ is an independent variable or is quantized depends on the degree of rigidity of the wave function along the ring. In the case of a superconducting ring, the wave function is rigid over a macroscopic scale because of the condensate of Cooper pairs [19], while the one-particle wave function in a normal metal ring cannot have such rigidity. An interesting case would be where such rigidity was varied continuously by means of the proximity effect of superconductivity as shown in Fig. 1.4.8 [20].

1.4.4 Fluctuations of Orbital Susceptibility

As seen in the preceding section, the external magnetic field induces electric current without dissipation leading to a finite orbital susceptibility χ, which can be either diamagnetic or paramagnetic. In the presence of random potential, the spatial pattern of such current flows is very complex, leading to fluctuations of susceptibility, though the effects are expected to be small. Actually, in the case of three-dimensional free electron gas, it has been shown

Fig. 1.4.9 Electron motion in a magnetic field in a spatially varying potential

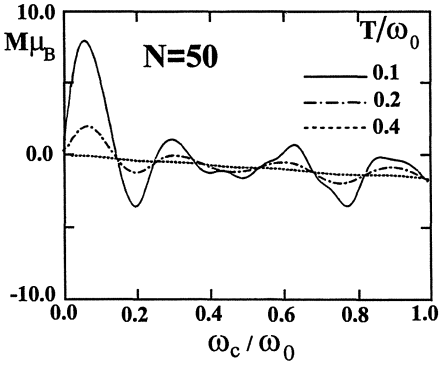


Fig. 1.4.10 Magnetic field dependence of the magnetization of a system with a finite number of electrons confined in a harmonic potential

that the random potential fluctuations do not affect average χ in the bulk [21]. However, it has become clear recently that the fluctuations $\delta\chi$ of the susceptibility have a singular dependence on randomness [22,23].

$$\frac{\langle \delta\chi^2 \rangle}{\chi L^2} \sim \begin{cases} (\varepsilon_F \tau)^2 & d = 2 \\ (\ell/L)^2 & d = 3 \end{cases} \tag{1.4.6}$$

where χ_L is the Landau diamagnetism, and ε_F and τ are the Fermi energy and the relaxation time, respectively. Equation (1.4.6) for $d=3$ implies that the limiting procedures of the macroscopic system size ($L \to \infty$) and the cleanness ($\ell \to \infty$) are not interchangeable, while for $d = 2$ fluctuations are larger than the average once $\varepsilon_F \tau > 1$. This anomalous behavior in $d = 2$, which implies that there will be a singular dependence of the magnetization on the magnetic field, is due to the possible closed orbits encircling a large area in both directions depending on whether the encircled area has a potential maximum or minimum as shown schematically in Fig. 1.4.9.

Associated with this interesting possibility, the field dependence of the magnetization M of a finite system under the confining potential $V(r)$ has been studied [24]. The case of $V(r) = (1/2)m\omega_0^2(x^2+y^2)$ with 50 electrons is shown in Fig. 1.4.10 for several choices of temperature T scaled by ω_0. As seen, there is a strong irregular dependence on magnetic field if the temperature is low, which is easily reduced to a smooth linear dependence once the temperature is raised.

There are still many puzzles associated with orbital currents in the presence of spatially varying potential including the boundaries, which await further investigations.

References

1 Y. Aharonov and D. Bohm, Phys. Rev. **115**, 485 (1959)
2 A. Tonomura, Rev. Mod. Phys. **59**, 639 (1987); Adv. Phys. **41**, 59 (1992)

3 R.A. Webb, S. Washburn, C.P. Umbach, and R.B. Laibowitz, Phys. Rev. Lett. **54**, 2696 (1985)
4 A.D. Stone, Phys. Rev. Lett. **54**, 2692 (1985)
5 C.P. Umbach, S. Washburn, R.B. Laibowitz, and R.A. Webb, *Proc. of LT17* (Elsevier, 1984), p. 717
6 C.P. Umbach, C. Van Hansendock, R.B. Laibowitz, S.Washburn, and R.A. Webb, Phys. Rev. Lett. **56**, 386 (1986)
7 B.I. Altshuler, A.G. Aronov, and B.Z. Spivak, JETP Lett. **33**, 101 (1981)
8 A.D. Stone and Y. Imry, Phys. Rev. Lett. **56**, 189 (1986)
9 B.L. Altshuler, JETP Lett **41**, 648 (1985)
10 P.A. Lee and A.D. Stone, Phys. Rev. Lett. **55**, 1622 (1985)
11 P.A. Lee, A.D. Stone, and H. Fukuyama, Phys. Rev. B **35**, 1039 (1987)
12 N. Byers and C.N. Yang, Phys. Rev. Lett. **7**, 46 (1961)
13 M. Büttiker, Y. Imry, and R. Landauer, Phys. Lett. **96A**, 365 (1963)
14 H.H. Cheung, Y. Gefen, E.K. Riedel, and W.H. Shih, Phys. Rev. B **37**, 6050 (1988)
15 B.L. Altshuler, Y. Gefen, and Y. Imry, Phys. Rev. Lett. **66**, 88 (1991)
16 L.P. Levy, G. Dolan, J. Dunsmuir, and H. Bouchiat, Phys. Rev. Lett. **64**, 2074 (1990)
17 V. Chandrasekhar, R.A. Webb, M.J. Bradly, M.B. Ketchen, W.J. Gallagher, and A. Kleinsadder, Phys. Rev. Lett. **67**, 3578 (1991)
18 For example, Y. Imry, in *Quantum Coherence in Mesoscopic Systems, NATO ASI Series B 254*, edited by B. Kramer (Plenum, 1991), p. 221; and in *Introduction to Mesoscopic Physics* (Oxford University Press, 1997)
19 F. Bloch, Phys. Rev. B **2**, 109 (1970)
20 H. Yoshioka and H. Fukuyama, Physica C **185-189**, 625 (1991); H. Fukuyama and H. Yoshioka, *New Horizons in Low Dimensional Electron Systems* (Kluwer Academic Publishers, 1992), p. 369
21 R.B. Dingle, Proc. Roy. Soc. A **211**, 517 (1952)
22 H. Fukuyama, J. Phys. Soc. Jpn. **58**, 47 (1989)
23 R.A. Serota and S. Oh, Phys. Rev. B **41**, 10523 (1990); S. Oh, A.Yu. Zyuzin, and R.A. Serota, Phys. Rev. B **44**, 8858 (1991)
24 D. Yoshioka and H. Fukuyama, J. Phys. Soc. Jpn. **61** 2368 (1992)

1.5 Ballistic Electron Transport

1.5.1 Quantization of Conductance

Some years ago, van Wees et al. [1] and Wharam et al. [2] found that the conductance of electrons through a narrow channel is quantized in the unit of $2e^2/h$, where e is the absolute value of the electronic charge and h is the Planck constant.

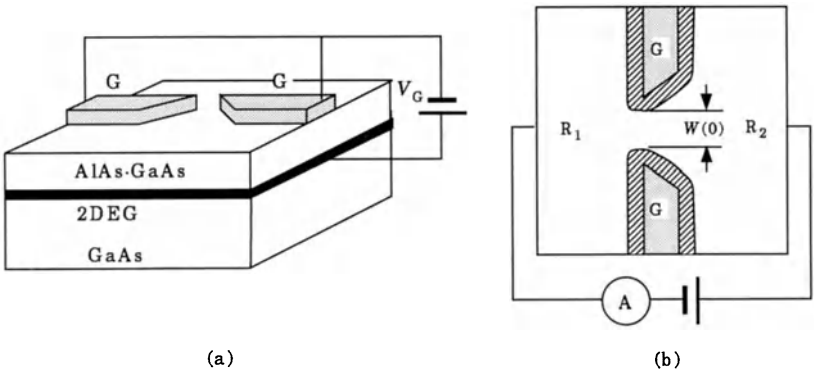


Fig. 1.5.1 (**a**) Schematic view of the sample used in the experiments of quantization on conductance by van Wees et al. [1]. (**b**) View of the sample from above. The shaded areas are the depleted regions

In Fig. 1.5.1a we show the sample used in the experiments by van Wees et al. A two-dimensional electron gas (2DEG) is formed on the interface between GaAs and AlAs/GaAs. Two gate electrodes (Gs in Fig. 1.5.1a) are deposited on the sample and we apply voltage V_G on them relative to 2DEG. If V_G is negative, a narrow channel is formed on 2DEG underneath the gap between the two electrodes because of a valley of the electrostatic potential due to the charge on the electrodes. This narrow channel is called the quantum point contact. In the regions where the potential is higher than the Fermi energy, 2DEG is depleted: such regions are shown in Fig. 1.5.1b schematically as the shaded areas.

The conductance through the narrow channel was measured as a function of the gate voltage V_G. As shown in Fig. 1.5.2, the conductance changes stepwise in the unit of $2e^2/h$. This behavior of the conductance can be understood in terms of Landauer's formula [3], the details of which are reviewed in Sect. 1.3. According to the formula, the conductance of a one-dimensional system is given by

$$G = \frac{2e^2}{h}\mathcal{T}, \tag{1.5.1}$$

where $\mathcal{T}$ is the transmission coefficient of the sample.

As a model of the sample in Fig. 1.5.1, we consider 2DEG confined in a region $|y| \leq W(x)/2$ in the xy plane (Fig. 1.5.3a) by infinitely large potential walls at $y = \pm W(x)/2$. For simplicity, we assume that the potential energy in

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

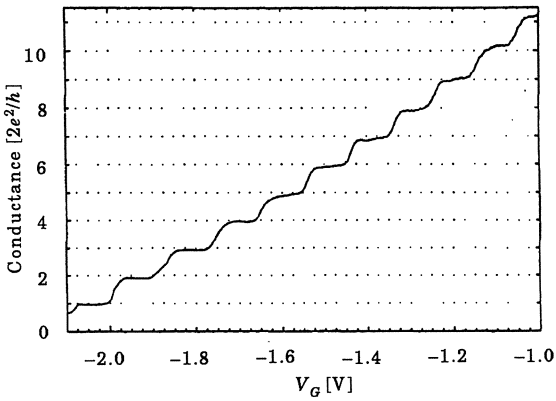


Fig. 1.5.2 Conductance through a narrow channel in Fig. 1.5.1b as a function of gate voltage V_G (from [1])

these regions is 0. Then the eigenstates of the model system can be obtained by solving the Schrödinger equation

$$-\frac{\hbar^2}{2m}\Delta\Psi(x,y) = E\Psi(x,y), \tag{1.5.2}$$

with a boundary condition

$$\Psi(x, \pm W(x)/2) = 0. \tag{1.5.3}$$

For the moment, we assume that the function $W(x)$ is slowly varying, and consider a trial function

$$\Psi(x,y) = \psi_n(x)\sqrt{\frac{2}{W(x)}}\sin\left[\frac{\pi n\{y + W(x)/2\}}{W(x)}\right], \tag{1.5.4}$$

as a solution of (1.5.2), where n is a positive integer. We easily find that this function satisfies the boundary condition (1.5.3) and that if $|dW(x)/dx|$ and $|W(x)d^2W(x)/dx^2| \ll 1$, (1.5.2) reduces to a one-dimensional Schrödinger equation

$$-\frac{\hbar^2}{2m}\frac{d^2}{dx^2}\psi_n(x) + V_n(x)\psi_n(x) = E\psi_n(x), \tag{1.5.5}$$

where $V_n(x)$ is the effective potential (see Fig. 1.5.3b):

$$V_n(x) = \frac{\hbar^2}{2m}\left(\frac{n\pi}{W(x)}\right)^2. \tag{1.5.6}$$

Suppose we set a voltage difference across the system in the x direction. Within this approximation, electrons in the states specified by each n (mode

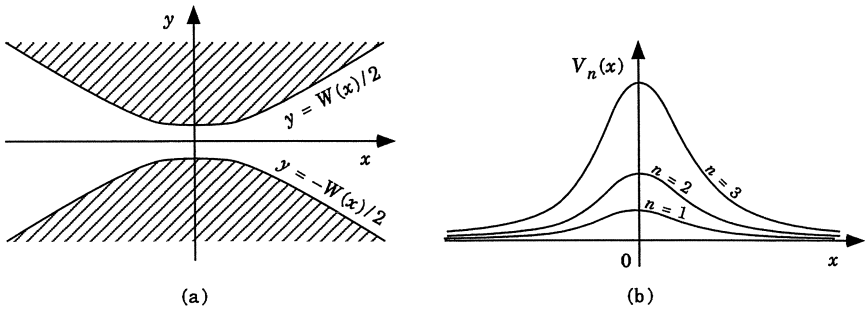


Fig. 1.5.3 (**a**) The model of the sample of Fig. 1.5.1. (**b**) The effective potential $V_n(x)$ given by (1.5.6)

n) independently contribute to the current. Therefore, according to Landauer's formula, the conductance of the system is given by

$$G = \frac{2e^2}{h} \sum_{n=1}^{\infty} \mathcal{T}_n, \tag{1.5.7}$$

where $\mathcal{T}_n$ is the transmission probability of electrons on the Fermi level in the mode n.

Let E_F be the Fermi energy. Since the function $W(x)$ is slowly varying, $V_n(x)$ is slowly varying, too, and the dependence of $\mathcal{T}_n$ on E_F is almost classical, i.e.,

$$\mathcal{T}_n \approx \begin{cases} 1 & (E_F > V_n(0)), \\ 0 & (E_F < V_n(0)), \end{cases} \tag{1.5.8}$$

where we have assumed that $W(x)$ takes the minimum value at $x = 0$.

Here we go back to the sample of Fig. 1.5.1. Suppose we set the gate voltage $|V_G|$ large enough so that the shaded areas (depleted regions) of Fig. 1.5.1b merge into one. In this case, there is no channel between the regions R_1 and R_2, which correspond to the reservoirs in Landauer's theory, and the conductance of the sample is 0.

As we decrease $|V_G|$, a channel opens on 2DEG underneath the gap between the electrodes G in Fig. 1.5.1b, and from (1.5.7) and (1.5.8) it follows that the contribution of the electrons in mode 1 to the conductance becomes $\approx 2e^2/h$, when the width of the channel becomes large so that $V_1(0) < E_F$. In the same way, we have

$$G \approx \frac{2e^2 N}{h}, \tag{1.5.9}$$

where N is defined by $V_N(0) < E_F < V_{N+1}(0)$, and we can explain the quantization of the conductance seen in Fig. 1.5.2.

If we apply the above theory to the sample of Fig. 1.5.1, however, we find that the conditions $|dW(x)/dx|$ and $|W(x)d^2W(x)/dx^2| \ll 1$ do not seem to hold in all the regions of x, although the details of the shapes of the depleted

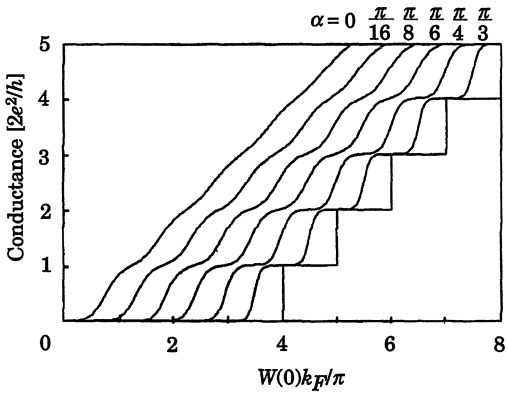


Fig. 1.5.4 Conductance through a narrow channel with a hyperbolic boundary as a function of the width of the channel (from [5])

regions are not known.

In order to calculate the conductance for a realistic shape of the boundary, one has to solve the Schrödinger equation (1.5.2) with the boundary condition (1.5.3). Although the equation is very simple, generally it is not easy to solve. Just after the experiments, Kawabata calculated the conductance for a hyperbolic boundary [4,5]:

$$\frac{y^2}{c^2\cos^2\alpha} - \frac{x^2}{c^2\sin^2\alpha} = 1. \tag{1.5.10}$$

Using elliptic coordinates (u, v) defined by

$$x = c\sinh u \sin v, \tag{1.5.11a}$$

$$y = c\cosh u \cos v, \tag{1.5.11b}$$

one can map the boundary to $v=\alpha$, $\pi-\alpha$. Then, putting $\Psi(x,y)=\psi(u)\phi(v)$, we easily find that the Schrödinger equation (1.5.2) reduces to

$$-\frac{d^2\psi(u)}{du^2} - 4q\psi(u)\cosh^2 u = -A\psi(u), \tag{1.5.12a}$$

$$-\frac{d^2\phi(v)}{dv^2} + 4q\phi(v)\cos^2 v = A\phi(v), \tag{1.5.12b}$$

with the boundary condition $\phi(\alpha)=\phi(\pi-\alpha)=0$, where $q=c^2mE/2\hbar^2$ and A is the parameter to be determined as the eigenvalue of (1.5.12b). The solution for each value of A corresponds to a mode discussed in Sect. 1.5.1.

Equations (1.5.12a) and (1.5.12b) cannot be solved analytically, but it is much easier to solve them numerically than to solve two-dimensional problem directly. In Fig. 1.5.4 we show the results, where $W(0)\equiv c\cos\alpha$ is the width of the narrowest part of the channel and k_F is the Fermi wave number. We observe that the conductance is quantized well enough for $\alpha \geq \pi/4$. At the same time, note that traces of the conductance steps are observed even for $\alpha=0$.

The calculations were done also for a different shape of boundary [5], and it was found that the conductance is determined mainly by the structures of the narrowest part of the channel: it is quantized well enough if the ratio of the channel width to the radius of curvature of the boundary of the depleted regions is less than 1, i.e., $|W(x)d^2W(x)/dx^2|_{x=0} < 1$ for the model of Fig. 1.5.3a. This condition may be satisfied in the case of the sample of Fig. 1.5.1.

1.5.2 Interaction Effects on Conductance Quantization

According to the arguments in the previous section, ideal quantization of conductance will be realized when the function $W(x)$ which specifies the boundary is a constant. In this limit the system is quasi one-dimensional and is called quantum wire. When the width of a quantum wire is small enough, only one mode (subband) is occupied and theories of purely one-dimensional systems are applicable.

A theory of one-dimensional interacting electron system was developed by Tomonaga many years ago [6]. He did not investigated the transport properties, but the fundamental idea, i.e., bosonization technique, was adoped in various studies [7]. It was argued that the conductance of one-dimensional systems is renormalized as

$$G = K\frac{2e^2}{h}, \tag{1.5.13}$$

with

$$K = \sqrt{\frac{1}{1+2\hat{V}(0)/(\pi\hbar v_F)}}, \tag{1.5.14}$$

where v_F is the Fermi velocity and $\hat{V}(0)$ is the Fourier component of the interaction potential.

It has been found, however, that the renormalization factor is the result of the wrong definition of conductance [8]. In deriving (1.5.13), the current $J(x,t)$ was calculated as a linear response to the externally applied electric potential $\Phi_0(x,t)$ and the conductance was defined as the ratio of the current to the electric field:

$$G \equiv \frac{J(x,t)}{-\partial\Phi_0(x,t)/\partial x}, \tag{1.5.15}$$

Izuyama noted that the conductance should be defined with respect to the self-consistent electric potential $\Phi(x,t)$, which contains the contributions of the polarization of electrons [9]. In fact, according to elementary electromagnetism, the electric field contains the contributions of the polarization. Therefore the conductance has to be defined as

$$G \equiv \frac{J(x,t)}{-\partial\Phi(x,t)/\partial x}. \tag{1.5.16}$$

Recently, Kawabata has shown that with this definition the factor K in (1.5.13) does not appear and the conductance is not affected by the interaction [8]. This idea can be extended to quantum wires in which more than

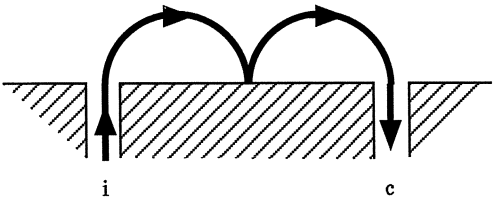


Fig. 1.5.5 A schematic view of magnetic focusing

two subbands are occupied. See Sects. 2.1 and 2.2 for more details on effects of electron-electron interaction on transport in quantum wires.

1.5.3 Magnetic Focusing

Magnetic focusing is one of the methods to observe directly the ballistic motions of electrons. We inject an electron beam into two-dimensional space with a magnetic field perpendicular to it. The electrons are detected at the corrector channel c. The number of electrons entering c per unit time is expected to show peaks as a function of the magnetic field when the distance between c and i is equal to the cyclotron diameter times an integer, provided that the electron motion is ballistic and that the reflection of electrons at the boundary is specular.

Figure 1.5.6a shows experiments by van Houten et al. [10]. Here the number of electrons entering c is measured as the voltage difference between c and i. The values of the magnetic field indicated by the arrows are those for which the distance between c and i is equal to the cyclotron diameter times the integers shown above the arrows. We find peaks of voltage at those values of magnetic field. We also find that some peaks split into two or more small peaks at low temperatures. Figure 1.5.6b shows the results of the numerical calculation of the probability for the electrons to enter the channel c [11]. We find the splitting of the peaks, but the correspondence with the experiments is unclear.

1.5.4 Bend Resistance and Transfer Resistance

For bulk samples we assume that the local current is determined by the local electric field. This is the case on a scale larger than the mean free path of electrons. When the electron motion is ballistic, the nonlocal natures of electron transport show up. An example of the quantities that characterize the nonlocal transport is the bend resistance. Consider a sample of crossed narrow channels fabricated on 2DEG shown in the inset of Fig. 1.5.7a. The bend resistance R_B is defined as $R_B \equiv V_{43}/I_{12}$, where V_{43} is the voltage of probe 4 relative to 3 and I_{12} is the current injected to probe 1 and corrected at 2.* Within classical circuit theory, R_B vanishes because the current between probes 3 and 4 is 0 (or it should be positive if we take into account the finite

* Generally this resistance is written as $R_{12,43}$.

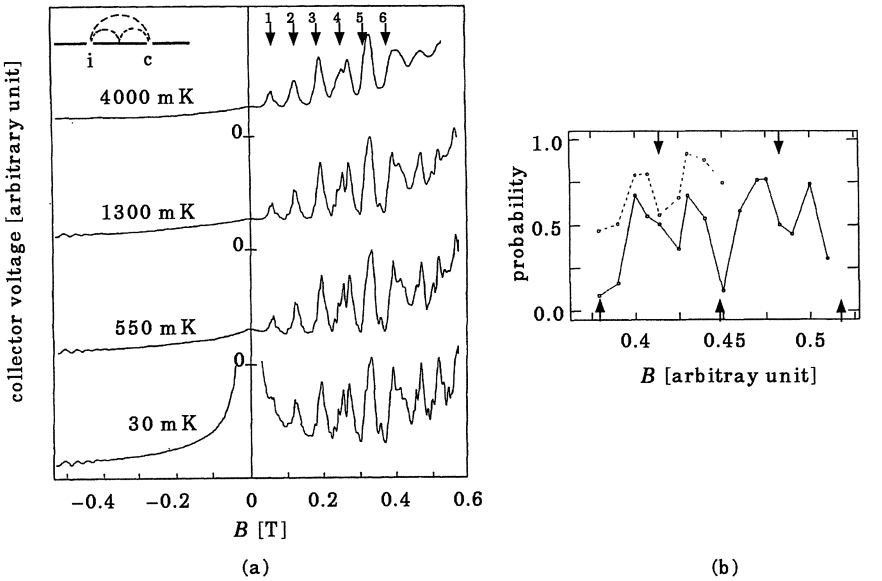


Fig. 1.5.6 (**a**) The voltage difference between c and i vs. magnetic field. The arrows indicate the values of the magnetic field for which peaks are expected from geometrical considerations (from [10]). (**b**) The probability for an electron to enter the channel c calculated by Okiji et al. The downward allows correspond to those (4 and 5) in (a) (from [11])

width of the channel). However, we observe negative R_B when the electron motion is ballistic. In the present configuration, the electrons (not the current) are injected from probe 2 and some of them fail to turn to the direction of probe 1 and go into probe 4. Then probe 4 is negatively charged and we observe negative bend resistance.

Figure 1.5.7a shows the experiment by Takagaki et al. [12]. The magnetic field dependence indicates that the electrons do not go into probe 4 under the strong magnetic field for which the cyclotron radius of the electron at the Fermi level is smaller than the width of the channel.

Nonlocal resistance is observed even for the sample shown in the inset of Fig. 1.5.7b for which the voltage probes are far from the classical current path. The figure shows the transfer resistance $R_T \equiv R_{21,43}$ as a function of the distance L [13]. We find that it decays nearly exponentially with L and we can estimate the effective mean free path from this data.

1.5.5 Anomaly in Weak-Field Hall Effect

Another example of the anomalous characteristics of ballistic transport is the quenching of the Hall effect in the weak-field regions. The ideal voltage probes in Hall-effect measurements are those for which the contact with the sample is infinitesimally weak. In real experiments, however, one uses a cross channel structure like the one in the inset of Fig. 1.5.7a. In this case, probes 1 and 3 are the current probes and 2 and 4 are the voltage probes.

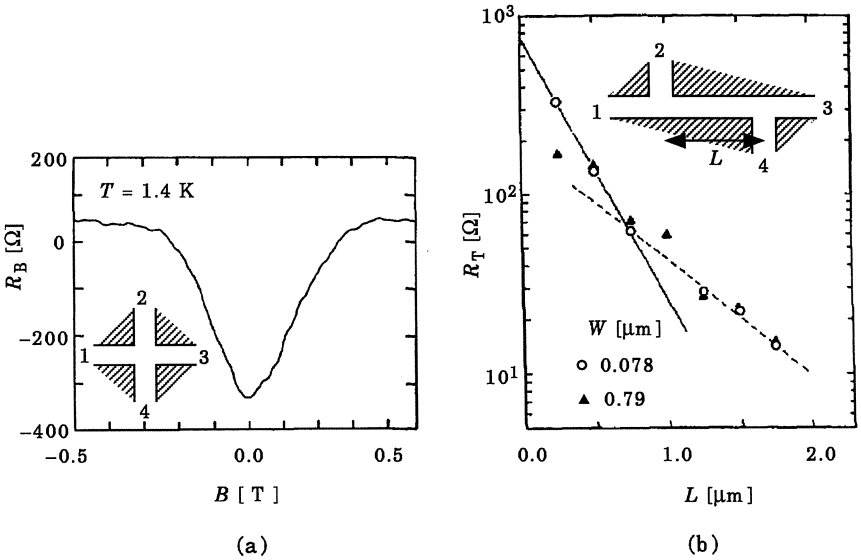


Fig. 1.5.7 (**a**) Bend resistance vs. magnetic field (from [12]). (**b**) Transfer resistance vs. distance L (from [13])

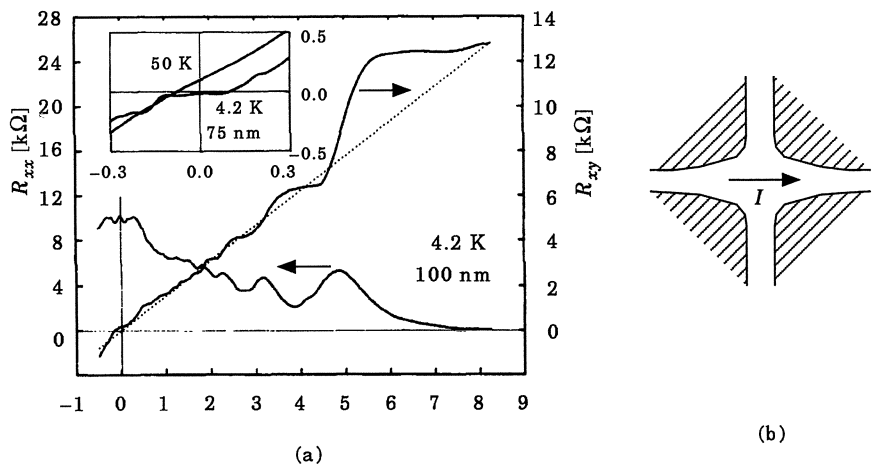


Fig. 1.5.8 (**a**) Hall resistance R_{xy} and longitudinal resistance R_{xx} for 100 nm wide channel at 4.2 K. Inset: R_{xy} for 75 nm wide channel at 50 K and 4.2 K. (**b**) The structure of the sample favorable for quenching the Hall effect

Roukes et al. [14] observed that the Hall resistance R_{xy} vanishes in the weak magnetic field regions for thin channels whose width is less than 100 nm (see Fig. 1.5.8a). This quenching of the Hall effect is interpreted as follows. The Hall voltage is not observed unless the electrons go into the voltage probes. If the electron motion is ballistic and the width of the Hall probe lead

is smaller than the cyclotron radius, the probability for an electron to go into the Hall probe might be very small.

Many theoretical investigations have been made on this problem. For instance, Baranger and Stone [15] calculated the Hall resistance for the structure shown in the inset of Fig. 1.5.7a. They found that the Hall resistance is very much dependent on the Fermi energy of the electrons and that it is quenched in limited regions of the Fermi energy. They argued that quenching of the Hall effect can be observed if the current channel is widened near the crossing and the crossing is not sharp, as is shown in Fig. 1.5.8b. For such a structure, the effective potential (see Sect. 1.5.1) is smaller in the wider part, and the electrons are accelerated along the channel at the expense of their transverse momentum. Therefore, the current density is small near the boundary (collimation) and the probability for an electron to be scattered into the voltage probe is small.

References

1 B.J. van Wees, H. van Houten, C.W.J. Beenakker, J.G. Williamson, L.P. Kouwenhoven, D. van der Marel, and C.T. Foxon, Phys. Rev. Lett. **60**, 848 (1988)

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

3 M. Büttiker, Y. Imry, R. Landauer, and S. Pinhas, Phys. Rev. B **31**, 6207 (1985)

4 A. Kawabata, J. Phys. Soc. Jpn. **58**, 372 (1989)

5 A. Kawabata, *Proceedings of the 3rd International Symposium on Foundations of Quantum Mechanics in the Light of New Technology, Tokyo, 1989* (Physical Society of Japan, 1990), p. 221

6 S. Tomonaga, Prog. Theor. Phys. **5**, 544 (1950)

7 J. Sólyom, Adv. Phys. **28**, 201 (1979)

8 A. Kawabata, J. Phys. Soc. Jpn. **65**, 30 (1996)

9 T. Izuyama, Prog. Theor. Phys. **25**, 964 (1961)

10 H. van Houten, B.J. van Wees, J.E. Mooij, C.W.J. Beenakker, J.G. Williamson, and C.T. Foxon, Europhys. Lett. **5**, 721 (1988)

11 A. Okiji, N. Negishi, and A. Nakamura, J. Phys. Soc. Jpn. **61**, 1145 (1992)

12 Y. Takagaki, K. Gamo, S. Namba, S. Takaoka, K. Murase, S. Ishida, K. Ishibashi, and Y. Aoyagi, Solid State Commun. **69**, 811 (1989)

13 T. Sakamoto, Y. Takagaki, K. Gamo, S. Namba, S. Takaoka, and K. Murase, Solid State Commun. **80**, 535 (1991)

14 M.L. Roukes, A. Scherer, S.J. Allen, Jr., H.G. Craighead Jr., R.M. Ruthen, E.D. Beebe, and J.P. Harbison, Phys. Rev. Lett. **59**, 3011 (1987)

15 H.U. Baranger and A.D. Stone, Phys. Rev. Lett. **63**, 414 (1989)

1.6 Coulomb Blockade

1.6.1 Introduction

Mesoscopic physics has started with the study of various kinds of interference effects of the electronic wave, but the features of electrons as a particle also show up in various interesting phenomena. For example, consider a small tunnel junction connected to an electrical source of constant voltage V (Fig. 1.6.1a). Suppose an electron tunnels from one electrode to the other through the insulator between them. If the junction is very small, its capacitance is small too, unless the insulator between the electrodes is extremely thin. Then the charging energy $e^2/2C$ can be as large as temperature times the Boltzmann constant, and is not negligible. Therefore, the tunneling is not realized unless the voltage V is large enough for this energy to be compensated by the energy eV.

Generally, in small systems the motion of electrons can be blocked by the charging energy; this phenomenon is called the Coulomb blockade. It has attracted much interest because of its possibility of application, and various interesting Coulomb blockade phenomena have been observed. In the following we review some topics among them.

1.6.2 Single Electron Tunneling

In the above we have seen that, in a small junction, tunneling occurs only if $eV > e^2/2C$. Therefore the current-voltage characteristics of a junction shows nonlinear behavior as is shown in Fig. 1.6.1b [1]. In this way we can control the tunneling of one electron; this phenomenon is called single-electron tunneling (SET).

The observation of SET is very sensitive to environmental effects. For example, even if the capacitance of the junction itself is very small, the effects of the leads connected to the junction have to be accounted for in the effective capacitance. Note that the effective capacitance is dependent on the time scale. After the tunneling, the charges diffuse into the leads in a characteristic time τ, and the charging energy reduces to 0. Therefore, in order that the electron "feel" the charging energy, the charging energy must be larger than the uncertainty of the energy, i.e.,

$$e^2/2C > \frac{\hbar}{\tau}. \tag{1.6.1}$$

Suppose that the effective capacitance is negligible if the charge is distributed on the leads over a certain length. We let R be the resistance of these parts of the leads. Then, the characteristic time τ can be estimated to be:

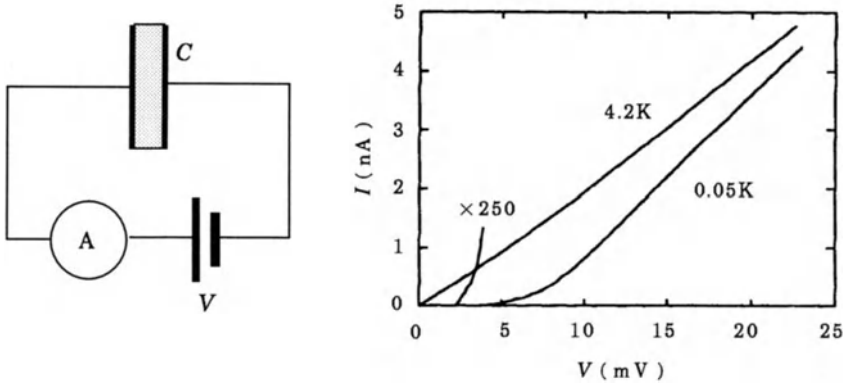
$$\tau \approx CR. \tag{1.6.2}$$

Although this estimation is based on classical electromagnetism, it should be good enough as long as the order of magnitude is concerned.

Thus, from (1.6.1) and (1.6.2), we find that one of the conditions to observe SET is

$$R > \frac{e^2}{\hbar}. \tag{1.6.3}$$

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998



(a) (b)

Fig. 1.6.1 (a) A tunneling junction connected to a constant voltage source. (b) $I-V$ characteristics of small tunneling junction (from [1])

In the experiment in Fig. 1.6.1b, the 19 junctions are connected in series in order to avoid the effects of the leads. The current begins to grow at $V \approx 7$ mV, and the voltage for one junction is 0.36 mV. The capacitance of a junction is estimated to be 0.28 fF, and the threshold voltage is in good agreement with $e/2C = 0.29$ mV.

1.6.3 SET Oscillation

Next, we consider the case when the junction is connected to a source which supplies constant current I. Suppose the junction is neutral at time $t = 0$. Then the charge on the junction is $Q = It$, and the tunneling of an electron becomes possible only when the electrostatic energy becomes smaller after the tunneling, i.e., $(Q-e)^2/2C < e^2/2C$, or $Q > e/2$. After the tunneling, Q becomes almost equal to $-e/2$, and it takes a time e/I until the Q becomes larger than $e/2$. Therefore, although the current source provides constant current, the current in the junction is pulse-like with average period e/I. This phenomenon is called SET oscillation. It should be noted that the period does not depend on the structure of the system.

SET oscillation can be observed utilizing the resonance with microwaves. Usually the constant current source is realized by a constant-voltage source with a large resistor R_0 connected in series, and the current I is controlled by source voltage V. Then, when junction is irradiated with microwaves of frequency ν, the SET oscillation resonates with the microwaves under the condition $\nu = I/e = V/eR_0$. If V deviates slightly from $eR_0\nu$, the frequency of the SET oscillation tends to be locked at ν and hence the current is kept constant. Therefore, SET oscillation is observed as the peak of the differential resistance dV/dI.

Figure 1.6.2 shows an example of experimental results. These measure the differential resistance of the sample used in Fig. 1.6.1 under microwaves of

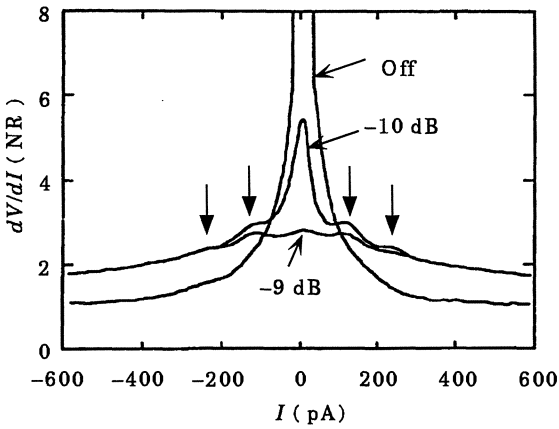


Fig. 1.6.2 Differential resistance of small tunneling junctions vs. current under microwaves of 0.75 GHz. Off, −10 dB, and −9 dB indicate the power of the microwaves. The arrows indicate the SET oscillation for which $I = \pm e\nu$, and the higher harmonics (from [1])

0.75 GHz. The arrows indicate the peaks at $I = \pm e\nu = \pm 120$ pA and the peaks due to the higher harmonics at $2e\nu$.

In order to apply this device as a standard of the current, it is desirable to have a sharp peak of differential resistance. One of the origins of the broadening of the peak is the stochastic feature of the tunneling. The time intervals between successive tunnelings are not constant but fluctuate about e/I, and this causes the broadening of the resonance. For details on SET, the reader is referred to [2].

1.6.4 Tunneling in Superconducting Junctions

Here we investigate the charging effects in small tunneling junctions of superconductors [3]. Let n_1 and n_2 be the number of electrons in the electrodes 1 and 2, respectively, and let us represent the ground state of the system with $|n_1\rangle|n_2\rangle$. Suppose the electrodes are neutral when $n_1 = N_1$ and $n_2 = N_2$. Then the states defined by

$$\Phi_n \equiv |N_1 + 2n\rangle\, |N_2 - 2n\rangle. \tag{1.6.4}$$

are degenerate if the charging energy is neglected and the chemical potentials on the electrodes are equal.

If we neglect the excited states, the system can be described by the Hamiltonian $\mathcal{H}$ whose matrix elements are

$$\langle \Phi_n | \mathcal{H} | \Phi_m \rangle = -E_J\, \delta_{n,m\pm 1}. \tag{1.6.5}$$

Here E_J is the coupling constant which represents the second order process with respect to the tunneling Hamiltonian. Let t be the tunneling matrix element of one electron. Then, the intermediate states of the second-order process are those in which an electron has tunneled from one electrode to the other. Because of the unpaired electrons, the energies of these states are at

least 2Δ higher than those of the ground state, where Δ is the energy gap. The intermediate states which mainly contribute to E_J are those in which the energies of the unpaired electrons are of order Δ. Therefore we find that $E_J \approx t^2 D^2 \Delta^2/\Delta$, where D is the state density. On the other hand, we easily find that the normal state tunneling resistance of the junction is given by $R_N \equiv \hbar/4\pi(etD)^2$. Thus we find that

$$E_J \approx \frac{\hbar\Delta}{e^2 R_N}. \tag{1.6.6}$$

It is easy to see that the eigenstates and the energy eigenvalues of the Hamiltonian (1.6.5) are given by

$$\Psi(\phi) \equiv \sum_n e^{in\phi}\Phi_n, \tag{1.6.7a}$$

$$E(\phi) \equiv -2E_J \cos\phi, \tag{1.6.7b}$$

where ϕ is a quantum number to label these states. Let $\hat{n}$ be the operator for n. Then, since Φ_n is the eigenstate of $\hat{n}$, from (1.6.7b) we find that $\hat{n}$ in ϕ representation is

$$\hat{n} = -i\frac{\partial}{\partial\phi}. \tag{1.6.8}$$

From this expression follows the commutation relation

$$[\phi, \hat{n}] = i. \tag{1.6.9}$$

The operator for the tunneling current from electrode 2 to 1 is given by

$$\hat{I} = -2e\frac{d\hat{n}}{dt} = \frac{2e}{\hbar}\left[\frac{\partial}{\partial\phi}, \mathcal{H}\right]. \tag{1.6.10}$$

Therefore, from (1.6.7b) it follows that $\Psi(\phi)$ is the eigenstate of $\hat{I}$ and its eigenvalue is

$$I(\phi) = \frac{2e}{\hbar}\frac{d}{d\phi}E(\phi) = \frac{4e}{\hbar}E_J \sin\phi. \tag{1.6.11}$$

Note that we have assumed that the chemical potentials of electrodes 1 and 2 are equal, and hence the above equation indicates that the tunneling current can be finite without the voltage difference between the electrodes. This is the ordinary Josephson effect [4].

The charging energy is $(2en)^2/2C$ with C being the capacitance of the junction, and from (1.6.7b) and (1.6.8) the Hamiltonian of the system in ϕ representation is

$$\mathcal{H} = -\frac{2e^2}{C}\frac{d^2}{d\phi^2} - 2E_J \cos\phi. \tag{1.6.12}$$

If we regard ϕ as the coordinate of a particle in one-dimensional space, this is just the Hamiltonian for a particle of mass $C\hbar^2/4e^2$ in a periodic potential $V(\phi) = -2E_J \cos\phi$. Then the eigenstates of the system are labeled by the

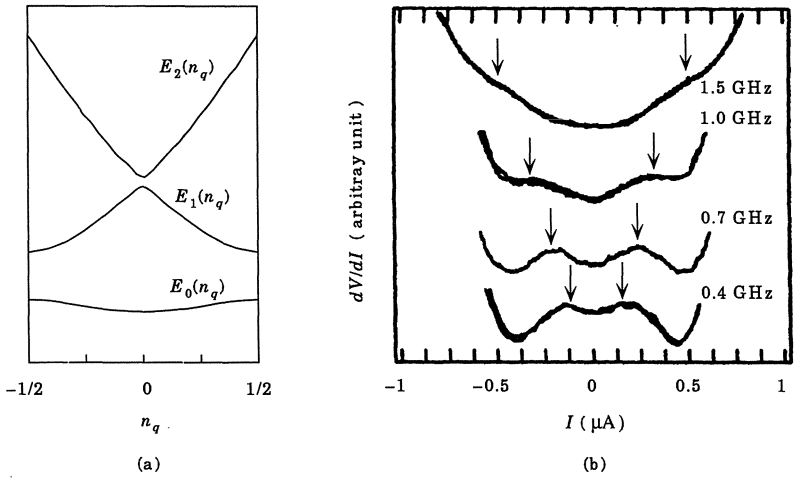


Fig. 1.6.3 (**a**) The energy band for a small superconducting junction (schematic). (**b**) The differential resistance of the junction vs. current in microwave. The peaks are observed at $I = \pm 2e\nu$ with the microwave frequency ν indicated on the curves (from [6])

"quasi-momentum" n_q and the band index N, and the energy eigenvalues $E_N(n_q)$ behave as in Fig. 1.6.3a.

We will consider the case when the current through the junction is driven by a current source of constant current I. The current source is represented by a Hamiltonian $\mathcal{H}_I$ such that

$$\frac{d\hat{n}}{dt} = \frac{i}{\hbar}[\mathcal{H}_I, \hat{n}] = \frac{I}{2e}. \tag{1.6.13}$$

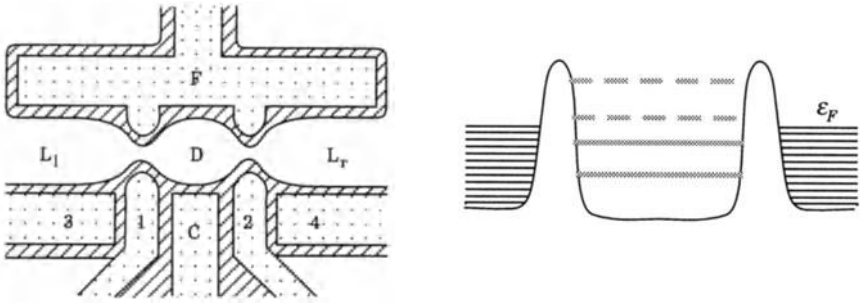
From the commutation relation (1.6.9) we easily find that this relation is satisfied by

$$\mathcal{H}_I \equiv -\frac{\hbar I}{2e}\phi. \tag{1.6.14}$$

If we regard ϕ as the coordinate of a particle in a periodic potential, as was mentioned above, $\mathcal{H}_I$ is a Hamiltonian for applying a constant force $\hbar I/2e$. Then n_q obeys the "acceleration equation" [5]:

$$\frac{dn_q}{dt} = \frac{I}{2e}. \tag{1.6.15}$$

Suppose the system is in a state specified by $N = 0$ and some value of n_q. Then, according to (1.6.15), n_q increases until it reaches the edge of the Brillouin zone, i.e., $n_q = 1/2$, and jumps to the other side of the Brillouin zone. The jump of n_q correspond to the tunneling of an electron from electrode 2 to electrode 1. The tunneling take places with period $2e/I$, and we observe a pulse current between the electrodes. This phenomenon is called Bloch oscillation.



(a) (b)

Fig. 1.6.4 (a) A typical structure of a quantum dot. The depleted (shaded) areas are controlled by electrodes 1–4, C, and F. Electrode C also controls the electrostatic potential in the dot. (b) A model of a quantum dot

Note that the tunneling does not take place in an instant. If $E_J \ll e^2/C$, the periodic potential is very weak and the wave function of the eigenstate is approximately equal to $\exp(in_q\phi)$ except near the edge of the Brillouin zone. Near the edge, the wave function is a linear combination of $\exp(in_q\phi)$ and $\exp\{i(n_q-1)\phi\}$ and their amplitude becomes equal at the edge. Thus in this case, the tunneling takes place during a time much shorter that $2e/I$ and we observe sharp pulse current.

On the other hand, if $E_J \gg e^2/C$, the wave function is not a pure plane wave everywhere in the Brillouin zone, and the tunneling takes place gradually, i.e., the current between the electrodes is almost constant. This limit corresponds to the ordinary Josephson effect.

Bloch oscillation can be observed with the use of resonance with microwaves as in SET oscillation, and an experimental result is shown in Fig. 1.6.3b [6]. Peaks of the differential resistance are seen at $I = \pm 2e\nu$ with ν being the frequency of the microwave.

The Bloch oscillation is not a mere superconducting version of SET oscillation. In contrast to SET, the tunneling of the electron is not stochastic but deterministic. Therefore, within the model of the above arguments there is no fluctuation of the time interval between successive tunnelings, and we can expect sharper resonance. Historically, the Bloch oscillation was observed earlier than SET oscillation.

1.6.5 Coulomb Blockade in Quantum Dots

Most quantum dots are made on a two-dimensional electron system on the interface of the GaAs with gate electrodes. A typical structure of a dot is shown in Fig. 1.6.4a. Here the dot D is connected to the leads L_l and L_r through the constriction controlled by the electrodes 1 and 2.

We consider the case when the width of the leads is smaller than the Fermi wavelength. Then a quantum dot can be modeled with double potential barrier as shown in Fig. 1.6.4b, where the barriers correspond to the constrictions. The energy levels in the dot are discrete but are broadened because of the

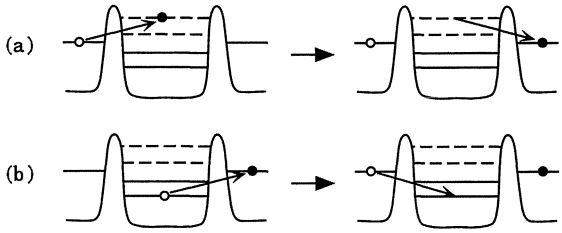


Fig. 1.6.5 The processes of the transmission of an electron through a quantum dot. The closed and open circles represent electrons and holes, respectively

tunneling through the barriers. Within one-electron theory, the widths of the levels have to be smaller than the spacing between the levels in order that the discreteness of the energy levels shows up in the conduction through the dot.

Under such conditions, in the low bias regions, i.e., when the voltage difference at the leads is small enough, the conductance through the dot is much smaller than $2e^2/h$, except when one of the levels in the dot coincides with the Fermi level in the leads and resonant transmission occurs. Therefore, if the bottom of the potential energy in the dot is lowered, for instance using electrode C in Fig. 1.6.4a, the energy levels go down and the conductance shows a peak every time one of the levels crosses the Fermi level.

For a very small dot, we have to take into account the charging energy of the dot. Let E_N be the ground state energy of the dot when N electrons are in the dot. Suppose the dot is neutral for $N = M$. The transmission of an electron through the dot occurs as a second-order process in the tunneling matrix element t between the states in the dot and the leads (see Fig. 1.6.5). The effective matrix element for the transmission is given by

$$P = \frac{|t|^2}{E_M - E_{M+1}} + \frac{|t|^2}{E_M - E_{M-1}}, \tag{1.6.16}$$

where the first and the second term correspond to processes (a) and (b), respectively, in Fig. 1.6.5.

The matrix element P diverges when $E_M = E_{M\pm1}$, and it is clear that this situation corresponds to resonant transmission. Therefore, if we lower the bottom of the potential energy, we will observe peaks of the conductance every time the conditions $E_M = E_{M+1}$, $E_{M+1} = E_{M+2}$, $\cdots$ are fulfilled. Such behavior of the conductance is called Coulomb blockade oscillation or Coulomb oscillation. We show an example of experiments in Fig. 1.6.6 [7].

The capacitance C of the dot is estimated to be 3×10^{-16}F, and hence the charging energy due to excess charge e is $e^2/2C = 0.3$ meV, while the spacing of one-electron energy is estimated to be 0.03 meV. Therefore, the charging energy plays an important role in the Coulomb blockade oscillations in Fig. 1.6.6.

1.6.6 Resonant Transmission and Kondo Effect

It is instructive to analyze resonant transmission in terms of phase shift. For the moment we neglect the effects of electron-electron interaction. We consider

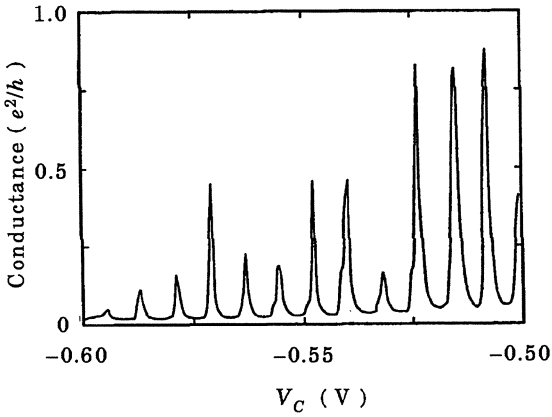


Fig. 1.6.6 Conductance of a quantum dot vs. the voltage of gate electrode C (from [7])

the case when the potential energy in Fig. 1.6.4b is symmetric with respect to $x=0$, the center of the dot. Then the one-electron eigenstates are symmetric or antisymmetric, and the wave functions of the states with energy $\hbar^2k^2/2m$ can be written in the form

$$\begin{aligned} \psi_{sk}(x) &= \cos[k|x| + \delta_s(k)], \\ \psi_{ak}(x) &= \mathrm{sgn}(x)\sin[k|x| + \delta_a(k)], \end{aligned} \tag{1.6.18}$$

outside the dot. Here $\delta_s(k)$ and $\delta_a(k)$ are called phase shifts.

It is easy to see the wave function $\psi_{t,k}(x) \equiv \exp[i\delta_s(k)]\psi_{sk}(x) + i\exp[i\delta_a(k)]\psi_{ak}(x)$ has only an outgoing wave in the regions $x>0$, and that the transmission coefficient is given by [8]:

$$\mathcal{T} = \cos^2[\delta_s(k_F) - \delta_a(k_F)]. \tag{1.6.19}$$

Suppose we raise the Fermi energy and it passes one of the discrete energy level in the dot, which is, say, symmetric. At the resonant transmission the symmetric wave function is connected to that level, and exhibits a considerable change with a small change of the Fermi energy. When the Fermi energy is well below the level, $\mathcal{T}$ is small. Then the value of the wave function at the barrier is small, and outside the dot it looks like curve a in Fig. 1.6.7a. As the Fermi energy approaches the level, $\mathcal{T}$ and the wave function changes like curves b→c, where c corresponds to the resonant transmission. If we let the Fermi energy increase, $\mathcal{T}$ decreases and the wave function changes like curve d→e.

Thus we find that the phase shift $\delta_S(k_F)$ increases by π when the Fermi energy passes through a quasi discrete symmetric energy level. If the level is very sharp and the change of the Fermi energy is very small, we can neglect the change of $\delta_a(k_F)$. This observation is consistent with (1.6.19). The transmission coefficient changes like $\mathcal{T} \approx 0 \to 1 \to 0$ and the phase shift changes

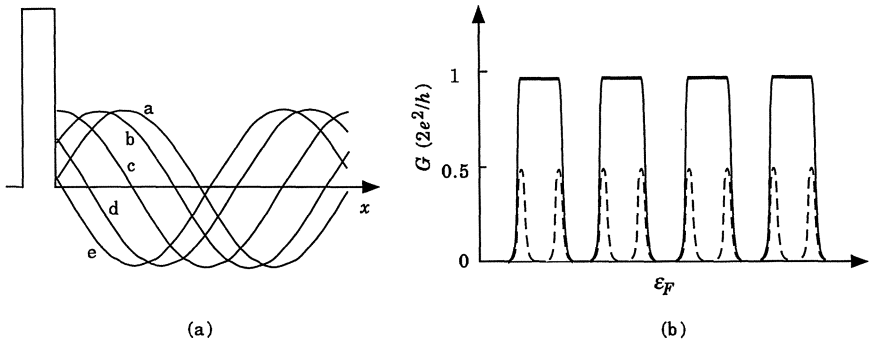


Fig. 1.6.7 (**a**) The change of the wave function outside the barrier of a dot when the Fermi energy passes through a quasi discrete level in the dot. (**b**) Conductance through a dot versus Fermi energy for $T<T_K$ (solid line) and for $T>T_K$ (dotted line)

like $\delta_s(k_F)-\delta_a(k_F)\approx(n+1/2)\pi \to (n+1)\pi \to (n+3/2)\pi$ with n an integer.

In fact we can show that the relations [9]

$$\begin{aligned}\Delta\delta_s(k_F) &= \frac{\pi}{2}\Delta N_s,\\ \Delta\delta_a(k_F) &= \frac{\pi}{2}\Delta N_a,\end{aligned} \tag{1.6.20}$$

hold when we change the Fermi energy, where ΔN_s and ΔN_a are the change of the numbers of the electrons in the symmetric and antisymmetric quasi discrete levels, respectively, and $\Delta\delta_s(k_F)$ and $\Delta\delta_a(k_F)$ are the change of the phase shifts.* This set of relations is a special form of the Friedel sum rule [10].

Hereafter we take into account the electron-electron interaction. It has been shown by Langer and Ambegaokar that the Friedel sum rule holds for interacting electrons [11].

Consider the case when the Fermi energy is below the lowest quasi discrete energy level in the dot, which is assumed to be symmetric. Then the transmission coefficient is small and we have $\delta_s(k_F)-\delta_a(k_F)\approx(n+1/2)\pi$ with n an integer. If we let the Fermi energy be just above the level, only one electron occupies the level because of the repulsive Coulomb energy, and from (1.6.20) we find that $\delta_s(k_F)$ increases by $\pi/2$. Therefore we have $\delta_s(k_F)-\delta_a(k_F)\approx(n+1)\pi$, and hence $\mathcal{T}\approx 1$. If the level is occupied by another electron when the Fermi energy is raised further, from (1.6.20) we have $\delta_s(k_F)-\delta_a(k_F)\approx(n+3/2)\pi$ and $\mathcal{T}\ll 1$.

Extending the above argument to the case when there are more electrons in the dot, we find that $\mathcal{T}\approx 1$ when the number of electron in the dot is odd [9]. This conclusion is not consistent with the arguments in the preceding subsection, in which many-body effects are taken into account only within

* Two electrons enter the quasi discrete level when Fermi energy passes through the level.

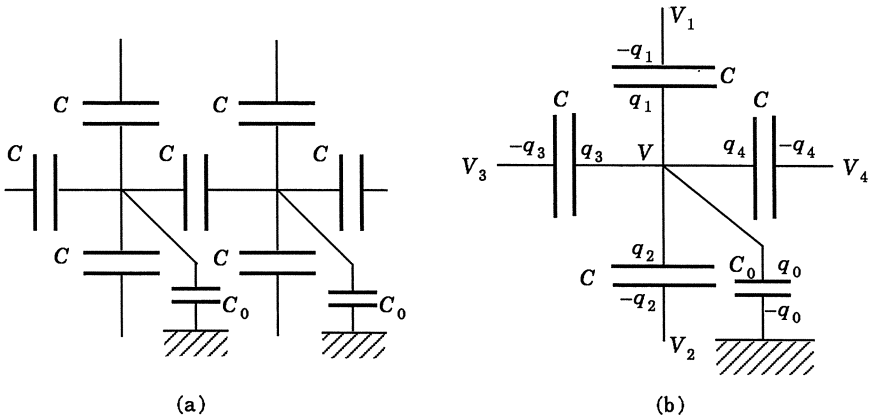


Fig. 1.6.8 (a) A model of the junction network. (b) Charges on the junctions

Hartree–Fock theory.

The large transmission coefficient when odd number of electrons are in the dot is due to the Kondo effect [12]. The total spin in the dot is half-integral and the dot has a magnetic moment. It couples with the electrons outside the dot and forms a resonant state at the Fermi level. In order to observe the Kondo effect, however, the temperature must be lower than the Kondo temperature T_K, and above T_K the conductance behaves in accordance with Hartree–Fock theory, as shown in Fig. 1.6.7b. For realistic values of the parameters characterizing the dot, the Kondo temperature is of order 1 mK, and within the authors' knowledge Kondo-resonant transmission through a quantum dot has not yet been observed. Perfect transmission in one dimension due to the Kondo effect was first discussed in [13,14] on a specific model.

1.6.7 KTB Transition in Junction Network

Consider a network of small tunneling junctions with capacitance C. The model of the system is shown in Fig. 1.6.8a, where the capacitance between the junction and the substrate is modeled by C_0.

Let V_i and q_i $(i = 0,\ 1 \cdots 4)$ be the electric potential and the charge on junction i connected to one of the lattice points, respectively (see Fig. 1.6.8b). Then the electric potential V and the charge q on the lattice point is given by

$$V = V_i + \frac{q_i}{C} = \frac{q_0}{C_0}, \quad (i = 1, 2, \cdots 4), \tag{1.6.21}$$

and

$$q = \sum_{i=0}^{4} q_i. \tag{1.6.22}$$

From these equations we obtain

$$(V_1 - V) - (V - V_3) + (V_2 - V) - (V - V_4) - \frac{C_0}{C} V + \frac{q}{C} = 0. \tag{1.6.23}$$

Here we assume that the differences between electric potentials on neighboring lattice points are small. We write $V(\mathbf{r}_i) \equiv V_i$, where $\mathbf{r}_i$ is the position of the ith lattice point. Then, expanding $V(\mathbf{r}_i) - V(\mathbf{r}) \equiv V_1 - V$ to the second order in the distance d between the nearest lattice points, we obtain

$$\Delta V(\mathbf{r}) - \kappa^2 V(\mathbf{r}) + \frac{q(\mathbf{r})}{C} = 0, \tag{1.6.24}$$

where $q(\mathbf{r})$ is the the charge density at $\mathbf{r}$ and

$$\kappa = \frac{1}{d}\sqrt{\frac{C_0}{C}}. \tag{1.6.25}$$

We consider the case when $q(\mathbf{r}) = Q\delta(\mathbf{r})$, namely, when there is a point charge at $\mathbf{r} = 0$. Then the solution of (1.6.24) is

$$V(\mathbf{r}) = A\,K(\kappa r). \tag{1.6.26}$$

where A is a constant and $K(r)$ is the modified Bessel function. The constant A is determined by Gauss's theorem

$$2\pi r \frac{\partial}{\partial r} V(\mathbf{r}) = \frac{Q}{C}. \tag{1.6.27}$$

Using the asymptotic form $K(\kappa r) \approx -\log(\kappa r)$ $(\kappa r \ll 1)$ we obtain

$$A = \frac{Q}{2\pi C}. \tag{1.6.28}$$

Thus, if $C_0 \ll C$, we find that

$$V(\mathbf{r}) = -A \log(\kappa r), \tag{1.6.29}$$

for a wide range of r. Hereafter we consider an ideal case when it holds for $0 < r < \infty$.

Even if the lattice points are neutral on the average, pairs of electrons and holes are thermally excited. The potential energy $U(r)$ between an electron and a hole is given by from (1.6.29) and (1.6.28) as

$$U(r) = \frac{e^2}{2\pi C} \log(\kappa r). \tag{1.6.30}$$

Because of the long range nature of this attractive potential, the electron and hole form a bound state at least at low temperatures. In order to investigate the feature of the bound pairs, we calculate the electric polarizability of a pair. Suppose we apply a uniform electric field E on the system. Then the dipole moment of a pair is given by

$$p = \frac{\int_{d<r} ex \exp\{-\beta(U(r) - exE)\} d\mathbf{r}}{\int_{d<r} \exp\{-\beta(U(r) - exE)\} d\mathbf{r}}. \tag{1.6.31}$$

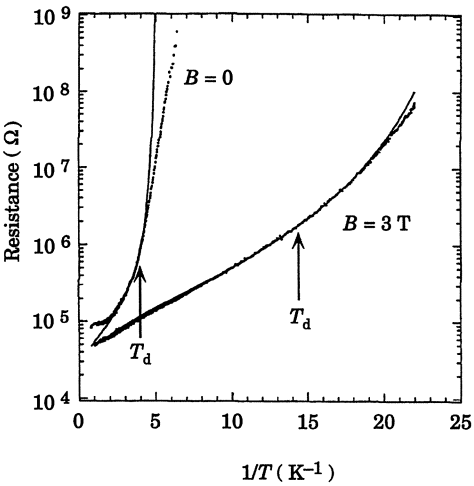


Fig. 1.6.9 The resistance of a junction network. The measurement were made with and without magnetic field (from Ref. [19])

We expand the right hand side of this equation to the lowest order in E and the polarizability α is given by

$$\alpha = \frac{p}{E} = \frac{\frac{e^2}{3}\int_{d<r} \beta r^2 \exp\{-\beta U(r)\} d\mathbf{r}}{\int_{d<r} \exp\{-\beta U(r)\} d\mathbf{r}}. \tag{1.6.32}$$

From (1.6.30) it follows that the integral in the numerator on the right hand side of this equation is

$$2\pi \int_d^\infty (\kappa r)^{-\gamma} r^3 dr, \tag{1.6.33}$$

where $\gamma = e^2/(2\pi C k_B T)$. Therefore we find that the polarizability diverges for $3-\gamma \geq -1$, i.e., for

$$T \geq T_c \equiv e^2/8\pi C k_B. \tag{1.6.34}$$

This means the system exhibits a transition from insulator to conductor at $T = T_c$ [15]. This transition is characteristic to systems in which particles interact via a logarithmic potential and is called KTB (Kosteritz–Thouless–Berezinskii) transition [16,17,18].

The conductance of the junction network is approximately proportional to the density of the dissociated electron-hole pairs. In calculating the temperature dependence of the density, it is important to take into account the interaction between the pairs. In [18], this effect was considered in the form of the dielectric constant, which was determined self-consistently. The reader is referred to the original paper for the details of the theory and here we give

only the result: The temperature dependence of the conductance is given by

$$\sigma \propto \exp\left(-\frac{2b}{\sqrt{T/T_c - 1}}\right), \tag{1.6.35}$$

where b is a numerical constant of order 1.

In analyzing experimental results, it is important to note that this equation is valid only in a limited region of T. On the one hand, it is valid only in the critical region, i.e., when T is close to T_c. On the other hand, when T is very close T_c, the number of pairs where the distance between electron and hole is very large increases, and, as (1.6.30) is valid only for $r < 1/\kappa$, the theory is no longer valid. The lower limit T_d of the temperature for which (1.6.35) holds is given by the relation [19]

$$\frac{4C_0}{\pi C} = K \exp\left(\frac{-2b}{\sqrt{T_d/T_c - 1}}\right), \tag{1.6.36}$$

where K is a numerical constant of order 1.

Such temperature dependences have been observed by Kobayashi et al. [19] on the Al junction network. The results are shown in Fig. 1.6.9. The data were fitted to (1.6.35) and they obtained $T_c = 0.05$ K and 0.19 K with and without a magnetic field, respectively. Without a magnetic field, (1.6.34) has to be modified because junctions are superconducting: Instead of electron-hole pairs, we have to consider the pairs of a Cooper pair and its hole. As a result, we have to replace e in (1.6.34) with $2e$. This is consistent with the observation that T_c without a magnetic field is about four times as large as that in a magnetic field.

References

1 P. Delsing, T. Claeson, K.K. Likharev, and L.S. Kuzmin, Phys. Rev. B **42**, 7439 (1990)
2 M.H. Devoret and H. Grabert, *Single Charge Tunneling. Coulomb Blockade Phenomena in Nanostructures* (Proceedings of a NATO Advanced Study Institute, Les Houches, 1991 (Plenum, London, 1992), p. 1
3 K.K. Likharev and A.B. Zorin J. Low Temp. Phys. **59**, 347 (1985)
4 B.D. Josephson, Phys. Lett. **1**, 251 (1962)
5 C. Kittel, *Quantum Theory of Solids* (John Wiley & Sons), p. 190
6 D.B. Haviland, L.S. Kuzmin, P. Delsing, K.K. Likharev, and T. Claeson, Z. Phys. B **85**, 339 (1991)
7 L.P. Kouwenhoven, N.C. van der Vaart, A.T. Johnson, W. Kool, C.J.P.M. Harmans, J.G. Williamson, A.A.M. Staring, and C.T. Foxon, Z. Phys. B **85**, 367 (1991)
8 A. Kawabata, J. Phys. Soc. Jpn. **58**, 372 (1989)
9 A. Kawabata, J. Phys. Soc. Jpn. **60**, 3222 (1991)
10 J. Friedel, Phil. Mag. **43** (1952) 153; Advance in Physics **3** (1954) 446
11 J.S. Langer and V. Ambegaokar, Phys. Rev. **121**, 1090 (1961)
12 J. Kondo, Prog. Theor. Phys. **32**, 37 (1964)
13 T.K. Ng and P.A. Lee, Phys. Rev. Lett. **61**, 1768 (1988)
14 L.I. Grazman and M.E. Raikh, JETP Lett. **44**, 452 (1988)
15 R. Fazio and G. Schön, Phys. Rev. **43**, 5307 (1991)
16 V.L. Berezinskii, Soviet Phys. JETP **32**, 493 (1971)
17 J.M. Kosterlitz and D.J. Thouless, J. Phys. C **6**, 1181 (1973)
18 J.M. Kosterlitz, J. Phys. C **7**, 1046 (1974)
19 S. Kobayashi, A. Kanda, and R. Yamada, Jpn. J. Appl. Phys. **34**, 4548 (1995)

2. Transport in Quantum Structures

2.1 Tomonaga–Luttinger Liquid in Quantum Wires

2.1.1 Introduction

Landauer's formula discussed in Sect. 1.3.3 gives the quantized conductance in an ideal quantum wire without impurities,

$$G = \frac{2e^2}{h} N, \tag{2.1.1}$$

where N is the number of subbands which cross the Fermi energy E_F. This formula was derived in noninteracting case. In one-dimensional systems, it has long been known that the mutual interaction has strong effects. In fact, each order of perturbation calculations has a divergence and thus the ground state is a non-perturbative one. Such a state is described as a Tomonaga–Luttinger liquid [1], which differs from the conventional Fermi liquid. It exhibits several anomalous behaviors such as a power-law decay of physical quantities and a spin-charge separation [2–5]. It is thus very interesting and important to study quantum wires and to find out such anomalous phenomena characteristic of the Tomonaga–Luttinger liquid.

There is also an ideal one-dimensional electron systems consisting of edge states in the quantum Hall regime [6] in addition to quantum wires [7–9]. The edge states can be specified as a special case of the Tomonaga–Luttinger liquid. In discussing quantum wires, the actual sample geometry, such as contacts between the quantum wire and leads, is known to play important roles. In the following, we assume just a smooth connection between sample and leads for simplicity.

2.1.2 Tomonaga–Luttinger Liquid

There are two anomalous behaviors different from the conventional Fermi liquid: power-law behavior of various physical quantities and the spin-charge separation. For example, the spin and density correlation functions at $T=0$ behave as

$$\begin{aligned} \langle \mathbf{S}_i \cdot \mathbf{S}_{i+r} \rangle &\sim -\frac{1}{\pi^2 r^2} + A \frac{e^{2ik_F r}}{r^{1+K_\rho}}, \\ \langle n_i n_{i+r} \rangle &\sim n^2 - \frac{K_\rho}{\pi^2 r^2} + B \frac{e^{2ik_F r}}{r^{1+K_\rho}} + C \frac{e^{4ik_F r}}{r^{4K_\rho}}, \end{aligned} \tag{2.1.2}$$

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

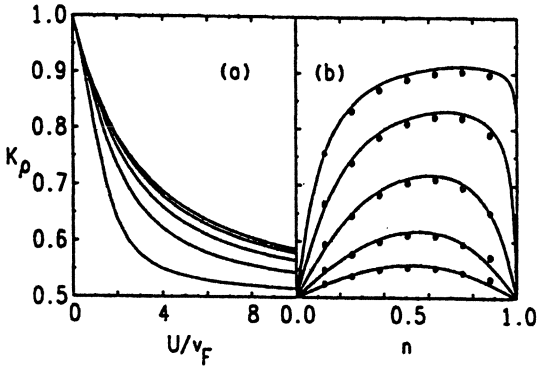


Fig. 2.1.1 Correlation exponent K_ρ in the Hubbard model (**a**) as a function of U/v_F for different values of the electron density (n=0.1, 0.3, 0.5, 0.7, and 0.9 for the top to bottom curves) and (**b**) as a function of electron density for various values of U (U/t=1, 2, 4, 8, and 16 for the top to bottom curves) [10]. Here U and t are the on-site Coulomb energy and and the hopping energy between nearest neighbor sites, respectively. Note the rapid decrease to $K_\rho \to 1/2$ near half-filling (n=1)

and the superconducting pair correlation behaves as

$$\langle \Delta_i \Delta_{i+r} \rangle \sim \frac{D}{r^{1+1/K_\rho}}, \tag{2.1.3}$$

with model-dependent coefficients A–D and a correlation exponent K_ρ. Since the power-law decay depends on K_ρ, it determines the most dominant fluctuation. For the noninteracting case, $K_\rho = 1$ corresponds to Fermi-liquid behavior where every physical quantity has an integer power-law dependence. In the presence of interactions, K_ρ has an arbitrary value depending on the interaction strength as well as the electron density. When the interaction is repulsive, $K_\rho < 1$ and thus $2k_F$ spin-density wave and charge-density wave in (2.1.2) is the dominant fluctuation. When the interaction is attractive, $K_\rho > 1$ and superconducting correlation (2.1.3) is dominant. The actual value of K_ρ has been obtained in one-dimensional solvable models such as the Hubbard model [10–12] and a supersymmetric t-J model [13] using the Bethe ansatz equation. For non-integrable models, K_ρ can be determined numerically [14]. Here for example, K_ρ in the Hubbard model is shown in Fig. 2.1.1.

The power-law behavior manifests itself as a dependence on temperature or frequency. For example, when the nuclear relaxation time is determined mainly from the $2k_F$ component of the spin-spin correlation, we expect $1/T_1T \propto T^{K_\rho - 1}$ [15], which is different from the Korringa relation in the Fermi liquid. Another example is the resistivity due to impurity scattering. Since the impurity backward scattering is related to the $2k_F$ component of the density-density correlation function, we obtain

$$\frac{1}{\tau_{\rm tr}} \sim \frac{1}{\tau_{\rm tr}^0} \left(\frac{k_B T}{\epsilon_F} \right)^{K_\rho - 1}, \tag{2.1.4}$$

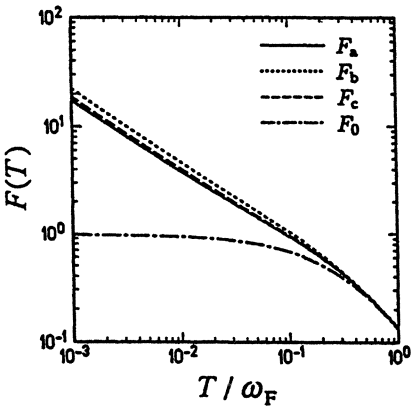


Fig. 2.1.2 Temperature dependence of $F(T)$, defined as $1/\tau_{tr} = F(T)/\tau^0_{tr}$, using realistic parameters of GaAs for three choices of the charge distribution (F_a, F_B, and F_c) in the cross section [16]. The dash-dotted line (F_0) is for the noninteracting electrons

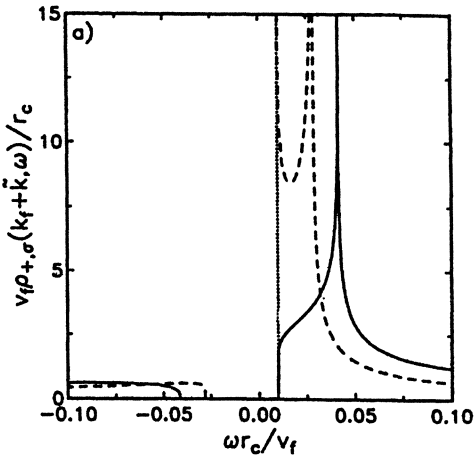


Fig. 2.1.3 Total spectral function (= Im $G(k,\omega)$) of the Tomonaga-Luttinger model as a function of ω for $K_\rho = 0.351$ (dotted curve), and $K_\rho = 0.240$ (full curve) [18]. The momentum is fixed at $k = k_F + 0.01/\Lambda$ with cut-off $1/\Lambda$, and the frequency is normalized by $\omega_F = v_F/\Lambda$. The curve for $K_\rho = 0.240$ is scaled up by a factor of 3

where τ^0_{tr} is the transport relaxation time for noninteracting electrons (Born approximation). For the repulsive interaction, K_ρ is generally less than 1, which indicates that the strength of impurity scattering is renormalized to larger values as temperature decreases. Using realistic parameters for GaAs, Fukuyama et al. [16] calculated the relaxation rate $1/\tau_{tr}$ as a function of T shown in Fig. 2.1.2. Using the renormalization group argument, Giamarchi and Schulz [17] also discussed the temperature dependence of resistivity.

The second characteristic of Tomonaga–Luttinger liquid is the spin-charge separation. Since the Fermi surface of one-dimensional system consists of only two points in k space (k_F and $-k_F$), the so-called collective mode in the Fermi liquid theory is nothing but particle-hole excitation. In other words, all the excitations in one-dimensional systems are described as collective modes. As shown in Fermi liquid theory, the collective mode is strongly affected by the mutual interactions, so that spin and charge excitations behave differently.

Furthermore, a quasiparticle is no longer an eigenstate: Elementary excitations are decomposed into spin- and charge-excitations propagating at velocities v_σ and v_ρ, respectively. As a result, Green's function for the has a two-peak structure at $\omega = v_\sigma(k - k_F)$ and $\omega = v_\rho(k - k_F)$. The behavior of the Green's function [18,19] is shown schematically in Fig. 2.1.3. To observe this kind of breakdown of the quasiparticle picture is another interesting challenge in experiments.

To describe these features of the Tomonaga–Luttinger liquid, it is convenient to use a bosonization technique [1] or a phase Hamiltonian [20]. By

linearizing the energy dispersion around $k=k_F$ and $k=-k_F$ and defining the density operators $\rho_{1\sigma}(x)$ and $\rho_{2\sigma}(x)$, each representing the electrons with spin σ around k_F (branch 1) and $-k_F$ (branch 2), boson operators are defined as

$$b^\dagger_{p\sigma} = \begin{cases} i\sqrt{\frac{2\pi}{pL}}\rho_{1p\sigma} & p>0\ , \\ i\sqrt{\frac{2\pi}{|p|L}}\rho_{2p\sigma} & p<0\ . \end{cases} \tag{2.1.5}$$

Since the forward scattering is represented as $\int dx\rho_1\rho_1$ or $\int dx\rho_1\rho_2$, the interaction can be diagonalized in terms of these boson operators. Note here that backward scattering terms are not included; they are discussed shortly.

The fermion operator is represented as [21,22]

$$\psi_{i\sigma}(x) = \frac{1}{\sqrt{2\pi\alpha}} e^{\pm ik_F x} U_{i\sigma} e^{i\Phi_{i\sigma}(x)}, \tag{2.1.6}$$

where the phase $\Phi_{i\sigma}(x)$ is a Fourier transform of $b_{p\sigma}$ defined as

$$\begin{aligned} \Phi_{1\sigma}(x) &= \sum_{p>0}\sqrt{\frac{2\pi}{pL}} b_{p\sigma} e^{ipx} + \text{h.c.} + \frac{2\pi i}{L} N_{1\sigma} x, \\ \Phi_{2\sigma}(x) &= \sum_{p<0}\sqrt{\frac{2\pi}{|p|L}} b_{p\sigma} e^{ipx} + \text{h.c.} - \frac{2\pi i}{L} N_{2\sigma} x, \end{aligned} \tag{2.1.7}$$

and $U_{i\sigma}$ is a unitary operator which adds one electron with spin σ on the Fermi surface of branch i $(i=1,2)$. This operator ensures the fermion commutation relations.

The charge and spin degrees of freedom are decoupled in the Hamiltonian

$$\begin{aligned} \mathcal{H}_\rho &= \int dx \left\{ A_\rho (\nabla\theta(x))^2 + C_\rho P(x)^2 \right\}, \\ \mathcal{H}_\sigma &= \int dx \left\{ A_\sigma (\nabla\phi(x))^2 + C_\sigma M(x)^2 \right\}, \end{aligned} \tag{2.1.8}$$

where $\theta(x)$ and $P(x)$ are defined as

$$\theta(x) = \frac{1}{2}\sum_\sigma \big(\Phi_{1\sigma}(x) - \Phi_{2\sigma}(x)\big), \quad P(x) = -\frac{1}{4\pi}\sum_\sigma \nabla\big(\Phi_{1\sigma}(x) + \Phi_{2\sigma}(x)\big), \tag{2.1.9}$$

and they satisfy commutation relations

$$[\theta(x), P(y)] = i\delta(x-y), \tag{2.1.10}$$

where $\theta(x)$ describes the fluctuation of charge degrees of freedom. Similarly the spin degrees of freedom are described by $\phi(x)$ and $M(x)$, which are defined as (2.1.9) with different spin summation. The charge and spin velocities are $v_\rho = 2\sqrt{A_\rho C_\rho}$ and $v_\sigma = 2\sqrt{A_\sigma C_\sigma}$, respectively, and the correlation exponent is $K_\rho = (1/2\pi)\sqrt{A_\rho/C_\rho}$.

The phase Hamiltonian (2.1.8) describes the Tomonaga–Luttinger behaviors in the low temperature and low frequency region. Since it is bilinear, various physical quantities can be calculated analytically using the form of $\psi_{i\sigma}(x)$. Finally, in terms of the phase variable, the backward scattering terms becomes

$$\int dx B_\sigma \cos(2\phi(x)). \tag{2.1.11}$$

This term is not bilinear but fortunately the backward scattering term is renormalized to zero in the renormalization group argument [2].

2.1.3 Conductance of Finite-Length Quantum Wire

In the pioneering work by Apel and Rice [23], the conductance per channel was obtained as

$$G = \frac{2e^2}{h} K_\rho, \tag{2.1.12}$$

where K_ρ is the correlation exponent. For the noninteracting case, $K_\rho = 1$ and thus (2.1.12) coincides with the original Landauer formula. Recently there have been several arguments that the coefficient K_ρ should not appear in (2.1.12), which will be discussed in the next section.

A characteristic feature of the Tomonaga–Luttinger liquid lies in the power-law dependences. In the preceding section it was shown that impurity scattering gives an anomalous $1/\tau_{\rm tr}$ with a power-law dependence on T. This T dependence also shows up in the conductance. Apel and Rice [23] showed that impurity scattering is renormalized to strong coupling for $K_\rho < 2$ in the presence of finite impurity density and that G goes to zero as temperature decreases. The same effect of a single impurity was also discussed [24], but in this case the critical K_ρ is 1 because the impurity density vanishes. On the other hand, the conductivity behaves as $\sigma \propto \tau_{\rm tr} \propto T^{1-K_\rho}$ as shown in Fig. 2.1.2. This finite conductivity means that the conductance g $(= \sigma/L)$ is proportional to $1/L$. Apparently the relation between the conductance and conductivity is not trivial.

A formula relating these two quantities was proposed by Ogata and Fukuyama as follows [25,26]:

$$\sigma = \frac{ne^2}{m^*}\left(\frac{2v_F}{L} + \frac{1}{\tau_{\rm tr}}\right)^{-1}, \quad \frac{1}{\tau_{\rm tr}} = \frac{1}{\tau_{\rm tr}^0} F(T, L). \tag{2.1.13}$$

This equation can be obtained from a usual Drude formula with a replacement $-i\omega \to 2v_F/L$. Usually in discussing the static conductivity, the term $-i\omega$ is neglected, but in the mesoscopic systems it is important to keep this term, replacing it with a site-dependent term because $\omega \ll v_F/L$. The formula (2.1.13) naturally interpolates the quantization of conductance for small systems $(2v_F/L \gg 1/\tau_{\rm tr})$ and the usual conductivity for longer systems $(2v_F/L \ll 1/\tau_{\rm tr})$. Note that $n = 2k_F/\pi = 4m^* v_F/h$ and $G = \sigma/L$.

The quantity $F(T, L)$ represents the effect of the impurity backward scattering. In the limit of a long wire $(L \gg \hbar v_F/k_B T)$, $F(T, L)$ behaves as

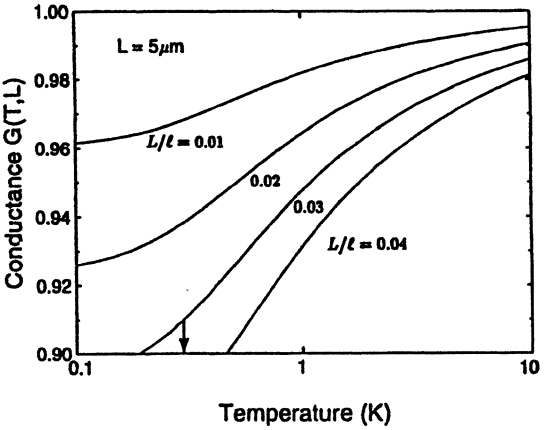


Fig. 2.1.4 Temperature dependence of the dimensionless conductance, $G(T,L)$, for a typical choice of L/ℓ and $T_L = v_F/L$: $T_L = 0.3$ K which corresponds to $L = 5$ μm in the case of $v_F = 2\times10^5$ m/s. The crossover temperature T_L is shown by an arrow [26]

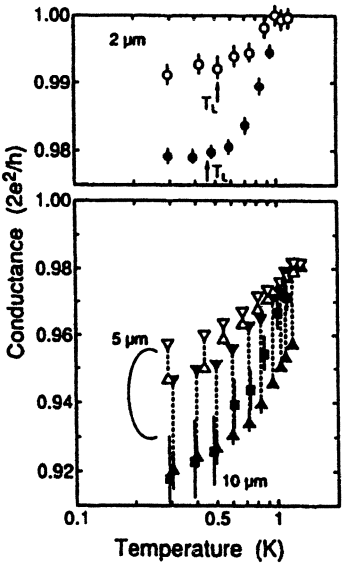


Fig. 2.1.5 The conductance of split-gate quantum wire in GaAs as a function of temperature for various wires [7]. The open and closed symbols represent the wires in the dark and after illumination, respectively

$F(T,L) \propto T^{K_\rho - 1}$. When T is small so that L is comparable to $\hbar v_F/k_B T$, we have to take account of the effect of a finite length. Below the temperature $T_L = \hbar v_F/k_B L$, the thermal length becomes larger than the system size and thus the divergence of $F(T,L)$ as $T \to 0$ is terminated.

Using this formula we obtain the conductance

$$G = \frac{2e^2}{h}\frac{1}{1+(L/2\ell)F(T,L)} = \frac{2e^2}{h}G(T,L), \tag{2.1.14}$$

where $\ell = v_F \tau_{\rm tr}^0$ is elastic mean-free path. The dimensionless conductance $G(T,L)$ thus obtained is shown in Fig. 2.1.4 for several choices of L/ℓ and T_L.

Experimental results in quantum wires were explained by this formula as shown in Fig. 2.1.5 and the Tomonaga–Luttinger parameter K_ρ was determined as $K_\rho \sim 0.7$ [7].

2.1.4 Quantized Value of Conductance

The experimental data in Fig. 2.1.5 shows that G approaches $2e^2/h$ as the temperature increases. This means that the coefficient appearing in (2.1.12) is unity in contrast to $K_\rho \sim 0.7$ determined from the temperature dependence. There have been several theoretical considerations that lead to $G = (2e^2/h)$ instead of $G = (2e^2/h)K_\rho$.

The importance of leads attached to both ends of the quantum wire has been claimed [27–29]. The wire is connected to leads in which the Coulomb interaction is absent. In such an inhomogeneous one-dimensional system, the current-current response function

$$K^R(x,y;t) = -\frac{i}{\hbar}\Theta(t)\langle[j(x,t),j(y,0)]\rangle, \tag{2.1.15}$$

is calculated by solving the equation of motion with boundary conditions at the interfaces between wire and leads ($x=0,L$). In the limit of $\omega \to 0$, the response function is assumed to be

$$K^R(x;y;\omega) = \begin{cases} A\exp(-i\frac{\omega}{v_F}x), & (x<0), \\ B\exp(-i\frac{\omega}{v_\rho}x) + C\exp(+i\frac{\omega}{v_\rho}x) - i\omega\frac{2e^2}{h}K_\rho\exp(i\frac{\omega}{v_\rho}|x-y|), & (0\le x,y\le L), \\ D\exp(+i\frac{\omega}{v_F}x), & (L<x). \end{cases} \tag{2.1.16}$$

Here, inside the quantum wire, $-i\omega(2e^2/h)K_\rho e^{i\omega|x-y|/v_\rho}$ is the same as the response function of Tomonaga–Luttinger liquid in the thermodynamic limit used by Apel and Rice [23]. The coefficients A, B, C, and D are determined from the boundary conditions as $A=D=-i\omega(2e^2/h)K_L$ with K_L being the exponent in the leads, and $B=C=-i\omega(e^2/h)(K_L-K_\rho)$. The conductivity is given by $\sigma(x,y;\omega)=-(i\omega)^{-1}\big(K^R(x,y;\omega)-K^R(x,y;0)\big)$ and in the static limit by $\sigma(x,y;\omega)\to 2e^2K_L/h$ independent of x and y. This shows that the coefficient of the conductance is K_L instead of K_ρ. In the absence of the interaction in the leads, i.e., $K_L=1$, the conductance is the same as than in the noninteracting case.

Kawabata [30] showed that if the renormalization of the electric field due to the mutual interaction is taken into account correctly, K_ρ in the conductance does not show up. Diagrammatically the effective electric field is screened by the polarization bubbles. Therefore, the conductivity is not given by the full current-current response function but by an irreducible one. This irreducible response function becomes exactly the same as that of a free electron gas due to a Ward identity, leading to the conclusion $G=2e^2/h$. See Sect. 1.5 also on this subject.

The same conclusion has been obtained also using Landauer's formula [31,32]. Since the electrons are classified into two branches as right-going (branch 1) and left-going (branch 2), a current-carrying state is realized by choosing $N_1 \neq N_2$, where N_i $(i=1,2)$ is the number of excess electrons on each branch. The uniform current is $I=ev_J(N_1-N_2)/L$ [22]. Here v_J is a velocity for the current, $v_J=v_\rho K_\rho$, which generally depends on the interaction strength. On the other hand, the chemical potentials of right- and left-going electrons are determined from

$$\mu_1 = \frac{\partial\langle H\rangle}{\partial N_1}, \quad \mu_2 = \frac{\partial\langle H\rangle}{\partial N_2}. \tag{2.1.17}$$

Since the charge degrees of freedom in the uniform part of the Hamiltonian

($p=0$ component) are written as [22]

$$\frac{\pi\hbar}{4L}\left[v_N(N_1+N_2)^2+v_J(N_1-N_2)^2\right], \tag{2.1.18}$$

the difference of the chemical potentials is $\mu_1-\mu_2=(\pi\hbar/L)v_J(N_1-N_2)$. Thus, the conductance becomes $G=I/V=eI/(\mu_1-\mu_2)=2e^2/h$ in which the interaction-dependent parameter v_J has canceled out exactly.

2.1.5 Mott–Hubbard Insulator

In usual quantum wires, the theoretical considerations given in the previous section show that the prefactor K_ρ does not appear in the conductance. The situation may change in 1D lattices near a half-filling with a strong on-site repulsion like systems described by a Hubbard model because Umklapp scattering and also backward scattering may play important roles.

In the Hubbard model, $K_\rho \to 1/2$ toward half-filling and in the strong coupling limit ($U\to\infty$) as shown in Fig. 2.1.1. In this limit the conductance $(2e^2/h)K_\rho$ is reduced to e^2/h. This has a clear intuitive meaning that electrons are described as spinless fermions because of a strong repulsion and as a result the conductance is reduced to half of the noninteracting value corresponding to a lack of spin degeneracy.

Such a system near half-filling can be realized in a quantum wire modulated with a periodic potential [26]. Quite recently, a split gate quantum wire with an extra periodic potential was fabricated on GaAs [33].

The conductance of two-terminal measurement was calculated [34] by taking account of the leads attached to both ends of quantum wire. Since the Umklapp term is nonlinear as shown in (2.1.11), the Luther–Emery method [35] was adopted for the charge degrees of freedom [36,37], which maps the phase Hamiltonian into a spinless fermion model. It turns out that at $T=0$ the conductance is still given by $G=2e^2/h$ near half-filling, even though Umklapp scattering opens a Mott gap in the energy spectra. This is because the system behaves as a Tomonaga–Luttinger liquid except just at half-filling. For $T\neq 0$ or in the low-temperature region, however, the conductance behaves as [34]

$$G=\left[1-f(\Delta)\right]\frac{2e^2}{h}, \tag{2.1.19}$$

where $f(\epsilon)$ is the Fermi distribution function given by $f(\epsilon)=[1+\exp(\epsilon/k_BT)]^{-1}$ and Δ is the energy difference between the Fermi energy and the top of the lower Hubbard band. If we extrapolate this result to higher temperatures as $k_BT>\Delta$, the conductance approaches $G=e^2/h$ since $f(\Delta)\to 1/2$ as $\Delta/k_BT\to 0$, in agreement with the intuitive picture discussed above.

References

1 S. Tomonaga, Prog. Theor. Phys. **5**, 544 (1950); J.M. Luttinger, J. Math. Phys. **4**, 1154 (1963); D.C. Mattis and E. Lieb, J. Math. Phys. **6**, 304 (1965)
2 J. Sólyom, Adv. Phys. **28**, 201 (1979)

3 V.J. Emery, *Highly Conducting One-Dimensional Solids*, edited by J.T. Devreese, R.P. Evrard, and V.E. van Doren (Plenum, New York, 1979), p. 247
4 H. Fukuyama and H. Takayama, *Electronic Properties of Inorganic Quasi-One-Dimensional Materials, I*, edited by P. Monceau (Reidel, 1985) p. 41
5 G.D. Mahan, *Many-Particle Physics* (Plenum, New York, 1990), p. 324
6 F.P. Milliken, C.P. Umbach, and R.A. Webb, Solid State Commun. **97**, 309 (1996); A.M. Chang, L.N. Pfeiffer, and K.W. West, Phys. Rev. Lett. **77**, 2538 (1996)
7 S. Tarucha, T. Honda, and T. Saku, Solid State Commun. **94**, 413 (1995)
8 A. Yacoby, H. L. Stormer, N.S. Wingreen, L.N. Pfeiffer, K.W. Baldwin, and K.W. West, Phys. Rev. Lett. **77**, 4612 (1996)
9 J.I. Pascual, J. Méndez, J. Gómez-Herrero, A.M. Baró, N. García, and V.T. Binh, Phys. Rev. Lett. **71**, 1852 (1993)
10 H.J. Schulz, Phys. Rev. Lett. **64**, 2831 (1990)
11 H. Frahm and V.E. Korepin, Phys. Rev. B **42**, 10553 (1990); **43**, 5653 (1991)
12 N. Kawakami and S.-K. Yang, Phys. Lett. A **148**, 359 (1990)
13 N. Kawakami and S.-K. Yang, Phys. Rev. Lett. **65**, 2309 (1990)
14 M. Ogata, M.U. Luchini, S. Sorella, and F.F. Assaad, Phys. Rev. Lett. **66**, 2388 (1991)
15 C. Bourbonnais *et al.*, J. de Phys. Lettres **45**, L755 (1984)
16 H. Fukuyama, H. Kohno, and R. Shirasaki, J. Phys. Soc. Jpn. **62**, 1109 (1993)
17 T. Giamarchi and H.J. Schulz, Phys. Rev. B **37**, 325 (1988)
18 V. Meden and K. Schönhammer, Phys. Rev. B **46**, 15753 (1992)
19 J. Voit, Phys. Rev. B **47**, 6740 (1993)
20 Y. Suzumura, Prog. Theor. Phys. **61**, 1 (1979)
21 A. Luther and I. Peschel, Phys. Rev. B **9**, 2911 (1974)
22 F.D.M. Haldane, J. Phys. C **14**, 2585 (1981)
23 W. Apel and T.M. Rice, Phys. Rev. B **26**, 7063 (1982)
24 C.L. Kane and M.P.A. Fisher, Phys. Rev. Lett. **68**, 1220 (1992); A. Furusaki and N. Nagaosa, Phys. Rev. B **47**, 4631 (1993); A. Furusaki, Thesis (University of Tokyo, 1993)
25 M. Ogata and H. Fukuyama, Phys. Rev. Lett. **73**, 468 (1994)
26 M. Ogata and H. Fukuyama, Jpn. J. Appl. Phys. **34**, 4284 (1995)
27 D.L. Maslov and M. Stone, Phys. Rev. B **52**, R5539 (1995)
28 V.V. Ponomarenko, Phys. Rev. B **52**, R8666 (1995)
29 I. Safi and H.J. Schulz, Phys. Rev. B **52**, R17040 (1995)
30 A. Kawabata, J. Phys. Soc. Jpn. **65**, 30 (1996)
31 Y. Oreg, and A.M. Finkel'stein, Phys. Rev. B **54**, R14265 (1996); Phys. Rev. Lett. **74**, 3668 (1995)
32 A. Shimizu, J. Phys. Soc. Jpn. **65**, 1162 (1996); **65**, 3096 (1996) (Errata)
33 S. Tarucha, private communication
34 M. Mori, M. Ogata and H. Fukuyama, preprint
35 A. Luther and V.J. Emery, Phys. Rev. Lett. **33**, 589 (1974)
36 T. Giamarchi, Phys. Rev. B **44**, 2905 (1991)
37 M. Mori and H. Fukuyama, J. Phys. Soc. Jpn. **65**, 3604 (1996)

2.2 Quantum Wires

2.2.1 Magnetoresistance and Boundary-Roughness Scattering

The conductance of a perfect quantum wire in which N subbands ($\equiv$ modes) are occupied is $2e^2N/h$, as discussed in Sect. 1.5.2, but it will be reduced because of the scattering of electrons by impurities or boundary roughness (fluctuations of the width of the wire). Magnetoresistance provides useful information on the scattering mechanism in quantum wires. For instance, when a magnetic field is applied perpendicular to the wire, the electrons are pushed toward the boundaries and hence we can expect a considerable change in the scattering probability if the boundary roughness is a dominant scattering mechanism.

We consider a two-dimensional electron gas confined in the region $|y| \leq W_0/2$. The wave functions of the eigenstates can be written in the form

$$\psi_{nk}(x,y) = \eta_n(y)\exp(ikx), \tag{2.2.1}$$

with

$$\eta_n(y) = \sqrt{\frac{2}{W_0}}\sin\left(\frac{y+W_0/2}{W_0}n\pi\right). \tag{2.2.2}$$

The energy eigenvalue of this state is $E_{nk}=(\hbar^2/2m)[k^2+(n\pi/W_0)^2]$, and for a given value of the Fermi energy E_F, N is given as the largest n which satisfies $E_{n0} < E_F$. The Fermi wave number of the nth subband k_n is given as the value of k satisfying $E_{nk}=E_F$.

Here we introduce boundary roughness by assuming a boundary of the form

$$y = \pm W_0/2 + \Delta_\pm(x). \tag{2.2.3}$$

We also assume the Gaussian correlation for $\Delta_\pm(x)$, i.e., $\langle\Delta_-(x)\rangle=\langle\Delta_+(x)\rangle=\langle\Delta_-(x)\Delta_+(x')\rangle=0$ and

$$\langle\Delta_-(x)\Delta_-(x')\rangle = \langle\Delta_+(x)\Delta_+(x')\rangle = \Delta^2\exp\left[-\frac{(x-x')^2}{\Lambda^2}\right]. \tag{2.2.4}$$

The conductivity is given by [1]

$$\sigma = \frac{4e^2}{h}\sum_{n=1}^{N}\ell_n, \tag{2.2.5}$$

where the mean free path ℓ_n of the subband n is the solution of the coupled equations

$$\sum_{n'=1}^{N}(K_{n-,n'+}-K_{n+,n'+})\ell_{n'} = 1 \quad (n=1,2,\cdots,N). \tag{2.2.6}$$

Here, the scattering probability $K_{n-,n'+}$ is defined as

$$K_{n\pm,n'+} = \frac{\sqrt{\pi}}{4}\Lambda\Delta^2\frac{(Y^+_{n,n'})^2+(Y^-_{n,n'})^2}{k_nk_{n'}}\exp\left[-\frac{1}{4}\Lambda^2(k_n\mp k_{n'})^2\right], \tag{2.2.7}$$

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

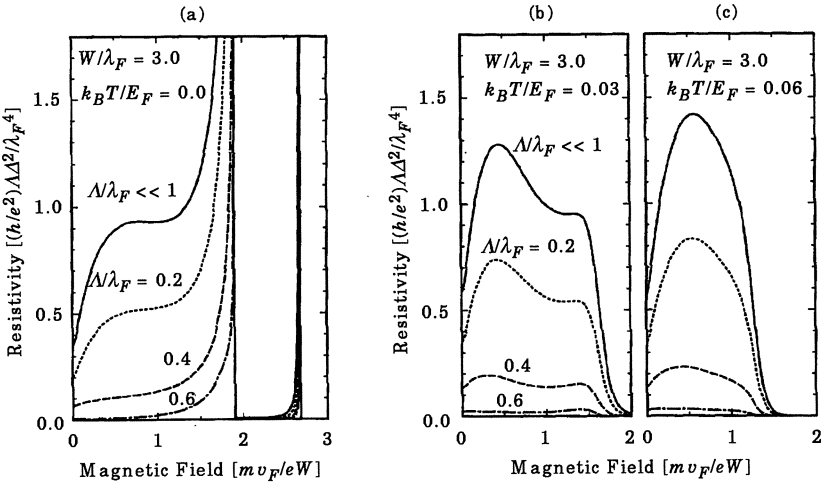


Fig. 2.2.1 Magnetic field dependence of the resistivity of a quantum wire. Here $v_F \equiv \hbar k_F/m$ is the Fermi velocity, and the abscissa is W/R_c with R_c being the cyclotron radius. After [1]

except for $K_{n+,n+}$, with

$$Y^{\pm}_{n,n'} = \mp \frac{d\eta_n(y)}{dy}\frac{d\eta_{n'}(y)}{dy}\bigg|_{y=\pm W_0/2}, \tag{2.2.8}$$

and

$$K_{n+,n+} = -\sum_{n'\neq n} K_{n'+,n+} - \sum_{n'} K_{n'-,n+}. \tag{2.2.9}$$

In the presence of a magnetic field, some modifications are needed for the above expressions. For details, the reader is referred to [1]. Figure 2.2.1 shows the magnetic field dependence of the resistivity for $W/\lambda_F = 3$ and various values of temperatures and the correlation length Λ, where k_F and λ_F are defined by $E_F = \hbar^2 k_F^2/2m$ and $\lambda_F = 2\pi/k_F$ for the Fermi energy E_F.

In the absence of a magnetic field, electrons in lower subbands (small subband index n) have a larger mean free path, because the amplitude of the wave function near the boundary is smaller and the scattering is less frequent.* Therefore, the current is carried mostly by electrons in lower subbands. In a magnetic field, electrons are pushed toward the boundaries by a Lorentz force and the scattering is enhanced. This is the explanation of the increase of the resistance in small field regions. In large field regions where $W/R_c \gtrsim 2$ (R_c is the cyclotron radius), edge states are formed. An edge state is an analog of

* Classically, electrons in lower subbands have a larger kinetic energy for motion along the wire, i.e., moving straight along the wire, and therefore are scattered by the boundary less frequently.

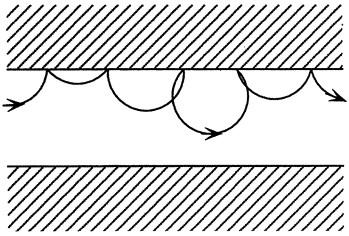


Fig. 2.2.2 A classical interpretation of edge states

the classical skipping orbit shown in Fig. 2.2.2. If $R_c < W/2$, the backward scattering never occurs only with boundary scattering and the resistance vanishes. The rapid decrease of the resistance for large magnetic fields in (b) and (c) can be interpreted in this way.

The large peaks in Fig. 2.2.1a reflect the divergent density of states at the bottom of each subband, which behaves like $1/|k|$. The number of states below E_F in each subband increases with the increase of the magnetic field. Therefore, the number of electrons in the highest occupied subband decreases and the density of states at E_F increases. It gives rises to large scattering probabilities into this subband from others, which results in a large resistance. The resistance drops suddenly just after this subband becomes empty. Thus, the position of these peaks is dependent not only on W/R_c but also on W/λ_F. Those peaks are smeared out at finite temperatures. This characteristic magnetoresistance due to boundary roughness scattering has been observed experimentally [2].

We find that the resistance is smaller for larger Λ. The reason is clear from (2.2.7). When the boundary varies slowly, the Fourier components of the boundary function $W(x)$ for backward scattering become small. The theory reviewed above is based, however, on the lowest-order Born approximation. In the following we will show that this approximation becomes worse for slower variation of the boundaries.

2.2.2 One-Dimensional Electron in Slowly Varying Potential

In Sect. 1.5.1 we have seen that slowly varying fluctuations of boundaries can be transformed into a form of an effective potential depending on subbands. When only a single subband is occupied by electrons, in particular, a quantum wire with a slowly varying boundary roughness can be replaced by a one-dimensional system in a slowly varying random potential.

It is instructive to investigate the scattering by a Gaussian potential

$$U(x) = \frac{\hbar^2 u}{2m} \exp\left(-\frac{x^2}{2\Lambda^2}\right). \tag{2.2.10}$$

Here slowly varying means that $\Lambda k_F \gg 1$. The reflection coefficient $\mathcal{R}$ can be expressed with T-matrix $\mathcal{T}$ defined by

$$\mathcal{R} = \frac{1}{4k_F^2}|\mathcal{T}|^2, \tag{2.2.11}$$

where k_F is the Fermi wave number.

To the lowest order in $U(x)$ (the first Born approximation), $\mathcal{T}$ is given by

$$\mathcal{T}_1 = \hat{U}(-2k_F) = \sqrt{2\pi}\Lambda u \exp(-2\Lambda^2 k_F^2), \tag{2.2.12}$$

where $\hat{U}(k)$ is the Fourier transform of $U(x)$. This shows that the reflection coefficient decreases rapidly as Λ increases. The second order term is given by

$$\begin{aligned}\mathcal{T}_2 &= \int \frac{dk}{2\pi} \frac{\hat{U}(-k-k_F)\hat{U}(k-k_F)}{k_F^2-k^2} \\ &= 2\pi\Lambda^2 u^2 \int \frac{dk}{2\pi} \frac{\exp[-\Lambda^2(k_F^2+k^2)]}{k_F^2-k^2}.\end{aligned} \tag{2.2.13}$$

In this equation only the regions $|k| \lesssim 1/\Lambda \ll k_F$ mainly contribute to the integrals because of the condition $\Lambda k_F \gg 1$. Therefore, in the denominator of the integral we can neglect k^2 and we can easily carry out the integral:

$$\mathcal{T}_2 = \frac{\sqrt{\pi}\Lambda u^2}{k_F^2} \exp(-\Lambda^2 k_F^2). \tag{2.2.14}$$

Comparing this with (2.2.12), we find that the first Born approximation is correct only in extremely limited regions $u \ll k_F^2 \exp(-\Lambda^2 k_F^2)$, and we expect that the higher-order terms become more important for larger Λ.

We can extend the calculation to higher orders and have the following result [3]

$$\mathcal{T} = \sqrt{2\pi}\Lambda \sum_{n=1}^{\infty} \frac{u^n}{\sqrt{n}} \Big[\frac{n^n}{(2k_F)^{n-1}n!}\Big]^2 \exp\Big(-\frac{2\Lambda^2 k_F^2}{n}\Big). \tag{2.2.15}$$

This series is convergent for $|u| < 4k_F^2 e^{-2\Lambda^2 k_F^2}$. We can pick up the largest contributions using the saddle point method, and from (2.2.11) we obtain

$$\mathcal{R} = e^{-8\Lambda k_F \beta}, \tag{2.2.16}$$

with $\beta = \sqrt{\log(2k_F/\sqrt{u}e)}$ $(u > 0)$. This result can be compared with the numerical calculation and the agreements are fairly good for certain ranges of parameters. Thus we find that the reflection coefficient does not decrease as rapidly as in (2.2.12).

The above idea can be extended to a random potential with a Gaussian correlation

$$\langle U(x)U(x')\rangle = U_0^2 \exp\Big[-\frac{(x-x')^2}{\Lambda^2}\Big], \tag{2.2.17}$$

corresponding to the boundary roughness discussed in Sect. 2.2.1. The resistivity is given by

$$\rho = \frac{2\sqrt{2}}{\pi\Lambda}(C_2\beta S^2)^{\Lambda k_F \beta/\sqrt{2}} \sqrt{\frac{\sqrt{2}\pi\Lambda k_F}{4\beta-\beta^3}}, \tag{2.2.18}$$

where

$$S = \frac{\sqrt{\pi}\Lambda}{k_F^3}\left(\frac{m}{\hbar^2}U_0\right)^2, \tag{2.2.19}$$

and

$$\beta = \sqrt{\frac{2}{\log(1/C_1 S)}}, \tag{2.2.20}$$

with $C_1 = e^4/4\sqrt{2\pi} \approx 5.45$ and $C_2 = e^7/32\pi \approx 10.9$. The result shows that the resistivity does not decrease with the increase of Λ as rapidly as expected from the first Born approximation. For details of the derivation of these formula, the reader is referred to [3].

2.2.3 Interaction Effects in Quantum Wires

The theory of 1D interacting electron systems can be applied to quantum wires in which only the lowest subband is occupied, as discussed in Sect. 2.2.1. It has to be modified for multi-subband quantum wires.

First, we review the case of one-dimensional systems as a model of single-subband wires. For simplicity we consider scattering of electrons by a potential $U(x)$ which vanishes for $|x| > a$ with a constant a. The potential gives rise to fluctuations of the electron density, which behave like

$$n(x) \propto \frac{\sin(2k_F|x|+\delta)}{|x|}, \tag{2.2.21}$$

for $|x| \gg a$, where δ is a constant dependent on the structure of $U(x)$ (a Friedel oscillation [4]). The potential energy of the Coulomb interaction with those charge fluctuations behaves in the same way, and we easily find that its matrix element between the plane-wave states with wave number k_F and $-k_F$ is logarithmically divergent at $x \to \pm\infty$. Thus even if the amplitude of the electron density fluctuation is small, the backward scattering amplitude of electrons on the Fermi level is infinite within the lowest order Born approximation. This is the origin of the anomalous interaction effects in one dimension.

Such anomalies show up even in the perturbation theory. To the first order in the interaction, the backward scattering amplitude of electrons with wave number k is given by*

$$S_{k,-k} = \hat{U}(-2k_F)\left[1+(\alpha_2-2\alpha_1)\log\left|\frac{k+k_F}{k-k_F}\right|\right], \tag{2.2.22}$$

where $\hat{U}(q)$ is the Fourier transform of $U(x)$ and

$$\alpha_1 \equiv \frac{m\hat{V}(2k_F)}{2\pi\hbar^2 k_F}, \qquad \alpha_2 \equiv \frac{m\hat{V}(0)}{2\pi\hbar^2 k_F}, \tag{2.2.23}$$

* We consider the case $k \approx k_F$ and replace k with k_F except in the singular term.

with $\hat{V}(q)$ being the Fourier transform of the potential $V(x)$ for the electron-electron interaction. In (2.2.22) the α_2 term represents the interaction of an electron with the charge fluctuations and the α_1 term arises from the exchange interaction. If $\alpha_2 > 2\alpha_1$, $S_{k,-k}$ diverges for $k \to k_F$. This divergence is spurious, and we have to include higher order terms in order to obtain convincing results. This can be done by a renormalization group method [5] or a parquet graph method [6,7]. The result is

$$S_{k,-k} = \frac{\hat{U}(-2k_F)}{\left[1+2\alpha_1 \log\left|\frac{k+k_F}{k-k_F}\right|\right]^{3/4}} \left|\frac{k+k_F}{k-k_F}\right|^{\alpha_2-\alpha_1/2}. \tag{2.2.24}$$

In fact, if we expand the right hand side of this equation to the first order in α_1 and α_2, we obtain (2.2.22).

At low temperatures, electrons with wave number k such that $|(\hbar^2k^2/2m) - (\hbar^2k_F^2/2m)| \lesssim k_BT$ take parts in the conduction. Therefore, the temperature dependence of the scattering amplitude relevant to conduction can be obtained by putting $|k_F - k| \approx mk_BT/\hbar^2k_F$. Then, when $\alpha_2 > \alpha_1/2$, we find that it diverges at 0 K. In order to obtain a proper expression for the scattering amplitude, we must take into account higher order terms in $U(x)$. The reflection coefficient of an electron by the potential is given by [5]

$$\mathcal{R}(T) = \frac{\mathcal{R}_0}{\mathcal{R}_0 + (1-\mathcal{R}_0)\left[1+2\alpha_1 \log\left(\frac{T_F}{T}\right)\right]^{3/2}\left(\frac{T}{T_F}\right)^{2\alpha_2-\alpha_1}}, \tag{2.2.25}$$

where $T_F = E_F/k_B$ and $\mathcal{R}_0$ is the reflection coefficient without the interaction effects.

If the interaction potential $V(x)$ is the Coulomb interaction, $\hat{V}(q)$ diverges logarithmically at $q=0$. In this case we can show that $V(x)$ has to be replaced by a screened Coulomb interaction and have $\alpha_2 \approx 1/2 \gg \alpha_1$ for a realistic model [7]. Then we find that at low temperatures the conductance behaves like

$$G(T) = \frac{2e^2}{h}[1-\mathcal{R}(T)] \propto T. \tag{2.2.26}$$

When the scatterer is distributed uniformly over the wire, the Friedel oscillation is disturbed by them and this temperature dependence will be modified at very low temperatures, although the conductance of infinitely long wire vanishes at $T=0$ because of the localization.

Next, we briefly discuss the interaction effects in multi-subband quantum wires. We use the notations used in Sect. 2.2.1. To the lowest order in the interaction, the scattering amplitude $S_{k,-k'}$ from $k \approx k_n$ to $k' \approx -k_{n'}$ is divergent like $\ln[k-k'-(k_n+k_{n'})]$. Using a parquet graph method or a renormalization group method, we can show that [8]

$$S_{k_n,-k_{n'}} \propto \left(\frac{T_F}{T}\right)^{\alpha_{nn'}}, \tag{2.2.27}$$

where $\alpha_{nn'} \approx 1/2N$ for a realistic model with the Coulomb interaction and N

is the number of occupied subbands. In order to calculate the conductivity, we have to solve the coupled equations (2.2.6), where $K_{n-,n'+}$ is dependent on T like $T^{-2\alpha_{nn'}}$ and $K_{n-,n'+}$ is weakly dependent on T. Thus the temperature dependence of the conductance is given by

$$G(T) \propto T^{1/N}, \tag{2.2.28}$$

The exponent $1/N$ is mainly due to the screening of the interaction. Generally, the interaction effects characteristic of quasi one-dimensional systems are weaker for larger N and are dominated by other interaction effects which are important in higher dimensions.

References

1 H. Akera and T. Ando, Phys. Rev. B **43**, 11676 (1991)
2 T.J. Thornton, M.L. Roukes, A. Scherer, and B.P. Van de Graag, Phys. Rev. Lett. **63**, 2128 (1989)
3 A. Kawabata, J. Phys. Soc. Jpn. **61**, 3220 (1992)
4 J. Friedel, Phil. Mag. **43**, 153 (1952)
5 K.A. Matveev, Y. Dongxiao, and L.I. Glazman, Phys. Rev. Lett. **71**, 3351 (1993)
6 Yu.A. Bychkov, L.P. Gor'kov, and I.E. Dzyaloshinskii, Sov. Phys. JETP **23**, 489 (1966)
7 A. Kawabata, J. Phys. Soc. Jpn. **63**, 2047 (1994)
8 A. Kawabata and T. Brandes, J. Phys. Soc. Jpn. **65**, 3712 (1996)

2.3 Magnetophonon Resonance in Quantum Wires

2.3.1 Introduction

The magnetophonon resonance (MPR) effect has been used since the mid-1960s to investigate effective mass, optical–phonon energy and electron-phonon interaction in bulk III–V compounds [1–5]. Several authors have also studied MPR effect in two-dimensional (2D) systems, such as GaAs/AlGaAs, GaInAs/InP, and GaInAs/AlInAs heterostructures [6]. The MPR effect manifests itself as an oscillatory behavior of transverse–conductivity σ_{xx} as a function of applied magnetic field B. At high magnetic fields, electrons move perpendicularly to both electric and magnetic fields, and the transverse-current is carried by electron-hopping motion induced by some scattering mechanisms. The MPR effect arises from the resonant absorption or emission of longitudinal-optical (LO) phonons by electrons when the Landau level is sharp and well defined, and scattering by LO phonons makes a significant contribution to limiting electron mobility. As the magnetic field increases, a small part of σ_{xx} resonantly increases each time the integral multiple of the Landau level spacing becomes equal to the LO phonon energy, because the scattering of the electrons takes place with resonant absorption or emission of LO phonons. In bulk materials and 2D systems, the oscillatory part of σ_{xx} is therefore proportional to the Fröhlich coupling constant α, and the resonance condition where σ_{xx} becomes maximal is written as

$$\hbar\omega_{\mathrm{LO}} = P\hbar\omega_c \quad (P = 1, 2, 3, \cdots), \tag{2.3.1}$$

where $\hbar\omega_{\mathrm{LO}}$ is the LO phonon energy, P the resonance index, and $\hbar\omega_c = \hbar eB/m$ the cyclotron energy. From the measurement of the MPR effect, effective mass m or LO phonon energy $\hbar\omega_{\mathrm{LO}}$ can be deduced, provided that one of the two is known. In order to observe the MPR effect, the temperature must be high enough to have a sufficient phonon population, but not too high, since the thermal broadening of the Landau levels reduces the oscillation amplitude. Optimal temperatures are known to be usually in the range 100–250 K for III–V materials.

In one-dimensional systems, Vasilopoulous et al. [7] theoretically predicted that additional electrical confinement leads to a shift of the resonance magnetic fields to the lower side. For quantum wire with a parabolic confining potential of $V(y) = (1/2)m\Omega^2 y^2$, they have shown that the usual resonance condition (2.3.1) becomes

$$\hbar\omega_{\mathrm{LO}} = P\hbar\widetilde{\omega}_c \quad (P = 1, 2, 3, \cdots) \tag{2.3.2}$$

with $\widetilde{\omega}_c = ({\omega_c}^2 + \Omega^2)^{1/2}$. This shift in resonance fields was experimentally observed by Berthold et al. in GaAs/AlGaAs quantum wires [8] showing that the MPR is an alternative magnetotransport tool for the determination of the lateral confinement of quantum wires at higher temperatures. Recently Ploner et al. have reported systematic investigation of the MPR effect in quantum wires fabricated by shallow etching on GaAs/AlGaAs heterostructures [9]. They analyzed the observed resonance fields by using (2.3.2), and found that

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

the subband spacing increases with decreasing carrier density and the magnetophonon mass increases rapidly with increasing confinement. Although resonance fields shift to lower values for quantum wires, the qualitative nature of magnetoconductivity in 1D systems is the same in bulk or 2D systems as was shown in [7]. However, a natural question arises when one considers very narrow quantum wires. In the limit of strong confinement, electrons are forced by the confining potential to flow along the quantum wire, and LO phonon scattering disturbs the motion. In such a case, one can expect a qualitative difference in magnetic field dependence of σ_{xx}, i.e., a minimum of σ_{xx} at resonance. This problem was first treated by Mori et al. [10] showing a qualitative change in σ_{xx} with confining potential, and the analytical expression for the oscillatory part of σ_{xx} was obtained by Ryu and O'Connell [11]. The theory of the MPR effect in quantum wires [10] is briefly described next.

2.3.2 Theory

A simple model for a quantum wire [7] is adopted for evaluating magnetoconductivity, in which a 2D electron gas formed in heterostructures is confined by narrow gates or split gates, and electrons are free only along the x direction. The heterointerface is defined to be normal to the z axis. The confinement in the y direction is assumed to be characterized by a parabolic potential of frequency Ω. When the magnetic field is applied in the z direction, the one-electron Hamiltonian can be written as

$$H_{\mathrm{e}} = \frac{1}{2m}(\mathbf{p}+e\mathbf{A})^2 + V(y) + H(z), \tag{2.3.3}$$

in which $\mathbf{A}=(-By,0,0)$ is the vector potential of the magnetic field using the Landau gauge, $V(y)=(1/2)m\Omega^2y^2$, and $H(z)$ is the confining potential due to the heterointerface. The eigenstates $|\lambda\rangle$ and eigenvalues ε_λ can be obtained to be

$$\langle \mathbf{r}|\lambda\rangle = \langle \mathbf{r}|nks\rangle = e^{ikx}\phi_n(\widetilde{\ell};\, y-\widetilde{Y})\,\xi_s(z), \tag{2.3.4}$$

$$\varepsilon_\lambda = \varepsilon_{nks} = \left(n+\frac{1}{2}\right)\hbar\widetilde{\omega}_c + \frac{\hbar^2k^2}{2\widetilde{m}} + \varepsilon_s, \tag{2.3.5}$$

where n $(=0,1,2,\cdots)$ is the Landau level index, k the wave vector in the x direction, s the subband index associated width the quantized z motion, $\widetilde{Y} = (\omega_c/m\widetilde{\omega}_c{}^2)k$, and $\widetilde{m} = m(\widetilde{\omega}_c/\Omega)^2$. The wave functions $\phi_n(\widetilde{\ell};y)$ are simple harmonic-oscillator-like solutions with effective magnetic length $\widetilde{\ell} = (m\widetilde{\omega}_c)^{-1/2}$. In the following, only the lowest subband in the z direction $(s=0)$ is considered, and the wave function $\xi_0(z)$ is assumed to be $|\xi_0(z)|^2=\delta(z)$ for the sake of simplicity.

Using the above basis of (2.3.4) and the Fröhlich electron-phonon interaction potential $M(q)$, the Hamiltonian for electron and phonon system is written as the usual form [10], and σ_{xx} can be evaluated by using the Kubo formula

$$\sigma_{xx} = -\lim_{\omega\to 0}\frac{\operatorname{Im}\pi(\omega+i0)}{\omega}, \tag{2.3.6}$$

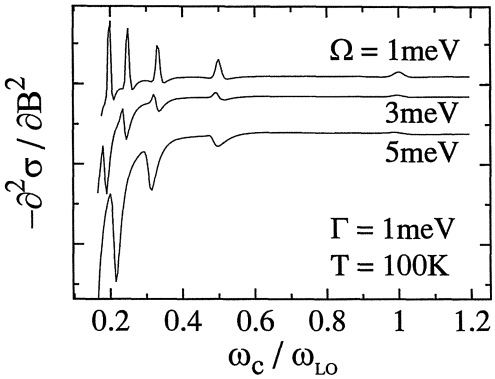


Fig. 2.3.1 The second derivative of magnetoconductivity $\sigma_{xx} = \sigma_{\rm ep} + \sigma_{\rm po}$ for quantum wires with confining potential $\hbar\Omega = 1$, 3, and 5 meV as a function of the normalized cyclotron frequency

where $\pi(\omega)$ is the current-current correlation function

$$\pi(i\omega) = -\int_0^{\hbar\beta} \frac{1}{\hbar} \mathrm{Tr}\left[T_\tau j_x(\tau) j_x(0)\right] e^{i\omega\tau} d\tau, \tag{2.3.7}$$

and $\beta = 1/k_B T$. At high magnetic fields, σ_{xx} is determined by the change in the center of cyclotron motion X, and can be calculated by (2.3.6) and (2.3.7) with j_x equal to

$$j_x = -e \int \Psi^+(\mathbf{r}) \dot{X} \Psi(\mathbf{r}) d^3\mathbf{r}, \tag{2.3.8}$$

where $\Psi(\mathbf{r})$ $(= \sum_\lambda \langle \mathbf{r}|\lambda\rangle\, a_\lambda)$ is the field operator for electrons. Since $\dot{X}$ is determined by the gradient of potential along the y direction (perpendicular to the quantum wire), the term $\dot{X}$ for electrons in the quantum wires is caused by the following two factors. One is the electron-phonon interaction potential $U_{\rm ep}(\mathbf{r}) = \sum_{\mathbf{q}} M(q)\, e^{i\mathbf{q}\cdot\mathbf{r}}\, (a_{\mathbf{q}} + a^+_{-\mathbf{q}})$ which leads to the current component $j_{\rm ep}$:

$$\begin{aligned} j_{\rm ep} &= -e \int \Psi^+(\mathbf{r}) \frac{1}{eB} \frac{\partial U_{\rm ep}(\mathbf{r})}{\partial y} \Psi(\mathbf{r}) d^3\mathbf{r} \\ &= -\frac{1}{B} \sum_{\mathbf{q}} \sum_{\lambda,\lambda'} i q_y\, M(q) \langle \lambda' | e^{i\mathbf{q}\cdot\mathbf{r}} | \lambda \rangle c^+_{\lambda'} c_\lambda (a_{\mathbf{q}} + a^+_{-\mathbf{q}}), \end{aligned} \tag{2.3.9}$$

and the other is the confining potential $V(y)$ which leads to the current component $j_{\rm po}$:

$$\begin{aligned} j_{\rm po} &= -e \int \Psi^+(\mathbf{r}) \frac{1}{eB} \frac{\partial V(y)}{\partial y} \Psi(\mathbf{r}) d^3\mathbf{r} \\ &= -\frac{1}{B} \sum_{\lambda,\lambda'} m\Omega^2 \langle \lambda' | y | \lambda \rangle\, c^+_{\lambda'} c_\lambda. \end{aligned} \tag{2.3.10}$$

Since the cross term between $j_{\rm ep}$ and $j_{\rm po}$ vanishes in the correlation function

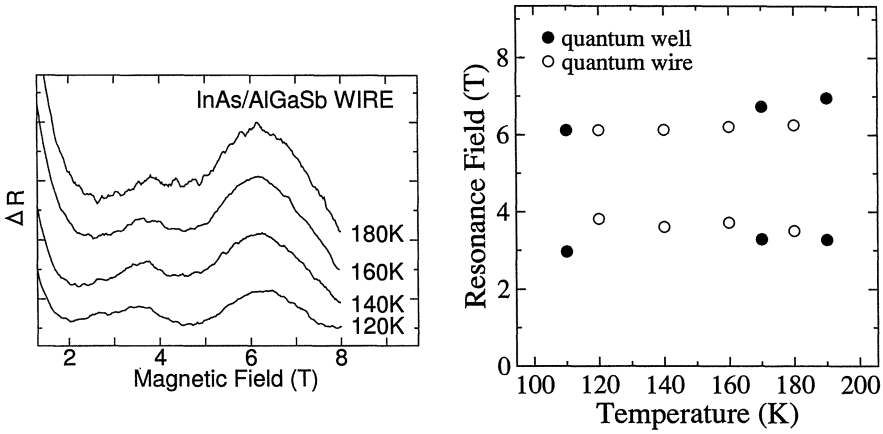


Fig. 2.3.2 Experimental results of the oscillatory part of the magnetoresistance of InAs /AlGaSb quantum wire (left). Temperature dependence of resonance fields for quantum well (solid circles) and wires (open circles) (right)

$\pi(\omega)$, σ_{xx} can also be divided into two parts as follows

$$\sigma_{xx} = \sigma_{\rm ep} + \sigma_{\rm po}. \tag{2.3.11}$$

Figure 2.3.1 shows the calculated results of magnetic field dependence of the second derivative of σ_{xx} $(= \sigma_{\rm ep}+\sigma_{\rm po})$ for GaAs quantum wires with confining potential $\hbar\Omega = 1$, 3, and 5 meV at $T = 100$ K. In the calculation, the low-density case is considered and the distribution function for electrons is approximated by the Boltzmann distribution function, retaining only the term related to electrons in the lowest Landau level $(n=0)$. As can be seen in the figure, for weak confining potential, $\sigma_{\rm ep}$ is dominant and total magnetoconductivity σ_{xx} shows maxima at resonances while for strong confining potential, σ_{xx} shows minima at resonances.

2.3.3 Experiments

Despite the existence of detailed theoretical work [7–11], there is only very little experimental work on the MPR effect in quantum wires. As was mentioned in Sect. 2.3.1, there is no conclusive evidence about the crossover between weak and strong confinement. Experimental magnetoresistance of GaAs/AlGaAs quantum wires reported so far [8,9] is well described by the theory for weak confining potential [7]. By carrying out systematic investigations of the MPR effect in GaAs/AlGaAs quantum wires, however, Ploner et al. demonstrated that the MPR effect can be applied as an alternative transport characterization method when the carrier density is too small to determine the 1D subband spacing from a standard magnetic depopulation technique [9].

Recently, Osako et al. [12] have succeeded in observing the MPR effect in InAs/AlGaSb quantum wires with high electron density compared with quantum well samples, and found that the resonance magnetic field for $P=2$

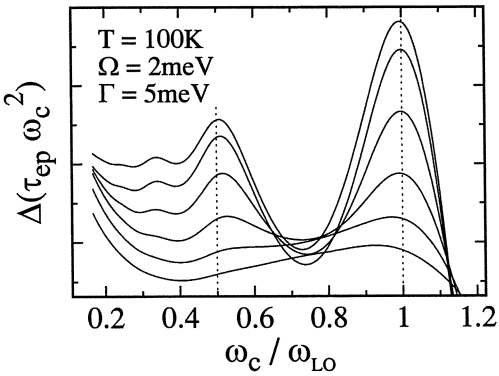


Fig. 2.3.3 The oscillatory part of $\tau_{ep}\omega_c^2$ as a function of ω_c/ω_{LO} for several Fermi energies between $E_F = 10$ meV (top curve) and $E_F = 60$ meV (bottom curve) in 10 meV steps. Vertical dotted lines show $\omega_c/\omega_{LO} = 1/2$ and 1

(at around 3.5 T) in quantum wires is higher than that in quantum wells (Fig. 2.3.2). They attributed this shift to the effects of high electron density and asymmetric density of states for 1D systems, as shown in Fig. 2.3.3, where the relaxation time τ_{ep} due to electron-phonon interaction is evaluated and the oscillatory part of $\tau_{ep}\omega_c^2$, which is proportional to magnetoresistance, is plotted as a function of ω_c/ω_{LO}.

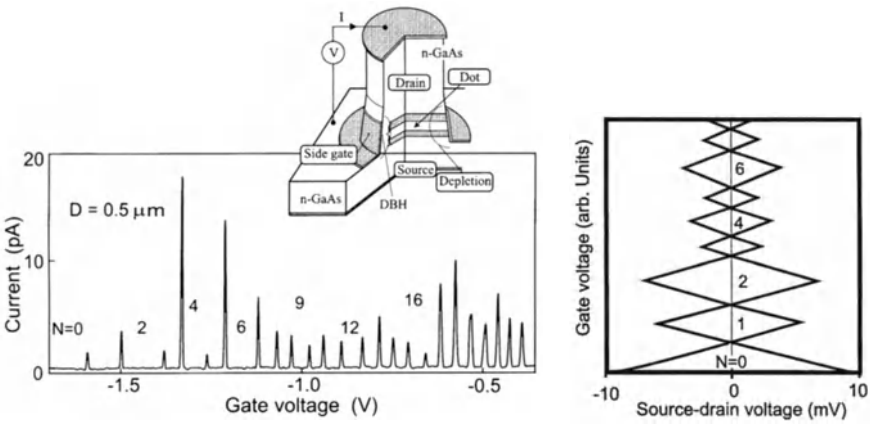
References

1 V.L. Gurevich and Yu.A. Firsov, Sov. Phys. JETP **13** 137 (1961)
2 R.A. Stradling and R.A. Wood, J. Phys. C **1**, 1711 (1968)
3 R.L. Peterson, *Semiconductors and Semimetals*, edited by R.K. Willardson and A.C. Beer (Academic Press, New York, 1975), Vol. 10, p. 221.
4 R.J. Nicholas, Prog. Quantum Electron. **10**, 1 (1985)
5 C. Hamaguchi and N. Mori, Physica B **164** 85 (1990)
6 D.C. Tsui et al., Phys. Rev. Lett. **44**, 341 (1980); G. Kido et al., J. Phys. Soc. Jpn. **51**, 2168 (1982); J.C. Portal et al., J. Phys. C **16**, 573 (1983); M.A. Brummell et al., J. Phys. C **16**, 579 (1983); H. Noguchi et al., Phys. Rev. B **45**, 12148 (1992); D.R. Leadley et al., Phys. Rev. Lett. **73**, 589 (1994)
7 P. Vasilopoulos, P. Warmenbol, F.M. Peeters, and J.T. Devreese, Phys. Rev. B **40**, 1810 (1989)
8 G. Berthold, J. Smoliner, C. Wirner, E. Gornik, G. Böhm, G. Weimann, M. Hauser, C. Hamaguchi, N. Mori, and H. Momose, Semicond. Sci. Technol. **8**, 753 (1993)
9 G. Ploner, J. Smoliner, G. Strasser, M. Hauser, and E. Gornik, preprint
10 N. Mori, H. Momose, and C. Hamaguchi, Phys. Rev. B **45**, 4536 (1992)
11 J-Y. Ryu and R.F. O'Connell, Phys. Rev. B **48**, 9126 (1993)
12 S. Osako, T. Hamano, K. Moriyasu, N. Mori, C. Hamaguchi, S. Sasa, and M. Inoue, unpublished

2.4 Quantum Dots and Artificial Atoms

2.4.1 Quantum Dots Containing a Few Electrons

When a small dot is weakly coupled to reservoirs via small tunnel junctions, addition of an extra electron into the dot raises the electrochemical potential of the dot. The one-by-one change of the number of electrons N in the dot leads to a conductance oscillation as a function of a gate voltage (called the Coulomb oscillation or Coulomb blockade oscillation). The oscillation period is usually constant for a system containing many electrons. However, in a small dot containing just a few electrons, both electron-electron interactions and quantum confinement effects become sufficiently strong to cause a significant modification of the Coulomb oscillation [1–3]. Such a system can be regarded as an artificial atom [4]. There have been several experiments on quantum dots containing only a few electrons. These include transport through a two-terminal asymmetric double-barrier tunneling structure [5–8] and through a gated double-barrier tunneling structure [2,3,9–11], and capacitance measurements of a vertically gated modulation-doped heterostructure [1,12]. In this section we describe transport measurements on a sub-micron gated double-barrier structure.



(a) (b)

Fig. 2.4.1 (**a**) Schematic diagram of a gated quantum dot device and a Coulomb oscillation in the absence of a magnetic fields T [3]. (**b**) dI/dV_{sd} plotted in the V_{sd}-V_g plane [13]

Figure 2.4.1a shows a schematic diagram of the gated double-barrier structure and the Coulomb oscillations (current I versus gate voltage V_g) observed in the absence of a magnetic field [3]. A circular Schottky gate is placed on the side of the mesa close to the dot region and is used to squeeze the dot sideways [11]. The dot is made from InGaAs, which has a narrower band gap than GaAs in the contacts. The inclusion of In in the well reduces the conduction band bottom edge in the dot below the Fermi level in the contacts.

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

This allows us to study linear transport through the vertical quantum dot since electrons are accumulated in the dot at the zero bias condition. The I-V_g characteristic measured at a small bias $V_{\rm sd}$ shows clear current oscillations for $V_g < -1.6$ V with each period corresponding to a change of exactly one electron in the dot. The energy needed to change N just by one is derived from measurements of a Coulomb gap ($V_{\rm sd}$ at which the current starts to flow). Figure 2.4.1b shows the differential conductance $dI/dV_{\rm sd}$, plotted in the $V_{\rm sd}$-V_g plane, and reveals a series of diamond-shaped regions associated with the Coulomb blockade [13]. The vertical $V_{\rm sd}=0$ axis directly corresponds to the data in Fig. 2.4.1a. The irregular $dI/dV_{\rm sd}$ peak spacings at $V_{\rm sd}=0$ lead to similarly irregular sizes for the diamond-shaped Coulomb gaps. Note that the Coulomb gap in the $N=0$ region never closes when we continue to make V_g more negative, implying that the dot is empty. From this observation the absolute values of N can be identified and used to label the spaces between the current peaks.

2.4.2 Atom-like Properties – Shell Filling

The three-dimensional spherically symmetric potential around atoms gives rise to a shell structure consisting of electron orbitals $1s$, $2s$, $2p$, ... The electron affinity has a large maximum when the shells of $1s$, $2s$, $2p$, ... are respectively filled for atomic numbers 2 (He), 10 (Ne), ... These shells are filled sequentially and Hund's rule determines whether a spin-down or a spin-up electron is added. This rule favors the filling of parallel spin states until each orbital is half filled, and at the resulting atomic numbers such as 7 (N in $2p$ orbitals) and 15 (P in $3p$ orbitals) the electron affinity has a maximum. These properties are also expected for a quantum dot which is laterally confined by a harmonic potential. A shell structure for a circular 2D potential was predicted by self-consistent calculations [14]. Here, we discuss the formation and electron filling of a shell structure by comparing a non-interacting model with those including Coulomb and exchange interactions.

If a disk-shaped dot is laterally confined by a 2D harmonic potential, the single-particle state with a radial quantum number n (= 0,1,2,...) and an angular momentum quantum number $\ell = 0, \pm 1, \pm 2, \ldots$) has energy $E_{n\ell}$ given by

$$E_{n\ell} = (2n + |\ell| + 1)\hbar\omega_0, \tag{2.4.1}$$

where $\hbar\omega_0$ is the electrostatic confinement energy. The energy levels consist of sets of degenerate states, which are separated by $\hbar\omega_0$ from each other. These single-particle states is sequentially filled with noninteracting electrons. For the actual filling of electrons in the dot, however, this single-particle picture can be significantly modified by electron-electron interactions. Nevertheless, if $\hbar\omega_0$ is no smaller than the interaction energy, each set of degenerate states are sequentially filled by 2, 6, 12, 20, ... electrons. This sequence in the electron number is just like the case for that in the atomic number in natural atoms, and can be regarded as a "magic number" for a 2D harmonic dot so that this can give rise to a shell structure of electronic states. In addition, the electron-electron interactions determine the filling of spin states in each shell

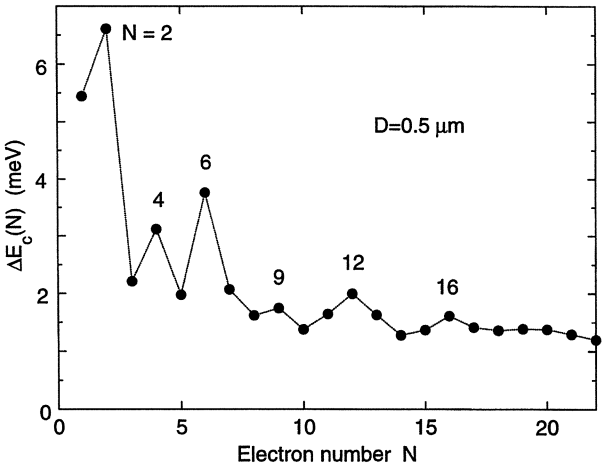


Fig. 2.4.2 The addition energy $\Delta E_c(N)$ vs. N obtained for the dot device whose Coulomb oscillation is shown in Fig. 2.4.1a [3]

in line with Hund's rule.

Here we define an addition energy $\Delta E_c(N) = \mu(N+1) - \mu(N)$, where $\mu(N)$ is the electrochemical potential for an N-electron dot. The potential $\mu(N)$ actually corresponds to the position in energy of the Nth current peak in Fig. 2.4.1a and $\Delta E_c(N)$ corresponds to the spacing between the $N+1$th and Nth current peaks. In the linear transport regime, $\Delta E_c(N)$ can be directly determined from half widths of the Coulomb diamonds in Fig. 2.4.1b, and plotted as a function of N in Fig. 2.4.2. It has large maxima for $N=2$, 6, and 12, and also relatively large maxima for $N=4$, 9, and 16. The first series of $N=2$, 6, and 12 is associated with the complete filling of the first, second, and third shells respectively, while the second series of $N=4$, 9, and 16 is linked to the half-filling of the second, third, and fourth shells, respectively. This N dependence of $\Delta E_c(N)$ is well reproduced by the Hartree–Fock calculation for interacting electrons [15,16].

2.4.3 Atom-like Properties – Spin Effects

The electronic states are expected to be significantly modified by a magnetic field B applied parallel to the tunnel current. Figure 2.4.3a shows the magnetic-field dependence of the position of the current oscillations in Fig. 2.4.1a. The positions of the first three peaks depend monotonicly on B, whereas the other peaks oscillate back and forth a number of times. The number of "wiggles" increases with N. A close inspection of the figure reveals that the current peaks generally shift in pairs with B. The spacing between the peaks when N is odd is nearly constant, whereas the peak spacing varies strongly with B when N is even. Most of these features can be explained within a single-particle framework.

The single-particle states in magnetic fields are known as the Fock-Darwin

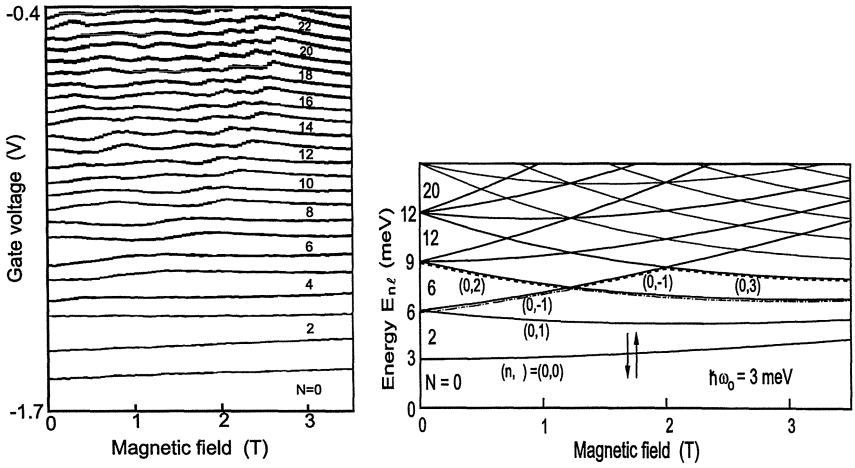


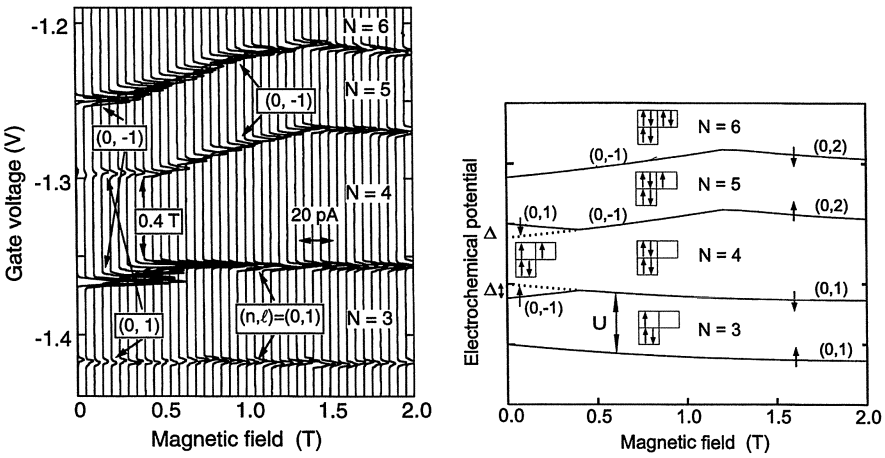
Fig. 2.4.3 (**a**) The gate voltage positions of the current peaks as a function of a magnetic field [3]. (**b**) Calculated single-particle energy as a function of a magnetic field [3]

states given by

$$E_{n\ell} = (2n+|\ell|+1)\hbar\sqrt{\frac{1}{4}\omega_c^2+\omega_0^2} - \frac{1}{2}\ell\hbar\omega_c, \tag{2.4.2}$$

where $\hbar\omega_c$ is the cyclotron energy. Figure 2.4.3b shows $E_{n\ell}$ versus B calculated for $\hbar\omega_0$=3 meV. The electron spin is neglected and each state is twofold degenerate. The degeneracies at $B=0$ are lifted for $B>0$ in such a way that a single-particle state with a positive or negative angular momentum shifts to the lower or higher energy side, respectively, in proportion to B for small B. This dependence gives rise to $\Delta E_c(N)$ for even N that is strongly dependent on B. On the other hand, $\Delta E_c(N)$ for odd N is determined only by the effect of the electron-electron interaction, which is responsible for lifting the spin degeneracy. This should lead to a pairing of the current peaks. In Fig. 2.4.3b the energy curve for the fifth and sixth electrons (dot-dashed line) predicts that these electrons undergo transitions in their quantum numbers; (n,ℓ) goes from $(0,-1)$ to $(0,2)$ at 1.3 T. The energy curve for the seventh and eighth electrons (dashed line) predicts that these electrons undergo transitions in (n,ℓ) from $(0,2)$ to $(0,-1)$ at 1.3 T and then to $(0,3)$ at 2 T. These transitions are also seen in Fig. 2.4.3a.

Figure 2.4.4a shows the B field dependence of the third, fourth, fifth, and sixth current peaks, i.e., peaks which are related to the filling of electrons into the second shell. The pairing of the third and fourth peaks, and the fifth and sixth peaks above 0.4 T, is clear. From the evolution of the respective pairs with magnetic field, we identify the quantum numbers of $(n,\ell)=(0,1)$ with antiparallel spins for the lower-lying pair and $(0,-1)$ with antiparallel spins for the higher pair. However, this pairing is rearranged for $B<0.4$ T. In this field range the third and sixth peaks shift smoothly with B, whereas



(a) (b)

Fig. 2.4.4 (a) Evolution of the third, fourth, fifth, and sixth current peaks in magnetic fields [3]. The original data consist of I versus V_g traces in different magnetic fields, which are offset and rotated by 90 °. (b) Calculation of $\mu(N)$ versus B for $N=3$ to 6

the fourth peak switches pairing partner from the third to the sixth peak, and similarly the fifth peak switches partner from the sixth to the third peak. This means that as B is increased from 0 T exceeding 0.4 T, the fourth electron undergoes an angular momentum transition from $\ell=-1$ to 1, while the fifth electron undergoes the opposite transition from $\ell=1$ to -1.

This behavior can be understood in terms of Hund's rule. This is modeled in the calculation of $\mu(N)$ versus B shown in Fig. 2.4.4b. In this figure, $\mu(N)$ reflects the Nth single-particle energy state as identified by the quantum numbers (n,ℓ) and the diagrams illustrate the spin configurations. In the constant interaction model, $\mu(N)$ can be written as a constant interaction energy U added to $E_{n\ell}$. To include Hund's rule in the calculation we introduce an energy Δ, which represents the reduction due to the exchange interaction between electrons with parallel spins. Specifically, for $N=4$, the ground state energy is reduced by exchange effects if the outer two electrons have parallel spins with different angular momenta in comparison with the case that they have anti-parallel spins with the same angular momentum. Thus, $\mu(4)$ is reduced by an amount Δ and there is a corresponding increase in $\mu(5)$ by Δ. This exchange effect is canceled in the presence of a magnetic field when the $(0,\pm1)$ states, which are degenerate at $B=0$ T, are split by an energy exceeding Δ. This is a simple way to include exchange effects in a constant interaction model. For small N we find a remarkable agreement between what is seen in Fig. 2.4.4a and that predicted in Fig. 2.4.4b with $U=3$ meV and $\Delta=0.7$ meV. In this model, $\Delta E_c(N)$ for $N=4$ at $B=0$ is expected to be larger by 2Δ than that for $N=3$ and 5, and this is indeed observed in Fig. 2.4.2. This simple Hund's rule model is a first correction to the constant interaction model. Recently, Eto [17] obtained in an exact

diagonalization approach $\mu(N)$ versus B in close agreement with the data given in Fig. 2.4.4(a).

References

1 R.C. Ashoori, H.L. Stoermer, J.S. Weiner, L.N. Pfeiffer, K.W. Baldwin, and K.W. West, Phys. Rev. Lett. **71**, 613 (1993)
2 S. Tarucha, D.G. Austing, and T. Honda, Superlattices & Microstructures **18**, 121 (1995)
3 S. Tarucha, D.G. Austing, T. Honda, R.J. van der Hage, and L.P. Kouwenhoven, Phys. Rev. Lett. **77**, 3613 (1996)
4 M.A. Kastner, Phys. Today **46**, 24 (1993)
5 Bo Su, V.J. Goldman, and J.E. Cunningham, Science **255**, 313 (1992)
6 M. Tewordt, L. Martin-Moreno, J.T. Nicholls, M. Pepper, M.J. Kelly, V.J. Law, D.A. Ritchie, J.E.F. Frost, and G.A.C. Jones, Phys. Rev. B **45**, 14407 (1992)
7 S. Tarucha, T. Honda, T. Saku, and Y. Tokura, Surf. Sci. **305**, 547 (1994)
8 T. Schmidt, M. Tewordt, R.H. Blick, R.J. Haug, D. Pfannkuche, K. von Klitzing, A. Forester, and H. Lueth, Phys. Rev. B **51**, 5570 (1995)
9 M.W. Dellow, P.H. Beton, M. Henini, P.C. Main, and L. Eaves, Electron. Lett. **27**, 134 (1991)
10 P. Gueret, N. Blanc, R. Germann, and H. Rothuizen, Phys. Rev. Lett. **68**, 1896 (1992)
11 D.G. Austing, T. Honda, and S. Tarucha, Semicond. Sci. Technol. **11**, 388 (1996)
12 R.C. Ashoori, Nature (London) **379**, 413 (1996)
13 L.P. Kouwenhoven, R.J. van der Hage, S. Tarucha, D.G. Austing, and T. Honda, unpublished
14 M. Macucci, K. Hess, and G.J. Iafrate, J. Appl. Phys. **77**, 3267 (1995)
15 H. Tamura, private communications
16 Y. Tanaka and H. Akera, J. Phys. Soc. Jpn. **66**, 15 (1997)
17 M. Eto, Jpn. J. Appl. Phys. (in press)

2.5 Antidot Lattices – Classical and Quantum Chaos

2.5.1 Antidot Lattices

The two-dimensional (2D) system modulated by a periodic strong repulsive potential is called an antidot lattice. Various interesting phenomena have been observed in antidot lattices in uniform perpendicular magnetic fields. They are the quenching of the Hall effect [1,2], Altshuler–Aronov–Spivak oscillation near vanishing field [3–6], so-called commensurability peaks in magnetoresistance [1,7–16], and fine oscillations around them [17,18].

Figure 2.5.1 shows a schematic illustration of the potential distribution of a square antidot lattice. It is characterized by the period a and the diameter d of each antidot. We have typically $d \gtrsim 1000$ Å, which is larger than the Fermi wavelength of the 2D electron system, $\lambda_F \sim 500$ Å for a typical electron concentration less than 5×10^{11} cm^{-2} in a GaAs/AlGaAs single heterostructure. Correspondingly, the period of the antidot is usually $a \geq 2000$ Å. This means that the antidot lattice is in a semiclassical regime.

The model antidot potential used often for a square antidot is given by

$$U(\mathbf{r}) = U_0 \left| \cos\left(\frac{\pi x}{a}\right) \cos\left(\frac{\pi y}{a}\right) \right|^{2\beta}, \tag{2.5.1}$$

where U_0 is a potential maximum and the parameter β characterizes its steepness. The antidot diameter can be defined as $d = |\mathbf{r}|$ with $U(\mathbf{r}) = E_F$, where E_F is the Fermi energy and $\mathbf{r}$ is chosen in the direction of a nearest neighbor.

A self-consistent calculation in quantum wires fabricated at GaAs/AlGaAs heterostructures [19,20] suggests that the potential is nearly parabolic for a wire with small width and consists of a flat central region and a parabolic increase near the edge for a wider wire. The width of the region where the potential increases from the bottom to the Fermi energy is of the same order as the Fermi wavelength for typical electron concentrations. This leads to roughly $\beta \sim 1$ for $d/a \sim 0.5$, $\beta \sim 2$ for $d/a \sim 0.4$, and $\beta \sim 4$ for $d/a \sim 0.3$ in the case $d \sim 1000$ Å.

2.5.2 Commensurability Peaks

Figure 2.5.2 shows an example of observed diagonal and Hall resistivity in a square antidot lattice [1]. Two prominent peaks appear in the diagonal resistivity ρ_{xx} in weak magnetic fields and correspondingly some steplike structures are present also for the off-diagonal Hall resistivity ρ_{xy}. Cyclotron orbits can be classified into those affected and not affected by antidots [1]. At certain magnetic fields electrons can move on the commensurate classical orbit encircling a specific number of antidots as shown in Fig. 2.5.3 (pinned orbits). The magnetoresistance was expected to increase at this magnetic field.

On the other hand, a numerical simulation based on the classical electron motion showed that the change in the volume of the pinned orbits in the phase space is not enough to cause the commensurability oscillation and suggested that chaotic orbits which have a character close to the periodic orbits give larger contribution [21]. The importance of the "runaway" orbit, which skips

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

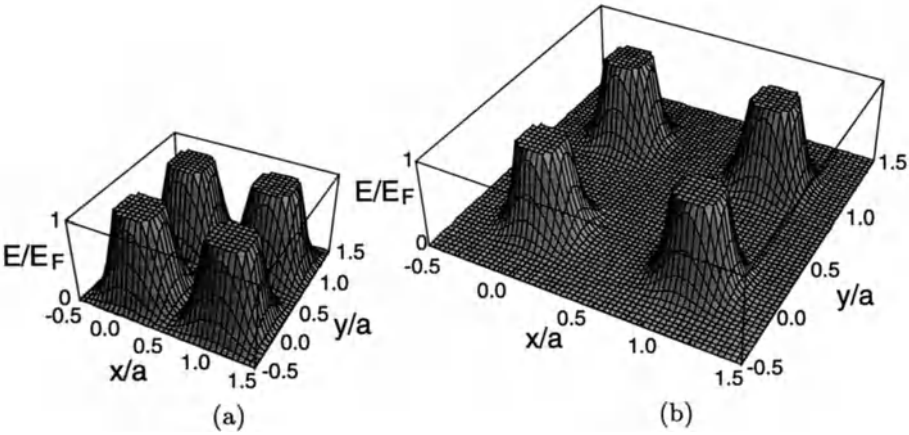


Fig. 2.5.1 A model antidot potential in a square lattice. (a) $d/a = 0.5$ and $\beta = 1$. (b) $d/a = 0.3$ and $\beta = 4$. The potential is cut off at the Fermi energy

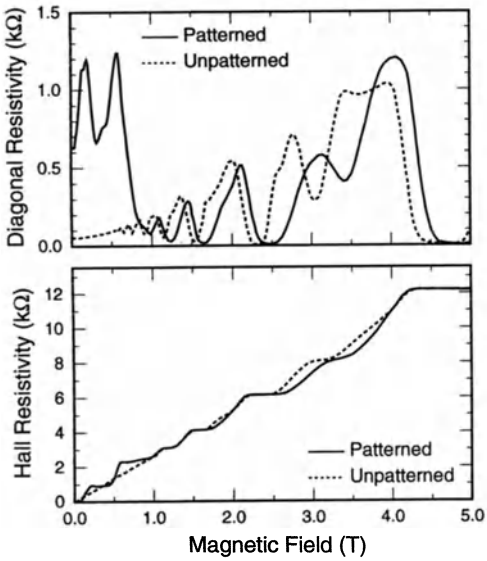


Fig. 2.5.2 An example of ρ_{xx} and ρ_{xy} observed in patterned (solid lines) and unpatterned (dotted lines) 2D systems [1]. Two prominent peaks and step structures appear in the diagonal and Hall resistivity, respectively, in antidot lattices

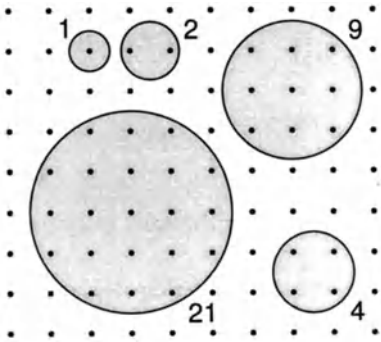


Fig. 2.5.3 Some examples of pinned orbits. It was suggested that the resistivity exhibits a peak at the magnetic fields where the classical cyclotron orbit encircles 1, 2, 4, 9, 21, ... antidots

regularly from an antidot to its neighboring antidot in the same direction, and to its neighboring orbits in the phase space, was also proposed [22]. Experiments on rectangular and/or disordered antidots [23–27] and numerical simulations [27–29] provided evidence showing the importance of such orbits.

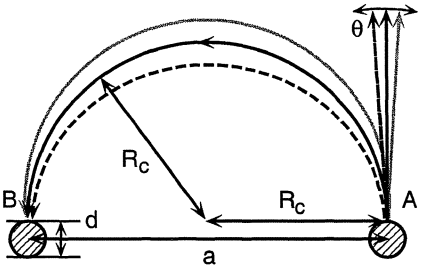


Fig. 2.5.4 Schematic illustration of magnetic focusing leading to the fundamental commensurability peak

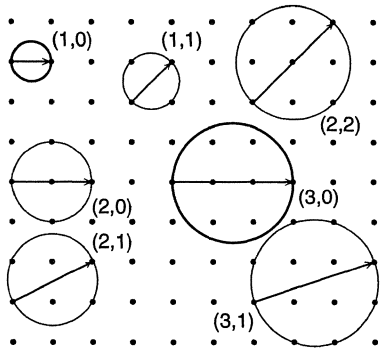


Fig. 2.5.5 Some examples of diffusive cyclotron orbits contributing to the migration of the guiding center

It is easy to understand the origin of the commensurability peaks in the limit of small aspect ratio $d/a \ll 1$. In this case, the electron loses its previous memory of the direction of the velocity when it collides with an antidot, and therefore successive scattering with antidots can be approximately regarded as independent of each other [30]. This means that antidots are nothing but independent scatterers.

In magnetic fields, transport is possible through the migration of the center of the cyclotron motion and therefore the conductivity vanishes in the absence of scattering. When $2R_c < a$, the scattering of an electron from an antidot cannot give rise to diffusion or conduction because the electron is trapped by the antidot.

Scattering from antidots starts to contribute to the conductivity when $2R_c > a - d$. The migration of the center of the cyclotron orbit occurs most frequently due to successive scattering from nearest-neighbor antidots at the magnetic field corresponding to $2R_c = a$. At this magnetic field the measure of such orbits becomes maximum in the phase space due to a kind of magnetic focusing effect. In fact, a slight change θ in the direction from the direction normal to the line connecting neighboring antidot leads to a change only of the order of $2R_c\theta^2$ in position when the electron collides with a neighboring antidot as is shown in Fig. 2.5.4. This leads to an increase of the phase-space volume of the orbits contributing to the increase of the diffusion coefficient at $2R_c \approx a$. This magnetic focusing is different from mechanisms suggested in other works [25,27,31].

As shown in Fig. 2.5.5, the orbit corresponding to $2R_c = a$ can be denoted as $(n_x, n_y) = (\pm 1, 0)$ or $(0, \pm 1)$, where the line segment connecting the point (n_x, n_y) and the origin constitutes the diameter of the circle. With a further decrease in the magnetic field, successive scattering with next nearest-neighbor antidots becomes possible and the conductivity has a peak around $2R_c = \sqrt{2}a$ corresponding to $(\pm 1, \pm 1)$. This contribution becomes less prominent, however, because the orbit passes through the position of a nearest-neighbor antidot. The next peak arises from $(\pm 2, 0)$ or $(0, \pm 2)$ and $(\pm 2, \pm 1)$ or $(\pm 1, \pm 2)$, which lie close to each other. The latter contribution should be

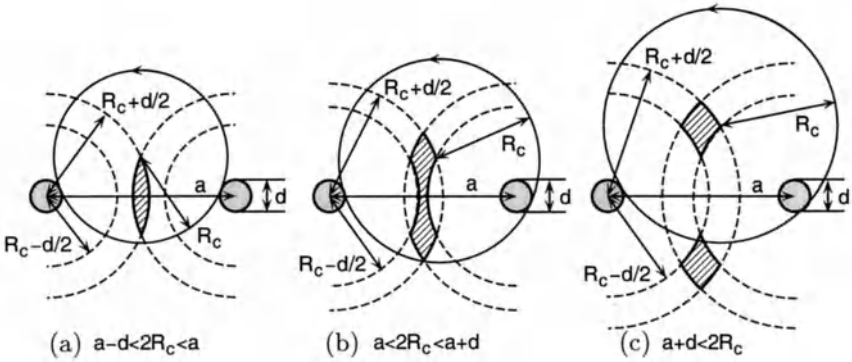


Fig. 2.5.6 The region (shaded) of the center of the cyclotron orbits passing through two nearest neighbor antidots

larger because its measure is twice as large as that of the former and therefore the peak occurs roughly around $2R_c \sim \sqrt{5}a$. This peak should be weaker, however, as these orbits also pass through the position of other antidots. The next prominent peak is expected to be given by $(\pm 2, \pm 2)$, $(\pm 3, 0)$, $(0, \pm 3)$, $(\pm 3, \pm 1)$, and $(\pm 1, \pm 3)$. The peak is expected to be at $2R_c \approx 3a$, because the orbits $(\pm 3, 0)$ and $(0, \pm 3)$ are not disturbed by other antidots. This explains most of the commensurability peaks observed experimentally.

When the diameter d is no longer negligible in comparison with the period a, the commensurability peak in the diagonal conductivity σ_{xx} deviates from the condition $2R_c = (n_x^2 + n_y^2)^{1/2} a$. In the following, we confine ourselves to the fundamental peak given by $(\pm 1, 0)$ and $(0, \pm 1)$. A cyclotron orbit starting at an antidot and colliding with a neighboring antidot has a center in the region schematically illustrated in Fig. 2.5.6, to which the phase-space volume of such orbits is proportional. Figure 2.5.7 shows the area as a function of the magnetic field. The area has a maximum at $2R_c = a + fd$ with $f \lesssim 0.5$ being a constant weakly dependent on d/a. This means that the commensurability peak is shifted to the weak-field side roughly in proportion to d/a.

For a nonnegligible d/a, the correlation among successive scattering from antidots becomes important. In fact, it manifests itself most directly in the Hall conductivity. As illustrated in Fig. 2.5.8, it tends to enhance the left circular motion of electrons in the weak-field region $2R_c \gtrsim a$ and reduce it in the high-field region $2R_c \lesssim a$. This leads to an enhancement of $|\sigma_{xy}|$ for $2R_c \lesssim a$ and a reduction for $2R_c \lesssim a$, giving rise to a steplike structure around $2R_c \approx a$.

With the increase of d/a, the measure of the diffusive orbits increases as shown in Fig. 2.5.7, and the diagonal conductivity becomes larger and comparable to σ_{xy} (numerical calculations show that $\sigma_{xx} \sim |\sigma_{xy}|$ around $d/a \sim 0.5$ [30]). Therefore, the commensurability peaks in the resistivity become different from those in the diagonal conductivity. Let $\Delta\sigma_{xx}$ and $\Delta\sigma_{xy}$ be a small change around the field corresponding to the commensurability peak

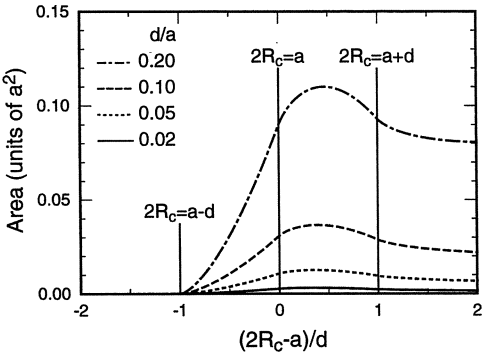


Fig. 2.5.7 The area of the region where cyclotron orbits passing through two nearest neighbor antidots. The diagonal conductivity is roughly proportional to the area at $2R_c \sim a$

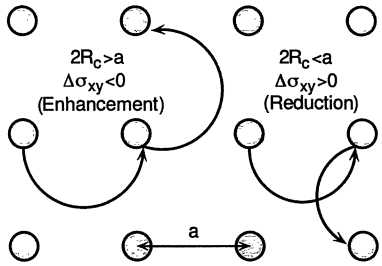


Fig. 2.5.8 A schematic illustration of scattering from two adjacent antidots. This leads to an enhancement of $|\sigma_{xy}|$ for $2R_c \lesssim a$ and a reduction for $2R_c \lesssim a$

and $\bar{\sigma}_{xx}$ and $\bar{\sigma}_{xy}$ be the average ($\Delta\sigma_{xx} \ll \bar{\sigma}_{xx}$, $|\Delta\sigma_{xy}| \ll |\bar{\sigma}_{xy}|$). Then, we have

$$\Delta\rho_{xx} = \frac{(\bar{\sigma}_{xy}^2 - \bar{\sigma}_{xx}^2)\Delta\sigma_{xx} - 2\bar{\sigma}_{xx}\bar{\sigma}_{xy}\Delta\sigma_{xy}}{(\bar{\sigma}_{xx}^2 + \bar{\sigma}_{xy}^2)^2}. \tag{2.5.2}$$

This means that $\Delta\sigma_{xy}$ determines the structure of ρ_{xx} when $\bar{\sigma}_{xx} \sim |\bar{\sigma}_{xy}|$, i.e., the reduction of $|\sigma_{xy}|$ in the high field side gives the peak in ρ_{xx} for large d/a.

Figure 2.5.9 shows an example of the results of recent numerical simulations in classical mechanics [30]. It shows that the peak in σ_{xx} is broadened considerably with the increase of d/a and at the same time shifted to the lower magnetic field side. A step structure corresponding to the first derivative of σ_{xx} with the magnetic field appears in σ_{xy}. It is interesting, however, that the commensurability peak in the resistivity ρ_{xx} remains around $2R_c = a$ independent of d/a. For $d/a \geq 0.5$, the structures in σ_{xx} and σ_{xy} disappear almost completely and ρ_{xx} does not show a clear commensurability peak.

A large aspect ratio usually corresponds to a small antidot period a and to a narrow spacing between the nearest-neighbor antidot comparable to the Fermi wavelength λ_F. Therefore quantum effects can be important for antidot lattices with large d/a. In fact, a quantum mechanical calculation shown in Fig. 2.5.10 in a self-consistent Born approximation (SCBA) gave the result for $d/a = 0.5$ that the diagonal conductivity σ_{xx} exhibits essentially no structure and the off-diagonal Hall conductivity has a small dip at $2R_c \sim a$, leading to the fundamental commensurability peak at $2R_c \sim a$ [32]. Some experiments have shown that σ_{xx} even has a dip at $2R_c \sim a$ for antidot lattices with larger d/a [27,33]. The structure appearing in the off-diagonal Hall resistivity at the fundamental commensurability peak was analyzed experimentally and theoretically [34].

2.5.3 Aharonov–Bohm Type Oscillation

A fine oscillation was observed superimposed on commensurability peaks of

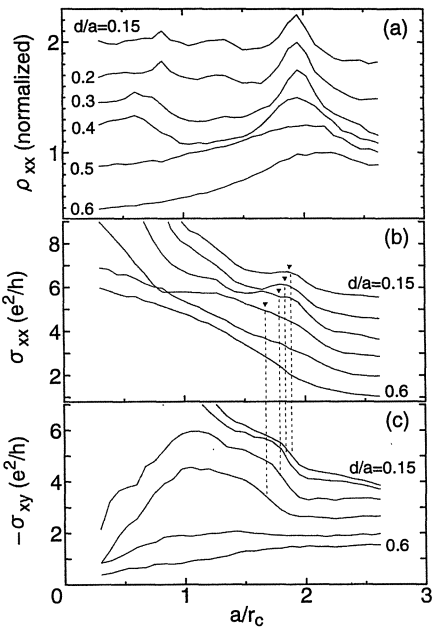


Fig. 2.5.9 Some examples of (a) ρ_{xx}, (b) σ_{xx}, and (c) σ_{xy} calculated in classical mechanics. Each curve of ρ_{xx} is normalized by its maximum value and shifted successively. The results for σ_{xx} are shifted successively by e^2/h. The peak positions of σ_{xx} are indicated by triangles and vertical dotted lines. The mean free path is $\Lambda = 3.3a$

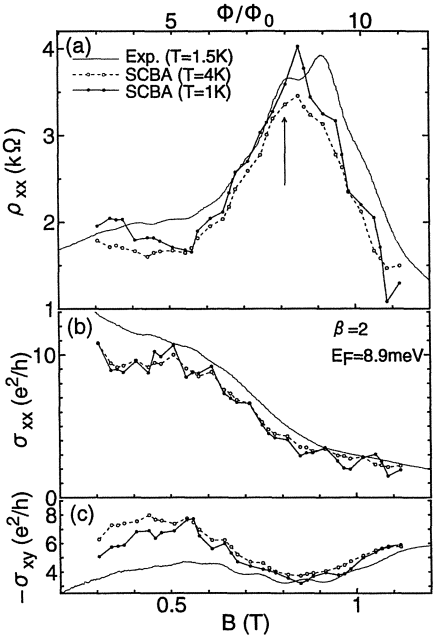


Fig. 2.5.10 The resistivity ρ_{xx}, σ_{xx}, and σ_{xy} calculated in SCBA as a function of the magnetic field for $\beta = 2$. The experimental data [17] are also shown for comparison

the magnetoresistance [5,17,18]. Figures 2.5.11 and 2.5.12 give examples of such Aharonov-Bohm (AB) type oscillation reported in [17,18]. The period is roughly given by $\Delta B \sim \Phi_0/a^2$ with $\Phi_0 = ch/e$ being the magnetic flux quantum. Numerical calculations have been performed on this AB type oscillation [32,35,36]. In the following some of the results [32,37] obtained in SCBA are reviewed briefly.

Figure 2.5.10 shows an example of the calculated σ_{xx}, σ_{xy}, and ρ_{xx} for $\hbar/\tau = 0.18$ meV, $\beta = 2$, and $U_0 = 35.4$ meV, together with the experimental data given in [17]. The strength of scattering is characterized by the relaxation time τ in the absence of a magnetic field.

The calculated diagonal and Hall conductivity at $T = 4$ K oscillate with an exact period of Φ_0 over the entire range of magnetic fields considered. This oscillation originates from Hofstadter's butterfly band structure [38] as has been suggested theoretically [39]. At a lower temperature ($T = 1$ K), another kind of oscillation appears in addition to this butterfly oscillation. The difference between this new oscillation and the butterfly oscillation becomes much clearer for the more realistic parameter $\beta = 1$.

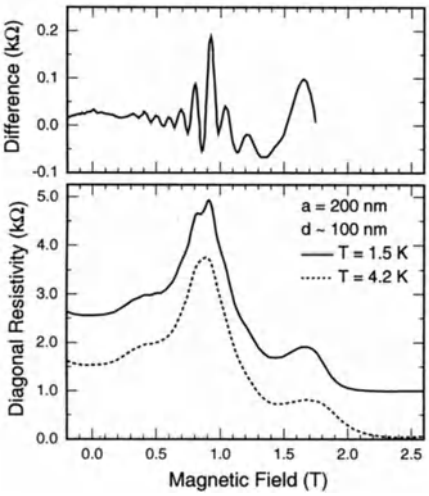


Fig. 2.5.11 An example of observed quantum oscillations superimposed on the fundamental commensurability peak [17]. The oscillation becomes clear in the difference of the resistivity at $T = 4.2$ K and that at $T = 1.5$ K (shifted vertically by 1 kΩ)

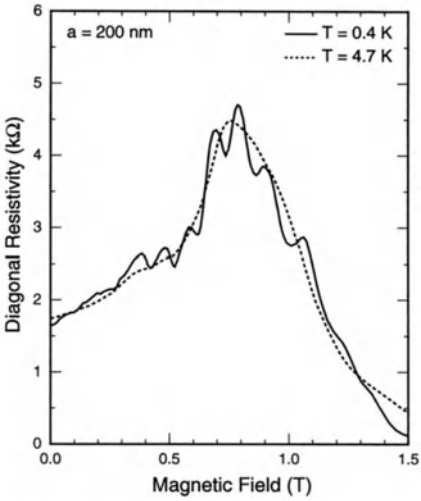


Fig. 2.5.12 An example of observed quantum oscillation superimposed on the fundamental commensurability peak [18]

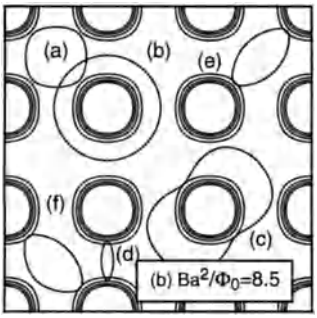


Fig. 2.5.13 Some examples of periodic orbits having a short trajectory in a square antidot in magnetic field $Ba^2/\Phi_0 = 8.5$ roughly corresponding to the main commensurability peak. The parameters are $a/\lambda_F = 4.31$, $U_0/E_F = 8.4$, $d/a = 0.6$, and $\beta = 2$

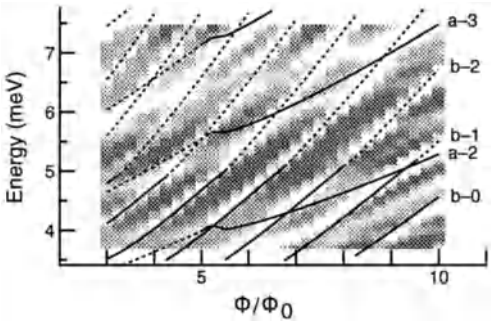


Fig. 2.5.14 The density of states calculated in SCBA for $\hbar/\tau = 0.18$ meV and $\beta = 1$. Its values are large in the light region and small in the dark region. The energy levels for the orbits (a) and (b) calculated semiclassically with the use of the trace formula for $m = 0$ are given by lines (solid lines for stable orbits and dotted lines for unstable orbits)

The density of states in classically chaotic systems is given semiclassically in the so-called periodic orbit theory [40–42]. According to the trace formula, the density of states is given by the sum of the classical contribution and the quantum correction. The latter is given by the sum over the contribution of

each periodic orbit. The contribution of a particular periodic orbit can be written as the summation of δ functions for a stable orbit,

$$D_{\rm osc}^{\rm po}(E) = \frac{T}{\hbar}\sum_{m=0}^{\infty}\left[\sum_{n=-\infty}^{\infty}\delta\Big(\frac{S_{\rm c}}{\hbar}-(m+\frac{1}{2})v-2n\pi\Big)-\frac{1}{2\pi}\right], \tag{2.5.3}$$

and the summation of Lorentzian functions for an unstable orbit,

$$D_{\rm osc}^{\rm po}(E) = \frac{T}{\hbar}\sum_{m=0}^{\infty}(\pm 1)^m\left[\sum_{n=-\infty}^{\infty}\frac{1}{\pi}\frac{(2m+1)u/2}{[S_{\rm c}/\hbar-2\pi(\mu/4+n)]^2+[(2m+1)u/2]^2}-\frac{1}{2\pi}\right], \tag{2.5.4}$$

where $S_{\rm c}$ is the classical action, T is the period, and v is the stability angle. For unstable orbits, u is the Lyapunov exponent, μ is the Maslov index, and the sign becomes $\pm$ according as the neighborhood of a periodic orbit is hyperbolic or inverse hyperbolic. The integer m is considered as a quantum number for a motion transverse to the periodic orbit [43]. The quantization conditions are

$$S_{\rm c} = \begin{cases} 2\pi\hbar\Big[n+\dfrac{1}{2\pi}\Big(m+\dfrac{1}{2}\Big)v\Big] & \text{for stable orbit,} \\ 2\pi\hbar\Big(n+\dfrac{\mu}{4}\Big) & \text{for unstable orbit,} \end{cases} \tag{2.5.5}$$

with integer n and nonnegative integer m.

Figure 2.5.13 shows some examples of periodic orbits having a short trajectory [44]. The orbit (a) is localized in the potential-well region surrounded by four neighboring antidots. It turns continuously into a normal cyclotron orbit in sufficiently high magnetic fields. The orbit (b) circling around an antidot is considered to be most important in giving rise to the quantum oscillation.

Figure 2.5.14 shows an example of the calculated density of states together with quantized energy levels associated with the orbits (a) and (b) for $m=0$. The density of states is correlated well with the semiclassical energy levels, although many features remain to be explained.

A semiclassical expression can be obtained for the conductivity tensor in a manner similar to that of the derivation of the density of states [32,45,46]. Figure 2.5.15 gives an example of the comparison between the SCBA results and those obtained in the semiclassical analysis in which the contribution of orbits (a) and (b) is included. Although the semiclassical result corresponds reasonably well to the quantum result for the density of states, the correspondence becomes worse for the conductivity tensor. In fact, the semiclassical diagonal conductivity exhibits an oscillation with a phase completely opposite to that of the SCBA result in the region of weak magnetic fields. Further, the Hall conductivity cannot be reproduced by a semiclassical theory at all.

As discussed in Sect. 2.5.2, those orbits which jump around due to successive scattering from neighboring antidots have a dominant contribution to the transport in the magnetic-field region near the fundamental commensurability peak. These orbits are perturbed by the presence of quantized periodic orbits, which is likely to lead to a quantum oscillation. Therefore, the semiclassical

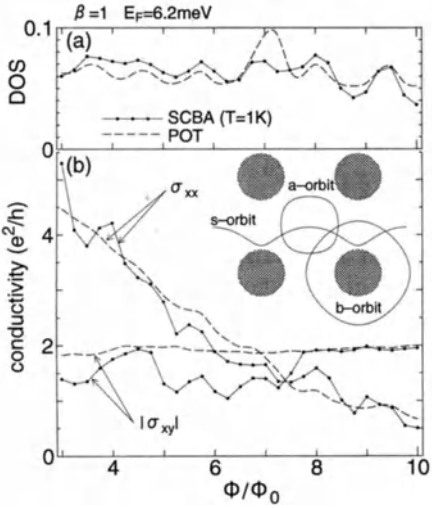


Fig. 2.5.15 Comparison of the density of states and the diagonal and Hall conductivity calculated in SCBA and using the trace formula in which the contributions of the orbits (a) and (b) orbit are included. The inset shows the orbits (a) and (b) at the field corresponding to the main commensurability peak and also a runaway orbit denoted as s-orbit

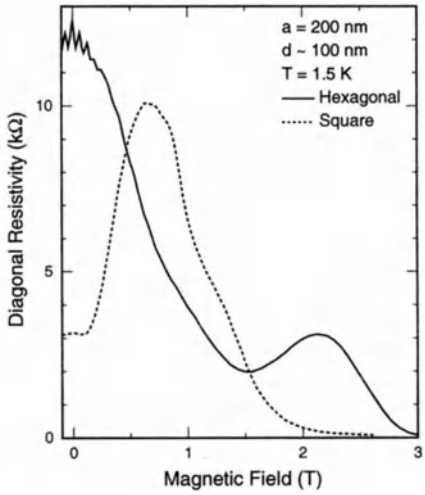


Fig. 2.5.16 Comparison of the resistivity in a hexagonal (solid line) and square (dotted line) antidot observed experimentally. A prominent AAS oscillation with period determined by $\Phi_0/2$ appears in the hexagonal lattice

expression for the conductivity is not as good as that for the density of states, and other quantum effects such as level repulsion effects between periodic and runaway orbits become dominant in the conductivity.

2.5.4 Altshuler–Aronov–Spivak Oscillation

The conductance of a conductor in the form of a hollow cylinder oscillates as a function of the magnetic flux passing through its cross section with period given by half of the flux quantum $\Phi_0 = ch/e$. This oscillation is a result of interference effects and is called the Altshuler–Aronov–Spivak (AAS) oscillation [47]. This AAS oscillation was experimentally observed in cylindrical metal films [48] and networks of thin metallic wires [49] in the diffusive regime where the mean free path is much smaller than the sample size or the wire width.

AAS oscillations have been observed also in antidot lattices [3–6]. The essential features of the experiments can be summarized as follows: (1) The oscillations are observed clearly in triangular lattices but are almost invisible in square lattices [5,6]. (2) The negative magnetoresistance is observed together with the AAS oscillation. In the following, a numerical study [50,51] is reviewed briefly.

Real antidot lattices have inherent randomness arising from fabrication processes in addition to impurities whose strength is characterized by the mean free path Λ. This disorder is introduced through fluctuations in antidot

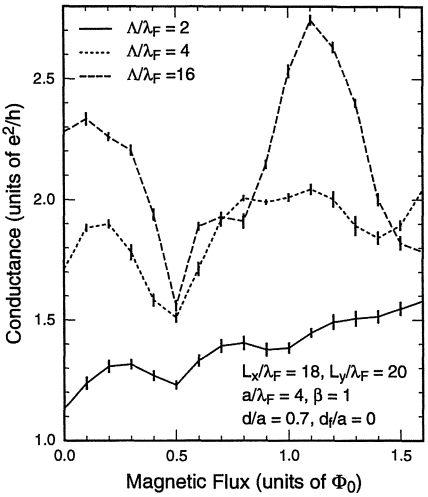


Fig. 2.5.17 Conductance of a triangular antidot lattice as a function of magnetic flux passing through a unit cell for different values of the mean free path

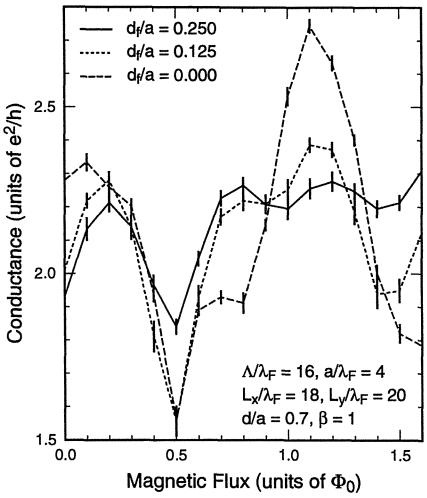


Fig. 2.5.18 Conductance of a triangular antidot lattice in the presence of fluctuations in antidot diameter

diameters distributed uniformly with width d_f around the mean value d. The conductance is calculated by a Green's function technique [52] for systems with length L_x and width L_y containing a finite number of antidots (~ 25).

Figure 2.5.17 gives calculated conductance of a triangular antidot lattice in the absence of disorder in the antidot potential. In the ballistic regime $\Lambda/\lambda_F = 16$, where Λ is larger than the antidot period, the conductance exhibits an irregular oscillation. This is a consequence of the fact that an AB type oscillation with period Φ_0 due to interference effects has a dominant contribution. In fact, both phase and form of the oscillation vary sensitively depending on parameters like d/a, a/λ_F, and β. With the decrease in the mean free path, the amplitude of the AB oscillation becomes smaller and that of an AAS oscillation characterized by period $\Phi_0/2$ becomes more and more dominant. In particular, a beautiful AAS oscillation manifests itself in the diffusive regime where $\Lambda/\lambda_F = 2$ or $\Lambda/a = 1/2$.

Similarly, the contribution of the AB type oscillation diminishes with the increase of disorder in the antidot potential, as shown in Fig. 2.5.18. The oscillation with period $\Phi_0/2$ becomes dominant for disorder as large as $d_f/a = 0.25$. Figure 2.5.18 shows that the AAS oscillation is accompanied by a weak negative magnetoresistance or a slight increase in the conductance as a function of a magnetic field. This is in agreement with one of the features of the experiments discussed above. The amplitude of the AAS oscillations saturates for $d_f/a > 0.25$.

In the square lattice the AAS oscillation has a smaller amplitude than that in the triangular lattice, again in excellent agreement with experiments. Note that the conductance of the triangular lattice is much smaller than that of the

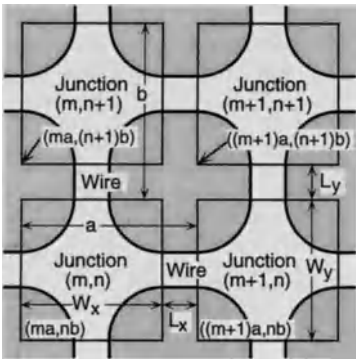


Fig. 2.5.19 A two-dimensional array of quantum wire junctions

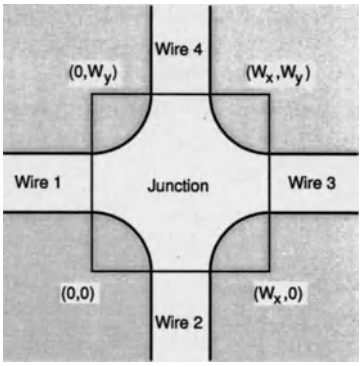


Fig. 2.5.20 A quantum wire junction for calculation of an S matrix

square lattice. Electrons perform more complicated motion and stay longer in the triangular lattice than in the square lattice [5,6]. This is presumably the main origin of the enhancement of the oscillation in the triangular lattice.

2.5.5 Scattering Matrix Formalism

Difficulties of quantum mechanical calculations of energy levels and transport coefficients in antidot lattices lie in the fact that we have to consider a magnetic unit cell which can become very large depending on the magnetic flux passing through a unit cell. In this section a new method of full quantum calculations of energy bands and transport quantities [53,54] is reviewed.

We shall replace an antidot lattice by a 2D array of quantum-wire junctions with the width W_x in the x direction and W_y in the y direction, as illustrated in Fig. 2.5.19. Each junction is connected to neighboring junctions through a quantum wire with length L_x and L_y in the x and y directions, respectively. The lattice period in the x direction is a and that in the y direction is b. For the calculation of an S matrix, we consider the model shown in Fig. 2.5.20. We divide this system into five regions, i.e., four wires and a junction.

The vector potential is chosen as $(-By, 0)$. The modes in the wires 1–4 are given by

$$\begin{aligned}
\psi^{\pm}_{1\nu}(x,y) &= (v^x_\nu)^{-1/2}\exp[i(\pm\kappa^x_\nu + W_y/2l^2)(x+L_x/2)]\eta_{\nu\pm}(y),\\
\psi^{\pm}_{2\nu}(x,y) &= (v^y_\nu)^{-1/2}\exp[i(\pm\kappa^y_\nu - W_x/2l^2)(y+L_y/2)]\exp(-ixy/l^2)\xi_{\nu\pm}(x),\\
\psi^{\pm}_{3\nu}(x,y) &= (v^x_\nu)^{-1/2}\exp[i(\pm\kappa^x_\nu + W_y/2l^2)(x-W_x-L_x/2)]\eta_{\nu\pm}(y),\\
\psi^{\pm}_{4\nu}(x,y) &= (v^y_\nu)^{-1/2}\exp[i(\pm\kappa^x_\nu - W_x/2l^2)(y-W_y-L_y/2)]\exp(-ixy/l^2)\xi_{\nu\pm}(x),
\end{aligned} \tag{2.5.6}$$

where $l = \sqrt{eB/c\hbar}$ is the magnetic length, ν denotes modes, v_ν corresponding velocities, $\eta(y)$ is an eigenfunction for a wire in the x direction, and $\xi(x)$ is that for a wire in the y direction for vector potential $(0, Bx)$.

There are two kinds of modes. One of them is a traveling mode whose wave number κ is real and the other is an evanescent mode whose wave number

has an imaginary part. The sign + (−) denotes a mode traveling or decaying in the positive (negative) direction. The wave function in a wire region is represented using expansion coefficients C with respect to modes in the wires.

$$\Psi_i(x,y) = \sum_\nu \left[C^+_{i\nu}\psi^+_{i\nu}(x,y) + C^-_{i\nu}\psi^-_{i\nu}(x,y)\right]. \tag{2.5.7}$$

The S matrix gives the following relation between incoming and outgoing waves:

$$\begin{pmatrix} \mathbf{C}_1^- \\ \mathbf{C}_2^- \\ \mathbf{C}_3^+ \\ \mathbf{C}_4^+ \end{pmatrix} = S(E) \begin{pmatrix} \mathbf{C}_1^+ \\ \mathbf{C}_2^+ \\ \mathbf{C}_3^- \\ \mathbf{C}_4^- \end{pmatrix}, \tag{2.5.8}$$

where $\mathbf{C}$ consists of expansion coefficients C and therefore its dimension is the number of modes.

We can calculate energy bands of an infinitely large array of junctions using Bloch's theorem. In the following, we confine ourselves to the case in the absence of the magnetic field for simplicity. The extension to the case in magnetic fields is discussed in [54]. The Bloch condition for $\Psi_i(x,y)$ gives the conditions for expansion coefficients, $C_{1\nu} = \exp(-ik_x a)C_{3\nu}$ and $C_{2\nu} = \exp(-ik_y b)C_{4\nu}$. These are rewritten as the following relation between incoming and outgoing waves.

$$\begin{pmatrix} \mathbf{C}_1^- \\ \mathbf{C}_2^- \\ \mathbf{C}_3^+ \\ \mathbf{C}_4^+ \end{pmatrix} = P(k_x,k_y) \begin{pmatrix} \mathbf{C}_1^+ \\ \mathbf{C}_2^+ \\ \mathbf{C}_3^- \\ \mathbf{C}_4^- \end{pmatrix}, \tag{2.5.9}$$

where

$$P(k_x,k_y) = \begin{pmatrix} 0 & 0 & P_x^{-1} & 0 \\ 0 & 0 & 0 & P_y^{-1} \\ P_x & 0 & 0 & 0 \\ 0 & P_y & 0 & 0 \end{pmatrix}, \tag{2.5.10}$$

with $P_x = \exp(ik_x a)$ and $P_y = \exp(ik_y b)$. Energy bands are given by the condition that (2.5.8) and (2.5.9) have nontrivial solutions for $\mathbf{C}$:

$$\det[P(k_x,k_y) - S(E)] = 0. \tag{2.5.11}$$

This equation gives the energy bands as equi-energy lines in $\mathbf{k}$ space at the energy E.

Figure 2.5.21 shows some examples of calculated energy bands for a square lattice ($b = a$). The parameters are $a/\lambda_F = 4.31$, $U_0/E_F = 8.4$, $d/a = 0.6$, and $\beta = 2$ (this choice of β is likely to be too large for realistic antidots). The corresponding equipotential lines are given in Fig. 2.5.13. The magnetic field $Ba^2/\Phi_0 = 4$ corresponds to about 0.4 T for a typical antidot lattice with $a \sim$ 2000 Å. The bands are calculated by including all traveling modes and a single evanescent mode having the smallest imaginary wave vector. The horizontal dotted lines denote the bottom of 1D subbands in the wire region.

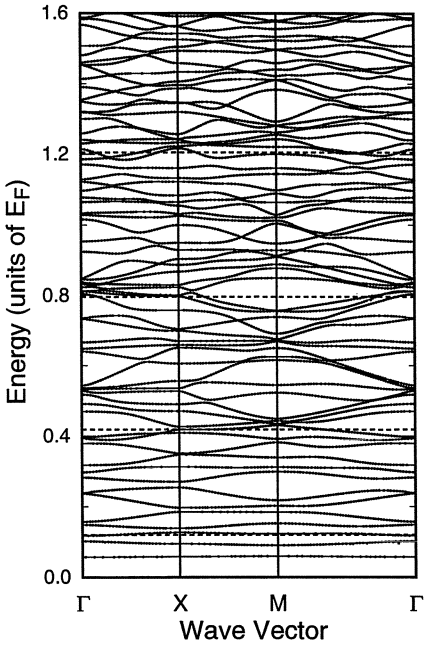


Fig. 2.5.21 Calculated energy bands in a magnetic field $Ba^2/\Phi_0 = 4$. In the S matrix method, all traveling modes and only a single evanescent mode are taken into account

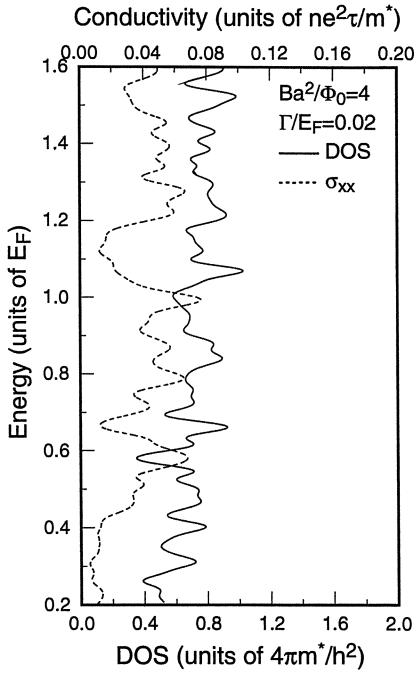


Fig. 2.5.22 The energy dependence of the density of states (solid line) and the conductivity (dotted line) in the magnetic field $Ba^2/\Phi_0 = 4$

The results using the S matrix formalism with only a single evanescent mode agree quite well with the exact results in which all the evanescent modes are included. This shows that the inclusion of a few evanescent modes is sufficient for accurate description of energy bands of antidot lattices. The bands for $Ba^2/\Phi_0 = 4$ corresponding to $2R_c/a \sim 2$ may be regarded as a mixture of a band with a large dispersion (runaway orbit) and those having a small dispersion for $0.4 \lesssim E/E_F \lesssim 1.0$ [32]. The energy bands become more complicated for higher energies.

Figure 2.5.22 shows the corresponding energy dependence of the density of states and the conductivity. The density of states is measured in units of the 2D value $m^*/\pi\hbar^2$ and the conductivity in units of $ne^2\tau/m^*$ where n is the electron concentration in the absence of antidot potential and m^* is the effective mass. They are averaged over the Gaussian distribution with broadening $\Gamma/E_F = 0.02$. Some structures of the conductivity correspond to those of the density of states, but most of them do not.

A semiclassical analysis has been made for the density of states by considering six periodic orbits (a)–(f) given in Fig. 2.5.13 [54,55]. As a cutoff of a quantum number m in (2.5.3), the factor $\exp(-\alpha m)$ has been introduced, where m is the quantum number for a motion transverse to each periodic

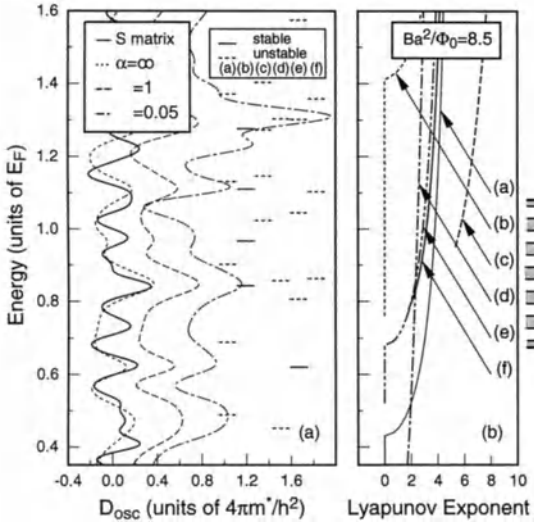


Fig. 2.5.23 (**a**) The oscillatory part of the density of states at $Ba^2/\Phi_0 = 8.5$ corresponding to $2R_c \approx 2$. The dashed line is shifted to the right by 0.4 and the dot-dashed line is shifted by 0.8. (**b**) The corresponding Lyapunov exponents

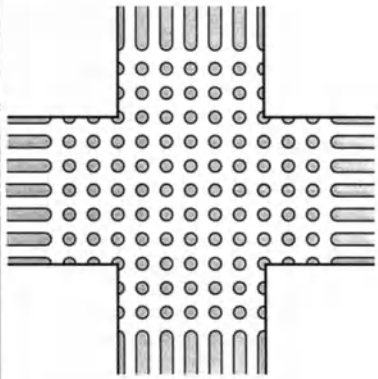


Fig. 2.5.24 Schematic illustration of a finite antidot lattice for which the numerical calculation is performed. It has 8 sites in the direction of the wire width and 16 sites in the direction of the wire length

orbit. Figure 2.5.23 shows some examples in the magnetic field corresponding to $Ba^2/\Phi_0 = 8.5$.

The orbits (a), (b), and (e) become stable in a certain energy range. The semiclassically quantized energy levels associated with such stable orbits give clear peaks in the semiclassical density of states, all of which are in agreement with those of the quantum mechanical density of states. The agreement becomes worse in the energy region higher than the Fermi energy. There are some features in the semiclassical density of states which appear only when higher transverse modes are included. For example, the inclusion of higher transverse modes broadens the peak at $E/E_F = 0.47$ and makes the density of states closer to (but still different from) the quantum-mechanical result. A more elaborate work in which many periodic orbits are taken into account is necessary for a full understanding of the correspondence between the quantum-mechanical and semiclassical results for the density of states.

In order to calculate transport quantities, we take a quantum wire junction array as a model of an antidot lattice as shown in Fig. 2.5.24 [36]. This is a kind of four-terminal system where constrictions at edges are connected respectively to four reservoirs through ideal leads. Transport coefficients are calculated based on the Büttiker–Landauer formula [56,57] from an S matrix for a junction array obtained by connecting S matrices for a single junction.

The following model potential $U(\mathbf{r})$ is used for an antidot with its center

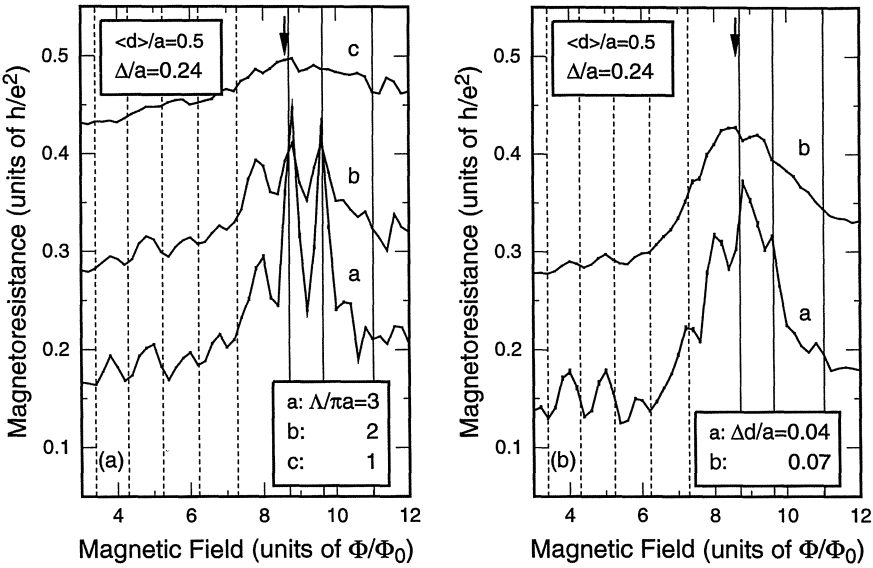


Fig. 2.5.25 (a) Magnetoresistance in the presence of impurity scattering calculated for systems consisting of a finite number of antidots in the S matrix formalism. The result for $\Lambda/\pi a = 2$ is shifted vertically by 0.1 and that for $\Lambda/\pi a = 1$ by 0.2. (b) Dependence on the disorder in the antidot diameter. The result for $\Delta d/a = 0.07$ is shifted by 0.15. The vertical lines represent quantized energy levels associated with the periodic orbit (b) calculated by the trace formula (solid lines for stable and dashed lines for unstable)

at the origin:

$$U(\mathbf{r}) = E_{\mathrm{F}}\Big(\frac{d/2+\Delta-r}{\Delta}\Big)^2\theta(d/2+\Delta-r), \tag{2.5.12}$$

where E_{F} is the Fermi energy, r is the distance from the center, d is the diameter at the Fermi energy, and $\theta(t)$ is the step function defined by $\theta(t)=1$ for $t>0$ and 0 for $t<0$. In the following, square antidot lattices are considered and parameters are chosen as $\langle d\rangle/a=0.5$, where $\langle d\rangle$ is the averaged diameter. The parameter Δ characterizing the steepness of the potential is chosen as $\Delta/a=0.24$ through comparison with self-consistent calculations [19,20]. This model potential becomes close to that given by (2.5.1) for $\beta=1$ and $d/a\sim 0.5$. Effects of impurities are introduced through random short-range scatterers and we shall consider fluctuations in the diameter of antidots.

Figure 2.5.25a shows the dependence of the calculated magnetoresistance on the mean free path Λ determined by impurity scattering in the absence of the antidot disorder. The vertical arrow indicates the magnetic field corresponding to the fundamental commensurability peak, i.e., $2R_c=a$. The AB type oscillation is clearly seen superimposed on the commensurability peak for $\Lambda/\pi a=2$ and 3. The vertical solid and dashed lines indicate semiclassical quantized levels of the stable and unstable periodic orbits, respectively, encircling a dot, which are calculated based on (2.5.5). These seem to correspond to the quantum oscillation obtained in SCBA discussed in Sect. 2.5.3.

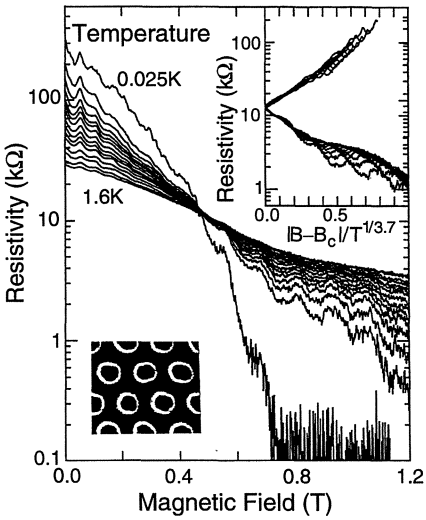


Fig. 2.5.26 Resistivity as a function of a magnetic field in a hexagonal antidot lattice at low temperatures (0.025, 0.15, 0.2, 0.3, 0.4, 0.6, 1.0, and 1.6 K). After [58]

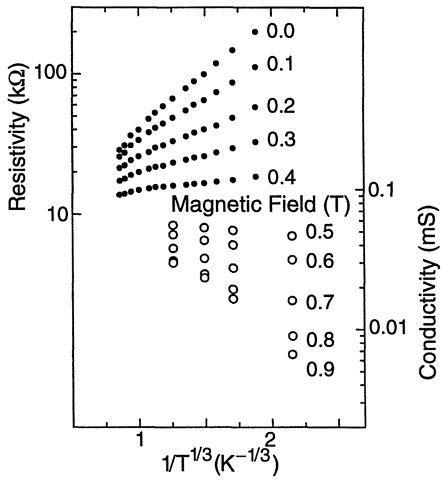


Fig. 2.5.27 Temperature variation of the resistivity at lower magnetic fields (black dots) and diagonal conductivity at higher magnetic fields (open circles) can be both explained by variable range hopping. After [58]

The calculation gives an oscillation much more prominent than experimentally observed. Therefore, mean free path much smaller than that of an unpatterned 2D system is required to obtain oscillation amplitude comparable to that observed experimentally. Figure 2.5.25b shows the dependence on disorder in the antidot diameter without impurities. It shows that the AB type oscillation is quite sensitive to antidot disorder, particularly in the magnetic-field region corresponding to the fundamental commensurability peak, demonstrating again the importance of antidot disorder.

For the amount of disorder characterized by $\Delta d/a \sim 0.07$ with the root-mean-square fluctuation $\Delta d = \sqrt{\langle (d - \langle d \rangle)^2 \rangle}$, the amplitude of AB type oscillation has become quite small; further, some peaks even turn into dips in the vicinity of the field corresponding to $2R_c = a$. Note that $\Delta d/a \sim 0.07$ corresponds to $d_f \sim 0.25$, sufficient to cause the AAS oscillation, as discussed in Sect. 2.5.4.

2.5.6 Anderson Localization

One important feature of a hexagonal antidot lattice is the appearance of an AAS oscillation in weak magnetic fields at low temperatures (1.6 K). Quite recently, behavior suggesting a strong localization was observed at lower temperatures [58]. Figure 2.5.26 shows an example of observed resistivity below 1.6 K. The resistivity increases considerably with decreasing temperature at low magnetic fields, while it increases at high magnetic fields. At a critical magnetic field, $B_c \approx 0.47$ T, the resistivity remains almost independent of the

temperature and is nearly equal to h/e^2.

Figure 2.5.27 shows that the resistivity for $B < B_c$ exhibits a temperature dependence consistent with variable range hopping below 1 K. Further, the diagonal conductivity also shows a temperature dependence reminiscent of variable range hopping for $B > B_c$. The localization length estimated from the fitting to variable range hopping turns out to be close to the lattice constant of the hexagonal lattice, showing that electrons may become completely confined in each "quantum dot" surrounded by six antidots.

Another interesting feature of the results given in Fig. 2.5.26 is the fine quantum oscillation. At lower magnetic fields ($B < B_c$) the period is determined by half of the flux quantum per unit cell and by the flux quantum at higher magnetic fields ($B > B_c$). This leads to the conclusion that the localization length oscillates with a period given by half of the flux quantum at $B < B_c$ and by the flux quantum at $B > B_c$. No theoretical work has been performed so far to understand these interesting experimental results.

References

1 D. Weiss, M.L. Roukes, A. Menschig, P. Grambow, K. von Klitzing, and G. Weimann, Phys. Rev. Lett. **66**, 2790 (1991)
2 R. Fleischmann, T. Geisel, and K. Ketzmerick, Europhys. Lett. **25**, 219 (1994)
3 G.M. Gusev, Z.D. Kvon, L.V. Litvin, Yu.V. Nastaushev, A.K. Kalagin, and A.I. Toropov, JETP Lett. **55**, 123 (1992)
4 G.M. Gusev, P. Basmaji, Z.D. Kvon, L.V. Litvin, Yu.V. Nastaushev, and A.I. Toropov, J. Phys.: Condens. Matter **6**, 73 (1994)
5 K. Nakamura, S. Ishizaka, and F. Nihey, Physica B **197**, 144 (1994)
6 F. Nihey, S.W. Hwang, and K. Nakamura, Phys. Rev. B **51**, 4649 (1995)
7 A. Lorke, J. P. Kotthaus, and K. Ploog, Phys. Rev. B **44**, 3447 (1991); Superlattices Microstruc. **9**, 103 (1991)
8 T. Yamashiro, J. Takahara, Y. Takagaki, K. Gamo, S. Namba, S. Takaoka, and K. Murase, Solid State Commun. **79**, 885 (1991)
9 J. Takahara, T. Kakuta, T. Yamashiro, Y. Takagaki, T. Shiokawa, K. Gamo, S. Namba, S. Takaoka, and K. Murase, Jpn. J. Appl. Phys. **30**, 3250 (1991)
10 T. Deruelle, K. Ensslin, P.M. Petroff, A.L. Efros, and F.G. Pikus, Phys. Rev. B **45**, 9082 (1992)
11 G. Berthold, J. Smoliner, V. Rosskopf, E. Gornik, G. Böhm, and G. Weimann, Phys. Rev. B **47**, 10383 (1993)
12 R. Schuster, K. Ensslin, J.P. Kotthaus, M. Holland, and C. Stanley, Phys. Rev. B **47**, 6843 (1993)
13 R. Schuster, K. Ensslin, D. Wharam, S. Kühn, J.P. Kotthaus, G. Böhm, W. Klein, T. Tränkle, and G. Weimann, Phys. Rev. B **49**, 8510 (1994)
14 D. Weiss, K. Richter, E. Vasiliadou, and G. Lütjering, Surf. Sci. **305**, 408 (1994)
15 T. Deruelle, B. Meurer, Y. Guldner, J.P. Vieren, M. Riek, D. Weiss, K. von Klitzing, K. Eberl, and K. Ploog, Phys. Rev. B **49**, 16561 (1994)
16 K. Tsukagoshi, S. Wakayama, K. Oto, S. Takaoka, K. Murase, and K. Gamo, Superlatt. Microstruct. **16**, 295 (1994)
17 F. Nihey and K. Nakamura, Physica B **184**, 398 (1993)
18 D. Weiss, K. Richter, A. Menschig, R. Bergmann, H. Schweizer K. von Klitzing, and G. Weimann, Phys. Rev. Lett. **70**, 4118 (1993)
19 A. Kumar, S.E. Laux, and F. Stern, Phys. Rev. B **42**, 5166 (1990)
20 T. Suzuki and T. Ando, J. Phys. Soc. Jpn. **62**, 2986 (1993); Physica B **201**, 345 (1994); **227**, 46 (1996)
21 R. Fleischmann, T. Geisel, and R. Ketzmerick, Phys. Rev. Lett. **68**, 1367 (1992)

22 E.M. Baskin, G.M. Gusev, Z.D. Kvon, A.G. Pogosov, and M.V. Entin, JETP Lett. **55**, 678 (1992)
23 K. Ensslin, S. Sasa, T. Deruelle, and P.M. Petroff, Surf. Sci. **263**, 319 (1992)
24 J. Takahara, A. Nomura, K. Gamo, S. Takaoka, K. Murase, and H. Ahmed, Jpn. J. Appl. Phys. **34**, 4325 (1995)
25 K. Tsukagoshi, S. Wakayama, K. Oto, S. Takaoka, K. Murase, and K. Gamo, Phys. Rev. B **52**, 8344 (1995)
26 K. Tsukagoshi, M. Haraguchi, K. Oto, S. Takaoka, K. Murase, and K. Gamo, Jpn. J. Appl. Phys. **34**, 4335 (1995)
27 K. Tsukagoshi, M. Haraguchi, S. Takaoka, and K. Murase, J. Phys. Soc. Jpn. **65**, 811 (1996)
28 T. Nagao, J. Phys. Soc. Jpn. **64**, 4097 (1995)
29 T. Nagao, J. Phys. Soc. Jpn. **65**, 2606 (1996)
30 S. Ishizaka and T. Ando, Phys. Rev. B **55**, 16331 (1997)
31 W.-C. Lu and O.K. Anderson, *Proceedings of 23rd International Conference on Physics of Semiconductors*, edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore, 1996), p. 1497
32 S. Ishizaka, F. Nihey, K. Nakamura, J. Sone, and T. Ando, Jpn. J. Appl. Phys. **34**, 4317 (1995)
33 R. Schuster, G. Ernst, K. Ensslin, M. Entin, M. Holland, G. Böhm, and W. Klein, Phys. Rev. B **50**, 8090 (1994)
34 K. Tsukagoshi, T. Nagao, M. Haraguchi, S. Takaoka, K. Murase, and K. Gamo, J. Phys. Soc. Jpn. **65**, 1914 (1996)
35 H. Silberbauer and U. Rössler, Phys. Rev. B **50**, 11911 (1994)
36 S. Uryu and T. Ando, *Proceedings of 23rd International Conference on Physics of Semiconductors*, edited by M. Scheffler and R. Zimmermann (World Scientific, Singapore, 1996), p. 1505
37 S. Ishizaka, F. Nihey, K. Nakamura, J. Sone, and T. Ando, Phys. Rev. B **51**, 9881 (1995)
38 D. Hofstadter, Phys. Rev. B **14**, 2239 (1976)
39 R.B.S. Oakshott and A. MacKinnon, J. Phys: Condens. Matter **5**, 6971 (1993); **5**, 6983 (1993); **5**, 6991 (1993); **6**, 1519 (1994)
40 M.C. Gutzwiller, J. Math. Phys. **8**, 1979 (1967); **10**, 1004 (1969); **11**, 1791 (1970); **12**, 343 (1971)
41 M.C. Gutzwiller, *Path Integrals*, edited by G.J. Papadopoulos and J.T. Devrees (Plenum, New York, 1978), p. 163
42 M.C. Gutzwiller, *Chaos in Classical and Quantum Mechanics* (Springer, New York, 1990)
43 W.H. Miller, J. Chem. Phys. **63**, 996 (1975)
44 K. Richter, Europhys. Lett. **29**, 7 (1995)
45 M. Wilkinson, J. Phys. A **20**, 2415 (1987)
46 G. Hackenbroich and F. von Oppen, Europhys. Lett. **29**, 151 (1995)
47 B.L. Altshuler, A.G. Aronov, and B.Z. Spivak, JETP Lett. **33**, 1255 (1981)
48 D.Yu. Sharvin and Yu.V. Sharvin, JETP Lett. **34**, 272 (1981)
49 G.J. Dolan, J.C. Licini, and D.J. Bishop, Phys. Rev. Lett. **56**, 1493 (1986)
50 T. Nakanishi and T. Ando, Phys. Rev. B **54**, 8021 (1996)
51 T. Nakanishi and T. Ando, Physica B **227**, 127 (1996)
52 T. Ando, Phys. Rev. B **44**, 8017 (1991)
53 S. Uryu and T. Ando, Jpn. J. Appl. Phys. **34**, 4295 (1995)
54 S. Uryu and T. Ando, Phys. Rev. B **53**, 13613 (1996)
55 S. Uryu and T. Ando, Physica B **227**, 138 (1996)
56 R. Landauer, IBM J. Res. Dev. **1**, 223 (1957); Philos. Mag. **21**, 863 (1970)
57 M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986); IBM J. Res. Dev. **32**, 317 (1988)
58 F. Nihey, M.A. Kastner, and K. Nakamura, Phys. Rev. B **55**, 4085 (1997)

2.6 Electric and Magnetic Lateral Superlattices

2.6.1 Lateral Modulation

One can impose artificial potential on a two-dimensional electron gas (2DEG) at the heterointerface of GaAs/AlGaAs by depositing a suitably patterned gate electrode on the surface. A periodic lateral potential can give rise to a rich variety of transport phenomena in 2DEG, in which the electron mean free path can be much larger than the lateral superlattice period. Magnetotransport in 2DEG under a periodic potential shows what can be generally called "commensurability effects," which arise from the geometrical resonance depending on the commensurability condition between the lattice periodicity and the cyclotron orbit. An example of this class of transport phenomena, discussed in the preceding section, is magnetotransport in 2DEG subject to a 2D potential modulation. In this section, we discuss a related phenomenon, called Weiss oscillation [1,2], which involves 2DEG in a 1D superlattice potential.

2.6.2 Weiss Oscillation

We choose the coordinate axes such that 2DEG is in the xy plane, the potential modulation is along the x axis, and a uniform magnetic field is applied along the z axis. Taking a Landau gauge for the vector potential, the one-electron Schrödinger equation can be written as

$$\left\{\frac{\hbar^2}{2m^*}\left[\left(\frac{\partial}{i\partial x}\right)^2+\left(\frac{\partial}{i\partial y}-eBx\right)^2\right]+V_0\cos(\kappa x)\right\}\psi(x,y)=\varepsilon\psi(x,y), \quad (2.6.1)$$

where $\kappa=2\pi/a$ is the modulation wavevector. In the absence of the potential term, this gives a series of Landau levels $\varepsilon=\hbar\omega_c(n+1/2)$, each of which has a degeneracy $1/2\pi l^2=eB/h$ reflecting the independence of the energy from the guiding center coordinate $X=-k_y/l^2$. The presence of the potential lifts the degeneracy of the Landau levels and turns them into Landau subbands. For a weak modulation, the subband dispersion may be calculated perturbatively,

$$\begin{aligned}\varepsilon(n,k_y)&=\hbar\omega_c\left(n+\frac{1}{2}\right)+\langle n,X|V_0\cos(\kappa x)|n,Y\rangle\\&\approx\hbar\omega_c\left(n+\frac{1}{2}\right)+V_0J_0(\kappa R_c)\cos(\kappa X),\end{aligned} \quad (2.6.2)$$

where $R_c=k_Fl^2=\hbar k_F/eB$ is the cyclotron radius for an electron on the Fermi surface with k_F being the Fermi wave vector. The subband dispersion vanishes at the zeros of the Bessel function, which by recalling the asymptotic form $J_0(\kappa R_c)=(2/\pi\kappa R_c)^{1/2}\cos(\kappa R_c-\pi/4)$, can be approximately written as

$$\frac{2R_c}{a}=n-\frac{1}{4}, \qquad n=1,\,2,\,3,\,\ldots. \quad (2.6.3)$$

When the cyclotron radius and the potential period fulfill the above commensurability condition, the subbands near the Fermi level become flat, which implies vanishing of the velocity of the electrons at the Fermi level. This makes

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

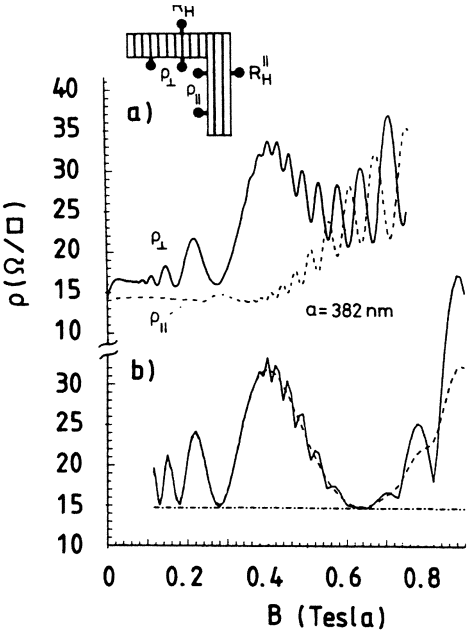


Fig. 2.6.1 Mangetoresistance traces of a sample shown in the inset. The solid and broken curves represent the data with current perpendicular and parallel to the stripes, respectively. After [1]

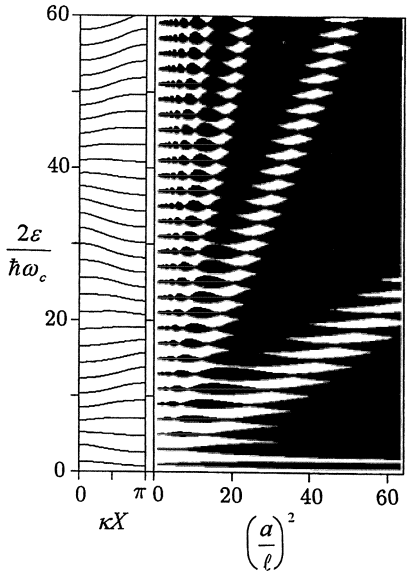


Fig. 2.6.2 Energy spectrum of 2DEG in the presence of a modulated magnetic field

the conductivity component σ_{yy} minimum, which in turn results in a series of minima in the longitudinal magnetoresistance $\rho_{xx}(B) = \sigma_{yy}/(\sigma_{xx}\sigma_{yy} - \sigma_{xy}\sigma_{yx})$ as shown in Fig. 2.6.1. The potential modulation in this particular experiment was generated by means of a laser interference pattern and persistent photoconductivity [1].

The physical origin of the resistance minima can be also interpreted in the following semiclassical picture [3]. The potential modulation gives rise to a spatially alternating electric field parallel to the x direction. An electron subject to this spatially alternating electric field and the magnetic field along the z axis acquires a (local) $\mathbf{E}\times\mathbf{B}$ drift velocity, whose y component is given by $v_y(x) = (\kappa V_0/B)\sin(\kappa x)$. The drift velocity in the y direction of an electron with the guiding center at (X, Y) is given by averaging over a full cyclotron motion:

$$\langle v_y \rangle = \frac{1}{2\pi}\int_0^{2\pi} \frac{\kappa V_0}{B} \sin[\kappa(X + R_c \sin\theta)] d\theta. \tag{2.6.4}$$

This quantity is enhanced or suppressed depending on whether the local electric fields at the two end points, $X \pm R_c$, have the same or opposite direction. This condition for the suppression of the guiding-center drift is given, as it should be, by (2.6.3).

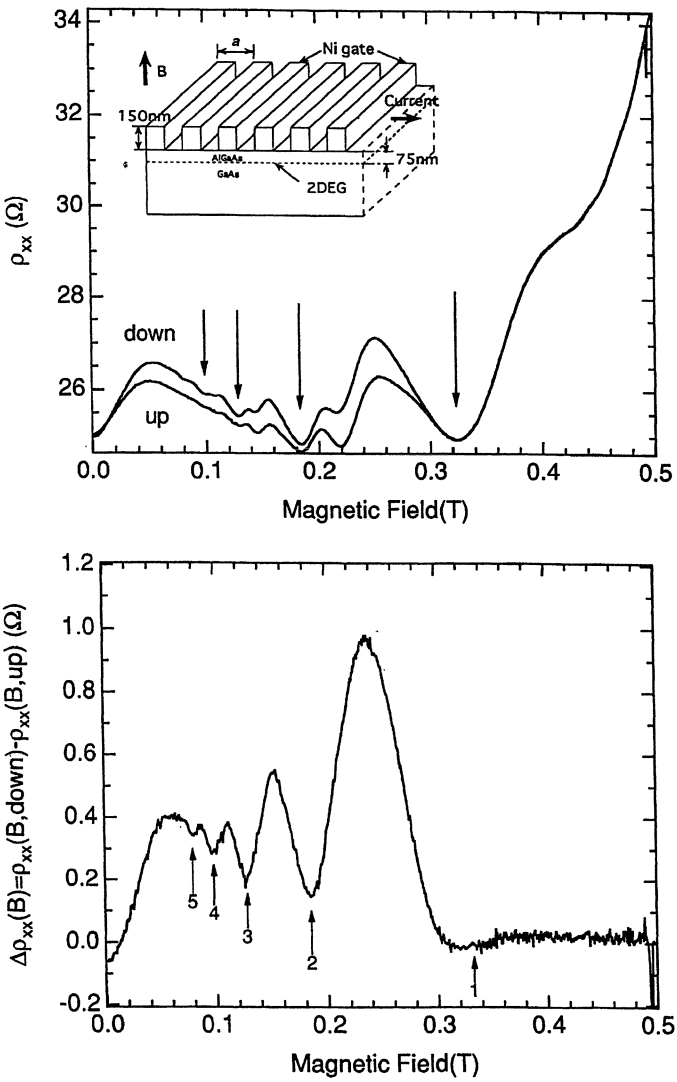


Fig. 2.6.3 Upper panel: Magnetoresistance of a sample illustrated in the inset. Lower panel: Hysteretic part of the magnetoresistance extracted from the data in the upper panel. The arrows indicate the expected position of resistance minima for the magnetic Weiss oscillation

2.6.3 Magnetic Weiss Oscillation

Let us now turn to the case of a spatially modulated magnetic field [4–6]. We consider a magnetic field parallel to the z direction whose strength varies periodically with x, i.e., $B_z = B_0 + B_1 \cos(\kappa x)$. The corresponding Landau

gauge is $\mathbf{A}=(0, B_0x+(B_1/\kappa)\sin(\kappa x))$ and the Schrödinger equation is

$$\left\{\frac{\hbar^2}{2m^*}\left[\left(\frac{\partial}{i\partial x}\right)^2+\left(\frac{\partial}{i\partial y}-eB_0x+\frac{B_1}{\kappa}\sin(\kappa x)\right)^2\right]+V_0\cos(\kappa x)\right\}\psi(x,y)=\varepsilon\psi(x,y), \tag{2.6.5}$$

Assuming that the modulation is weak ($B_1 \ll B_0$), we only need to consider the first order term in B_1 given by $U(x) = -(e^2B_1B_0/m^*\kappa)x\sin(\kappa x)$. The energy spectrum corresponding to (2.6.2) is given by

$$\varepsilon(n,k_y) \approx \hbar\omega_c\left(n+\frac{1}{2}\right)+\frac{e^2}{m^*}\frac{B_0B_1}{\kappa}XJ_1(\kappa R_c)\cos(\kappa X). \tag{2.6.6}$$

The coefficient of the second term contains the Bessel function $J_1(\kappa R_c)$ in place of $J_0(\kappa R_c)$ in (2.6.2). The flat subband condition for the case of magnetic modulation is

$$\frac{2R_c}{a}=n+\frac{1}{4}, \qquad n=1,\,2,\,3,\ldots. \tag{2.6.7}$$

Figure 2.6.2 shows an example of an energy spectrum calculated as a function of the ratio of the modulation period and the cyclotron radius [5].

Such a magnetic-field modulation can be realized by placing a patterned ferromagnetic material on the surface of a 2DEG sample and applying a uniform magnetic field. However, it is difficult in practice to produce a purely magnetic modulation without any potential modulation. The presence of the patterned gate electrode generally creates potential modulations by various mechanisms, including strain-induced deformation potential or piezoelectric potential modulation and spatially-dependent surface band bending. Izawa et al. [7] have shown that an appropriate choice of the gate bias can counterbalance such a "built-in" potential modulation, leaving an almost purely magnetic modulation.

The upper panel of Fig. 2.6.3 shows an example of the magnetoresistance trace that shows the magnetic Weiss oscillation. The hysteresis is associated with that of the ferromagnetic gate material (Ni in this case), i.e., the difference in the magnetization in the up and down sweep of the external magnetic field. The hysteretic part of the magnetoresistance shown in the lower panel of Fig. 2.6.3, which is of purely magnetic origin, has a minimum position in excellent agreement with the condition given by (2.6.7). The magnetic Weiss oscillation was also observed by Ye et al. [8] and by Carmona et al. [9]. The latter group used magnetic flux exclusion from superconducting strips to generate a modulated magnetic field.

Magnetoresistance under the combined influence of electric and magnetic modulations has been calculated by Peeters and Vasilopoulos [6]. As one goes from a purely electric modulation limit to a purely magnetic modulation limit, the magnetoresistance minima change to maxima. When the electric

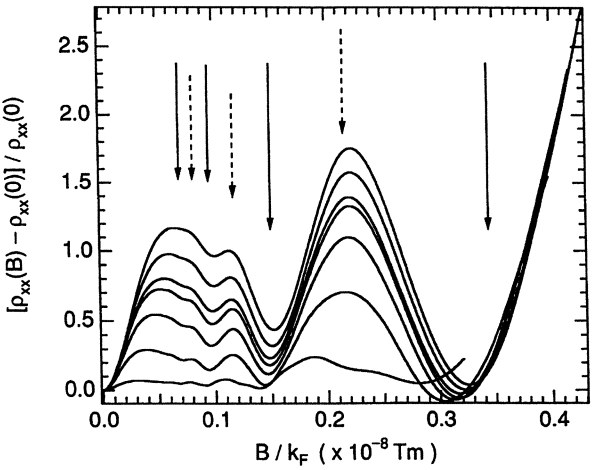


Fig. 2.6.4 Magnetoresistance in the presence of electric and magnetic modulations. The gate bias changes the strength of the electric potential modulation

and magnetic modulations are in phase, the magnetoresistance is given by

$$\begin{aligned}\frac{\rho_{xx}(B)-\rho_{xx}(0)}{\rho_{xx}(0)} &= \frac{ak_F}{2\pi^2}\frac{\hbar\omega_1}{\hbar\omega_0}\frac{\hbar\omega_1}{E_F}(\omega_0\tau)^2(1+\delta^2) \\ &\times\left[1-A\left(\frac{T}{T_a}\right)+A\left(\frac{T}{T_a}\right)\sin^2\left(\frac{2\pi R_c}{a}-\frac{\pi}{4}+\phi\right)\right],\end{aligned} \tag{2.6.8}$$

where $\omega = eB_0/m^*$, $\omega_1 = eB_1/m^*$, $k_BT_a = \hbar\omega_0/4\pi^2akF$, $A(x) = x/\sinh x$, and E_F is the Fermi energy. The quantity $\delta = 2\pi V_0/ak_F\hbar\omega_0\tan\phi$ is a dimensionless parameter characterizing the relative amplitude of the electric and magnetic modulations. The phase of the oscillating component (the third term in the square bracket of (2.6.8)) and hence the position of the resistance minimum shift with δ.

When the electric and magnetic modulations are out of phase, the positions of the minima and maxima do not shift, instead minima simply turn to maxima and vice versa when the gate bias is changed [10]. Figure 2.6.4 shows an example of the shift of the resistance minima as the gate bias is varied, corresponding to the in-phase case. Here the horizontal axis is taken as B/k_F, so as to normalize the change in the value of k_F with the gate bias. Setting the bias condition so as to realize a purely magnetic modulation and fitting (2.6.8), the amplitude of the magnetic modulation at the plane of 2DEG is estimated as $B_1 \approx 20$ mT for this sample.

One can obtain a larger amplitude of the magnetic modulation by using materials with a larger saturation moment. Often it helps to use a tilted magnetic field, fully magnetizing the gate material by the total magnetic field, while measuring the magnetoresistance as a function of the normal component. In the extreme case of a parallel magnetic field, an alternating magnetic

field with zero average value may be realized. Such a situation is of great interest in the context of the composite fermion picture of a half-filled Landau level, where a spatial variation of the electron density can be translated to a magnetic-field variation and a new type of electron trajectory called a snake orbit may play an important role in the transport. Also, there is great theoretical interest in the problem of electron localization in a spatially random magnetic field. Although it is rather difficult to actually generate a random magnetic field without simultaneously inducing any random potential, 2DEG with a microfabricated magnetic material seems to be the best candidate for an experimental approach to such problems.

References

1 R.R. Gerhardt, D. Weiss, and K. von Klitzing, Phys. Rev. Lett. **62**, 1173 (1989)
2 R.W. Winkler and J.P. Kotthaus, Phys. Rev. Lett. **62**, 1177 (1989)
3 C.W.J. Beenakker, Phys. Rev. Lett. **62**, 2020 (1989)
4 D. Yoshioka and Y. Iye, J. Phys. Soc. Jpn. **56**, 448 (1987)
5 R. Yagi and Y. Iye, J. Phys. Soc. Jpn. **62**, 1279 (1993)
6 F.M. Peeters and P. Vasilopoulos, Phys. Rev. B **46**, 4667 (1992); F.M. Peeters and P. Valilopoulos, Phys. Rev. B **47**, 146 (1993)
7 S. Izawa, S. Katsumoto, A. Endo, and Y. Iye, J. Phys. Soc. Jpn. **64**, 706 (1995)
8 P.D. Ye, D. Weiss, R.R. Gerhardt, M. Seeger, and K. von Klitzing, Phys. Rev. Lett. **74**, 3013 (1995)
9 H.A. Carmona, A.K. Geim, A. Norgaret, P.C. Main, T.J. Foster, and A. Henini, Phys. Rev. Lett. **74**, 3009 (1995)
10 A. Endo, S. Izawa, S. Katsumoto, and Y. Iye, Surf. Sci. **361/362**, 333 (1996)

2.7 Terahertz Spectroscopy of Nanostructures

2.7.1 Introduction

Many of the characteristic energy scales in semiconductor quantum nanostructures such as quantized subband energies are typically between 1 meV and 100 meV. On the other hand, characteristic time scales in nanostructures such as scattering times and tunneling times are typically 0.1–10 ps. All these energy and time scales, if converted to frequencies of electromagnetic radiation, lie in the terahertz (THz) range. Therefore, the investigation of interaction between electrons and THz electromagnetic radiation gives us important information on the electronic states and dynamical transport properties of electrons in quantum nanostructures, which is not accessible by dc transport experiments. In this section, basic techniques of THz spectroscopy are briefly introduced and some recent topics on the THz spectroscopy of quantum nanostructures are discussed.

2.7.2 Swept-Frequency THz Spectroscopy

Since the THz frequency range corresponds to the far-infrared (FIR) range, frequency-domain THz spectroscopy is often called FIR spectroscopy as well. Swept-frequency FIR spectroscopy with a Fourier transform spectrometer is the most extensively used technique for spectroscopy of low-dimensional electron systems. Broadband FIR radiation from a Hg lamp source is transmitted through a Michaelson interferometer with one moving mirror and focused onto the sample. FIR spectroscopy experiments are usually done in a transmission configuration. The radiation transmitted through the sample is detected with a cryogenic bolometer. Good reviews can be found in [1,2]. Swept-frequency FIR spectroscopy has been extensively used to investigate electronic states of two-dimensional (2D) electron systems in metal-oxide-semiconductor field-effect transistors (MOSFETs) and quantum wells [2,3].

Recently, this technique has been applied also for the investigation of electronic states of low-dimensional electron systems in quantum wires (QWRs) [2,4,5] and quantum dots [6–8]. The advantage of FIR spectroscopy lies in the fact that direct information on the quantized electronic states in quantum nanostructures can be obtained without making electrical contacts, which may perturb the electronic states in the system. For FIR transmission experiments, it is necessary to fabricate nominally identical submicron-sized nanostructures over an area of approximately 10 mm^2 in order to obtain measurable FIR signals. Review articles on the spectroscopy of multiple quantum wires and quantum dots can be found in [2,8] and references therein.

2.7.3 Electronic States in Single Quantum Wire Structure

Electrostatic confinement by split-gate geometry [9] is a common method to realize quasi-one-dimensional (1D) electron systems. However, since in such structures both channel widths and electron densities are varied by changing the split-gate voltage V_g, it is highly desirable to establish accurate and reliable methods to determine the electronic states in such structures. Numerical calculations on split-gate QWR structures [10] have shown that the

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

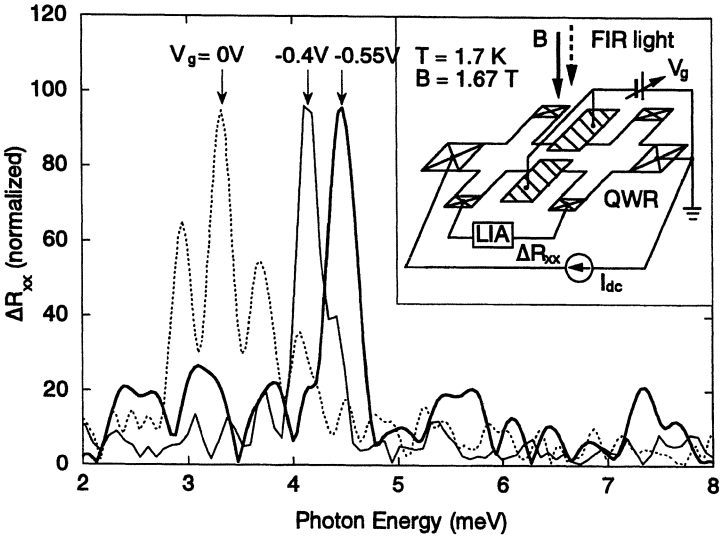


Fig. 2.7.1 Spectra of the FIR photoresponse of the magnetoresistance of a split-gate QWR ($L = 20$ μm, $W_0 = 0.3$ μm) measured at $B = 1.67$ T. The inset shows a schematic diagram of the experimental setup for the FIR photoinduced resistance measurements

self-consistent potential $V_{sc}(x)$ in the constricted region is mainly determined by two factors: the bare confinement potential $V_b(x)$ imposed by the split-gate electric field and the Hartree potential $V_H(x)$ due to Coulomb repulsion of electrons in the channel. It is known that $V_b(x)$ is well approximated by a parabolic shape in the central part of the QWR structure, while $V_{sc}(x)$ has the shape of a truncated parabola with a flat bottom [10]. Experimentally, magnetoresistance measurements on the magnetic depopulation (MD) effect [11] have been so far the most commonly used method to estimate $V_{sc}(x)$ in QWR structures. Another frequently used method is the FIR transmission spectroscopy as described above. For FIR transmission measurements, however, multi-parallel QWR structures fabricated uniformly over an area of approximately 3×3 mm^2 are necessary, which is not an easy task. The signal obtained from such measurements is an average of those of thousands of QWRs. Therefore, in general, conventional FIR transmission spectroscopy is not suitable to investigate the electronic states in single (or a few) quantum nanostructures.

However, this problem can be overcome by using a nanostructure itself as an FIR detector. The inset of Fig. 2.7.1 shows a schematic diagram of the experimental setup for investigating the electronic structures of a single QWR structure. Broadband FIR light from a Hg lamp and a magnetic field B are incident normally to the surfaces of a split-gate GaAs/AlGaAs QWR sample (the Faraday configuration). The photoinduced resistance change ΔR_{xx} of the diagonal magnetoresistance of the samples is detected by a conventional lock-in technique. Figure 2.7.1 shows the spectra of ΔR_{xx} of a split-gate QWR sample with a split-gate gap $W_0 = 0.30$ μm measured at $B = 1.67$

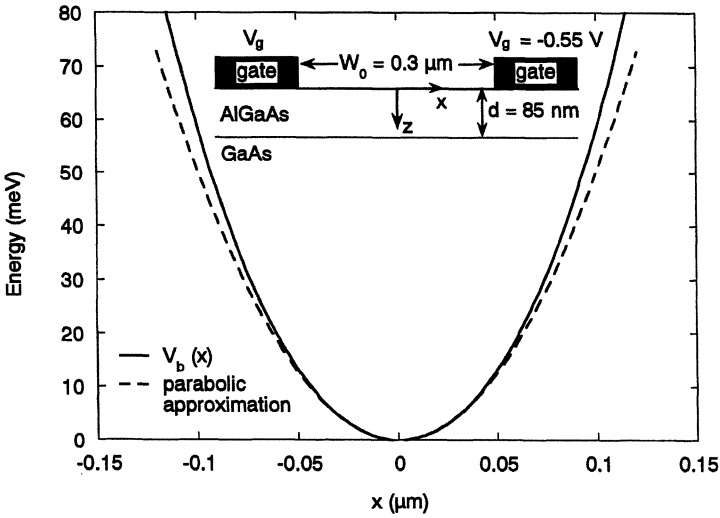


Fig. 2.7.2 The calculated bare lateral confinement potential $V_b(x)$ (solid line) in the split-gate quantum wire with $W_0 = 0.3$ mm when $V_g = -0.55$ V. The dotted line is a parabolic approximation as described in the text. The inset gives a cross-sectional view of the split-gate structure

T [12]. When $V_g < -0.25$ V and a quasi-1D channel is formed, ΔR_{xx} starts exhibiting a single narrow peak, which shifts to higher energies with increasing magnitude of negative V_g. The photoresponse of R_{xx} is attributed to electron heating due to resonant FIR absorption corresponding to transitions between the magnetoelectric subbands.

From simple electrostatics, the shape of the bare confinement potential $V_b(x)$ imposed at the heterointerface by the split-gate voltage V_g is expressed by [13]

$$V_b(x) = \frac{V_g}{\pi} \tan^{-1}\left[\frac{W_0 d}{d^2+x^2-W_0^2/4}\right], \tag{2.7.1}$$

where d is the distance between the sample surface and the electron layer. It can be well approximated by a parabola whose curvature is expressed as $m^*\omega_0^2$ in a relatively wide region near the gap center, as seen in Fig. 2.7.2. This fact indicates that in the presence of a constant magnetic field B, the total bare magnetoelectric confinement in the QWR region can be expressed by a parabolic potential with a characteristic frequency

$$\omega = \sqrt{\omega_c^2+\omega_0^2}, \tag{2.7.2}$$

where $\omega_c = eB/m^*$ is the cyclotron frequency. Kohn's theorem states that an electron system in a parabolic confinement potential absorbs light at the total bare harmonic oscillator frequency ω given by (2.7.2) independent of the electron-electron interaction [14]. Since ω_c at fixed B is known, this implies that $V_b(x)$ or ω_0 in the QWR can be directly determined by measuring the

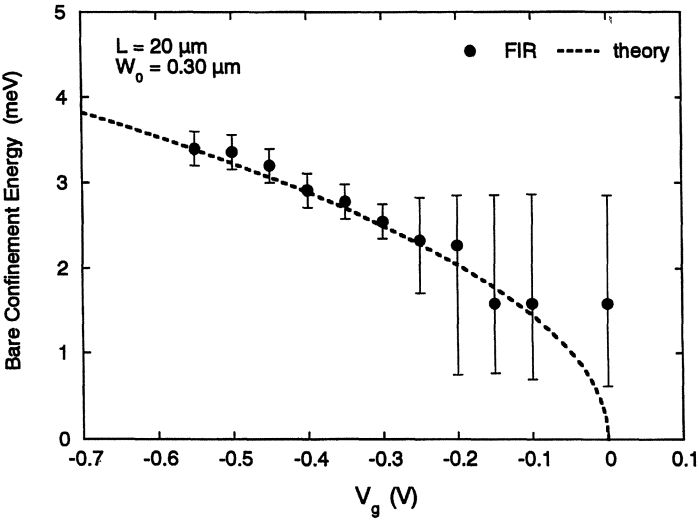


Fig. 2.7.3 Lateral bare confinement energy in the split-gate QWR with $W_0 = 0.30$ μm. Circles: FIR result; dashed line: the calculated bare confinement energy

resonant absorption frequency ω for each V_g by using FIR spectroscopy. By analyzing the shift of ΔR_{xx} shown in Fig. 2.7.1, it is possible to determine $\hbar\omega_0$ as a function of the split-gate voltage V_g as plotted in Fig. 2.7.3 by full circles. The results show that $\hbar\omega_0$ gradually increases with decreasing V_g.

The characteristic frequency of the bare confinement potential ω_0 can be calculated by taking the second derivative of (2.7.1) with respect to x at $x = 0$. The calculated $\hbar\omega_0$ is also displayed in Fig. 2.7.3, which shows that ω_0 determined by FIR spectroscopy is in excellent agreement with the theoretical results. This good agreement indicates that the FIR spectroscopy is a powerful tool for determining the electronic states even in individual quantum nanostructures. The ability to detect weak FIR radiation from a Hg lamp by a split-gate QWR structure also suggests the possibility of realizing voltage-tunable FIR detectors by using quantum nanostructures.

2.7.4 Blackbody Radiation from Hot Carriers

High-mobility electron systems in AlGaAs/GaAs heterostructures and quantum wires are suitable for realizing high-speed field-effect transistors. However, the high-mobility electrons in the channels of such transistors are easily heated up by an external electric field, leading to a rapid rise in the electron temperature and a reduction of electron mobility. Such a "hot electron" effect is one of the most important issues in practical device applications. To learn the electron temperature, one can study their blackbody emission spectrum in the FIR range.

An electron system kept at a temperature T emits radiation whose spectral emission intensity per solid angle is given by

$$i_b(\omega)d\omega d\Omega = i_{bb}(\omega)A(\omega,\theta)\cos\theta d\omega d\Omega, \tag{2.7.3}$$

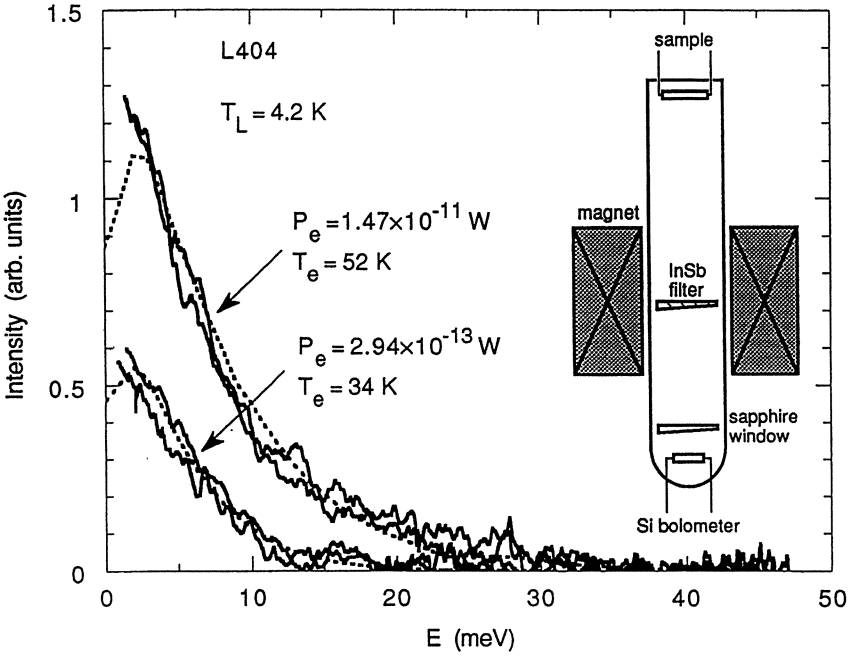


Fig. 2.7.4 FIR emission spectra from hot 2D electrons. Also plotted by dotted lines are the theoretical curves for blackbody radiation. The inset shows a schematic illustration of the experimental setup for the FIR emission spectroscopy

where $i_{bb}(\omega)$ is the spectral emission intensity of a real blackbody, which can be found in ordinary textbooks, Ω the solid angle, and $A(\omega,\theta)$ the incident-angle dependent spectral absorptivity of the electron system. For example, for an AlGaAs/GaAs single interface 2D electron system, $A(\omega,\theta)$ for normal incidence is expressed as

$$A(\omega,\theta=0) = 4\mathrm{Re}(F)/|1+\sqrt{\kappa}+F|^2, \tag{2.7.4}$$

where $F = \sigma(\omega)\varepsilon_0 c$ with $\sigma(\omega)$ being the dynamical conductivity of the 2D electrons, ε_0 the dielectric permittivity of the vacuum, and κ the dielectric constant of the material.

Experimental difficulties in performing spectroscopy on FIR radiation from nanostructures arise from the fact that such thermal radiation is usually very weak (at most of the order of nW). In order to overcome such difficulties, it is necessary to use a special experimental setup [14,15] as schematically shown in the inset of Fig. 2.7.4. The FIR radiation emitted from the sample is detected by a broadband detector (bolometer). Between the sample and the detector, a wedged semiconductor slab is inserted at the center of the superconducting solenoid. This semiconductor slab serves as a tunable notch filter whose blocking frequency is the cyclotron frequency of the carrier in the filter material. The cyclotron frequency can be swept by changing the

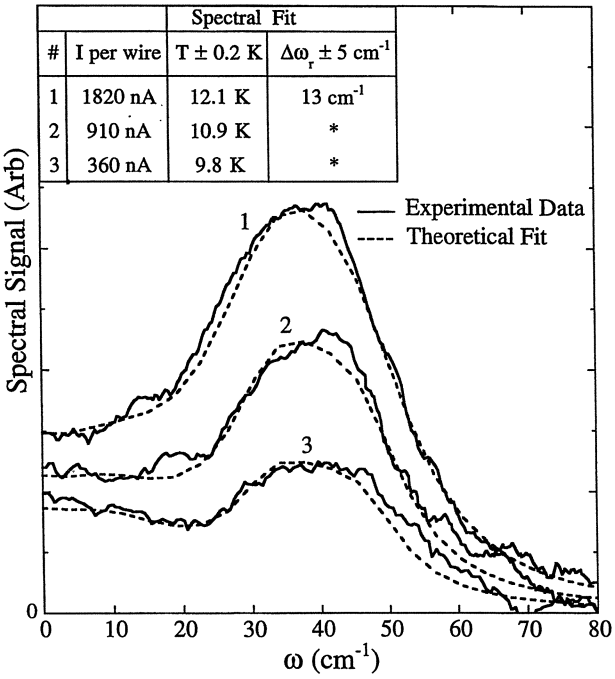


Fig. 2.7.5 FIR emission spectra from hot electrons in the QWR array. The dashed lines are the theoretical fits to the experimental data. The table lists the determined electron temperatures. After [18]

magnetic field. A suitable filter material is selected by considering the needed frequency range (i.e., the carrier effective mass) and spectral resolution (i.e., the carrier mobility).

Figure 2.7.4 shows typical FIR radiation spectra from hot 2D electrons in an AlGaAs/GaAs heterostructure excited by applying pulsed electric fields. As seen in the figure, the FIR emission spectra from hot 2D electrons are broad and, for higher excitations, both intensity and width of the emission increase. The dotted lines in the figure are the theoretical fits of the blackbody emission, assuming that $\sigma(\omega)$ is expressed by classical Drude conductivity. The good agreement between experiment and theory confirms that the observed FIR radiation is the blackbody radiation from the hot 2D electron system. By analyzing the intensity and the spectral shape of the emission, it is possible to sensitively determine the thermodynamic temperature of the electron system for a wide temperature range [15].

When a grating coupler is formed on the surface of the 2D electron system, conductivities with a finite wavevector (i.e., plasma oscillations) become accessible by FIR spectroscopy [16]. Narrowband radiation peaks due to thermally excited plasmons have been observed by using a similar setup with a high-mobility GaAs filter and the thermodynamical temperature of the hot plasmon system has also been discussed [17].

The same technique can be applied for investigating hot electron temperatures of quasi-1D electron system [18]. Figure 2.7.5 shows the FIR emission spectra from hot electrons in 300 nm-wide multiple quantum wire structures fabricated by x-ray lithography and wet chemical etching over an area of 3×3 mm^2. As seen in the figure, the FIR emission spectra consist of two components; one is a broadband background emission extending up to 80 cm^{-1} and the other a relatively narrowband emission centered at 40 cm^{-1} (the observed line width is dominated by the filter resolution). The broadband emission is the blackbody radiation due to the longitudinal motion of electrons along the quantum wires. On the other hand, the relatively-narrowband emission is due to electron motion perpendicular to the wires, i.e., thermally excited intersubband transitions between quantized 1D subbands. Note again that the observed intersubband emission frequency is determined by the bare confinement potential [19] and not the self-consistent subband spacings.

The dotted lines in the figure show the theoretical fits calculated by assuming that conductivity along the wire can be expressed by Drude conductivity and that the intersubband transition can be modeled as a Lorentzian oscillator [18]. Comparison between theory and experiment can be used to determine the temperature of the 1D hot electron system, as listed in the table in the figure. In general, the intersubband emission peak is much more sensitive to temperature than the Drude emission, although both components imply the same temperature.

2.7.5 Summary

So far, we have seen that the frequency-domain THz spectroscopy is a powerful tool for investigating the electronic structures and dynamical transport properties of semiconductor nanostructures. However, recent technological development of laser quantum electronics has enabled the generation of ultrashort optical pulses whose width is approximately 100 fs and even made this a standard laboratory tool. Such ultrafast optical technologies have opened possibilities of investigating dynamical motion of electron wave packets in the time domain. In the last ten years, time-resolved THz spectroscopy with photoconductive dipole antennas [20] has emerged as a powerful technique to study transient ultrafast phenomena in semiconductor nanostructures. For example, impulsive THz emission of electromagnetic waves due to quantum beats in coupled quantum wells [21], Bloch oscillations in semiconductor superlattices [22], and plasma oscillations of 2D electrons [23] have been observed. Such time-domain spectroscopy provides direct information on the dynamics of low-dimensional electron wave packets, which is complimentary to the frequency-domain data. Furthermore, it is expected that the development of such techniques will bridge the gap between high-frequency electronics and ultrafast optics.

References

1 E. Batke and D. Heitmann, Infrared Phys. **24**, 189 (1984)
2 J.P. Kotthaus, *Interface, Quantum Wells, and Superlattices, NATO ASI Series B*, edited by C.R. Leavens and R. Taylor (Plenum, London), p. 95

3 D. Heitmann, *Physics and Applications of Quantum Wells and Superlattices, NATO ASI Series B*, edited by E.E. Mendez and K. von Klitzing (Plenum, London), p. 317
4 S.J. Allen, Jr., F. DeRosa, G.J. Dolan, and C.W. Tu, *Proceedings of the 17th International Conference on the Physics of Semiconductors*, edited by J.D. Chadi and W.A. Harrison (Springer, New York, 1985), p. 313
5 W. Hansen, M. Horst, J.P. Kotthaus, U. Merkt, and Ch. Sikorski, Phys. Rev. Lett. **58**, 2586 (1987)
6 S.J. Allen, Jr., H.L. Stormer, and J.C.M. Hwang, Phys. Rev. B **28**, 4875 (1983)
7 C. Sikorski and U. Merkt, Phys. Rev. Lett. **62**, 2164 (1989)
8 D. Heitmann and J.P. Kotthaus, Phys. Today, June, 56 (1993)
9 T.J. Thornton, M. Pepper, H. Ahmed, D. Andrews, and G.J. Davies, Phys. Rev. Lett. **56**, 1198 (1986)
10 T. Suzuki and T. Ando, J. Phys. Soc. Jpn. **62**, 2986 (1993)
11 K. F. Berggren, G. Roos, and H. van Houten, Phys. Rev. B **37**, 10118 (1988)
12 S.-N. Wang, K. Yamanaka, K. Hirakawa, M. Noguchi, and T. Ikoma, Jpn. J. Appl. Phys. **35**, pt. 2, L1251 (1996)
13 J.H. Davies, Semicond. Sci. Technol. **3**, 995 (1988)
14 M. Helm, E. Colas, P. England, F. DeRosa, and S.J. Allen, Jr., Appl. Phys. Lett. **53**, 1714 (1988)
15 K. Hirakawa, M. Grayson, D.C. Tsui, and C. Kurdak, Phys. Rev. B **47**, 16651 (1993)
16 D.C. Tsui, E. Gornik, and R. A. Logan, Solid State Commun. **35**, 875 (1980)
17 K. Hirakawa, K. Yamanaka, M. Grayson, and D.C. Tsui, Appl. Phys. Lett. **67**, 2326 (1995)
18 M. Grayson, D.C. Tsui, M. Shayegan, K. Hirakawa, R. A. Ghanbari, and H.I. Smith, Appl. Phys. Lett. **67**, 1564 (1995)
19 W. Kohn, Phys. Rev. **123**, 1242 (1961)
20 P.R. Smith, D.H. Auston, and M.C. Nuss, IEEE J. Quantum Electron. **JQE-24**, 255 (1988)
21 H.G. Roskos, M.C. Nuss, J. Shah, K. Leo, D.A.B. Miller, A.M. Fox, S. Schmitt-Rink, and K. Kohler, Phys. Rev. Lett. **68**, 2216 (1992)
22 C. Waschke, H.G. Roskos, R. Schwendler, K. Leo, H. Kurz, and K. Kohler, Phys. Rev. Lett. **70**, 3319 (1993)
23 K. Hirakawa, I. Wilke, K. Yamanaka, H.G. Roskos, M. Vosseburger, F. Wolter, C. Waschke, H. Kurz, M. Grayson, and D.C. Tsui, Surf. Sci. **361/362**, 368 (1996)

2.8 Wannier–Stark Effect in Transport

2.8.1 Wannier–Stark Effect

Since the first prediction by G.H. Wannier, the energy spectrum of a crystalline solid in an electric field, which is called a Wannier–Stark state or a Stark ladder, has attracted much interest. In 1960, Wannier studied electronic states in the presence of a uniform electric field theoretically and found that eigenstates are localized along the direction of the electric field and have quantized energy levels

$$E_n = neFd, \tag{2.8.1}$$

where n is an integer, F the electric field, and d the lattice period along the electric field [1]. The localized Stark ladder states are associated with Bloch oscillation. In an electric field $\mathbf{F}$, temporal motion of an electronic states is described by $\dot{\mathbf{k}} = e\mathbf{F}/\hbar$. At the edge of the Brillouin zone, electrons are Bragg reflected by the periodic crystal potential and their motion becomes periodic in space, as well as in the $\mathbf{k}$ space.

There have been a number of controversial arguments on the Stark ladder states. Zak claimed that such Stark ladder states cannot exist because they will decay rapidly due to Zener tunneling caused by interband mixing [2]. Some authors claimed that such interband effects are totally absent and no Zener tunneling is observable [3]. Others claimed that Zener tunneling is possible in Stark ladders [4,5].

A number of experimental challenges were carried out in order to observe a Stark ladder in bulk materials, but no clear evidence was found. Only weak oscillation of conductivity observed in ZnS was attributed to the hopping motion of electrons between Stark ladder states [6,7]. The failure of observation is attributed to scattering of electrons by impurities or phonons, which prevent the acceleration of electrons up to a Brillouin-zone edge.

Step like photocurrent spectra in a GaAs/AlAs superlattice in electric fields were observed in 1988 and considered to give evidence for Stark ladders [8]. In superlattices, electrons can be accelerated easily to the edge of the Brillouin zone before they are scattered because of a large lattice constant (accordingly small Brillouin zone). Since then, a number of optical measurements on Wannier–Stark effects have been performed [9–12] using various techniques such as photocurrent and electroreflectance (ER).

Figure 2.8.1a shows electroreflectance spectra observed in a GaAs(40Å)/ $Al_{0.3}Ga_{0.7}As$(20Å) superlattice [12]. The signals in the photon energy region 1.5∼1.8 eV shift in proportion to the applied voltage toward both lower and higher energies. This behavior arises from oblique transitions between states localized in adjacent quantum wells and clearly indicates formation of Stark ladder states. Complex signals are also observed in higher energy region (1.8–2.2 eV). Calculated energy levels in a tight-binding model are shown in Fig. 2.8.1b. By comparing experimental results with theoretical transition energies, the ER signals at lower energy region are well assigned to transitions from Stark ladder states of heavy and light holes to conduction states. The ER signals in the higher energy region 1.9–2.2 eV are attributed to transitions

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

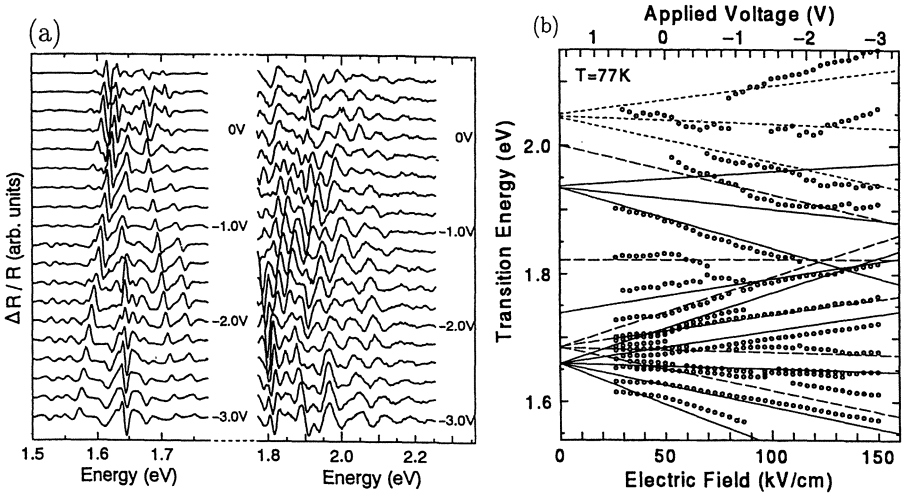


Fig. 2.8.1 (a) Electroreflectance spectra observed in a GaAs(40Å)/$Al_{0.3}Ga_{0.7}As$ (20Å) superlattice. Lower and higher shifts of critical energies proportional to applied voltage shown in the left part of the figure indicate oblique transitions between Stark ladder states in adjacent quantum wells. Complicated signals in higher energy region given in the right part of this figure also show shift. (**b**) Calculated transition energies between valence and conduction Stark ladder states. The solid, dashed, and dotted lines are transition energies from heavy-hole states, light-hole states, and spin-orbit split-off states, respectively. Experimental results are shown by open circles

from spin-orbit split-off (SO) band. The tight-binding calculation indicates that the SO states are localized in electric fields although they have energies larger than the AlGaAs barrier.

Stark ladders associated with the X valleys in AlAs barriers have been also observed [13–15]. Recently, a Bloch oscillation was also observed by using the four-wave mixing measurements [16,17]. Although optical properties of the Stark ladder states have been well understood, their transport properties remain still controversial.

2.8.2 Zener Tunneling and Wannier–Stark States

The Zener tunneling through a p-i-n diode under a reverse voltage has been studied in a tight-binding model [18,19]. The theory predicts that the current is modulated as a result of resonant coupling with localized Stark ladder states (the Wannier-Start resonance). In contrast to the conventional method based on the effective mass approximation, the tight-binding model properly accounts for interband mixing induced by an electric field between the valence and conduction bands. The tunneling current is calculated in this method for the theoretical analysis discussed below.

Figures 2.8.2a illustrates a Wannier–Stark resonance of the Zener tunneling. In a p-i-n diode with reverse bias, electrons in the valence band of n-region can tunnel to the conduction band of p-region through the forbidden band gap. The tunneling current is a smooth curve in weaker fields (the

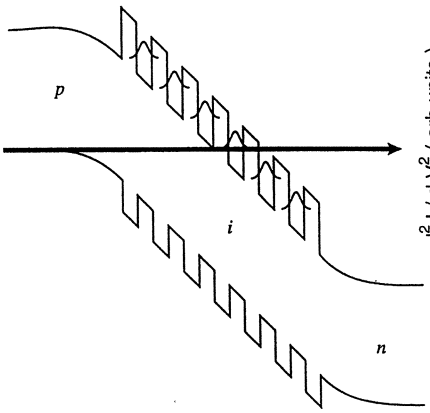


Fig. 2.8.2 Schematic diagram for mechanism of Wannier–Stark resonance of Zener tunneling. With increasing reverse voltage, Zener tunneling electrons and localized Stark ladder states crossover in energy and resonance between them takes place, resulting in modulation of Zener current

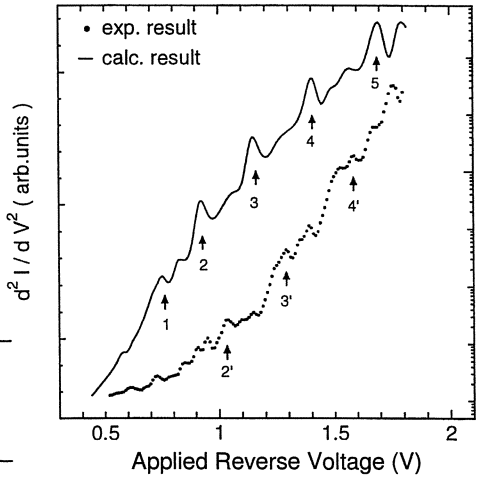


Fig. 2.8.3 Measured second derivative of Zener current from p-i-n diode with 10 periods of $(GaAs)_5/(AlAs)_2$ superlattice in the intrinsic region is shown by a dotted curve. Calculated curve for the corresponding structure is shown by a solid line After [20–23]

so-called Zener regime). As the reverse voltage increases, tunneling electrons and localized Stark ladder states crossover in energy and resonance takes place between them. In this Stark regime, the number of localized states relevant to the resonance increases with the voltage. As a result, the Zener current is periodically modulated. The modulation is quite small because the resonance is masked with the larger amount of the direct Zener tunneling current and can be seen only in the derivative of current-voltage characteristics.

2.8.3 Measurements of Zener Current through a p-i-n Diode

Measurements of the Zener current were carried out at 4.2 K for a p-i-n diode whose intrinsic region consists of a thin $(GaAs)_5/(AlAs)_2$ superlattice [20–23]. The n- and p-regions are doped with 5×10^{18} cm^{-3}, and 30 Å non-doped buffer layers are put between the superlattice and the doped regions.

The second derivative of the *I-V* characteristics was measured with the standard lock-in technique by picking up the second harmonic component of Zener current with a lock-in amplifier. The frequency and amplitude of the modulation voltage are typically 1 kHz and 20 mV, respectively. In Fig. 2.8.3, a result of the measurements is shown with a curve calculated for the corresponding structure, where internal voltage due to the p-n junction is assumed as 1.52 V. Over-all features of both curves are similar with several peaks from 0.7~1.7 V. Figure 2.8.4 shows ER spectra of the sample in reverse voltages. As the $(GaAs)_5/(AlAs)_2$ superlattice is so-called type II, ER signals arising from the X states of AlAs barriers as well as the Γ states in GaAs well are observed. Shifts of peaks proportional to the applied reverse voltage toward

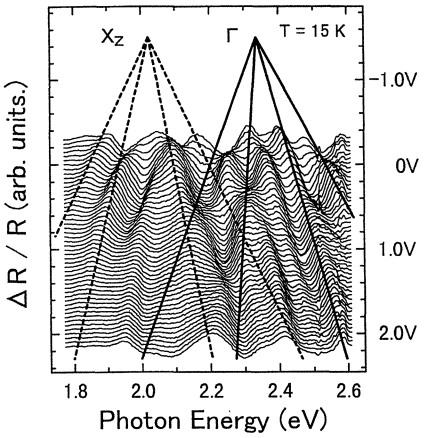


Fig. 2.8.4 Electroreflectance spectra from the p-i-n diode with 10 periods of $(GaAs)_5/(AlAs)_2$ superlattice. Transition related to the X- and Γ- Stark ladder states are observed. Higher and lower shifts of peaks proportional to the applied voltage indicate formation of Stark ladders in the intrinsic region of the superlattice

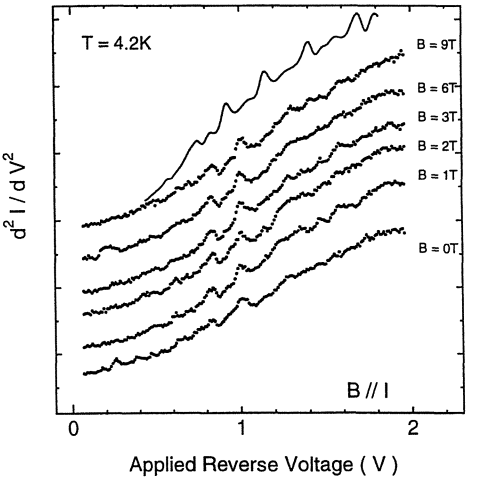


Fig. 2.8.5 Measured second derivative of Zener current from the p-i-n diode with 10 periods of $(GaAs)_5/(AlAs)_2$ superlattice in magnetic fields parallel to the direction of Zener current up to 9 T

lower and higher energies clearly indicate that Stark ladders are formed in the intrinsic region. Electric fields estimated from the peak spacing of the ER spectra well correspond to those of the calculated curve. There remain slight disagreement between the experiments and the theory, which may originate from various causes including diffusion of dopant atoms and interface roughness.

The second derivative curves measured in magnetic fields parallel to the direction of Zener current up to 9 T are shown in Fig. 2.8.5. Although the change of the curves due to magnetic fields is not significant, peaks become slightly sharp in magnetic field. This may be explained by one-dimensional character of tunneling states in the magnetic fields. Further experimental and theoretical studies are necessary to understand fully the properties of the Wannier–Stark states in transport phenomena.

References

1 G.H. Wannier, Phys. Rev. **117**, 432 (1960)
2 J. Zak, Phys. Rev. Lett. **20**, 1477 (1968)
3 D. Emin and C.F. Hart, Phys. Rev. B **36**, 7353 (1987)
4 P.N. Argyres and S. Sfiat, J. Phys. C: Condens. Matter. **2**, 7089 (1990)
5 J. Leo and A. MacKinnon, J. Phys. C: Condens. Matter. **1**, 1449 (1989)
6 S. Maekawa, Phys. Rev. Lett. **24**, 1175 (1970)
7 M. Saitoh, J. Phys. C: Solid State Phys. **5**, 914(1972)
8 F. Agulló -Rueda, E.E. Mendez, and J.M. Hong, Phys. Rev Lett. **40**, 1357 (1989)
9 K. Fujiwara, Jpn. J. Appl. Phys. **28**, L1718 (1989)

10 M. Nakayama, I. Tanaka, H. Nishimura, K. Kawashima, and K. Fujiwara, Phys. Rev. B **44**, 5935 (1991)
11 H. Schneider, A. Fischer, and K. Ploog, Phys. Rev. B **45**, 6329 (1992)
12 C. Hamaguchi, M. Yamaguchi, M. Morifuji, H. Kubo, K. Taniguchi, C. Gmachl, and E. Gornik, Semicond. Sci. & Technol. **9**, 1994 (1994)
13 A.J. Shields, P.C. Klipstein, M.S. Skolnick, G.W. Smith and C.R. Whitehouse, Phys. Rev. B **42**, 5879 (1990)
14 M. Yamaguchi, H. Nagasawa, M. Morifuji, H. Kubo, K. Taniguchi, C. Hamaguchi, C. Gmachl, and E. Gornik, Semicond. Sci. & Technol. **9**, 1810 (1994)
15 M. Morifuji, M. Yamaguchi, K. Taniguchi, and C. Hamaguchi, Phys. Rev. B **50**, 8722 (1994)
16 J. Feldmann, K. Leo, J. Shah, D.A.B. Miller, J.E. Cunningham, T. Meier, G. von Plessen, P. Thomas, and S. Schmitt-Rink, Phys. Rev. Lett. **46**, 7252 (1992)
17 T. Dekorsy, R. Ott, H. Kurz, and K. Köhler, Phys. Rev. B **51**, 17275 (1995)
18 A. Di Carlo, P. Vogl, and W. Pötz, Phys. Rev. B**50**, 8358 (1994)
19 J.N. Schulman and Y-C.Chang, Phys. Rev. B**27**, 2346 (1983)
20 C. Hamaguchi, M. Yamaguchi, H. Nagasawa, M. Morifuji, A. Di Carlo, P. Vogl, G. Böhm, G. Tränkle, G. Weimann, Y. Nishikawa, and S. Muto, Jpn. J. Appl. Phys. **34**, 4519 (1995)
21 H. Nagasawa, K. Murayama, M. Yamaguchi, M. Morifuji, C. Hamaguchi, A. Di Carlo, P. Vogl, G. Böhm, G. Tränkle, and G. Weimann, Solid State Electron. **40**, 2346 (1996)
22 H. Nagasawa, K. Murayama, M. Morifuji, A. Di Carlo, P. Vogl, G. Böhm, G. Tränkle, G. Weimann, and C. Hamaguchi, Physica B **227**, 206 (1996)
23 K. Murayama, H. Nagasawa, S. Ozaki, M. Morifuji, C. Hamaguchi, A. Di Carlo, P. Vogl, G. Böhm, and G. Weimann, Superlattices and Microstructures, **20**, 493 (1996)

3. Quantum Hall Effect

3.1 Crossover from Quantum to Classical Regime

The quantum Hall effect is a manifestation of quantum effects in macroscopic transport phenomena. In particular, the integer quantum Hall effect has provided various fundamental problems in solid state physics such as localization in high magnetic fields, a topological winding number, edge channels, etc. It has been used also as a resistance standard and has opened up a way to examine the validity of quantum electrodynamics through comparison of the fine structure constant.

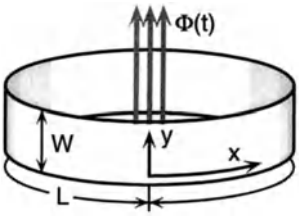
Two different pictures have been proposed for the explanation of the integer quantum Hall effect. In the bulk-current picture, the Hall electric field is essential and the current is carried by extended states existing in the vicinity of the center of each broadened bulk Landau level. In the edge-current picture, on the other hand, the current is carried by edge states localized along system boundaries and the Hall field is not required. The relationship between these two pictures remains still unclear. In this section, a brief review is given of some of theoretical work performed for the purpose of shedding some light on this problem.

3.1.1 Bulk Versus Edge Current Picture

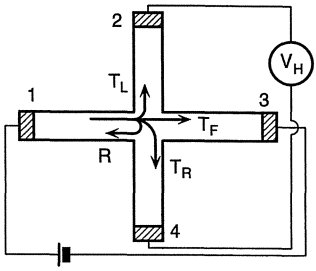
In the bulk-current picture, a key ingredient of the quantum Hall effect is the Anderson localization of Landau states in the presence of random potential fluctuations. It has been shown in various different ways that the Hall conductivity is quantized into integer multiples of $-e^2/h$ when states around the Fermi level are all localized. Several reviews have been published already [1–3].

Let us consider a two-dimensional system on a cylinder surface as shown in Fig. 3.1.1. The circumference is chosen in the x direction and its length is L. The cylinder height or the system size in the y direction is given by W. A time-dependent magnetic flux $\Phi(t)$ passing along the cylinder axis gives an electric field in the x direction $E_x = -(1/cL)(\partial\Phi(t)/\partial t)$. For localized states, effects of the magnetic flux can be fully incorporated as a change in the phase of their localized wave function. When the Fermi level lies in the energy gap of extended states, the system changes periodically as a function of the flux with period Φ_0 as long as the flux is varied sufficiently slowly, where $\Phi_0 = ch/e$ is the magnetic flux quantum. The only change allowed between Φ and $\Phi+\Phi_0$ is a transfer of some electrons through boundaries with a reservoir. Let n be

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998



3.1.1 A two-dimensional system on a cylinder surface. An electric field E_x is induced in the x direction along the circumference (length L_x) by a time-dependent flux $\Phi(t)$ passing along the cylinder axis. The boundaries $y=0$ and $y=L_y$ are connected to a reservoir or an infinitely high barrier depending on the geometry



3.1.2 A four-terminal geometry to measure the Hall effect

the number of electrons going from edge $y=0$ to $y=W$. Then the current density in the y direction is given by

$$j_y = (-e)n\frac{1}{L}\frac{\partial\Phi(t)}{\partial t}\frac{1}{\Phi_0} = \frac{e^2}{h}\,n\,E_x, \tag{3.1.1}$$

which leads to the quantized Hall conductivity. This is essentially the same as the original argument of Laughlin [4].

A more mathematical explanation of the quantized Hall conductivity is based on a topological winding number [5–7]. In particular, it can be shown that the Kubo formula for the Hall conductivity in a two-dimensional system on a torus is written as a sum of topological winding numbers corresponding to eigenstates after being averaged over Aharonov–Bohm magnetic fluxes through the openings of the torus [8,9].

It should be noticed that the current j_y given by (3.1.1) is the average of the Hall current over the period Φ_0. A current $j_y(\Phi)$ at certain Φ is not necessarily quantized and can vary as a function of Φ. This is true also for the Hall conductivity in the torus geometry. It becomes quantized after being averaged over different boundary conditions or fictitious Aharonov–Bohm fluxes. In the limit of infinitely large systems, the boundary conditions should not play a role and the quantity for a given Φ is expected to approach the average. As discussed in the following, the Hall conductivity for a system with a finite width does depend on a fictitious flux, however, and quantities measured in actual experiments should rather be regarded as those averaged over the flux or boundary conditions.

There have been various theoretical and experimental investigations on the localization in the quantum Hall regime. It is now well established that states are all localized exponentially except those just at the center of each broadened Landau level. The inverse localization length exhibits a power-law energy dependence away from the center, i.e., $\alpha(E) \propto |E|^s$, with s close to 2.3 for the lowest Landau level, where energy E is measured from the center of the Landau level. There remain some discussions on the universality on

different Landau levels, on nature of dominant scatterers, on the strength of the magnetic field, etc., but these are beyond the scope of this section.

The quantized Hall effect can also be explained by the transmission approach based on the Büttiker–Landauer formula [10,11]

$$I_i = \frac{e^2}{\pi\hbar}\Big[(1-R_{ii})V_i - \sum_{j\neq i} T_{ij}V_j\Big], \tag{3.1.2}$$

where I_i is the current flowing into the system through terminal i, R_{ii} is the reflection probability for an electron flowing into terminal i, T_{ij} is the transmission probability from terminal j to i, and V_i is the voltage of terminal i. In fact, if we consider a four-terminal geometry used in Hall-effect experiments shown in Fig. 3.1.2 and assume that $T_L=1$, $R=T_F=T_R=0$, then we immediately obtain the quantized Hall resistance $R_H=h/2e^2$, where R, T_L, T_F, and T_R are transmission probabilities for being reflected back, turning left, going straight, and turning right at the junction of two wires.

In a two-terminal geometry, the conductance of a wire is given by Landauer's formula [10],

$$G = \frac{e^2}{\pi\hbar}\sum_{\mu\nu} T_{\mu\nu}, \tag{3.1.3}$$

where G is the conductance, ν is the index for incident channels, and μ is that for outward channels. The channels correspond to edge states localized at a system boundary in the quantum Hall regime. This transmission probability is related to the diagonal conductivity σ_{xx} of the wire [12–15] and does not depend on the Hall conductivity σ_{xy}. This is quite in contrast to the classical galvanomagnetic effect in which σ_{xy} plays a decisive role.

Consider a wire with width W in the y direction and infinitely long in the x direction. The usual textbook procedure is to start with the transport relations, $j_x=\sigma_{xx}E_x+\sigma_{xy}E_y$ and $j_y=\sigma_{yx}E_x+\sigma_{yy}E_y$, and impose the condition $j_y=0$, where $\mathbf{j}=(j_x,j_y)$ is the current density and $\mathbf{E}=(E_x,E_y)$ is the electric field. Then, we have $E_y=-(\sigma_{yx}/\sigma_{yy})E_x$ and $j_x=(\sigma_{xx}-\sigma_{xy}\sigma_{yx}/\sigma_{yy})E_x$. In high magnetic fields, the diagonal elements of the conductivity tensor are much smaller than the off-diagonal ones and the total current is dominated by the Hall current determined by the Hall field E_y.

A quantum analogue of the classical procedure is to consider the linear relation:

$$\begin{pmatrix} j_x(y) \\ j_y(y) \end{pmatrix} = \int dy' \begin{pmatrix} \tilde{\sigma}_{xx}(y,y') & \tilde{\sigma}_{xy}(y,y') \\ \tilde{\sigma}_{yx}(y,y') & \tilde{\sigma}_{yy}(y,y') \end{pmatrix} \begin{pmatrix} E_x(y') \\ E_y(y') \end{pmatrix}, \tag{3.1.4}$$

with nonlocal conductivity tensor $\tilde{\sigma}_{\mu\nu}$. Unfortunately, this procedure does not work because we usually get vanishing $\tilde{\sigma}_{yy}$ and nonvanishing $\tilde{\sigma}_{yx}$ for wires with finite width. This leads to the conclusion that the current flows in the y direction perpendicular to the wire direction.

Let us consider the two-dimensional system on a cylinder surface as shown in Fig. 3.1.1. The Aharonov–Bohm flux Φ changes the center coordinate of

the cyclotron motion by $(2\pi l^2/L)(\Phi/\Phi_0)$, thus leading to a shift in the average electron position in the y direction, where l is the magnetic length. The only exceptions are the case that all the states below the Fermi level are localized for which the flux is absorbed fully by a phase change of the wave function. The current calculated in the linear response is nothing but the time derivative of this average position and therefore need not vanish at each value of Φ.

As discussed previously, the system comes back to the same state after a change of the flux by the flux quantum Φ_0. In the system with a finite width, the average y position oscillates with period Φ_0 and the net current in the y direction vanishes after a time average over the period corresponding to flux change from Φ to $\Phi+\Phi_0$, leading to a vanishing average current [16].

The above suggests that the current calculated in the conventional linear response theory does not represent an observable quantity and that of a flux average should be regarded as a more appropriate quantity corresponding to experimental situations. The flux average of $\bar{j}_y$ vanishes identically and also leads to the conclusion that the flux average of $\bar{j}_x$ is determined by σ_{xx} alone as Landauer's formula, where $\bar{j}_x$ and $\bar{j}_y$ are the average of j_x and j_y in the y direction.

Note that the average over flux should be taken under the condition that the total number of electrons in the system remains the same. If the chemical potential is fixed instead of the electron number, the averaged Hall conductivity becomes nonzero. It is usually not quantized but slowly approaches the quantized value with increasing width [17,18]. Presumably, the fixed chemical potential implicitly assumes electron flows between the system and reservoirs across the boundaries $y=0$ and $y=W$.

Note also that the above does not necessarily mean that the current is carried only by edge states in actual systems. It may just show that the total current is independent of the Hall field and/or the Hall conductivity σ_{xy} in quantum wires. Note also that the Hall polarization field due to charges accumulated spatially [19–26] has not been included in the above discussion. This polarization field does not influence the total Hall current although it may strongly modify the local current distribution.

Many different experiments reported so far have demonstrated the importance of the edge current in the integer quantum Hall effect. In particular, a nonequilibrium distribution of current among edge and bulk states is known to give rise to a huge nonlocal resistance [27]. This subject is discussed in Sect. 3.2. A recent study of the width dependence of the breakdown of the quantum Hall effect shows that a major part of the Hall current is carried by bulk states rather than edge states [28].

3.1.2 Edge Transport and Bulk States

We consider both two- and four-terminal geometries, schematically illustrated in Fig. 3.1.3. For the two-terminal geometry, a wire with width W in the y direction and length L in the x direction is connected to an infinitely long ideal wire which is eventually connected to a reservoir. For the four-terminal geometry, we shall consider a junction of two wires with length L and width

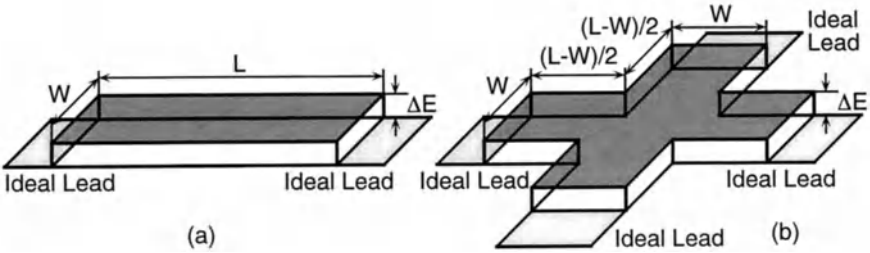


Fig. 3.1.3 Schematic illustration of a quantum wire (a) and its crossing (b) with width W and length L. The wire is confined by infinitely high barriers in one direction and connected to an infinitely long ideal wire which is eventually connected to a reservoir. The bottom energy is raised by $\Delta E = \hbar\omega_c/2$ from that in the ideal wire

W. Ideal wires are connected at all ends.

The strength of scatterers with short-range delta-potentials is characterized by broadening Γ of the Landau level, which is chosen as $\Gamma/\hbar\omega_c = 0.2$ with the cyclotron frequency ω_c. The transmission probabilities are calculated using recursive Green's function technique based on a lattice model [29] and the conductance can easily be given by the Büttiker–Landauer formula [10,11]. Numerical calculations are performed for $L/W = 1/2$, 1, and 2 in the two-terminal case and for $L/W = 3$ in the crossed-wires geometry.

Figure 3.1.4 gives some examples of calculated energy dependence of the averaged conductance given by $\langle G\rangle = (e^2/\pi\hbar)\langle T\rangle$ for the lowest Landau level, where $\langle\cdots\rangle$ denotes a sample average and T is the transmission probability. The dotted lines represent the transmission probability for square systems with $L/W = 1$ in the absence of randomness. It is a step function rising at the energy of the Landau level in two dimensions, which is slightly smaller than $\hbar\omega_c/2$ because of the small nonparabolicity present in the present lattice model. The energy dependence of the transition probability in the presence of randomness becomes steeper with increasing system size. The effective energy width over which the transmission probability increases from zero to unity is estimated as the inverse of the maximum derivative $\Gamma(d\langle T\rangle/dE)_{\max}$. The threshold energy $E_{\rm th}$ is obtained as $\langle T(E_{\rm th})\rangle = 1/2$.

Figure 3.1.5 gives the maximum derivative as a function of the width in logarithmic scales. It is clear that the results can be fitted to a straight line, except those for $W/l = 5$, showing that $\Gamma(d\langle T\rangle/dE)_{\max} \propto W^{1/s}$. The exponent is almost independent of the form of the system and is obtained as $s = 2.2 \pm 0.1$. This exponent is consistent with that of almost all the existing numerical studies because actual errors can be larger than those estimated statistically.

Figure 3.1.4 contains also the conductance fluctuation $\sqrt{\langle (G - \langle G\rangle)^2\rangle}$. The fluctuation is a universal function of the average conductance and the distribution function of the conductance is also universal.

In the four-terminal geometry we calculate reflection (R) and transmission probabilities, turning right (T_R), turning left (T_L), and going straight (T_F), as a function of energy for the lowest Landau level. Figure 3.1.6 shows an

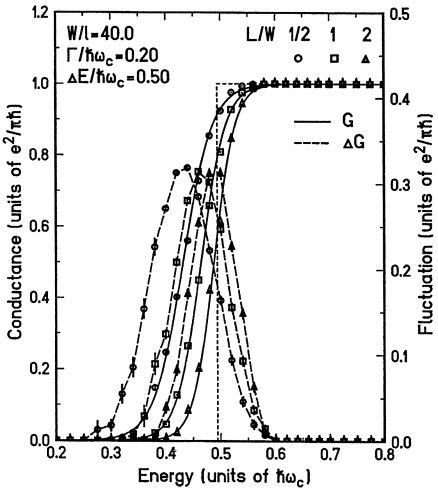


Fig. 3.1.4 Calculated conductance and its fluctuation (the open symbol) as a function of energy for the lowest Landau level. The circles denote the results for a short rectangular system ($L/W = 1/2$), squares those for a square ($L/W = 1$), and triangles those for a long rectangular system ($L/W = 2$). The dotted line represents the conductance in the absence of randomness. $W/l = 40$

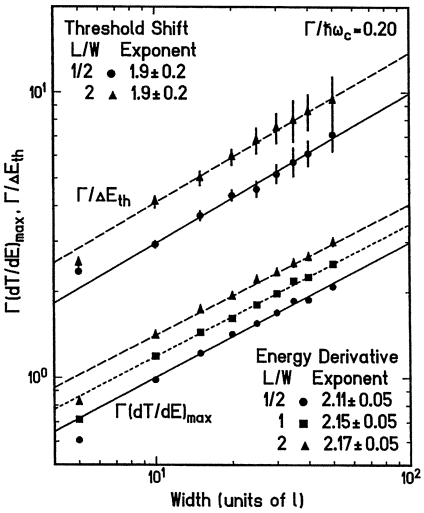


Fig. 3.1.5 Log-log plot of the maximum energy derivative of the two-terminal conductance as a function of the wire width. The straight lines represent $\Gamma(d\langle T\rangle/dE)_{\max} \propto W^{1/s}$ with exponent s listed at the bottom-right corner. The dots denoted as $\Gamma/\Delta E_{\rm th}$ represent the threshold energy difference between $L/W = 1$ and $1/2$, and the squares the difference between $L/W = 2$ and 1. The exponents defined by $\Delta E_{\rm th} \propto W^{-1/s}$ are listed at the top-left corner

example for $W/l = 40$. With increasing wire width, the transition region, where $\langle R\rangle$ and $\langle T_L\rangle$ change between zero and unity and where $\langle T_R\rangle$ and $\langle T_F\rangle$ are nonzero, becomes narrower. The effective width of this transition region can be deduced also from $-\Gamma(d\langle R\rangle/dE)_{\max}$ and $\Gamma(d\langle T_L\rangle/dE)_{\max}$.

The critical exponent can be calculated also from the dependence of the maximum derivative $-\Gamma(d\langle R\rangle/dE)_{\max}$ and $\Gamma(d\langle T_L\rangle/dE)_{\max}$ on W, and further from $\Delta E_{\rm th}^{TR} = E_{\rm th}^{T} - E_{\rm th}^{R}$, where $E_{\rm th}^{T}$ and $E_{\rm th}^{R}$ are the threshold for T_L and R, respectively. Within the numerical accuracy, the obtained exponents have turned out to agree with those obtained in different methods.

Figure 3.1.7 gives the average transmission probabilities as a function of the average reflection probability for systems with different widths. It is evident that the average transmission probabilities are determined by $\langle R\rangle$ alone within the numerical accuracy. The universality prevails also for the fluctuations and correlation functions although not explicitly shown here. This is consistent with a scaling relation between σ_{xx} and σ_{xy} obtained numerically for systems without a boundary [30]. This universality suggests that $\alpha(E)$ is the only relevant parameter in the present system.

Various experiments have been reported on the width of the transition region where the Hall conductivity changes from an integer to another. In

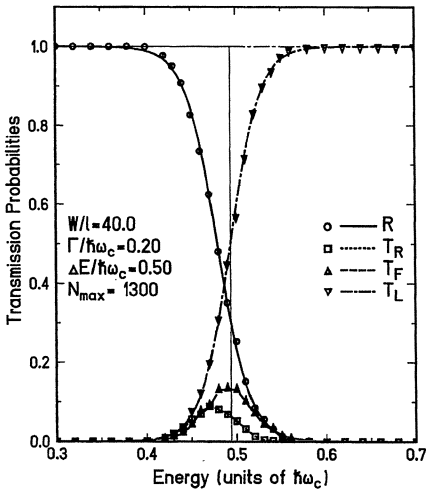


Fig. 3.1.6 Average reflection and transmission probabilities at wire junctions with width $W/l = 40$

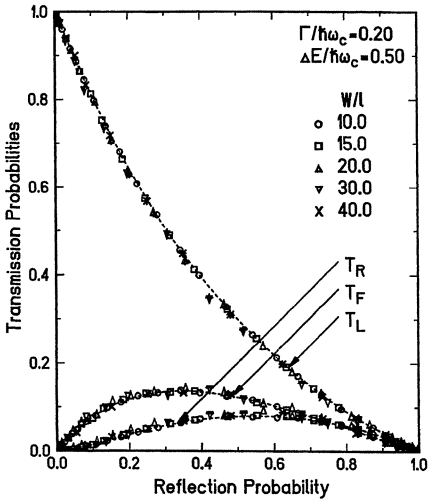


Fig. 3.1.7 Average transmission probabilities versus average reflection probability at wire junctions with different width. The dotted lines are guides to the eye

particular, Koch et al. [31] measured $(d\rho_{xy}/dB)_{\max}$ as a function of temperature for wires fabricated at GaAs/AlGaAs heterostructures with width lying between 9 and 64 μm. If we assume that $(d\rho_{xy}/dB)_{\max}$ is determined by $\Gamma(d\langle T_L\rangle/dE)_{\max}$ calculated above, we get a value for $(d\rho_{xy}/dB)_{\max}$ about half of that observed experimentally. One possible reason for this slight disagreement lies in the difference in the effective range of dominant scatterers. In GaAs/AlGaAs heterostructures, scatterers are known to have a long-range potential instead of the short range assumed above. In fact, a numerical study has demonstrated that the localization becomes stronger for long-range scatterers though the exponent itself remains the same [32]. Another possibility is the difference between $(d\rho_{xy}/dB)_{\max}$ and $\Gamma(d\langle T_L\rangle/dE)_{\max}$. The Hall resistance is written as $R_H = (h/2e^2)(T_L - T_R)/[(T_L - T_R)^2 + (T_L + T_R + 2t_F)^2]$ and has a strong tendency to rise much more quickly than T_L as a function of energy.

3.1.3 Voltage Distribution

To introduce inelastic scattering and calculate the voltage distribution in the system, we shall extend Büttiker's method [33–36]. The system is schematically illustrated in Fig. 3.1.8, in which a wire with a single current channel is attached and the wire is connected to a reservoir with a certain voltage. The voltage is determined by the condition that the net current passing through the wire vanishes. This voltage probe introduces inelastic scattering which can be adjusted freely by changing in the coupling of the wire to the system. The coupling is characterized by the imaginary part $-i\hbar/2\tau_\phi$ in the diagonal site energy with inelastic scattering time, which is related to the phase coherence length L_ϕ through $L_\phi = \sqrt{2\tau_\phi D^*}$ with the diffusion coefficient D^* given

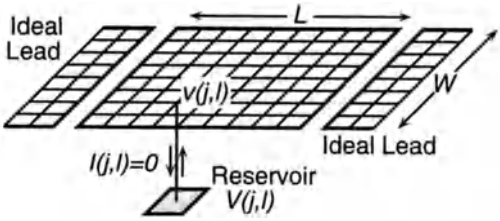


Fig. 3.1.8 A schematic illustration of a quantum wire in the presence of inelastic scattering. Each lattice point is connected to a reservoir through an ideal wire characterized by the transfer integral $-\hbar/2\tau_\phi$. The voltage of each reservoir is determined in such a way that the current along the wire should vanish

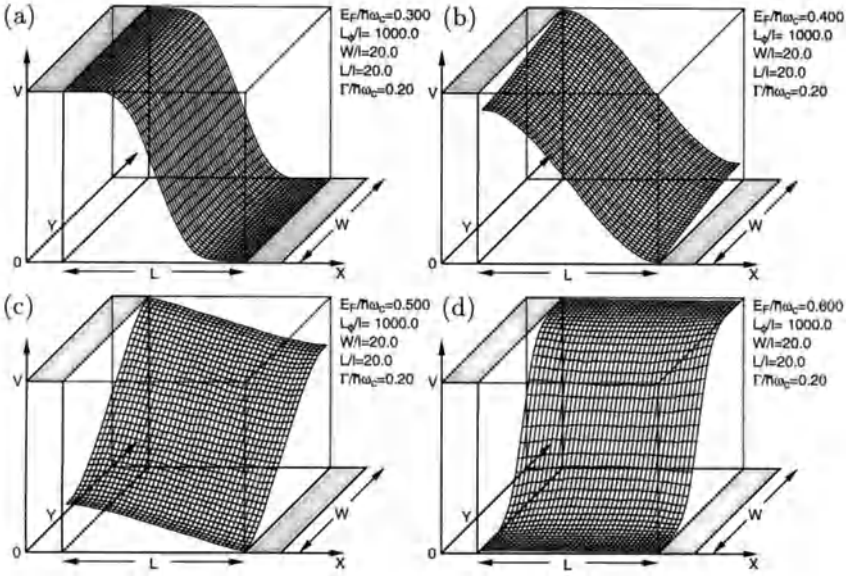


Fig. 3.1.9 Calculated voltage distribution in the case of weak inelastic scattering

by $l^2\Gamma/\hbar$.

Numerical calculations have been performed for the lowest Landau level and for $\Gamma/\hbar\omega_c = 0.2$ [37]. Figure 3.1.9 shows some examples of the results for $L/l = W/l = 20$ and $L_\phi/l = 1000$ and Fig. 3.1.10 shows those for $L_\phi/l = 5$. In the former case the phase coherence length is larger than the system size, while it is smaller in the latter case.

When the energy E lies near the center of the Landau level, i.e., (b) and (c) in Figs. 3.1.9 and 3.1.10, the voltage distribution is independent of the strength of inelastic scattering. The voltage distribution is approximated well by a flat plane determined classically by the local conductivity tensor σ_{xx} and σ_{xy} except in the region close to the system boundaries ($LE_x + WE_y = V$ and $|\sigma_{xy}|E_x + \sigma_{xx}E_y = 0$ where V is the voltage drop between left and right reservoirs and E_x and E_y are fields in the x and y directions, respectively).

The situation changes drastically when the Fermi level moves away from

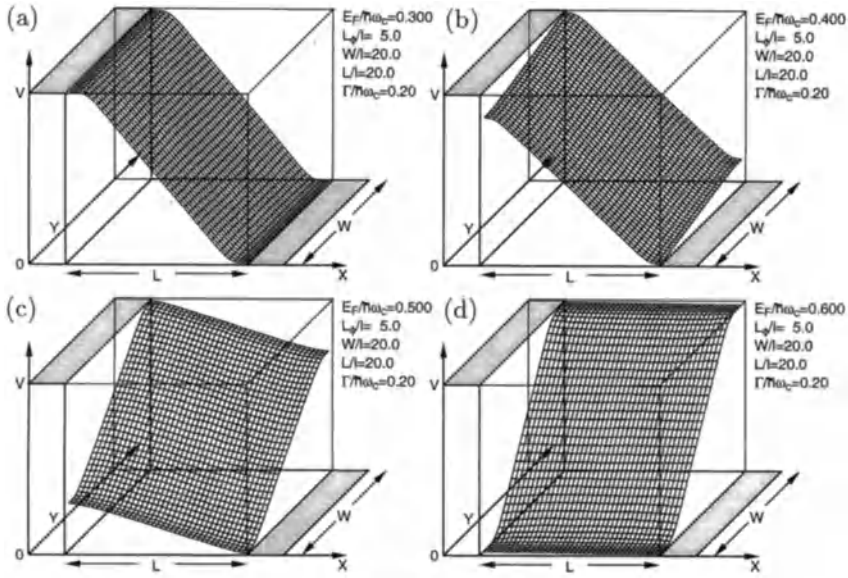


Fig. 3.1.10 Calculated voltage distribution in the case of strong inelastic scattering

the center. When E lies in the region corresponding to the high-energy tail of the broadened Landau level, (d) in Figs. 3.1.9 and 3.1.10, the voltage exhibits a step-function-like change in the y direction for $L_\phi \gg L, W$, while the voltage decays linearly in the y direction for $L_\phi \ll L, W$. The same is applicable in the case that E lies in the region of the low-energy tail, i.e., (a) in Figs. 3.1.9 and 3.1.10, except that the voltage change now occurs in the x direction.

Figure 3.1.11 shows the effective field at the center $(L/2, W/2)$ of the system for different values of L_ϕ/l. At the center of the broadened Landau level, $E/\hbar\omega_c \sim 0.46$ (shifted from $(1/2)\hbar\omega_c$ to the low energy side due to the nonparabolicity of the lattice model and due to level repulsion from higher Landau levels), the field in the x and y directions becomes independent of the strength of inelastic scattering. This again shows that the field distribution at the Landau-level center is the same as that determined classically by the local conductivity tensor σ_{xx} and σ_{xy}. Note that the sum $E_x + E_y$ is slightly larger than the classical value V/L in the present $L = W$ case.

Effects of inelastic scattering on the voltage distribution become stronger for states with energy away from the center. Furthermore, they become more appreciable with increasing system size even for a fixed value of L_ϕ as shown in Fig. 3.1.12. In this figure the decrease of the field for states with energy away from the center is quite prominent, while that of the field for states near the Landau-level center of the Landau level is roughly explained by the classical field. The sum $E_x + E_y$ approaches that of the classical field independent of energy with increasing system size for a fixed L_ϕ.

The width of the region where the voltage distribution changes is determined by the localization length in the absence of inelastic scattering. When

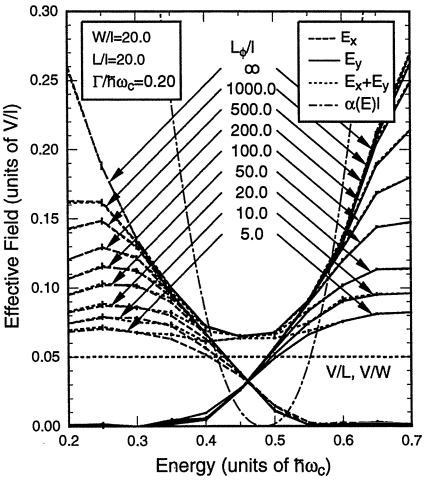


Fig. 3.1.11 Calculated effective field at the center of the wire. $W/l = L/l = 30$

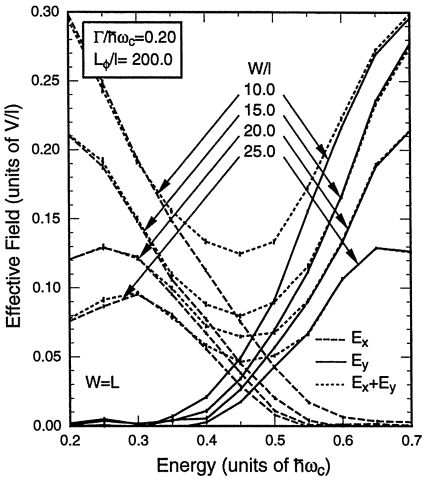


Fig. 3.1.12 Calculated effective field at the center of the wire. $L_\phi/l = 200$

inelastic scattering is introduced, the voltage distribution in the central region becomes linear. The distance from the boundary where the distribution starts to deviate from this linear dependence is determined by the condition that the escape rate of an electron through boundaries becomes comparable to τ_ϕ^{-1}. Therefore, the width of this "boundary region" is determined only by the localization length and τ_ϕ. The crossover occurs when the system width exceeds the width of the boundary region.

Note that the localization length determined from Fig. 3.1.13 by an extrapolation $L_\phi/l \to \infty$ is larger than the known localization length in two dimensions [38–39]. Although the localization length of a two-dimensional system without boundaries is a single relevant length scale [40], the actual localization length can be modified in the presence of confinement potential leading to edge states.

3.1.4 Summary

A review has been given of two pictures for understanding the integer quantum Hall effect: the edge current picture and the bulk current picture. A numerical study based on the transmission approach performed for the purpose of understanding the relation between the two pictures has also been reviewed briefly. The results all seem to be consistent with those predicted in the bulk-current picture, including the critical exponent of localization of bulk states and the universal behavior of transport coefficients.

Effects of inelastic scattering have also been discussed. The calculated voltage distribution shows a clear crossover from the quantum case where the phase coherence length is much larger than the system size to the classical case where the phase coherence length is much smaller and the conventional description based on local σ_{xx} and σ_{xy} is valid.

It should be stressed that one important ingredient of the quantum Hall effect lies in its insensitivity to various parameters. In fact, high-accuracy measurements are usually made in the presence of a considerable current for which the linear response may no longer be valid. At present we are still far from a real understanding of the quantum Hall effect.

References

1 T. Ando, Prog. Theor. Phys. Suppl. **84**, 69 (1985)
2 *The Quantum Hall Effect*, edited by R.E. Prange and S.M. Girvin (Springer, New York, 1987)
3 H. Aoki, Rev. Progr. Phys. **50**, 655 (1987)
4 R.B. Laughlin, Phys. Rev. B **23**, 5632 (1981)
5 D.J. Thouless, M. Kohmoto, M.P. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982)
6 J. Avron, R. Seiler, and B. Simon, Phys. Rev. Lett. **51**, 51 (1983)
7 B. Simon, Phys. Rev. Lett. **51**, 2167 (1983)
8 Q. Niu, D.J. Thouless, and Y.-S. Wu, Phys. Rev. B **31**, 3372 (1985)
9 H. Aoki and T. Ando, Phys. Rev. Lett. **57**, 3093 (1986)
10 R. Landauer, IBM J. Res. Dev. **1**, 223 (1957); Phil. Mag. **21**, 863 (1970)
11 M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986)
12 E.N. Economou and C.M. Soukoulis, Phys. Rev. Lett. **46**, 618 (1981)
13 D.S. Fisher and P.A. Lee, Phys. Rev. B **23**, 6851 (1981)
14 D.C. Langreth and E. Abrahams, Phys. Rev. B **24**, 2978 (1981)
15 D.J. Thouless, Phys. Rev. Lett. **47**, 972 (1981)
16 T. Ando and H. Aoki, Physica B **184**, 365 (1983)
17 J. Hajdu, M. Janßen, and O. Viehweger, Z. Phys. B **66**, 433 (1987)
18 T. Ohtsuki and Y. Ono, Solid State Commun. **68**, 787 (1988)
19 A.H. MacDonald, T.M. Rice, and W.F. Brinkman, Phys. Rev. B **28**, 3648 (1983)
20 J. Riess, J. Phys. C **17**, L847 (1984)
21 D.J. Thouless, J. Phys. C **18**, 6211 (1985)
22 O. Heinonen and P.L. Taylor, Phys. Rev. B **32**, 633 (1985)
23 D. Pfannkuche and J. Hajdu, Phys. Rev. B **46**, 7032 (1992)
24 D.J. Thouless, Phys. Rev. Lett. **71**, 1879 (1993)
25 C. Wexler and D.J. Thouless, Phys. Rev. B **49**, 4815 (1994)
26 H. Hirai and S. Komiyama, Phys. Rev. B **49**, 14012 (1994)
27 See, for example, S. Komiyama and H. Nii, Physica B **184**, 7 (1993) and references cited therein
28 S. Kawaji, K. Hirakawa, and M. Nagata, Physica B **184**, 17 (1993)
29 T. Ando, Phys. Rev. B **44**, 8017 (1991)
30 T. Ando, J. Phys. Soc. Jpn. **35**, 3199 (1986)
31 S. Koch, R.J. Haug, K. von Klitzing, and K. Ploog, Phys. Rev. Lett. **67**, 883 (1991)
32 T. Ando and H. Aoki, J. Phys. Soc. Jpn. **54**, 2238 (1985)
33 M. Büttiker, Phys. Rev. B **32**, 1846 (1985)
34 M. Büttiker, Phys. Rev. B **33**, 3020 (1986)
35 M. Büttiker, IBM J. Res. Develop. **32**, 63 (1988)
36 J.L. D'Amato and H.M. Pastawski, Phys. Rev. B **41**, 7411 (1990)
37 T. Ando, Surf. Sci. **361/362**, 270 (1996)
38 T. Ando, J. Phys. Soc. Jpn. **52**, 1893 (1983); **53**, 310 (1983); **53**, 3126 (1983); Phys. Rev. B **40**, 9965 (1989)
39 B. Huckestein and B. Kramer, Phys. Rev. Lett. **64**, 1437 (1990); Solid State Commun. **71**, 445 (1989)
40 T. Ando, Phys. Rev. B **49**, 4679 (1994)

3.2 Edge States and Nonlocal Effects

3.2.1 What Is Edge Current?

Various phenomena have been successfully analyzed by an edge-channel formulation of integer quantum Hall effect (IQHE) [1–3], which describes the conduction only in terms of the "edge current" that flows along the edge of a two-dimensional electron gas (2DEG) layer. On the other hand, a number of experiments strongly suggest that the conduction is dominated by "bulk Hall current" [4–8]. Consideration is given to an implication of the edge current in this section.

Consider a long, disorder-free 2DEG conductor connected to electron reservoirs at opposite ends. A strong magnetic field B is applied along the z direction, where the magnetic length $\ell = \sqrt{h/eB}$ is far smaller than the width of the conductor. The conductor is in equilibrium when the electrochemical potentials, μ_L and μ_R, of the reservoirs are equal to each other, i.e., $\mu_L = \mu_R = \varepsilon_F$. Let $U_0(x)$ be the electrostatic potential energy for the 2DEG and $\varepsilon^0_{n,k}$ the energy of an electron state in this equilibrium condition as schematically shown in the upper panel of Fig. 3.2.1a. Here, $\varepsilon^0_{n,k}$ is labeled with the Landau level index n and the wave vector k along the y direction, which determines the center coordinate $X = -\ell^2 k$ of the wavefunction. So long as $U_0(x)$ is independent of the y coordinate the fraction of current dI due to the electrons filling an infinitesimal energy interval $d\varepsilon$ is given as [9]

$$dI = (-e/h)d\varepsilon. \tag{3.2.1}$$

It follows from (3.2.1) that the current density in the equilibrium state is given by

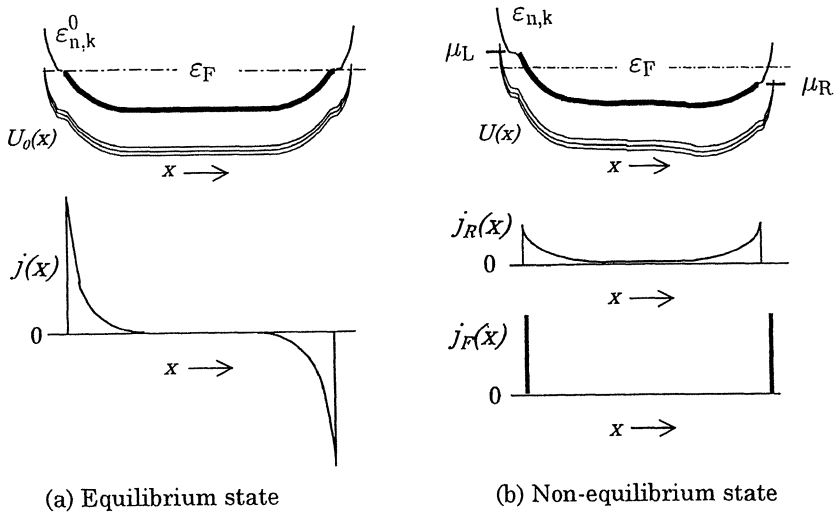
$$j_0(x) = -N_0(n,k)(-e/h)(\partial\varepsilon^0_{n,k}/\partial X), \tag{3.2.2}$$

where $N_0(n,k)$ is the occupation number of the state $|n,k\rangle^0$ with the energy eigenvalue $\varepsilon^0_{n,k}$ and is unity if $\varepsilon^0_{n,k} < \varepsilon_F$ and zero otherwise. The slope of the Landau level $\partial\varepsilon^0_{n,k}/\partial X$ is positive and negative in the opposite edges, and therefore the current density along the right and left edges has opposite signs as schematically shown in the lower panel of Fig. 3.2.1a. The integral of (3.2.2) across the conductor from the left edge ($\varepsilon = \varepsilon_F$) to the right edge ($\varepsilon = \varepsilon_F$) vanishes, assuring that net current is zero.

When we impose a small voltage difference, $(-e)V = \mu_L - \mu_R \ll \hbar\omega_c$, between the reservoirs, the opposite edge states are unequally filled with electrons, respectively, up to μ_L ($> \varepsilon_F$) and μ_R ($< \varepsilon_F$). The small net current $I = (-e/h)(\mu_L - \mu_R)$ flows in the direction normal to the plane of Fig. 3.2.1a. In this condition, a certain number of extra electrons are added to the edge states at the left side, and a certain number of electrons are removed from the edge states at the right side.

If the electrostatic potential $U_0(x)$ is unaffected by those added (removed) extra electrons (holes), the Landau-level energy spectrum $\varepsilon^0_{n,k}$ would be also kept unchanged and the conduction would be envisioned as arising solely from those added extra electrons (holes), which carry current only along the edges.

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998



(a) Equilibrium state (b) Non-equilibrium state

Fig. 3.2.1 Upper panels: Landau-level energies and the electrostatic potential. The thick solid lines indicate occupation by electrons. Lower panels: (**a**) j_0 and (**b**) j_R and j_F

This is the simplified description of the edge current, which has been adopted by many researchers but is misleading.

3.2.2 Halperin's Edge Current

In reality, a negative (positive) line charge due to the added (removed) electrons along the left (right) edge causes the electrostatic potential $U_0(x)$ to increase at the left edge and to decrease at the right edge. (Though smaller in amplitude, polarization charges develop in the interior region as well.) Because of the long-range nature of the Coulomb interaction, the electrostatic-potential change penetrates into the interior region of the conductor. Let $U(x)$ be the modified profile of the electrostatic potential as schematically shown in the top panel of Fig. 3.2.1b. The Landau level energy accordingly increases (decreases) at the left (right) edge, to yield a reorganized Landau-level energy spectrum $\varepsilon_{n,k}$ in the presence of the net conduction. Therefore, with a given electrochemical potential difference between the opposite edges, $\Delta\mu = \mu_L - \mu_R$, the number of electrons added to (removed from) the edge states is less than what would be expected in the hypothetical condition where $U_0(x)$ were unaffected. To be explicit, the amount of added (removed) electrons is determined so that the given $\Delta\mu$ is equal to the sum of the induced electrostatic potential change $dU_{L,R}$ and the change in the Fermi energy $d\varepsilon_F^{L,R}$ due to additional occupation of states at the edges [10–12], i.e.,

$$d\mu_{L,R} = dU_{L,R} + d\varepsilon_F^{L,R} \tag{3.2.3}$$

where $d\mu_{L,R} = \mu_{L,R} - \varepsilon_F$ are the respective changes of the electrochemical potential at the opposite edges.

It follows from (3.2.3) that the total current is the sum of the edge current $I_{\mathrm{H-edge}} = (-e/h)(d\varepsilon_F^L - d\varepsilon_F^R)$ due to the added (removed) electrons and the Hall current $I_{\mathrm{Hall}} = (-e/h)(dU_L - dU_R)$, i.e., [11–13]

$$I = I_{\mathrm{Hall}} + I_{\mathrm{H-edge}}. \tag{3.2.4}$$

Below we designate this edge current $I_{\mathrm{H-edge}}$ as Halperin's edge current [14]. The fraction of Halperin's edge current is approximated as [11]

$$I_{\mathrm{H-edge}}/I = (h/e^2)v_{n,k}^0 C, \tag{3.2.5}$$

where $v_{n,k}^0$ is the velocity of electrons at the edges and C is the electrostatic capacitance [12] between the opposite edges. According to the evaluation by Hirai and Komiyama, the ratio $I_{\mathrm{H-edge}}/I$ is as small as 2–8% in realistic conditions [11].

Several experiments suggest that $I_{\mathrm{H-edge}}/I$ is indeed small. Among them, a Hall potential across a Hall bar was shown to be dominant through an opto-electric effect [4]. However, the studies were made in a condition with relatively large currents. The studies of edge magnetoplasmon excitations should be particularly mentioned [7,8]. The velocity at which edge magnetoplasmon excitations are propagated is a sensitive function of the fraction of the change in the Fermi energy in (3.2.3) [8]. The experiments definitely indicate that the electrostatic potential part in (3.2.3) dominates, and is regarded as a definite experimental evidence for $I_{\mathrm{H-edge}}/I \ll 1$ in the limit of small I.

3.2.3 Local Current Distribution

Irrespective of the detailed profile of $U(x)$ or $\varepsilon_{n,k}$ formed in the presence of a net current, (3.2.1) remains valid to assure that the current density is represented as

$$j(x) = N(n,k)(-e/h)(\partial\varepsilon_{n,k}/\partial_X), \tag{3.2.6}$$

which is proportional to the slope of the reorganized Landau-level energy $\partial\varepsilon_{n,k}/\partial X$. Here, $N(n,k)$ is the occupation number of the state $|n,k\rangle$ with the energy eigenvalue $\varepsilon_{n,k}$ formed in the modified electrostatic potential $U(x)$ and is unity if $\varepsilon_{n,k} < \mu_{L,R}$ for the respective edges and zero otherwise.

Conventionally, the local current distribution is defined as the difference between this current density $j(x)$ and the equilibrium current density $j_0(x)$ in (3.2.2),

$$j_R(x) = j(x) - j_0(x), \tag{3.2.7}$$

which we will refer to as the response-current density [13]. Because the electrostatic potential change is nonzero also in the interior region, the modification of the slope of the Landau-level energy occurs not only in the vicinity of the edges but also in the interior region of the conductor. It follows that $j_R(x)$ in (3.2.7) is generally finite in the interior region of a conductor, as schematically depicted in the middle panel of Fig. 3.2.1b. In particular, if we ignore a small contribution from Halperin's edge current in a slowly varying potential, (3.2.7) is approximated by the Hall current density, $j_R(x) = (e^2/h)E_H(x)$, which is

the local response against the Hall electric field, $E_H(x) = -e^{-1}\partial(U-U_0)/\partial x$, where $I = I_{\text{Hall}} = \int j_R dx$.

The study of the local distribution of j_R in a given conductor is by itself an interesting subject, which has been investigated both theoretically [9,15,16] and experimentally [4,6]. However, it does not have much relevance to the issue of edge current.

3.2.4 Büttiker's Edge Current

The "edge current" referred to by the edge-channel picture of the IQHE [1–3] should be regarded as different from Halperin's edge current. Below we designate the edge current in the edge-channel picture as Büttiker's edge current [13].

Even in the presence of a net current, all the states $|n,k\rangle$ with energies $\varepsilon_{n,k}$ lower than μ_R are completely occupied with electrons in the conductor shown in Fig. 3.2.1b. Consequently, (3.2.1) assures that the contribution to the current from the completely occupied states below μ_R adds up to zero. Therefore, the net current I is given by only the electrons distributed in the small energy interval $\mu_R < \varepsilon_{n,k} < \mu_L$. Since in the conductor of Fig. 3.2.1b these electrons exist only at the edges, the current due to these electrons flows only along the edges.

Generally, it can be argued that, even if a conductor is disordered, the conduction is determined completely by only those electrons distributed in the small energy domain bounded by the lowest $\mu_{\min}$ and the highest $\mu_{\max}$ of the electrochemical potentials of the electron reservoirs attached to the conductor. Moreover, if inelastic scattering processes are absent, each of those electrons in the energy interval of $\mu_{\min} < \varepsilon < \mu_{\max}$ is regarded as in a scattering-wave state that forms an exact eigenstate of the system with the modified electrostatic potential $U(x)$. Since each scattering-wave state describes a state incident on the conductor from a given electron reservoir, different scattering-wave states are occupied with electrons up to different energies. Therefore, the conduction can be viewed as arising from such exact eigenstates that are unequally occupied in the energy interval of $\mu_{\min} < \varepsilon < \mu_{\max}$. Such conduction current carried by the electron waves in the energy interval $\mu_{\min} < \varepsilon < \mu_{\max}$ is referred to as the Fermi-surface current. The Büttiker–Landauer formulation of transport [17,18] can be interpreted as addressing the Fermi-surface current [13].

The distribution of Fermi-surface current $j_F(x)$ is generally different from the distribution of the response current, $j_R(x)$. In particular, $j_F(x)$ in the conductor of Fig. 3.2.1b is represented by

$$j_F(x) = j(x) - j_{\varepsilon_F}(x), \tag{3.2.8}$$

where

$$j_{\varepsilon_F}(x) = N_{\varepsilon_F}(n,k)(e/h)(\partial\varepsilon_{n,k}/\partial X). \tag{3.2.9}$$

Here, $N_{\varepsilon_F}(n,k)$ is unity for $\varepsilon_{n,k} < \varepsilon_F$ and zero for $\varepsilon_{n,k} > \varepsilon_F$. Note that $j_{\varepsilon_F}(x)$ in (3.2.9) refers to the reorganized energy spectrum $\varepsilon_{n,k}$ in the presence of

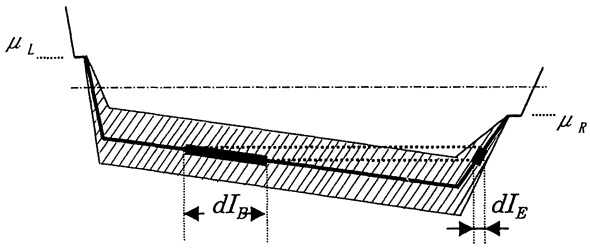


Fig. 3.2.2 A Landau level in a disordered conduction with a net current

transport, while $j_0(x)$ in (3.2.7) refers to the initial energy spectrum $\varepsilon^0_{n,k}$ in the absence of transport. The difference is subtle but crucial for distinguishing between $j_F(x)$ and $j_R(x)$. These two currents would be identical if the conduction occurred without the change in the electrostatic potential. It is for this reason that the difference between $j_F(x)$ and $j_R(x)$ was not explicitly recognized in the original arguments for the edge-channel picture, which did not consider the electrostatic-potential change in a conductor.

The Fermi-surface current of $I_L = (-e/h)(\mu_L - \varepsilon_F)$ is carried by the electrons with energies $\varepsilon_F < \varepsilon < \mu_L$ along the left edge and $I_R = (-e/h)(\varepsilon_F - \mu_R)$ is carried by the holes with energies $\mu_R < \varepsilon < \varepsilon_F$ along the right edge, as schematically represented in the bottom panel of Fig. 3.2.1b,

When a conductor is disordered, as shown in Fig. 3.2.2, electron states with energies around the Fermi level ($\mu_R < \varepsilon < \mu_L$) exist not only at the edges but also in the interior region. In the IQHE regime, however, these states are all localized [19] and do not contribute to Fermi-surface-current. In other words, the electron waves incident on the conductor from the reservoir at the left end with energies $\mu_R < \varepsilon < \mu_L$ are propagated along the left edge of the conductor without being backscattered to the opposite edge [3]. Even in this case, the true physical current j (arising from all the states) is generally finite in the interior region of the conductor due to the existence of bulk extended states lying well below μ_R. However, as indicated in Fig. 3.2.2, each fraction of the bulk current, dI_B, is compensated by a fraction of the current, dI_E, due to electrons in the vicinity of one edge and is not counted in the Fermi-surface current density.

Thus, Fermi-surface current flows only along the edges in the IQHE regime, whereas bulk Hall current is dominant in the conventional sense of local current distribution. The Fermi-surface current in the IQHE regime can thus be called Büttiker's edge current.

The approach of considering Fermi-surface current may not be particularly useful for the description of the simple examples shown in Figs. 3.2.1 and 3.2.2. However, it serves as an extremely instructive and convenient way of looking at transport problems in some experimental conditions. An example is shown in Figure 3.2.3. A potential barrier is introduced in a Hall conductor, in which N Landau levels are occupied. Two electron reservoirs with μ_L and μ_R are attached to both ends.

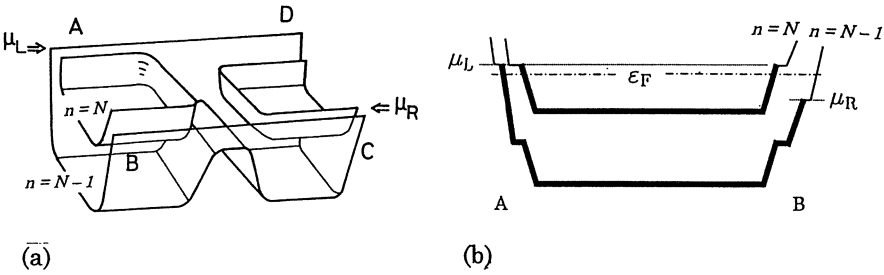


Fig. 3.2.3 (a) Landau levels of $n=N-1$ and $n=N$ in a Hall conductor with a potential barrier. (b) Representation of unequally occupied Landau levels on a cross section A–B

The equilibrium current distribution $j_0(x)$ in this conductor is already highly nonuniform due to the presence of the potential barrier [20]. In the presence of net conduction, the Hall potential distribution becomes even more complicated. These facts make the discussion in terms of the response-current density, $j_R(x)$, an inconvenient approach. In contrast, a simple and instructive physical insight is obtained from the consideration of the Fermi-surface current, because the transport through this conductor is determined only by the influence of the potential barrier on the motion of electrons around the Fermi level ($\mu_R < \varepsilon < \mu_L$). The two-terminal conductance is expressed as $G=-eI/(\mu_L-\mu_R)=(e^2/h)\sum T_n$, where T_n has a definite physical meaning as the transmission coefficient of the electron waves incident on the conductor through the nth edge state.

In general, it is of great importance for practical application of the approach that, when the net current is small, the trace of the Fermi-surface current is substantially unaffected by a small change of the electrostatic potential $U_0(x)-U(x)$, and can therefore be approximated by the one expected for the incoming waves with ε_F in the equilibrium potential $U_0(x)$.

When the net current increases so that $\mu_L-\mu_R$ exceeds $\hbar\omega_c$, the Fermi-surface current begins to flow in the interior region as well [13]. This gives rise to remarkable nonlinear effects [10,21].

3.2.5 Nonlocal Resistance

Experimentally, it is possible to adjust the potential barrier height in the conductor shown in Fig. 3.2.3a so that $T_n=0$ for $n=N$ and $T_n=1$ for $n=1$, $2,\cdots,N-1$. If the edge state of $n=N$ is completely decoupled from the other edge states, the $n=N$ edge state will be occupied with electrons up to μ_L along the boundary from A to B, while the edge states of $n=1,2,\cdots,N-1$ are filled up to μ_R along the side BC. Thus the $n=N$ edge state and the other edge states are occupied up to the different electrochemical potentials at the side B as shown in Fig. 3.2.3b. In this situation, these edge states carry different Fermi-surface currents along the boundary of B; that is, the edge states of $n=1,2,\cdots,N-1$ respectively carry the current of $I_F=(2e/h)(\varepsilon_F-\mu_R)$ due to holes whereas that of $n=N$ carries current $-I_F=(2e/h)(\mu_L-\varepsilon_F)$ in the opposite direction due to electrons, where the spin-up and spin-down states are

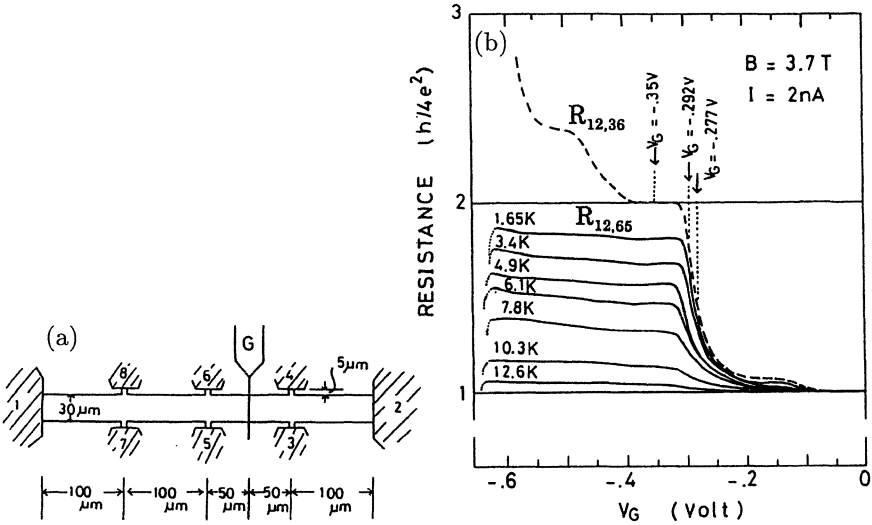


Fig. 3.2.4 (a) A Hall device with a Shottky cross gate. (b) The Hall resistance $R_{12,65} = V_{65}/I_{12}$ versus gate-bias voltage V_G. Potential probe contact 5 is disordered to selectively probe the electrochemical potential of the lower Landau levels

assumed to be degenerate. It should be noted that such adiabatic transport, reported in a number of experiments [22–25], makes explicit the existence of Büttiker's edge-current but does not provide by itself any information about the existence of Halperin's edge current. In fact, the fraction of Halperin's edge current has been estimated to remain small in such conditions [26].

Selective scattering of electron waves in different edge states occurs not only at well defined potential barriers but also in disordered contacts [3,27]. Resistance measurements with disordered contacts thus encounter deviation of the Hall resistance as well as anomalous nonlocal resistance due to the adiabatic transport [28].

Adiabatic transport is destroyed if all the edge states are strongly coupled to establish local equilibrium via inter-edge-state scattering processes. Figure 3.2.4 shows an example for the case of $N = 2$ in a GaAs/AlGaAs single heterostructure device and displays deviations of the Hall resistance from the quantized value $h/4e^2$ due to adiabatic transport, where the magnitude of deviation is proportional to the amplitude of unequal occupation [10]. The inter-edge-state equilibration is promoted to recover the Hall resistance quantization as the temperature is elevated.

The inverse of the inter-edge-state equilibration-length derived from the data of Fig. 3.2.4 is plotted with solid dots against temperature in Fig. 3.2.5. The equilibration length reaches about 0.5 mm at the limit of low temperature. The remarkable suppression of the inter-edge-state scattering processes is supposed to arise from (i) a small overlapping integral between relevant edge-state wavefunctions [29] and (ii) the long range nature of random potentials [30]. According to model calculations assuming a half-parabolic confining

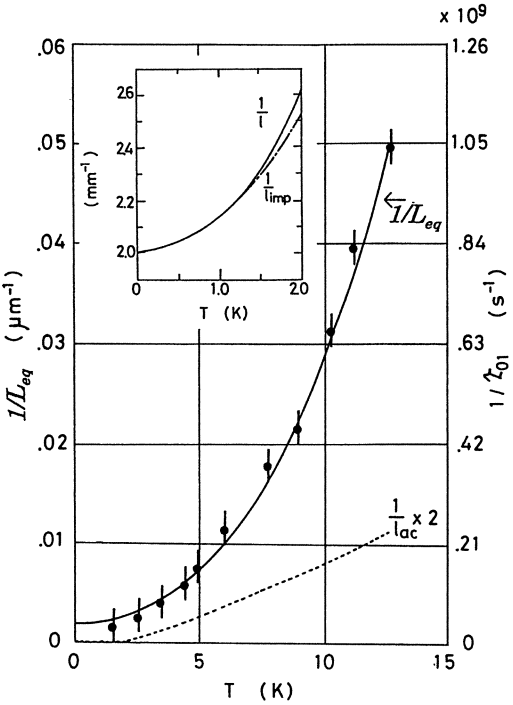


Fig. 3.2.5 Temperature dependence of inter-edge-state equilibration length L_{eq}

potential, $U_0(x) = (1/2m^*)\omega_0^2 x^2$ for $x > 0$ and 0 for $x < 0$, and considering impurity scattering due to Gaussian potential $V_i(x) \propto \exp[-(r - r_i)^2/A^2]$ and acoustic phonon scattering due to a deformation potential, a good fit with the experimental results is obtained as shown with the solid line in Fig. 3.2.5 with parameter values of $\omega_0 = 1.7 \times 10^{12}$ s $= 0.17\omega_c$ and $A = 157$ Å. Although the assumed structureless confining potential is inconsistent with the self-consistent edge-potential expected for the absolute zero temperature [31], the assumption is supposed reasonable for the experimental temperatures [32].

The exponential-type decrease of the equilibration length with increasing temperature has been reproduced in more recent work [33], and explained theoretically [30]. According to the analysis, the inter-edge-state equilibration process is dominated by impurity scattering over the temperature range up to 13 K, and the strong temperature dependence comes from the broadening of the Fermi distribution of electrons in the edge states.

Adiabatic transport manifests itself also in standard Hall bars without a potential barrier when they are in transition regions between successive IQHE states. In the transition region, the electron wave incident on the Nth (highest) Landau level is backscattered due to random potentials while the electron waves incident on the lower Landau levels remain fully transmitted through a conductor without being backscattered.

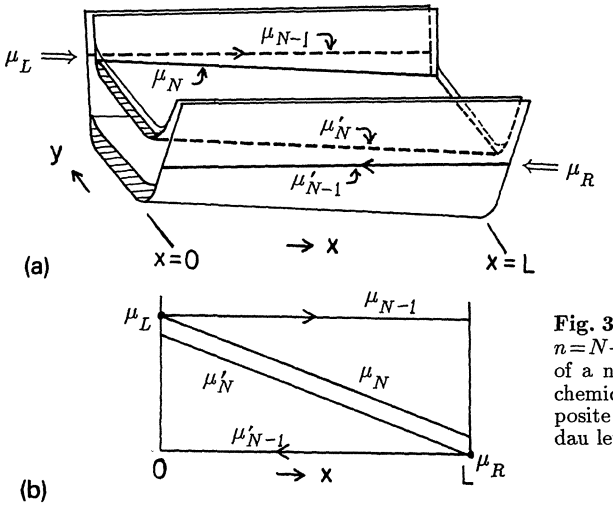


Fig. 3.2.6 (a) Landau levels of $n=N-1$ and N in the presence of a net current. (b) Electrochemical potentials at the opposite edges of respective Landau levels

Suppose that there are no scattering processes between the Nth Landau level and the edge states originating from the lower Landau levels. When a Hall conductor in such conditions is attached at both ends to electron reservoirs with different electrochemical potentials of μ_L and μ_R, the edge states of $n=1, 2, \cdots, N-1$ are occupied up to μ_L and μ_R, respectively, along the opposite boundaries to form equi-electrochemical-potential lines, while the electrochemical potential of the Nth Landau level has a slope in the lengthwise direction in the conductor as shown in Fig. 3.2.6. Unequal occupation thus develops between the Nth Landau level and the lower edge states at each side of the conductor.

When one attaches a voltage probe to a conductor boundary in such a condition, the voltage probe selectively senses the electrochemical potential of the edge states of $n=1, 2, \cdots, N-1$ because the electrons in the Nth Landau level are partially backscattered in the lead region of the voltage probe and only a part of them can be absorbed by the reservoir of the probe. This has remarkable consequences in resistance measurements, such as strong suppression of Shubnikov–de Haas oscillations [25,34,35], anomalous widening of Hall-resistance plateaus, occasionally accompanied by an overshoot [34,36,37], and the occurrence of nonlocal resistance [34,35,38,39]. McEuen has proposed a model to analyze these effects within the framework of the edge-channel formulation on the assumption that the Nth level is completely decoupled [34]. The model was later improved to include finite coupling and successfully explained the effects [35].

Figure 3.2.7a displays the longitudinal resistance and the Hall resistance as a function of magnetic field B in a Hall bar fabricated on a GaAs/AlGaAs single heterostructure crystal [35]. Figure 3.2.7b displays a nonlocal resistance studied in the same device. All the data show that the anomaly due to the adiabatic nature of transport develops as the current decreases below 1 μA. In other words, these data, as well as other available data, strongly suggest

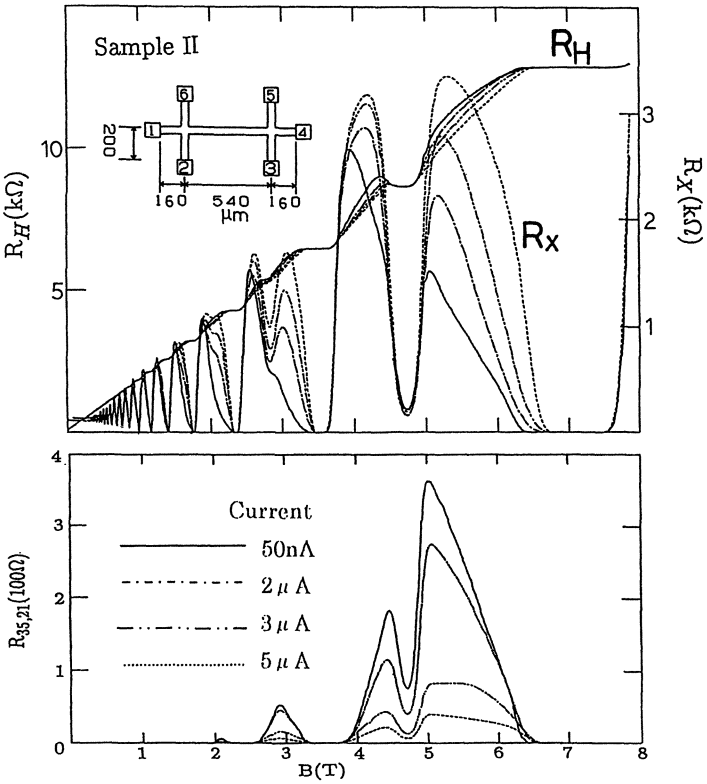


Fig. 3.2.7 Upper panel: Hall resistance $R_H = R_{14,53}$ and longitudinal resistance $R_{14,23}$. Lower panel: nonlocal resistance $R_{35,21}$

that the equilibration among different edge states and between edge and bulk states is remarkably promoted to destroy adiabatic transport with increasing current. The threshold current approximately corresponds to the condition where the amplitude of unequal occupation between edge- and bulk-states reaches $\hbar\omega_c$ [10,35], at which the Fermi-surface current in the lower Landau levels of $n=1, 2, \cdots, N-1$ begins to flow in the interior region.

Solid lines in Fig. 3.2.8 are replots of the curves of the longitudinal resistance and the nonlocal resistance in the transition region from the filling factor of Landau levels $\nu = 4$ ($B = 3.5$ T) to $\nu = 2$ ($B = 7$ T) at the lowest current level in Fig. 3.2.7. The data are respectively compared with theoretical values (dotted lines) calculated by taking the equilibration length $L_{\rm eq}$ between the highest ($n = 2$) Landau level and the lower edge state ($n = 1$) as a parameter [35]. The equilibration length $L_{\rm eq}$ determined by the intersections of the theoretical curves with the experimental ones in Figs. 3.2.7a and (b) is shown as a function of B in Fig. 3.2.9, where similar data in different transition regions are shown together. The sizes of $L_{\rm eq}$ are macroscopic,

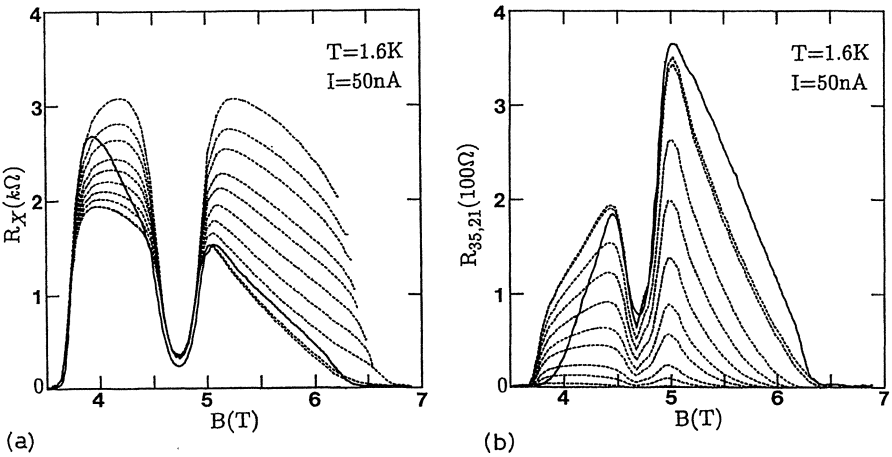


Fig. 3.2.8 Comparison of experimental curves in Fig. 3.2.7 (solid lines) with theoretical values (dotted lines). (a) From the top, $L_{\rm eq}=0.5$, 0.75, 1, 2, 3, 5, 10, 100, and 1000 mm. (b) From the top, $L_{\rm eq}=1000$, 100, $\cdots$, 0.75, and 0.5 mm

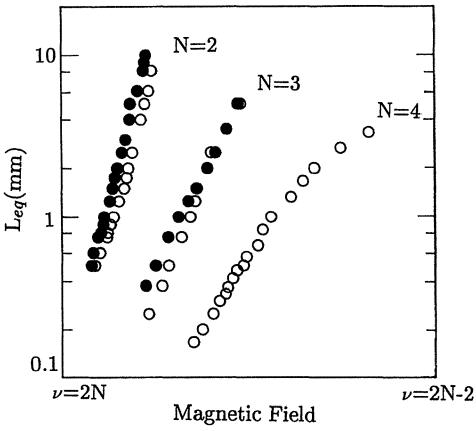


Fig. 3.2.9 Magnetic-field dependence of $L_{\rm eq}$ derived from R_x (open dots) and $T_{35,21}$ (solid dots) for the transition regions of $\nu=2\rightarrow 4$ ($N=2$), $\nu=4\rightarrow 6$ ($N=3$), and $\nu=6\rightarrow 8$ ($N=3$)

and increase exponentially with increasing B. In other words, the decoupling of the highest Landau level from the lower Landau levels is indeed remarkable. The exponential dependence of $L_{\rm eq}$ on B has been reported also by Alphenaar et al. [24] and Faist [39]. The origin of the dependence has not been fully analyzed [33], but is supposed to relate to the increasing distance between the topmost Landau level and the adjacent edge state as suggested from the theoretical studies of the dispersion of edge states [31,40].

References

1 P. Streda, J. Kucera, and A.H. MacDonald, Phys. Rev. Lett. **59**, 1973 (1987)
2 J.K. Jain and S.A. Kivelson, Phys. Rev. B **37**, 4276 (1988)

3 M. Büttiker, Phys. Rev. B **38**, 9375 (1988)
4 P.F. Fontein, P. Hendriks, F.A.P. Blom, J.H. Wolter, L.J. Giling, and C.W. Benakker, Surf. Sci. **263**, 91 (1992)
5 S. Kawaji, K. Hirakawa, M. Nagata, T. Okamoto, T. Goto, and T. Fukase, J. Phys. Soc. Jpn. **63**, 2303 (1994)
6 J. Wakabayashi, K. Sumiyoshi, T. Nagashima, T. Mochiku, and K. Kadowaki, J. Phys. Soc. Jpn. **66**, 413 (1997)
7 V.I. Talyanskii, Sov. Phys. JETP **65**, 1036 (1987)
8 M. Wassermeier, J.Harris, Phys. Rev. B **41**, 10287 (1990)
9 A.H. MacDonald, T.M. Rice, and W.F. Brinkman, Phys. Rev. B **28**, 3648 (1983)
10 S. Komiyama, H. Hirai, M. Ohsawa, and Y. Matsuda, Phys. Rev. B **45**, 11085 (1992)
11 H. Hirai and S. Komiyama, Phys. Rev. B **49**, 14012 (1994); D.J. Thouless, Phys. Rev. Lett. **71**, 1879 (1993)
12 T. Christen and M. Büttiker, Phys. Rev. B **53**, 20064 (1996)
13 S. Komiyama and H. Hirai, Phys. Rev. B **54**, 2067 (1996)
14 B.I. Halperin, Phys. Rev. B **25**, 2185 (1982)
15 J. Riess, J. Phys. C **17**, L847 (1984)
16 D. Pfannkuche and J. Hajdu, Phys. Rev. B **46**, 7032 (1992)
17 R. Landauer, IBM J. Res. Dev. **1**, 223 (1957)
18 M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986)
19 H. Aoki and T. Ando, Surf. Sci. **170**, 249 (1986)
20 T. Ohtsuki and Y. Ono, J. Phys. Soc. Jpn. **50**, 637 (1990)
21 G. Müller et al., Surf. Sci. **263**, 280 (1992)
22 B.J. van Wees, E.M.M. Willems, C.J.P.M. Harrmans, C.W.J. Beenakker, H. van Houten, J.G. Wiiliamson, C.T. Foxon, and J.J. Harris, Phys. Rev. Lett. **62**, 1981 (1989)
23 S. Komiyama, H. Hirai, S. Sasa, and S. Hiyamizu, Phys. Rev. B **40**, 12566 (1989)
24 B.W. Alphenaar, P.L. McEuen, R.G. Wheelar, and R.N. Sacks, Phys. Rev. Lett. **64**, 677 (1990)
25 R.J. Haug, K. von Klitzing, K. Ploog, and P. Streda, Surf. Sci. **229**, 229 (1990)
26 H. Hirai and S. Komiyama, Jpn. J. Appl. Phys. **34**, 4321 (1995)
27 S. Komiyama and H. Hirai, Phys. Rev. B **40**, 7767 (1989)
28 S. Komiyama, H. Hirai, S. Sasa, and T. Fujii, Solid State Commun. **73**, 91 (1990); Surf. Sci. **229**, 224 (1990)
29 T. Martin and S. Fend, Phys. Rev. Lett. **64**, 1971 (1990)
30 D.L. Maslov, Y.B. Levinson, and S.M. Badalian, Phys. Rev. B **46**, 7002 (1992)
31 D.B. Chklovski, B.I. Shklovskii, and L.I. Glazman, Phys. Rev. B **46**, 4026 (1992)
32 T. Machida, H. Hirai, and S. Komiyama, Phys. Rev. B **54**, R14261 (1996)
33 H. Hirai, S. Komiyama, S. Fukatsu, T. Osada Y. Shiraki, and H. Toyoshima, Phys. Rev. B **52**, 11159 (1995)
34 P.L. McEuen, S. Szafer, C.A. Richter, B.W. Alphenaar, J.K. Jain, A.D. Stone, R G. Wheeler, and R.N. Sacks, Phys. Rev. Lett. **64**, 2062 (1990)
35 S. Komiyama and H. Nii, Physica B **184**, 7 (1993)
36 S. Komiyama, H. Nii, M. Ohsawa, S. Fukatsu, Y. Shiraki, R. Itoh, and H. Toyoshima, Solid state Commun. **80**, 157 (1991)
37 C.A. Richter, R.G. Wheeler, and R.N. Sacks, Surf. Sci. **263**, 270 (1992)
38 K. Tsukagoshi, K. Oto, S. Takaoka, K. Murase, Y. Takagaki, K. Gano, and S. Namba, Solid State Commun. **80**, 797 (1991)
39 J. Faist, Europhys. Lett. **15**, 331 (1991)
40 T. Suzuki and T. Ando, J. Phys. Soc. Jpn. **62**, 2986 (1993): Physica B **201**, 345 (1994); **227**, 46 (1996)

3.3 Magnetocapacitance and Edge States

3.3.1 Spatial Dispersion of Edge States

The spatial distribution of edge states at quantum Hall (QH) plateaus has been studied both experimentally and theoretically. In the classical simple picture, the width of the edge states is expected to be of the order of the cyclotron radius, since the electrons move along the sample edge by a skipping motion as shown in Fig. 3.3.1. In the one-electron quantum mechanical picture, the width of the edge states is also of the order of the magnetic length $l = \sqrt{\hbar/eB}$.

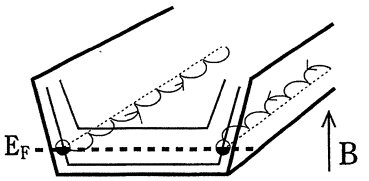


Fig. 3.3.1 One-electron picture of edge states. Electrons move along the edge with a skipping motion

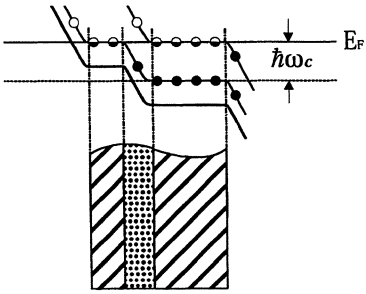


Fig. 3.3.2 The confinement potential and the Landau level when the screening effect is considered. The shaded region is the compressible strip where the filling factor is not integral. The dotted region is the incompressible strip where the Landau level is fully occupied

A model assuming the presence of compressible and incompressible strips was first proposed by Beenakker for the fractional QH regime [1]. The spatial distribution of the edge states becomes flat in the compressible region as shown in Fig. 3.3.2, where the filling factor of the Landau level varies and the confinement potential is screened. The width of the edge states was theoretically evaluated to be of the order of the depletion width at the sample boundary [2]. The depletion width in the GaAs/AlGaAs samples is usually of the order of 1 μm. A self-consistent quantum mechanical calculation also showed the flat dispersion of the edge states [3].

Some experimental results on dispersion of edge states were reported in resistance measurement with point contacts [4], a time resolved measurement of the edge magneto-plasmon [5], and nonlinearity of the Hall resistance through the nonequilibrium edge states [6]. All the results show the spread (flat) feature of the edge state dispersion and a wider edge state width than the cyclotron radius or the magnetic length. In the following, studies by means of magnetocapacitance are described.

3.3.2 Edge States Width and Magnetocapacitance

The capacitance between a two-dimensional electron system (2DES) and a metal plate (gate) covering a GaAs/AlGaAs heterostructure is considered as

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

shown in Fig. 3.3.3. A solution of the Poisson equation in the heterostructure yields [7]:

$$e\phi_g = \frac{e^2 d_b}{\varepsilon_b} N_S + E_F + K, \tag{3.3.1}$$

where ϕ_g is an electric potential of the gate, d_b the thickness of a barrier layer and ε_b the effective dielectric constant of the barrier layer, N_S the electron density of 2DES, and K comes from a term due to fixed charges in the barrier layer and barrier heights at interfaces as shown in Fig. 3.3.3b. At low temperatures carriers in the barrier layer are frozen out and K is regarded as a constant. The capacitance measured by a capacitance bridge with a small ac voltage is regarded as the differential capacitance with respect to the gate voltage. The differential capacitance (C) is expected to be given as

$$C = \frac{dQ}{dV_g} = \frac{d(SeN_S)}{d(\phi_g - E_F/e)} = \frac{\varepsilon_b S}{d_b}, \tag{3.3.2}$$

where V_g is the gate voltage, Q the total charge in 2DES, and S the gated area. It should be noted that V_g is equal not to ϕ_g but to $\phi_g - E_F/e$. This capacitance is determined by the geometry of the capacitor (d_b and S) and is independent of magnetic fields. Actually, however, the oscillations of the capacitance are observed to be as shown in Fig. 3.3.4.

In the QH plateau regime, the capacitance shows minima, and the values are not zero. At the ideal QH plateau ($\sigma_{xx} = 0$) the bulk state is fully localized and electronic conduction occurs only through the edge states. The induced charge in 2DES caused by change of V_g is supplied not to the bulk area but to the edge state region from the electrode. At the QH plateaus, the effective area of the capacitor (S in (3.3.2)) should be reduced from the total gated area to the conductive (edge) region.

The widths of edge states can be evaluated from the bottom values of the capacitance. Assuming that the edge state is a metallic thin strip, the capacitance between an edge state (whose area is S^*) and a metal gate is approximately given as

$$C_e = \frac{\varepsilon_b S^*}{d_b}. \tag{3.3.3}$$

By substituting the measured bottom values for C_e in (3.3.3), the width of the conductive region (W_e) along the edge is estimated by assuming that $S^* = W_e L_e$, where L_e is the total edge length under the gate.

The magnetocapacitance of the samples with different gated areas and with the same longitudinal edge length of the gated area is shown in Fig. 3.3.5 [8]. The bottom values at various filling factor of Landau levels (ν) are plotted with respect to the gated area in Fig. 3.3.5c. The bottom values of the capacitances are almost the same. There was another model for the capacitance oscillations which is directly related to the density of states at the Fermi level (DOS model [9]). The above result cannot be explained by the DOS model, where the minima of the capacitance must be proportional to the gated area.

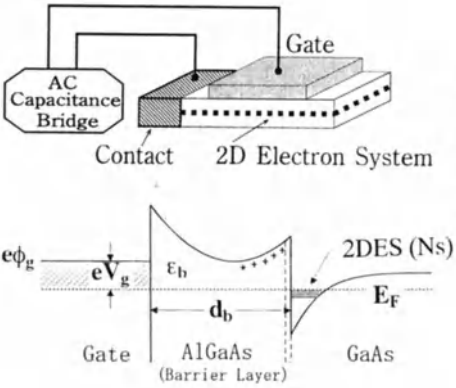


Fig. 3.3.3 (a) Schematic view of samples used in magnetocapacitance measurement. The ohmic contact is connected to the 2DES. (b) Energy level of the gated GaAs/AlGaAs heterostructure. The gate voltage is $\phi_g - e^{-1}E_F$

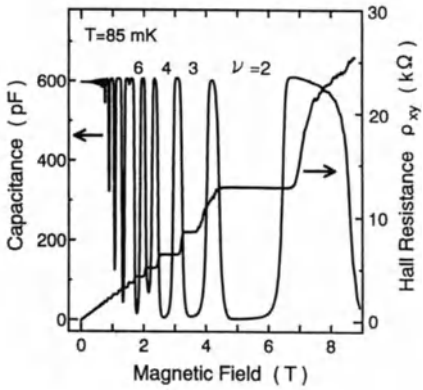


Fig. 3.3.4 Magnetocapacitance and Hall resistance measured with the samples made form the same wafer. At each QH plateau, the capacitance shows minimum, while the Hall resistance is quantized

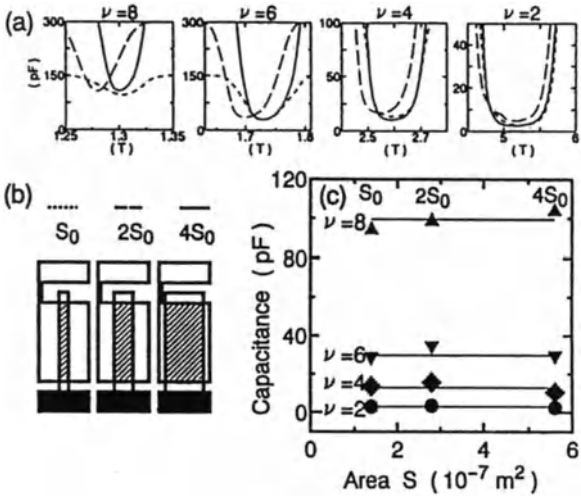


Fig. 3.3.5 (a) Magnetocapacitance of the samples with various gated areas and same edge length of the gate as shown schematically in (b). The area S_0 is 0.1×1.4 mm^2. (c) Bottom values of magnetocapacitance at various filling factor (ν) are plotted with respect to the gated area. Bottom values are nearly independent of the gate area. Solid lines are provided to guide the eye

The magnetocapacitance with different edge lengths ($L_e = 2nL$ with $n =$ 1,2,4, and 8) due to fins and with the same sample length ($L = 1.4$ mm) and width ($D = 0.4$ mm) is shown in Fig. 3.3.6. In Fig. 3.3.6(c), the bottom values at various n are plotted against the edge length. The bottom capacitance

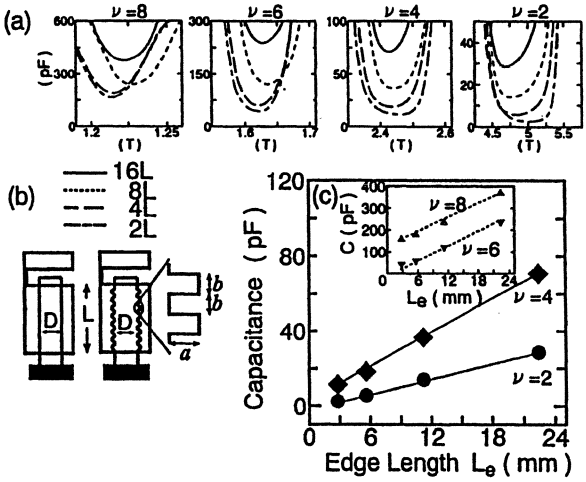


Fig. 3.3.6 (a) Magnetocapacitance of the samples with various edge lengths ($L_e = 2nL$) due to fins as shown schematically in (b). Main sample length L, width D, and fin width b is 1.4, 0.4, and 20 mm, respectively. (c) Bottom values of magnetocapacitance at various ν are plotted with the edge lengths. Bottom values at $\nu = 2$ and 4 are proportional to the edge length. Solid lines are provided to guide the eye. At $\nu = 6$ and 8, as shown in the inset of (c), the capacitance is not proportional to the edge length, since the residual σ_{xx} is high enough to contribute to the capacitance

values become larger by more than several times along with the edge length. The values at $n = 2$ and 4 are nearly proportional to the edge lengths showing that the capacitance at the QH plateaus is governed by the "edge."

To evaluate the area of edge states more precisely, the residual contribution of the bulk state to the capacitance should be considered in a model [10] as shown in Fig. 3.3.7. The edge state capacitance C_e is assumed to be purely capacitive. The bulk state capacitance C_b (complex) is treated as a distributed circuit of the resistive plate [11] with the conductivity σ_{xx}. The total complex capacitance C is given by:

$$C = C_e + C_b(\sigma_{xx}, f) = C_e + (C_0 - C_e)\frac{\tanh(Z)}{Z}, \tag{3.3.4}$$

where f is the ac frequency and C_0 is the capacitance at $f = 0$ for which the zero magnetic field capacitance is substituted, because ρ_{xx} in the absence of the magnetic field is large enough and the frequency dependence can be neglected. The parameter Z is given by $Z = \sqrt{i\omega(C_0 - C_e)\alpha/\sigma_{xx}}$, where $\omega = 2\pi f$ and α is a constant determined by the sample shape. Since σ_{xx} becomes very small at the QH plateaus, the bulk state under the gate may be connected primarily to the edge states. Assuming that the edge states at both sides are treated as the electrodes (length $2L$) for bulk state, and the localized electrons move toward the midpoint ($W/2$) of the sample width, α becomes $(W/2)/(2L)$. By using this model, the residual contribution of the bulk state to the measured capacitance can be determined by the frequency dependence

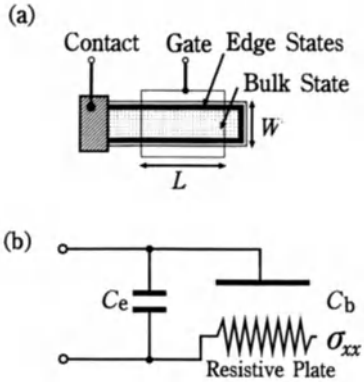


Fig. 3.3.7 (a) Schematic top view of the sample. (b) Equivalent circuit model. Complex capacitance C_b is treated as a distributed system. C_e is edge channel capacitance (real)

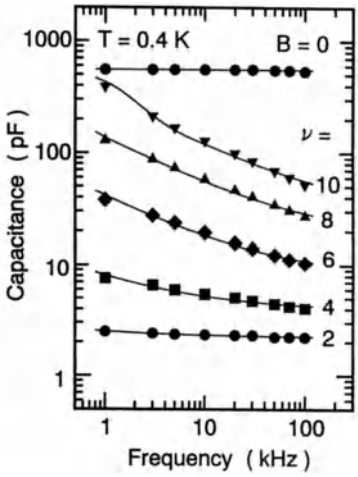


Fig. 3.3.8 Frequency dependence of the capacitance minima at each ν. Solid curves are calculated (see text)

of the capacitance minima at the QH plateaus, and the width of edge states can be obtained.

Table 3.3.1 The parameters used in the fitting. The diagonal conductivity σ_{xx} and the capacitance of the edge channels C_e are listed. The total edge state width W_e are calculated from C_e on the assumption of a two-plate capacitor between edge channels and gate

Filling factor ν	2	4	6	8
σ_{xx} (S)	1.3×10^{-13}	2.9×10^{-11}	1.9×10^{-9}	2.5×10^{-8}
C_e (pF)	2.2	3.9	7.5	17
W_e (μm)	0.8	1.4	2.7	6.0

The frequency dependence of capacitance (real part) at each filling factor ν is plotted in Fig. 3.3.8. Since an appreciable contribution from the bulk state to the capacitance remains at $f = 1$ kHz except for $\nu = 2$, the capacitance decreases markedly with increasing f. At $\nu = 2$, the f dependence is very weak due to very low σ_{xx}. The curves calculated using (3.3.4) can be fitted to the measured one as shown in Fig. 3.3.8. The parameters obtained by the fitting and the evaluated edge state width W_e from C_e are listed in Table 3.3.1. The total width of edge states at $\nu=2$ is 0.8 μm, which is much larger than the order of a magnetic length ($\sim$10 nm) expected from the one-electron picture. The evaluated edge state width W_e (0.8 μm) at $\nu=2$ is comparable to a depletion width 0.2–0.5 μm at the 2DES edge. The calculation by the compressible-incompressible strip model [2] also predicted that the total width

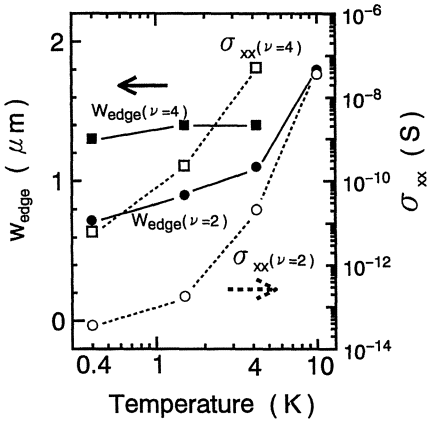


Fig. 3.3.9 Edge state width (closed symbols) and bulk conductivity σ_{xx} (open symbols) by the frequency dependence of the capacitance at $n = 2$ (circle) and 4 (square). Lines are guides for the eyes

of edge states is of the order of the depletion width, and it becomes wider with increasing ν.

The temperature dependence of the edge state width and σ_{xx} evaluated by the frequency dependence of the capacitance at $\nu = 2$ and 4 are plotted in Fig. 3.3.9. At 10 K, both edge state width and σ_{xx} become large due to thermally excited electrons. Below 4.2 K, the width is almost constant in contrast with the drastic change in σ_{xx} of up to two orders of magnitude. The observed wide edge states (μm order) may not be attributed to the residual bulk conductivity due to the thermal activation.

References

1 C.W.J. Beenakker, Phys. Rev. Lett. **64**, 216 (1990)
2 D.B. Chklovskii, B.I. Shklovskii, and L.I. Glazman, Phys. Rev. B **46**, 4026 (1992)
3 L. Bery, J.J. Palacios, and C. Tejedor, Phys. Rev. B **47**, 13884 (1993).
4 S.W. Hwang, D.C. Tsui, and M. Shayegan, Phys. Rev. B **48**, 8161 (1993)
5 N.B. Zhitenev, R.J. Haug, K. von Klitzing, and K. Eberl, Phys. Rev. Lett. **71**, 2292 (1993)
6 T. Machida, H. Hirai, S. Komiyama, T. Osada, and Y. Shiraki, Phys. Rev. B **54**, 14261 (1996)
7 V. Mosser, D. Weiss, K. von Klitzing, K. Ploog, and G. Weimann, Solid State Commun. **58**, 5 (1986)
8 S. Takaoka, K. Oto, H. Kurimoto, K. Murase, K. Gamo, and S. Nishi, Phys. Rev. Lett. **72**, 3080 (1994)
9 T.P. Smith, B.B. Goldberg, P.J. Stiles, and M. Heiblum, Phys. Rev. B **32**, 2696 (1985)
10 K. Oto, S. Takaoka, H. Kurimoto, and K. Murase, *Proceedings of 11th International Conference on High Magnetic Fields in Semiconductor Physics* (World Scientific, Singapore, 1995) p. 142
11 R.K. Goodall, R.J. Higgins, and J.P. Harrang, Phys. Rev. B **31**, 6597 (1985)

4. Electron-Photon Interaction in Nanostructures

4.1 Introduction

A reduction of the dimensionality of the electron motion in quantum wells, wires, and dots is expected to realize new device concepts or improvements in the performance of existing devices, such as transistors and lasers. In 1982, Arakawa and Sakaki proposed the use of quantum wires or dots as an active medium in semiconductor lasers, suggesting a significant improvement of lasing characteristics [1]. The quantum wire effects were realized by placing double heterostructure lasers in a high magnetic field [2]. This report is the first in which the concept of three-dimensionally confined electrons in semiconductor quantum dots is discussed. The target is still distant, because sufficiently small and uniform structures must be achieved.

A full confinement of photons in microcavitites is considered to be able to realize zero-threshold lasers in which electron-hole transitions are coupled with only a single spontaneous emission mode. There are two coupling regimes in the interaction between electrons and photons, the strong and weak coupling regime. In conventional semiconductor lasers, the weak coupling regime is usually used. The strong coupling regime is achieved in microcavities with a high Q-value and leads to coherent interaction between electrons and photons [3,4].

In this chapter, we discuss various electron-photon interactions in nanostructures and microcavities. In Sect. 4.2 fundamentals of the electron-photon interaction are discussed. In Sect. 4.3 effects of a microcavity, in particular, a spontaneous-emission control and cavity-polariton effect, are described. In Sect. 4.4 the electron-photon interaction in photonic crystals is discussed together with some experimental results. In Sect. 4.5, we discuss microcavity surface emitting lasers for practical device applications. Finally in Sect. 4.6, some efforts toward the ultimate semiconductor lasers using fully quantized electrons and photons are discussed.

References

1 Y. Arakawa and H. Sakaki, Appl. Phys. Lett. **40**, 939 (1982).
2 Y. Arakawa, M. Nishioka, H. Sakaki, and H. Okamoto, Jpn. J. Appl. Phys. **22**, 12, L804 (1983)
3 C. Weisbuch M. Nishioka, A.Ishikawa, and Y. Arakawa, Phys. Rev. Lett. **69**, 3314 (1992)
4 T. Norris, J.-K. Rhee, C.-Y. Sung, Y. Arakawa, M. Nishioka, and C. Weisbuch, Phys. Rev. B **50**, 14663 (1994)

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

4.2 Theory of Electron-Photon Interaction

4.2.1 Electron and Hole Operators in Insulating Solids

First, the theory for the bulk crystals [1–3] is described. According to Bloch's theorem, a single-body state of an electron in a solid crystal of volume V takes the following form:

$$\phi_{b\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k}\cdot\mathbf{r}} u_{b\mathbf{k}}(\mathbf{r}), \tag{4.2.1}$$

where b labels bands (conduction band, heavy-hole band, and so on), and $u_{b\mathbf{k}}$ is the cell-periodic function. The spin index is omitted for the sake of simplicity. We can construct the Wannier function which is localized around cell ℓ as

$$w_b(\mathbf{r}-\mathbf{R}_\ell) \equiv \sqrt{\frac{v_0}{V}} \sum_{\mathbf{k}}{}' e^{-i\mathbf{k}\cdot\mathbf{R}_\ell} \phi_{b\mathbf{k}}(\mathbf{r}), \tag{4.2.2}$$

where $\mathbf{R}_\ell$ is the center position of the cell and $\sum'$ denotes the summation over the first Brillouin zone. The inverse transformation is

$$\phi_{b\mathbf{k}}(\mathbf{r}) = \sqrt{\frac{v_0}{V}} \sum_{\ell} e^{i\mathbf{k}\cdot\mathbf{R}_\ell} w_b(\mathbf{r}-\mathbf{R}_\ell), \tag{4.2.3}$$

where v_0 is the volume of a unit cell. Both Bloch and Wannier functions form complete sets. To describe many-body states, we introduce the electron field operator $\hat{\psi}(\mathbf{r})$ which satisfies

$$\{\hat{\psi}(\mathbf{r}), \hat{\psi}^\dagger(\mathbf{r}')\} = \delta(\mathbf{r}-\mathbf{r}'), \quad \{\hat{\psi}(\mathbf{r}), \hat{\psi}(\mathbf{r}')\} = 0. \tag{4.2.4}$$

We can expand $\hat{\psi}(\mathbf{r})$ as

$$\hat{\psi}(\mathbf{r}) = \sum_b \sum_{\mathbf{k}}{}' \phi_{b\mathbf{k}}(\mathbf{r}) \hat{c}_{b\mathbf{k}} = \sum_b \sum_\ell w_b(\mathbf{r}-\mathbf{R}_\ell) \hat{c}_{b\ell}. \tag{4.2.5}$$

We can easily show that $\hat{c}_{b\mathbf{k}}$ and $\hat{c}_{b\ell}$ satisfy

$$\{\hat{c}_{b\mathbf{k}}, \hat{c}^\dagger_{b'\mathbf{k}'}\} = \delta_{bb'}\delta_{\mathbf{k}\mathbf{k}'}, \quad \{\hat{c}_{b\mathbf{k}}, \hat{c}_{b'\mathbf{k}'}\} = 0, \tag{4.2.6}$$

$$\{\hat{c}_{b\ell}, \hat{c}^\dagger_{b'\ell'}\} = \delta_{bb'}\delta_{\ell\ell'}, \quad \{\hat{c}_{b\ell}, \hat{c}_{b'\ell'}\} = 0. \tag{4.2.7}$$

We can interpret $\hat{c}^\dagger_{b\mathbf{k}}$ and $\hat{c}^\dagger_{b\ell}$ as the creation operators of an electron whose wavefunction is $\phi_{b\mathbf{k}}(\mathbf{r})$ and $w_b(\mathbf{r}-\mathbf{R}_\ell)$, respectively.

We here consider a non-doped semiconductor or dielectric material. It is customary to define the creation operator of a hole ($b=v$) by

$$\hat{h}^\dagger_{v\mathbf{k}\uparrow} \equiv \hat{c}_{v-\mathbf{k}\downarrow}, \tag{4.2.8}$$

where only in this equation we restore the spin index for the reader's convenience. The ground state is the vacuum of the holes and the conduction electrons:

$$\hat{c}_{c\mathbf{k}}|g\rangle = \hat{h}_{v\mathbf{k}}|g\rangle = 0, \tag{4.2.9}$$

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

electrons:

$$\hat{c}_{c\mathbf{k}}|g\rangle = \hat{h}_{v\mathbf{k}}|g\rangle = 0, \tag{4.2.9}$$

where c and v stand for conduction and valance bands, respectively. Hereafter, we call a conduction electron simply an electron.

It is convenient to decompose the fundamental field $\hat{\psi}$ into the electron and hole parts, $\hat{\psi}_e$ and $\hat{\psi}_h$, respectively, as

$$\hat{\psi}(\mathbf{r}) = \hat{\psi}_e(\mathbf{r}) + \hat{\psi}_h^\dagger(\mathbf{r}), \tag{4.2.10}$$

$$\hat{\psi}_e(\mathbf{r}) \equiv \sum_c \sum_{\mathbf{k}}{}' \phi_{c\mathbf{k}}(\mathbf{r})\hat{c}_{c\mathbf{k}}, \tag{4.2.11}$$

$$\hat{\psi}_h^\dagger(\mathbf{r}) \equiv \sum_v \sum_{\mathbf{k}}{}' \phi_{v\mathbf{k}}(\mathbf{r})\hat{c}_{v\mathbf{k}} = \sum_v \sum_{\mathbf{k}}{}' \phi_{v\mathbf{k}}(\mathbf{r})\hat{h}_{v-\mathbf{k}}^\dagger. \tag{4.2.12}$$

The anti-commutation relations are evaluated as

$$\{\hat{\psi}_e(\mathbf{r}), \hat{\psi}_e^\dagger(\mathbf{r}')\} = \sum_c \sum_{\mathbf{k}}{}' \phi_{c\mathbf{k}}(\mathbf{r})\phi_{c\mathbf{k}}^*(\mathbf{r}') \equiv \delta_e(\mathbf{r}-\mathbf{r}') = \delta_e^*(\mathbf{r}'-\mathbf{r}), \tag{4.2.13}$$

$$\{\hat{\psi}_h(\mathbf{r}), \hat{\psi}_h^\dagger(\mathbf{r}')\} = \sum_v \sum_{\mathbf{k}}{}' \phi_{v\mathbf{k}}^*(\mathbf{r})\phi_{v\mathbf{k}}(\mathbf{r}') \equiv \delta_h(\mathbf{r}-\mathbf{r}') = \delta_h^*(\mathbf{r}'-\mathbf{r}), \tag{4.2.14}$$

and other anti-commutators are zero. Note that

$$\delta(\mathbf{r}-\mathbf{r}') = \delta_e(\mathbf{r}-\mathbf{r}') + \delta_h^*(\mathbf{r}-\mathbf{r}'). \tag{4.2.15}$$

Hence, δ_c and δ_v work as projection operators onto subspaces spanned by $\phi_{c\mathbf{k}}(\mathbf{r})$ and $\phi_{v\mathbf{k}}(\mathbf{r})$, respectively. It is only in each subspace that they work as delta functions and $\hat{\psi}_e^\dagger(\mathbf{r})$ and $\hat{\psi}_h^\dagger(\mathbf{r})$ can be interpreted as operators which create an electron and a hole, respectively, at the position $\mathbf{r}$. One should always keep this limited meaning of δ_e, δ_h, $\hat{\psi}_e$, and $\hat{\psi}_h$ in mind to avoid inconsistencies.

4.2.2 Effective-Mass Approximation

A single-electron state can be obtained by operating $\hat{c}_{c\mathbf{k}}^\dagger$ on $|g\rangle$. It is an eigenstate of the mean-field Hamiltonian $\hat{\mathcal{H}}_{\mathrm{mf}}$ of the perfect crystal:

$$\hat{\mathcal{H}}_{\mathrm{mf}}\hat{c}_{c\mathbf{k}}^\dagger|g\rangle = \varepsilon_c(\mathbf{k})\hat{c}_{c\mathbf{k}}^\dagger|g\rangle, \tag{4.2.16}$$

where we have taken $\hat{\mathcal{H}}_{\mathrm{mf}}|g\rangle = 0$ and $\varepsilon_c(\mathbf{k})$ is called the band energy. A single-hole state, on the other hand, is obtained by operating $\hat{h}_{v\mathbf{k}}^\dagger$ on $|g\rangle$, and its energy is $-\varepsilon_v(\mathbf{k})$:

$$\hat{\mathcal{H}}_{\mathrm{mf}}\hat{h}_{v\mathbf{k}}^\dagger|g\rangle = -\varepsilon_v(\mathbf{k})\hat{h}_{v\mathbf{k}}^\dagger|g\rangle. \tag{4.2.17}$$

Suppose that $\varepsilon_c(\mathbf{k})$ takes the minimum value at $\mathbf{k} = \mathbf{k}_c$, and $\varepsilon_v(\mathbf{k})$ is maximum at $\mathbf{k} = \mathbf{k}_v$. The minimum energy cost for creating an electron in band c and a hole in band v is

$$\varepsilon_c(\mathbf{k}_c) - \varepsilon_v(\mathbf{k}_v) \equiv E_{Gcv}, \tag{4.2.18}$$

which is called the energy gap between bands c and v. In particular, the energy gap between the bottom (lowest) conduction band c_m and the top (highest) valance band v_m is simply called the energy gap, and is denoted by E_G. For simplicity, we hereafter assume that this gap is located at $\mathbf{k}=0$, i.e., $\mathbf{k}_{c_m}=\mathbf{k}_{v_m}=0$:

$$\varepsilon_{c_m}(0)-\varepsilon_{v_m}(0)\equiv E_G. \tag{4.2.19}$$

We are interested in states with small k in bands c_m and v_m. Hereafter, we write c_m and v_m simply as c and v.

When an additional single-body potential V_1, which may be due to impurities, external potential, and so on, are present the additional term appears in the Hamiltonian;

$$\hat{\mathcal{H}}=\hat{\mathcal{H}}_{\mathrm{mf}}+\int_V d^3r\ \hat{\psi}^\dagger(\mathbf{r})V_1(\mathbf{r})\hat{\psi}(\mathbf{r})\equiv\hat{\mathcal{H}}_{\mathrm{mf}}+\hat{\mathcal{H}}_1. \tag{4.2.20}$$

In this case, $\hat{c}^\dagger_{c\mathbf{k}}|g\rangle$ and $\hat{h}^\dagger_{v\mathbf{k}}|g\rangle$ are no longer eigenstates, and we must take their linear combinations to obtain true eigenstates. The linear combination must be taken over different bands and different $\mathbf{k}$, in general. However, the band mixing may be negligible if (i) the spatial variation of V_1 is slow (i.e., the characteristic length of the variation $\gg$ lattice constants) and if (ii) the band of interest is well separated around $\mathbf{k}=0$ from other bands. (This condition is usually satisfied for the conduction band. For the valence band, however, the degeneracy of heavy and light hole bands causes band mixing in many semiconductors. This point will be discussed later.)

In such a case, a single-electron eigenstate,

$$\hat{\mathcal{H}}|e\eta\rangle=\varepsilon_{e\eta}|e\eta\rangle, \tag{4.2.21}$$

which is labeled by some set of quantum numbers η, takes the following form:

$$|e\eta\rangle=\int_V d^3r\ \varphi_{c\eta}(\mathbf{r})\hat{\psi}^\dagger(\mathbf{r})|g\rangle=\int_V d^3r\ \varphi_{c\eta}(\mathbf{r})\hat{\psi}^\dagger_e(\mathbf{r})|g\rangle, \tag{4.2.22}$$

with

$$\varphi_{c\eta}(\mathbf{r})=\sum_{\mathbf{k}}{}'\tilde{F}_{c\eta}(\mathbf{k})\phi_{c\mathbf{k}}(\mathbf{r})=\sqrt{v_0}\sum_{\ell}F_{c\eta}(\mathbf{R}_\ell)w_c(\mathbf{r}-\mathbf{R}_\ell), \tag{4.2.23}$$

where

$$F_{c\eta}(\mathbf{r})\equiv\frac{1}{\sqrt{V}}\sum_{\mathbf{k}}{}'e^{i\mathbf{k}\cdot\mathbf{r}}\tilde{F}_{c\eta}(\mathbf{k}). \tag{4.2.24}$$

The probability amplitude of finding an electron at the position $\mathbf{r}$ in the subspace spanned by $\phi_{c\mathbf{k}}(\mathbf{r})$ is evaluated as

$$\langle g|\hat{\psi}_e(\mathbf{r})|e\eta\rangle=\varphi_{c\eta}(\mathbf{r}). \tag{4.2.25}$$

Therefore, we can interpret $\varphi_{e_\eta}(\mathbf{r})$ as the single-body wavefunction of the

electron in the subspace. In a similar manner, a single-hole eigenstate takes the following form:

$$\hat{\mathcal{H}}|h\eta\rangle = \varepsilon_{h\eta}|h\eta\rangle, \tag{4.2.26}$$

$$|h\eta\rangle = \int_V d^3r\ \varphi^*_{v\bar{\eta}}(\mathbf{r})\hat{\psi}(\mathbf{r})|g\rangle = \int_V d^3r\ \varphi^*_{v\bar{\eta}}(\mathbf{r})\hat{\psi}^\dagger_h(\mathbf{r})|g\rangle, \tag{4.2.27}$$

$$\varphi_{v\eta}(\mathbf{r}) = \sum_{\mathbf{k}}{}' \tilde{F}_{v\eta}(\mathbf{k})\phi_{v\mathbf{k}}(\mathbf{r}) = \sqrt{v_0}\sum_{\ell} F_{v\eta}(\mathbf{R}_\ell)w_v(\mathbf{r}-\mathbf{R}_\ell), \tag{4.2.28}$$

$$F_{v\eta}(\mathbf{r}) \equiv \frac{1}{\sqrt{V}}\sum_{\mathbf{k}}{}' e^{i\mathbf{k}\cdot\mathbf{r}}\tilde{F}_{v\eta}(\mathbf{k}), \tag{4.2.29}$$

where $\bar{\eta}$ is the "hole-conjugate" of η. That is, we define $\bar{\eta}$ in such a way that it can be naturally interpreted as a set of quantum numbers of the hole: for a hydrogen-like state, for example, $\eta=(n,\ell,m,\sigma)$, where (n,ℓ,m) are the orbital quantum numbers and σ denotes spin, we then take $\bar{\eta}=(n,\ell,-m,-\sigma)$. The probability amplitude of finding a hole at the position $\mathbf{r}$ in the subspace spanned by $\phi_{v\mathbf{k}}(\mathbf{r})$ is given by

$$\langle g|\hat{\psi}_h(\mathbf{r})|h\eta\rangle = \varphi^*_{v\bar{\eta}}(\mathbf{r}). \tag{4.2.30}$$

Therefore, we can interpret $\varphi^*_{v\bar{\eta}}(\mathbf{r})$ as the single-body wavefunction of the hole which has the set of quantum numbers $\bar{\eta}$.

Under our assumptions (i) and (ii), $F_{b\eta}(\mathbf{r})$ $(b = c,\ v)$ becomes a slowly varying function of $\mathbf{r}$ and hence called the envelope function. We note that if a function $A(\mathbf{r})$ is slowly varying and if a function $B(\mathbf{r})$ is a cell-periodic function, then

$$\int_V d^3r\ A(\mathbf{r})B(\mathbf{r}) \simeq \sum_{\ell} A(\mathbf{R}_\ell)\int_{v_0} d^3\vec{\rho}\ B(\vec{\rho}) \simeq \Big[\int_V d^3r A(\mathbf{r})\Big]\cdot\Big[\frac{1}{v_0}\int_{v_0} d^3\vec{\rho}\ B(\vec{\rho})\Big], \tag{4.2.31}$$

where $\int_{v_0} d^3\vec{\rho}$ means the integral over a unit cell. Using this relation and the assumptions (i) and (ii), we can derive the approximate eigenvalue equations for the envelope functions as

$$\Big[\varepsilon_c\Big(-i\vec{\nabla}\Big) + V_1(\mathbf{r})\Big]F_{c\eta}(\mathbf{r}) = \varepsilon_{e\eta}F_{c\eta}(\mathbf{r}), \tag{4.2.32}$$

$$\Big[-\varepsilon_v\Big(-i\vec{\nabla}\Big) - V_1(\mathbf{r})\Big]F_{v\eta}(\mathbf{r}) = \varepsilon_{h\bar{\eta}}F_{v\eta}(\mathbf{r}). \tag{4.2.33}$$

In particular, when the bands have parabolic dispersions,

$$\varepsilon_c(\mathbf{k}) = \varepsilon_c(0) + \frac{\hbar^2k^2}{2m_e}, \qquad \varepsilon_v(\mathbf{k}) = \varepsilon_v(0) - \frac{\hbar^2k^2}{2m_h}, \tag{4.2.34}$$

then we obtain the Schrödinger equations for the envelope functions:

$$\Big[-\frac{\hbar^2}{2m_e}\nabla^2 + \varepsilon_c(0) + V_1(\mathbf{r})\Big]F_{c\eta}(\mathbf{r}) = \varepsilon_{c\eta}F_{c\eta}(\mathbf{r}), \tag{4.2.35}$$

$$\Big[-\frac{\hbar^2}{2m_h}\nabla^2-\varepsilon_v(0)-V_1(\mathbf{r})\Big]F_{v\eta}(\mathbf{r})=\varepsilon_{h\bar{\eta}}F_{v\eta}(\mathbf{r}). \tag{4.2.36}$$

These eigenvalue equations are much easier to solve than the original Schrödinger equations (4.2.21) and (4.2.26), because the complicated crystal potentials in $\mathcal{H}_{\rm mf}$ are absorbed in the effective masses m_e, m_h and the bottom and top energies of the bands, $\varepsilon_c(0)$ and $\varepsilon_v(0)$, respectively. Once the envelope functions are thus obtained, the wavefunctions can be obtained from (4.2.23) or (4.2.27). However, care should be taken in discussions of optical properties, as we will explain in the subsequent subsections.

We next consider a state in which an electron-hole pair is excited. In this case we must take account of the strong electron-hole interaction. If the band mixing caused by the interaction is negligible a pair state (exciton state), which is labeled by a set of quantum numbers η, can be written as

$$|x\eta\rangle=\int_V d^3r_e\int_V d^3r_h\ \varphi_{x\eta}(\mathbf{R}_e,\mathbf{R}_h)\hat{\psi}_e^\dagger(\mathbf{R}_e)\hat{\psi}_h^\dagger(\mathbf{R}_h)|g\rangle, \tag{4.2.37}$$

with

$$\varphi_{x\eta}(\mathbf{R}_e,\mathbf{R}_h)=\sum_{\mathbf{k}_e}{}'\sum_{\mathbf{k}_h}{}'\tilde{F}_{x\eta}(\mathbf{k}_e,\mathbf{k}_h)\phi_{c\mathbf{k}_e}(\mathbf{R}_e)\phi^*_{v\mathbf{k}_h}(\mathbf{R}_h) \tag{4.2.38}$$

$$=v_0\sum_j\sum_\ell F_{x\eta}(\mathbf{R}_j,\mathbf{R}_\ell)w_c(\mathbf{R}_e-\mathbf{R}_j)w_v^*(\mathbf{R}_h-\mathbf{R}_\ell), \tag{4.2.39}$$

where

$$F_{x\eta}(\mathbf{R}_e,\mathbf{R}_h)=\frac{1}{V}\sum_{\mathbf{k}_e}{}'\sum_{\mathbf{k}_h}{}'e^{i\mathbf{k}_e\cdot\mathbf{R}_e-i\mathbf{k}_h\cdot\mathbf{R}_h}\tilde{F}_{x\eta}(\mathbf{k}_e,\mathbf{k}_h). \tag{4.2.40}$$

The probability amplitude of finding an electron at the position $\mathbf{R}_e$ and a hole at $\mathbf{R}_h$ is evaluated as

$$\langle g|\hat{\psi}_h(\mathbf{R}_h)\hat{\psi}_e(\mathbf{R}_e)|\varphi_{x\eta}\rangle=\varphi_{x\eta}(\mathbf{R}_e,\mathbf{R}_h). \tag{4.2.41}$$

Therefore, $\varphi_{x\eta}(\mathbf{R}_e,\mathbf{R}_h)$ can be interpreted as the wavefunction of the pair. Note that $\varphi_{x\eta}(\mathbf{R}_e,\mathbf{R}_h)$ is not required to be antisymmetric under interchange of $\mathbf{R}_e$ and $\mathbf{R}_h$. That is, we can treat an electron and a hole as if they were different particles (as far as we can neglect the band mixing).

The forms (4.2.38)–(4.2.39) of the wavefunction can describe both the Wannier exciton (for which the effective Bohr radius $a_B^*\gg$ lattice constant a) and the Frenkel exciton (for which $a_B^*\sim a$). We here consider the case of the Wannier exciton, and assume, for simplicity, the parabolic dispersions (4.2.34). Then, in a perfect crystal, the total crystal momentum $\mathbf{k}$ of the pair becomes a good quantum number, and we write $\eta=(\mathbf{K},\nu)$, where ν denotes the set of the other quantum numbers. Considering that $\mathbf{k}$ should be shared between the electron and hole in proportion to their masses, we can take $\tilde{F}_{x\eta}$

in the following form:

$$\tilde{F}_{x\eta}(\mathbf{k}_e, \mathbf{k}_h) = \sum_{\mathbf{k}}{}' \tilde{U}_\nu(\mathbf{k}) \delta_{\mathbf{k}_e, \mathbf{k}+\frac{m_e}{M}\mathbf{k}} \delta_{\mathbf{k}_h, \mathbf{k}-\frac{m_h}{M}\mathbf{k}} \equiv \tilde{F}_{\mathbf{k}\nu}(\mathbf{k}_e, \mathbf{k}_h), \tag{4.2.42}$$

where $\tilde{U}_\nu$ is a function to be determined and $M \equiv m_e + m_h$. Inserting this form into (4.2.40), we obtain

$$F_{x\eta}(\mathbf{R}_e, \mathbf{R}_h) = \frac{1}{\sqrt{V}} \exp[i\mathbf{k} \cdot \mathbf{R}]\ U_\nu(\mathbf{R}_e - \mathbf{R}_h) \equiv F_{\mathbf{k}\nu}(\mathbf{R}_e, \mathbf{R}_h), \tag{4.2.43}$$

where we have introduced the center-of-mass coordinate

$$\mathbf{r} \equiv \frac{m_e \mathbf{R}_e + m_h \mathbf{R}_h}{M}, \tag{4.2.44}$$

and

$$U_\nu(\mathbf{r}) \equiv \frac{1}{\sqrt{V}} \sum_{\mathbf{k}}{}' e^{i\mathbf{k}\cdot\mathbf{r}} \tilde{U}_\nu(\mathbf{k}). \tag{4.2.45}$$

We can interpret (4.2.43) as the product of the center-of-mass envelope function $e^{i\mathbf{k}\cdot\mathbf{r}}/\sqrt{V}$ and the relative-motion envelope function U_ν.

The functional form of U_ν depends crucially on the electron-hole interaction. The form of the e-h interaction was reviewed, for example, in [1]. It was shown there that the form is nontrivial for $r \lesssim a_B^*$ because of various exchange potentials and the finite size of the electron and hole. Since we are considering Wannier excitons, the long-range parts of the interactions are more important than the short-range parts. The dominant e-h interaction would then be the screened Coulomb attraction $-e^2/\epsilon r$, where ϵ is the dielectric constant. When the e-h interaction is approximated by it,

$$\left[-\frac{\hbar^2}{2\mu}\nabla^2 - \frac{e^2}{\epsilon r} \right] U_\nu = E_\nu U_\nu, \tag{4.2.46}$$

where μ is the reduced mass, $1/\mu = 1/m_e + 1/m_h$. By solving the simple eigenvalue equation (4.2.46), we can calculate the eigenenergy $E_{\mathbf{k}\nu}$ and the wavefunction $\varphi_{\mathbf{k}\nu}$ of the state $|\mathbf{k}\nu\rangle$. That is,

$$E_{\mathbf{k}\nu} = E_G + E_\nu + \frac{\hbar^2 K^2}{2M}. \tag{4.2.47}$$

We see that the exciton energy consists of the band gap energy E_G, the energy of the e-h relative motion E_ν and the energy of the center-of-mass motion $\hbar^2 K^2/2M$. The binding energy of the e-h pair is $-E_\nu$. We can obtain the wavefunction $\varphi_{x\eta}(\mathbf{R}_e, \mathbf{R}_h) \equiv \varphi_{\mathbf{k}\nu}(\mathbf{R}_e, \mathbf{R}_h)$ by inserting the eigenfunction U_ν into (4.2.43), and the resulting expression into (4.2.39):

$$\begin{aligned}\varphi_{\mathbf{k}\nu}(\mathbf{R}_e, \mathbf{R}_h) = v_0 \sum_j \sum_\ell \frac{1}{\sqrt{V}} \exp\left[i\mathbf{k} \cdot \frac{m_e \mathbf{R}_j + m_h \mathbf{R}_\ell}{M} \right] U_\nu(\mathbf{R}_j - \mathbf{R}_\ell) \\ \times w_c(\mathbf{R}_e - \mathbf{R}_j) w_v^*(\mathbf{R}_h - \mathbf{R}_\ell).\end{aligned} \tag{4.2.48}$$

Or, we may insert $\tilde{U}_\nu$ into (4.2.38) to get

$$\varphi_{\mathbf{k}\nu}(\mathbf{R}_e, \mathbf{R}_h) = \frac{1}{\sqrt{V}} \exp[i\mathbf{k}\cdot\mathbf{R}]\ S_{\mathbf{k}\nu}(\mathbf{R}_e, \mathbf{R}_h), \tag{4.2.49}$$

$$S_{\mathbf{k}\nu}(\mathbf{R}_e, \mathbf{R}_h) = \frac{1}{\sqrt{V}} {\sum_{\mathbf{k}}}' \exp[i\mathbf{k}\cdot(\mathbf{R}_e - \mathbf{R}_h)]\ \tilde{U}_\nu(\mathbf{k})\ u_{c\mathbf{k}+\frac{m_e}{M}\mathbf{k}}(\mathbf{R}_e) u^*_{v\mathbf{k}-\frac{m_h}{M}\mathbf{k}}(\mathbf{R}_h). \tag{4.2.50}$$

Note that we can replace $u_c(\mathbf{R}_e)$ and $u_v(\mathbf{R}_h)$ with $u_c(\vec{\rho}_e)$ and $u_v(\vec{\rho}_h)$, respectively, where $\vec{\rho}_e$ and $\vec{\rho}_h$ denote the position vectors which are reduced in a unit cell (i.e., they are $\mathbf{R}_e$ and $\mathbf{R}_h$ measured from the center of the unit cells in which $\mathbf{R}_e$ and $\mathbf{R}_h$ are located, respectively).

Actually, except for high-quality crystals kept at a low enough temperature, the crystal momentum $\mathbf{k}$ of the center-of-mass motion is not conserved because the exciton is scattered by phonons, impurities, and/or defects, and the exciton wavefunctions are modified. For example, when the total potential of impurities varies slowly (i.e., the characteristic length of the variation $\gg a_B^*$) the wavefunction of an exciton state $|N\nu\rangle$ trapped by the impurities may be obtained as the superposition of $\varphi_{\mathbf{k}\nu}(\mathbf{R}_e, \mathbf{R}_h)$ over various $\mathbf{k}$,

$$\varphi_{N\nu}(\mathbf{R}_e, \mathbf{R}_h) = {\sum_{\mathbf{k}}}' \tilde{G}_{N\nu}(\mathbf{k}) \varphi_{\mathbf{k}\nu}(\mathbf{R}_e, \mathbf{R}_h) \tag{4.2.51}$$

$$= v_0 \sum_j \sum_\ell G_{N\nu}\Big(\frac{m_e\mathbf{R}_j + m_h\mathbf{R}_\ell}{M}\Big) U_\nu(\mathbf{R}_j - \mathbf{R}_\ell) \times w_c(\mathbf{R}_e - \mathbf{R}_j) w_v^*(\mathbf{R}_h - \mathbf{R}_\ell), \tag{4.2.52}$$

where $\tilde{G}_{N\nu}$ is a superposition coefficient, with N a quantum number, and

$$G_{N\nu}(\mathbf{r}) \equiv \frac{1}{\sqrt{V}} {\sum_{\mathbf{k}}}' \tilde{G}_{N\nu}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{r}} \tag{4.2.53}$$

can be regarded as the envelope function of the trapped center-of-mass motion. It is normalized as

$$\int_V G^*_{N\nu}(\mathbf{r}) G_{N'\nu}(\mathbf{r}) d^3R = {\sum_{\mathbf{k}}}' \tilde{G}^*_{N\nu}(\mathbf{k}) \tilde{G}_{N'\nu}(\mathbf{k}) = \delta_{N,N'}. \tag{4.2.54}$$

The trapping may be viewed as the Anderson localization of the exciton, where the size ℓ_G of the spatial extension of $G_{N\nu}(\mathbf{r})$ is the localization length $\ell_{\rm loc}$. We may also use the form (4.2.51) or (4.2.52) as an approximate wavefunction when the quantum coherence of the center-of-mass motion is broken by scatterings by phonons or other excitons. In this case ℓ_G is the phase breaking length ℓ_ϕ. Both the localization and the phase breaking may be important in real systems, and ℓ_G would be given roughly by

$$\ell_G \simeq \min[\ell_{\rm loc},\ \ell_\phi]. \tag{4.2.55}$$

It will be shown later that the oscillator strength of the exciton is proportional to ℓ_G^3 when ℓ_G is less than the wavelength of the optical field.

4.2.3 Optical Matrix Elements

The envelope functions should not be taken as the true wavefunctions because optical phenomena are frequently accompanied with transitions between the cell-periodic parts $u_{b\mathbf{k}}$ of the Bloch functions. That is, optical matrix elements depend not only on the envelope functions but also on the functional forms of $u_{b\mathbf{k}}$ (or, equivalently, of the Wannier functions). The latter dependence becomes crucial when nonlinear optical processes or intraband transitions are discussed. Unfortunately, this point was disregarded in much literature, and unphysical results were sometimes reported. Here, we present typical matrix elements for the states described in the previous subsection.

The interaction with an external optical field $\mathbf{A}(\mathbf{r},t)$ in the Coulomb gauge takes the following form:

$$\hat{\mathcal{H}}_I = \int_V d^3r\ \hat{\psi}^\dagger(\mathbf{r})\Big[-\frac{\hbar e}{imc}\mathbf{A}\cdot\vec{\nabla} + \frac{e^2}{2mc^2}A^2\Big]\hat{\psi}(\mathbf{r}) = \hat{\mathcal{H}}_I^{(1)} + \hat{\mathcal{H}}_I^{(2)}, \quad (4.2.56)$$

where $\hat{\mathcal{H}}_I^{(1)}$ is linear in $\mathbf{A}$, whereas $\hat{\mathcal{H}}_I^{(2)}$ is quadratic. To evaluate various matrix elements of $\hat{\mathcal{H}}_I^{(1)}$ and $\hat{\mathcal{H}}_I^{(2)}$, it is useful to derive formulas for matrix elements of a general operator of the form

$$\hat{Q} = \int_V d^3r \int_V d^3r'\ Q(\mathbf{r},\mathbf{r}')\hat{\psi}^\dagger(\mathbf{r})\hat{\psi}(\mathbf{r}'). \quad (4.2.57)$$

In some cases, its expectation value in the ground state $|g\rangle$ does not vanish. Such a nonzero expectation value in the ground state can be absorbed by redefinition of the operator. This means a change of the zero of energy when $\hat{Q}$ is a part of the Hamiltonian. The redefinition is most conveniently accomplished by the "normal order" with respect to $\hat{\psi}_e$ and $\hat{\psi}_h$. The normally-ordered $\hat{Q}$ is

$$\begin{aligned} :\hat{Q}: &= \int_V d^3r \int_V d^3r'\ Q(\mathbf{r},\mathbf{r}') :\hat{\psi}^\dagger(\mathbf{r})\hat{\psi}(\mathbf{r}'): \\ &= \int_V d^3r \int_V d^3r'\ Q(\mathbf{r},\mathbf{r}')\Big[\hat{\psi}_e^\dagger(\mathbf{r})\hat{\psi}_e(\mathbf{r}') + \hat{\psi}_e^\dagger(\mathbf{r})\hat{\psi}_h^\dagger(\mathbf{r}') \\ &\qquad + \hat{\psi}_h(\mathbf{r})\hat{\psi}_e(\mathbf{r}') - \hat{\psi}_h^\dagger(\mathbf{r}')\hat{\psi}_h(\mathbf{r})\Big], \end{aligned} \quad (4.2.58)$$

which differs from $\hat{Q}$ only by a c-number term. Using (4.2.9) and the anticommutation relations, we can easily show that

$$\langle g| :\hat{Q}: |g\rangle = 0, \quad (4.2.59)$$

$$\langle x\eta| :\hat{Q}: |g\rangle = \int_V d^3r_1 \int_V d^3r_2\ \varphi_{x\eta}^*(\mathbf{R}_1,\mathbf{R}_2)Q(\mathbf{R}_1,\mathbf{R}_2), \quad (4.2.60)$$

$$\begin{aligned} \langle x\eta| :\hat{Q}: |x\eta'\rangle = &\int_V d^3r_1 \int_V d^3r_2 \int_V d^3r_3\ [\varphi_{x\eta}^*(\mathbf{R}_1,\mathbf{R}_3)Q(\mathbf{R}_1,\mathbf{R}_2)\varphi_{x\eta'}(\mathbf{R}_2,\mathbf{R}_3) \\ &- \varphi_{x\eta}^*(\mathbf{R}_3,\mathbf{R}_1)Q(\mathbf{R}_2,\mathbf{R}_1)\varphi_{x\eta'}(\mathbf{R}_3,\mathbf{R}_2)]. \end{aligned} \quad (4.2.61)$$

These formulas enable us to evaluate various matrix elements from the two-body wavefunction $\varphi_{x\eta}(\mathbf{R}_e, \mathbf{R}_h)$.

For an optical field whose photon energy is of the order of E_G or less, the wavelength λ of $\mathbf{A}$ is hundreds of nm or longer. When the spatial extension of $\varphi_{x\eta}(\mathbf{R}_e, \mathbf{R}_h)$ is smaller than λ we can approximate matrix elements of $:\mathcal{H}_I:$ as

$$\langle x\eta| : \mathcal{H}_I^{(1)} : |g\rangle \simeq -\frac{\hbar e}{imc}\mathbf{A}(\mathbf{R}_{x\eta}) \int_V d^3r \ \langle x\eta| : \hat{\psi}^\dagger(\mathbf{r}) \ \vec{\epsilon}\cdot\vec{\nabla} \ \hat{\psi}(\mathbf{r}) : |g\rangle, \tag{4.2.62}$$

$$\langle x\eta| : \mathcal{H}_I^{(2)} : |g\rangle \simeq \frac{e^2}{2mc^2}A^2(\mathbf{R}_{x\eta}) \int_V d^3r \ \langle x\eta| : \hat{\psi}^\dagger(\mathbf{r})\hat{\psi}(\mathbf{r}) : |g\rangle, \tag{4.2.63}$$

and similarly for other matrix elements. Here, $\mathbf{R}_{x\eta}$ denotes the center position around which $\varphi_{x\eta}$ is localized. We may use, for example, $\varphi_{N\nu}$ for $\varphi_{x\eta}$, and evaluate the above integrals using (4.2.60) by putting $Q(\mathbf{r},\mathbf{r}') = \vec{\epsilon}\cdot\vec{\nabla}\delta(\mathbf{r}-\mathbf{r}')$ and $Q(\mathbf{r},\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}')$, respectively. From (4.2.49), (4.2.50), and (4.2.51), we find

$$\begin{aligned}\int_V d^3r \ \langle N\nu| : \hat{\psi}^\dagger(\mathbf{r})\vec{\epsilon}\cdot\vec{\nabla}\hat{\psi}(\mathbf{r}) : |g\rangle &= \int_V d^3r \ \left[\vec{\epsilon}\cdot\vec{\nabla}'\varphi^*_{N\nu}(\mathbf{r},\mathbf{r}')\right]_{\mathbf{r}'\to\mathbf{r}} \\ &= \tilde{G}_{N\nu}(0)\sum_{\mathbf{k}}{}' \tilde{U}^*_\nu(\mathbf{k})(c\mathbf{k}| \ \vec{\epsilon}\cdot\vec{\nabla} \ |v\mathbf{k}),\end{aligned} \tag{4.2.64}$$

$$\int_V d^3r \ \langle N\nu| : \hat{\psi}^\dagger(\mathbf{r})\hat{\psi}(\mathbf{r}) : |g\rangle = \int_V d^3r \ \varphi^*_{N\nu}(\mathbf{r},\mathbf{r}) = 0, \tag{4.2.65}$$

where we have utilized the normalization conditions of the Bloch functions and Wannier functions, (4.2.31), and

$$(c\mathbf{k}| \ \vec{\epsilon}\cdot\vec{\nabla} \ |v\mathbf{k}) \equiv \frac{1}{v_0}\int_{v_0} d^3\rho \ u^*_{c\mathbf{k}}(\mathbf{r}) \ \vec{\epsilon}\cdot\vec{\nabla} \ u_{v\mathbf{k}}(\mathbf{r}) \tag{4.2.66}$$

represents contribution from each cell. If this integral is independent of $\mathbf{k}$ in the range of $\mathbf{k}$ for which $\tilde{U}_\nu(\mathbf{k})$ is not small, then we denote the integral by $(c|\vec{\epsilon}\cdot\vec{\nabla}|v)$, and (4.2.64) is reduced to

$$\int_V d^3r \ \langle N\nu| : \hat{\psi}^\dagger(\mathbf{r}) \ \vec{\epsilon}\cdot\vec{\nabla} \ \hat{\psi}(\mathbf{r}) : |g\rangle = \sqrt{V} \ \tilde{G}^*_{N\nu}(0)U^*_\nu(0)(c| \ \vec{\epsilon}\cdot\vec{\nabla} \ |v). \tag{4.2.67}$$

At first sight this result might look unphysical because of the anomalous factor $\sqrt{V}$. However, this factor is canceled by the factor $1/\sqrt{V}$ in $\tilde{G}_{N\nu}(0)$. This can be seen as follows. Recall that we have assumed that the exciton state is well localized, $\ell_G < \lambda$. For $\varphi_{N\nu}$ to be well localized $\tilde{G}_{N\nu}(\mathbf{k})$ must be a smooth function of $\mathbf{k}$, and we can replace the summation in (4.2.54) with the integral as

$$1 = \sum_{\mathbf{k}}{}' |\tilde{G}_{N\nu}(\mathbf{k})|^2 = \frac{V}{(2\pi)^3}\int d^3K \ |\tilde{G}_{N\nu}(\mathbf{k})|^2. \tag{4.2.68}$$

Considering also that the spatial extension of $G_{N\nu}$ is of the order of ℓ_G, we

find that the maximum value of $|\tilde{G}_{N\nu}(0)|$ should behave as

$$\max_N |\tilde{G}_{N\nu}(0)| \sim \sqrt{\frac{V_G}{V}} \tag{4.2.69}$$

where we have put, apart from an unimportant numerical factor of order unity, $V_G \sim \ell_G^3$. Therefore, the matrix element (4.2.67) is independent of V, as it should be. The matrix element is actually proportional to $\sqrt{V_G}$, which is sometimes called the coherence volume [4,5] because it gives a measure of the volume of the space region over which the exciton wavefunction is coherent. An important consequence of the V_G dependence of the matrix element is that the radiative recombination rate $\Gamma_{N\nu}^{\rm rad}$ of the exciton is proportional to V_G in the region where $\ell_G < \lambda$ [4,5]. This phenomenon has been observed in quantum wells [4] and in quantum dots [6,7], as the dependence of $\Gamma_{N\nu}^{\rm rad}$ on temperature (on which ℓ_ϕ depends) and on the dot size ($\simeq \ell_{\rm loc}$), respectively (see (4.2.55)).

Note that the result (4.2.67) can also be obtained when we neglect the $\mathbf{k}$ dependences of u_c and u_v^*

$$S_{\mathbf{k}\nu}(\mathbf{R}_e, \mathbf{R}_h) \simeq U_\nu(\mathbf{R}_e - \mathbf{R}_h) u_{c0}(\mathbf{R}_e) u_{v0}^*(\mathbf{R}_h) \equiv S_\nu^{(0)}. \tag{4.2.70}$$

In this approximation, the wavefunction $\varphi_{N\nu}$, for example, is decoupled into the envelope function of the center-of-mass motion, that of relative motion, and cell-periodic parts as

$$\varphi_{N\nu}(\mathbf{R}_e, \mathbf{R}_h) \simeq G_{N\nu}(\mathbf{r}) U_\nu(\mathbf{R}_e - \mathbf{R}_h) u_{c0}(\mathbf{R}_e) u_{v0}^*(\mathbf{R}_h) \equiv \varphi_{N\nu}^{(0)}(\mathbf{R}_e, \mathbf{R}_h). \tag{4.2.71}$$

This approximation is appropriate in simple nondegenerate cases and has been used in much literature because of its simplicity. In studies of nonlinear optical phenomena, however, we must also evaluate other matrix elements. For such general matrix elements, the approximate forms (4.2.70) and (4.2.71) often give wrong and unphysical results, as we now explain.

As a typical example let us investigate the transition matrix elements between different exciton states. When the spatial extension of both the initial and final states are smaller than λ, we find from formula (4.2.61)

$$\begin{aligned} \langle x\eta| : \mathcal{H}_I^{(1)} : |x\eta'\rangle \simeq & -\frac{\hbar e}{imc} \mathbf{A}(\mathbf{R}_{x\eta}) \int_V d^3 r_e \int_V d^3 r_h \\ & \times \varphi_{x\eta}^*(\mathbf{R}_e, \mathbf{R}_h) \vec{\epsilon} \cdot (\vec{\nabla}_e + \vec{\nabla}_h) \varphi_{x\eta'}(\mathbf{R}_e, \mathbf{R}_h). \end{aligned} \tag{4.2.72}$$

Here, one can arbitrarily replace $\mathbf{A}(\mathbf{R}_{x\eta})$ with $\mathbf{A}(\mathbf{R}_{x\eta'})$, because $\mathbf{A}(\mathbf{R}_{x\eta}) \simeq \mathbf{A}(\mathbf{R}_{x\eta'})$ from our assumption $\lambda < \ell_G$. The point is that derivatives of both the envelope functions and cell-periodic parts contribute to this matrix element. For the latter, we can use the group-velocity formula;

$$\int_V d^3 r \phi_{b\mathbf{k}}^*(\mathbf{r}) \frac{\hbar}{im} \vec{\nabla} \phi_{b\mathbf{k}}(\mathbf{r}) = \frac{1}{\hbar} \frac{\partial \varepsilon_b(\mathbf{k})}{\partial \mathbf{k}}, \qquad (b = c, v) \tag{4.2.73}$$

where m denotes the free-electron mass. For the parabolic dispersion (4.2.34), this yields

$$(c\mathbf{k}|\vec{\nabla}|c\mathbf{k}) = i\Big(\frac{m}{m_e}-1\Big)\mathbf{k}, \quad (v\mathbf{k}|\vec{\nabla}|v\mathbf{k}) = i\Big(-\frac{m}{m_h}-1\Big)\mathbf{k}. \tag{4.2.74}$$

It is seen that the derivatives of the cell-periodic parts can never be neglected unless $m_e = m$ and $m_h = -m$. For example, when $m_e \simeq m_h \simeq m$ the derivatives are $(c\mathbf{k}|\vec{\nabla}|c\mathbf{k}) \simeq 0$ and $(v\mathbf{k}|\vec{\nabla}|v\mathbf{k}) \simeq -2i\mathbf{k}$. Physically, these derivatives represent modification of velocities (whose operator is $\hbar\vec{\nabla}/im$ in the absence of electromagnetic fields) due to Bragg reflections by the crystal potential. The modification is larger for holes than for electrons because usually both m_e and m_h are positive. Use of the simplified forms (4.2.70) or (4.2.71) leads to the total neglect of these derivatives (because they vanish at $\mathbf{k}=0$), and often results in unphysical results. For example, when $\varphi_{x\eta} = \varphi_{N\nu}$ we find

$$\langle N\nu| : \mathcal{H}_I^{(1)} : |N'\nu'\rangle \simeq -\frac{\hbar e}{imc}\mathbf{A}(\mathbf{R}_{x\eta})\Big[\,\delta_{\nu,\nu'}\int_V d^3R\; G^*_{N\nu}(\mathbf{r})\;\vec{\epsilon}\cdot\vec{\nabla}\;G_{N'\nu}(\mathbf{r})$$
$$+\frac{m}{\mu}\int_V d^3R\; G^*_{N\nu}(\mathbf{r})G_{N'\nu'}(\mathbf{r})\int_V d^3r\; U^*_\nu(\mathbf{r})\;\vec{\epsilon}\cdot\vec{\nabla}\;U_{\nu'}(\mathbf{r})\Big]. \tag{4.2.75}$$

Note that the first and second terms of the right hand side become finite for different values of N, ν, N', and ν'. Note also that the typical value of the first term is much smaller than that of the second term because $a_B^* \ll \ell_G$ (by our assumption on $\varphi_{N\nu}$) and $\mu < m$ (e.g., $m/\mu \sim 10$ for GaAs). On the other hand, if we took $\varphi_{x\eta} = \varphi^{(0)}_{N\nu}$ we could obtain the first term only;

$$\langle N\nu^{(0)}| : \mathcal{H}_I^{(1)} : |N'\nu'^{(0)}\rangle \simeq -\frac{\hbar e}{imc}\mathbf{A}(\mathbf{R}_{x\eta})\delta_{\nu,\nu'}\int_V d^3R\; G^*_{N\nu}(\mathbf{r})\;\vec{\epsilon}\cdot\vec{\nabla}\;G_{N'\nu}(\mathbf{r}). \tag{4.2.76}$$

It is clear that the use of the simplified form $\varphi^{(0)}_{N\nu}$ would lead to wrong results for the matrix element, both in magnitude and in the selection rules. Therefore, we cannot use the simplified forms (4.2.70) and (4.2.71) for physical processes in which exciton scatterings by optical fields are involved. Such processes include, for example, second-order nonlinear responses and third-order processes in which the exciton scatterings by optical fields are involved as intermediate (virtual) states.

It seems worthwhile to point out another difficulty arising from the use of the simplified forms. Equations (4.2.74) suggest that terms $u_{b\mathbf{k}}$ play important roles in the gauge invariance of the theory. However, the use of the simplified forms is equivalent to the fixing of $u_{b\mathbf{k}}$ on u_{b0} which are defined in some particular gauge. As a result, the gauge invariance of such a theory is violated,* and physical predictions of the theory depend on the gauge [9,10].

* There is another possible origin of the violation of the gauge invariance of the theory: some nonlocal interaction, such as the exchange interaction, may appear as a result of a mean-field approximation [8]. In our case, however, such an interaction has been neglected because we are considering Wannier excitons.

No definite conclusions can be drawn from such a theory, of course.

Both of the above difficulties become severe if one proceeds from linear optical processes to nonlinear processes. Unfortunately, however, the simplified forms (4.2.70) and (4.2.71) have been used in much of literature discussing nonlinear optical processes, without any consideration on these difficulties. The most appropriate way of calculation is to use the correct wavefunctions (4.2.50) and (4.2.51) together with the original interaction Hamiltonian (4.2.56). There are alternative ways to use the simplified form (4.2.71). One way is to use the "length formula" for the interaction [8,9]:

$$\hat{\mathcal{H}}_I^{\text{length}} = e \int_V d^3r \; \hat{\psi}^\dagger(\mathbf{r}) \; \mathbf{r} \cdot \mathbf{E} \; \hat{\psi}(\mathbf{r}). \tag{4.2.77}$$

This expression is rather anomalous because the value of $\mathbf{r}$ extends to infinity as $V \to \infty$. Therefore, the normal-order form $: \hat{\mathcal{H}}_I^{\text{length}} :$ should be used in actual calculations. It is not difficult to show that this interaction together with the simplified form (4.2.71) give correct results. The other way is to perform the calculation according to the following prescription: In calculations of the envelope functions and matrix elements of intraband transitions (exciton-exciton scatterings), use the following Luttinger–Kohn Hamiltonian as an effective "intraband Hamiltonian" for the envelope function $G_{N\nu}U_\nu$ [9]:

$$H_{intra}^{env} \equiv \frac{1}{2m_e}\Big(\frac{\hbar}{i}\vec{\nabla}_e + \frac{e}{c}\mathbf{A}\Big)^2 + \frac{1}{2m_h}\Big(\frac{\hbar}{i}\vec{\nabla}_h - \frac{e}{c}\mathbf{A}\Big)^2 - \frac{e^2}{\epsilon|\mathbf{R}_e - \mathbf{R}_h|}. \tag{4.2.78}$$

In calculations of matrix elements of interband transitions (exciton creation and annihilation), on the other hand, one must use the bare interaction $\hat{\mathcal{H}}_I$. Physically, this prescription indicates that slowly varying phenomena are determined by the effective masses whereas rapidly varying phenomena are determined by the free-electron mass [1].

4.2.4 Quantum States in Nanostructures

The above formulation must be modified to treat nanostructures, such as quantum wells (QWs) and quantum wires (QWRs), because the band structure is modified by the quantum confinement and also because nanostructures are usually made of semiconductors which have degenerate valance bands. One way to take account of these effects is to combine the $\mathbf{k}\cdot\mathbf{p}$ theory with the exciton theory, which would enable us to determine (though perturbatively) both the band structure and the exciton wavefunction simultaneously.

Here, we describe another, simpler way – a rigid-band approximation. This approximation may be valid (i) in the case where the quantum confinement is much stronger than the exciton binding and also (ii) in the case where the quantum confinement is very weak and both the c and v bands are nondegenerate. We can take the band structure as that of the nanostructure and the bulk crystal for cases (i) and (ii), respectively, because the band structure would be rigid against the e-h interaction. The electron and hole operators can be defined according to the rigid band structure. We can then evaluate

exciton states taking the e-h interaction into account and evaluate various optical properties such as nonlinear susceptibilities.

In case (ii), in particular, the rigid bands are the bands of the bulk crystal, and the only important effect of the quantum confinement would be confinement of the center-of-mass motion. Hence we may employ $\varphi_{N\nu}$ of (4.2.51) as the exciton wavefunction, and the formulations of the previous subsections can be applied directly. A typical example is the $1s$ exciton in a quantum dot whose size is larger then a_B^* and which is made of non-degenerate semiconductors. For higher energy states ($2p$, $3d$, and so on), the weak-confinement assumption becomes worse because the spatial extension of the e-h relative motion becomes larger for higher energies.

In case (i), on the other hand, the band structure would be modified by the quantum confinement, but the modified band structure would not be affected strongly by the e-h interaction. In GaAs/AlGaAs nanostructures, for example, the two fold degeneracy at the top of the valance bands is lifted by the quantum confinement and two series of subbands are formed, which are called the heavy-hole subbands and light-hole subbands. The subbands are no longer degenerate at the top of the valence bands. When the intersubband energy separation is larger than the exciton binding energy, the e-h interaction would not alter the subband structure strongly and we may take the single-body eigenstates of the mean-field Hamiltonian of the nanostructure as basis functions. For concreteness, we assume for a nanostructure a quantum well structure of well width L and area S, and $a \ll L \ll a_B^*$ (a is the lattice constant). The cases of a quantum wire and quantum dot can be formulated similarly. Then, a single-body eigenstate (of low energies) may be approximated by the form

$$\phi_{bn\mathbf{k}_\parallel}(\mathbf{r}) = \sqrt{\frac{v_0}{S}} \sum_\ell \exp[i\mathbf{k}_\parallel \cdot \mathbf{R}_{\ell\parallel}] F_{bn}(Z_\ell) w_b(\mathbf{r} - \mathbf{R}_\ell; \mathbf{R}_\ell), \tag{4.2.79}$$

where n denotes a quantum number, $\mathbf{k}_\parallel$ and $\mathbf{R}_{\ell\parallel}$ denote 2D vectors representing, respectively, the wavenumber and the cell position in the QW plane (xy plane), and Z_ℓ is the z coordinate of the cell position, i.e., $\mathbf{R}_\ell = (\mathbf{R}_{\ell\parallel}, Z_\ell)$. Note that b here labels the bands which are modified (the degeneracy is lifted) by the quantum confinement and $w_b(\mathbf{r}-\mathbf{R}_\ell; \mathbf{R}_\ell)$ denotes the Wannier function of the modified band in the well (when $\mathbf{R}_\ell \in$ well region) or barrier (when $\mathbf{R}_\ell \in$ barrier region). The Wannier functions and the envelope function F_{bn} may be obtained, e.g., by using the $\mathbf{k}\cdot\mathbf{p}$ perturbations. In particular, for bands which are non-degenerate in the bulk crystal, we may approximate w_b by that of the bulk crystal and F_{bn} may be obtained by solving the effective-mass equation

$$\varepsilon_b\Big(\mathbf{k}_\parallel, -i\frac{\partial}{\partial z}; z\Big) F_{bn\mathbf{k}_\parallel}(z) = \varepsilon_{bn\mathbf{k}_\parallel} F_{bn\mathbf{k}_\parallel}(z), \tag{4.2.80}$$

where $\varepsilon_b(\mathbf{k}_\parallel, k_z; z)$ is the dispersion of the bulk band of the semiconductor which composes the well (when z is in the well region) or barrier (when z is

in the barrier region). For a parabolic band, in particular, (4.2.80) becomes

$$\Big[\frac{\hbar^2 k_{\|}^2}{2m_{b\|}(z)} - \frac{\hbar^2}{2m_{bz}(z)}\frac{\partial^2}{\partial z^2} + \varepsilon_b(0;z)\Big] F_{bn\mathbf{k}_{\|}}(z) = \varepsilon_{bn\mathbf{k}_{\|}} F_{bn\mathbf{k}_{\|}}(z). \tag{4.2.81}$$

Here we have considered the fact that in general the effective mass m_b becomes anisotropic by the band mixing. The $\mathbf{k}_{\|} \to 0$ limit of this equation is the familiar form:

$$\Big[-\frac{\hbar^2}{2m_{bz}(z)}\frac{\partial^2}{\partial z^2} + \varepsilon_b(0;z)\Big] F_{bn}(z) = \varepsilon_{bn} F_{bn}(z), \tag{4.2.82}$$

where $F_{bn} \equiv F_{bn0}$ and $\varepsilon_{bn} \equiv \varepsilon_{bn0}$. Equations (4.2.81) and (4.2.82) are basically particle-in-a-box problems. Note, however, that the boundary conditions on $F_{bn\mathbf{k}}$ or F_{bn} at heterojunctions depend on the functional forms of $w_b(\mathbf{r}-\mathbf{R}_\ell;\mathbf{R}_\ell)$. For example, when w_b takes different values across a heterojunction (because the semiconductors are different on both sides) then $F_{bn\mathbf{k}}$ must be discontinuous in order for the wavefunction to be continuous across the heterojunction [11].

For a typical GaAs/AlGaAs QW, the conduction band is non-degenerate and we may use (4.2.80), whereas the valance bands are degenerate and we should use another method such as the $\mathbf{k}\cdot\mathbf{p}$ perturbation. However, (4.2.80)–(4.2.82) have been widely used even for the valance bands. This may be justified if we replace w_b with that of the QW (i.e., w_b which is evaluated with the $\mathbf{k}\cdot\mathbf{p}$ perturbation) and simultaneously adjust parameters in $\varepsilon_b(\mathbf{k}_{\|}, k_z; z)$ of (4.2.80) in such a way that the wavefunction agrees with the result of the $\mathbf{k}\cdot\mathbf{p}$ perturbation.

We can also construct the wavefunction of an exciton, using the Wannier functions of QW, as

$$\varphi_{x\eta}(\mathbf{R}_e, \mathbf{R}_h) = v_0 \sum_j \sum_\ell F_{x\eta}(\mathbf{R}_j, \mathbf{R}_\ell) w_c(\mathbf{R}_e - \mathbf{R}_j; \mathbf{R}_j) w_v^*(\mathbf{R}_h - \mathbf{R}_\ell; \mathbf{R}_\ell). \tag{4.2.83}$$

The envelope function $F_{x\eta}(\mathbf{R}_e, \mathbf{R}_h)$ can be evaluated from, say, an effective-mass equation. Using the $\varphi_{x\eta}$ thus obtained, we can evaluate various optical matrix elements in a similar manner to the previous subsections. The most important point in this case is that the Wannier functions (and therefore the cell periodic functions) become anisotropic when the band degeneracy is lifted by the quantum confinement [12]. For example, when the valance band degeneracy is lifted, the matrix elements between the cell periodic functions become anisotropic, and, for small k,

$$(v\mathbf{k}|\vec{\nabla}|v\mathbf{k}) = i\Big(-\frac{m}{m_h(\hat{\mathbf{k}})} - 1\Big)\mathbf{k}, \tag{4.2.84}$$

where $\hat{\mathbf{k}}$ denotes the unit vector parallel to $\mathbf{k}$. The interband matrix element also becomes anisotropic, which for small k we denote as

$$\lim_{\mathbf{k}\to 0}(c\mathbf{k}|\vec{\epsilon}\cdot\vec{\nabla}|v\mathbf{k}) = (c|\vec{\epsilon}\cdot\vec{\nabla}|v\hat{\mathbf{k}}). \tag{4.2.85}$$

These anisotropies of the cell-periodic parts have been observed as anisotropies of various optical properties of QWs. As in the case of the bulk crystal, optical matrix elements must be evaluated carefully. Linear as well as nonlinear optical properties of nanostructures have been calculated using the above formulation [13–17] and some of the results have been confirmed experimentally [18–20].

4.2.5 Quantum Optical Phenomena in Nanostructures

So far the photon field has been treated as a classical field, but sometimes quantum optical phenomena manifest themselves in nanostructures.

The first example is the quantum nondemolition measurement of the photon number using an electron interferometer composed of double quantum wires [21,22]. It was shown that one can measure the photon number of a quantum state of photons using the electron interferometer as a photodetector in such a way that the post-measurement state of the photons carries the same information on the photon number as the pre-measurement sate. When the pre-measurement state is an eigenstate of the photon number, in particular, the wavefunction is unchanged by the measurement, whereas the observer obtains information on the photon number. In contrast, when the observer uses conventional photodetector the information on the photon number is lost by the measurement: when he uses a photodiode, for example, the post-measurement photon state is the zero-photon state, which has no information on the pre-measurement state.

The second example is photon creation from a false vacuum of semiconductors [23]. In quantum field theory, particles are created when one or more parameters in the Lagrangian (or, equivalently, in the field equations) have time dependence even if the initial state is the vacuum state [24]. Consequently, photons may be created if material parameters varies as a function of time [25,26]. Conventional discussions on photon creation in condensed matter, which are based on phenomenological quantization of the macroscopic photon field, break down near the exciton resonance [23]. To treat the exciton resonance correctly, the use of a two-field model has been proposed, in which both photon field and exciton field are taken into account as microscopic degrees of freedom [23].

The third example is the radiative lifetime of an excited atom in an inhomogeneous and absorptive cavity [27]. The lifetime has been shown to depend on such material parameters as the absorption coefficient and the cavity structure. The fourth example is that of photon statistics of light emitting diodes at a low injection level [28].

References

1 R.S. Knox, *Theory of Excitons* (Solid State Physics, Academic Press, New York, 1963)
2 D.C. Reynolds and T.C. Collins, *Excitons – Their Properties and Uses* (Academic Press, New York, 1981)
3 S. Nakajima, Y. Toyozawa, and R. Abe, *The Physics of Elementary Excitations* [Springer Series in Solid-State Sciences Vol. 12] (Springer, Berlin, 1980)
4 J. Feldmann et al., Phys. Rev. Lett. **59**, 2337 (1987)

5 E. Hanamura, Phys. Rev. B **38**, 1228 (1988)
6 A. Nakamura, H. Yamada, and T. Tokizaki, Phys. Rev. B **40**, 8585 (1989)
7 T. Itho et al., Solid State Commun. **73**, 271 (1990)
8 A.F. Starace, Phys. Rev. A **3**, 1242 (1971)
9 E. De Salvo, R. Girlanda, and A. Quattropani, Nuovo Cimento **5**, 63 (1985)
10 A. Pasquarello and A. Quattropani, Phys. Rev. B **38**, 6206 (1988)
11 T. Ando and S. Mori, Surf. Sci. **113**, 124 (1982)
12 M. Yamanishi and I. Suemune, Jpn. J. Appl. Phys. **23**, L35 (1984)
13 A. Shimizu, Phys. Rev. **B40**, 1403 (1989)
14 A. Shimizu, T. Ogawa, and H. Sakaki, Phys. Rev. B **45**, 11338 (1992)
15 T. Ogawa and A. Shimizu, Phys. Rev. B **48**, 4910 (1993)
16 A. Shimizu, M. Kuwata-Gonokami, and H. Sakaki, Appl. Phys. Lett. **61**, 399 (1992)
17 A. Shimizu and M. Yamanishi, Phys. Rev. Lett. **72**, 3343 (1994)
18 K. Fujii et al., Phys. Rev. Lett. **65**, 1808 (1990)
19 K. Tai et al., Phys. Rev. Lett. **62**, 1784 (1989)
20 R. Cingolani et al., Phys. Rev. Lett. **69**, 1276 (1992)
21 A. Shimizu, Phys. Rev. A **43**, 3819 (1991)
22 A. Shimizu, *Nanostructures and Quantum Effects*, edited by H. Sakaki and H. Noge [Springer Series in Material Science Vol. 31] (Springer, Berlin, 1994), p. 35
23 T. Okushima and A. Shimizu, Jpn. J. Appl. Phys. **34**, 4508 (1995); T. Okushima and A. Shimizu, unpublished
24 For a review, see for example, S.A. Fulling, *Aspects of Quantum Field Theory in Curved Space-Time* (Cambridge University Press, 1989)
25 E. Yablonovitch, Phys. Rev. Lett. **62** (1989) 1742
26 J. Schwinger, Proc. Natl. Acad. Sci. USA **90** (1993) 958
27 K. Koshino and A. Shimizu, Phys. Rev. A **53**, 4468 (1996)
28 H. Fujisaki and A. Shimizu, J. Phys. Soc. Jpn. **66**, 34 (1997); H. Fujisaki and A. Shimizu, Phys. Rev. A (submitted for publication)

4.3 Electron-Photon Interaction in Microcavities

4.3.1 Concept of Spontaneous Emission Control

a. Outline

Spontaneous emission is a physical phenomenon observed not only in laboratories but also in daily life where various materials emit light spontaneously. The typical phenomenon is the fluorescence of materials excited in some way. Spontaneous emission is known as an elementary process which is required, together with the process of stimulated emission, to establish thermal equilibrium between matter and a light field. In free space, spontaneous emission is an irreversible process, i.e., once a photon is emitted by an atom, it never returns to the emitting atom and, ultimately, escapes to remote space or is dissipated in obstacles. Spontaneous emission occurs when an atom is excited to a high-lying level and then releases the energy in the form of light. The emission of light is accompanied by a decay to a lower-lying level.

Classically, emission of light has its origin in an oscillating electric dipole or some other forms of electric oscillation. However, in an atom every energy level corresponds to a stable solution of the Schrödinger equation that describes the atom, which means that the atom in an excited state can exist at that level for a long time if no perturbation is applied. In reality, however, an excited atom decays in a finite time. The explanation requires a quantized description of the light field. The origin of spontaneous emission is the so-called vacuum fluctuation, which is a quantum-mechanical concept. Even if there is no measurable light field in the free space surrounding the atom, there always exists an electromagnetic fluctuation in space. This has to exist by the requirement of the uncertainty principle. This fluctuation is accompanied by an oscillating electric field that can mix the wavefunctions of the atomic excited state and a lower-lying state. The mixture of the state wavefunctions creates an oscillating electric dipole, etc. This causes the atom to emit light with some finite probability. This emission process, caused by the vacuum fluctuation, is spontaneous emission. This is a quantum-mechanical effect that can be traced back to the uncertainty principle. The emitted energy of light in an emission process appears as a light quantum, called a photon, with an energy equal to the energy difference between the two energy states involved, divided by Planck's constant h.

In free space, the vacuum fluctuation can be thought to be composed of continuously distributed field modes where a mode may be described by its frequency ν, the propagation direction, and the polarization direction. Each mode has vacuum fluctuation energy $(1/2)h\nu$ and can drive the atom. Thus the spontaneous emission occurs uniformly in all directions if the atomic dipole has a random direction. As the number of modes to which the emitting atom can have access is virtually infinite, the atom is said to be coupled to an electromagnetic field reservoir, and the emission process is irreversible because there is no chance that the emitted photon will return to the atom from the reservoir. The phenomenon of spontaneous emission has long been considered to be an immutable characteristic of each atom or material medium that cannot be eliminated or altered. Only recently has the possibility of controlling

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

spontaneous emission been discussed and become a subject of investigation in laboratories. The central concept for controlling spontaneous emission is to change the vacuum fluctuation field in its spatial and/or spectral content.

A means to change vacuum fluctuation is to use a cavity that is very small and can confine the emitted photon. That is, one confines the emitting atom and the emitted photon in a small cavity structure. Under the presence of the cavity walls the characteristics of the vacuum fluctuation are altered from that in free space and the resulting spontaneous emission process is also changed in various ways. A simple way to interpret this change in spontaneous emission is to trace the path of photon that is reflected at the walls and goes back to the emitting atom many times, to cause the atom to emit differently than in free space. As a result of the effects of the cavity walls, both the angular distribution of the emitted photon and the decay rate with which the atom decays to a lower-lying state are altered.

The stimulated emission process, which is another fundamental elementary process of light emission, is deeply related to the process of spontaneous emission and can also be altered by cavity wall effects. The stimulated emission rate in a given field mode increases in proportion to the spontaneous emission rate in that mode.

These effects are useful for improving lasers where usually spontaneous emission is a cause of energy loss and noise. Microcavity lasers are attracting attention as ultra-low threshold lasers. Ideally, atoms in a microcavity can couple with a single mode of the optical field. Therefore, there will be no loss of energy due to emission into unwanted field modes. This is considered to result in a thresholdless laser. As a matter of fact, coupling to only one mode is difficult to realize. The important parameter in a microcavity is therefore the coupling coefficient of the atom to the lasing mode, which we will call the spontaneous emission coupling ratio β.

Another important potential of a microcavity is to increase the total spontaneous emission rate as compared with the rate in free space. Since in a single mode the spontaneous emission rate is proportional to the stimulated emission rate, this increased emission rate can lead to enhanced laser activity. We call the ratio of the total spontaneous emission rate in a microcavity to that in free space the spontaneous emission enhancement factor and designate it as F.

As discussed in Sect. 4.3.2 several types of microcavities have been investigated so far. They have differing β, F, and Q values.

b. Spontaneous Emission in a Single Mode

Here we consider the interaction of an atom and a single mode of the electromagnetic field. Assume that there is only one mode of the field in a space of volume V. The mode function $\mathbf{U}(\mathbf{r})$ which describes the spatial distribution of the field may be normalized so that it satisfies

$$\int_V \varepsilon_0 |\mathbf{U}(\mathbf{r})|^2 dv = 1, \tag{4.3.1}$$

where ε_0 is the vacuum dielectric constant. Thus for a standing wave and for a traveling wave we have

$$\begin{aligned} \mathbf{U}(\mathbf{r}) &= \sqrt{\frac{2}{\varepsilon_0 V}}\,\sin(\mathbf{k}\cdot\mathbf{r}), \\ \mathbf{U}(\mathbf{r}) &= \sqrt{\frac{1}{\varepsilon_0 V}}\,\exp(i\mathbf{k}\cdot\mathbf{r}). \end{aligned} \tag{4.3.2}$$

For the field mode quantization procedure the reader is referred to Sect. 4.2. The electric field operator is

$$\hat{\mathbf{E}} = i\sqrt{\frac{\hbar\omega}{2}}\,[\mathbf{U}(\mathbf{r})\hat{a} - \mathbf{U}^*(\mathbf{r})\hat{a}^+], \tag{4.3.3}$$

where a and a^+ are the annihilation and creation operators, respectively, of the field mode. The Hamiltonian for the field mode of frequency $\omega = 2\pi\nu$ is

$$\hat{H}_f = \hbar\omega\Big(a^+a + \frac{1}{2}\Big). \tag{4.3.4}$$

The eigenstate $|n\rangle$ of the Hamiltonian satisfies

$$\hat{H}_f|n\rangle = \hbar\omega\Big(n + \frac{1}{2}\Big)|n\rangle, \tag{4.3.5}$$

where n is the quantum number corresponding to the measurable number of photons in the state and $\hbar\omega = h\nu$ is the photon energy.

The atom may be modeled as a two-level system with the Hamiltonian $\hat{H}_a$ having eigenvectors $|u\rangle$ and $|\ell\rangle$. The states $|u\rangle$ and $|\ell\rangle$ are the upper and the lower energy states, respectively, of the atom. They satisfy

$$\hat{H}_a|u\rangle = E_a|u\rangle, \qquad \hat{H}_a|\ell\rangle = E_\ell|\ell\rangle, \tag{4.3.6}$$

with the energy difference $E_u - E_\ell = \hbar\omega_A$.

Under the dipole approximation the interaction of the atom and the field mode is described by the interaction Hamiltonian

$$\hat{H}_i = -\hat{\mathbf{d}}\cdot\hat{\mathbf{E}}(\mathbf{r}_A) = -i\sqrt{\frac{\hbar\omega}{2}}\,[\mathbf{U}(\mathbf{r}_A)\hat{a} - \mathbf{U}^*(\mathbf{r}_A)\hat{a}^+], \tag{4.3.7}$$

where $\hat{\mathbf{d}}$ is the electric dipole operator of the atom and $\mathbf{r}_A$ is the location of the atom. We assume that the atom is initially (at $t = 0$) at the upper energy level and the field mode is in the $n = 0$ state, i.e., the vacuum state. The interaction (4.3.7) will alter the states of both the atom and the field for $t > 0$. Assuming that the total energy of the atom-field system is conserved, the total wave function of the system may be written as

$$|\Psi(t)\rangle = C_u(t)|u\rangle|0\rangle e^{-i\omega_A t} + C_\ell(t)|\ell\rangle|1\rangle e^{-i\omega t}, \tag{4.3.8}$$

where $|u\rangle|0\rangle$ denotes the initial state at $t = 0$ while $|\ell\rangle|1\rangle$ denotes the state

where the atom is in the lower state and the field mode is in the first excited state with $n=1$. The factors C_u and C_ℓ are respective probability amplitudes with the initial conditions $C_u(0)=1$ and $C_\ell(0)=0$. In (4.3.8) we have set $\mathbf{E}_\ell = 0$ and omitted the vacuum field energy $(1/2)\hbar\omega$ because its inclusion does not alter the results.

The Schrödinger equation for the coupled system reads

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = (\hat{H}_a + \hat{H}_f + \hat{H}_i)|\Psi(t)\rangle. \tag{4.3.9}$$

Equations (4.3.8) and (4.3.9) yield

$$\begin{aligned} \dot{C}_u(t) &= -\sqrt{\frac{\omega}{2\hbar}}\,\mathbf{d}\cdot\mathbf{U}(\mathbf{r}_A)e^{-i(\omega-\omega_A)t}C_\ell(t), \\ \dot{C}_\ell(t) &= \sqrt{\frac{\omega}{2\hbar}}\,\mathbf{d}^*\cdot\mathbf{U}^*(\mathbf{r}_A)e^{i(\omega-\omega_A)t}C_u(t), \end{aligned} \tag{4.3.10}$$

where $\mathbf{d}=\langle u|\hat{\mathbf{d}}|\ell\rangle$ is the electric dipole matrix element of the atom. Under the above initial conditions these equations can be solved to yield the probability $P_\ell(t)$ of finding the system in the state $|\ell\rangle|1\rangle$. We find

$$P_\ell(t) = |C_\ell(t)|^2 = \frac{\Omega_0^2}{\Omega^2}\sin^2\left(\frac{\Omega}{2}t\right), \tag{4.3.11}$$

where the frequency Ω is called the vacuum Rabi frequency and

$$\Omega = \sqrt{\Omega_0^2+(\omega-\omega_A)^2} = \sqrt{\frac{2\omega}{\hbar}|\mathbf{d}\cdot\mathbf{U}|^2+(\omega-\omega_A)^2}. \tag{4.3.12}$$

We see that the probability that the atom is in the lower state and the field is in one-photon state oscillates with a fixed frequency. The important point is that, if the field has only one mode, the emission process is reversible, which can be seen if we note that $|C_u(t)|^2 = \Omega^{-2}[\Omega_0^2\cos^2(\Omega t/2)+(\omega-\omega_A)^2]$. The last equation implies that the initial state of the atom-field system is recovered periodically. This is called the vacuum Rabi oscillation, which will be discussed again in Sect. 4.3.3.

c. Spontaneous Emission in Free Space

Next, we consider the case where the atom is located in free space that has a continuum of field modes. In this case we will label the modes by j. Then the electric field operator is rewritten as

$$\hat{\mathbf{E}} = i\sum_j\sqrt{\frac{\hbar\omega_j}{2}}\,[\mathbf{U}_j(\mathbf{r})\hat{a}_j - \mathbf{U}_j^*(\mathbf{r})\hat{a}_j^+]. \tag{4.3.13}$$

In this case the atom can emit into various field modes so that the system mode function will read, instead of (4.3.8),

$$|\Psi(t)\rangle = C_u(t)|u\rangle|0\rangle e^{-i\omega_A t} + \sum_j C_{\ell j}(t)|\ell\rangle|1_j\rangle e^{-i\omega_j t}. \tag{4.3.14}$$

Here, the symbol $|0\rangle$ now denotes the state of the field where no photon exists in any of the field modes and the symbol $1_j\rangle$ denotes the state where one photon is in the jth mode and no photons are in the other modes. The initial condition is $C_u(0)=1$ and $C_{\ell j}(0)=0$. Also, (4.3.10) becomes

$$\begin{aligned}\dot{C}_u(t) &= -\sum_j \sqrt{\frac{\omega_j}{2\hbar}}\,\mathbf{d}\cdot\mathbf{U}_j(\mathbf{r}_A)e^{-i(\omega_j-\omega_A)t}C_{\ell j}(t),\\ \dot{C}_{\ell j}(t) &= \sqrt{\frac{\omega_j}{2\hbar}}\,\mathbf{d}^*\cdot\mathbf{U}_j^*(\mathbf{r}_A)e^{i(\omega_j-\omega_A)t}C_u(t),\end{aligned} \tag{4.3.15}$$

This time we do not have simple solutions because of the infinite sum over j. A standard recipe for this problem is to assume that the interaction is weak enough to regard the effects of the field modes on the atom as small perturbations. Then the probability amplitude of the initial state C_u is set to be $C_u(t)=1$ and the second of (4.3.15) is integrated to obtain

$$C_{\ell j}(t) = \sqrt{\frac{\omega_j}{2\hbar}}\mathbf{d}^*\cdot\mathbf{U}_j^*(\mathbf{r}_A)\frac{e^{i(\omega_j-\omega_A)t}-1}{i(\omega_j-\omega_A)}. \tag{4.3.16}$$

Thus the probability of the atom being in the lower state is

$$P_\ell(t) = \sum_j |C_{\ell j}(t)|^2 = \sum_j \frac{\pi\omega_j}{\hbar}|\mathbf{d}\cdot\mathbf{U}_u(\mathbf{r}_A)|^2\delta(\omega_j-\omega_A)\,t, \tag{4.3.17}$$

where the last equation applies for $t\to\infty$. By differentiating with time t we have a time-independent transition rate A which is the spontaneous emission rate:

$$A = \sum_j \frac{\pi\omega_j}{\hbar}|\mathbf{d}\cdot\mathbf{U}_j(\mathbf{r}_A)|^2\delta(\omega_j-\omega_A). \tag{4.3.18}$$

This is an example of Fermi's golden rule.

In a free space with light velocity c the density of modes per unit volume is given by

$$\rho(\omega)d\omega = \frac{\omega^2}{\pi c^3}d\omega. \tag{4.3.19}$$

Therefore, if the direction of the atomic dipole is random, the spontaneous emission rate A_0 in free space is

$$A_0 = \frac{\omega_A^3}{3\pi\varepsilon_0\hbar c^3}|\mathbf{d}|^3, \tag{4.3.20}$$

where we have set $\sum_j = V\int\rho(\omega)d\omega$ in (4.3.18).

d. Controlled Spontaneous Emission

Equation (4.3.18) suggests that the spontaneous emission rate A can be modified if (1) the density of modes is changed at $\omega_j=\omega_A$ or (2) the mode amplitude $\mathbf{U}_j(\mathbf{r}_A)$ is changed at $\omega_j=\omega_A$. The first idea of changing the density

of field modes (number of modes per unit volume per unit angular frequency range) was proposed as early as 1946 by Purcell [1] who considered radio wave emission in a cavity of quality factor Q and of volume V. Since there is one field mode in the volume V and in the frequency spread ω/Q in this case, the effective mode density is

$$\rho_c(\omega)d\omega = \frac{1}{(\omega/Q)V}d\omega = \frac{Q}{\omega V}d\omega. \tag{4.3.21}$$

Therefore, the spontaneous emission rate will be enhanced by a factor F given by

$$F = \frac{\rho_c(\omega)}{\rho(\omega)} = \frac{\lambda^3 Q}{8\pi V}. \tag{4.3.22}$$

Purcell obtained a similar result with a numerical factor $3/4\pi^2$ instead of $1/8\pi$. Kleppner [2] proposed suppression of spontaneous emission in a microwave waveguide where density of modes, as derived from the field dispersion relation, is very small below the cutoff frequency. The second way of changing the spontaneous emission rate, i.e., changing the mode amplitude $\mathbf{U}_j(\mathbf{r}_A)$ at $\omega_j = \omega_A$, applies when the confining cavity is partially transparent and the field inside the cavity is continued to that of free space by the boundary conditions imposed by the semi-transparent mirrors of the cavity. In this case one assumes that the cavity is embedded in an infinitely large, imaginary box (quantization box) and that the density of modes is dictated by the quantization box and is given by (4.3.19). The field amplitude outside the cavity will be the same as that in free space. However, the amplitude inside the cavity is altered so as to be consistent with the boundary conditions at the semi-transparent mirrors. Thus the spontaneous emission will be enhanced or suppressed according to the magnitude of $\mathbf{U}_j(\mathbf{r}_A)$ at $\omega_j = \omega_A$.

This idea was pursued by DeMartini et al. [3], Ujihara et al. [4], and Björk et al. [5] all for a planar, Fabry–Perot type structure of the microcavity. The former two attacked the problem by quantum-mechanical method outlined above while the latter took a classical approach. The problem here is to obtain the expressions for the mode functions $\mathbf{U}(\mathbf{r})$ that satisfy the mirror boundary conditions. For a planar microcavity, (4.3.18) results in the following form of spontaneous emission rate

$$A = \sum \frac{A_0}{4\pi} \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \, \sin\theta \, \frac{3}{8}\varepsilon_0 V |\mathbf{U}(\mathbf{r}_A)|^2, \tag{4.3.23}$$

where θ and ϕ are polar coordinates with the z axis perpendicular to the planar mirrors and the summation is over the polarization and symmetry of the field modes [4]. Inside the planar microcavity the function $\mathbf{U}(\mathbf{r}_A) = \mathbf{U}_{\phi,\theta,\omega_A}(\mathbf{r}_A)$ reflects the characteristics of a Fabry–Perot resonance. If the cavity is resonant at ω_A its magnitude is large in the direction normal to the mirrors ($\theta = 0$) and is small in other directions, resulting in an emission lobe normal to the mirrors. This lobe may be thought to be released from a cavity resonance mode.

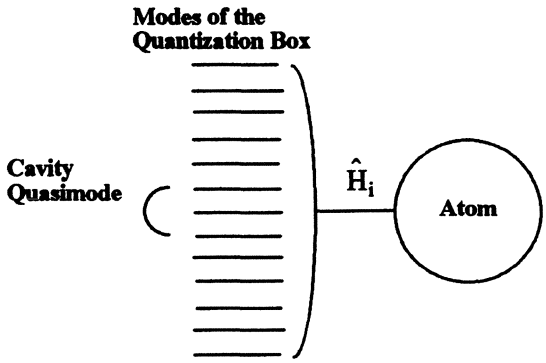


Fig. 4.3.1 Concept of the cavity quasimode

Here the concept of cavity resonance mode is useful. This mode is not a true eigenmode of the cavity because it has a finite loss due to the mirror transmission and it is sometimes called a quasimode. Figure 4.3.1 shows the relation between the modes of the quantization box and the cavity quasimode. We can define the resonance frequency $\omega_c = m\pi c/\ell$, resonance band width $\gamma = (c/2n\ell)\ln(1/R)$, and mode volume of the quasimode $V = \pi r_{\text{eff}}^2 \ell$. In particular the quasimode radius is given by

$$r_{\text{eff}} = \sqrt{\frac{\pi/\lambda}{8n(1-R)}} = \frac{\lambda}{4n}\sqrt{\frac{m\pi}{1-R}} = \frac{\lambda}{4n}\sqrt{Q}, \tag{4.3.24}$$

where ℓ is the cavity length (mirror separation), λ is the resonant wavelength measured outside the cavity, m is the order of longitudinal resonance, n is the refractive index of the medium inside the cavity relative to the index outside the cavity, and R is the power reflectivity of the mirrors. This quasimode radius is defined by the angular spread $\Delta\theta$ of the spontaneous emission lobe through the diffraction relation $\Delta\theta = \lambda/2r_{\text{eff}}$. If the atom is at an antinode of the resonance and if the atomic frequency ω_A is equal to the cavity resonance frequency, the emission rate in the direction normal to the cavity mirrors is enhanced by a factor of $4/(1-R)$.

The enhancement factor in this case can be obtained by integrating the angle-dependent emission rate, evaluating the contributions from the oblique resonance modes and the guided modes, and finally dividing the total rate by the free-space emission rate A_0 to find $F \sim 1$. If on the other hand we used (4.3.24) in (4.3.22) we would have $F = 4/\pi^2 m$ where $m = n\ell/(\lambda/2)$ is the order of the longitudinal resonance. Thus we have $F \sim 1/m$ in contradiction with the above result ($F \sim 1$) . This discrepancy is solved if we consider the presence of oblique resonance modes for $m \geq 2$.

The important factor β, the spontaneous emission coupling ratio, in a microcavity with relatively large volume V ($\gamma > \lambda^3$) can be obtained as the

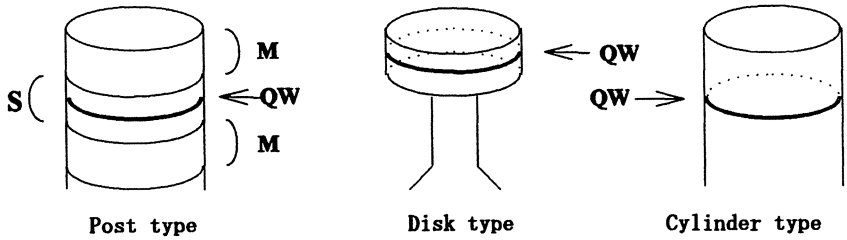


Fig. 4.3.2 Types of microcavity. M: multilayer mirror, S: spacer, QW: quantum well layer

inverse of the number of modes M in the volume V and within the fluorescence linewidth $\Delta\omega$ of the emitting atom. Thus using (4.3.19) we have

$$\beta = \frac{1}{M} = \frac{1}{\rho(\omega)V\Delta\omega} = \frac{\lambda^4}{8\pi n^3 V\Delta\lambda}. \tag{4.3.25}$$

In the case of a planar microcavity it can be shown that, for an atom with very small linewidth, the factor β is given by $\beta = \lambda/8n\ell$. For an atom with a linewidth greater than the cavity width, this factor should be multiplied by $\gamma/\Delta\omega$. Thus, we have

$$\beta = \frac{\lambda}{8n\ell}\frac{\gamma}{\Delta\omega}, \tag{4.3.26}$$

which reproduces (4.3.25) within a numerical factor, if (4.3.24) and the expression for γ are used.

4.3.2 Experimental Results of Spontaneous Emission Control

Microcavities studied so far include a planar type (a Fabry–Perot type) [3–5], a post type [6,7], a disk type [8,9], a cylinder type [10], a hemispherical type [11], and a spherical type [12]. Some of them are shown in Fig. 4.3.2. Microcavities in photonic band structures are also being investigated [13,14] and will be discussed in Sect. 4.4.

DeMatini et al. [15] claimed that they discovered zero-threshold oscillation using a planar half-wavelength cavity with dye sulforhodamine-640. Also, using DCM dye in a planar microcavity, they found the transverse correlation length with respect to the interaction of two operating regions on the cavity surface and interpreted it by the uncertainty principle [16]. This length is equivalent to the mode radius in (4.3.24). Björk et al. [5] considered the changes in the spontaneous emission characteristics in a planar microcavity and experimentally observed the angular pattern of the spontaneous emission lobe from a GaAs quantum well in a planar microcavity. Yokoyama et al. [17] found a spontaneous-emission dominant regime and a stimulated-emission dominant regime in the emission from Rhodamine 6G in a planar microcavity as distinguished by the emission time after pulsed excitation. Wang et al. [18] observed unpolarized spontaneous emission from a planar microcavity containing a dye solution. The emission changed to a linear polarization at

the threshold of oscillation. They determined the β value of their cavity to be about 10^{-3}.

Tezuka et al. [6] made a post type (a pillar type) microcavity after etching planar layers constituting a planar microcavity with one-wavelength cavity layer with a GaAs quantum well at the center. post type microcavitypillar type microcavityThe post diameter was 1.3–2 μm and the cavity resonance frequency changed from sample to sample. When the cavity resonance frequency coincided with the quantum well emission frequency, they observed a clear shortening of the emission time after a pulsed optical excitation, which can be explained as a spontaneous emission enhancement by the microcavity.

Levi et al. [8] constructed a disk-shaped microcavity where a multiple quantum well structure of InGaAs was built in parallel to the disk surface. The optical mode is the whispering-gallery mode which, by means of total internal reflection, travels along the circular edge of the disk. The disk diameter was 2–19 μm and the 0.1 μm thick semiconductor disk was supported by a semiconductor pedestal. By an electric current excitation a number of modes were observed to be emitted simultaneously.

Levi et al. [10] also made a cylinder-shaped microcavity similar in shape to the post type except that it had no planar mirrors. The optical mode here is also a whispering-gallery mode that travels along the edge of the InGaAs quantum wells that extend vertically to the cylinder axis. The 6 μm diameter quantum wells were excited by electric current. Emission of many whispering-gallery modes was observed. The mode spacing given by $\Delta\lambda = \lambda^2/\pi dn$ was much smaller than the emission bandwidth of the quantum well. Here d is the cylinder diameter and n is the refractive index of the cylinder.

A hemispherical microcavity was made by Matinaga et al. [11] using a planar semiconductor reflector and a curved dielectric mirror. The curvature of the mirror was 5–10 μm and the cavity length was 2.5 μm. An InGaAs quantum well layer was located inside the cavity. They found the β value to be 0.01 as compared to the theoretical value of 0.019.

Microspheres of liquid and dielectric materials have also been studied as very high-Q microcavities. Kuwata-Gonokami et al. [12] doped polystyrene microspheres of 10–92 μm diameter with the dye Nile Red and, by optical excitation, observed a broadband spontaneous emission corresponding to many lines of whispering-gallery modes which were more densely distributed than those in a microdisk or a microcylinder due to the spherical rather than circular symmetry.

4.3.3 Cavity-Polariton Effects

So far we have described spontaneous emission in a microcavity in the weak coupling regime, i.e., the regime where the perturbation approximation leading to (4.3.16)–(4.3.18) is appropriate. In this section we describe phenomena in strong coupling regime where reversibility, or vacuum Rabi oscillation, as described by (4.3.11) is relevant.

Spontaneous emission in a strong coupling regime exhibits oscillatory time developement which will result in a double-peaked emission spectrum. When

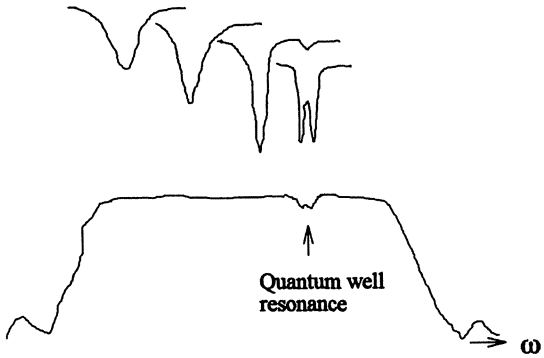


Fig. 4.3.3 Schematic drawing of the reflectivity curves with various detunings. After [19]

the number of atoms coupled to the field mode is N, and when they cooperate, the Rabi frequency is multiplied by $\sqrt{N}$. In a semiconductor quantum well microcavity discussed below, where excitons play the role of the atoms, a similar picture may apply. If the coupling is strong in these cases, the emission spectrum can also be double peaked. In a three-dimensional solid an exciton couples to a photon mode of the same wave vector and energy as the exciton itself. The resultant coupled excitation is called a polariton. In a planar microcavity containing a quantum well the photon mode has no wave vector component vertical to the mirrors. The exciton is also confined in a two dimensional quantum well parallel to the mirrors. The 2D exciton couples to the cavity photon mode with the same transverse wave vector to constitute a cavity polariton. Note that the cavity mode here is a quasimode. The cavity polariton is created when the resonance frequencies of the exciton and the cavity mode coincide and there appear two normal modes of slightly different frequency.

On the one hand, this frequency splitting is an indication of quantum-mechanical Rabi oscillation. On the other, the same splitting is classically explained as a result of normal mode splitting of coupled oscillators.

The splitting also appears in the absorption or reflection spectrum of the coupled oscillators. A classical explanation for the fact that two absorption or reflection peaks appear can be given by considering the Fabry–Perot characteristics including the phase shift associated with the optical wave propagation. This phase shift is influenced by the dispersion of the medium inside the cavity. In particular, when the medium has a resonance at the cavity resonance frequency, the absorption or transmission spectrum can have two peaks because of the medium dispersion. In the case of a cavity polariton it can be shown that the splitting is determined by the total oscillator strength of the exciton and the cavity size and that it is independent of the cavity mirror reflectivity [20].

Weisbuch et al. [19] performed an experiment to observe Rabi splitting in the reflection spectrum from a semiconductor planar microcavity containing

GaAs quantum wells as the active medium. They used a microcavity with mirrors of GaAlAs/AlAs quarter-wave layers and a one-wavelength cavity of GaAlAs in which the quantum well layers of GaAs were located. At a low temperature of 5 K they observed the reflection spectrum changing the detuning of the cavity from the exciton. When the detuning was large, a single dip was obtained at the cavity resonance frequency. When the detuning was zero, the dip showed a splitting that is characteristic of a two-mode coupling as depicted in Fig. 4.3.3. They explained the splitting as Rabi splitting. Norris et al. [21] observed Rabi oscillations in the time domain by use of the same microcavity sample as in [19].

References

1 E.M. Purcell, Phys. Rev. **69**, 681 (1946)
2 D. Kleppner, Phys. Rev. Lett. **47**, 233 (1981)
3 F. DeMartini, M. Marrocco, P. Mataloni, L. Crescentini, and R. Loudon, Phys. Rev. A **43**, 2480 (1991)
4 K. Ujihara, A. Nakamura, O. Manba, and X.P. Feng, Jpn. J. Appl. Phys. **30**, 3388 (1991)
5 G. Bjork, S. Machida, Y. Yamamoto, and K. Igeta, Phys. Rev. A **44**, 669 (1991)
6 T. Tezuka, S. Nunoue, H. Yoshida, and T. Noda, Jpn. J. Appl. Phys. **32**, L54 (1993)
7 T. Numai, T. Kawakami, T. Yoshikawa, M. Sugimoto, Y. Sugimoto, H. Yokoyama, K. Kasahara, and K. Asakawa, Jpn. J. Appl. Phys. **32**, L1533 (1993)
8 A.F. Levi, R.E. Slusher, S.L. McCall, T. Tanbun-Ek, D.L. Coblentz, and S.J. Pearton, Electron. Lett. **28**, 1010 (1992)
9 R.E. Slusher, A.F.J. Levi, U. Mohideen, S.L. McCall, S.J. Pearton, and R.A. Logan, Appl. Phys. Lett. **63** 1310 (1993)
10 A.F.J. Levi, R.E. Slusher, S.L. McCall, S.J. Pearton, and W.S. Hobson, Appl. Phys. Lett. **62**, 2021 (1993)
11 F.M. Matinaga, A. Karlsson, S. Machida, Y. Yamamoto, T. Suzuki, Y. Kadota, and M. Ikeda, Appl. Phys. Lett. **62**, 443 (1993)
12 M. Kuwata-Gonokami, K. Takeda, H. Yasuda, and K. Ema, Jpn. J. Appl. Phys. **31**, L99 (1992)
13 E. Yablonovitch, Phys. Rev.Lett. **58**, 2059 (1987)
14 H. Hirayama, T. Hamano, and Y. Aoyagi, Appl. Phys. Lett. **69**, 791 (1996)
15 F. DeMartini and G.R. Jacobovitz, Phys. Rev. Lett. **60**, 1711 (1988)
16 F. DeMartini, M. Marrocco, and D. Murra, Phys. Rev. Lett. **65**, 1853 (1990)
17 H. Yokoyama, M. Suzuki, and Y. Nambu, Appl. Phys. Lett. **58**, 2598 (1991)
18 X. Wang, R.A. Linke, G. Devlin, and H. Yokoyama, Phys. Rev. A **47**, R2488 (1993)
19 C. Weisbuch, M. Nishioka, A. Ishikawa, and Y. Arakawa, Phys. Rev. Lett. **69**, 3314 (1992)
20 C. Weisbuch, R. Houdre, and R.P. Stanley, *Spontaneous emission and laser oscillation in microcavities* edited by H. Yokoyama and K. Ujihara (CRC Press, Boca Raton, 1995), p. 109
21 T.B. Norris, J.-K. Rhee, C.-Y. Sung, Y. Arakawa, M. Nishioka, and C. Weisbuch, Phys. Rev. B **50**, 14663 (1994)

4.4 Photonic Crystals

4.4.1 Photonic Crystals and Spontaneous Emission Control

Photonic crystals are artificial nanostructures constructed from optical atoms arranged in a background medium with a period on the order of half the optical wavelength [1]. They are of great interest since those made of semiconductors have the possibility of spontaneous emission control, which allows thresholdless operation of laser diodes. A large refractive-index contrast between semiconductor and air provides a wide photonic band gap, which means a frequency range that inhibits the existence of modes. Figure 4.4.1 schematically illustrates a photonic crystal of various dimensions, and the corresponding wavevector space of cavity modes and that of the emission spectrum, both inhibited by each photonic band gap.

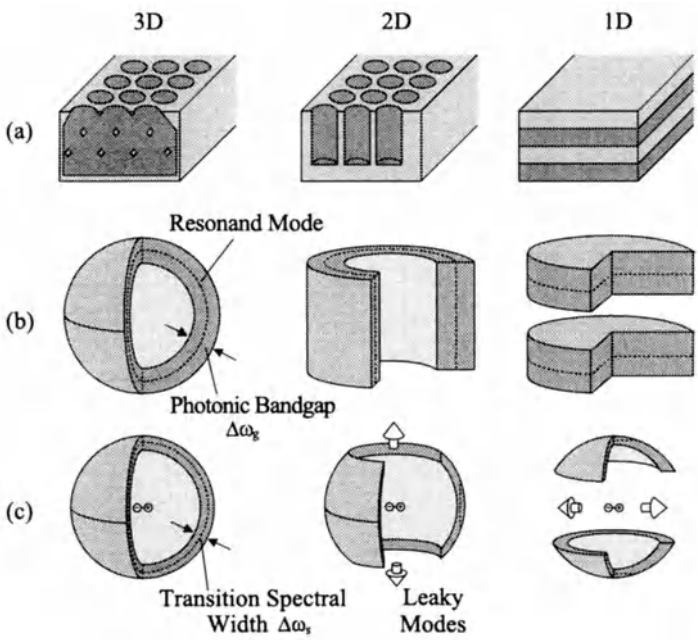


Fig. 4.4.1 Schematic structure of a photonic crystal (**a**), the wavevector space of modes inhibited by photonic band gap (**b**), and corresponding emission spectrum (**c**)

The 1D crystal is a simple multilayer stack and its photonic band gap is equivalent to the stopband of a multilayer filter. Many experiments have been performed using such 1D crystals [2–4] to avoid the difficult fabrication process for 2D and 3D. However, the inhibition of the emission is not perfect, because light oriented parallel to the multilayer plane is permitted. Typically, 2D crystals consist of vertical holes or pillars formed on a substrate. Although this permits the light to be oriented normally to the substrate plane, the total solid angle of the inhibited orientation is much larger than that in 1D. The

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

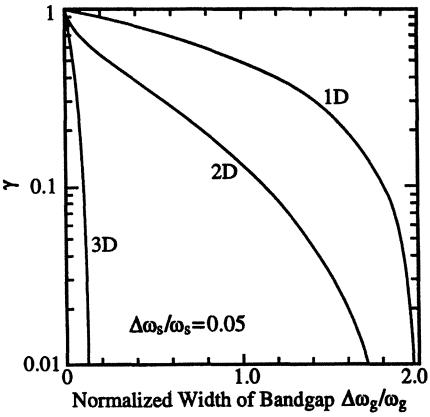


Fig. 4.4.2 Decrease of spontaneous emission rate by photonic band gap. $\Delta\omega_s/\omega_s$ is the normalized width of the transition spectrum determined by electron states. Typical values for GaInAsP quantum wells are assumed

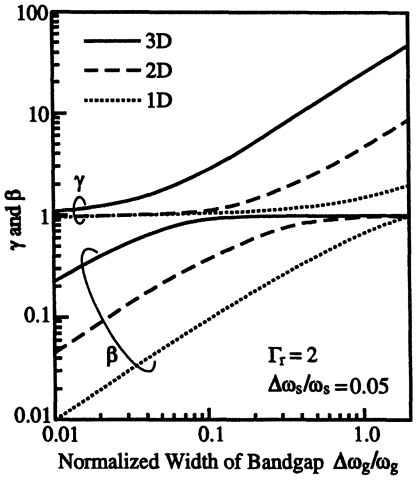


Fig. 4.4.3 Increase of spontaneous emission rate and spontaneous emission coupling efficiency in photonic crystals with a defect mode. Γ_r is the relative confinement factor indicating how efficiently the resonant mode overlaps with the active laser region

3D crystal is a complicated mosaic and all the orientations are inhibited in the ideal case.

Consider the condition that the center frequency of a photonic band gap is tuned to that of the emission spectrum. The spontaneous emission rate $R_{\rm sp}$ should become smaller by the decrease of allowed modes. Figure 4.4.2 shows the ratio of $R_{\rm sp}$ in photonic crystals to that in a large bulk semiconductor, which is denoted as γ. It is estimated from the ratio of the radiation solid angle of allowed light to the total solid angle 4π. It has been shown that such a simple estimation gives a good approximation for 1D crystals [5]. As seen in Fig. 4.4.2, the change of $R_{\rm sp}$ is more marked in 3D than in 1D and 2D.

When photonic crystals have an irregular element called defect [6], a resonant mode whose field is strongly localized around the defect appears. A large γ and a large coupling efficiency β of the spontaneous emission into such a defect mode are expected. Figure 4.4.3 shows rough estimations of γ and β in photonic crystals with a defect. Here, the polarization degeneracy and the change of mode effective index are ignored. The spontaneous emission energy coupled into the defect mode and the photonic band gap are approximated by the transmission energy of a resonant mode in a high-Q Fabry–Perot cavity and the free spectral range between resonant modes, respectively [7]. In a 3D crystal, γ is increased by a factor of 10 at $\Delta\omega_g/\omega_g = 0.3$, where the gap is assumed to be present in the frequency range $\omega_g - \Delta\omega_g/2 < \omega < \omega_g + \Delta\omega_g/2$. Because of the imperfect band gap in 1D and 2D, some amount of the emission couples into leaky modes widespread outside the band gap, and hence the change of γ is small. The efficiency β is calculated as the ratio of the emission

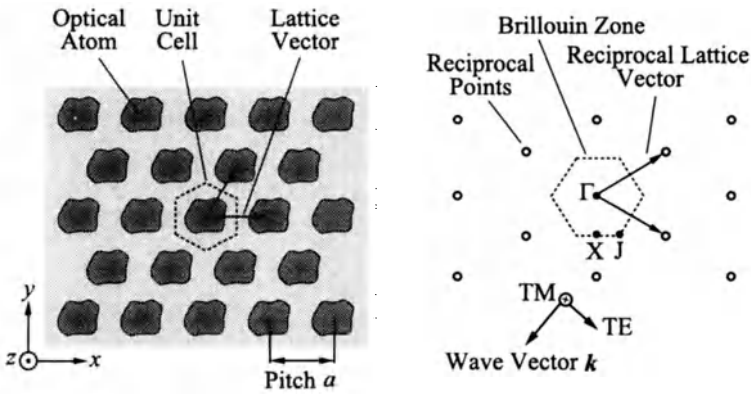


Fig. 4.4.4 Schematic veiw of 2D photonic crystal with triangular lattice (left) and corresponding reciprocal lattice (right)

energy coupled into defect modes to the total energy given by the sum of those coupled into the defect modes and the leaky modes. As seen in Fig. 4.4.3, $\beta \sim 1$ is possible in 3D crystals with $\Delta\omega_g/\omega_g > 0.1$ and $\beta > 0.6$ even in 2D crystals with $\Delta\omega_g/\omega_g = 0.15$. Spontaneous emission factor C [5] comparable to the large β will be realized in 3D crystals when the defect mode does not have a polarization degeneracy. It will also be realized in 2D if a cavity mode is defined by 3D boundaries and an unwanted polarization is suppressed. The 3D boundary will be realized by making a 2D crystal with a defect into a high index contrast multilayer. A nearly single-polarized emission can be obtained by the dipole anisotropy in strained quantum wells.

4.4.2 Band Structure of Photonic Crystals

The photonic band is the dispersion characteristic of light obtained by the Bloch wave approximation in infinite photonic crystals. Those for 1D crystals are easily obtained using a characteristic matrix method. On the other hand, the following photonic band calculation is used for 2D and 3D [8–10]. Similar to those for solid crystals, a periodic potential distribution (for light, refractive index distribution) $n^2(\mathbf{r}) = n^2(\mathbf{r}-\mathbf{R})$, a unit cell, a Brillouin zone, lattice vectors $\mathbf{R}$, and reciprocal lattice vectors $\mathbf{G}$ are considered, as illustrated in Fig. 4.4.4. The magnetic field $\mathbf{H}(\mathbf{r})$ is expressed as

$$\mathbf{H}(\mathbf{r}) = \sum_{\mathbf{G}} \mathbf{H}_{\mathbf{G}} \exp[i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}], \tag{4.4.1}$$

where $\mathbf{H}_{\mathbf{G}}$ is the Fourier component of $\mathbf{H}(\mathbf{r})$ and $\mathbf{k}$ is the wavevector inside the Brillouin zone. Substituting (4.4.1) into the wave equation for $\mathbf{H}(\mathbf{r})$ gives

$$(\mathbf{k}+\mathbf{G})\times\sum_{\mathbf{G}'}(\mathbf{k}+\mathbf{G}')\times\mathbf{H}_{\mathbf{G}'} + \omega^2\mathbf{H}_{\mathbf{G}} = 0, \tag{4.4.2}$$

where $\mathbf{h}_{\mathbf{G}-\mathbf{G}'}$ is the Fourier component of $n^2(\mathbf{r})$. Equation (4.4.2) is the

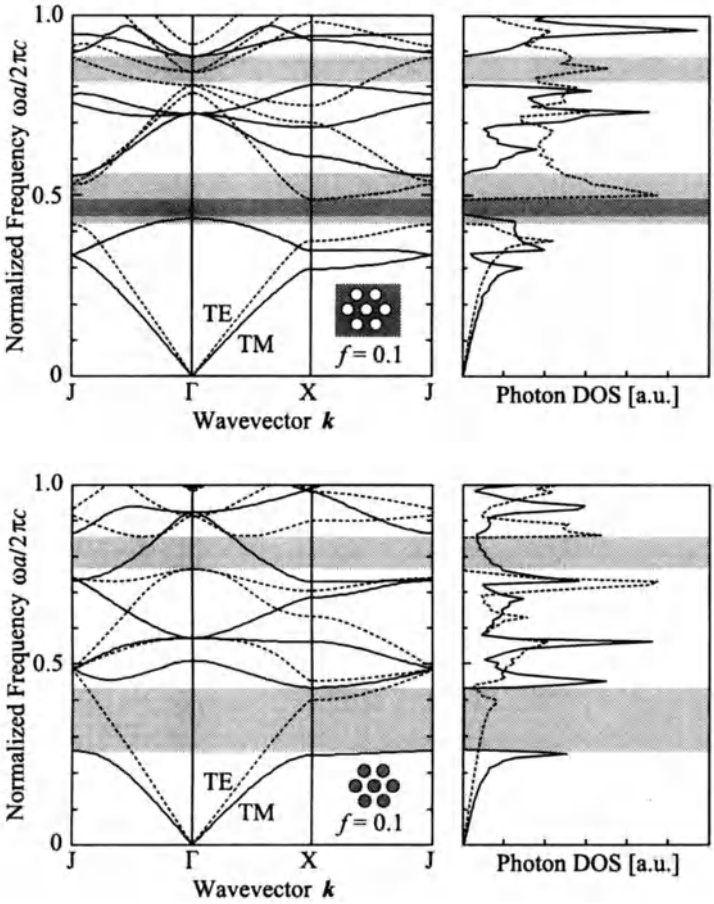


Fig. 4.4.5 Photonic band diagrams calculated for crystals shown in insets and corresponding photon DOS. Dark and bright regions indicate semiconductor ($n = 3.46$) and air, respectively

eigenvalue equation with respect to $\mathbf{H_G}$. Photonic bands and magnetic field $\mathbf{H}(\mathbf{r})$ are obtained by solving this equation numerically with a sufficiently large number of $\mathbf{G}'$. Photonic bands can also be obtained from the similar eigenvalue equation for the electric field $\mathbf{E}(\mathbf{r})$. For this case, however, the numerical task is more difficult due to the complicated form of the equation.

Consider a 2D crystal constructed by circular optical atoms arranged as shown in Fig. 4.4.4. When $\mathbf{k}$ is inside the xy plane, photonic bands are obtained individually for the xy components (TM polarization) and z component (TE polarization) of $\mathbf{H_G}$, as shown in Fig. 4.4.5, where the filling fraction f of the semiconductor is fixed to 0.1. For circular holes, the overlap of band gaps for two polarizations is observed, whose normalized width $\Delta\omega_g/\omega_g$ is 0.1. For columns, $\Delta\omega_g/\omega_g$ of the lowest TM band gap is as wide as 0.46. As seen in Fig. 4.4.3, $\beta > 0.9$ is expected for this band gap if a resonant mode is pro-

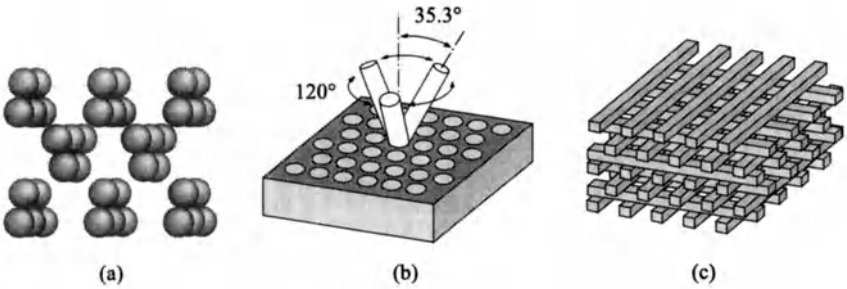


Fig. 4.4.6 Typical structures of 3D photonic crystal with diamond lattice. (**a**) Spherical atoms, (**b**) symmetric holes (Yablonovite structure), and (**c**) stacked rods

duced by one defect and the polarization is restricted to TM by some dipole anisotropy in strained quantum wells. The photon density of states (DOS) for each polarization is also obtained by integrating $dk/d\omega$ inside the Brillouin zone. The distribution of band gaps for various atoms and various filling fractions f has been examined [10,11]. In general, structures constructed from holes are likely to exhibit TE band gaps and those from pillars TM ones. These originate from the abrupt change of the displacement field in continuous structures like holes and that of the magnetic field in discrete structures like pillars, respectively, at band gaps [10]. The absolute band gap occurs for circular holes with $0.09 < f < 0.2$, and $\Delta\omega_g/\omega_g$ is maximally 0.15 for $f = 0.13$. For columns, several wide TM band gaps and narrow TE band gaps occur for a wide range of f. For hexagonal atoms, the distribution of band gap is similar to that for circular atoms. However, an absolute band gap is not observed, since hexagonal holes do not separate the semiconductor with any fraction f. For square holes, TE and TM band gaps exist simultaneously but do not overlap with each other.

Similar calculation can be performed for 3D crystals. For this case, (4.4.2) is solved directly without the separation for polarizations. It has been found that no absolute band gaps are opened for the fcc lattice but a large band gap with $\Delta\omega_g/\omega_g < 0.3$ is opened for the diamond lattice. Typical structures with the diamond lattice, as illustrated in Fig. 4.4.6, have been proposed [12–14]. Symmetric holes (Yablonovite structure) and stacked rods allow maximally $\Delta\omega_g/\omega_g = 0.19$.

To simulate defect modes in 2D and 3D crystals, a variational method [10], a super-cell method [15,16], a Green's function method [17], and a time domain method [18] have been developed. Mode frequencies, field distribution, Q factor of resonance, transmission intensity, and temporal behavior can be obtained by these methods. However, there are no reports on precise evaluations of β or C factors, which require a large amount of calculation for all modes including defect modes and leaky modes.

4.4.3 Technologies of Photonic Crystals

Some trial fabrications of 2D and 3D crystals using etching [19–21] and selective growth [22] techniques have been reported. However, the emission

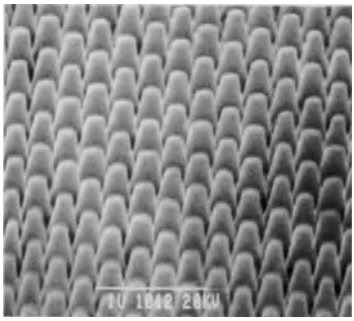


Fig. 4.4.7 Formed columns of 0.2 μm in diameter and 0.6 mm in height

characteristic is difficult to evaluate for GaAs/AlGaAs systems due to the serious surface recombination at etched sidewalls. The GaInAsP/InP system is advantageous since it has a surface recombination velocity 1–2 orders of magnitude slower than that of GaAs/AlGaAs. Another advantage of this system is the long emission wavelength, which makes the designed structure 1.4–1.8 times larger than that for GaAs/AlGaAs ones. Let us check the actual size of 2D crystals exhibiting photonic band gaps at emission wavelength λ_s of 1.55 μm. To obtain an absolute band gap for circular holes with $f = 0.13$, their pitch and diameter must be 0.698 and 0.683 μm, respectively, and hence semiconductor walls with a minimum thickness of 15 nm are required between holes. For columns with $f = 0.1$, the lowest TM band gap requires pitch and diameter of 0.527 and 0.179 μm, respectively. To operate 2D crystals properly as expected, the height of those structures must be much larger than the pitch. The minimum aspect ratio required for the thin walls between holes is over 10 : 1 and for columns 3 : 1.

A 2D crystal constructed by columns has been fabricated using CH_4 based reactive ion beam etching (RIBE), and evaluated using photoluminescence (PL) and transmission measurements at room temperature [7]. A 1.55 μm GaInAsP/InP wafer having a compressive-strained multiple quantum-well (CS-MQW) active region is used for the experiment. The typical etching condition for the material is gas flow rate of $CH_4 : H_2 : Ar = 0.65 : 0.9 : 0.3$ sccm, gas pressure of 8×10^{-5} Torr, acceleration voltage of 200 V, microwave power of 100 W, and no substrate heating. On this condition, etch rate for InP is maximally 80 nm/min. Figure 4.4.7 shows the SEM view of formed columns, where a titanium film is used as a mask. The sidewall roughness is evaluated from a topographic SEM image to be less than 5 nm, but the sidewall is inclined with angle of 79–87 degrees due to the insufficient sharpness of the formed mask. From the degradation of PL intensity, the damage induced in the etched surface and that in sidewalls is estimated to be 200 nm and < 100 nm, respectively. Upon slight wet etching with HCl: $H_2O = 4 : 1$ solution, the PL intensity multiplied by fraction f is recovered to almost the same level as that from the as-grown wafer. For formed columns with diameters of 0.35 and 0.40 μm, polarization characteristics of PL intensity

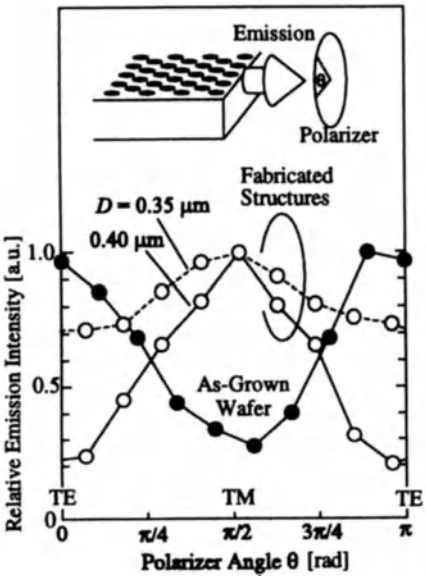


Fig. 4.4.8 Polarization characteristic of spontaneous emission from formed columns with diameter D

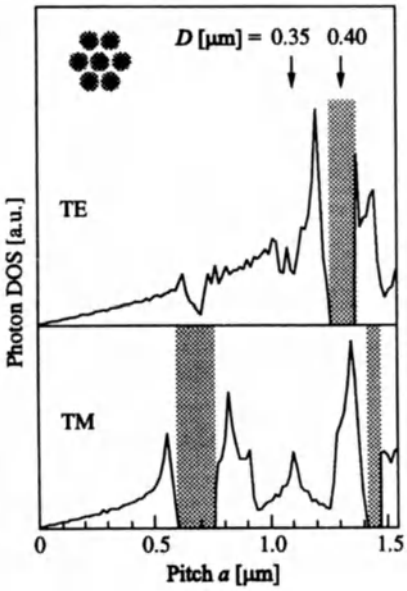


Fig. 4.4.9 Photon DOS calculated for formed columns

is measured after the wet etching in the Γ-X direction inside the substrate plane, as shown in Fig. 4.4.8. The emission from the edge is collimated and focused into a multimode fiber by a pair of lenses. The polarization of the emission is analyzed by a polarizer inserted between the lenses. The emission from as-grown wafer is TE-polarized, while that from formed columns is TM-polarized. The TE-polarized emission is caused by dipole anisotropy in the CS-MQW wafer used. The TM-polarized emission can be explained as an effect of the photonic crystal. Figure 4.4.9 shows the photon DOS calculated for formed structures. Columns with diameter 0.4 μm almost satisfy the TE band gap condition. Columns with diameter 0.35 μm do not overlap with any band gaps, but rather overlap with a small peak of photon DOS for the TM polarization.

The measurement of transmission spectra is effective for a direct evaluation of the photonic band gap [13]. To avoid measurement difficulties for such small columns inside the 2D plane, an equivalent transmission measurement is performed. Various columns are prepared on the same wafer satisfying $f = 0.1$, i.e., 0.3–0.5 μm in diameter and 0.9–1.5 μm in pitch. To avoid the deformation of columns, no wet etchings are performed after the RIBE. The setup for the transmission measurement is illustrated in Fig. 4.4.10. As a light source, a large square region neighboring the columns is used, which is not removed by the RIBE. This region is photo-pumped from the top of the wafer. The emission in the lateral direction is detected directly by a single-mode fiber from a distant position and analyzed by an optical spectrum analyzer after

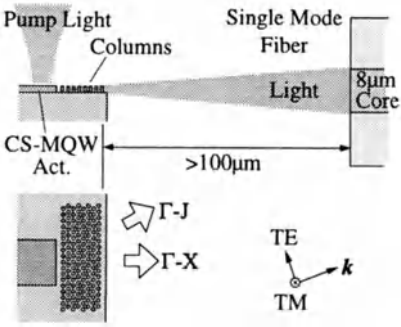


Fig. 4.4.10 Measurement setup of equivalent transmission measurement

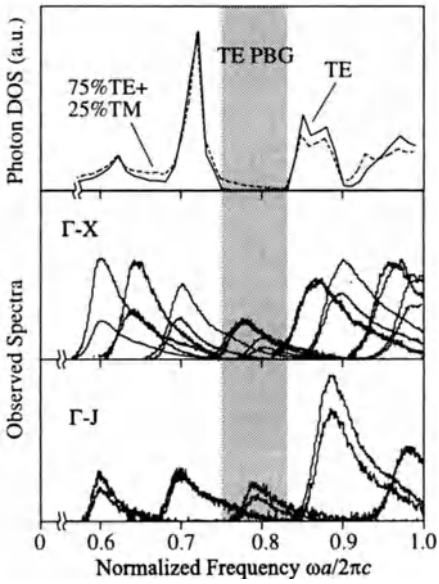


Fig. 4.4.11 Calculated photon DOS and observed transmission spectra

passing through the columns. Lenses and polarizer are not inserted between the wafer and the fiber to avoid reducing the light power and to minimize variation of the detecting conditions. By repeating the measurement for various columns and plotting all the measured spectra on the axis of normalized frequency $\omega a/2\pi c$, an equivalent transmission spectrum is obtained as an envelope of the spectra. Such measurements are performed in both Γ-X and Γ-J directions. The results are summarized in Fig. 4.4.11. Here, calculated photon DOS for TE polarization and that for the sum of 75% TE and 25% TM are also plotted for comparison. The latter polarization ratio simulates the polarization anisotropy observed for the wafer used. Theoretically, a TE band gap lies around $\omega a/2\pi c = 0.8$. In the Γ-X direction, spectra almost overlapping with the band gap show nearly −4 dB decrease of transmission intensity. Although it is ambiguous in the Γ-J direction, at least −1 dB decrease could be recognized at the same frequency range. The frequency range of low transmission intensity is blue-shifted and the decrease of transmission intensity at the band gap becomes more and more ambiguous when the fiber is slightly tilted upward. This is considered to be influenced by the sidewall inclination of columns.

To achieve perfect inhibition of the spontaneous emission at photonic band gaps and a high-Q resonance applicable to lasers, fully uniform, precise, and smooth structures are necessary. The key point is that the etched sidewall be precisely normal to the substrate plane, which may be obtained by further improvement of the etching mask. For 3D crystals, finer etching techniques are required. So far, a Yablonovite structure has been fabricated into GaAs using chemically assisted ion beam etching and a photonic band gap at lightwave

frequency range has been demonstrated by a direct transmission measurement [21]. Problems that prevent application for emitting devices lie in the material itself, as mentioned before, and the difficulty of making a defect inside the structure. An interesting approach using a fusion technique is also being developed [23]. It shows promise for realizing a GaInAsP/InP 3D crystal constructed by stacked rods with a defect, although it needs a complicated process of repeated fusions and alignments.

References

1 E. Yablonovitch, Phys. Rev. Lett. **58**, 2059 (1987)
2 F. De Martini, G. Innocenti, G. Jacobovitz, and P. Mataloni, Phys. Rev. Lett. **59**, 2955 (1987)
3 Y. Yamamoto, S. Machida, K. Igeta, and G. Bjork, *Coherence, Amplification and Quantum Effects in Semiconductor Lasers* edited by Y. Yamamoto (John Wiley & Sons, New York, 1991)
4 H. Yokoyama, K. Nishi, T. Anan, H. Yamada, S. Brorson, and E. Ippen, Appl. Phys. Lett. **57**, 2814 (1990)
5 T. Baba, T. Hamano, F. Koyama, and K. Iga, IEEE J. Quantum Electron. **27**, 1347 (1991)
6 E. Yablonovitch, T. Gmitter, R. Meade, A. Rappe, K. Brommer, and J. Joannopoulos, Phys. Rev. Lett. **67**, 3380 (1991)
7 T. Baba, IEEE J. Selected Topics on Quantum Electron. **3** (1997)
8 K. Leung and Y. Liu, Phys. Rev. Lett. **65**, 2646 (1990)
9 Z. Zhang and S. Satpathy, Phys. Rev. Lett. **65**, 2650 (1990)
10 J. Joannopoulos, R. Meade, and J. Winn, *Photonic Crystals* (Princeton University Press, 1995)
11 T. Baba and T. Matsuzaki, Jpn. J. Appl. Phys. **34**, 4496 (1995)
12 K. Ho, C. Chan, and C. Soukoulis, Phys. Rev. Lett. **65**, 3152 (1990)
13 E. Yablonovitch, T. Gmitter, and K. Leung, Phys. Rev. Lett. **67**, 2295 (1991)
14 K. Ho. C. Chan, C. Soukoulis, R. Biswas, and M. Sigalas, Solid State Commun. **89**, 413 (1994)
15 R. Meade, A. Rappe, K. Brommer, and J. Joannopoulos, Phys. Rev. B **44**, 13772 (1991)
16 X. Feng and Y. Arakawa, IEEE J. Quantum Electron. **32**, 535 (1996)
17 K. Leung, J. Opt. Soc. Am. B **10**, 283 (1993)
18 P. Villeneuve, S. Fan, and J.D. Joannopoulos, *Microcavities and Photonic Bandgaps: Physics and Applications* (1996), p. 133
19 T. Krauss, Y. Song, S. Thoms, C. Wilkinson, R. DelaRue, Electron. Lett. **30**, 1444 (1994)
20 T. Baba and T. Matsuzaki, Electron. Lett. **31**, 1776 (1995)
21 C. Cheng, A. Scherer, V. Engles, and E. Yablonovitch, J. Vac. Sci. Technol. B **14**, 4110 (1996)
22 T. Hamano, H. Hirayama, and Y. Aoyagi, Proc. Topical Meet. Quantum Optoelectronics **9**, 29 (1997)
23 S. Noda, N. Yamamoto, and A. Sasaki, Jpn. J. Appl. Phys. **35**, L909 (1996)

4.5 Microcavity Surface Emitting Lasers

In this section, we review progress on microcavity vertical cavity surface emitting lasers by featuring their materials and performance including ultimate thresholds, power output, efficiency, and so on. Advanced technologies for control of polarization, lasing wavelength, spontaneous emission, etc., will be introduced. We discuss potential applications such as parallel optical interconnects and links, opto-electronic equipments including laser printers, etc., and generation of the field of ultra-parallel opto-electronics.

4.5.1 Overview

Technical progress on surface emitting (SE) lasers [1] and vertical cavity SE lasers (VCSELS) has enabled us to utilize these kinds of device in various optical systems, such as LANs and WANs, optical parallel links, laser printers, and ultra-parallel opto-electronics [2]. We review recent advances in SE lasers for wide range of optical spectra based on GaInAsP, GaInAs, GaAlAs, AlGaInP, ZnSe, GaN, and related compounds.

(a) 1550 nm: The epitaxial bonding of GaInAsP-InP/GaAs-AlAs mirrors has exhibited CW operation up to 64°C [3]. Further development may be necessary to compete with edge emitters employing DFB and spot-size transformer structures.

(b) 1300 nm: A buried structure is utilized together with a thermally conductive MgO/Si mirror demonstrating the first room temperature CW operation [4]. The hottest operation is up to 34°C [5].

(c) 980 nm: The GaInAs/GaAs strained pseudomorphic system grown on GaAs exhibits a high laser gain and has been introduced into SE lasers together with use of GaAlAs/AlAs DBRs [6] as shown in Fig. 4.5.1.

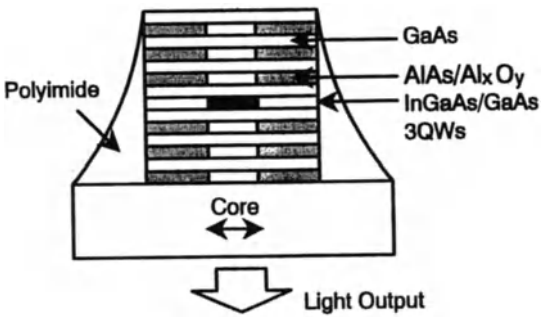


Fig. 4.5.1 An oxide-confined microcavity surface emitting laser

More recently, an even lower threshold (10 mA, 40 mA) has been reported [7,8]. Some commercial devices have already been developed for optical interconnect applications. A power conversion efficiency of 50% was reported. Also, near 200 mW has been demonstrated. threshold current

However, due to the unavailability of low cost detectors for this wavelength, it is difficult to find real applications. This wavelength might be useful for a "header" device.

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

Table 4.5.1 Comparison of physical dimensions of VCSELs and stripe lasers

Parameters	Symbol	Stripe lasers	VCSELs
Active layer thickness	d	10–100 nm	6–500 nm
Reflectivity	R_m	0.3	0.99–0.999
Cavity length	L	300 μm	0.3 μm ($\lambda/2-\lambda$)
Optical confinement factor	ξ	0.5%	3% $\times$ 2
Area of active region	S	3×300 μm^2	5×5 μm^2
Photon lifetime	τ_p	0.1 ps	0.1 ps
Relaxation frequency	f_r	5 GHz	10 GHz

(d) 850 nm: Sub-mA thresholds and 10 mW outputs have been achieved. A power conversion efficiency of 57% has been demonstrated [9]. Some commercial optical links have already been marketed. The price of low-skew multimode fiber ribbons may be a key issue for inexpensive multimode-fiber-based data links.

(e) 780 nm: A device with 200 mA threshold and 1.1 mW output has been reported intended for CD player application [10]. VCSEL-loaded pickups are being commercialized.

(f) 670–650 nm: AlGaInP/GaAs SE lasers are developed and room temperature operation exhibiting sub-mA threshold and a few mW outputs have been obtained [11]. The best application of red-color VCSELs may be laser printers and first-generation DVDs.

(g) 555 nm: The ZnSe system is still the only material to provide CW operation of blue-green semiconductor lasers operating beyond 100 hours. Some trials of green surface emitters have been attempted.

(h) 470–340 nm: GaN and related materials can cover wide spectral ranges from green to UV. The reported reliability of GaN based LEDs and LDs [12] appears to indicate good potential for surface emitting lasers as well [13]. Estimation of thresholds for GaN/$Al_{0.1}Ga_{0.9}N$ quantum well lasers is carried out using density matrix theory with intraband broadening [14].

4.5.2 Technology for Low Threshold Surface Emitting Lasers

The physical difference between VCSELs and conventional stripe geometry lasers is summarized in Table 4.5.1. The main points are that the cavity length of VCSELs is in wavelength order which is much smaller than that of stripe lasers, with about 300 mm and several mm of lateral size. Those provide us with substantial differences in laser performance.

By simply assuming that the threshold current depends on the lateral confinement factor and carrier diffusion in the active layer, the threshold current is calculated as shown in Fig. 4.5.2 [15]. From this figure, we can expect that the minimum threshold could be as low as a few μA.

In order to achieve further substantial improvements in surface emitting laser performance, the following technical issues need to be solved or optimized.

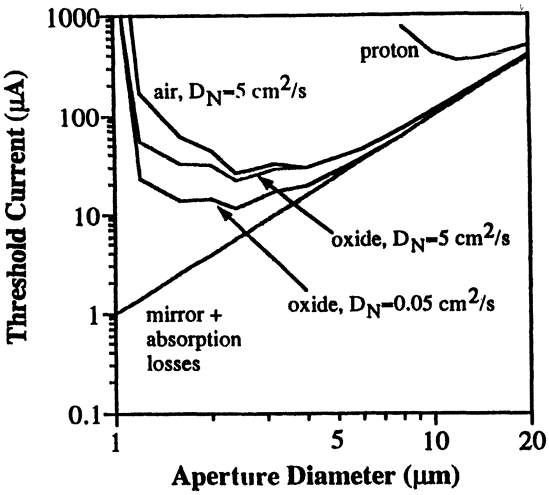


Fig. 4.5.2 Estimation of threshold currents in microcavity vessels

(a) AlAs oxidation in GaAs and InP based structures for application to current confinement and optical beam focusing. Device lifetime is now being positively examined.
(b) Modulation and delta doping, p-type and n-type, to quantum wells and barriers and to DBRs.
(c) Carbon auto-doping technique.
(d) Quantum wires and dots for active engines.
(e) Strained quantum wells and strain compensation.
(f) High index substrates such as (311A), (311B), (411), etc., for higher gains and polarization control [15].
(g) Transparent mirrors to increase quantum efficiency and output power.
(h) Multi-quantum barriers (MQBs) to prevent carrier leakage to p-cladding layer.
(i) Also, we have to develop some moduling and packaging method specific to surface emitting lasers.

The native oxide of AlAs is successfully used for current confinement to achieve ultra-low thresholds as already shown in Fig. 4.5.1. A typical size [16] is a 20 μm core starting from a 30 μm mesa diameter. We have achieved about 1 mW of power output and sub-μA threshold. A smaller device having 5 μm starting from a 20 μm mesa has exhibited a threshold of 70 mA at room temperature CW operation as shown in Fig. 4.5.3. This value is the third world record for SE lasers.

4.5.3 New Materials for Surface Emitting Lasers

Some choices for the materials for semiconductor lasers are shown in Fig. 4.5.4. The availability of substrates against lattice constant related to the systems

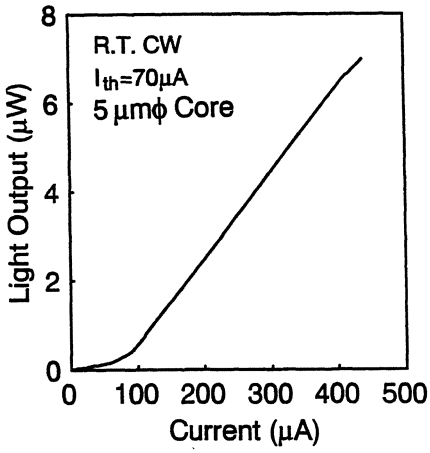


Fig. 4.5.3 The threshold of 70 μA achieved by an oxide-confined vessel

is also shown. We first list problems below which should be considered when making vertical cavity SE lasers.

(a) Design of resonant cavity and mode-gain matching.
(b) Multi-layered distributed Bragg reflectors (DBRs) to realize high reflective mirrors.
(c) Optical losses such as Auger recombination, intravalence band absorption, scattering loss, and diffraction loss.
(d) p-type doping to reduce the resistivity in p-type materials for CW and high efficiency operation. If we wish to form multilayer DBRs, this will become much more critical.
(e) Heat sinking for high-temperature, high-power operation.
(f) Catastrophic optical damage level is very important for high power operation.
(g) Crystal growth at reasonably high temperatures (i.e., higher than half of melting temperatures).

We mention some trials on developing new materials for surface emitting lasers. The introduction of GaInNAs/GaAs and AlGaInAs/InP having an AlAs oxidation capability, the use of GaAs/AlAs DBRs, etc., may be possible solutions to remove the material bottleneck of GaInAsP/InP systems, i.e., temperature characteristics. It would be extremely important to develop 1300 nm VCSELs for single-mode fiber based optical links due to low chip and moduling costs. It may replace edge emitters for optical interconnects and links, if 1 mW devices can be realized.

A trial for realizing blue-UV SE lasers has just started. Some optical pumping experiments have been reported. We have made a preliminary study to search for selective growth and dry etch of a GaN system by a Cl based reactive ion beam etch. These may be crucial technologies for GaN based devices. The application area of blue-UV VCSELs is huge. It may provide the next-generation DVD, laser printers incorporating blue-sensitive drums,

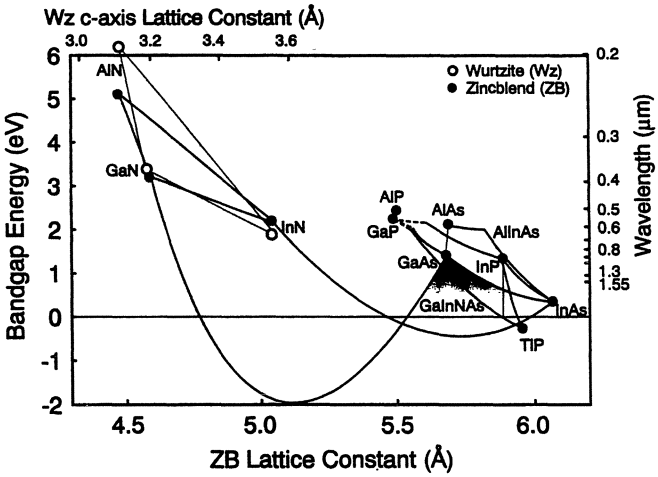


Fig. 4.5.4 A mapping of semiconductors after Tomoyuki and Miyamoto

full color displays, high efficiency illuminations together with green and red devices, and so on.

References

1 K. Iga, F. Koyama, and S. Kinoshita, IEEE J. Quant. Electron. **24**, 1845 (1988)
2 K. Iga, Gakujutu-Geppo (Monthly J. of MECSS) **49**, 42 (1996)
3 D.I. Babic, K. Streubel, R.P. Mirin, J. Pirek, N.M. Margalit, J.E. Bowers, E.L. Hu, D.E. Mars, L. Yang, and K. Carey, IPRM'96, ThA1-2 (1996)
4 T. Baba ,Y. Yogo, K. Suzuki, F. Koyama, and K. Iga, Electron. Lett. **29**, 913 (1993)
5 S. Uchiyama, N. Yokouchi, and T. Ninomiya, The 43th Spring Meeting of Jpn. Soc. Appl. Phys. 26p-C-7 (1996)
6 Y. Hayashi, T. Mukaihara, N. Hatori, N. Ohnoki, A. Matsutani, F. Koyama, and K. Iga, Electron. Lett. **31**, 560 (1995)
7 G.M. Yang, M.H. Macdougal, and P.D. Dapkus, CLEO'95, Post Deadline Papers, CPD4-1, Baltimore (1995)
8 D.L. Huffaker, D.G. Deppe, C. Lei, and L.A. Hodge, CLEO'96 (Anaheim), JTuH5 (1996)
9 B. Weigl, G. Reiner, M. Grabherr, and K. Ebeling, CLEO'96 (Anaheim), JTuH2 (1996)
10 H.E. Shin, Y.G. Zoo, and E.H. Lee, 16th Conference on Lasers and Electro-Optics (Anaheim), JTuH7 (1996)
11 J.A. Lott and R.P. Shneider, Jr., Electron. Lett. 29, 830 (1993)
12 S. Nakamura, M. Senoh, S. Nagahama, N. Iwasa, T. Yamada, T. Matsushita, H. Kiyoku, and Y. Sugimoto, Jpn. J. Appl. Phys. **35**, L14 (1996)
13 K. Iga, Int. Symposium on Blue Laser and Light Emitting Diodes, Th-11 (1996)
14 T. Honda, F. Koyama, and K. Iga, *MRS 1996 Fall Meeting Symposia Proceedings, Material Research Society, Pittsburgh, 1997*
15 H. Bissessur, T. Mukaihara, M. Abe, F. Koyama, and K.Iga, IEEE/LEOS Summer Top. Meet. on Smart Pixels, WA3 (1996)
16 K. Takahashi, P.O. Vaccaro, T. Watanabe, T. Mukaihara, F. Koyama, and K. Iga, Jpn. J. Appl. Phys. **35**, 6102 (1996)

4.6 Toward Lasers of the Next Generation

4.6.1 Quantum Dot Lasers

One very feasible nanostructure optical device is a quantum dot (QD) laser. When quantum wire (QWR) and QD lasers were proposed in 1982 [1], a suppression of the temperature dependence of the threshold current in QWR and QD lasers was discussed theoretically. A significant improvement of lasing characteristics such as a low threshold current, fine modulation dynamics, and good spectral properties were predicted theoretically [2–4]. In this section, we focus on QD lasers with emphasis on lasing characteristics and microcavity effects [5].

Since the density of states has a peaked structure (i.e., a δ-function-like shape) in QDs, an extremely narrow gain profile $g(E)$ can be obtained. A simple expression for $g(E)$ for a QD laser is given by the following equation if the Lorentzian energy broadening due to the dephasing of the dipole moment is assumed:

$$g(E) = \sum_i 2A \frac{\hbar/\tau_{\mathrm{in}}}{(\hbar/\tau_{\mathrm{in}})^2 + (E - E_i)^2}, \tag{4.6.1}$$

where A is a constant determined by optical matrix elements, E_i is the energy difference between quantized electrons and holes, and τ_{in} denotes the dephasing time of the optical dipole moment. Because the gain profile is narrow, the Fermi statistics of carriers is not important in the shape of the gain profile, leading to a suppression of temperature dependence of the threshold current and a wide modulation bandwidth.

The threshold current J_{th} is expressed by

$$J_{\mathrm{th}} = \frac{eV_a n_{\mathrm{th}}}{\tau_r}, \tag{4.6.2}$$

where n_{th} is the carrier concentration per unit volume, τ_r is the recombination lifetime, and V_a is the total volume of the active region. In the case of a highly peaked gain, J_{th} can be reduced considerably by reducing n_{th}. Usually, however, decreasing V_a is more significant in QWR and QD lasers. In fact, because the ratio of V_a for QD lasers to that for quantum well lasers can be as small as 10^4, even a sub-μA threshold current may be achievable. Various experiments have been reported demonstrating a reduction of J_{th} in QWR and QD lasers, but most of the results can be understood as a volume effect only. In order to fully use the advantage of the narrow gain profile, the size of dots should be less than 20 nm and their fluctuations should be less than 5%.

The modulation dynamics of semiconductor lasers is an important issue for their practical application to high-speed optical communication systems. The important parameter for the modulation bandwidth is the relaxation resonant frequency f_r, which can be derived by conventional rate equations as follows:

$$f_r = \frac{1}{2\pi}\sqrt{\frac{g' P_0}{\tau_p}}, \tag{4.6.3}$$

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

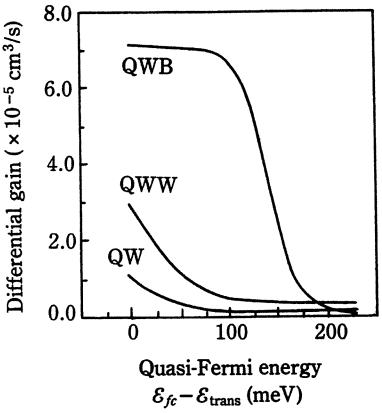


Fig. 4.6.1 The differential gain g' of a QD laser plotted as a function of the quasi-Fermi level ε_{fc} measured from the quasi-Fermi level $\varepsilon_{\text{trans}}$ at which the transparency condition is satisfied

where P_0 is the stationary photon density in the cavity, τ_p is the photon lifetime, and g' is the differential gain defined as $\delta g/\delta n$. In a QD laser, we can expect that g' is much enhanced compared to double heterostructure (DH) lasers since the gain profile is extremely narrow. Therefore, the use of the QD structure in semiconductor lasers is expected to lead to ultra-high-speed modulation.

Figure 4.6.1 shows g' of the QD laser plotted as a function of the quasi-Fermi energy level ε_{fc} measured from the quasi-Fermi level $\varepsilon_{\text{trans}}$ at which the transparency condition is satisfied. For comparison, g' values for a quantum well and QWR laser are also plotted. The calculated results for g' show two important features. The first feature is that g' is enhanced in a QD laser compared to quantum well lasers by a factor of ten in the region of low quasi-Fermi level. The maximum value of g' in a QD laser reaches over 7×10^{-5} cm^3s^{-1}. The second is that g' is strongly dependent on the quasi-Fermi level. This dependence is more pronounced with decreasing dimensionality of the electron motion. Therefore, a wide modulation bandwidth can be achieved by increasing the number of quantum dots because the quasi-Fermi level can be reduced by increasing the number.

The spectral linewidth of the semiconductor laser is enhanced by a factor of $(1+\alpha^2)$ compared to the Schawlow-Townes linewidth $\Delta\nu_{\text{ST}}$, where α is the linewidth enhancement factor given by

$$\alpha = \frac{\partial \chi_R}{\partial n}\left(\frac{\partial \chi_I}{\partial n}\right)^{-1}, \tag{4.6.4}$$

where χ_R and χ_I are the real and imaginary part of the complex susceptibility, respectively. In bulk DH lasers, α has been measured to be in the range 2.2–7. In QD lasers we can expect to have $\alpha \approx 0$. The elimination of the linewidth enhancement factor is considered as a big advantage for not only laser devices but also optical modulators with low chirping effects.

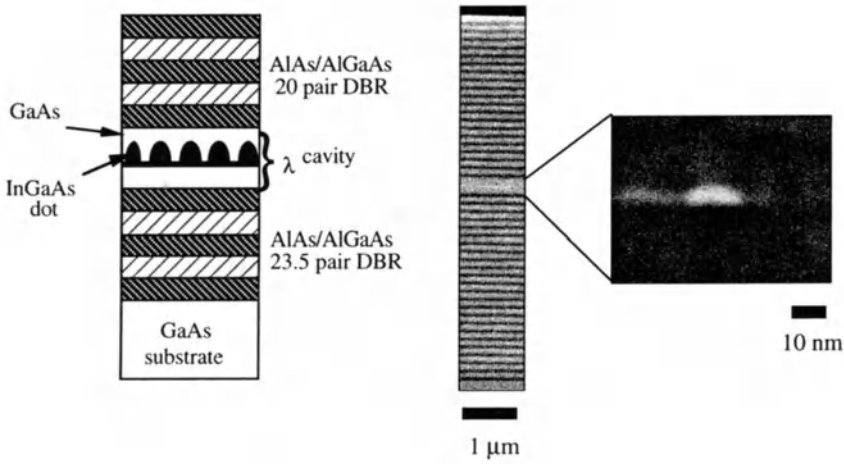


Fig. 4.6.2 A schematic illustration and a cross-sectional SEM image of a laser structure

In order to bring QD lasers into the real world, various issues must be solved. These include the size, size fluctuation, quality, density, and many others. Concerning the size, small QDs of less than 20 nm have already been obtained. The problem of the size fluctuation is very serious. In fact, it is equivalent to inhomogeneous broadening in the gain profile and therefore reduces the gain.

4.6.2 Microcavity Quantum Dot Lasers

Microcavity lasers have been investigated from the viewpoint of controlling spontaneous emission and reducing threshold current. At present, because of technological limitations, mainly vertical cavity structures are being investigated. The ultimate structure is of course a 3D microcavity in which photon modes are completely controlled, resulting in single lasing mode operations. In QD structures, electrons are confined three-dimensionally and a single mode is realized for the electron wave. When the two concepts are combined, a new type of laser in which both electron and photon waves are completely controlled will be generated. In fact, a laser oscillation may be achieved with only a few QDs in ideal cases.

There are two coupling regimes of the electron-photon interaction. In the weak-coupling regime, the concentration of spontaneous emission into a cavity mode opens a way to realize better light emitters. A microcavity quantum dot laser is included in this category. In the strong-coupling regime, coherent interaction between electrons and photon leads to a Rabi oscillation [6]. A very short lifetime of photons in the cavity may be used to generate a short light pulse determined by the Rabi frequency.

A few trials have been made of lasing from InGaAs quantum dot Fabry–Perot type laser structures. InGaAs quantum dots were grown by a Stranski–Krastanov (SK) growth mode with MOCVD [7]. Figure 4.6.2 shows a schematic illustration and a cross-sectional SEM image of the laser structure. The

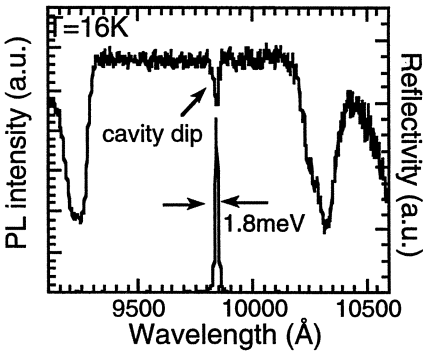


Fig. 4.6.3 Reflection and photoluminescence spectra

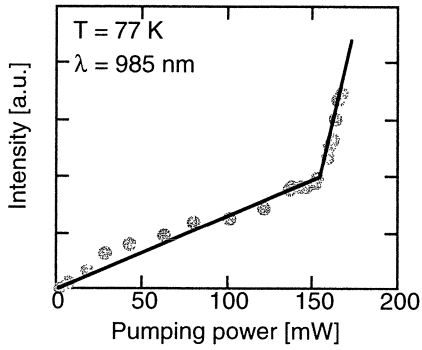


Fig. 4.6.4 Light input-output characteristics from a vertical microcavity quantum dot laser

microcavity consists of an InGaAs quantum dot layer sandwiched by GaAs layers, located between two AlAs/$Al_{0.2}Ga_{0.8}As$ distributed Bragg reflector (DBR) mirrors. InGaAs quantum dots were embedded in GaAs at an antinode of the cavity. When the growth of the first GaAs layer in the cavity was started, the growth temperature was gradually decreased from 700°C to 500°C. Then, InGaAs quantum dots were formed by an SK growth mode at 500°C. After the deposition of InGaAs was finished, the growth was interrupted for 10 min and the temperature was simultaneously lowered to 450°C. The upper GaAs cap layer was grown while raising the growth temperature from 450°C to 700°C so that the sample surface becomes smooth.

First, the photoluminescence (PL) spectrum of quantum dots without DBR mirrors was measured with the use of a He-Ne laser with a wavelength of 632.8 nm. At a low excitation intensity a single PL peak appears at 1.255 eV (988 nm), while with increasing excitation power other peaks also appear at 1.315 eV (943 nm) and 1.365 eV (908 nm). These peaks can be attributed to higher subbands of the quantum dots.

Optical properties of the vertical microcavity quantum dot laser structure were measured by an optical pumping at 77 K. The resonance energy of the microcavity structures was tuned to coincide with the ground state of the PL peak. To measure the reflection characteristics of the cavity structures, an external source of collimated white light was used. This light was superimposed on the laser beam using a beam splitter. Figure 4.6.3 shows spectra of both reflected white light and PL. The reflection spectrum displays a 90 nm wide stop band. The asymmetry of the spectrum is probably due to the wavelength dependence of the response characteristics of the measurement system. Since the wavelength of the pumping laser light corresponds to an energy lower than that of the band gaps of GaAs and AlGaAs used, no serious losses occur in the DBR structure. The cavity resonance was tuned to the PL peak of the lowest subband of the quantum dot structures. The lasing wavelength was clamped at this wavelength owing to the cavity effect even when higher subbands were excited. As shown in the diagram, a sharp PL linewidth with a full width

at half maximum of less than 1.8 meV was observed. By comparing spectral broadening with and without the DBR structure, the cavity effect was clearly demonstrated. The light input-output characteristics indicate a lasing oscillation as shown in Fig. 4.6.4.

References

1 Y. Arakawa and H. Sakaki, Appl. Phys. Lett. **40**, 939 (1982).
2 Y. Arakawa, K. Vahala, and A. Yariv, Appl. Phys. Lett. **40**, 950 (1984)
3 Y. Arakawa and A. Yariv, IEEE J. of Quantum Electron. **QE-22**, 1887 (1986)
4 Y. Arakawa and T. Takahashi, Optoelectronics **3**, 155 (1988)
5 Y. Arakawa, *Proc. of 23rd International Conference on the Physics of Semiconductors, 1996* (World Scientific, Singapore, 1996), p. 1349
6 C. Weisbuch M. Nishioka, A. Ishikawa, and Y. Arakawa, Phys. Rev. Lett. **69**, 3314 (1992)
7 Y. Arakawa, M. Nishioka, H. Nakayama, and M. Kitamura, IEICE Trans. Electron. **E79-C**, 1487 (1996)

5. Quantum-Effect Devices

5.1 Introduction

The long history of the field of electronics shows that it takes more than a decade to mature a new technology after planting its seed. LSI technology is growing steadily at present but the growth cannot last forever and it is now time for the seed of new electronics and post-LSI technology to be planted. The wave nature of the quantum-mechanical electron motion gives rise to various phenomena such as interference, multiple reflection, Bragg reflection, diffraction, etc. The discrete electronic charge gives rise to fascinating phenomena like the Coulomb blockade, single-electron tunneling, etc. These quantum-mechanical aspects of the electron may be used for the control of electron transport and have high potential as new quantum-effect devices.

As to information processing at present, the device function required is simple on-off switching. Integrating huge number of switches, one may build a system which performs a function up to a certain degree of complexity. However, it has been anticipated that a system as complex as a human brain is not built with digital switches because of the heat removal problem [1]. In the neural system of the brain, exponential response to a stimulus often appears. To simulate such a system with an artificial circuit, the exponential response may be realized by using a natural Boltzmann-type electron distribution in a single device instead of many switching devices and many operation steps. Such an analog scheme seems promising for realization of large and low-power systems [1].

Future electronic devices are expected to provide not only a simple switching function but also some sophisticated functions using appropriate phenomena effectively. Quantum-effect devices may provide such functions better than silicon LSI [2]. Various new functions have been proposed that can be realized by quantum-effect devices, including noise elimination in image data [2], pattern processing such as Fresnel–Fourier transforms, an analog solver of the Schrödinger equation, signal transmission without wiring using electron wave diffraction, etc. In this chapter a discussion is given of research and development toward such quantum-effect devices.

References

1 C. Mead, *Analog VLSI and Neural System* (Addison-Wesley, Reading, 1989).

2 Y. Amemiya, *Study of functional devices based on non-Boolean logic*, Journal of Institute of Electronics, Information and Communication Engineers (IEICE), June 1993 (in Japanese).

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*

© Springer-Verlag Berlin Heidelberg 1998

5.2 Electron-Wave Reflection and Resonance Devices

5.2.1 Introduction

Metal/insulator ultrathin heterostructures are good candidates for high-speed electron devices, because the high carrier density of the metal and the low dielectric constant of the insulator are suitable for size reduction and high-speed operation [1,2]. In addition, due to a very large conduction-band discontinuity at the heterointerface, the interference of the electron wave is expected to become significant in multilayer structures, which result in high transconductance and multifunctionality of the quantum-effect devices [3,4].

The epitaxial growth of the metal/insulator heterostructure system [5] and its application to quantum-effect electron devices [6,7,8,9] have been studied. As the materials, $CoSi_2$ and CaF_2 have been chosen because they are nearly lattice-matched to Si, with mismatches of -1.2% and $+0.6\%$, respectively. In this section, we describe epitaxial growth of $CoSi_2/CaF_2$ nanometer-thick layered structures and quantum-effect devices, such as resonant tunneling transistors and quantum interference transistors. We also discuss field-effect tunneling transistors with $Si/CdF_2/CaF_2$ and its epitaxial growth, which can be extended to field-effect quantum devices on Si.

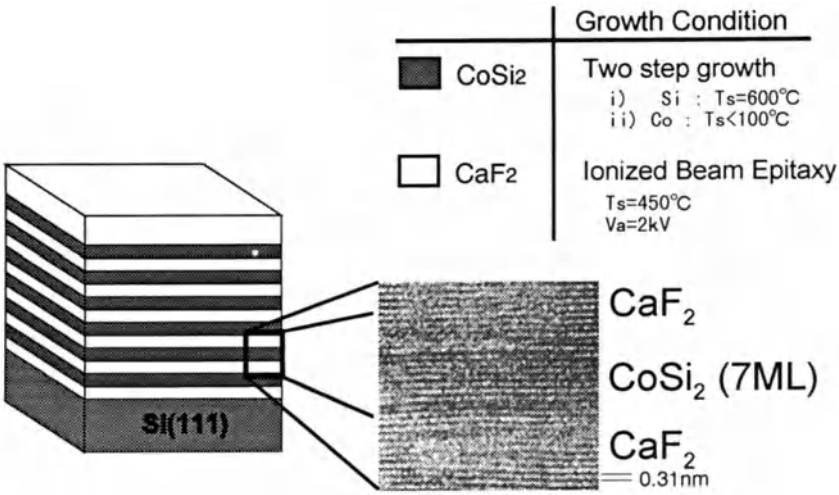


Fig. 5.2.1 Lattice image of transmission electron microscopy of metal ($CoSi_2$) / insulator (CaF_2) superlattice on Si substrate

5.2.2 Epitaxial Growth of $CoSi_2/CaF_2$ on Si

In order to create $CoSi_2/CaF_2$ epitaxial multilayers, epitaxial growth conditions were studied separately for CaF_2 on $CoSi_2$ and $CoSi_2$ on CaF_2 [5]. In the case of $CoSi_2$ on CaF_2, the agglomeration of Co on CaF_2 is a critical problem if we use co-deposition of Si and Co with the conventional MBE technique. This problem was overcome by the following two-step growth technique: First,

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

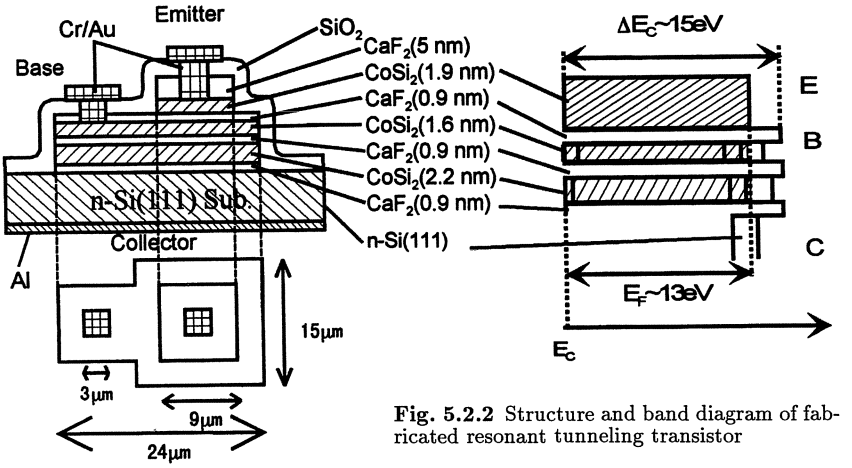


Fig. 5.2.2 Structure and band diagram of fabricated resonant tunneling transistor

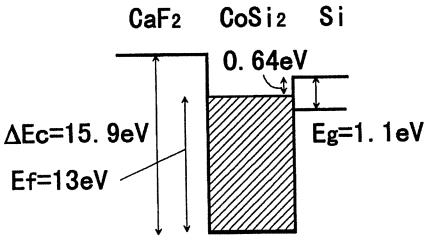


Fig. 5.2.3 Band lineup of $CoSi_2/CaF_2/Si$ heterostructure

four flat monolayers of Si were formed on CaF_2 at the substrate temperature $T_s = 600°C$. Then, two monolayers of Co were deposited on the Si layers by solid phase epitaxy at a temperature lower than about 100°C. Because of the difference of surface energy between Co and CaF_2, two flat monolayers of $CoSi_2$ were grown on CaF_2 without agglomeration by this technique. $CoSi_2$ layers thicker than two monolayers were obtained by repeating this process with a Si growth temperature of 550°C.

In the case of CaF_2 on $CoSi_2/CaF_2$, T_s must be lower than about 550°C to avoid thermal deformation and agglomeration of $CoSi_2$ layers. We used partially ionized beam epitaxy to give external kinetic energy for migration to the evaporated particles. CaF_2 was grown on $CoSi_2$ at $T_s = 450°C$ with the acceleration voltage $V_a = 2$ kV and the ionization current $I_e = 400$ mA, at which conditions the ionization ratio of the CaF_2 beam was about 2%.

Using these growth techniques for CaF_2 and $CoSi_2$ layers, we made nanometer-thick $CoSi_2/CaF_2$ multilayers on Si(111). First, a 1 nm-thick $CoSi_2$ layer was formed by the deposition of Co on a Si(111) substrate and solid phase epitaxy near room temperature (<100°C), and then the above technique for CaF_2 and $CoSi_2$ layers was iterated. The cross-sectional transmission electron microscopy (TEM) lattice image of the multilayer obtained is shown in Fig. 5.2.1. The epitaxial growth between $CoSi_2$ and CaF_2 is clearly confirmed, though the relation between each layer, i.e., type A or B, is not yet clear.

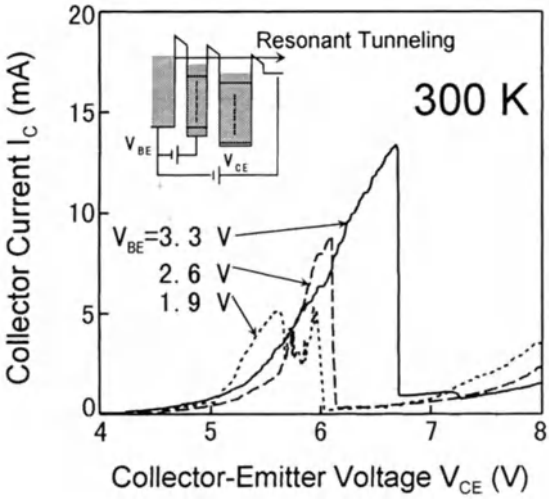


Fig. 5.2.4 Common emitter characteristics of metal ($CoSi_2$) / insulator (CaF_2) resonant tunneling transistor measured at room temperature

To improve the crystalline quality and interface of the $CoSi_2/CaF_2$ multilayer, in situ postannealing was carried out at 860°C for about 15 min. The resistivity was reduced from 100 $\mu\Omega$cm to 30 $\mu\Omega$cm for 2 nm-thick $CoSi_2$ layers sandwiched between CaF_2 layers without agglomeration.

5.2.3 Resonant Tunneling Transistor

We fabricated $CoSi_2/CaF_2$ resonant tunneling transistors, shown in Fig. 5.2.2. The transistor is composed of a few nm-thick triple CaF_2 barriers and double $CoSi_2$ wells on n-Si(111). As shown in Fig. 5.2.3, the conduction band offset at the $CoSi_2/CaF_2$ interface is about 15 eV and the Fermi level is about 13 eV from the conduction band bottom of $CoSi_2$. These values were estimated from the related values of the bulk materials [7]. The emitter, base, and collector electrodes are connected, respectively to the top $CoSi_2$ layer, one of the $CoSi_2$ quantum wells near the emitter contact layer, and the substrate. Electrons pass through the resonant levels in two quantum wells from the emitter to the collector if these levels are aligned. The condition of the alignment between the resonant levels is controlled with the collector-base voltage V_{CB}, and the number of electrons from the emitter is controlled with the base-emitter voltage V_{BE}. Thus, the collector current I_C varies with V_{BE} and exhibits a negative differential resistance (NDR) with respect to V_{CB} for constant V_{BE}. As a different case, if the base electrode is connected to the other quantum well near the collector, I_C exhibits NDR with respect to V_{BE} for a constant V_{CB}. The operation is similar to that of semiconductor resonant tunneling hot electron transistor (RHET) [10].

The fabrication process is as follows. After the growth and anealing of the $CoSi_2/CaF_2$ multilayer by the technique described in the previous subsection, a mesa structure with an emitter size of 9×9 μm^2, emitter and base contact holes of 3×3 μm^2, and device isolation were made by photolithography with an AZ-1350 resist and selective wet chemical etching. For the ohmic contact

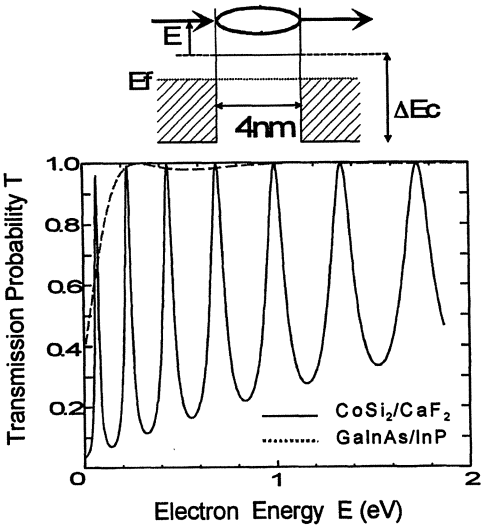


Fig. 5.2.5 Transmission coefficient of hot electron passing through the potential profile formed by metal/insulator/metal structure. Dashed curve is the case for semiconductor heterostructure

to the 1.6 nm-thick $CoSi_2$ base metal layer, a 10 nm-thick CaF_2 cap layer and 1.9 nm-thick $CoSi_2$ emitter contact layer were removed completely from the base area by careful wet chemical etching at the formation of the emitter mesa structure. Then, a 0.9 nm-thick CaF_2 emitter barrier layer was etched to make the contact hole. CaF_2 layers were etched by $H_2O : H_2SO_4 = 4 : 1$ and $CoSi_2$ layers were etched by 1 wt% buffered HF. The etching rate at 0°C was about 1 nm/s for both of the materials. Concerning etch selectivity, we have confirmed, by a scanning electron microscope, that the CaF_2 layer is not etched by 1 wt% buffered HF at least in 1 min, and neither is the $CoSi_2$ by the above etchant of CaF_2. The emitter and base electrodes of 200×200 μm^2 were formed with a Au/Cr nonalloy contact by the lift-off process and the collector electrode was formed with Al at the bottom of the substrate.

The common-emitter characteristics measured at room temperature are shown in Fig. 5.2.4. We observed NDR in the I_C-V_{CE} curve, in which the peak current is controlled by V_{BE}. The peak-to-valley ratio of NDRs was 13–19 at room temperature. The transfer efficiency of electrons injected from the emitter toward the collector through the quantized levels has been estimated from NDR in the collector current and corresponding change in the base current [8]. The transfer efficiency more than 0.99 was obtained at 77 K. In order to estimate the transfer efficiency more exactly, a detailed discussion is necessary on the transit time and the scattering rate in the base region as well as the three-dimensional structure of the device. However, the large value obtained implies at least that the scattering rate is very small in a nanometer-thick epitaxial metal layer.

5.2.4 Observation of Hot Electron Interference

As a quantum effect in metal/insulator system other than resonant tunneling, the interference of a hot electron wave with energy higher than the barrier

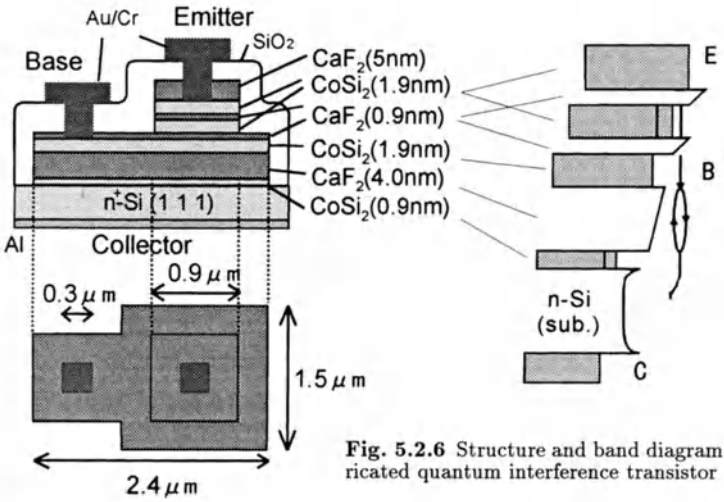


Fig. 5.2.6 Structure and band diagram of fabricated quantum interference transistor

height in a metal/insulator/metal structure is applicable to multifunctional electron devices [4]. The transmission coefficient of hot electrons passing normally through this structure is strongly modulated by this interference, as shown in Fig. 5.2.5. This effect is the same as that in an optical Fabry–Perot interferometer. The interference is much larger in a metal/insulator system compared to conventional semiconductor heterostructures because of the large band offset.

To observe this interference, we fabricated a quantum interference transistor structure, as shown in Fig. 5.2.6. The structure is composed of a double-barrier resonant tunneling emitter, base layer, collector barrier, and collector region on a Si substrate [9]. Hot electrons with collimated energy are emitted from the resonant tunneling emitter, and injected into the conduction band of the collector barrier in which the interference takes place. The interference can be observed as multiple NDRs in the collector current I_C as a function of V_{BE} or V_{CB}. The fabrication process is almost the same as that of the resonant tunneling transistors described in the previous subsection except for the use of electron-beam lithography to make a small-size device with the 0.3 μm design rule as shown in Fig. 5.2.6.

In the fabricated transistors, we observed multiple NDRs in the I_C-V_{CB} curve at room temperature, as shown in Fig. 5.2.7. The voltage interval between the adjacent NDRs was about 1.5 V, in reasonable agreement with theoretical calculation (about 1 V) [9]. A stepwise current variation was observed in the $I_B - V_{CB}$ curve at V_{CB} values of NDRs in I_C. This can be attributed to additional electrons in I_B at NDRs produced by reflection at the collector barrier interference region.

Although room-temperature operation was obtained, the NDRs are very small due to the leak current in or around the device, and also due to the emitter-parasitic resistance produced during the fabrication process. Improvement of the characteristics is possible by optimizing the device.

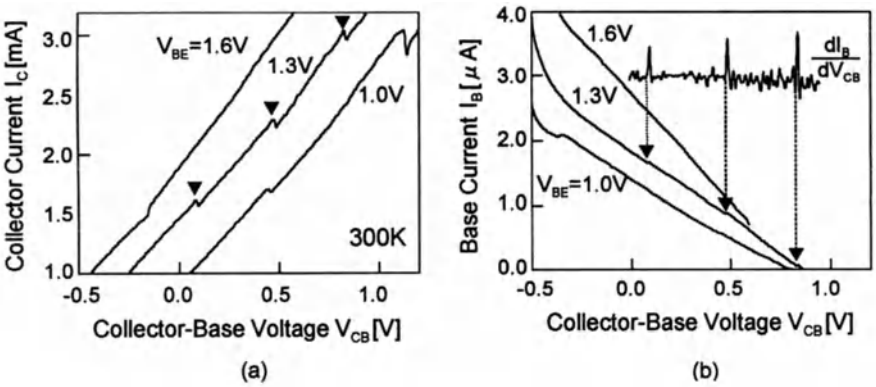


Fig. 5.2.7 Collector and base currents of quantum interference transistor measured at room temperature as a function of collector-base voltage

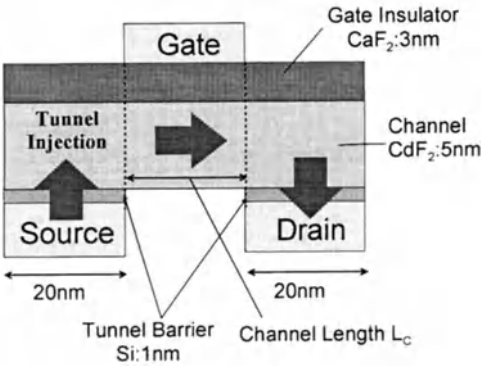


Fig. 5.2.8 Schematic structure of tunneling field effect transistor based on $CaF_2/CdF_2/Si$ heterostructures. This device can be extended to field effect quantum devices such as coupled waveguide devices

5.2.5 Field-Effect Quantum Device

From the viewpoint of input impedance, circuit design, and power consumption, a field-effect device is desirable also in quantum-effect devices. We have proposed a very short channel tunneling field-effect transistor with a $CoSi_2/Si/CdF_2/CaF_2$ heterostructure [11] as a basic device which can be extended to field-effect quantum devices such as coupled waveguide devices. The device structure is shown in Fig. 5.2.8. In the device operation, the current from the source ($CoSi_2$) to the drain ($CoSi_2$) is controlled by the gate electric field applied to the barrier at the source through the gate insulator (CaF_2) and the channel (CdF_2). A theoretical analysis shows that this transistor has characteristics with a saturated drain current in the I_D-V_{DS} curve even with channel lengths of 20 nm and 5 nm [11]. This device has a base structure which can be extended to field-effect quantum devices such as a coupled waveguide devices.

To realize this transistor, growth of $CoSi_2/Si/CdF_2/CaF_2$ on Si is necessary. $CoSi_2/Si$ and CdF_2/CaF_2 in these heterostructures have already been reported [12,13] and the growth of Si on CdF_2 has to be developed.

References

1 C.A. Mead, J. Appl. Phys. **32**, 646 (1961)
2 M. Heiblum, Solid State Electron. **24**, 343 (1981)
3 Y. Nakata, M. Asada, and Y. Suematsu, Electron. Lett. **22**, 58 (1986)
4 T. Sakaguchi, M. Watanabe, and M. Asada, IEICE Trans. **E74**, 3326 (1991)
5 M. Watanabe, S. Muratake, T. Suemasu, H. Fujimoto, S. Sakamori, M. Asada, and S. Arai, J. Electron. Mat. **21**, 783 (1992)
6 M. Watanabe, T. Suemasu, S. Muratake, and M. Asada, Appl. Phys. Lett. **62**, 300 (1993)
7 T. Suemasu, M. Watanabe, J. Suzuki, Y. Kohno, M. Asada, and N. Suzuki, Japan. J. Appl. Phys. **33**, 57 (1994)
8 T. Suemasu, Y. Kohno, W. Saitoh, M. Watanabe, and M. Asada, IEEE Trans. Electron Devices, **42**, 2203 (1995)
9 W. Saitoh, T. Suemasu, Y. Kohno, M. Watanabe, and M. Asada, Japan. J. Appl. Phys. **34**, 4481 (1995)
10 N. Yokoyama, K. Imamura, S. Muto, S. Hiyamizu, and H. Nishi, Japan. J. Appl. Phys. **24**, L853 (1985)
11 W. Saitoh, K. Yamazaki, M. Asada, and M. Watanabe, Japan. J. Appl. Phys. **35**, L1104 (1996)
12 R.T. Tung and F. Schrey, Appl. Phys. Lett. **54**, 852 (1989)
13 A. Izumi, K. Tsutsui, N.S. Sokolov, N.N. Faleev, S.V. Gastev, S.V. Novikov, and N.L. Yakovlev, J. Cryst. Growth **150**, 1115 (1995)

5.3 Electron-Wave Coherent Coupling Devices

5.3.1 Coherent Coupling in Double Quantum Well

One of the most fundamental problems of quantum mechanics is the time evolution of a nonstationary wave packet [1–6]. The motion of such wave packet is purely coherent if phase-breaking interactions with a thermal bath are absent. Experimentally, we can observe such coherent quantum dynamics if the dimension is comparable to or less than the mean free path of the electrons.

When two quantum wells are placed sufficiently close to each other, the quantum-mechanical interaction modifies the wave function into symmetric and antisymmetric modes. Due to the dynamical interference of these modes, the electron wave performs an oscillatory tunneling motion (coherent electron oscillation).

We can express an eigenstate $\psi(t)$ of an electron in the coupled quantum well as a linear combination of the wave function of the two isolated quantum wells, i.e.,

$$\psi(t) = a(t)\psi_a + b(t)\psi_b, \tag{5.3.1}$$

where ψ_a and ψ_b are the wave functions of a single quantum well, satisfying

$$H_{a,b}\psi_{a,b} = E_{a,b}\psi_{a,b}, \qquad H_{a,b} = -\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial x^2} + V_{a,b}, \tag{5.3.2}$$

where m^* is the carrier effective mass and V_a and V_b are the potential energy functions of the isolated quantum wells.

We consider the time evolution of the electron wave function after an electron is initially launched in one of the two wells. The time evolution of the wave function is treated by the time-dependent Schrödinger equation:

$$i\hbar\frac{\partial\psi}{\partial t} = H\psi(t), \tag{5.3.3}$$

$$H = -\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial x^2} + V_a + V_b - V_0, \tag{5.3.4}$$

and V_0 is the potential barrier height between the two quantum wells. Substituting (5.3.1) into (5.3.3) and using (5.3.2), we get coupled equations for the amplitudes $a(t)$ and $b(t)$ of (5.3.1) [4]:

$$\frac{da(t)}{dt} = -i\omega_a a(t) - i\kappa_{ab}b(t), \quad \frac{db(t)}{dt} = -i\omega_b b(t) - i\kappa_{ba}a(t), \tag{5.3.5}$$

where $\omega = E_a/\hbar$, $\omega_b = E_b/\hbar$, and κ_{ab} (κ_{ba}) stands for the wave coupling coefficients from well a (b) to well b (a) in angular-frequency units. The coupling coefficient is directly related to the tunneling coefficient between the wells and given by a quasi-classical estimation [1]. For a symmetric coupled quantum well, $\kappa_{ab} = \kappa_{ba} \equiv \kappa(\hbar k/Lm^*)e^{-kL}$ with $k \equiv 2\pi/\lambda$, where $\lambda = \hbar/\sqrt{2m^*(V_0 - E_0)}$ is the de Broglie wavelength of the tunneling electron and E_0 is the electron

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

energy, i.e., $E_a = E_b \equiv E_0$. When the static bias E_{dc} is applied perpendicular to the well layers, the energy difference of the two levels is given by $\omega_a - \omega_b = (E_a - E_b)/\hbar = eE_{dc}d/\hbar$ where $-e$ is the electron charge and d is the center-to-center separation between the wells.

Note that equations (5.3.5) have the same form as those for an equation appearing in the coupling of optical mode in space [7] and for the Schrödinger equation for a two-level atomic system interacting with the electromagnetic (laser) field [8].

For an assumed $\exp(-i\omega t)$ dependence,

$$(\omega - \omega_a)(\omega - \omega_b) - \kappa_{ab}\kappa_{ba} = 0, \tag{5.3.6}$$

with the solution $\omega_\pm = (\omega_a + \omega_b)/2 \pm \omega_0$, where $\omega_0 = \sqrt{[(\omega_a - \omega_b)/2]^2 + \kappa_{ab}\kappa_{ba}}$. Therefore, the two modes, ψ_a and ψ_b of the isolated quantum wells are coupled and form two new modes which oscillate independently from each other, one with ω_+ and the other with ω_-. The difference $\Delta\omega = \omega_+ - \omega_-$ causes the phase difference between the symmetric and antisymmetric states leading to constructive or destructive interference of the wavefunction in each well.

Suppose that the electron waves in each well are given as $a = a(0)$ and $b = b(0)$ at $t = 0$. For these conditions the solutions are given as follows:

$$a(t) = \left[a(0)\left(\cos\omega_0 t + i\frac{\omega_a - \omega_b}{2\omega_0}\sin\omega_0 t\right) - i\frac{\kappa_{ab}}{\omega_0}b(0)\sin\omega_0 t\right]e^{-i(\omega_a+\omega_b)t/2}, \tag{5.3.7}$$

$$b(t) = \left[-i\frac{\kappa_{ab}}{\omega_0}a(0)\sin\omega_0 t + b(0)\left(\cos\omega_0 t + i\frac{\omega_a - \omega_b}{2\omega_0}\sin\omega_0 t\right)\right]e^{-i(\omega_a+\omega_b)t/2}. \tag{5.3.8}$$

For the symmetric coupled quantum well structures, i.e., $\omega_a = \omega_b$ and $\kappa_{ab} = \kappa_{ba} = \kappa$, and for the initial conditions $a(0) = 1$ and $b(0) = 0$, (5.3.7) and (5.3.8) become simple and the probabilities of finding the electron in each well are given by

$$|a(t)|^2 = \cos^2\kappa t, \qquad |b(t)|^2 = \sin^2\kappa t. \tag{5.3.9}$$

These equations mean that an electron, initially in one of the two wells, is transferred to the other at time $t = (n + 1/2)\pi/\kappa$ ($n = 0, 1, 2, \ldots$). At time $t = n\pi/\kappa$ the electron transfers completely back to the well initially injected with the electron. For complete electron transfer, the subband energy levels (E_a and E_b) of the isolated wells should therefore have the same value. If $E_a \neq E_b$, complete transfer is never obtained and only a partial transfer is realized.

5.3.2 Electron Directional Coupler

Many kinds of quantum interference devices have been proposed up to now [9]. These devices differ in two major respects from ordinary electronic devices. They employ quantum effects and do not rely on drift and diffusion of carriers, but rather on ballistic, non-phase-destroying transport. The reason for interest in these devices is their potential for ultrafast signal processing in compact structures.

Del Alamo and Eugster [10] and Tsukada et al. [4] independently proposed the concept of an electron directional coupler. It basically consists of two electron waveguides that, over a certain length, come in very close proximity to each other so that electron tunneling from one waveguide to the other is possible [11]. If the phase of the electron is preserved throughout the whole coupled region, coherent tunneling takes place between the two electron waveguides.

Since the original proposal, there have been a number of theoretical studies of electron directional coupling [12–20]. A number of them established the feasibility of the device using existing technology [4,10,12]. Others showed that 100% transfer between waveguides is only achievable if both waveguides are symmetric and contain a single subband [4,10,12,19,20]. If either one of these conditions is broken, incomplete transfer will result.

Yang and Xu have derived coupled differential equations for propagation of guided waves in two coupled waveguides [13]. They write the electron wave $\psi(z)$ propagating in the z direction as a linear combination of the eigenfunctions of the individual waveguide ψ_a and ψ_b, i.e., $\psi(z)=a(z)\psi_a+b(z)\psi_b$, and obtain the coupled differential equations for $a(z)$ and $b(z)$ by use of the time-independent Schrödinger equation. These equations have the same form as (5.3.5), if we replace the differential of time t with the differential of space z (see (49) and (50) in [13]). Therefore, the general solutions of these equations are also given by (5.3.7) and (5.3.8) if we replace t with z. If we launch a wave packet into one of the waveguides, this wave packet travels down the coupled region at the Fermi velocity v_F with a probability density that oscillates back and forth between the two waveguides. It takes a certain transfer length, $L_t = v_F t_T$ ($t_T = 2\kappa$: tunneling time), for the wave packet to transfer from one waveguide to the other. This length is a function of the strength of the coupling between the two waveguides, i.e., L_t becomes shorter for a lower or thinner potential barrier.

In the quest for the electron directional coupling, del Alamo and Eugster [11] have fabricated split-gate dual electron waveguide structures. Their devices allow the independent modulation of the electron subband population as well as of the strength of the potential barrier that separates the waveguides. They have observed surprising phenomena such as 1D to 2D tunneling between an electron waveguide and a neighboring 2D electron gas and 1D to 1D tunneling between two electron waveguides. Electron directional coupling, however, has not been observed because strong coupling with sharp confining potential barriers between the waveguides were not obtained by a split-gate technology. One approach to realize strong coupling with sharp potential barriers between the waveguides is to use a stacked structure [21] or a buried structure [22]. Noguchi and Ikoma [21] proposed a stacked structure, which has two waveguides grown vertically over one another and a gate voltage applied vertically. Wakaya et al. [22] fabricated a buried structure using in-situ focused ion beam implantation and regrowth by molecular beam epitaxy. Recently, irregularities of the conductance which are strongly related to the coherent tunneling and interference of electronic states in stacked double quantum waveguides were observed by Sawaki et al. [23].

In spite of tremendous efforts, unambiguous observation of coherent electron transfer (or electron directional coupling) remains to this date an elusive goal.

Recently, Tsukada et al. [24] proposed a nonlinear electron directional coupler. The device consists of two closely spaced, parallel electron waveguides with extremely small dimensions (or capacitances). For small input signals, the device behaves as a linear (ordinary) directional coupler as explained above. For large input signals, on the other hand, the electronic charge carried by the signals induces nonnegligible changes of the electrostatic potential in the waveguides and detunes the coupler.

By introducing the detuning (phase mismatching) term due to the Coulomb charging potential into the coupled equations (5.3.5), Tsukada et al. obtained nonlinear coupled equations with the same form as those for the nonlinear optical directional coupler, which was originally proposed by Jensen [25]. In the optical coupler, the nonlinear term arises from the interaction of guided modes with themselves through the nonlinear refractive index n_2. Similarly, in the electron directional coupler, the nonlinear coupling arises from the Coulomb charging (electrostatic) potential induced by the guided electrons themselves.

This device is capable of achieving exceedingly fast switching times, because the switching time of the device is limited not by the RC and transit times but only by the tunneling (transfer) time.

5.3.3 Coherent Oscillation Devices

An experiment has been proposed for direct observation of the time evolution in heterostructure barrier tunneling [3]. The idea of the experiment is to excite a coherent electron oscillation between coupled quantum wells and observe the luminescence signal from each well in a different frequency range. The proposed structure contains two quantum wells separated by a heterostructure barrier (see Fig. 1 in [3]); the wells have identical ground-state levels ($E_1 = E_2 \equiv E_0$) in the conduction band. The interband transition energies $h\nu_1$ and $h\nu_2$ are different in the two wells, which permits us to selectively excite electrons in one of the wells. If the excitation energy is tuned to the lower of the two frequencies, $h\nu_{ex} = h\nu_2$, then, immediately upon the excitation, electrons will be prepared in one of the wells. Subsequent evolution of the electron wave packet in time is given by (5.3.9), so that the electron densities in the two wells oscillate with the angular frequency $\omega_0 = \kappa$. The luminescence signals at frequencies ν_1 and ν_2 will oscillate 180° out of phase. Their intensities will be proportional to $\sin^2\kappa t$ and $\cos^2\kappa t$, respectively. However, this phenomenon has not been observed so far.

Recently, Leo et al. [26] have seen strong evidence for coherent electron oscillations in coupled quantum well structure from optical pump-probe and degenerate four-wave mixing experiments. A much more direct probe of such charge oscillations is the observation of electromagnetic dipole radiation emerging from these electron oscillations. Roskos et al. [27] have reported the first observation of oscillatory coherent submillimeter-wave emission from a GaAs/AlGaAs coupled quantum wells at a sample temperature of 10 K with

160 fs laser pulses. Using time-domain THz spectroscopy, they observed up to fourteen cycles of the electromagnetic radiation at 1.5 THz emitted from the oscillating electrons before phase relaxation destroyed the coherence of the electronic wave packet. A time-dependent polarization $P(t)$ is associated with the coherently oscillating wave packet, leading to dipole radiation with a radiated electromagnetic field $E(t) \approx d^2P(t)/dt^2$.

Soon afterward, Planken et al. [28] reported on the observation of coherent THz radiation, tunable from 1.4 to 2.6 THz, emerging from GaAs/AlGaAs single quantum wells from the coherent excitation of light hole (lh) and heavy hole (hh) excitons. They attributed the radiation to charge oscillations following coherent excitation of the excitons in an electric field. The measured THz frequencies were shown to be in excellent agreement with the measured hh-lh splittings.

Recently, quantum beats in a compositionally asymmetric GaInAs/GaAs quantum well in AlAs barriers have been observed by exciting the sample with 12 fs optical pulses [29]. Because of the larger transition energy, in the 100 meV range, the wave packet oscillation was observed even at room temperature. The oscillation period of 33 fs corresponds to an intersubband transition energy $\Delta\omega = 126$ meV, i.e., an emission wavelength of 9.8 μm.

All the experimental results [26–29] reported so far show that coherent oscillations in quantum wells can be observed only an extremely short time alter the optical excitation. From the practical point of view, continuous self-oscillations by exploiting some kind of feedback mechanism are desired. However, it is still premature to discuss such devices.

5.3.4 Bloch Oscillation Devices

In a pioneering paper, Esaki and Tsu [30] proposed to obtain high frequency oscillators by tailoring the nonlinear electronic properties of semiconductor superlattices. Electrons accelerated by a moderate, constant electric field E_{dc} can perform a repetitive motion of acceleration and Bragg reflection called Bloch oscillation, characterized by the Bloch frequency $\omega_B = eE_{dc}d/\hbar$, where d is the superlattice period. It took more than two decades before the first observation of Bloch oscillations was made in a degenerate four-wave mixing experiment [31] and in a transient THz emission experiment on optically excited electrons [32].

Waschke et al. [32] have directly observed the coherent electromagnetic radiation originating from Bloch oscillations of charge carriers in an electrically biased semiconductor superlattice structure. The oscillation frequency can be tuned with the applied bias field from 0.5 THz to more than 2 THz at a sample temperature of 15 K with 100 fs optical pulses. Increasing the bias from flatband conditions, one tunes the coupling between wells and the degree of localization of the electron and hole wave functions. Each energy miniband splits into a series of discrete levels forming a Wannier–Stark ladder.

All attempts to generate continuous Bloch emission from electrically injected carriers in a superlattice have failed. Ignatov et al. [33] reported on a theoretical study of the efficiency of a superlattice oscillator. Making use of an

analogy between the dynamics of carriers in a miniband and Josephson junction dynamics [34], they showed that emission of electromagnetic radiation from superlattices can occur due to an increase of the multiphoton transitions (see Sect. 5.3.5).

One of the current topics intimately related to this research field is to observe the Bloch oscillation and the Wannier–Stark ladders, two fundamental quantum effects in solid state physics, for ultracold atoms in a periodic optical potential [35–37]. A spatially periodic optical potential yields an energy spectrum of Bloch bands for the atoms. Dahan et al. [36] have successfully observed Bloch oscillations of the atoms driven by a constant inertial force. Wilkinson et al. [37] reported observation of Wannier–Stark ladders with ultracold atoms in an accelerating one-dimensional standing wave of light. Atoms are trapped in a far-detuned standing wave that is accelerated for a controlled duration of time.

5.3.5 Coherent Oscillations in ac-Field

The problem of coherent tunneling through the barrier of a double well or a superlattice in the presence of an external oscillating (ac or laser) field has received a considerable amount of attention recently [38–41].

The motion of a charged particle on a linear chain of sites m $(-\infty < m < +\infty)$ under the combined action of a time-dependent electric field $E(t) = E_{ac}\sin\omega t$ in the direction of the lattice and of nearest-neighbor coupling (or tunneling) V has been analyzed by Dunlap and Kenkre [42]. The Hamiltonian for the system is

$$H(t) = V\sum_{m=-\infty}^{\infty}\big(|m\rangle\langle m{+}1| + |m{+}1\rangle\langle m|\big) - eE(t)d\sum_{m=-\infty}^{\infty} m\big(|m\rangle\langle m|\big). \quad (5.3.10)$$

In this expression, $|m\rangle$ represents a Wannier state localized on lattice site m, d is the lattice constant, and e is the charge on the particle. By expressing the particle state $\psi(t)$ as a linear combination of Wannier states $|m\rangle$, $\psi(t) = \sum_m C_m(t)|m\rangle$, where $C_m(t)$ are the time-dependent amplitudes, one obtains the following evolution equation for the amplitudes $C_m(t)$:

$$\frac{dC_m}{dt} = \frac{imeE(t)d}{\hbar}C_m - iV(C_{m+1} + C_{m-1}). \quad (5.3.11)$$

This expression is identical to (5.3.5), if we consider the linear-chain with two sites and $\omega_a - \omega_b = eE(t)d/\hbar$. They have obtained exact results for arbitrary time dependence of the electric field. The electron is localized for the cases when the ratio of the field magnitude and the field frequency is a root of J_0 where J_n is the ordinary Bessel function. This localization could be used, in principle, for inducing anisotropy in the transport properties of an ordinarily isotropic material. Holthaus [40] has independently analyzed quasi-energy minibands in superlattices which interact with intense far-infrared radiation and shown that the miniband width becomes close to zero if the ratio of the Bloch frequency and the laser frequency approaches a zero of J_0. They called this phenomenon miniband collapse.

Grossmann et al. [38] reported on an interesting effect of a CW laser acting on an electron in a double well. If the electron is initially localized in one of the two wells, and if the laser power and frequency are chosen appropriately, the radiation field can prevent coherent electron oscillation. They called this phenomenon coherent destruction of tunneling. Bavli and Metiu [39] reported on a more complex situation. They showed that a semi-infinite laser field acting on a ground-state (delocalized) electron localizes it in one of the wells and then confines it there. They have given only some numerical results and not shown necessary conditions to realize the localization. Recently, Tsukada et al. [41] reported that the localization is realized if the ratio of the field magnitude and the field frequency is a root of J_n for n-photon-assisted resonances. They also pointed out that the miniband collapse of the superlattice and the destruction of coherent tunneling are physically identical.

In order to observe such phenomena, the two conditions $\omega\tau > 1$ and $eE_{ac}d \geq \hbar\omega$ should be satisfied. The former statement implies that the electrons should be submitted to at least one cycle of the applied ac field before being scattered in a mean time τ, while the latter condition requires that the high frequency electric field of magnitude E_{ac}, applied across a nanostructure with a width d, produces an ac voltage drop that exceeds the photon energy $\hbar\omega$. Semiconductor superlattices typically have scattering time $\tau \approx 10^{-12}$ s and a spatial period $d \approx 10$ nm. Near-infrared or higher frequency radiation can easily achieve $\omega\tau > 1$, but at these frequencies, high-power tunable sources are only achieved using free-electron lasers (FEL). Major contributions to the experiments for observation of photon-assisted tunneling have been made by using FEL at University of California, Santa Barbara (UCSB) [43–46]. These laser systems routinely provide tunable radiation at the kilowatt power level from 120 GHz to 5 THz. A typical pulse lasts for a few seconds with a repetition rate of 0.5 Hz, and can be focused down to a spot diameter of a few millimeters [47].

Recently, features associated with photon-assisted tunneling have been observed in semiconductor superlattices [43,44] and quantum dots [48]. By using the facilities of UCSB, photon-assisted tunneling in a resonant tunneling diode has been observed [45]. The dc I-V characteristic of a triple-barrier resonant tunneling diode integrated in a bow-tie antenna and driven by THz radiation displays up to five additional resonant tunneling channels whose voltage positions vary linearly with frequency in the investigated range between 1.0 and 3.4 THz. These peaks were attributed to photon-assisted tunneling processes corresponding to absorption and stimulated emission of up to three photons. The experiments suggest that such a device can be utilized to detect and generate THz radiation. The power emitted by their device was estimated to be 0.6 W/cm^2 at 2 THz.

Unterrainer et al. [46] have observed the same phenomenon in superlattices. They called it the inverse Bloch oscillation effect defined by resonant changes in the current-voltage characteristics of miniband semiconductor superlattices when the Bloch frequency is resonant with a THz field and its harmonics. The samples used in this study were homogeneously Si doped GaAs/AlGaAs superlattices consisted of 40 periods of 80 Å wide GaAs wells

and 20 Å thick AlGaAs barriers. Superlattice mesas with an area of 8 μm^2 were formed by dry etching and an ion (H^+) implantation-isolation process. Ohmic Au/Ge/Ni contacts were fabricated on the top and the bottom of the mesas. A coplanar broadband bow-tie antenna was used to couple the THz electric field to the mesa parallel to the growth direction and enhance the electric field by a factor of 50 to 100. The experiments were performed at 10 K in a temperature controlled flow cryostat with Z-cut quartz windows. The conductance of the superlattice was measured during the microsecond long pulses of THz radiation provided by the UCSB free electron lasers. At the highest intensities they observed an increase of the photocurrent in the negative dc region and attributed it to the stimulated emission of photons (electrons move down the Wannier–Stark ladder). The experiment demonstrates that a superlattice is a possible gain medium for THz oscillators; however, the gain (1.3×10^{-5}) is very small and would hardly overcome resonator losses.

Continuing progress on crystal growth technologies, providing high quality crystals with longer scattering times, and optimization of device structures will realize continuous THz oscillations by superlattice devices or by the coupled quantum well devices in the near feature. These devices may be useful for tunable quantum detectors and radiation sources in the THz range.

References

1 L.D. Landau and E.M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory, 3rd ed.* (Pergamon, London, 1977), Sect. 50.
2 A.J. Legett, S. Chakravarty, A.T. Dorsey, M.A. Fisher, A. Garg, and W. Zwerger, Rev. Mod. Phys. **59**, 1 (1987)
3 S. Luryi, Solid State Commun. **65**, 787 (1989); IEEE J. Quantum Electron. **QE27**, 54 (1991)
4 N. Tsukada, A.D. Wieck, and K. Ploog, Apl. Phys. Lett. **56**, 2527 (1990)
5 C. Juang and J.H. Chang, IEEE J. Quantum Electron. **28**, 2039 (1992)
6 V. Sankaran and J. Singh, Appl. Phys. Lett. **58**, 1590 (1991)
7 H. Kressel and J.K. Butler, *Semiconductor Lasers and Heterojunction LEDs* (Academic Press, New York, 1977), Chapter 15.
8 See, for example, L. Allen and J.H. Eberly, *Optical Resonance and Two-Level Atoms* (Wiley, New York, 1975).
9 See, for example, Phys. Today **43**, No. 2 (1990), Special Issue on Nanoscale and Ultrafast Devices
10 J.A. del Alamo and C.C. Eugster, Appl. Phys. Lett. **56**, 78 (1990)
11 J.A. del Alamo and C.C. Eugster, Jpn. J. Appl. Phys. **34**, 4439 (1995) and references therein
12 N. Dagli, G. Snider, J. Waldman, and E. Hu, J. Appl. Phys. **69**, 1047 (1991).
13 R.Q. Yang and J.M. Xu, Phys. Rev. B **43**, 1699 (1991)
14 J. Wang, J. Guo, and R. Harris, Appl. Phys. Lett. **59**, 3070 (1991)
15 L. Thyen and O. Sahen, IEEE J. Quantum Electron. **28**, 31 (1992)
16 P. Zhao, Phys. Rev. B **45**, 4301 (1992)
17 T. Singh and M. Cahay, Superlattices and Microstructures **11**, 253 (1992)
18 J.-Q. Wang, B.-Y. Gu, and G.-Z. Yang, J. Appl. Phys. **72**, 2299 (1992)
19 D.W. Wilson, E.N. Glytsis, and T.K. Gaylord, J. Appl. Phys. **73**, 3352 (1993)
20 S. Nara, Y. Hara, H. Totsuji, Y. Nomura, and N. Tsukada, Jpn. J. Appl. Phys. **32**, 1210 (1993)
21 M. Noguchi and T. Ikoma, Seisan-Kenkyu **45**, 148 (1993)
22 F. Wakaya, Y. Yuba, S. Takaoka, K. Murase, and K. Gamo, Jpn. J. Appl. Phys. **32**, 6242 (1993); F. Wakaya, A. Nozawa, J. Yanagisawa, Y. Yuba, S. Takaoka, K. Murase, and K, Gamo, Microelectron. Eng. **23**, 123 (1994)

23 M. Yamaguchi, K. Masukawa, N. Ikedo, N. Sawaki, K. Fujii, and H. Goto, Physica B **227**, 387 (1996)
24 N. Tsukada, M. Gotoda, M. Nunoshita, and T. Nishino, Phys. Rev. B **53**, R7603 (1996); See also, N. Tsukada, M. Gotoda, and M. Nunoshita, Phys. Rev. B **50**, R5764 (1994); N. Tsukada, M. Gotoda, M. Nunoshita, and T. Nishino, Phys. Rev. B **52**, R17005 (1995)
25 S.M. Jensen, IEEE J. Quantum Electron. **QE-18**, 1580 (1982)
26 K. Leo, J. Shah, E. O. Gobel, T. C. Damen, S. Schnitt-Rink, W. Schafer, and K. Kohler, Phys. Rev. Lett. **66**, 201 (1991); K. Leo, P.H. Bolivar, F. Bruggemann, R. Schwedler, and K. Kohler, Solid State Commun. **84**, 943 (1992)
27 H.G. Roskos, M.C. Nuss, J. Shah, K. Leo, D.A.B. Miller, A.M. Fox, S. Schmitt-Rink, and K. Kohler, Phys. Rev. Lett. **68**, 2216 (1992)
28 P.C. Planken, M.C. Nuss, I. Brener, K.W. Goossen, M.S.C. Luo, S.L. Chuang, and L. Pfeiffer, Phys. Rev. Lett. **69**, 3800 (1992)
29 A. Bonvalet, J. Nagle, V. Berger, A. Migus, J.-L. Martin, and M. Joffre, Phys. Rev. Lett. **76**, 4392 (1996)
30 L. Esaki and R. Tsu, IBM J. Res. Dev. **14**, 61 (1970)
31 J. Feldmann, K. Leo, J. Shah, D.A.B. Miller, J.E. Cunningham, T. Meier, G. von Plessen, A. Schulze, P. Thomas, and S. Schmitt-Rink, Phys. Rev. B **46**, 7252 (1992)
32 C. Waschke, H.G. Roskos, R. Schwedler, K. Leo, H. Kurz, and K. Kohler, Phys, Rev. Lett. **70**, 3319 (1993)
33 A.A. Ignatov, K.F. Renk, and E.P. Dodin, Phys. Rev. Lett. **70**, 1996 (1993)
34 P.K. Tien and J.P. Gordon, Phys. Rev. **129**, 647 (1963)
35 Q. Niu, X-G. Zhao, G.A. Georgakis, and M.G. Raizen, Phys. Rev. Lett. **76**, 4504 (1996)
36 M.B. Dahan, E. Peik, J. Reichel, Y. Castin, and C. Salomon, Phys. Rev. Lett. **76**, 4508 (1996)
37 S.R. Wilkinson, C.F. Bharucha, K.W. Madison, Q. Niu, and M.G. Raizen, Phys. Rev. Lett. **76**, 4512 (1996)
38 F. Grossmann, T. Dittrich, P. Jung, and P. Hanngi, Phys. Rev. Lett. **67**, 516 (1991)
39 R. Bavli and H. Metiu, Phys. Rev. Lett. **69**, 1986 (1992)
40 M. Holthaus, Phys. Rev. Lett. **69**, 301 (1992); **69**, 1696 (1992)
41 N. Tsukada, M. Gotoda, T. Isu, M. Nunoshita, and T. Nishino, Jpn. J. Appl. Phys. **35**, L1490 (1996)
42 D.H. Dunlap and V.M. Kenkre, Phys. Rev. B **34**, 3625 (1986); **37**, 6622 (1988); S. Raghavan, V.M. Kenkre, D.H. Dunlop, A.R. Bishop, and M.I. Salkola, Phys. Rev. A **54**, R1781 (1996)
43 P.S.S. Guimaraes, B.J. Keay, J.P. Kaminsky, S.J. Allen, P.F. Hopkins, A.C. Gossard, L.T. Florez, and J.P. Harbison, Phys. Rev. Lett. **70**, 3792 (1993)
44 B.J. Keay, S. Zeuner, S.J. Allen, Jr., K.D. Maranowski, A.C. Gossard, U. Bhattacharya, and M.J.W. Rodwell, Phys. Rev. Lett. **75**, 4102 (1990)
45 H. Drexler, J.S. Scott, S.J. Allen, K.L. Campman, and A.C. Gossard, Appl. Phys. Lett. **67**, 2816 (1995)
46 K. Unterrainer, B.J. Keay, M.C. Wanke, and S.J. Allen, Phys. Rev. Lett. **76**, 2973 (1996)
47 G. Ramian, Nucl. Instrum. Methods Phys. Res. A **318**, 225 (1992)
48 L.P. Kouwenhoven, S. Jauhar, J. Orenstein, P.L. McEuen, Y. Nagamune, J. Motohisa, and H. Sakaki, Phys. Rev. Lett. **73**, 3443 (1994)

5.4 Electron-Wave Diffraction Devices

5.4.1 Electron Wavefront and Its Manipulation

By analogy with optics, for example, Fourier transforms can be performed by the electron wave. Important for the realization of such devices is the spread of the electron wavefront, which is the subject of this section [1]. Figure 5.4.1 shows a conceptual electron wave device where the electron wave from a source passes through an input layer where the amplitude and/or the phase of the electron wave is modulated according to the input signal pattern $I(x)$. Then the electron wave forms a diffraction or interference pattern of the probability current density on the output layer. This conceptual device performs pattern transformation using the nature of the wave propagation. The wavefront spread is determined by the emitter and the propagation. Here we focus on the former using an effective-mass Schrödinger equation.

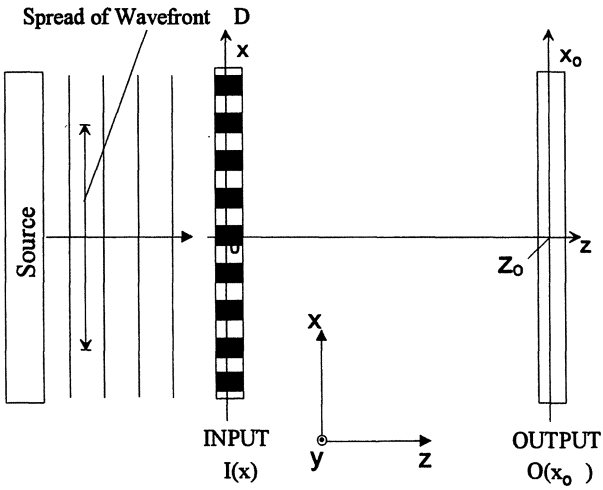


Fig. 5.4.1 Conceptual electron wave device and spread of wavefront

If the electron wave is expressed as a plane wave such as $\exp(i\mathbf{k}\cdot\mathbf{r})$ which is a momentum eigenstate, the spread of the electron wavefront is infinite. However, in the source region, an electron occupies various states of different wave vectors $\mathbf{k}$. The electron wave is expressed as $\sum_{\mathbf{k}} c(\mathbf{k})\exp(i\mathbf{k}\cdot\mathbf{r})$ where $C(\mathbf{k})$ is the expansion coefficient. We consider a tunnel emitter consist of a double heterojunction across which a voltage is applied. Electrons in thermal equilibrium in the cathode region pass through the tunnel barrier to become hot electrons in the propagation region. The wave function Φ is represented as:

$$\Phi(\mathbf{r},\mathbf{k}) = \begin{cases} N[\exp(i\mathbf{k}\cdot\mathbf{r}) + \text{reflecting wave}] & \text{(in the cathode region)} \\ Nt(\mathbf{k})\exp(i\mathbf{k}'\cdot\mathbf{r}) & \text{(in the propagation region)}, \end{cases} \tag{5.4.1}$$

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

where $\mathbf{r}$ is the position vector, $\mathbf{k}$ the wave vector in the cathode region, $t(\mathbf{k})$ the transmission coefficient at the barrier, and $\mathbf{k}'$ the wave vector in the propagation region. We have $k_z'^2 = k_z^2 + 2m^* eV_z/\hbar^2$ where V_z is the applied voltage and m^* is the effective mass. A reservoir is assumed to supply electrons so that one electron exists in the state $\mathbf{k}$ in the cathode region at the Fermi–Dirac occupation probability $f(E)$ where E is the energy of the state $\mathbf{k}$.

In the propagation region, the electron wave given by (5.4.1) passes through the aperture (the input plane) to cause the diffraction. Because of the coincidence of the equation form, the diffraction of a hot electron wave is analyzed based on Kirchhoff's diffraction theory for the electromagnetic wave to obtain the diffracted wavefunction $\Psi(\mathbf{r}_0;\mathbf{k})$. To calculate the probability current density $\mathbf{J}(\mathbf{r}_0)$ in the output plane, we calculate $\sum C^*(\mathbf{k}_1)\Psi^*(\mathbf{r}_0;\mathbf{k}_1)\times \sum C(\mathbf{k}_2)\vec{\nabla}\Psi(\mathbf{r}_0;\mathbf{k}_2)$. The double summation produces cross terms having the factor $C^*(\mathbf{k}_1)C(\mathbf{k}_2)$. The phase angle of $C^*(\mathbf{k}_1)C(\mathbf{k}_2)$ varies almost at random when $\mathbf{k}_1 \neq \mathbf{k}_2$ and the cross terms are small enough to be neglected. Then, the overall probability current density is given as $J_z(\mathbf{r}_0) = (\hbar k_z'/m^*)\sum |C(\mathbf{k})|^2|\Psi(\mathbf{r}_0;\mathbf{k})|^2$, where the summation is over all states $\mathbf{k}$ in the cathode region. $|C(\mathbf{k})|^2$ is the occupation probability of the state $\mathbf{k}$ and is represented by the Fermi–Dirac function $f(E)$, where E is the energy of the state $\mathbf{k}$ [2].

The interference pattern caused by a double slit uniform in the y direction will be calculated. As an idealized double slit, $S(x) = s[\delta(x+d/2)+\delta(x-d/2)]$ is used as the aperture function, where s is the product of the transmission coefficient and the aperture width and d is the spacing between two slits. The contrast of the interference pattern is evaluated using the visibility $(J_{\max} - J_{\min})/(J_{\max}+J_{\min})$, where $J_{\max}$ and $J_{\min}$ are the probability current density at the center peak and the trough nearest to the center peak, respectively. The visibility depends on the temperature and the Fermi energy of the cathode while it is almost independent of the voltage across the barrier. The visibility decreases when the slit spacing increases. When the spread of the wavefront is less than the spacing, interference does not occur. Thus we consider that the maximum spacing of the double-slit providing visibility beyond the critical value, for example $1/e$, gives the effective spread of the electron wavefront. As the result of the present analysis, the effective spread of the wavefront increases when the temperature and/or the Fermi energy decreases as shown in Fig. 5.4.2. For example, when the Fermi energy is 5 meV, the spread is 60 and 27 nm at 4 and 77 K, respectively.

5.4.2 Coherence of Electron Wave

a. Phase Breaking Time Required for Interference

Let us consider an interference switch. The electron wave is divided into two waves ψ_1 and ψ_2 at the branching point B which propagate along waveguides 1 and 2 and meet at the merging point M and the output is proportional to $|\psi_1+\psi_2|^2$. When phase shifts in both waveguides are identical, the interference is constructive to make the output maximum, while when the phase shifts differ by π, the interference is destructive to cause the minimum output.

To obtain a high extinction ratio, the probability of the phase breaking in the waveguides 1 and 2 must be sufficiently small. That is, the transit time $\tau_{\rm tr}$ should be much smaller than the phase breaking time $\tau_{\rm ph1}$. The minimum transit time of the interference switch is as follows. When the length of the waveguide is l, the potential change V given by $E[1-(1-\pi/k_0 l)^2]$ (k_0 is the wavenumber at $V=0$, E is the electron energy) makes the phase-shift difference of π. The transit time is $lk_0(\hbar/2E)/(1-\pi/k_0 l)$. As numerical examples, when E is 10 meV, $k_0 l$ is 2π (5π), the modulation potential V is 7.5 (3.6) meV and the transit time is 0.4 (0.6) ps. When E is 200 meV, these are 150 (72) meV and 0.02 (0.03) ps. Provided that the effective mass is $0.04m$ and $k_0 l$ is 2π (5π), l is 64 (160) nm and 14 (35) nm for E of 10 meV and 200 meV, respectively.

The phase breaking before the branching point B also degrades the extinction ratio. To obtain a high output ratio, the difference $\Delta\tau$ in the transit time of two waveguides must be much smaller than the phase breaking time $\tau_{\rm ph2}$ in the part before point B. The minimum difference $\Delta\tau_{\rm tr}$ is $\pi\hbar/2/(1-\pi/k_0 l)$ when two waveguides have equal length and the phase-shift difference is π. As numerical examples, when E is 10 meV and $k_0 l$ is 2π (5π), $\Delta\tau_{\rm tr}$ is 0.2 (0.1) ps. When E is 200 meV, $\Delta\tau_{\rm tr}$ is 0.01 (0.005) ps. As shown in these examples, $\tau_{\rm ph2}$ can be shorter than $\tau_{\rm ph1}$.

b. Energy Sharpness Required for Interference

The energy width of the electron wave also degrades the extinction ratio. When the potential change V is applied to one waveguide, the difference in the phase shift between two waveguides depends on the energy E as $\delta(E) = k_0 l[\sqrt{E}-\sqrt{E-V}]/\sqrt{E}$. Denoting an energy spectrum function by $G(E)$, the output of the switch is given by $1+\int G(E)\cos\delta(E)dE$. Then, the visibility, $(J_{\rm max}-J_{\rm min})/(J_{\rm max}+J_{\rm min})$, is given by $-(\pi/2)^2\Delta E^2/E_0^2$ where ΔE^2 is the second moment of $G(E)$. To obtain visibility higher than 0.9, the relative energy width $\sqrt{\Delta E^2}/E_0$ should be less than 0.2.

c. Phase Breaking Time Estimated

In metals and degenerate semiconductors, electron-electron scattering is the dominant phase breaking process and has been studied theoretically and experimentally [2–5]. Electrons interact with each other through the Coulomb potential. However, at 0 K and at the Fermi energy, since all final states are occupied, transitions are forbidden and, therefore, no scattering occurs. When the temperature deviates from 0 K or the energy is different from the Fermi energy, scatterings take place at a rate proportional to the square of the temperature or the energy deviation from the Fermi energy [6]. The phase breaking time in a modulation doped structure was measured to agree with the theory, i.e., 30 ps at 1.4 K and for energy 1 meV above $E_F = 13.7$ meV [4].

The phase breaking time in insulators or intrinsic semiconductors is determined by an LO phonon emission to be about 0.15 ps when the electron energy exceeds the LO phonon energy [7]. For example, the LO phonon energy of GaAs is 36 meV. The spontaneous LO phonon emission rate does not depend on the temperature.

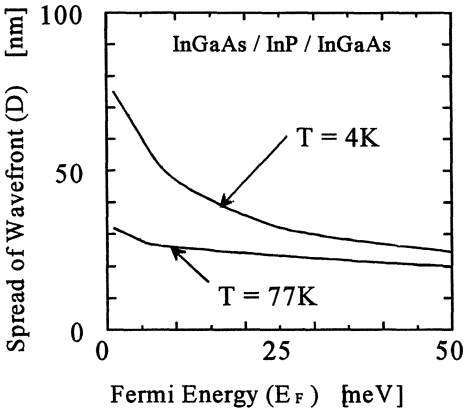


Fig. 5.4.2 Spread of wavefront versus the Fermi energy

Based on the above, below 1 K and at an energy close enough to the Fermi energy, the condition of a high contrast interference is satisfied for metals and degenerate semiconductors. In insulators or intrinsic semiconductors, when the electron energy is 200 meV, for example, the minimum transit time and the transit time difference in the interferometer are 0.02 and 0.01 ps which are much less than the phase breaking time 0.15 ps. This means that there is a good possibility to achieve interference devices. For thermal equilibrium electrons, the energy width is k_BT (k_B is the Boltzmann constant). For example, at a temperature of 20 K and an electron energy of 10 meV, the relative width of the electron energy is small enough to obtain a visibility of 0.9. For nonequilibrium electrons generated by a tunnel emitter, for example, parameters of the emitter can be designed to meet the condition.

d. Coherence of Electron Wave

A method for the estimation of the phase coherence length has been proposed using double-barrier resonant tunneling diodes (DBRTDs) [8,9] where the phase coherence length L_C is related to the resonant energy width ΔE. The ΔE is a full width at half maximum of the transmittance evaluated from the second derivatives of current-voltage (I-V) characteristics [10]. However, this method using DBRTDs is not applicable for temperatures above 10 K because of the bluntness of the energy distribution function at Fermi level in the emitter electrode. Furthermore the variable range of the Fermi level for the method is restricted to be larger than ΔE and less than $E_R - \Delta E$, where E_R denotes the resonant energy level. Since ΔE is evaluated from d^2I/dV^2 around a voltage where the Fermi level E_F in the emitter aligns to E_R, the method is not available for estimating L_C of non-equilibrium electrons such as hot electrons whose energies are far above the Fermi energy.

By using triple-barrier (TB) RTDs instead of DBRTDs, it is possible to overcome the above limitations. The typical conduction diagram of a TBRTD

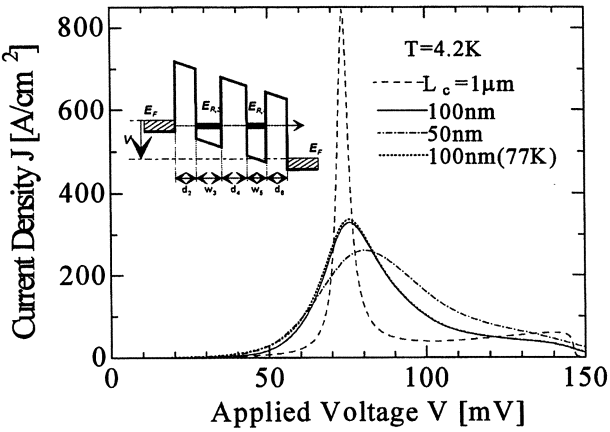


Fig. 5.4.3 Numerical calculation of J-V characteristics in GaInAs TBRTD. d_2, w_3, d_4, w_5, and d_6 are assumed to be 6.0, 9.0, 9.0, 5.5, and 6.0 nm respectively, and $E_{R,3}$, $E_{R,5}$, and E_F are assumed to be 44, 77, and 40 meV, respectively

is illustrated in the inset of Fig. 5.4.3. In TBRTD, I-V characteristics are almost governed by the alignment of energy levels in two wells. Therefore the I-V characteristics reveal a current peak at a certain applied voltage. The L_C can be estimated from the voltage width of peak current in I-V characteristics even at high temperatures when structural parameters of TBRTD are set appropriately. For TBRTDs, conditions for showing a resonant peak in I-V curve whose width is determined predominantly by the phase-breaking process are summarized as follows [11]. (1) The center barrier d_4 is thick enough to suppress excess broadening of the peak width due to the strong coupling between two wells (i.e., $d_4 > 7$ nm is applicable for GaInAs/InP TBRTDs). (2) The quantized energy level of region 3 is lower than that of region 5 (i.e., $E_{R,3} < E_{R,5}$). (3) E_F in the emitter is higher than the resonant level and E_F in the collector is lower than the level when a voltage is applied between the emitter and the collector so that two levels line up. The variability of the Fermi level selection in the emitter is freer in the TBRTD method than in the DBRTD method.

Numerical results of the current-density versus voltage (J-V) characteristics in a GaInAs/InP TBRTD are shown in Fig. 5.4.3. The transmission coefficient in the TB structure is calculated by a transfer matrix analysis taking phase breaking processes into account as an imaginary part of the wavenumber in the wells [11]. Both coherent and incoherent tunneling processes have been considered so as to reproduce experimental values [12]. The peak width of J-V characteristics does not depend on the temperature but strongly depends on L_C, while the product of the peak current and the peak voltage remains almost constant.

The phase coherence length L_C is related to the full voltage width ΔV at half maximum (FWHM) of the current by $L_C = 2\hbar v/[e(\eta_R - \eta_L)\Delta V]$ where v is the group velocity of the electron at the resonant level in the right well. In

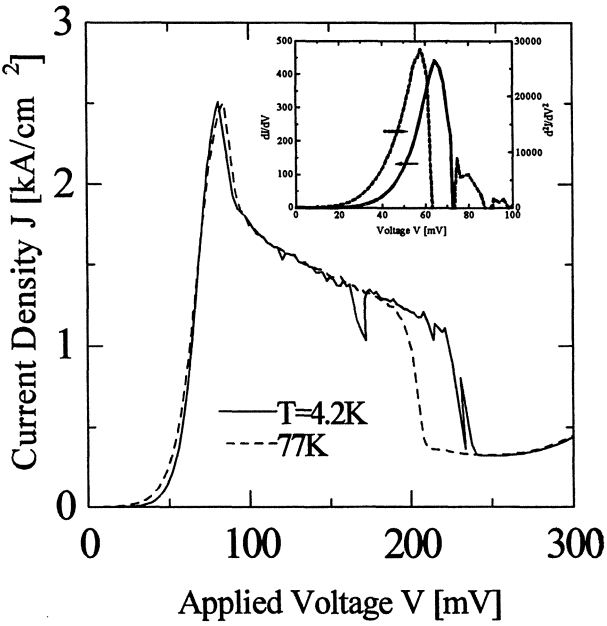


Fig. 5.4.4 Measured J-V characteristics in GaInAs/InP TBRTDs grown by OMVPE. Nominal values of structural parameters were same as Fig. 5.4.3. The derivatives are indicated in the inset

actual experiments, ΔV is often difficult to evaluate because of instabilities in the negative differential resistance (NDR) region. Therefore, a method using derivatives of J-V characteristics has been proposed [13]. In this case, ΔV is related to peak voltage values of dJ/dV and d^2I/dV^2 when the peak shape of J-V curve is assumed to be Lorentzian. The relation between L_C and the peak voltage difference of the first and the second derivatives can be derived as $L_C = (1-1/\sqrt{3})\hbar v/[e(\eta_R-\eta_L)\Delta V]$ where, $\eta_R = (d_2+w_3+d_4+w_5/2)/(d_2+w_3+d_4+w_5+d_6)$ and $\eta_L = (d_2+w_3/2)/(d_2+w_3+d_4+w_5+d_6)$. In this method ΔV can be evaluated even in the NDR region.

Measured J-V characteristics and derivatives in GaInAs/InP TBRTDs fabricated by organometallic vapor phase epitaxy (OMVPE) or metalorganic chemical vapor deposition (MOCVD) are shown in Fig. 5.4.4. In these experiments, ΔV was estimated to be 5.5 mV at 4.2 K and 9.0 mV at 77 K, as shown in an inset, and the values of L_C were estimated as 90 nm for 4.2 K and 55 nm for 77 K, respectively.

Spatial inhomogeneity of the heterointerface roughness or impurities in the TB structure and the electrode can make the resonant level width broaden and make the apparent L_C shorten. Thus the experiments give only a lower bound for L_C. By applying an atomic order flattening growth technique [14] for TBRTDs, it becomes possible to reduce the effects of interface roughness on apparent L_C.

The phase coherence length L_C estimated by using RTDs is that of elec-

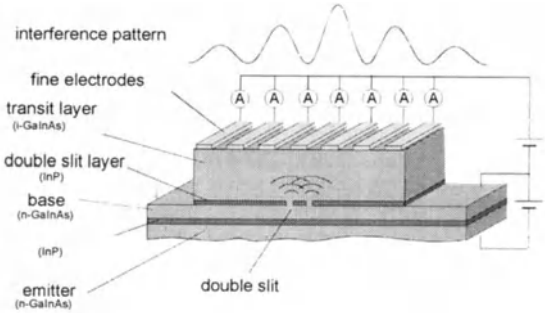


Fig. 5.4.5 Proposed device structure for hot electron interference by double slit

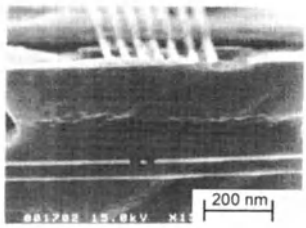


Fig. 5.4.6 SEM view of aligned stripe electrodes

trons confined in wells, i.e., 2D electrons. The estimation of L_C for 3D electrons requires other techniques.

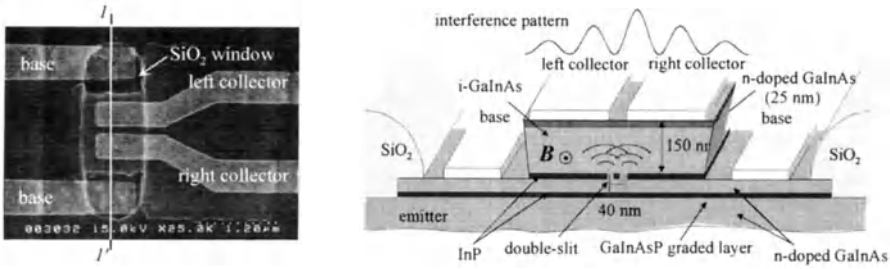
5.4.3 Diffraction of Hot Electron Wave

Devices using diffraction of hot electrons have been proposed [15]. A schematic view of the device is shown in Fig. 5.4.5. It consists of a buried double slit whose size t is comparable to the electron wavelength and to the fine electrodes for detection. A simulation shows that the pitch of the detection electrodes must be smaller than the period of the interference pattern. This means that the product of the slit pitch and the electrode pitch must be smaller than 2000 nm^2.

To fabricate such fine structures, electron beam lithography (EBL) has been employed [16]. A 40 nm pitch buried double slit has been fabricated by EBL and OMVPE regrowth. The formation of the fine semiconductor structure by wet etching is restricted by the patterning of e-beam resist rather than by the wet etching process. If we could obtain a finer resist pattern (~20 nm pitch), finer structures can be fabricated by our etching technique. Fine metal wires with 50 nm pitch on a mesa have been formed by the EBL and liftoff technique. The product of these pitches is 2000 nm^2, which satisfies the required size condition. It is important not only to fabricate these fine structures but also to align them for actual devices. Using $W/Cr/Au/SiO_2$ masks, 50 nm pitch stripe electrodes have been formed on a buried heterostructure of 40 nm pitch with ~20 nm of relative positional shift between two structures as shown in Fig. 5.4.6.

The contact currents of 100 nm pitch Cr/Au electrodes are measured to demonstrate electrical properties of the nanostructures [17]. The current passing through each metal wire is measured to show the absence of any disconnection in the wires. Six values of contact resistivity per area are obtained with a deviation within 10%. This shows the possibility of detecting hot electron interference by using these fine electrodes. However, the reproducibility in the measurement is not good for electrodes with a pitch less than 70 nm.

A magnetic field has been used for actual observation of diffraction effects of hot electrons [18]. The device (Fig. 5.4.7) used in this experiment has a



(a) (b)

Fig. 5.4.7 Device structure for a double-slit experiment on hot electrons in a semiconductor under a magnetic field. (**a**) Top SEM view of a fabricated device. Under a narrow mesa, a double slit is formed. (**b**) A schematic cross-sectional view at 1-1' in (**a**)

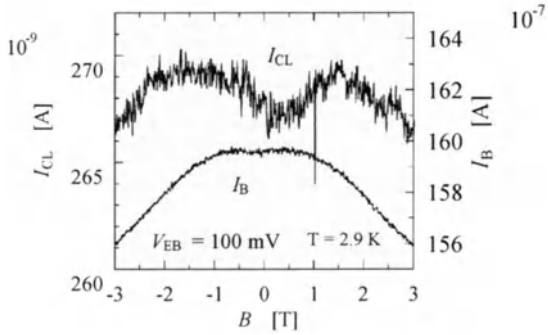


Fig. 5.4.8 Measured I_{CL} and I_B at $V_{EB}=0.1$ V at $T=2.9$ K

mesa of 1 μm width. A single barrier emitter of hot electrons is formed under the mesa. Base electrodes are besides the mesa and an emitter electrode is on the backside of the substrate. The device is made of an GaInAs/InP material system. To obtain coherent electrons, the emitter is lightly doped to have a Fermi level of 6 meV. The pitch of the double slit is 40 nm and the width of each collector is 350 nm. The energy of electrons can be varied with the voltage V_{EB} applied on the single barrier. To prevent an accumulation at the emitter-base interface, the emitter has a GaInAsP graded layer designed to be in the flat band condition at $V_{EB} = 0.1$ V. Most electrons emitted from under the mesa become base current I_B, but a part of them reach the collector passing through the slits. At the collectors, current I_{CL} and I_{CR} are detected which are proportional to the intensity of the interference pattern integrated over each electrode area. To observe the interference effect, a magnetic field B is applied perpendicularly to the plane of the electron motion. The magnetic field changes the electron trajectory and consequently the phase, which leads to a change of the interference pattern.

In actual measurements the current satisfies the relation $I_{CR} > I_B > I_{CL}$ and I_{CR} is thought to have some leak component. All measurement has been performed at a temperature below 2.9 K. When V_{EB} is 0.1 mV, I_{CL} shows

a clear minimum at $B = 0$ (Fig. 5.4.8) and its variation is larger than the noise level by an order of magnitude. I_B shows a small minimum also, but the current change is about the same as the noise level. The decrease of I_{CL} at $B = 0$ agrees reasonably with a theoretical estimation. This demonstrates double-slit interference of hot electrons.

References

1 H. Matsuura and K. Furuya, Jpn. J. Appl. Phys. **34**, 3689 (1995)
2 G.F. Giuliani and J.J. Quinn, Phys. Rev. B, **26**, 4421 (1982)
3 T. Hiramoto, K. Hirakawa, Y. Iye, and T. Ikoma, Appl. Phys. Lett. **54**, 2103 (1989)
4 A. Yacoby, U. Sivan, C.P. Umbach, and J.M. Hong, Phys. Rev. Lett. **66**, 1938 (1991)
5 G. Fasol, Extended abstracts 1992 Int. Conf. Solid State Devices and Materials, 747 (1992)
6 N.W. Ashcroft and N.D. Mermin, *Solid State Physics* (Holt, Rinehart and Winston, New York, 1976), Chapter 17
7 S. Krishnamurthy, M.A. Berding, A. Sher, and A.B. Chen, J. Appl. Phys. **63**, 4540 (1988)
8 Y.C. Kang, K. Furuya, M. Suhara, and Y. Miyamoto, Jpn. J. Appl. Phys. **33**, 6491 (1994)
9 Y.C. Kang, M. Suhara, K. Furuya, and R. Koizumi, Jpn. J. Appl. Phys. **34**, 4417 (1995)
10 M. Tsuchiya and H. Sakaki, Jpn. J. Appl. Phys. **30**, 1164 (1991)
11 M. Suhara, R. Takemura, and K. Furuya, Jpn. J. Appl. Phys. **34**, 4452 (1995)
12 R. Takemura, M. Suhara, Y. Miyamoto, K. Furuya, and Y. Nakamura, Trans. IEICE E79-C **11**, 1525 (1996)
13 R. Takemura, M. Suhara, T. Oobo, Y. Miyamoto, and K. Furuya, Jpn. J. Appl. Phys. **36**, 3B 1846 (1997)
14 M. Suhara, C. Nagao, H. Honji, Y. Miyamoto, K. Furuya, and R. Takemura, J. Cryst. Growth **179**, 18 (1997)
15 H. Hongo, Y. Miyamoto, J. Suzuki, M. Funayama, T. Morita, and K. Furuya, Jpn. J. Appl. Phys. **33**, 925 (1994)
16 H. Hongo, J. Suzuki, M. Suhara, Y. Miyamoto, and K. Furuya, Jpn. J. Appl. Phys. **34**, 4436 (1995)
17 H. Hongo, H. Tanaka, Y. Miyamoto, J. Yoshinaga, and K. Furuya, Jpn. J. Appl. Phys. **35**, L964 (1996)
18 H. Hongo, Y. Miyamoto, M. Suhara, and K. Furuya, Appl. Phys. Lett. **70**, 93 (1997)

5.5 Devices Using Ultimate Silicon Technology

5.5.1 Future of VLSI Device Technology

The size of VLSI devices is becoming smaller and smaller for higher integration and higher performance. Figure 5.5.1 shows the trend for VLSI devices, taking DRAMs for example. The density of DRAMs has increased 4 times during every three years, while the gate length of MOSFETs has decreased 0.7 times during every three years. This device trend has been valid for last 25 years. If the trend continued in the future, the device size would be 0.1 μm in 10 years and 0.03 μm in 20 years. These sizes are almost comparable with the electron wavelength in semiconductors, suggesting that quantum effects or single-electron phenomena would play an important role in electron transport even in VLSI devices. Various problems and issues that limit device scaling are being discussed, including size fluctuations, statistical fluctuations of dopants in the channel, limited lithography resolution, and difficulty of manufacturing.

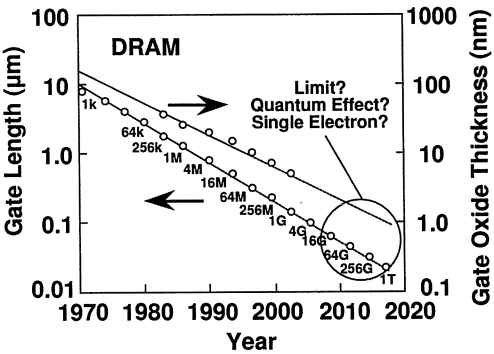


Fig. 5.5.1 MOS device trend for DRAMs. Variations of the gate length and gate oxide thickness with year are plotted. DRAM capacities are also indicated

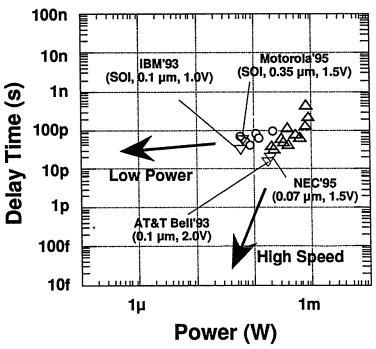


Fig. 5.5.2 The relation between the delay time and power of MOS devices. The circles show high-speed designs and the triangles show low-power designs. Recent publication data are also shown

The scaled device size achieves not only higher integration and faster circuit speed but also lower energy dissipation. Low-power VLSI devices have become very important for recent portable instruments applications. However, the design methodologies for high-speed devices and low-power devices are not exactly the same. VLSI device design has thus split into two directions: high-speed applications and low-power applications. Figure 5.5.2 shows the delay time versus power of future VLSI devices. The circles are for high-speed designs and triangles for low-power designs. As shown in the figure, the directions of the two designs are clearly splitting. Note that recent publications on silicon-on-insulator (SOI) devices point in the low-power direction.

The energy consumption E per gate or per bit storage is roughly given by

$$E = C_{\text{load}} V_{dd}^2, \tag{5.5.1}$$

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

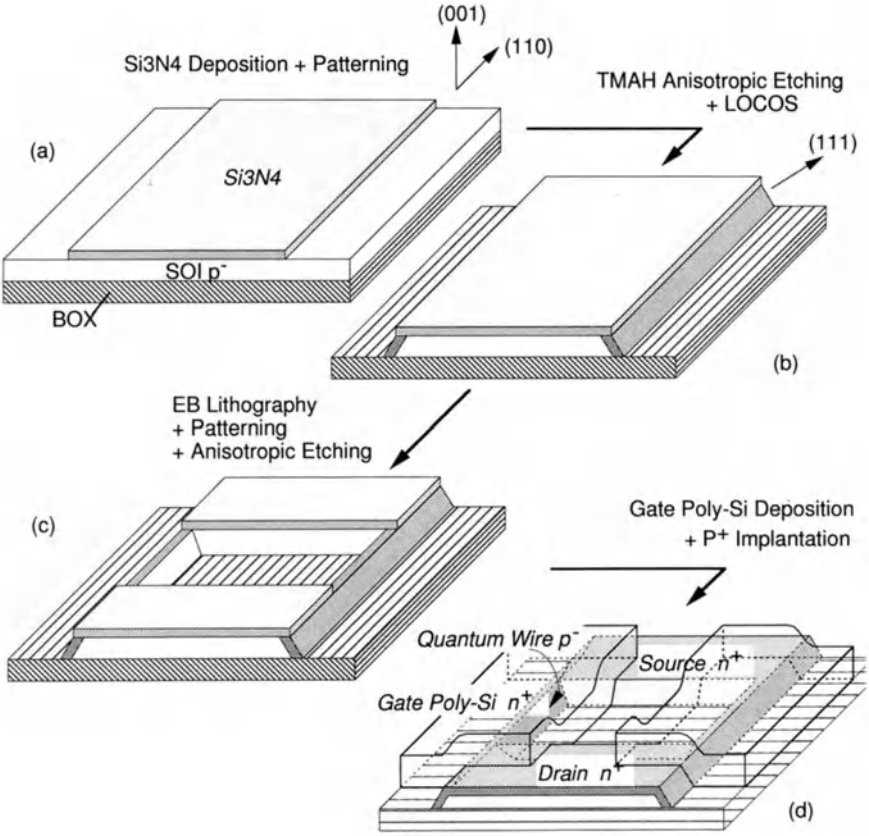


Fig. 5.5.3 Fabrication steps for an extremely narrow silicon wire MOSFET using anisotropic etching technique. The device acts as a multiple-dot single-electron device. The width of the wire is not dependent on the lithography resolution

where C_{load} is the load capacitance including the wiring capacitance and parasitic capacitance of the device, and V_{dd} the supply voltage. On the other hand, the number of electrons N that are required for one gate transition or one bit storage is given by

$$N = e^{-1}\, C_{\text{load}}\, V_{dd}, \tag{5.5.2}$$

where e is the electron charge. Therefore, the energy consumption is closely correlated with the number of electrons and a reduction in energy consumption is achieved by reducing the electron number. There are two ways to reduce the energy: decreasing V_{dd} and decreasing C_{load}. In present VLSI design, reduction of V_{dd} is commonly used to reduce power. However, V_{dd} has a lower limit. Since the subthreshold leakage current should be suppressed for portable applications, V_{dd} should not be less than 0.5–1.0 V. Then, the only way to reduce the energy used (thus the electron number) is to reduce C_{load}.

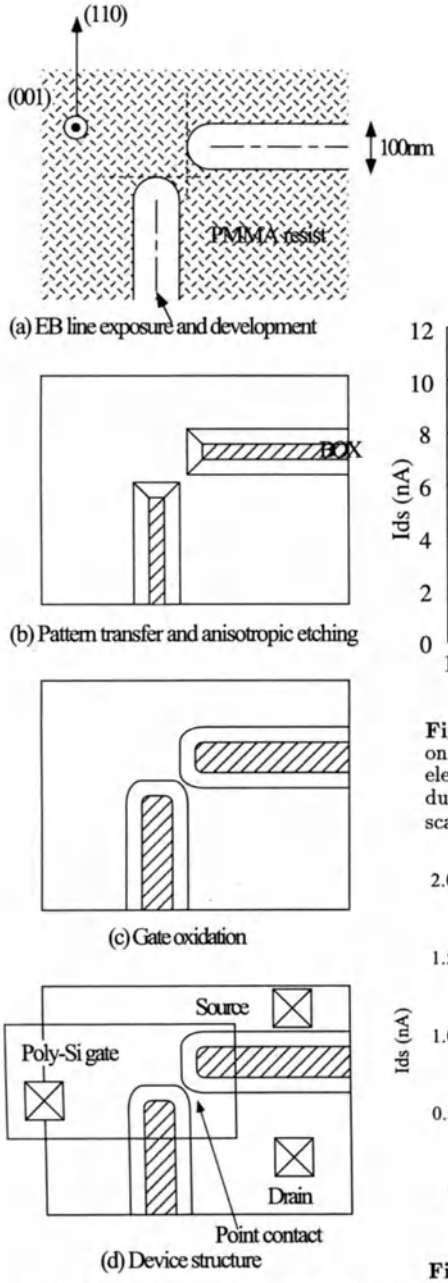


Fig. 5.5.4 Fabrication steps of a point contact MOSFET. The device act as a single-dot single-electron device

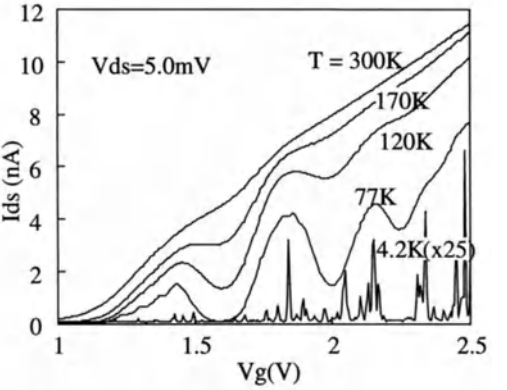


Fig. 5.5.5 Dependence of the drain current on gate voltage in the multiple-dot single-electron MOSFET. The temperature is reduced from room temperature to 77 K (linear scale) and from 77 K to 14 K (log scale)

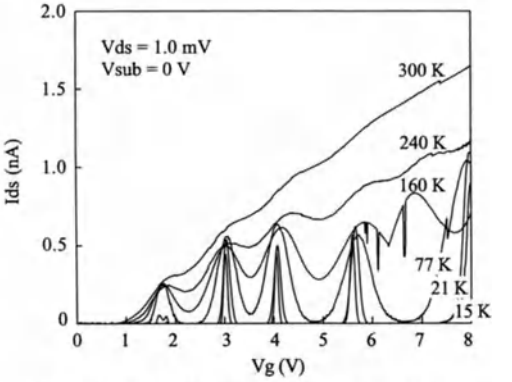


Fig. 5.5.6 I_{ds}-V_g characteristics at 4.2K in the multiple-dot single-electron MOSFET. The sharp drain current peaks are clearly observed

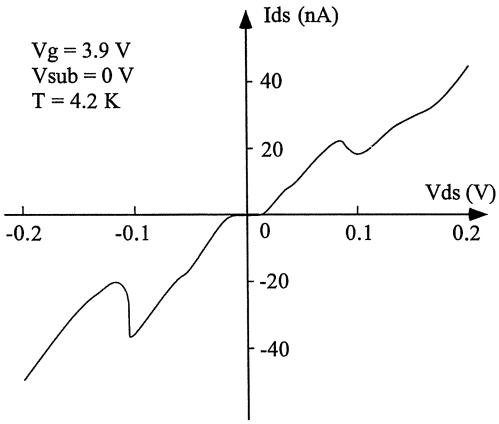


Fig. 5.5.7 The gate voltage dependence of the drain current in the single-dot single-electron MOSFET

5.5.2 Silicon Single-Electron Devices

When the capacitance is extremely reduced and the electron number in the device becomes only one, this is a single-electron device. Single-electron charging effects [1] have been of great interest for physicists. Much theoretical and experimental work has been reported. The experimental work has been done mainly in metals and compound semiconductors at very low temperatures, because it has been easy to fabricate small tunneling junctions in metals and compound semiconductors. Single-electron devices have attracted much attention also for extremely low-power device applications [2] and silicon single-electron devices have been reported [3–7]. In this section, the fabrication and characterization of two types of silicon single-electron devices are described.

Figure 5.5.3 shows schematic fabrication steps for a very narrow silicon wire MOSFET which acts as a multiple-dot single-electron device [7]. The process is based on anisotropic etching of a silicon-on-insulator (SOI) substrate. First, a SiN film is deposited on the SOI and patterned along the (110) direction. Then, the first anisotropic etch is carried out, resulting in (111) facets at the edges of the pattern. The (111) facets are coated by selective oxidation. After electron beam lithography to pattern the SiN film, the second anisotropic etching is performed. The etching automatically stops, resulting in the formation of very narrow triangular Si wires. The gate oxide is formed by thermal oxidation and the wire is further squeezed. Then, gate-poly is deposited and patterned. The channel width is less than 10 nm and the channel length is about 100 nm. The wire width is independent of the lithography and the etching time, and depends only on the SOI thickness. It is thus easy to fabricate narrower wires beyond the lithography limit.

Figure 5.5.4 shows schematic fabrication steps of a point contact MOSFET which acts as a single-dot single-electron device [8]. First, the thermal oxide is formed on an SOI substrates as a mask for the anisotropic etching. The point contact structures are defined by the EB lithography and the anisotropic

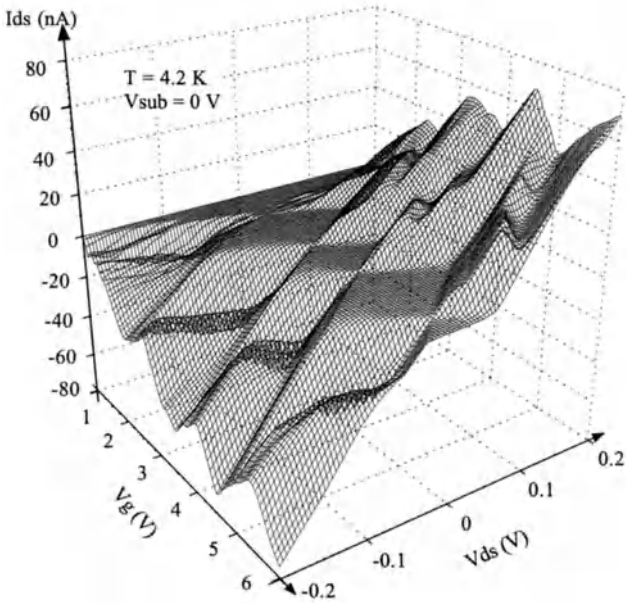


Fig. 5.5.8 Drain current vs. drain voltage in the single-dot single electron MOSFET. Clear NDCs and fine structures are observed

etching. Although the edges of the EB resist patterns are rounded, the edges become sharpened by the following anisotropic etching. Since the anisotropic etching stops automatically, the width of the point contact is precisely controlled by the EB lithography and can be narrower than the EB lithography resolution. Then, the gate oxide is formed by thermal oxidation. The edges of the point contact become rounded again and their width is further narrowed by thermal oxidation. The gate poly-Si is then deposited and patterned. The width of the point contact is less than 10 nm.

Figures 5.5.5 and 5.5.6 show the gate voltage dependence of the drain current just above the threshold voltage at different temperatures [7]. Clear oscillations are observed at room temperature which become larger oscillations as the temperature is decreased to 77 K. Note that the large oscillations split into several sharp peaks as the temperature is further decreased. Moreover, the height of the split peaks is modulated as if a beating effect took place, particularly at 4.2 K. These experimental results indicate that the observed fine peaks are not caused by a normal Coulomb blockade in a single dot. The splitting of the oscillations into sharp peaks strongly suggests that the channel should be separated into quantum dots [7,9–11]. This is reasonable because the channel is extremely narrow (less than 10 nm) and relatively long (100 nm) in this device. The charging energy in the dot is about 50 meV and larger than the thermal energy at room temperature. This is the reason why the oscillations are observed at room temperature. More detailed measurements of temperature dependence and magnetic field dependence revealed that not only resonant tunneling but also thermally activated hopping transport is

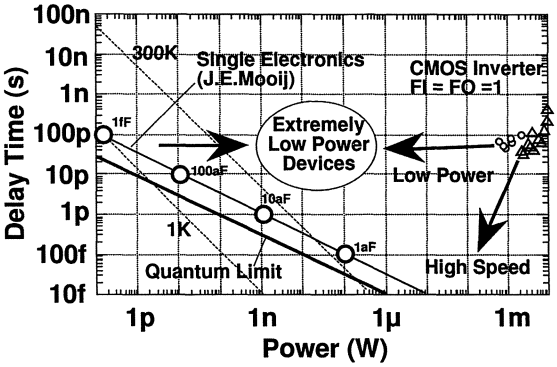


Fig. 5.5.9 Delay time vs. power. The extremely low-power devices of the future would be located just between the conventional MOS devices and the single-electron devices

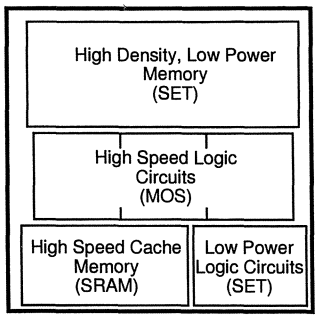


Fig. 5.5.10 A vision of a future VLSI chip. MOS and single-electron devices are integrated on one chip

dominant in multiple-dot systems at low temperatures [12].

Figure 5.5.7 shows the drain current as a function of V_{ds} in a point contact MOSFET [8,13]. Note that the intervals between the peaks are not constant. Figure 5.5.8 shows the V_{ds} dependence on I_{ds}. The negative differential conductances (NDCs) and fine structures are clearly observed as well as a suppressed I_{ds} region by Coulomb blockade at low drain voltage.

To grasp the entire property of the device, the I_{ds} dependence on the V_{ds} and V_g is plotted in Fig. 5.5.8. Clear rhombus shapes which are caused by the Coulomb blockade are observed. Although the Coulomb blockade effects would be caused by a single-dot system [8,13], the characteristics are quite different from the Coulomb blockade characteristics of the semi-classical model [1] and are attributed to quantum-mechanical effects [8,13].

In a small semiconductor dot (typically less than 10 nm), the separations between the quantized energy levels become comparable to the single-electron charging energy, which breaks the periodicity of the Coulomb blockade oscillations [1]. NDCs would be observed when the conduction band edge of the source passes the ground state level of the dot, and the fine structures should be observed when the Fermi level of the source corresponds to the excited state level of the dot [14,15].

From the experimental data, the total capacitance of the single dot is derived to be 1.4 aF, which corresponds to a charging energy of 58 meV. The separation of quantized energy levels in the dot is also estimated to be 30 meV from the position of the fine structures. The dot size would be about 6 nm. The quantized energy separations are comparable to the charging energy. Therefore, not only the single electron phenomena but also quantum mechanical effects affect the characteristics of such a small device at low temperatures. The quantum mechanical effects will become more prominent and will have a significant influence on the device characteristics as the quantum dot size in single-electron devices decreases.

5.5.3 Integration of MOS and Single-Electron Devices

Figure 5.5.9 shows the delay time versus power of devices. The single-electron devices is certainly in the direction of low power devices. It is suggested, however, that device operation by only one electron is not practical for VLSI. More than 100 or 1000 electrons are at least required in actual VLSI circuits. Therefore, we propose that the extremely low-power devices are in the regime between present VLSI devices and single-electron devices. Reduction of capacitance is essential to realize extremely low-power devices. Figure 5.5.10 shows a future vision of a VLSI chip [2]. Conventional MOS devices and single-electron devices are integrated on one chip. MOSFETs are used in the fast circuits and output interface circuits, while single-electron devices are used in circuits where a very low power consumption and high packing density are essential. In this case, the single-electron devices would operate using much more than a single electron.

References

1 *Single Charge Tunneling*, edited by H. Grabert and M.H. Devoret (Plenum, New York, 1992)
2 T. Hiramoto, H. Ishikuro, T. Fujii, T. Saraya, G. Hashiguchi, and T. Ikoma, Jpn. J. Appl. Phys. **35**, 6664 (1996)
3 K. Yano, T. Ishii, T. Hashimoto, T. Kobayashi, F. Murai, and K. Seki, IEEE Trans. Electron Devices **ED-41**, 1628 (1994)
4 Y. Takahashi, M. Nagase, H. Namatsu, K. Kurihara, K. Iwadate, Y. Nakajima, S. Horiguchi, K. Murase, and M. Tabe, Electron. Lett. **31**, 136 (1995)
5 H.I. Hanafi, S. Tiwari, and I. Khan, IEEE Trans. Electron Devices **ED-43**, 1553 (1996)
6 T. Hiramoto, H. Ishikuro, T. Fujii, T. Saraya, G. Hashiguchi, and T. Ikoma, Physica B **227**, 95 (1996)
7 H. Ishikuro, T. Fujii, T. Saraya, G. Hashiguchi, T. Hiramoto, and T. Ikoma, Appl. Phys. Lett. **68**, 3585 (1996)
8 H. Ishikuro and T. Hiramoto, 55th Device Research Conference, Colorado, USA, 84, June, 1997
9 G. Chen, G. Klimeck, S. Datta, G. Chen, and A. Goddard III, Phys. Rev. B **50**, 8035 (1994)
10 F.R. Waugh, M.J. Berry, C.H. Crouch, C. Livermore, D.J. Mar, R.M. Westervelt, K.L. Champman, and A.C. Gossard, Phys. Rev. B **53**, 1413 (1996)
11 C.A. Stafford and S.D. Sarma, Phys. Rev. Lett. **72**, 3590 (1994)
12 T. Hiramoto, H. Ishikuro, T. Fujii, G. Hashiguchi, and T. Ikoma, Jpn. J. Appl. Phys. **36**, 4139 (1997)
13 H. Ishikuro and T. Hiramoto, Appl. Phys. Lett. (submitted for publication)
14 G.W. Bryant, Phys. Rev. B **39**, 3145 (1989)
15 G.W. Bryant, Phys. Rev. B **44**, 3782 (1991)

5.6 Circuit Systems Using Quantum-Effect Devices

5.6.1 Information Processing Architectures

One of the goals in quantum electronics is to develop computing systems that can perform data processing by using quantum devices. There are two basic ways that we can use quantum-effect devices for data processing. One is to imitate existing silicon LSI circuit systems, using quantum devices as analogs of MOSFETs. The problem with this approach is that integrated circuits made of new transistor-type devices would have a hard time competing with silicon LSI, which is a well-established, mature technology. The more promising strategy is to reconsider the procedure for implementing data processing and to pursue a path different from that of ordinary LSI systems. This section reviews several approaches that are based on such a new-paradigm strategy.

The procedure for the commonly used processing method is as follows (Fig. 5.6.1): (1) devise first an algorithm for processing, (2) execute each step of the algorithm under a Neumann-type computing architecture, (3) express the algorithm as a sequence of Boolean operations on digital functions, (4) represent each digital function in terms of a Boolean expression, and (5) implement the Boolean expression with binary logic gates constructed from MOSFETs.

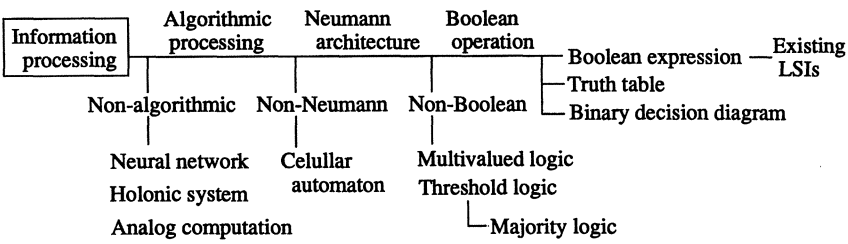


Fig. 5.6.1 Various methods for information processing

Here we consider creating computation systems based on a processing method different from the above conventional one (Fig. 5.6.1) [1]. The challenge of this interesting subject has already been taken up. This section reviews several attempts at developing such new-paradigm systems by using quantum devices.

5.6.2 Binary-Decision-Diagram Circuits

The binary decision diagram (BDD) is a method of representing digital functions by using a directed graph, instead of a Boolean expression. It was conceived and developed as a convenient tool for symbolic Boolean manipulation in computer-aided logic design. Recently an idea for relating BDD with quantum devices has been presented.

A BDD is a graph composed of many nodes and two terminals, with each node labeled by a variable. An example BDD graph, representing a four-variable logic $X_1X_2+X_3X_4$, is illustrated in Fig. 5.6.2a. In determining the

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

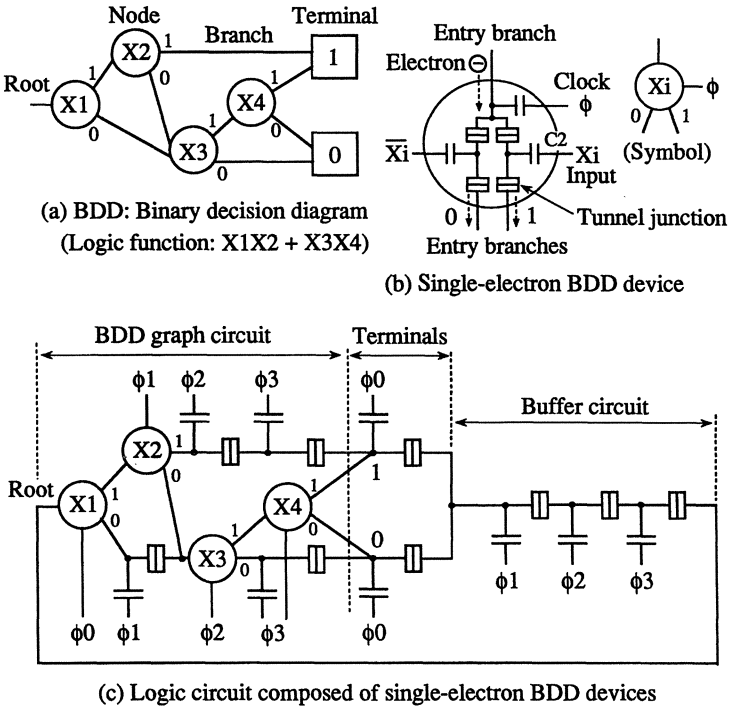


Fig. 5.6.2 Binary-decision-diagram device utilizing single-electron transport

value of the function for a given set of variables, we enter at the root and proceed downward to a terminal. At each node, we follow the branch corresponding to the value of the variable; that is, we follow the 1 branch if $X_i = \mathbf{1}$ and the 0 branch if $X_i = \mathbf{0}$. The value of the function is equal to the value of the terminal we reach; the function is **1** if we reach the 1 terminal and **0** if the 0 terminal.

BDD is composed of many identical nodes, so the node is the unit element of BDD. The function of the element is two-way switching controlled by an input variable. This function can be materialized by utilizing a physical effect that changes the coerce of a traveling medium in response to a voltage input. Figure 5.6.2b illustrates a unit device constructed from a single-electron switch circuit. Any combinational logic can be implemented by connecting identical unit devices into a cascade to build the tree of a BDD graph. Illustrated in Fig. 5.6.2c is the sample circuit that implements the BDD of Fig. 5.6.2a or the logic $X_1X_2 + X_3X_4$. It has been confirmed by computer simulation that the designed circuits will perform the logic operation correctly [2,3].

5.6.3 Local-Interaction Logic Circuits

The local-interaction logic circuit is a logic device that utilizes the energy-minimizing property of quantum structures. Its concept is as follows. To

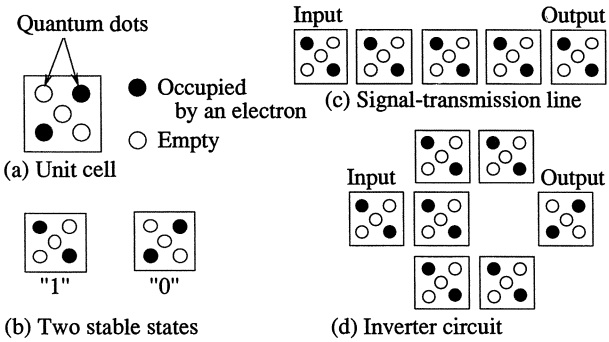


Fig. 5.6.3 QCA: unit element and circuit layout

perform binary logic computation, first a many-body electron system is prepared that consists of a large number of simple identical quantum elements (cells), where each cell has two stable states and changes its state through local interaction with its neighboring cells. In computing, some of the cells are used as input cells and others as output cells. The input binary data are provided for the system by fixing the states of the input cells. Under a given input, each cell adjusts its own state to those of all the others. After some interaction time, all the cells finally reach maximal consensus about their individual states, and the whole system then stabilizes in the minimum-energy state. By an ingenious arrangement or layout of the cells, the system can be constructed such that its minimum-energy state is achieved.

Several quantum structures have been proposed to materialize the local-interaction circuit [4-7]. Among them, the most famous is a quantum cell device called a quantum cellular automaton (QCA). It uses identical unit cells, each of which consists of five dots located at the corners and the center of a square and has two electrons that can tunnel between neighboring dots within the cell (Fig. 5.6.3a). Because of the Coulomb repulsion, two electrons tend to occupy diagonally opposed sites in the cell. As this gives the cell two states for polarization (Fig. 5.6.3b), we can encode a binary signal by manipulating the polarization. By combining identical cells in appropriate layouts, various logic circuits can be designed. Two sample circuits are illustrated in Figs. 5.6.3c and 5.6.4d [4,5].

5.6.4 Analog Computation Systems

Analog computation is a processing method that solves a mathematical problem by applying an analogous physical system to the problem. To solve the problem with this method, one prepares an appropriate physical system and represents each problem variable by a physical quantity in the system. If the mathematical relations between the physical quantities are analogous to those of the problem, then one can find the solution to the problem by observing the behavior of the system and measuring the corresponding physical quantities. A processing method based on this principle is called analog computation. A well-known (and classical) example of an analog computation system is the

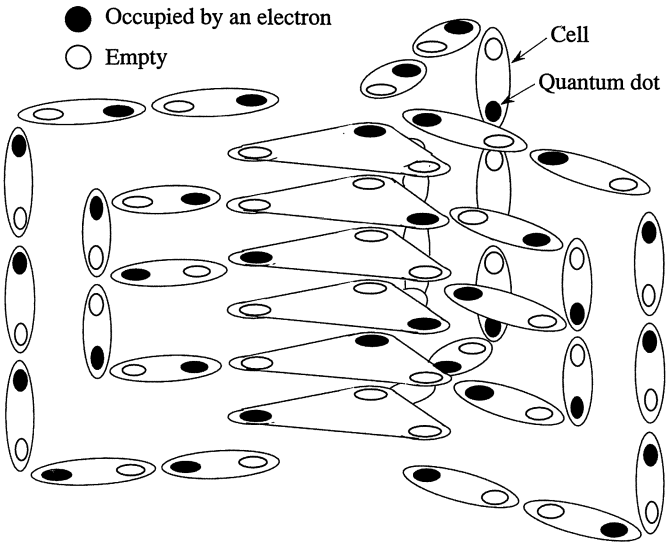


Fig. 5.6.4 Analog computation device for solving the graph-coloring problem

differential analyzer (it was employed in engineering until the 1950s).

A unique and important feature of analog computation is that it is based on real, physical matter and not on symbolic operations; therefore it is probable that quantum effects will be utilized well for implementing analog-computation systems. For example, the local-interaction logic circuit can probably be used for constructing analog computation systems for solving combinatorial problems. One such system, a QCA circuit for analyzing the graph-coloring problem, has been proposed and is illustrated in Fig. 5.6.4. For a detailed description of the circuit operation see [8].

5.6.5 MOBILE Circuit Systems

A resonant tunneling junction produces a clear negative-differential resistance at room temperature. It also provides ultrahigh-speed switching devices (a 1.5 ps switching time has been reported) [9]. The monostable-to-bistable transition logic element (MOBILE) is a logic gate device that utilizes such superior properties of resonant tunneling junctions.

MOBILE consists of a series of two resonant tunneling junctions with a structure for modulating the value of the peak current of a tunneling junction (Fig. 5.6.5a). MOBILE is driven by a clocked bias voltage (V_{bias}). It has one stable state (point S) if V_{bias} is less than twice the peak voltage (V_p) of the tunneling junction ($V_{\text{bias}} < 2V_p$, Fig. 5.6.5b) and has two stable states (points S1 and S2) if V_{bias} exceeds $2V_p$ ($V_{\text{bias}} > 2V_p$, Fig. 5.6.5c). Whether the circuit settle down to S_1 or S_2 is determined by a small difference between the peak currents of two resonant tunneling junctions. Various logic operations can be produced by utilizing this characteristic [10].

MOBILE has several advantages over conventional transistors. One is that

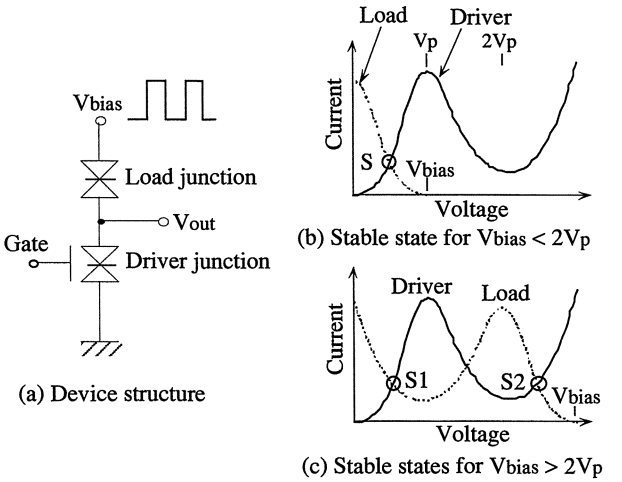


Fig. 5.6.5 MOBILE device structure and operation

it will provide multi-input and multi-output capabilities without sacrificing the high-speed properties of resonant tunneling junctions. This is because a large current-driving capability can be obtained, owing to the differential operation of the coupled resonant tunneling junctions. Another advantage is that MOBILE will provide a threshold logic device that can deal with both positive and negative weights for inputs. These advantages make MOBILE suitable for applications to parallel-processing systems such as artificial neural networks and cellular automaton systems. For example, a MOBILE cellular automaton for random-bit generation has been realized [11].

One of variant structures of MOBILE is a merged parallel connection of two resonant tunneling junctions and a high electron mobility transistor (HEMT). In this structure, the effective peak current of a tunneling junction is modulated by the gate voltage applied to HEMT. We can optimize device parameters independently for the resonant tunneling junction and for HEMT. This is advantageous for reproducibility and uniformity in device integration. For an account of device fabrication and a demonstration see [12].

Modified MOBILEs have been developed that include three or more resonant tunneling junctions in series. Various functional circuits with a simple configuration can be designed by using modified MOBILEs. The point in the operation of modified MOBILEs is that the sequence of switching in the tunneling junctions, during a critical period of increase in the clocked bias voltage, can be controlled by changing the effective peak currents through the use of the control gates. Using modified MOBILEs, various logic gates have been constructed, such as a literal gate (one of the elementary gates for multiple-valued logic), an analog-to-quaternary quantizer, and a programmable logic gate that can switch logics of NAND, NOR, AND, OR, and XOR. The operation of these functional circuits has been successfully demonstrated at room temperature [13,14].

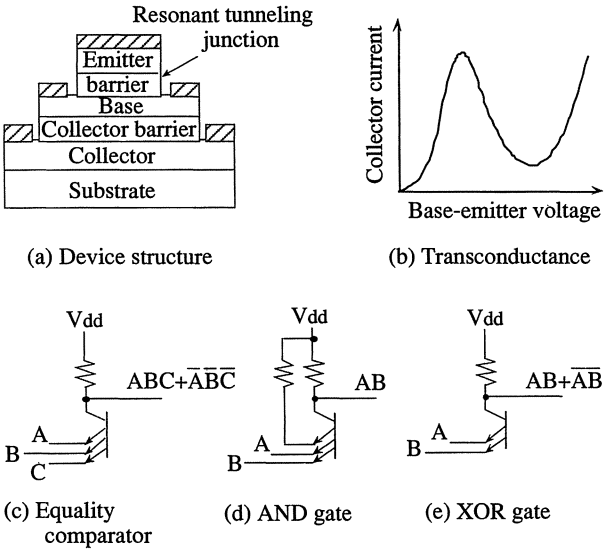


Fig. 5.6.6 RHET device and logic circuits

5.6.6 RHET Circuit Systems

The resonant tunneling hot electron transistor (RHET) is a functional device that utilizes a negative transconductance produced by a combination of hot electron transport and resonant tunneling. It is a transistor consisting of an emitter, a base, and a collector, in which the emitter-base junction is made of a resonant tunneling junction (Fig. 5.6.6a). Owing to the negative-resistance property of the resonant tunneling junction, RHET shows a negative transconductance in its base-voltage versus collector-current curve (Fig. 5.6.6b). Utilizing this characteristic, we can produce various logic functions with a simple circuit construction. For a detailed description of the device structure and operation see [15]. In the circuit applications, RHET is often designed as a multi-emitter transistor, in which two or more emitters are placed on a base layer that has no base electrode. In this structure, each emitter acts as a practical emitter if applied to a low-level voltage, and acts as a base electrode if applied to a high-level voltage. The basic logic circuit for a multi-emitter RHET is illustrated in Fig. 5.6.6c. It operates as an equality comparator that produces an output **1** if every input is **1** or every input is **0**. Several elementary logic gates using this function are illustrated in Fig. 5.6.6d and Fig. 5.6.6e. Various other logic circuits can be constructed in a simple manner. A static random-access memory circuit has also been reported [16,17].

References

1 Y. Amemiya, J. Intelligent Material Systems and Structures **5**, 418 (1994)
2 N. Asahi, M. Akazawa, and Y. Amemiya, IEEE Trans. Electron Devices **42**, 1999 (1995)
3 N. Asahi, M. Akazawa, and Y. Amemiya, IEEE Trans. Electron Devices (in press)

4 C.S. Lent and P.D. Tougaw, J. Appl. Phys. **74**, 6227 (1993)
5 C.S. Lent, P.D. Tougaw, and W. Porod, Appl. Phys. Lett. **62**, 714 (1993)
6 K. Obermayer, W.G. Teich, and G. Mahler, Phys. Rev. B **37**, 8096 (1988)
7 S. Bandyopadhyay and V. Roychowdhury, Jpn. J. Appl. Phys. **35**, 3350 (1995)
8 Y. Amemiya, IEICE Trans. Electronics **E79-C**, 1481 (1996)
9 N. Shimizu, T. Nagatsuma, T. Waho, M. Shinagawa, M. Yaita, and M. Yamamoto, IEEE Electronics Lett. **31**, 1695 (1995)
10 K. Maezawa, T. Akeyoshi, and T. Mizutani, IEEE Trans. Electron Devices, **41**, 148 (1994)
11 K. Maezawa, T. Akeyoshi, and T. Mizutani, IEEE Trans. Electron Devices, **42**, 1005 (1995)
12 K.J. Chen, K. Maezawa, and M. Yamamoto, IEEE Electron Device Lett. **17**, 127 (1996)
13 T. Waho, K.J. Chen, and M. Yamamoto, IEEE Electron Device Lett. **17**, 223 (1996)
14 K.J. Chen, T. Waho, and M. Yamamoto, IEEE Electron Device Lett. **17**, 309 (1996)
15 N. Yokoyama et al., Jpn. J. Appl. Phys. **24**, L853 (1985)
16 M. Takatsu et al., J. Solid-State Circuits **27**, 1247 (1992)
17 M. Takatsu et al., ISSCC Digest of Technical papers **TA7.2**, 124 (1994)

6. Formation and Characterization of Quantum Structures

6.1 Introduction

Semiconductor quantum structures, such as quantum wires (QWRs) and dots (QDs), are attracting much interest in the field of fundamental physics and novel device application, as described in the previous chapters. The realization of such properties requires extreme control of fabrication tolerance, and nonuniformity in structural dimensions often fatally degrades the advantageous properties. For instance, inhomogeneous broadening of the gain spectrum of QD lasers will result from fabricational variations in the dot size and shape. Quantum dot lasers do not offer significant advantages over conventional bulk lasers unless fabricational tolerances are tightly controlled [1].

Many attempts have been made to fabricate uniform quantum structures. Most of them require complicated processing of the structures, either patterning of the substrate prior to the growth, or various types of etching or implantation of quantum wells, including overgrowth and patterning of small stressor structures [2,3]. These quantum structures have been able to provide only limited evidence of low-dimensional quantum confinement because of the limit imposed by obtainable size and nonuniformity. Sometimes, these etching, patterning, and implantation processes induce defects, confusing the results obtained with such structures.

Recently, self-organized or self-assembling techniques have been proposed. Metalorganic chemical vapor deposition (MOCVD) or organometallic vapor phase epitaxy (OMVPE) and molecular beam epitaxy (MBE) were used, resulting in the formation of tilted superlattices, V-groove QWRs, Stranski–Krastanov QDs, etc. [4–6]. There remain serious problems in controlling the size, shape, and location of QWRs and QDs, however. In this chapter, we describe a variety of formation techniques for quantum structures, including those using these two epitaxial methods.

References

1 K.J. Vahala, IEEE J. Quantum Electro. **24**, 523 (1988)
2 H. Sasaki, Surf. Sci. **267**, 623 (1992)
3 B. Mener, D. Heitmann, and K. Ploog, Phys. Rev. Lett. **68**, 1371 (1992)
4 J.M. Gaines, P.M. Petroff, H. Kroemer, R.J. Simes, R.S. Geels and J.H. English, J. Vac. Sci. Technol. B **6**, 1378 (1988)
5 E. Kapon, D.M. Hwang, and R. Bhat. Phys. Rev. Lett. **63**, 430 (1989)
6 P. Leonard, M. Krishnamurthy, C.M. Reaves, S.P. Den Baas and P.M. Petroff, Appl. Phys. Lett. **63**, 3203 (1993)

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

6.2 Quantum Wires and Dots by MOCVD (I)

6.2.1 Quantum Wires on Vicinal Surfaces

The use of monoatomic and multi-atomic steps formed on vicinal (001) GaAs substrates has been proposed and demonstrated to realize two-dimensional quantum confinement structures such as quantum wires (QWRs) grown by MOCVD [1–4] which can be used as a new type of electron wave interference device [5]. Similar QWR structures have been observed in tilted GaAs/AlAs superlattices on vicinal (001) GaAs substrates grown by MBE [6], in a thin InGaAs layer at the edge of InP multiatomic steps on vicinal (001) InP surfaces grown by metalorganic MBE [7], and in a thin AlGaAs composition modulation layer at the edge of AlAs multiatomic steps on vicinal (110) GaAs substrates grown by gas-source MBE [8]. In this section we review some of these structures.

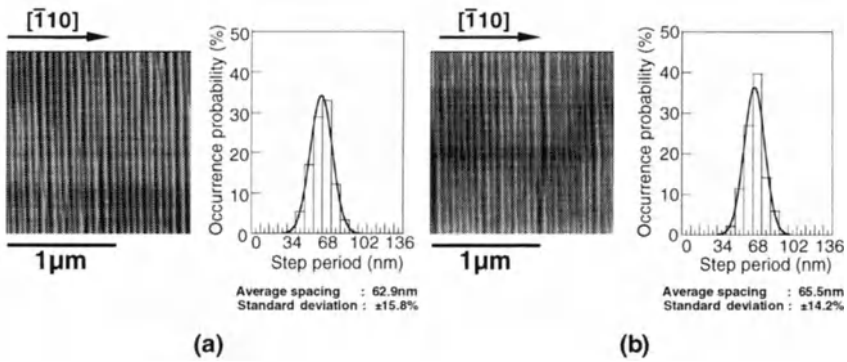


Fig. 6.2.1 AFM images of GaAs and AlAs grown surfaces on vicinal substrates misoriented by 5.0° toward the [-110] direction. The observed area is $1.7\times1.7\ \mu m^2$. (**a**) GaAs buffer layer surface and (**b**) AlAs lower barrier layer surface grown on coherent GaAs multiatomic steps at 600°C

Figure 6.2.1a shows a typical atomic-force-microscopy (AFM) image of GaAs grown surface after a 30 min thermal treatment at 600°C under AsH_3/H_2 atmosphere and the distribution of step spacing [9]. Coherent multiatomic steps with extremely straight edges are observed over a several-micron area. Average height and spacing of multiatomic steps are 5.5 nm and 63 nm, respectively, and the fluctuation of step spacing is ±16%. Next, an AlAs surface grown on GaAs multiatomic steps is observed by AFM. Figure 6.2.1b shows an AFM image of AlAs surface on 5.0°-misoriented substrate grown at 650°C and the distribution of step spacing within the observed area. In this sample, 3 nm-thick GaAs cap layers are grown on the AlAs layer for AFM observations in air. Average height and spacing of the multiatomic steps are 5.7 nm and 66 nm, respectively, and the uniformity of step spacing is ±14%. These results suggest that the step spacing of multiatomic steps on the AlAs surface is almost the same as that on the underlying GaAs surface.

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

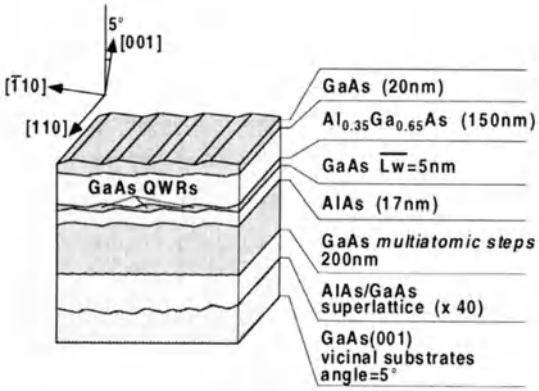


Fig. 6.2.2 Schematic illustration of QWR structures on GaAs vicinal substrate

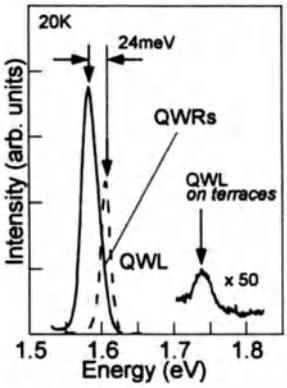


Fig. 6.2.3 PL spectrum of QWRs on 5.0°-misoriented substrate and that of QW on singular (001) substrate

Figure 6.2.2 shows a typical schematic illustration of the GaAs QWR structure. Photoluminescence (PL) spectra at 20 K are shown in Fig. 6.2.3. The average quantum well widths (L_w) are 5.5 nm (1.617 eV). PL peak position of QWRs on a 5.0°-misoriented substrate is 1.593 eV. It is found that the PL peak energy of QWRs is smaller than that of a quantum well (QW) formed on the singular (001) GaAs substrate. In MOCVD growth, the total amount of the grown material is basically the same for these structures, being independent of the substrate misorientation angle. Moreover, we have verified that the growth thickness is almost the same from wafer to wafer within the carbon suscepter in our MOCVD system, and the growth thickness difference from run to run is within 1%. Therefore, although peak energy should shift toward the higher energy region due to 2D quantum confinement, these energy shifts indicate that locally thick areas are formed in a GaAs well layer on a 5.0°-misoriented substrate, that is, quantum-wire-like structures are formed at the edges of multiatomic steps as schematically shown in Fig. 6.2.2.

In these PL spectra at 20 K, the full width at half-maximum (FWHM) of QWRs is 25 meV. The result shows that the size uniformity of QWRs with an AlAs lower barrier layer has been much improved over the previous report [2]. From AFM observations, this size uniformity of QWRs is probably due to the spacing uniformity of multiatomic steps on AlAs lower barrier layer surfaces. Moreover, in PL spectra of the sample with AlAs as lower barrier layer, we observe an additional weak spectrum at about 1.748 eV. This peak probably corresponds to QWs formed on (001) terraces connected with QWRs. The result indicates that most of the photo-excited electrons and holes diffuse to lower energy regions, that is, QWR regions. We also measure the polarization dependence of the PL intensity (I) from QWRs using AlAs as a lower barrier layer at 20 K. Since the misorientation direction is [−110] and the direction along QWRs is [110], we defined the degree of polarization as $(I_{[110]} - I_{[-110]})/(I_{[110]} + I_{[-110]})$. The degree of polarization for QWRs is found to be 0.12. For a QW on the singular (001) GaAs substrate, this value

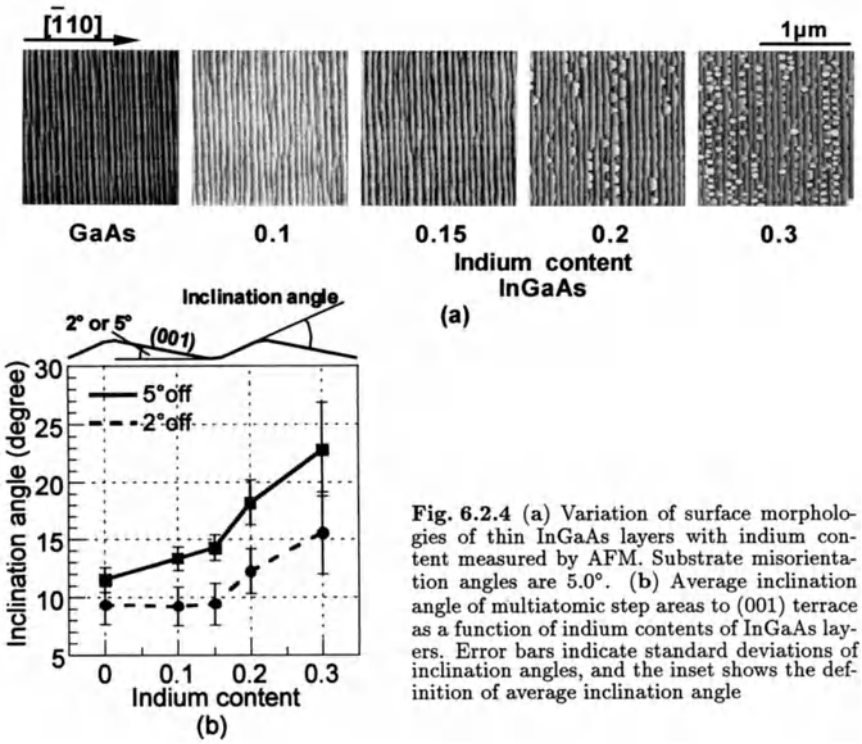


Fig. 6.2.4 (a) Variation of surface morphologies of thin InGaAs layers with indium content measured by AFM. Substrate misorientation angles are 5.0°. (b) Average inclination angle of multiatomic step areas to (001) terrace as a function of indium contents of InGaAs layers. Error bars indicate standard deviations of inclination angles, and the inset shows the definition of average inclination angle

is 0.04, and no large effect due to internal stress is expected in this material system. Therefore, the observed polarization anisotropy also supports the successful formation of QWRs at the edges of multiatomic steps. Furthermore, we observe fine-area PL spectra for QWRs using Au/Ge patterned mask. fine-area photoluminescence spectrumThe open area is 2 μm wide and 800μm long; thus we can obtain PL spectra from about thirty QWRs. It is found that the FWHM of PL spectra from about thirty QWRs is 17 meV, which is 8 meV smaller than that from the whole area. This result suggests that the size fluctuation of QWRs still remains for a wide area.

Next, InGaAs layers have been grown directly on coherent GaAs multiatomic steps. The variation of surface morphologies of thin InGaAs layers with the indium content is shown in Fig. 6.2.4a. The substrate misorientation angle is 5.0°. When the indium content is smaller than 0.15, InGaAs layers are grown in a step-flow growth mode. For $In_{0.2}Ga_{0.8}As$ growth, a prominent undulation of the multiatomic step front is partly observed due to the increase of the strain effect. For $In_{0.3}Ga_{0.7}As$ growth, many 3D InGaAs islands are formed at the edge of GaAs multiatomic steps in order to reduce the stress due to the difference of lattice constants between InGaAs and GaAs. Although no 3D island structures are formed on singular (001) GaAs substrates, a larger average density of these 3D islands is observed with increasing substrate misorientation angle. This is probably due to the in-

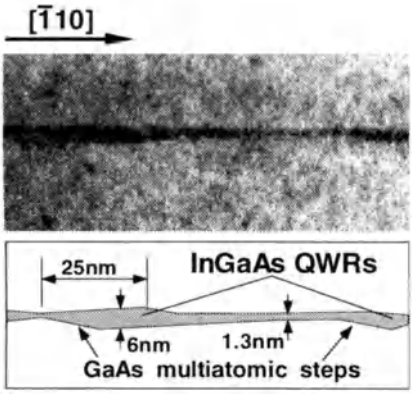


Fig. 6.2.5 Cross-sectional TEM image and schematic illustration of $In_{0.1}Ga_{0.9}As$ QWRs on 5.0°-misoriented substrate

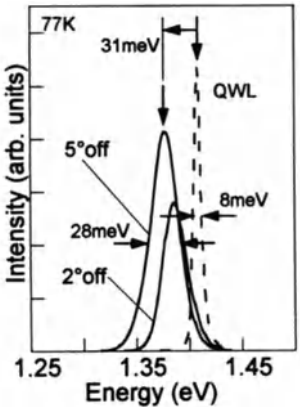


Fig. 6.2.6 PL spectra of QWRs on 5.0°- and 2.0°-misoriented substrates and a single QW on singular (001) substrate at 77 K. The In content of InGaAs layers is 0.15 and the average thickness of InGaAs QW layers is 3 nm

crease in the stress at the edges of multiatomic steps and the step density on vicinal substrates. For $In_{0.3}Ga_{0.7}As$ growth, we also find slight ordering of the island sizes and positions along the multiatomic steps. To control the size uniformity and the position of these 3D InGaAs islands, further optimization of the growth conditions is needed.

Figure 6.2.4b shows the average inclination angles measured from line scan profiles of AFM images. The growth temperature of all samples is 600°C and the thickness of thin InGaAs layers on GaAs multiatomic steps is about 3 nm. The inset shows the definition of the average inclination angle and the error bars indicate standard deviations of the inclination angles. These data exclude the 3D island-like area for $In_{0.2}Ga_{0.8}As$ and $In_{0.3}Ga_{0.7}As$. With the increase of the indium content of the InGaAs layer, the average inclination angle of the multiatomic step areas to the terrace increases for both 2.0°- and 5.0°-misoriented substrates. These results suggest that InGaAs QWR structures are successfully formed at the edge of GaAs multiatomic steps when the indium content is smaller than 0.15 and that, for an indium content larger than 0.2, 3D island structures are connected with wire-like structures.

Figure 6.2.5 shows a cross-sectional TEM image and schematic illustration of InGaAs QWRs on a 5.0°-misoriented substrate for an indium content of 0.1. The InGaAs layer (dark area) is thicker at the edge of GaAs (bright areas) multiatomic steps than on the (001) terrace region. However, the corrugation of multiatomic steps on the InGaAs top surface seems to decrease compared to the AFM image, which is probably due to a slight segregation of indium atoms during the upper GaAs barrier layer growth. We estimate from this TEM image that the lateral width and the vertical thickness of InGaAs QWRs are about 25 and 6 nm, respectively, and that the quantum well thickness on

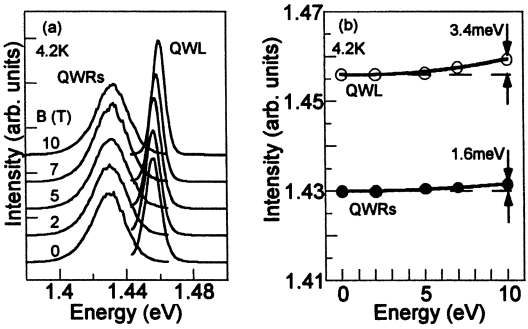


Fig. 6.2.7 (**a**) PL spectra of QWRs and QW at 4.2 K in magnetic fields up to 10 T. The In content and the average thickness of InGaAs layers are 0.15 and 3 nm, respectively. (**b**) Peak shifts as a function of magnitude of magnetic fields

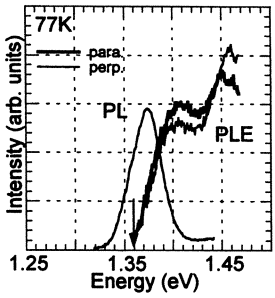


Fig. 6.2.8 Polarization anisotropy in PLE spectra of $In_{0.15}Ga_{0.85}As$ QWRs on 5.0°-misoriented substrates at 77 K. An arrow shows the detection energy

the (001) terrace is 1.3 nm. On the other hand, for the single quantum well on a singular (001) GaAs substrate, very abrupt InGaAs/GaAs heterointerfaces are obtained.

PL spectra of $In_{0.15}Ga_{0.85}As$ QWRs on vicinal substrates and reference QW on a singular (001) substrate at 77 K are shown in Fig. 6.2.6. The well layer width of the $In_{0.15}Ga_{0.85}As$/GaAs strained quantum well is about 3 nm. PL spectra of QWRs shift to a lower energy by 23 meV for the 2.0°-misoriented sample and 31 meV for the 5.0°-misoriented sample with respect to that of QW. FWHM of the PL spectra are about 27, 28, and 8 meV for 2.0°, 5.0°-misoriented, and singular (001) samples, respectively. For the $In_{0.1}Ga_{0.9}As$ QWRs sample, these red shifts and FWHM of the PL spectra are 15 and 29 meV for the 2.0°-misoriented sample, and 25 and 28 meV for the 5.0°-misoriented sample, respectively.

In a previous report, GaAs QWRs formed at the edge of the multiatomic steps on the 5.0°-misoriented substrate showed a lower energy shift (red shift) by 23 meV with respect to that of the QW, which agree well with the results from a numerical calculation using a finite difference method in the 2D Schrödinger equation for the GaAs QWR size obtained from cross-sectional TEM images [2]. Furthermore, the average indium content of a MOCVD-grown InGaAs layer on a vicinal GaAs surface was smaller than that on a singular (001) surface by energy-dispersive x-ray spectroscopy and PL spectra of InGaAs/GaAs QWs on misoriented surfaces shifted to a higher energy with respect to that on a singular (001) surface (blue shift) [10]. Therefore, the red shifts for the misoriented samples with respect to singular (001) samples are probably caused by a formation of a locally thick InGaAs QWR at the edge of GaAs multiatomic steps.

Figure 6.2.7a shows PL spectra of QWRs in magnetic fields B and shows peak shifts as a function of the magnitude of magnetic field in Fig. 6.2.7b. The magnetic field is in the growth direction of the samples. When $B = 10$ T, the PL peak position for QW on singular (001) GaAs substrates shifts

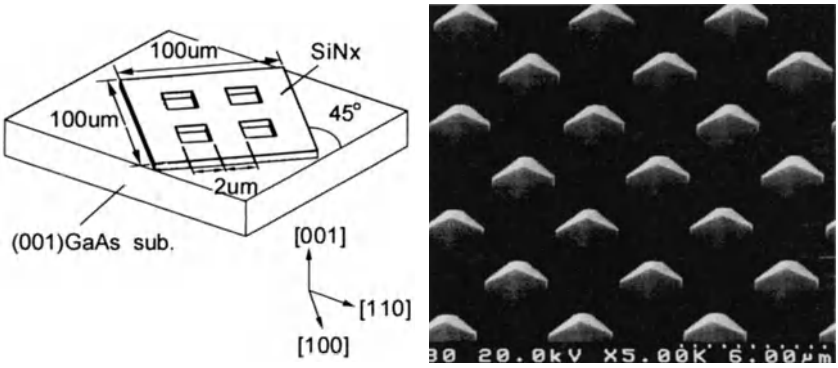


Fig. 6.2.9 Schematic illustration of the SiN_x mask pattern for selective area MOCVD and an SEM image of GaAs pyramidal structures

toward a higher energy by about 3.4 meV. On the other hand, the amount of such energy shifts for QWRs is only about 1.6 meV. This small energy shift suggests the effect of lateral confinement in QWRs and successful fabrication of QWRs at the edges of multiatomic steps, although we cannot evaluate the lateral confinement in our QWRs quantitatively. Figure 6.2.8 shows photoluminescence excitation (PLE) spectra of an InGaAs strained QWR sample with 0.15 indium content, performed with a CW titanium sapphire laser at 77 K, indicating two absorption peaks which probably correspond to heavy hole-electron and light hole-electron excitonic states in the first subband in QWRs. The detected energy is 1.359 eV, corresponding to the lower energy side of the PL spectrum from QWRs. PLE spectra show the polarization of excitation light parallel and perpendicular to QWRs.

6.2.2 Quantum Dot Formation on Masked Substrates

Selective area (SA) MOCVD growth seems to be one of the most promising approaches for formation of low-dimensional structures using crystal growth only. It has potentially the advantages of (1) formation of defect-free quantum structures without process-induced damage and contamination, (2) realization of quantum confinement sizes unobtainable by conventional process techniques by self-organized selective crystal growth, (3) realization of atomically smooth interfaces due to the appearance of crystalline facets, (4) controllability of the position of quantum structures by patterning with standard lithography and etching, and (5) realization of excellent size uniformity, not affected by lithographic size fluctuation, due to self-limited growth.

Various kinds of structures can be formed by adjusting the growth conditions as well as the substrate orientation and the mask pattern. For example, lateral QWRs and tetrahedral quantum dots (QDs) on a (111)B GaAs substrate [11] and pyramidal QDs on (001) GaAs substrates [12,13] have been reported. The structures are formed by utilizing differences of the surface energy between low index surfaces. Determination of critical width on the ridge structures by selective area MOCVD using diethyl gallium chloride [14]

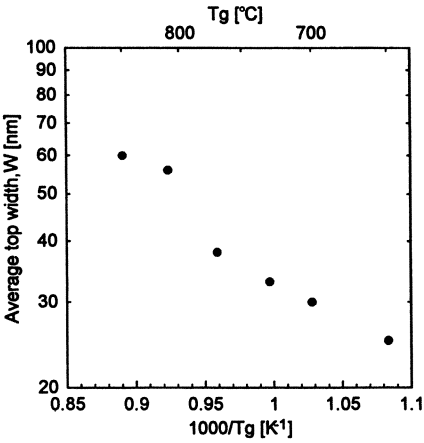


Fig. 6.2.10 Dependence of growth temperature dependence on the average size of the top portion of the GaAs pyramids

and MBE [15], and of the top size of truncated tetrahedrons on (111)B GaAs substrate [16] and pyramids on (001) GaAs substrate [13] are by changing the growth conditions, mainly the growth temperature. Another typical QD formation method is Stranski–Krastanow growth mode using self-assembling mechanism in an InGaAs/GaAs strained system [17,18], which still has difficulties in controlling the position and size.

Figure 6.2.9 shows a schematic illustration of the SiN_x mask pattern for selective area MOCVD. Using this mask pattern, GaAs is grown on this masked substrate at a high temperature of 730°C and low arsine partial pressure of 6.7×10^{-5} atm. It is clear that pyramidal structures surrounded by four {011} facet sidewalls are formed as shown in Fig. 6.2.9. The structures are of a truncated pyramidal shape and the top size of the structures becomes narrower as the growth proceeds.

The growth process of the pyramidal structures is as follows: In the early stage of the growth, a part of the reactants diffuse laterally in the vapor phase outside the 100×100 μm^2 island pattern, because the outside of the island becomes a large sink for the reactants. Then the amount of diffused reactants at the edge of the pattern is larger than at the center. Therefore, nonuniform growth occurs within the island. However, as the growth proceeds, the pyramidal formation is completed from the center to the edge of the island pattern. Once the pyramidal formation is completed, no growth occurs on the top region of the pyramids or {011} facet sidewalls, because the growth rate is much slower than that on any other surfaces. In addition, the desorption rate of Ga atoms at the top of the pyramids increases rapidly because of the weak bonding energy, so that the growth saturates and most Ga atoms start to diffuse out of island-like patterns. Since the shape and the top size of the pyramids are thought to be determined in thermal equilibrium, the widths of the tops of the pyramids become very uniform due to the self-limited growth

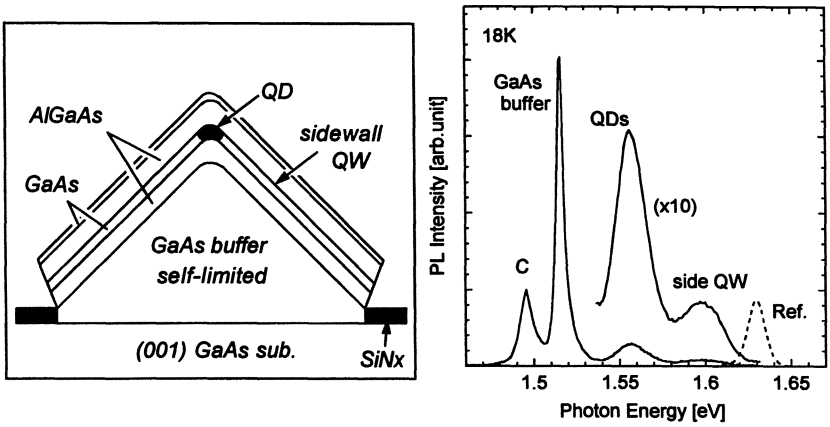


Fig. 6.2.11 Schematic view of a cleaved cross section of the quantum dot structure and PL spectrum of the quantum dot structures (solid curve). The broken curve represents a reference quantum well on the planar substrate

mechanism.

Figure 6.2.10 shows the temperature dependence of the average size of the top portion of the completed GaAs pyramids measured by AFM. The top size decreases as the growth temperature decreases. Therefore, it is possible to form uniform and size-controlled QDs by changing the growth conditions.

A self-limited mechanism is then utilized to fabricate a uniform QD array. First, GaAs is grown by a self-limited growth, at a high temperature of 730°C and low AsH_3 pressure of 6.7×10^{-5} atm. Under these growth conditions, the top of the pyramid is saturated at about 30 nm. Next, an AlGaAs/GaAs single quantum well is grown on these pyramids at a low temperature of 650°C and high AsH_3 pressure of 5.0×10^{-4} atm. AlGaAs and GaAs layers are grown both at the top of the pyramids and at the {011} facet sidewalls under the overgrowth condition. In this case, the growth rate of the top portion of the pyramid measured by SEM is 1.8 times higher than that of {011} facet sidewalls, that is, the quantum dots formed at the top portion of the pyramids are connected to quantum wells on the {110} sidewalls. A schematic view of a cleaved cross section of the quantum dot is shown in Fig. 6.2.11. The lateral size of QD is estimated to be 30 nm. Figure 6.2.11 shows PL spectra of the QD structures at 18 K. The reference QW width on the planar substrate is 4.5 nm. The peaks of 1.546 eV and 1.596 eV correspond to the emissions from QDs formed on the top of the pyramids and sidewall quantum wells and the calculated well widths from PL peak energies are 12 nm and 6.0 nm, respectively. From the SEM observation, the layer thickness is 10.4 nm for top QDs and 5.8 nm for sidewall QDs, and these values agree with those estimated from PL peaks. The sharp emission from QDs (FWHM 22 meV) suggests that a uniform QD array is formed using self-limited growth mode. Figure 6.2.12 shows an SEM image of a high-density QD array. The pattern is formed by electron beam lithography and the density of QDs is 2.0×10^9

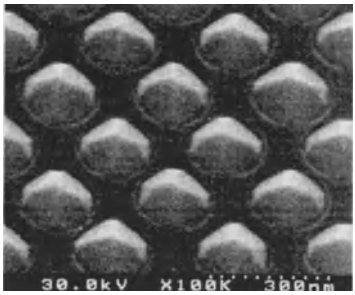


Fig. 6.2.12 SEM image of high-density quantum dot array. The density of the quantum dots is 2.0×10^9 cm^{-2}

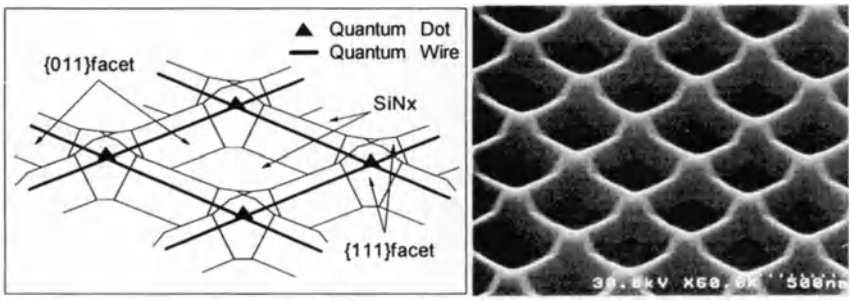


Fig. 6.2.13 Schematic illustration of idealized coupled quantum wire-dot structures and SEM image of the AlGaAs structure consisting of pyramidal dots connected to a wire grown on the masked substrate

cm^{-2}.

SA-MOCVD is not only useful for fabrication techniques for QDs, but also promising for the fabrication of optical integrated circuits [19]. Using this selective growth, the formation of future optical devices and optoelectronic integrated circuits (OEIC) will be possible. On the other hand, the application of QDs to electron transport devices, especially single-electron devices, is also a very important research field for future electron devices. For fabrication of these devices, these well-arranged quantum dots have to be connected to each other. Figure 6.2.13 shows idealized coupled quantum wire-dot structures, and an SEM image of an AlGaAs structure consisting of pyramidal dots connected to wire grown on the masked substrate. Using these coupled wire and dot structures, novel single-electron devices and electron wave interference devices as well as high-density matrices of QD memory or QD networks can be realized by further progress on the SA-MOCVD technique.

References

1 T. Fukui and H. Saito, Appl. Phys. Lett. **50**, 824 (1987)
2 S. Hara, J. Ishizaki, J. Motohisa, T. Fukui, and H. Hasegawa, J. Cryst. Growth **145**, 692 (1994)

3 S. Hara, J. Motohisa, T. Fukui, and H. Hasegawa, Jpn. J. Appl. Phys. **34**, 4401 (1995)
4 S. Hara, J. Motohisa, and T. Fukui, J. Cryst. Growth **170**, 579 (1997)
5 J. Motohisa, M. Akabori, S. Hara, J. Ishizaki, K. Ohkuri, and T. Fukui, Physica B **227**, 295 (1996)
6 P.M. Petroff, A.C. Gossard, and W. Wiegmann, Appl. Phys. Lett. **45**, 620 (1984)
7 M.J.S.P. Brasil, A.A. Bernussi, M.A. Cotta, M.V. Marquezini, J.A. Brum, R.A. Hamm, S.N.G. Chu, L.R. Harriott, and H. Temkin, Appl. Phys. Lett. **65**, 857 (1994)
8 K. Inoue, K. Kimura, K. Maehashi, S. Hasegawa, H. Nakashima, M. Iwane, O. Matsuda, and K. Murase, J. Cryst. Growth **127**, 1041 (1993)
9 J. Ishizaki, K. Ohkuri, and T. Fukui, Jpn. J. Appl. Phys. **35**, 1280 (1996)
10 K. Hiramoto, T. Tsuchiya, M. Sagawa, and K. Uomi, J. Cryst. Growth **145**, 133 (1994)
11 T. Fukui, S. Ando, T. Honda, and T. Toriyama, Surf. Sci. **267**, 236 (1992)
12 Y. Nagamune, S. Tsukamoto, M. Nishioka, and Y. Arakawa, J. Cryst. Growth **126**, 707 (1993)
13 K. Kumakura, K. Nakakoshi, J. Motohisa, T. Fukui, and H. Hasegawa, Jpn. J. Appl. Phys. **34**, 4387 (1995)
14 K. Yamaguchi and K. Okamoto, Appl. Phys. Lett. **59**, 3580 (1991)
15 S. Koshiba, H. Noge, H. Akiyama, T. Inoshita, Y. Nakamura, A. Shimizu, Y. Nagamune, M. Tsuchiya, H. Kano, H. Sakaki, and K. Wada, Appl. Phys. Lett. **64**, 363 (1994)
16 S. Ando, T. Honda, and N. Kobayashi, Jpn. J. Appl. Phys. **32**, L104 (1994)
17 J. Oshinowo, M. Nishioka, S. Ishida, and Y. Arakawa, Appl. Phys. Lett. **65**, 1421 (1994)
18 P.M. Petroff and S. P. DenBaars, Superlattice & Microstruct. **15**, 15 (1994)
19 M. Aoki, H. Sano, M. Suzuki, M. Takahashi, K. Uomi, and A. Takai, Electrn. Lett. **27**, 2138 (1990)

6.3 Quantum Wires and Dots by MOCVD (II)

6.3.1 Quantum Wires by Selective MOCVD

To fabricate quantum wires, various techniques such as wet chemical etching [1], reactive ion etching [2], ion beam implantation [3], and ion beam milling [4,5] have been investigated. These methods suffer from free surface effects, creation of a damage field during implantation, or interface problems due to various disordering mechanisms. To avoid these problems, growth techniques on masked substrates [6,7] and nonplanar substrates [8–10] have been investigated. Pioneering work by Kapon et al. successfully fabricated quantum wires on V-grooved (100) oriented GaAs substrates [8] or submicron gratings [9] by MOCVD. Vertically-stacked quantum wires on a single V-groove were also achieved [10]. In these works wet chemical etching was used.

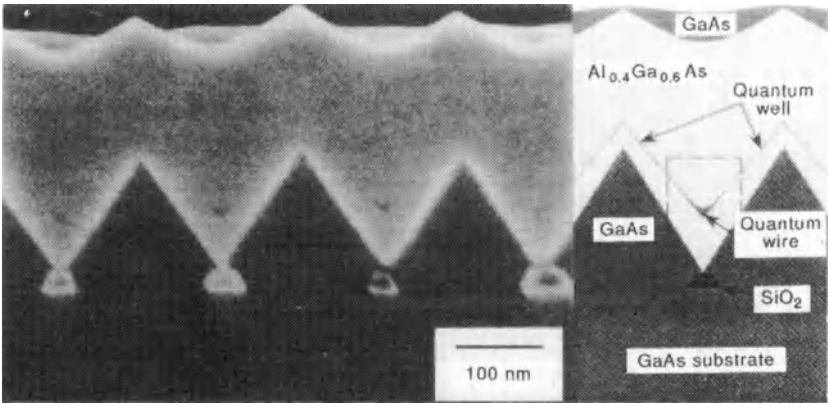


Fig. 6.3.1 A cross-sectional SEM image of quantum wires with 200 nm period

V-grooves can be also formed by utilizing MOCVD selective growth on a SiO_2 patterned substrate [11,12]. A fabrication procedure for quantum wires is as follows [11]. First, a SiO_2 layer with a thickness of 20 nm was formed by plasma chemical vapor deposition on a semi-insulating (100) GaAs substrate. PMMA was then patterned on the SiO_2 layer by electron beam lithography technique and wet chemical etching was used to pattern the SiO_2 film. Then, triangular shaped GaAs with (111)A facet sidewalls was grown on the masked substrate by MOCVD growth. The formation of the triangular structure is due to a large growth rate difference between (100) and (111) orientations. Further continuation of the growth leads to lateral growth above the SiO_2 mask, reducing the gap between the triangular prisms. The gap between the triangular prisms was then filled with an $Al_{0.4}Ga_{0.6}As$ layer by switching the growth material from GaAs to $Al_{0.4}Ga_{0.6}As$. In contrast to GaAs, orientation dependence of the growth rate is small in the AlGaAs layer, although the dependence is strongly affected by the temperature and the Al content. As a result, a sharp V-groove between the triangular prisms was formed. By

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

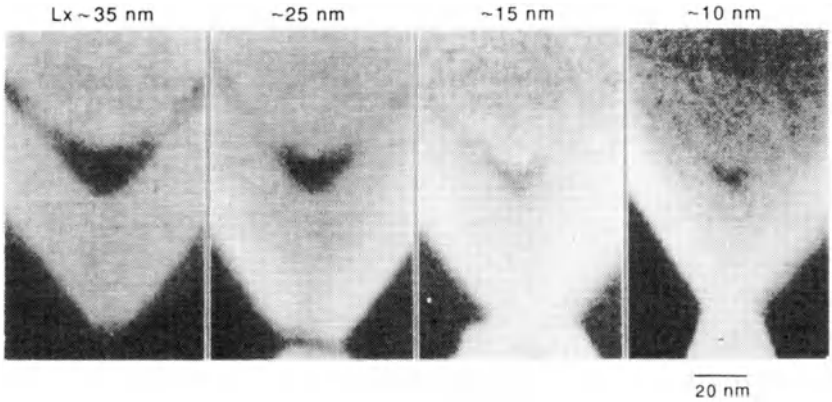


Fig. 6.3.2 SEM images of a quantum wire region with lateral widths of ~10, ~15, ~25, and ~35 nm

switching the growing layer again to GaAs, a triangular-shaped quantum wire was grown at the V-groove corner. This quantum wire was connected to thin quantum wells grown on (111)A sidewalls on large triangular GaAs prisms.

Figure 6.3.1 shows a cross-sectional SEM image of fabricated quantum wires with a 200 nm period. Even though there are ~30% lateral size fluctuations in the SiO_2 mask, quantum wires are formed uniformly owing to the lateral selective growth of the triangular prisms.

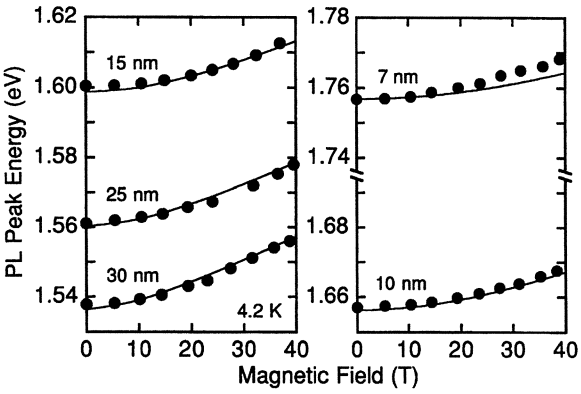


Fig. 6.3.3 Magneto-PL shift for the quantum wires with various widths

By changing the growth time of GaAs for the quantum wires, we can obtain wires with various lateral widths. Figure 6.3.2 shows SEM images of a quantum wire region with widths of ~10, ~15, ~25, an ~ 35nm [12]. Each quantum wire is smoothly connected to quantum well layers with ~2, ~3, ~5, and ~7 nm thickness.

A photoluminescence (PL) peak for these quantum wires shows a clear

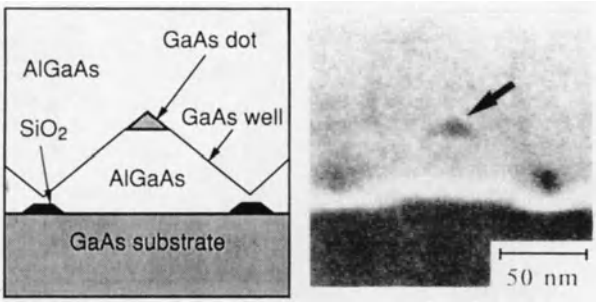


Fig. 6.3.4 SEM photograph of a GaAs quantum dot with a lateral width of 25 nm and its illustration

blue shift in comparison with bulk PL. PL was measured also for wires with various lateral widths in high magnetic fields with three different directions [13,14]. Figure 6.3.3 shows observed PL peak positions (filled circles) as a function of a magnetic field. The PL peak positions exhibit characteristic diamagnetic shift, which becomes smaller with decreasing the lateral width.

The solid curve in the figure shows the result of a variational calculation in which the electron-hole Coulomb interaction and a confinement potential corresponding to the V-groove quantum wire are included. The calculation is in good agreement with the experiments, which indicates that the diamagnetic shift is understood quite well by the formation of a 1D exciton in the quantum wire. The estimated exciton binding energy for wires with width less than 10 nm is larger than the maximum binding energy ($\sim$18 meV) observed in quantum wells.

6.3.2 Quantum Dots by Selective MOCVD

The selective growth by MOCVD also led to a successful fabrication of GaAs quantum dots on a SiO_2 patterned (100) GaAs substrate [15]. A mask consists of 100×100 nm^2 windows with a period of 140 nm. First, $Al_{0.4}Ga_{0.6}As$ plinths are formed on the SiO_2 mask. Then GaAs quantum dots are grown on the top of the AlGaAs plinths, followed by the growth of $Al_{0.4}Ga_{0.6}As$ so that the GaAs quantum dots are embedded in $Al_{0.4}Ga_{0.6}As$. Figure 6.3.4 shows a cross-sectional view of a GaAs dot surrounded by AlGaAs and its illustration. The lateral spread of the quantum dot is 25×25 nm^2.

A PL peak of the quantum dot is shifted to a higher energy by $\sim$18 meV due to the lateral confinement of carriers. The full width at half maximum (FWHM) of the PL peak is broad in comparison with that of quantum wires. This broadening is ascribed to size fluctuations of the dots (i.e., inhomogeneous broadening). In fact, in the measurement, more than 100 000 quantum dots are simultaneously excited.

A micro-PL technique was used for the measurement of a single GaAs/ $Al_{0.4}Ga_{0.6}As$ quantum dot structure at low temperatures [16]. The FWHM of the PL spectral line is plotted as a function of laser power in Fig. 6.3.5a, where the results for the reference quantum well with a thickness of 12.5

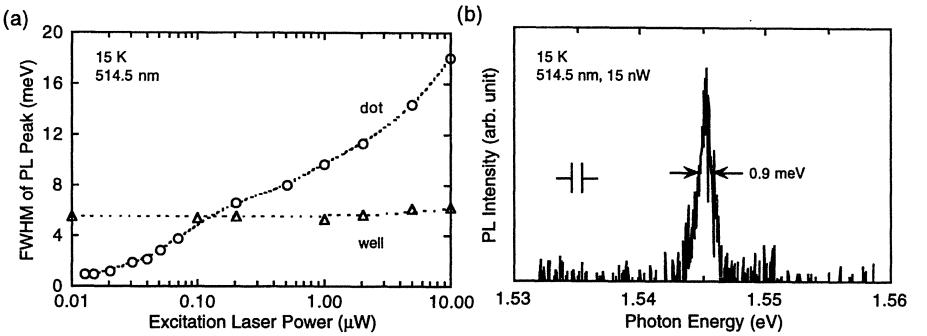


Fig. 6.3.5 (a) FWHM of the photoluminescence peak of the dot and that of the 12.5-nm thick quantum well as a function of laser power. (b) The photoluminescence spectra of the dot at the lowest laser power

nm grown under the same condition are also plotted for comparison. The FWHM of the dot becomes smaller with the decrease in the laser power, while that of the quantum well stays almost constant. At the lowest power, in which about one electron-hole pair exists in each dot on a time average, the FWHM is as low as 0.9 meV. This value is smaller than k_BT and the broadening corresponding to fluctuations of one monolayer (ML), where k_B is the Boltzmann constant and $T = 15$ K. This sharp PL peak of the dot shows that the radiative recombination of excitons in the dot occurs via the lowest energy state.

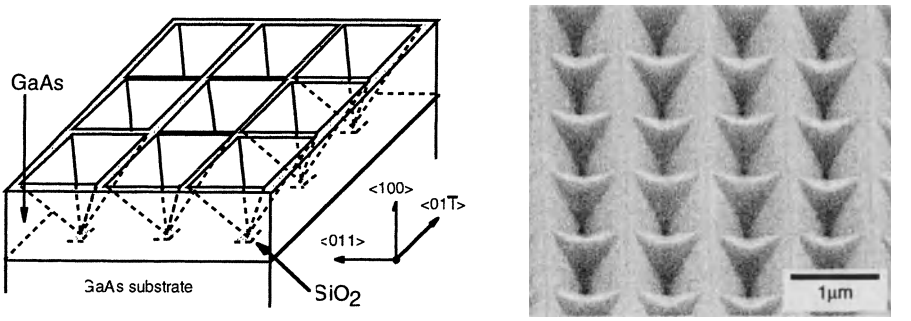


Fig. 6.3.6 Schematic illustration of 2DVG composed of (111)A and (111)B facets on GaAs (100) substrate

6.3.3 Quantum Dots in 2D V-Grooves

Quantum dots were also fabricated using 2D V-grooves (2DVGs) by selective MOCVD [17]. Figure 6.3.6 shows a schematic illustration of 2DVGs composed of (111)A and (111)B facets of GaAs. For a linear SiO_2 pattern in the $\langle 01\bar{1}\rangle$ direction, V-grooves composed of (111)A facets are formed, and when it is in the $\langle 011\rangle$ direction, V-grooves are formed by (111)B facets. 2DVGs are

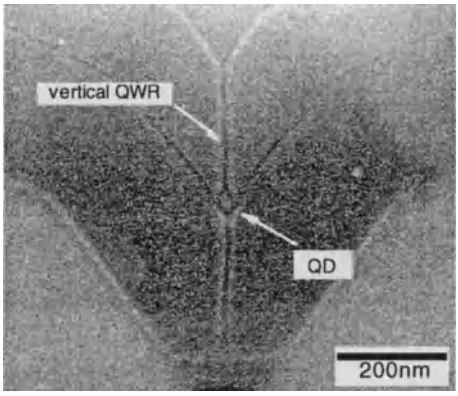


Fig. 6.3.7 SEM micrograph of a cross-section of GaAs quantum dots grown in a 2DVG. The cross-sectional plane is perpendicular to the $(0\bar{1}1)$ orientation

made from these two kinds of V-grooves. Many stacks of quantum dots can be formed by repeating the dot-barrier growth sequence.

The following summarizes the fabrication process. Before starting the MOCVD growth, a SiO_2 layer with a thickness of 20 nm was formed by plasma chemical vapor deposition on a (100) GaAs substrate. A PMMA photo-resist pattern with a thickness of 120 nm was defined on the SiO_2 layer by electron beam lithography. The size of the mask and period were 400 nm and 1 μm, respectively. The SiO_2 layer not covered by the PMMA-resist pattern mask was removed using wet chemical etching with buffered HF (17%).

The growth process by MOCVD is as follows. The GaAs is grown in the window region of the SiO_2 pattern selectively. Then, by switching the growth material from GaAs to AlGaAs, the barrier layer is grown. The size of SiO_2, which is at the bottom of the 2DVG, is larger than the size required to realize quantum structures. After the AlGaAs layer is grown, the bottom of the 2DVG is small enough to allow the growth of the quantum dot structures, because the migration length of Al is very short under a high partial pressure of AsH_3. Then, by switching the growth material from AlGaAs to GaAs, the quantum dot structures are grown at the bottom of the 2DVGs. By switching the material from GaAs to AlGaAs again, the quantum dots are buried. The growth temperature was 700°C and the V/III ratio was 100 for the growth of the AlGaAs barrier and the cap layers, and 10 for the growth of GaAs 2DVGs and the dot structures. The concentration of Al in the AlGaAs layer is 0.4, based on the flow ratio. Figure 6.3.7 shows a cross-sectional SEM image of a GaAs quantum dot. In this SEM image the sample was stain etched with NH_3-H_2O_2-H_2O for observation. The lighter parts are AlGaAs and the darker parts are GaAs.

Quantum dots with different sizes can be obtained by changing the growth time of GaAs. By using GaAs growth times of 20, 10, and 5 s, for example, we can obtain the dot sizes ~40, ~15, and ~10 nm, respectively. When we switch the material from AlGaAs to GaAs after the AlGaAs cap layer is grown and the first GaAs quantum dot is buried, we can grow a new quantum

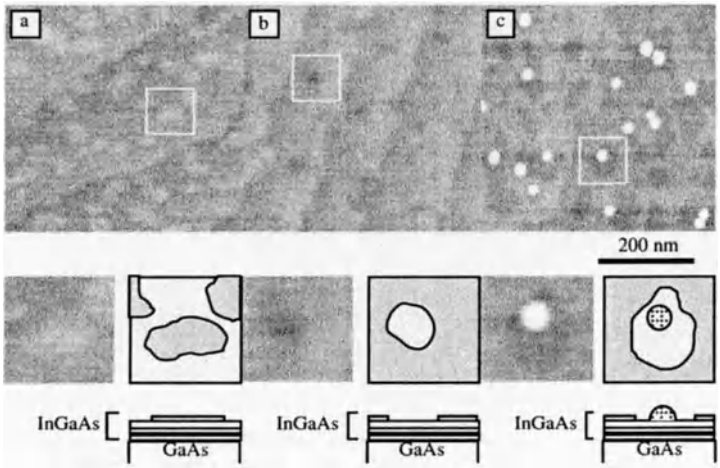


Fig. 6.3.8 AFM images and schematic illustrations of InGaAs surface. The deposition times are (**a**) 3.0 s (3.3 ML), (**b**) 3.5 s (3.9 ML), and (**c**) 3.7 s (4.1 ML)

dot. By repeating this cycle, we can grow arbitrary numbers of quantum dots vertically.

The SEM image given in Fig. 6.3.7 shows a vertical line in the center of the 2DVG. This is a vertical quantum wire (V-quantum wire) formed above the SiO_2 at the bottom of 2DVG because the content of Ga is higher than that in the surrounding AlGaAs.

6.3.4 Self-Assembled InGaAs Quantum Dots

Quantum dots can be implemented by the so-called Stranski–Krastanow (SK) growth mode of a highly strained InAs (InGaAs) layer on a GaAs unpatterned (100) substrate. The SK growth mode has been intensively investigated for the fabrication of self-assembled quantum dots. The first SK mode growth of InGaAs quantum dots by MOCVD was reported in 1994 [18–20].

In this growth mode, 3D islands are naturally formed by the strain effect after an initial 2D growth of a few MLs. The transition from a 2D layer (wetting layer) to a 3D island growth can be seen by RHEED during the growth of the strained layer and the critical thickness of InGaAs for the change in the growth mode has been measured.

Figure 6.3.8 shows AFM images and illustrations of an InGaAs surface grown on GaAs for 3.0, 3.5, and 3.7 s, respectively. As shown in Fig. 6.3.8a, there exist 2D islands on terraces in the sample grown for 3.0 s. The formation of the 2D islands is due to the wide terraces and the strain effects. Note that no such 2D island was observed when only GaAs was grown. In fact, in GaAs a step-flow growth always occurs at the step edge of the substrate. 2D islands were at the level of the fourth layer of InGaAs. When the growth time is 3.5 s (Fig. 6.3.8b), many holes appeared on the surface, because the fourth layer

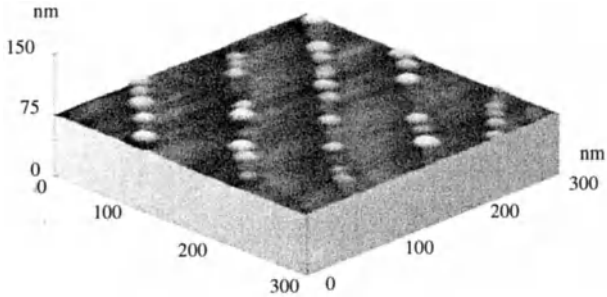


Fig. 6.3.9 AFM images of self-aligned quantum dots grown on a [010]-misoriented surface

occupied most of the surface.

When the growth time is 3.7 s, 3D islands (i.e., quantum dots) distributed randomly appear, as shown in Fig. 6.3.8c. Figure 6.3.8b,c suggests that the critical thickness is 4 MLs under this growth condition. The diameter and the height of the quantum dots were 20 nm and 4.5 nm, respectively. Figure 6.3.8c also shows that some areas of the third layer were exposed on the surface around dots. These results indicate that quantum dots are formed by desorbing some atoms nearby at the fourth layer. Photoluminescence spectra for these samples exhibit consistent results with this AFM observation.

Further, by choosing growth conditions properly, we can create straight multi-atomic step edges and align quantum dots linearly on them. Figure 6.3.9 shows AFM images of aligned quantum dots grown on a [010]-misorientated surface. The growth time for InGaAs dots was 3.5 s. The diameter of the dots is less than 20 nm.

6.3.5 Use of Spinodal Phase Separation

Self-assembly of quantum dots can be realized by utilizing a temperature gradient through a spinodal phase separation [21]. In this mode, quantum dots are grown in InGaAs layers with a nominal In content as low as $x = 0.03$. For the growth of quantum dots the temperature was lowered from an initial temperature of 550°C by 100°C during the deposition of InGaAs, with a rate of temperature decrease of 0.28°C/s, as indicated in Fig. 6.3.10. This method leads to a phase separation into island-like structures during the growth. A (100) GaAs substrate with no misorientation was used.

The sample structure is given in Fig. 6.3.10 together with an SEM image. First, a GaAs buffer is grown at a temperature between 500°C and 600°C. Then the gas flow is switched for the growth of InGaAs, while simultaneously the temperature is lowered by 50°C or 100°C. The quantum dot layer is then capped at the low temperature immediately after growth, while the flow rates for the precursors are kept constant. A quantum dot layer at the interface between GaAs and InGaAs is covered with a low-In-content cap layer ($x = 0.03$). The SEM image on the left shows a quantum dot at the interface between GaAs and InGaAs. The contrast in this image suggests that much

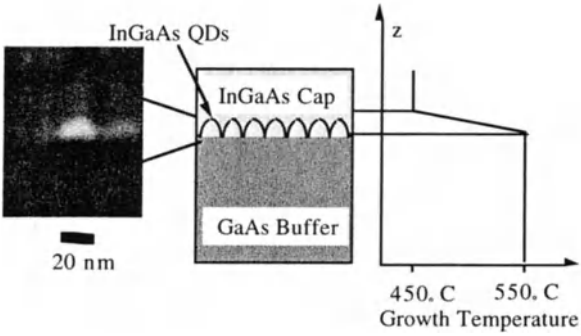


Fig. 6.3.10 Sketch of the sample structure and typical growth temperatures

of the In of the surrounding nominal $In_{0.03}Ga_{0.97}As$ material is concentrated in the quantum dots.

Figure 6.3.11 shows an AFM image of a sample in which the growth was interrupted after 10 nm. The structure was grown with a nominal In content of $x = 0.03$, and a rate of temperature decrease of 0.28°C/s at a growth temperature that was lowered from 550°C. The growth conditions for this sample were therefore similar to those for capped structures. The etched surface displays quantum dots with an average diameter of ~22 nm and an area density of $\sim 2 \times 10^{10}$ cm^{-2}. XPS measurements, taken together with the quantum dot coverage rate, allow an estimation of the In content of the quantum dots to be $x \sim 0.4$.

Excited states in the quantum dots are also investigated by using photoluminescence at moderately high excitation densities. Figure 6.3.12 shows semilogarithmic PL spectra from a sample with a nominal In content of $x = 0.03$ (temperature decrease rate: 0.28°C/s) at excitation densities P ranging from 0.3 to 600 mW/cm^2. At $P = 30$ mW/cm^2 a second peak appears at ~1.37 eV, which can be attributed to a higher subband transition. The peak at ~1.45 eV is probably due to 2D InGaAs structures at the GaAs/InGaAs interface. The excited state exhibits a larger inhomogenous broadening (~56 meV) than the ground state (~34 meV), as expected from the theory; with increased transition energies in the quantum dots an increased level separation is predicted. The inset in Fig. 6.3.12 depicts the PL intensity divided by the input power density on a linear scale for the ground state transition, measured at a different spot on the sample. The observed PL characteristics are comparable to those reported by others and indicate a discrete density of states in the optically active layer.

A crucial barrier for obtaining semiconductor nanostructures useful for devices is the controllability of the size of the nanostructures. For example, a typical value of the FWHM of the PL spectrum of self-assembling quantum dots grown by the SK mode is about 40 meV. Unless an FWHM smaller than 5 meV is achieved, the advantages of the zero-dimensional density of states cannot be utilized. This controllability of the size is one of the most serious issues.

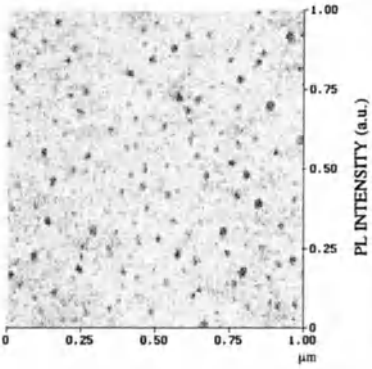


Fig. 6.3.11 AFM images of etched InGaAs surfaces

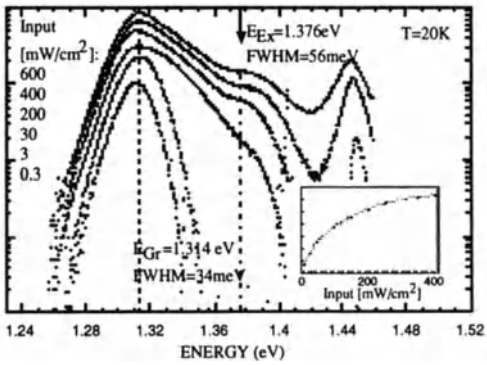


Fig. 6.3.12 PL spectra for different excitation densities given in mW/cm^2, on a semilogarithmic scale

References

1 B.I. Miller, A. Shahar, U. Koren, and P.J. Corvini, Appl. Phys. Lett. **54**, 188 (1989)
2 K. Kash, A. Scherer, J.M. Worlock, H.G. Craighead, and M.C. Tamargo, Appl. Phys. Lett. **49**, 1043 (1986)
3 J. Cibert, P.M. Petroff, G.J. Dolan, S.J. Pearton, A.C. Grossard, and J.H. English, Appl. Phys. Lett. **49**, 1275 (1986)
4 H. Temkin, G.L. Dolan, M.B. Panish, and S.N.G. Chu, Appl. Phys. Lett. **50**, 413 (1987)
5 D. Gershoni, H. Temkin, G.L. Dolan, J. Dunsmuir, S.N.G. Chu, and M.B. Panish, Appl. Phys. Lett. **53**, 995 (1988)
6 J.A. Lebens, C.S. Tsai, and K.J. Vahala, Appl. Phys. Lett. **56**, 2642 (1990)
7 T. Fukui, S. Ando, and Y.K. Fukai, Appl. Phys. Lett. **57**, 1209 (1990)
8 R. Bhat, E. Kapon, D.M. Hwang, M.A. Koza, and C.P. Yun, J. Cryst. Growth, **93**, 850 (1988)
9 E. Colas, S. Simhony, E. Kapon, R. Bhat, D.M. Hwang, and P.S.D. Lin, Appl. Phys. Lett. **57**, 914 (1990)
10 E. Kapon, S. Simhony, D.M. Hwang, E. Colas, and N.G. Stoffel, *Proceedings of 12th IEEE Interrnational Semiconductor Laser Conference, Switzerland, 1990*, p. 80
11 S. Tsukamoto, Y. Nagamune, M. Nishioka, and Y. Arakawa, J. Appl. Phys. **71**, 533 (1992)
12 S. Tsukamoto, Y. Nagamune, M. Nishioka, and Y. Arakawa, Appl. Phys. Lett. **63**, 310 (1993)
13 Y. Nagamune, Y. Arakawa, S. Tsukamoto, and M. Nishioka, Phys. Rev. Lett. **69**, 2963 (1992)
14 Y. Nagamune, T. Tanaka, T. Kono, S. Tsukamoto, M. Nishioka, and Y. Arakawa, Appl. Phys. Lett. **66**, 2502 (1995)
15 Y. Nagamune, S. Tsukamoto, M. Nishioka, Y. Arakawa, K. Uchida, and N. Mira, Appl. Phys. Lett. **64**, 2495 (1994)
16 Y. Nagamune, H. Watabe, M. Nishioka, and Y. Arakawa, Appl. Phys. Lett. **67**, 3257 (1995)
17 S. Ishida and Y. Arakawa, Appl. Phys. Lett. (submitted for publication)
18 J. Oshinowo, M. Nishioka, S. Ishida and Y. Arakawa, Appl. Phys. Lett. **65**, 1421 (1994)
19 J. Oshinowo, M. Nishioka, S. Ishida and Y. Arakawa, Jpn. J. Appl. Phys. **33**, L1634 (1994)
20 M. Kitamura, M. Nishioka, and Y. Arakawa, J. Cryt. Growth **43**, 9844 (1996)
21 R. Schur, F. Sogawa, M. Nishioka, S. Ishida, and Y. Arakawa, Jpn. J. Appl. Phys. **36** Part 2, L120 (1997)

6.4 Quantum Wires on Vicinal GaAs (110) Surfaces

6.4.1 Introduction

MBE growth of GaAs on GaAs (110) surfaces has been reported not to provide smooth surface morphology due to columnar growth of epitaxial layers [1]. Highly faceted surfaces are observed for MBE growth on (110) surfaces. Only vicinal (110) surfaces misoriented 6° toward (111)A have provided device quality GaAs layers [2]. However, when GaAs/AlGaAs superlattices are grown on vicinal (110) surfaces misoriented toward (111)A, quantum wire-like (QWR-like) structures have been observed [3,4]. These QWR-like structures are induced by the formation of coherently aligned giant growth steps and remarkable composition and thickness modulation of AlGaAs and GaAs layers at the giant step edges. Using these phenomena, AlGaAs and GaAs quantum wires (QWRs) have been formed on vicinal (110) surfaces misoriented toward (111)A.

6.4.2 Step Structures

Figure 6.4.1 shows RHEED (reflection high-energy electron diffraction) patterns during $Al_{0.5}Ga_{0.5}As$ growth on vicinal (110) surfaces misoriented 6° toward (111)A using solid arsenic (As_4) at 650°C and 580°C, respectively [5]. At a temperature of 650°C, the surface exhibits the (1×1) structure with splitting around the (110) surface integer-order rods as shown in Fig. 6.4.1a, which is characteristic of well-ordered arrays of steps. By analyzing the RHEED pattern, we deduce an average step height of 4.0 Å, which is very close to twice the (110) interplanar spacing, and an average terrace width of 39 Å. From these values a misorientation of 5.9° is calculated, which equals the off-angle of the substrate. This also indicates that the surface has a regular step array with double-layer height. At 580°C, an unsplit (1×1) streak pattern with a slash pattern is observed as shown in Fig. 6.4.1b. This RHEED pattern reveals that giant steps or multi-atomic steps are formed on the surface and (331) facets appear at the step edges from the angle (13°) between the slash streak and the (110) surface rod.

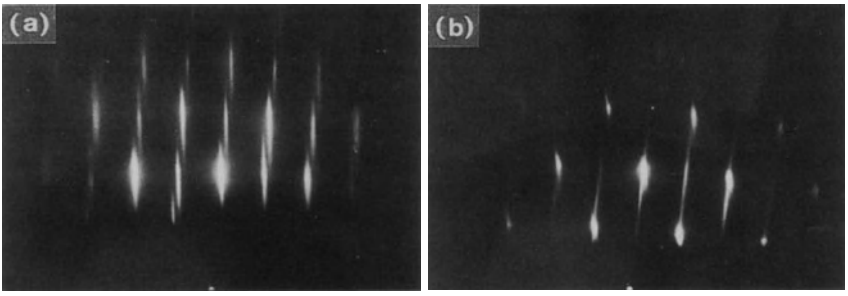


Fig. 6.4.1 RHEED patterns during $Al_{0.5}Ga_{0.5}As$ growth on vicinal (110) surface at (**a**) 650°C and (**b**) 580°C

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

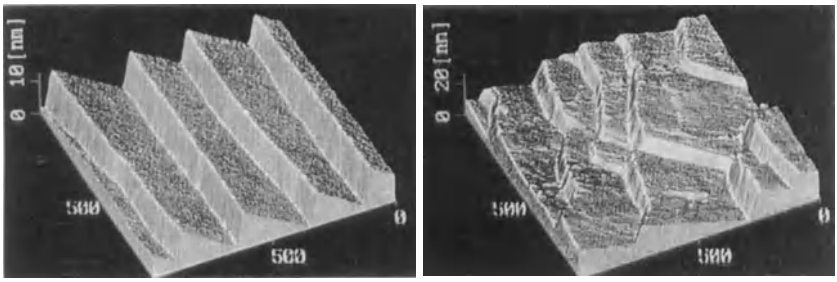


Fig. 6.4.2 AFM observations of the surface morphologies of samples using (**a**) AsH_3 and (**b**) solid As (1×1 μm^2)

From the RHEED observations at different growth temperatures, we find that for AlGaAs growth, there exists a transition temperature at about 600°C from giant steps to diatomic steps. This transition temperature from giant steps to diatomic steps is slightly different between the growths using solid As (630°C) and AsH_3 (580°C). Similar RHEED patterns are observed for AlGaAs growth using both solid As and AsH_3. However, the surface morphologies observed by AFM are quite different between the samples grown using solid As and AsH_3, as shown in Fig. 6.4.2. The giant step structure of the sample grown using solid As is not coherent compared with that using AsH_3. When we use a GaAs source, we get the same surface morphologies as those of the sample grown using AsH_3. The difference in surface morphologies is due to the difference between MBE growth using As_2 and As_4 [6].

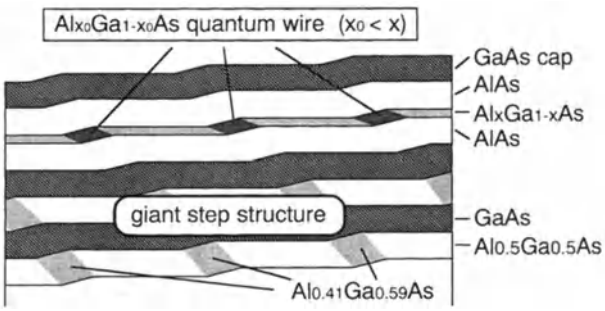


Fig. 6.4.3 Schematic illustration of AlGaAs quantum wire structure

AFM observations of the surface morphologies of samples grown using As_2 show that the giant step structure depends strongly on the growth temperature and weakly on the V/III ratio. The terrace width becomes wider as the growth temperature increases. However, the step edges are not so sharp as those at a lower growth temperature. The giant step structure also depends on the off-angles of substrates. Coherent giant step structures are observed for 3°-off and 6°-off samples which show a different terrace width, but not for an 8°-off sample. The step height and terrace width are controlled by controlling

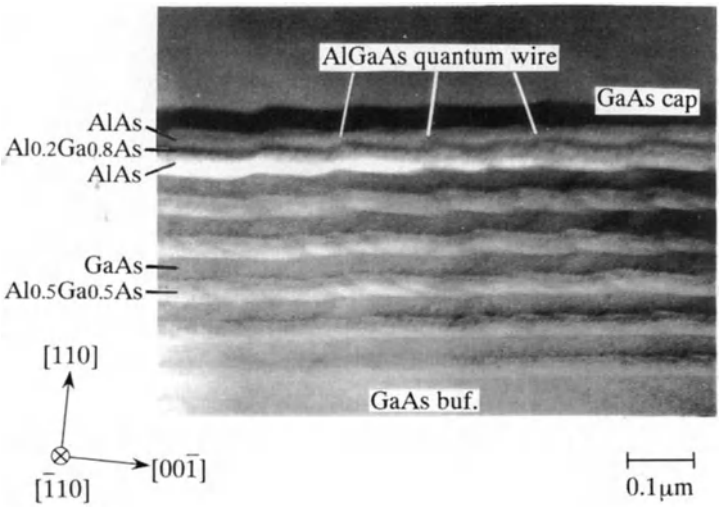


Fig. 6.4.4 Cross-sectional TEM micrograph of AlGaAs QWR structure

the growth conditions and off-angle of substrates. When we grow superlattices (SL) instead of AlGaAs layers, more coherent giant step structures are obtained [7].

6.4.3 AlGaAs Quantum Wires

AlGaAs QWRs are formed by following MBE growth procedure using AsH_3 [8]. First, GaAs (30nm)/$Al_{0.5}Ga_{0.5}As$(30nm) SL with 5 periods is grown on vicinal (110) surfaces misoriented 6° toward (111)A. The growth temperature and AsH_3 flow rate are 540°C and 1.5 sccm, respectively. During the SL growth, giant steps are coherently formed and at the step edges the region with low AlAs mole fraction (~0.41) are formed in $Al_{0.5}Ga_{0.5}As$ layers [7]. Next, an AlAs(30nm)/$Al_xGa_{1-x}As$(10nm)/AlAs(30nm) single quantum well is grown on the surface with giant steps. AlGaAs QWRs with AlAs mole fraction x_0 less than x are naturally formed at the giant step edges due to composition modulation, as schematically illustrated in Fig. 6.4.3. The AlAs mole fraction x is varied from 0.2 to 0.5. For comparison, the above structure is grown on (100) substrates which are placed side-by-side to the (110) samples.

Figure 6.4.4 shows a cross-sectional TEM image of the sample. The dark layers are GaAs and the bright layers are AlGaAs. The giant steps become higher and more coherent as the SL grows. Slightly darker areas are observed at the step edges in $Al_{0.5}Ga_{0.5}As$ layers, indicating the composition modulation at the giant step edges, which is confirmed by energy-dispersive x-ray spectroscopy [7]. Although the compositional modulation is not clear in the $Al_{0.2}Ga_{0.8}As$ quantum well, it is confirmed from photoluminescence (PL) results [8].

Figure 6.4.5 shows PL spectra of the (100) and (110) samples with $x = 0.2$. The (110) samples shows a peak remarkably red-shifted from that of the (100)

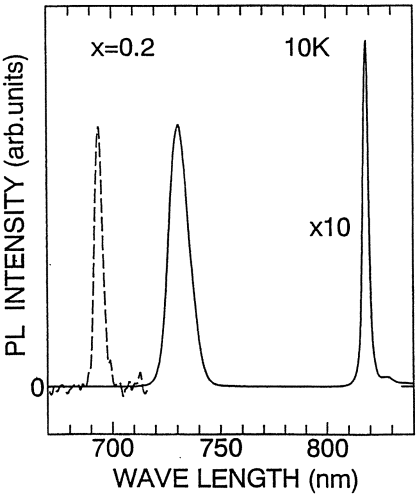


Fig. 6.4.5 PL spectra of the (110) sample (solid line) and (100) sample (dashed line) with $x = 0.2$

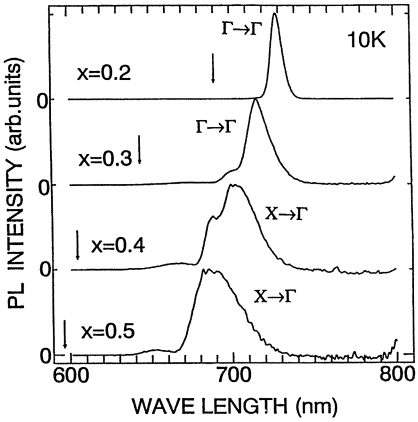


Fig. 6.4.6 PL spectra of the (110) samples. Arrows indicate band-gap positions of $Al_xGa_{1-x}As$

sample, indicating the formation of QWRs with x_0 less than 0.14. The intense peak at 817 nm comes from GaAs layers of the underlying SL. Figure 6.4.6 shows PL spectra of (110) samples with x ranging from 0.2 to 0.5. For all x values, PL peaks shift toward longer wavelengths than the band gap positions (arrows in the figure) showing the QWR formation. Two samples with $x = 0.4$ and 0.5 show broad PL bands with very weak intensities. The integrated intensity of the peak for $x = 0.4$ is three order of magnitude weaker than that for $x = 0.2$. Then, we assigned two PL bands of the samples with $x = 0.4$ and 0.5 to the indirect transition from the X point of AlAs barriers to the Γ holes of the wires, and two other PL bands to the $\Gamma \rightarrow \Gamma$ transition. The PL from these QWRs are strongly polarized parallel to the [110] wire direction as shown in Fig. 6.4.7 [8]. The polarization degree, $(I_{[110]} - I_{[001]}/(I_{[110]} + I_{[001]})$, is estimated to be 0.25, indicating hole confinement to QWRs.

We have taken a set of cathodoluminescence (CL) spot spectra. The excitation area is about 0.1 μm in diameter. The spot spectrum changes with position and shows a fine structure with several extremely sharp lines. In order to confirm that each sharp CL line originates from a specific single QWR, we took CL images at different wavelengths, as shown in Fig. 6.4.8 [9]. These monochromatic CL images clearly show that each QWR emits luminescence at a different wavelength. Longitudinal nonuniformity is observed in some QWRs.

6.4.4 GaAs Quantum Wires

GaAs quantum wires are formed by a similar MBE growth procedure to that used for AlGaAs QWR formation. The only difference is the growth of an $Al_{0.5}Ga_{0.5}As$(30 nm)/GaAs(t)/$Al_{0.5}Ga_{0.5}As$(30 nm) single quantum well on

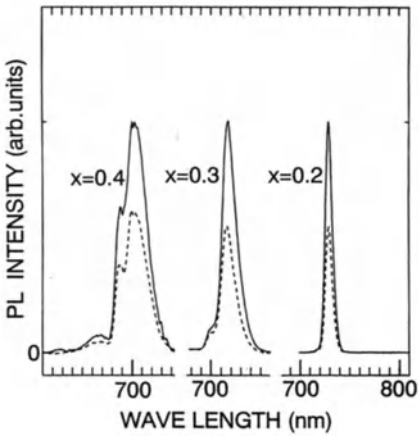


Fig. 6.4.7 Polarization dependence of PL of AlGaAs QWRs

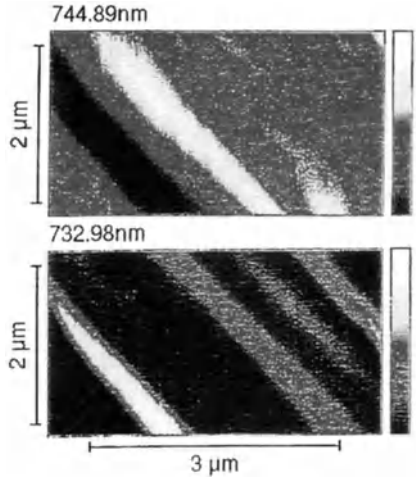


Fig. 6.4.8 CL images of AlGaAs QWRs at different wavelengths

the surface with giant steps, as shown in Fig. 6.4.9. Since the thickness of the GaAs layer at the giant step edges is thicker than that at the terraces (thickness modulation), GaAs QWRs are naturally formed at giant step edges. The thickness of the GaAs well is changed from 1 to 3 nm [10]. TEM, AFM, and PL observations confirmed the formation of GaAs QWRs. Although these QWRs are successfully formed, broad PL bands are observed for the GaAs QWRs, indicating nonuniformity of the QWR structures. Moreover, from the temperature dependence of the PL intensities the confinement energies between the step edge wires and terrace wells are found to be low (~150 meV) [11].

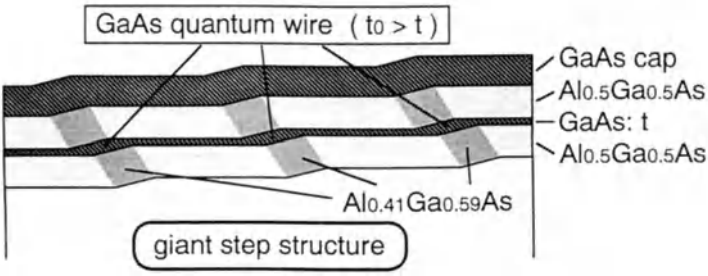


Fig. 6.4.9 Schematic illustration of GaAs QWR structure

A study of dependence on the growth temperature reveals that a lower growth temperature is suitable for the formation of sharp and clear giant steps and a higher growth temperature is better for the larger thickness modulation. The optimum temperature for the giant step formation is 500°C and for the

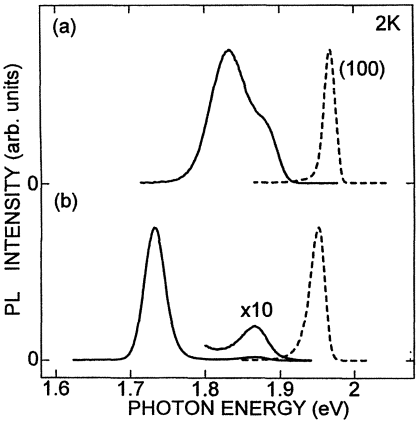


Fig. 6.4.10 PL spectra of (110) samples grown by (a) one-step and (b) two-step growth (solid line)

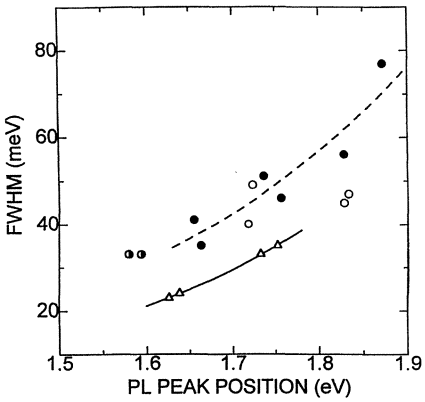


Fig. 6.4.11 PL peak position versus spectral width comparing one-step growth (open circle: 500°C, closed circle: 520°C, and half-closed circle: 540°C) and two-step growth (triangle: 500°C → 550°C)

thickness modulation 550°C. As a consequence, we grow the samples using two-step growth, i.e, giant steps at 500°C and QWRs at 550°C [12]. Figure 6.4.10 shows PL spectra of (110) samples grown by one-step and two-step growth and (100) samples which are placed side-by-side to the (110) samples during MBE growth for comparison. Comparing these two PL spectra of (110) samples, the sample grown by two-step growth shows a more red-shifted peak with narrower spectral width. This result indicates that at a higher growth temperature, the Ga or GaAs migration length becomes longer and the GaAs thickness at step edges becomes thicker. CL images showed that two PL peaks of the sample grown by one-step growth originate from two groups of wires with different size [11].

When the thickness increases, the spectral width becomes narrower, because the effect of monolayer thickness fluctuation is not large for thick wires. Consequently, we took the PL peak position versus FWHM (full width at half minimum) as shown in Fig. 6.4.11. This figure compares one-step growth at different temperatures and two-step growth. Obviously, the samples grown by two-step growth show narrower spectral width, indicating more uniform QWR structures, which is also confirmed by CL images [12].

Using this two-step MBE growth technique, stacked multi-QWRs are grown as shown in a cross-sectional TEM micrograph in Fig. 6.4.12 [13]. First, giant steps are formed at 500°C on vicinal (110) surfaces. Then, stacked QWRs are formed at 550°C on this surface with giant steps by alternately growing GaAs and $Al_{0.5}Ga_{0.5}As$ layers. The thickness of GaAs and AlGaAs layers are changed from 1 to 3 nm. These thickness are those of (100) samples placed side-by-side with the (110) samples. The QWRs are thicker than these values due to thickness modulation. When the thickness of the GaAs layers of the corresponding (100) sample is 3 nm, that of QWRs increases to 9 nm, as

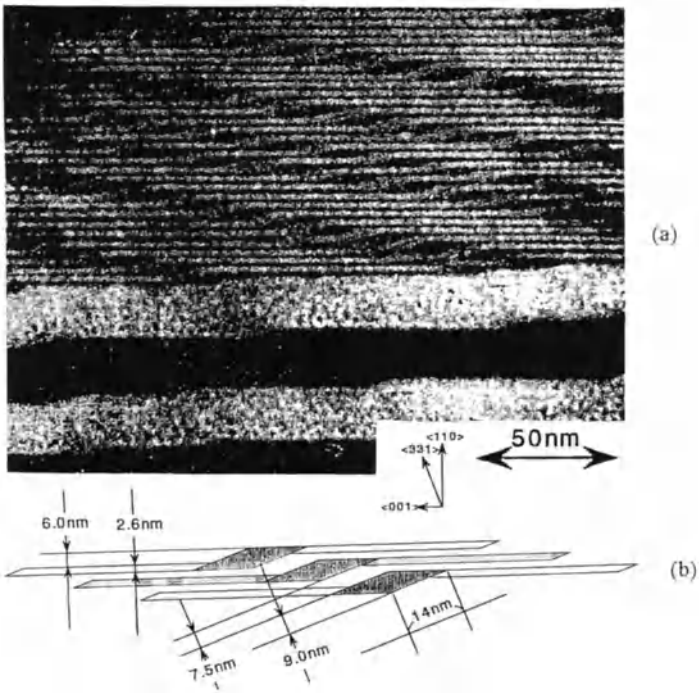


Fig. 6.4.12 Cross-sectional TEM micrograph of stacked QWRs. Dark layers are GaAs

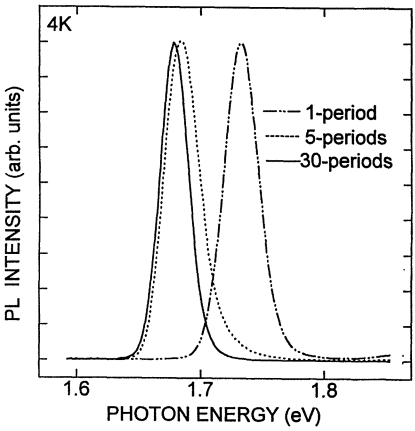


Fig. 6.4.13 PL spectra of stacked QWRs with different periods

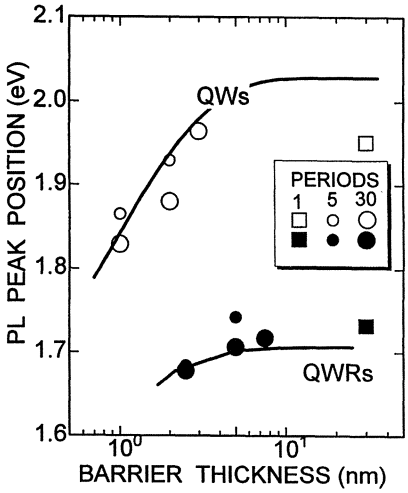


Fig. 6.4.14 PL peak position of quantum wells of (100) samples and wires of (110) samples as a function of AlGaAs barrier thickness. Solid lines are calculated

shown in Fig. 6.4.12b.

Figure 6.4.13 compares the PL of a single layer of QWRs with those of stacked ones with 5 and 30 periods. The thickness of GaAs of the (100) samples is 1 nm. The PL peak positions of the stacked samples are red-shifted from that of the single-layer one. The spectral width of the stacked samples is not broadened, showing that uniformity is conserved during the stacking process. The red shift of the PL peak position is considered to be due to coupling between the wires. Figure 6.4.14 shows the PL peak position of quantum wells of (100) samples and wires of (110) samples as a function of AlGaAs barrier thickness. The solid lines are calculated, assuming the coupling effects. For both cases, the peak energy decreases as the barrier decreases, showing the coupling between wells or wires.

References

1 W.I. Wang. J. Vac. Sci. Technol. B **1**, 630 (1983)

2 L.T.P. Allen, E. R. Weber, J. Washburn, and Y.C. Pao. Appl. Phys. Lett. **51**, 670 (1987)

3 M. Sato, K. Maehashi, H. Asahi, S. Hasegawa, and H. Nakashima. Superlattices and Microstructures, **7**, 279 (1990)

4 S. Hasegawa, M. Sato, K. Maehashi, H. Asahi, and H. Nakashima. J. Crystal Growth, **111**, 371 (1991)

5 S. Hasegawa, K. Kimura, M. Sato, K. Maehashi, and H. Nakashima, Surf. Sci. **267**, 5 (1992)

6 S. Hasegawa, K. Sato, S. Torii, and H. Nakashima, to be published in J. Crystal Growth

7 M. Takeuchi, K. Shiba, H. K. Huang, K. Inoue, K. Maehashi, and H. Nakashima, J. Crytal Growth **150**, 441 (1995)

8 K. Inoue , K. Kimura, K. Maehashi, S. Hasegawa, H. Nakashima, M. Iwane, O. Matsuda, and K. Murase, J. Crystal Growth **127**, 1041 (1993)

9 H. Nakashima, M. Takeuchi, K. Kimura, M. Iwane, H. K. Huang, K. Inoue, J. Christen, M. Grundmann and D. Bimberg, Solid-State Electron. **40**, 319 (1996)

10 M. Takeuchi, K. Shiba, K. Sato, H.K. Huang, K. Inoue, and H. Nakashima, Jpn. J. Appl. Phys. **34**, 4411 (1995)

11 H. Nakashima, M. Takeuchi, K. Inoue, T. Takeuchi, Y. Inoue, P. Fischer, J. Christen, M. Grundmann, and D. Bimberg, Physica B **227**, 291 (1996)

12 M. Takeuchi, T. Takeuchi, Y. Inoue, T. Kato, K. Inoue, H. Nakashima, K. Maehashi, P. Fischer, J. Christen, M. Grundmann, and D. Bimberg, Superlattices & Microstructures **22** (1997) (in press)

13 T. Kato, T. Takeuchi, Y. Inoue, K. Inoue, S. Hasegawa, and H. Nakashima, *Extended Abstract of 1996 Int. Conf. on Solid State Devices and Materials, Yokohama, 1996*, p. 872

6.5 Tilted T-Shaped and (775)B Quantum Wires

6.5.1 Introduction

We describe two kinds of high quality GaAs/AlGaAs quantum wires (QWRs) grown by MBE: (1) GaAs/$Al_{0.3}Ga_{0.7}As$ tilted T-shaped QWRs (T-QWRs) with extremely precisely controlled line width and cross section fabricated on reverse-mesa etched stripes on a (100) GaAs substrate by a two-step growth of glancing-angle MBE (GA-MBE) and normal MBE and (2) extremely high density (8×10^5 cm^{-1}) of GaAs/AlAs QWRs naturally formed in a GaAs/AlAs quantum well (QW) with a regularly corrugated AlAs/GaAs interface and a flat GaAs/AlAs interface grown on a (775)B-oriented GaAs substrate by MBE. The former is favorably compared with conventional T-QWRs fabricated on a (110) cleaved surface of a GaAs/AlGaAs multi-quantum well (MQW) layer by cleaved edge overgrowth (CEO), because many tilted T-QWRs can be made on a large area of usual (100) GaAs substrate surface. The latter shows high optical quality, high density, and high uniformity in addition to high reproducibility of fabrication, indicating its high application potential for optical devices.

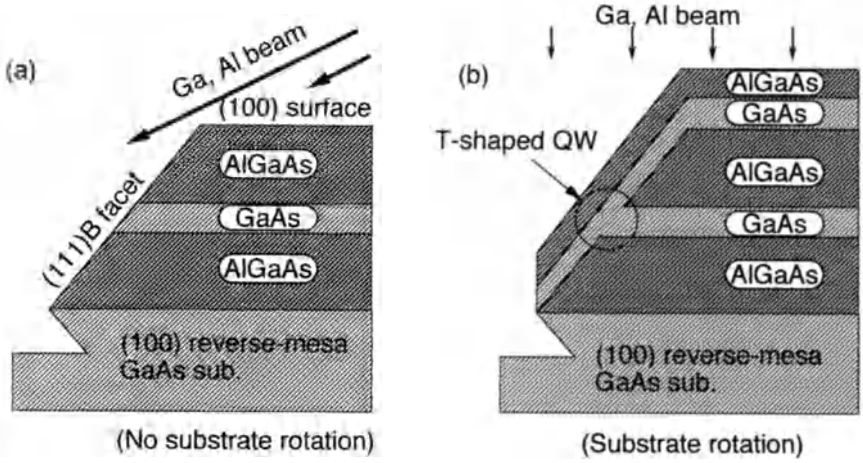


Fig. 6.5.1 Tilted T-shaped GaAs/$Al_{0.3}Ga_{0.7}As$ QWRs grown by glancing-angle molecular beam epitaxy (GA-MBE) (**a**) and overgrowth on (111) B facet (**b**). Tilted GaAs/AlGaAs T-shaped QWRs are fabricated by overgrowing GaAs QW and AlGaAs barrier layers on the (111) B facet while rotating the substrate (normal MBE mode)

6.5.2 GaAs/$Al_{0.3}Ga_{0.7}As$ Tilted T-shaped QWRs

Pfeiffer et al. [1] fabricated GaAs/$Al_{0.3}Ga_{0.7}As$ T-QWRs by overgrowing a GaAs QW layer and an $Al_{0.3}Ga_{0.7}As$ barrier layer on a (110) cleaved plane of a GaAs/$Al_{0.3}Ga_{0.7}As$ MQW layer grown on a (100) GaAs substrate by MBE (CEO method). The cross-sectional dimension of the T-QWRs was precisely controlled on the atomic scale, and high-quality T-QWRs have been obtained [1–4]. These T-QWRs, however, are fabricated only on a very limited area

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

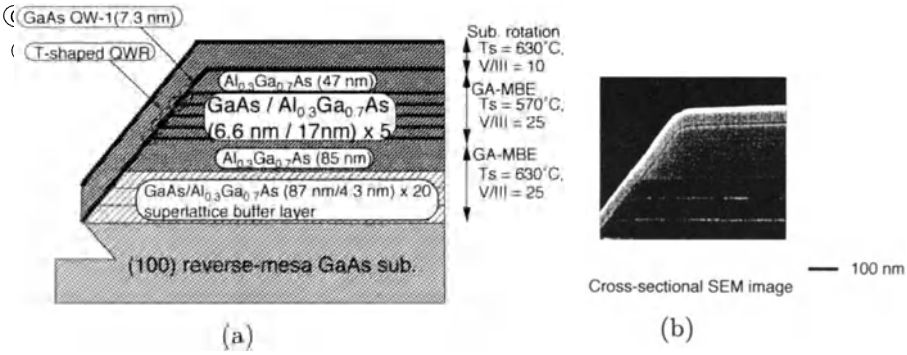


Fig. 6.5.2 Sample structure of $GaAs/Al_{0.3}Ga_{0.7}As$ tilted T-QWRs (**a**) and its cross-sectional SEM image (**b**). After [7]

of the cleaved surface of the MQW layer, making their device applications difficult. In order to avoid this inconvenience, $GaAs/Al_{0.3}Ga_{0.7}As$ tilted T-QWRs were fabricated by using (111) B facet planes, instead of the (110) cleaved plane, preferentially grown on a reverse-mesa etched GaAs substrate by GA-MBE [5–8]. In this case, many tilted T-QWRs can be easily fabricated on a lot of mesa stripes on a large (100) GaAs substrate surface, keeping the precise controllability in size for each QWR.

6.5.3 Fabrication of $GaAs/Al_{0.3}Ga_{0.7}As$ Tilted T-QWRs

During GA-MBE growth (no substrate rotation), a preferentially formed (111) B facet plane is not directly exposed to both Ga and Al beams due to a self-shadowing effect as shown in Fig. 6.5.1a. After the GA-MBE growth of a $GaAs/Al_{0.3}Ga_{0.7}As$ MQW layer, a cross-section of the MQW structure appears on the (111) B facet plane when we choose the right growth conditions where interfacet migration of Ga and Al atoms does not occur from the (100) surface to the (111) B facet. $GaAs/Al_{0.3}Ga_{0.7}As$ tilted T-QWRs are fabricated by overgrowing a GaAs QW layer and an $Al_{0.3}Ga_{0.7}As$ barrier layer on the (111) B facet plane by normal MBE, rotating the substrate only after the GA-MBE growth (Fig. 6.5.1b).

A fabricated structure of the $GaAs/Al_{0.3}Ga_{0.7}As$ tilted T-QWRs is shown in Fig. 6.5.2a [7]. A $GaAs/Al_{0.3}Ga_{0.7}As$ (87 nm/4.3 nm, 20 periods) superlattice buffer and an $Al_{0.3}Ga_{0.7}As$ layer (85 nm) are grown on a reverse-mesa etched (100) GaAs substrate by GA-MBE (T_s = 630°C, V/III=25). Next, a $GaAs/Al_{0.3}Ga_{0.7}As$ (6.6 nm/17 nm, 5 periods) MQW layer and an $Al_{0.3}Ga_{0.7}As$ barrier layer (30 nm) are grown by GA-MBE with a decreased T_s (570°C) and a high V/III ratio (25) to reduce the interfacet migration of Ga and Al atoms. Finally, a $GaAs/Al_{0.3}Ga_{0.7}As$ (7.3 nm/51 nm) single quantum well (SQW) and a 12 nm thick GaAs cap layer are overgrown on the (111) B facet under optimized overgrowth conditions (T_s=630°C, V/III=10) to obtain high optical quality of the $GaAs/Al_{0.3}Ga_{0.7}As$ SQW on the (111) B facet while rotating the substrate at 60 rpm. Figure 6.5.2b shows the cross-sectional SEM image of the grown sample [7]. Five tilted T-QWRs are formed

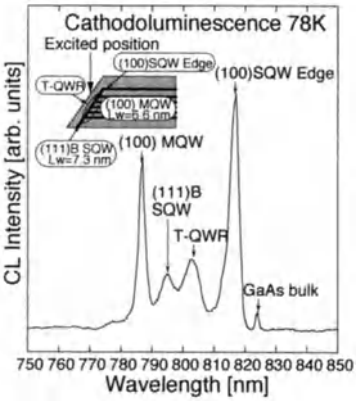


Fig. 6.5.3 CL spectrum (78 K) observed in the $GaAs/Al_{0.3}Ga_{0.7}As$ T-QWRs sample when the (111) B facet region is excited by the electron beam. After [7]

at the junction region of MQW on the (100) plane and the SQW on the (111) B facet plane.

a. Cathodoluminescence Measurements

A cathodoluminescence (CL) spectrum at 78 K [7] is shown in Fig. 6.5.3, which is obtained when a focused electron beam (0.2 μm diameter, 7 keV) excites the (111) B facet region (just above the tilted T-QWRs) with excitation intensity of nA. All CL peaks are well resolved. CL peaks at 787 nm and 795 nm come from the $GaAs/Al_{0.3}Ga_{0.7}As$ MQW on the (100) plane and the $GaAs/Al_{0.3}Ga_{0.7}As$ SQW on the (111) B facet, respectively. The luminescence from the tilted T-QWRs is observed at 803 nm. This peak energy is 1.544 eV, which is 31 meV lower than that of MQW on the (100) plane and 15 meV lower than that of SQW on the (111) B facet plane. This red shift is caused by the weaker confinement of carriers of the tilted T-QWRs. A peak at 817 nm comes from the edge of SQW on the (100) plane which is formed with a thicker QW due to the surface migration of Ga atoms from the (111) B facet plane to (100) plane when $GaAs/Al_{0.3}Ga_{0.7}As$ SQW is overgrown on the (111) B facet by normal MBE. When the excitation intensity is reduced to nA, the full width at half maximum (FWHM) of the peak (803 nm) from the tilted T-QWRs become as small as 10 meV [7], which is almost comparable to those (10 meV [2], 4–7 meV [3,4]) of conventional $GaAs/Al_{0.3}Ga_{0.7}As$ T-QWRs fabricated by the CEO method.

b. Calculation of Electron and Hole States in T-QWRs

Energy levels of electrons and holes are calculated by using the finite element method [7]. The z axis is taken to be perpendicular to the (100) substrate surface. The y axis is parallel to the tilted T-QWRs (i.e., [110] direction) and the x axis is perpendicular to the tilted T-QWRs in the (100) plane (i.e., $[1\bar{1}0]$ direction). In the effective-mass approximation, the envelope wave function

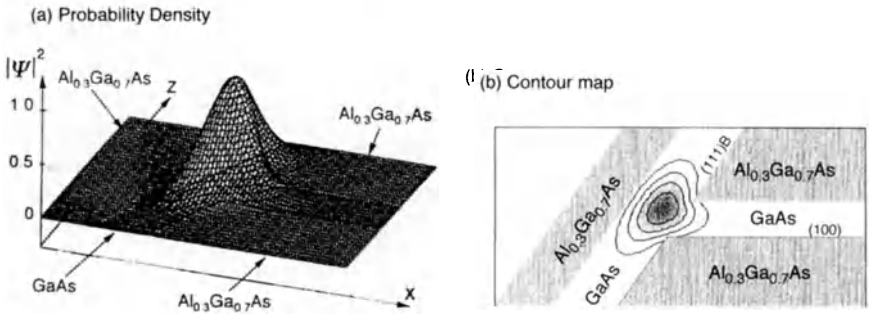


Fig. 6.5.4 Calculated probability density $|\Psi|^2$ (**a**) and contour map of $|\Psi|^2$ (**b**) for electrons confined in the tilted T-QWRs. After [7]

of conduction electrons is of the form $e^{iky}\Psi(x,z)$, where $\Psi(x,z)$ satisfies the 2D Schrödinger equation given by

$$-\frac{\hbar^2}{2m_e^*}\left(\frac{\partial^2\Psi}{\partial x^2}+\frac{\partial^2\Psi}{\partial z^2}\right)+U_c\Psi=E_e\,\Psi, \tag{6.5.1}$$

where E_e is an eigenenergy of an electron confined in the tilted T-QWR. In the GaAs and $Al_{0.3}Ga_{0.7}As$ regions, the effective mass becomes $m_e^*=0.067m_0$ and $m_e^*=0.071m_0$, respectively, where m_0 is the free electron mass. The conduction-band offset of 225 meV is used for $U_c(x,z)$. The Schrödinger equation for holes is similarly given by

$$-\frac{\hbar^2}{2m_h^*}\left(\frac{\partial^2\Psi}{\partial x^2}+\frac{\partial^2\Psi}{\partial z^2}\right)+U_v\Psi=E_h\Psi. \tag{6.5.2}$$

The valence band offset of 131 meV is used for $U_v(x,z)$. The common hole mass of $m_h^*=0.45m_0$ is used for GaAs and $Al_{0.3}Ga_{0.7}As$ regions. The electron-hole Coulomb interaction is neglected.

Figure 6.5.4a shows $|\Psi|^2$ for electrons confined in the tilted T-QWR, and a contour map of $|\Psi|^2$ is given in Fig. 6.5.4b [7]. The contours are lines of constant probability ($|\Psi|^2$=0.2, 0.4, 0.6, and 0.8). The top of calculated $|\Psi|^2$ shifts slightly from the cross point of the center of the QW on the (100) plane and that of QW on the (111) B facet due to the effectively thinner width of the upper $Al_{0.3}Ga_{0.7}As$ barrier on the (100) QW layer near the T-shaped region. In this calculation, a red-shift of 22 meV for the recombination energy of an electron and a hole is obtained in the tilted T-QWRs from the corresponding energy of MQW. The observed red shift is about 31 meV. The difference (9 meV) between the observed and calculated red shifts is considered to be an excess part of the exciton binding energy for the one-dimensional structure of tilted T-QWR from that of 2D QW. Taking into account an exciton binding energy of 11 meV expected for the 6.6 nm-wide MQW, we obtain an exciton binding energy of 20 meV for tilted T-QWRs, which agrees well with the value (17 meV) obtained for a GaAs/$Al_{0.3}Ga_{0.7}As$ T-QWR by Wegscheider et al. [2].

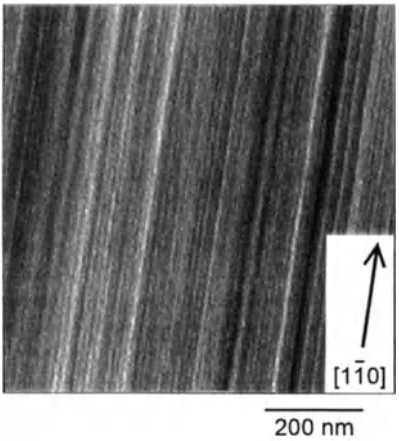


Fig. 6.5.5 AFM image of a GaAs surface of a GaAs(5 nm)/AlAs(5 nm) superlattice grown on (775)B GaAs substrate by MBE. After [9]

Recently, twenty GaAs/$Al_{0.3}Ga_{0.7}As$ tilted T-QWRs [8] were uniformly fabricated on a (111) B facet by improving growth conditions of GA-MBE to further reduce interfacet-migration of Ga and Al atoms from the (100) to the (111) B facet. A layer structure of the QWRs consists of a GaAs/$Al_{0.3}Ga_{0.7}As$ MQW layer with twenty GaAs wells ($L_w = 6.1$ nm) grown on reverse-mesa stripes on a (100) GaAs substrate by GA-MBE and a GaAs/$Al_{0.3}Ga_{0.7}As$ SQW ($L_w = 6.3$ nm) overgrown on the (111) B facet. The FWHM of a photoluminescence (PL) peak ($\lambda = 792$ nm) from the tilted T-QWRs is as small as 8.7 meV at 28K [8] which is comparable to those of conventional T-QWRs fabricated by the CEO method [2–4]. These results suggest that high-quality tilted T-QWRs can be uniformly fabricated on a large surface of (100) GaAs substrate by the two-step growth of GA-MBE/MBE.

6.5.4 Naturally Formed QWRs on (775)B GaAs Substrates

Naturally formed GaAs/AlGaAs QWRs on high index planes of GaAs substrates by MBE and MOCVD have attracted much interest [10–13], because of (1) high density of QWRs, (2) high crystal quality (no fabrication damage), and (3) high reproducibility (easy fabrication process). The uniformity of these QWRs, however, is still rather poor. The FWHM of a PL peak from these QWRs is still large (25 meV at 20 K [10], 34 meV at 4 K [11], and 20–40 meV at 2 K [12]). Recently, very uniform GaAs/$(GaAs)_m(AlAs)_n$ QWRs have been naturally formed on (775)B-oriented GaAs substrates by MBE [9–15].

a. MBE Growth of GaAs/$(GaAs)_m(AlAs)_n$ QWRs

The (775)B plane is of 8.5° off from the (111) B plane toward the (110) plane. A thin GaAs film grown on the (775)B GaAs substrate shows a regular surface corrugation. Figure 6.5.5 shows an AFM image of a GaAs surface of

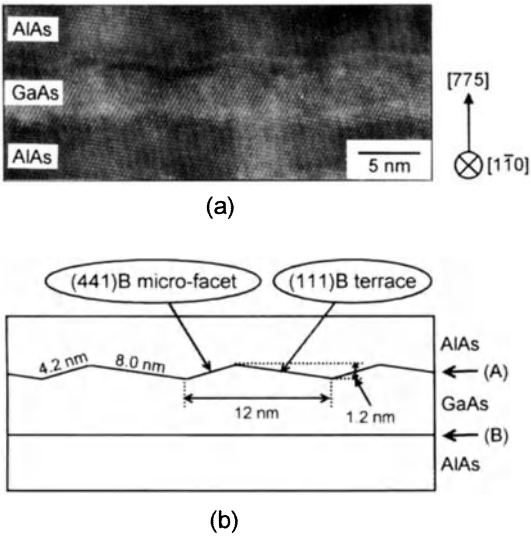


Fig. 6.5.6 (a) TEM micrograph of the (110) cross section of GaAs(5 nm)/AlAs(5 nm) superlattice grown on (775)B GaAs substrate (a) and schematic illustration of a GaAs/AlAs QW (b) with a corrugated AlAs/GaAs interface, and a flat GaAs/AlAs interface. After [9]

the GaAs/AlAs superlattice grown on the (775)B GaAs substrate by MBE at T_s=670°C and V/III=10 [9]. Regular corrugation with extremely straight step edges along the [1$\bar{1}$0] direction can be seen. The lateral period of the corrugation is about 12 nm. This corrugation of the GaAs layer is formed on the (775)B surface grown above 640°C.

Figure 6.5.6 shows a transmission electron microscopy image of the ($\bar{1}$10) cross section and its schematic illustration [9]. A regular corrugation can be seen on an AlAs/GaAs interface with a period of 12 nm (interface (A) in Fig. 6.5.6b) in contrast to a rather flat GaAs/AlAs interface (interface (B) in Fig. 6.5.6b) [9]. The corrugated interface consists of a (111) B terrace (8 nm) and a (441) B microfacet (4.2 nm). The lateral period and vertical amplitude of the corrugation are found to be 12 nm and 1.2 nm, respectively. The regularly corrugated AlAs/GaAs interface and the flat GaAs/AlAs interface result in regular modulation of the well thickness of AlAs/GaAs/AlAs QW grown on the (775)B GaAs substrate, and extremely high-density QWRs can be expected to be formed in the QW when the well thickness is significantly reduced.

GaAs/$(GaAs)_2(AlAs)_2$ QWRs [14], instead of GaAs/AlAs QWRs, were grown on the (775)B GaAs substrate, because the ground energy level of electrons in a GaAs/AlAs QW approaches the X-band edge of the AlAs barrier layers when the well thickness decreases to about 3 nm or less. Figure 6.5.7 shows a schematic illustration [14]. It is fabricated as follows. First, a 25-period (30 nm thick) $(GaAs)_2(AlAs)_2$ short-period superlattice (SPS) barrier

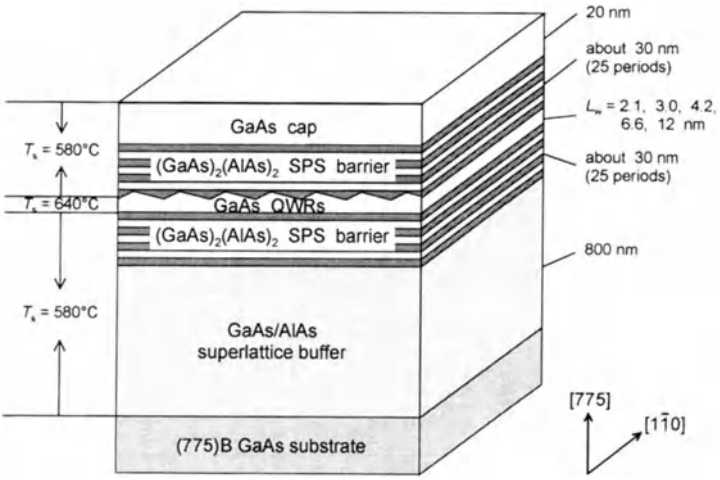


Fig. 6.5.7 Schematic illustration of GaAs/(GaAs)$_2$(AlAs)$_2$ QWRs grown on (775) B GaAs substrate. After [14]

layer is grown on a 800 nm thick GaAs/AlAs buffer layer at $T_s = 580°C$. The surfaces of GaAs and AlAs grown on the (775)B GaAs substrate at $T_s = 580°C$ are flat. Second, the substrate temperature is raised to 640°C and a GaAs QW (QWRs) layer is grown on the structure at $T_s = 640°C$. Regular corrugations are formed on the GaAs surface. The lateral period and vertical amplitude of the corrugations are 12 nm and 1.2 nm. Third, a (GaAs)$_2$(AlAs)$_2$ SPS barrier is grown after reducing the substrate temperature from 640°C to 580°C. The surface turns to become flat again. After repeating the second and the third processes with changing QW layer thickness ($L_w = 2.1$, 3.0, 4.2, 6.6, and 12 nm), a 20 nm thick GaAs cap layer is grown finally.

b. Photoluminescence Measurements

Figure 6.5.8 shows PL spectra at 14 K from GaAs/(GaAs)$_2$(AlAs)$_2$ QWs (QWRs) on the (775)B substrate and GaAs/(GaAs)$_2$(AlAs)$_2$ QWs on the (100) substrate for two different polarization directions [14]. For the sample on the (775)B substrate, a solid line shows the PL spectrum from the QWs for the polarization parallel to QWRs ([1$\bar{1}$0] direction, $I_{\parallel}$), and a dotted line shows that for the polarization perpendicular to QWRs ($I_{\perp}$). PL peaks at 670, 707, 741, 776, and 804 nm are from QWRs with average L_w of 2.1, 3.0, 4.2, 6.6, and 12 nm on the (775)B substrate, respectively. For the reference QWs grown on the (100) substrate, a solid line shows a PL spectrum for the polarization parallel to the cleaved surface ([011] direction, $I_{\parallel}$) and a dotted line shows that for the polarization perpendicular to the cleaved surface ([0$\bar{1}$1] direction, $I_{\perp}$).

The PL intensity of (775)B GaAs/(GaAs)$_2$(AlAs)$_2$ QWRs is almost the same to that of the QW simultaneously grown on the (100) GaAs substrate, indicating high optical quality of the (775)B QWRs. The polarization degrees

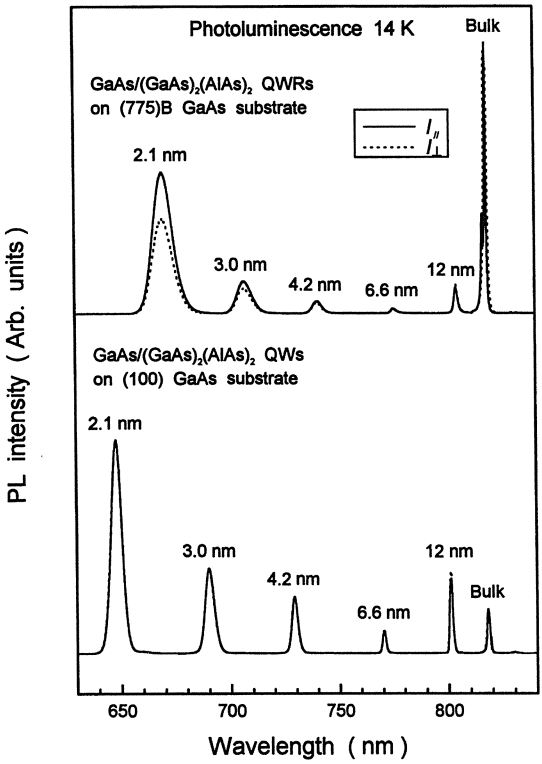


Fig. 6.5.8 Polarized PL spectra (14 K) from $GaAs/(GaAs)_2(AlAs)_2$ QWs (including QWRs) grown on (775)B GaAs substrate and conventional $GaAs/(GaAs)_2(AlAs)_2$ QWs grown on (100) GaAs substrate. After [14]

$P = (I_{\parallel} - I_{\perp})/(I_{\parallel} + I_{\perp})$ of PL peaks are plotted as a function of L_w in Fig. 6.5.9 [14]. With decreasing L_w from 12 nm to 2.1 nm, the polarization degree of PL from QWs (QWRs) on the (775)B substrate increases monotonicly. The PL peak from the QWRs with the average L_w of 2.1 nm on the (775)B substrate exhibits the strongest polarization degree ($P = 0.21$). On the other hand, the polarization degree is constantly almost zero for QWs on the (100) substrate. The FWHM of the PL peak with an average L_w of 2.1 nm is 30 meV, which is almost comparable with those of other self-organized QWRs [10–12]. These results are in good agreement with theoretical results that PL from a QWR is polarized parallel to the wire direction [16]. These indicate that high-density $GaAs/(GaAs)_2(AlAs)_2$ QWRs with a considerably strong one-dimensional confinement of carriers are realized on the (775)B substrate by MBE.

Very recently, the uniformity of $GaAs/(GaAs)_4(AlAs)_2$ QWRs grown on (775)B GaAs substrate have been much improved by optimizing the MBE growth process [15]. Strong polarization dependence and very narrow PL line width are obtained.

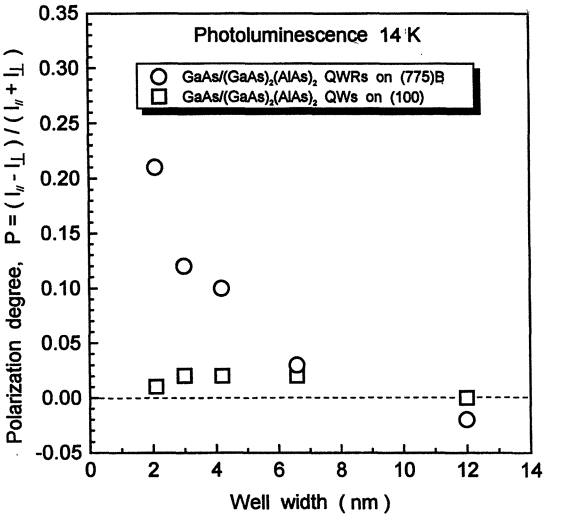


Fig. 6.5.9 Polarization degree, $P=(I_{\parallel}-I_{\perp})/(I_{\parallel}+I_{\perp})$, for GaAs/$(GaAs)_2(AlAs)_2$ QWs (QWRs) grown on (775)B GaAs substrate and GaAs/$(GaAs)_2(AlAs)_2$ QWs grown on (100) GaAs substrate as a function of average well width L_w. After [14]

References

1 L.N. Pfeiffer, K.W. West, H.L. Stormer, J.P. Eisenstein, K. W. Baldwin, D. Gershoni, and J. Spector, Appl. Phys. Lett. **56**, 1697 (1990)
2 W. Wegscheider, L.N. Pferffer, M.M. Dignam, A. Pinczuk, K.W. West, S.L. McCall, and R. Hull, Phys. Rev. Lett. **71**, 4071 (1993)
3 T. Someya, H. Akiyama, and H. Sakaki, Phys. Rev. Lett. **74**, 3664 (1995)
4 T. Someya, H. Akiyama, and H. Sakaki, Appl. Phys. Lett. **66**, 3672 (1995)
5 S. Shimomura, K. Inoue, M. Tanaka, N. Tomita, A. Adachi, M. Fujii, T. Yamamoto, T. Watanabe, N. Sano, K. Murase, and S. Hiyamizu, Solid-State Electron. **37**, 597 (1994)
6 N. Tomita, M. Tanaka, T. Saeki, K. Fujita, T. Watanabe, T. Higuchi, N. Sano, A. Adachi, S. Shimomura, and S. Hiyamizu, J. Vac. Sci. Technol. B **14**, 3550 (1996)
7 N. Tomita, T. Kishi, K. Takekawa, K. Fujita, T. Watanabe, A. Adachi, S. Shimomura, and S. Hiyamizu, to be published in J. Cryst. Growth (1997)
8 N. Tomita, K. Takekawa, K. Ohta, K. Fujita, N. Egami, Y. Okamoto, S. Shimomura, and S. Hiyamizu, J. Vac. Sci. Technol. B (submitted for publication)
9 M. Higashiwaki, M. Yamamoto, T. Higuchi, S. Shimomura, A. Adachi, Y. Okamoto, N, Sano, and S. Hiyamizu, Jpn. J. Appl. Phys. **35**, 606 (1996)
10 S. Hara, J. Motohisa, T. Fukui, and H. Hasegawa, Jpn. J. Appl. Phys. **34**, 4401 (1995)
11 M. Takeuchi, K. Shiba, K. Sato, H. K. Huang, K. Inoue, and H. Nakashima, Jpn. J. Appl. Phys. **34**, 4411 (1995)
12 M. Takeuchi, T. Takeuchi, Y. Inoue, T. Kato, K. Inoue, H. Nakashima, K. Maehashi, P. Fischer, J. Christen, M. Grundmann, and D. Bimberg, Superlattices & Microstructures **22**, 43 (1997)
13 R. Nötzel, N. N. Ledentsov, L. Däweritz, K. Ploog, and M. Hohenstein, Phys. Rev. B **45**, 3507 (1992)
14 M. Higashiwaki, M. Yamamoto, S. Shimomura, A. Adachi, and S. Hiyamizu, to be published in J. Cryst. Growth (1997)
15 M. Higashiwaki, M. Yamamoto, S. Shimomura, S. Hiyamizu, unpublished
16 C. R. McIntyre, and L. J. Sham, Phys. Rev. B **45**, 9443 (1995)

6.6 SiGe Quantum Structures

6.6.1 Band Modification by SiGe/Si Heterostructures

Si-based heterostructures are being intensively studied since various new functionalities which cannot be obtained with Si alone are expected owing to the predictable and controllable modification of the band structure [1]. In particular, SiGe/Si heterostructures are of great importance and have had significant impact on Si-based device applications as well as on the basic science of semiconductors. The built-in strain is not an obstacle but is now regarded as an advantage since it can give additional flexibility in the band engineering.

The unstrained $Si_{1-x}Ge_x$ alloys are formed in the whole Ge composition range and the band gap decreases monotonicly with increasing Ge composition. The band structure of the alloy is Si-like for $0 < x < 0.85$ and the conduction band minimum is six-fold degenerate at a point along the [100] direction (Δ) in the reciprocal space. On the other hand, Ge-like band structure is realized for $x > 0.85$ and the four-fold degenerate L along the [111] direction is the bottom of the conduction band. The crossover was clearly observed by photoluminescence (PL) spectroscopy [2] and the unstrained excitonic band gap of the $Si_{1-x}Ge_x$ alloy was obtained as

$$E_g^{(\Delta)}(x) = (1.155 - 0.43x + 0.206x^2)\ \text{eV} \quad (0<x<0.85), \tag{6.6.1}$$

$$E_g^{(L)}(x) = (2.017 - 1.27x)\ \text{eV} \quad (0.85<x<1). \tag{6.6.2}$$

Pseudomorphic epitaxial growth of $Si_{1-x}Ge_x$ alloys on Si (100) results in biaxial in-plane compressive strain. The strain can be divided into two components, hydrostatic and uniaxial, which respectively induce the shift of the average position and the splitting of the bands, and thus alter the band gap. The fundamental indirect band gap of strained SiGe is smaller than that of the unstrained bulk in the whole Ge composition. The excitonic band gap of strained SiGe on Si was derived with slight uncertainty by PL measurements [3], and the following equation has been obtained by the least squares fit

$$E_g(x) = (1.17 - 0.896x + 0.396x^2)\ \text{eV}. \tag{6.6.3}$$

This result is in good agreement with the former theoretical calculations [4]. It is important to note that the four-fold degenerate $\Delta(4)$ is the minimum of the conduction bands of pseudomorphically strained SiGe on Si in the whole alloy range, even in "pure" Ge.

The band lineup problem, which is of great importance, is closely coupled with the strain effect. In other words, various band lineups can be realized in SiGe/Si heterostructures by controlling the states of the strain. The most powerful tool to predict the band lineup would be the model-solid theory, developed by Van de Walle [5]. In the theory, the band offset is regarded as a difference between reference values determined for a "model-solid" and the band lineup of various heterostructures can be successfully predicted by using deformation potentials which describe the strain effect on the band modification.

T. Ando et al. (eds.), *Mesoscopic Physics and Electronics*
© Springer-Verlag Berlin Heidelberg 1998

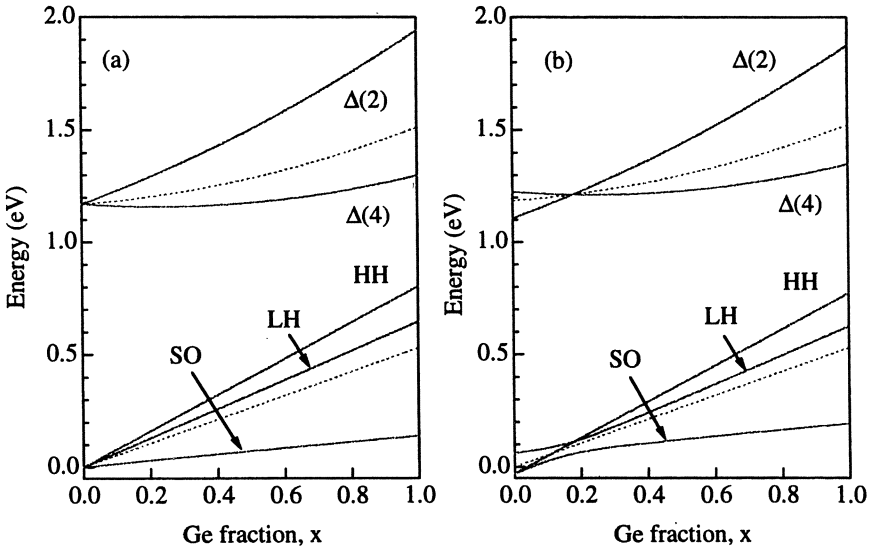


Fig. 6.6.1 Calculated valence and conduction bands in pseudomorphically strained SiGe alloys grown on (**a**) Si (100) and (**b**) relaxed $Si_{0.82}Ge_{0.18}$ (100) substrates. The averages of the valence and the conduction bands are plotted by dotted lines. HH, LH, and SO represent heavy hole, light hole, and split-off hole, respectively

Figure 6.6.1 shows calculated valence and conduction bands in pseudomorphically strained $Si_{1-x}Ge_x$ alloys grown on Si (100) and relaxed $Si_{0.82}Ge_{0.18}$ (100). It can be seen that a compressively strained $Si_{1-x}Ge_x$/Si heterointerface ($x \lesssim 0.4$) shows a type I character, i.e., the narrower SiGe band gap falls within the wider Si band gap where most of the band offset is consumed at the valence band side. The amount of the offset is sufficient to observe modulation doping effects for holes but not for electrons. On the other hand, tension-strained $Si_{1-x}Ge_x$/relaxed $Si_{1-y}Ge_y$ ($x < y$) constructs a type II heterointerface and confining electrons in a tension-strained layer is possible. Extraordinarily high electron mobility has been obtained in n-type modulation doped strained Si/relaxed SiGe heterostructures [6]. Rigorous knowledge of the band gap of the strained SiGe alloys and the strain-dependent band alignment, when combined with advanced growth techniques, allows us to obtain a number of interesting electrical and optical properties of strained SiGe/Si heterostructures. The following introduces recent experimental observations focusing on optical properties of SiGe/Si quantum structures.

6.6.2 SiGe Quantum Wells

Si has not been recognized as a promising candidate for light-emitting materials for a long time because of its poor radiative recombination probability due to the inherently indirect nature of the band structure. However, the pioneering works of PL from strained SiGe/Si quantum wells (QWs) [7–9] and superlattices [10] clearly demonstrated their potential as light-emitting materials and stimulated a lot of researchers.

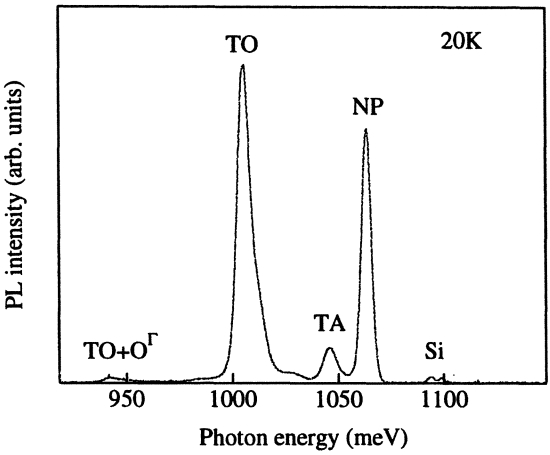


Fig. 6.6.2 PL spectrum of 20 periods of strained $Si_{0.82}Ge_{0.18}/Si$ multiple QWs grown by gas-source MBE measured at 20 K. In addition to PL peaks from the Si substrate at around 1097 meV, well-resolved PL from SiGe can be seen clearly

Figure 6.6.2 shows a typical PL spectrum of strained $Si_{0.82}Ge_{0.18}/Si$ QWs with well width of 34 Å grown by gas-source MBE. The spectrum contains a no-phonon (NP) transition, followed by momentum-conserving phonon replicas involving transverse-acoustic (TA) and three transverse-optical (TO) phonons. The strained SiGe layer provides an effective trap for holes owing to the comparatively large valence band offset, thus attracts also electrons through the Coulomb interaction in spite of the nearly zero conduction band offset. The absence of the conduction band offset is not a fatal problem as evidenced by room-temperature electroluminescence observations of SiGe/Si multiple QWs by several groups [11,12]. However, confining electrons with high potential barriers would be more favorable from a practical point of view.

An attempt to confine electrons in tension-strained Si QWs embedded between relaxed $Si_{0.82}Ge_{0.18}$ has been carried out [13]. A quantum confinement effect was clearly observed by a systematic PL peak shift with Si well width variation; however, the PL intensity is limited to be low since the transition is spatially-indirect across the Si/SiGe heterointerface and no special structures for hole confinement are considered.

In order to realize effective confinement of both electrons and holes, a new class of semiconductor heterostructures, neighboring confinement structure (NCS) [14], has been applied to SiGe/Si systems. NCS consists of a pair of type II neighboring QWs sandwiched by barrier layers, and electrons and holes are separately confined in the pair as shown in Fig. 6.6.3a. In spite of the indirect band structure in real- and k-spaces, radiative recombination is expected to be enhanced due to the penetration of wavefunctions by optimizing mutual layer widths. NCS can be realized in SiGe/Si systems by cladding a pair consisting of a tension-strained Si layer and a compressively strained $Si_{1-y}Ge_y$ layer between relaxed $Si_{1-x}Ge_x$ ($y > x$) layers. The strained Si and

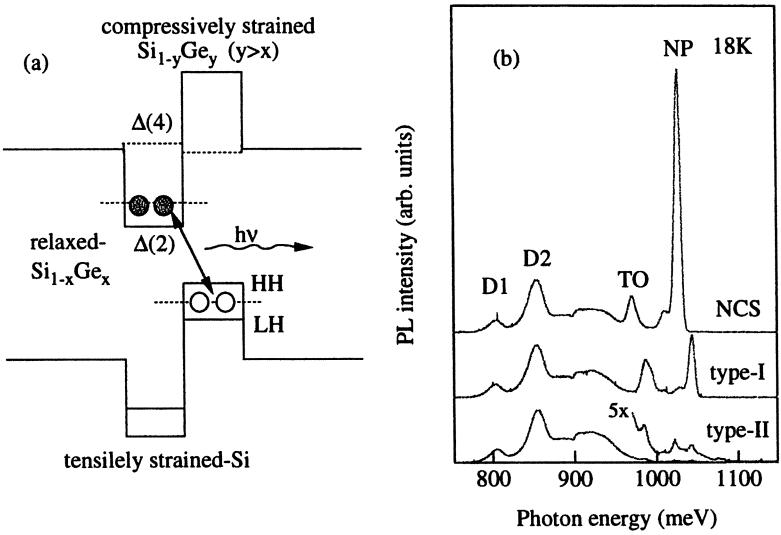


Fig. 6.6.3 (a) Schematic band alignment of NCS and (b) comparison of 18 K PL spectra from three different samples, type II 10 Å strained Si QWs, type I 10 Å strained $Si_{0.64}Ge_{0.36}$ QWs and 10 Å strained Si/10 Å strained $Si_{0.64}Ge_{0.36}$ NCS grown on relaxed $Si_{0.82}Ge_{0.18}$. Remarkable enhancement of integrated PL intensity and dramatic increase of relative NP intensity compared to TO can be seen in NCS

strained $Si_{1-y}Ge_y$ layers act as QWs for electrons and holes, respectively.

Figure 6.6.3b shows comparison of 18 K PL spectra from three different samples, 10 Å strained Si QWs, 10 Å strained $Si_{0.64}Ge_{0.36}$ QWs, and 10 Å strained Si/10 Å strained $Si_{0.64}Ge_{0.36}$ (NCS) grown on relaxed $Si_{0.82}Ge_{0.18}$. Note that the first sample has type II band alignment where Δ(2)-electrons are confined in the Si well, the second type I, with nearly zero conduction band offset. Remarkable enhancement of integrated PL intensity can be seen in NCS due to the improved spatial overlap of wavefunctions of electrons and holes. In addition, spectral feature is drastically changed and NP intensity is much stronger than phonon replicas.

It is worth mentioning that NCS can be also realized without using thick relaxed SiGe buffer layers. Strained SiC alloys on Si can be regarded as tension-strained Si if the amount of carbon is very small. Therefore, the direct growth of a SiC/SiGe pair on Si constructs NCS. In fact, similar PL spectra with enhanced NP intensity are observed from a pair of $Si_{0.988}C_{0.012}/Si_{0.84}Ge_{0.16}$ on Si [15]. The enhanced NP feature is very similar to PL from localized excitons in strain-symmetrized Si/Ge superlattices on relaxed SiGe [10]. In superlattices, localization of excitons is likely to occur at low temperatures since the inevitable monolayer fluctuation gives an inhomogeneous broadening of the ground state energy comparable to (or even larger than) the miniband width. However, the NP line rapidly decreases due to the delocalization of excitons above 20K. The origin of the NP enhancement in NCS was found to be also exciton localization due to in-plane energy variation owing to the interface

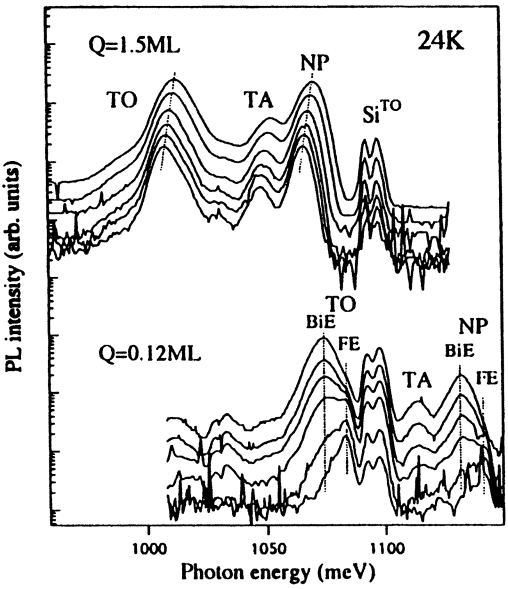


Fig. 6.6.4 PL spectra of the samples with $Q = 0.12$ and 1.5 ML with various excitation powers. The energy shift due to the band filling effect is invisible in Q=0.12 ML and a new peak (BiE) appears on the lower energy side of free-exciton (FE) with increasing excitation

roughness as evidenced by time-resolved PL [16]. In contrast to superlattices, however, the enhanced NP feature persists up to above 100 K due to the formation of rigorous localization potential. Therefore, NCS is believed to be a promising quantum structure to improve PL from indirect semiconductors.

6.6.3 SiGe Quantum Wires and Dots

SiGe/Si quantum wires (QWRs) were first fabricated by direct epitaxial growth of SiGe/Si QWs on V-groove patterned Si substrates by gas-source MBE [17]. Crescent-shaped SiGe layers are formed at the bottom of the V-groove as evidenced by transmission electron microscope (TEM). Three pairs of NP and TO are observed in the PL spectrum, and they are assigned as originating from (111) facet QWs, QWRs at the bottom of the V-groove, and planar (100) QWs between the neighboring grooves from the higher energy side, respectively. The spectral blue-shift of QWRs compared to control QWs is mainly due to the inhomogeneous strain in the V-groove rather than the lateral confinement effect. The relative intensity of the three pairs depends strongly on the temperature and the V-groove period. This is due to the redistribution of long-lived excitons through diffusion [18]. In order to collect excitons in QWRs, selective epitaxial growth (SEG) is exploited, which enables one to grow epitaxial layers only inside the V-groove. The cross-sectional luminescence from QWRs obtained by SEG shows no preferential polarization, in contrast to QWs, which have TE polarization [19].

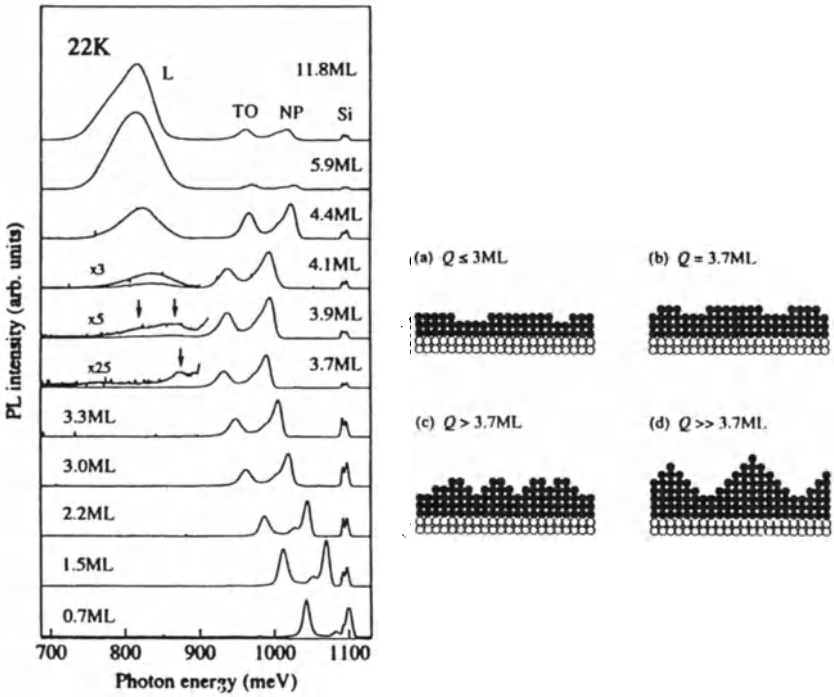


Fig. 6.6.5 PL spectra of Si/pure Ge/Si QWs with various Ge coverage and schematic diagram of Ge growth on Si. Two-dimensional growth persists up to the critical thickness for the three-dimensional growth. Then Ge islands grow, leaving a 3.0 ML wetting layer

Self-assembled QWRs are obtained by growing a submonolayer of Ge on Si(100) substrates with surface atomic steps [20]. Since Ge has a lower surface energy than Si, Ge atoms are preferentially grown at the step-edges of the Si substrate, leading to the formation of QWRs along the step-edge. Figure 6.6.4 compares the excitation power dependence of PL spectra of Si/pure Ge/Si with Ge coverage (Q) of 1.5 monolayers (ML) and 0.12 ML. The band-filling effect, which is normally observed in 2D SiGe/Si QWs, is observed for the $Q = 1.5$ ML sample but not for the $Q = 0.12$ ML sample. This implies realization of a bunched density of states at the band-edge for the latter sample reflecting the 1D QWR formation. The emergence of a new luminescence peak, which may come from bi-excitons in QWRs, is observed for the $Q = 0.12$ ML sample under high excitations. This is a consequence of an enhanced exciton population in QWRs with a small cross-section. The formation of bi-excitons in QWRs is also supported by the almost doubled decay rates of free excitons [21].

The Stranski–Krastanow (SK) growth mode, in which 2D layer-by-layer growth is followed by 3D islanded growth, is regarded as a promising method to fabricate quantum dots (QDs) without any processing. The size of islands can be smaller than the limitation of conventional lithography under appro-

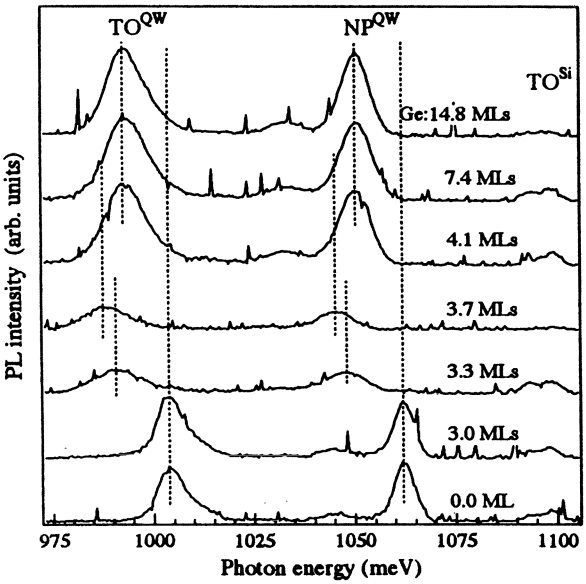


Fig. 6.6.6 PL spectra of $Si_{0.82}Ge_{0.18}$/Si QWs with various coverage of surface Ge. Ge islands generate an inhomogeneous strain field for buried layers and modulate the potential of buried SiGe/Si QWs

priate growth conditions. Figure 6.6.5 shows a series of PL spectra of Si/pure Ge/Si heterostructures with various Q grown by gas-source MBE at 700°C [22]. At $Q < 3.7$ ML, the spectral feature is almost the same as that of strained SiGe/Si QWs and peaks labeled NP and TO are assigned as originating from pure Ge/Si 2D QWs. The PL energy shifts to lower energies with increasing Q reflecting the quantum confinement effect. However, the further increase of Q brings anomalies in PL spectra. The shift of PL peaks from 2D QWs is suddenly quenched and a spectral blue shift occurs followed by a saturation of energy. In addition, broad PL labeled L appears at lower energies which grows in intensity with increasing Q. This phenomenon comes from a 2D-3D growth mode changeover as evidenced by cross-sectional TEM. The kinetic critical thickness of the 2D growth was determined as 3.7 ML, and Ge islands were found to grow, leaving a 2D wetting layer of 3.0 ML. Interestingly, the weak emissions indicated by arrows for $Q = 3.7$ ML transform into strong L bands by introducing growth interruption (GI) for 1 min after the growth of the Ge layer. On the other hand, GI gives no change for PL spectra of $Q \leq 3.0$ ML. This means that the equilibrium critical thickness of Ge on Si is 3.0 ML. It is noted that the Ge islands formed by the SK mode act as deep traps for holes and room-temperature PL can be observed [23].

It should be remarked that the formation of Ge islands is accompanied by partial strain relaxation and generates an inhomogeneous strain field which modulates the potential of buried layers. This phenomenon is also usable to obtain low-dimensional quantum structures. Figure 6.6.6 shows PL spectra of

strained $Si_{0.82}Ge_{0.18}/Si$ QWs with surface Ge islands as a function of Q [24]. It can be seen that the drastic spectral shift to lower energies corresponds to the 2D-3D growth mode changeover at Q of 3.3–3.7 ML, and further increase of Q results in a decrease of the red shift. This evidences strain-induced lateral potential modulation by the surface Ge islands. Namely, QD-like structures are formed by the Ge islands. Detailed analysis of this anomalous behavior would be helpful to clarify the mechanism of island formation and strain relaxation. Also, strain-induced QDs would be useful to investigate such fundamental physics in QDs as carrier relaxation mechanisms.

References

1 *Properties of strained and relaxed Silicon Germanium*, edited by E. Kasper (Institution of Electrical Engineers, London, 1995)
2 J. Weber and M.I. Alonso, Phys. Rev. B **40**, 5683 (1989)
3 D.J. Robbins, L.T. Canham, S.J. Barnett, A.D. Pitt, and P. Calcott, J. Appl. Phys. **71**, 1407 (1992)
4 R. People, Phys. Rev. B **32**, 1405 (1985)
5 C.G. Van de Walle, Phys. Rev. B **39**, 1871 (1989)
6 K. Ismail, M. Arafa, K.L. Saenger, J.O. Chu, and B.S. Meyerson, Appl. Phys. Lett. **66**, 1077 (1995)
7 J.C. Sturm, H. Monoharan, L.C. Lenchyshyn, M.L.W. Thewalt, N.L. Rowell, J.-P. Nöl, and D.C. Houghton, Phys. Rev. Lett. **66**, 1362 (1991)
8 J. Brunner, U. Menczigar, M. Gail, E. Friess, and G. Abstreiter, Thin Solid Films **222**, 27 (1992)
9 S. Fukatsu, Solid-State Electron. **37**, 817 (1993) and references therein
10 U. Menczigar, G. Abstreiter, S. Olajos, H. Grimmeiss, H. Kibbel, H. Presting, and E. Kasper, Phys. Rev. B **47**, 4099 (1993)
11 Q. Mi, X. Xiao, J.C. Sturm, L.C. Lenchyshyn, and M.L.W. Thewalt, Appl. Phys. Lett. **60**, 3177 (1992)
12 S. Fukatsu, N. Usami, Y. Shiraki, A. Nishida, and K. Nakagawa, Appl. Phys. Lett. **63**, 967 (1993)
13 D.K. Nayak, N. Usami, S. Fukatsu, and Y. Shiraki, Appl. Phys. Lett. **63**, 3509 (1993)
14 N. Usami, F. Issiki, D.K. Nayak, Y. Shiraki, and S. Fukatsu, Appl. Phys. Lett. **67**, 524 (1995)
15 K. Brunner, W. Winter, and K. Eberl, Appl. Phys. Lett. **69**, 1279 (1996)
16 N. Usami, Y. Shiraki, and S. Fukatsu, Appl. Phys. Lett. **68**, 2340 (1996)
17 N. Usami, T. Mine, S. Fukatsu, and Y. Shiraki, Appl. Phys. Lett. **63**, 2789 (1993)
18 N. Usami, H. Akiyama, S. Fukatsu, and Y. Shiraki, Phys. Rev. B **52**, 5132 (1995)
19 N. Usami, T. Mine, S. Fukatsu, and Y. Shiraki, Appl. Phys. Lett. **64**, 1126 (1994)
20 H. Sunamura, N. Usami, Y. Shiraki, and S. Fukatsu, Appl. Phys. Lett. **68**, 1847 (1996)
21 H. Sunamura, Y. Shiraki, and S. Fukatsu, Solid-State Electron. **40**, 693 (1996)
22 H. Sunamura, N. Usami, Y. Shiraki, and S. Fukatsu, Appl. Phys. Lett. **66**, 3024 (1995)
23 H. Sunamura, N. Usami, Y. Shiraki, and S. Fukatsu, J. Cryst. Growth **157** (1996)
24 E.S. Kim, N. Usami, and Y. Shiraki, Appl. Phys. Lett. **70**, 295 (1997)

Index

AAS oscillation 16, 80, 87
AB effect 15
AB oscillation 15
AB type oscillation 76, 81, 86
acceleration equation 35
addition energy 67
AFM 228, 243, 247
Ag 3
Aharonov–Bohm effect 1, 15
Aharonov–Bohm flux 110
Aharonov–Bohm oscillation 16
Aharonov–Bohm type oscillation 76
Al 43
AlAs 105, 176, 183, 199, 228, 249
AlAs/GaAs 22
AlGaAs 104, 184, 228, 238, 247
AlGaAs/GaAs 99
AlGaInAs/InP 179
AlGaInP 176
AlGaInP/GaAs 177
Altshuler–Aronov–Spivak oscillation 16, 72, 80
analog computation 222
Anderson localization 5, 8, 87, 109, 146
anisotropic etching 216
antidot disorder 80, 87
antidot lattice 8, 72
antidot potential 72
aperture function 205
artificial atom 1, 66
aspect ratio 73
atomic-force-microscopy 228
Au 15

backward scattering 56
ballistic motion 4, 27
ballistic regime 5
ballistic transport 1, 6, 28
band offset 190
BDD 220
bend resistance 27
Bethe ansatz 45
bi-exciton 268
binary decision diagram 220
blackbody emission 99
blackbody radiation 101
Bloch condition 83
Bloch frequency 199
Bloch oscillation 35, 102, 104, 199
Bloch's theorem 83, 140
Bohr radius 4
Boltzmann constant 5, 31
Boltzmann distribution 64
Boltzmann transport equation 3
Boolean expression 220
Boolean operation 220
Born approximation 46, 56, 58
bosonization technique 25, 47
boundary roughness 54
Bragg reflection 1, 104, 187, 199
breakdown of quantum Hall effect 112
Brillouin zone 35, 104
bulk 4

bulk-current picture 109
butterfly oscillation 77
butterfly spectrum 8
Büttiker–Landauer formula 85, 111

CaF_2 188
cathodoluminescence 250, 257
cavity polariton 164
cavity quasimode 161
cavity resonance mode 161
center coordinate 111, 120
CEO 255
chaos 1, 8
chaotic orbit 72
charging effect 1
charging energy 4, 6, 31, 34, 37
chemical potential 11, 51
CL 250, 257
classical action 79
cleaved edge overgrowth 255
Co 188
coherence volume 149
coherent destruction of tunneling 200
collimation 30
commensurability condition 90
commensurability oscillation 72
commensurability peak 72, 86
complex unitary matrix 9
composite fermion 94
compressible and incompressible strips 132, 136
conductance fluctuation 1
conductance quantization 1, 13, 22
constriction 36
Cooper pair 15, 18, 43
$CoSi_2$ 188
$CoSi_2/CaF_2$ 188
Coulomb blockade 1, 31, 66, 187, 218
Coulomb blockade oscillation 37, 66, 218
Coulomb diamond 67
Coulomb gap 66
Coulomb interaction 1, 4, 6, 50, 58, 67
Coulomb oscillation 37, 66
coupled quantum well 195
critical exponent of localization 114
critical randomness 9
Cu 3
current probe 28
current source 32
current standard 33
current-voltage characteristics 105
cyclotron diameter 27
cyclotron energy 61, 69
cyclotron frequency 116
cyclotron motion 63
cyclotron orbit 72, 79, 90
cyclotron radius 28, 55, 90, 132
cylinder-shaped microcavity 164
cylindrical metal film 80

DBR 179, 183
DBRTD 207
de Broglie wavelength 195
deep mesa etching 6
defect mode 168
deformation potential 93
delocalized state 9
density-density correlation function 46
density of states 9, 56, 64, 77, 84, 133
dephasing process 16
device trend 213
diagonal conductivity 75
diagonal resistivity 72
diamagnetic shift 239

differential capacitance 133
differential conductance 66
differential resistance 32
diffraction 187
diffusion coefficient 4, 9, 18, 74, 115
diffusive motion 4
diffusive regime 5, 13, 80
disk-shaped microcavity 164
disordered contact 126
distributed Bragg reflector 179, 183
distributed circuit 135
DOS model 133
double-barrier resonant tunneling diode 207
double-barrier structure 66
double slit 205, 210
DRAM 213
drift velocity 91
Drude conductivity 101
Drude formula 49
dynamical interaction 5

EBL 210
EB lithography 216
edge channel 109
edge current 112, 120
edge-current picture 109
edge magnetoplasmon 122
edge state 45, 55, 132
effective channel number 13
effective chemical potential 12
effective mass 144, 195
effective-mass approximation 105, 141, 257
Einstein relation 9
electrochemical potential 66, 120
electroluminescence 265
electron affinity 67
electron beam lithography 210
electron directional coupler 196
electron-electron interaction 26, 37, 66, 98
electron-electron scattering 5, 206
electron heating 97
electron-hole pair 42
electron holography 15
electron-phonon interaction 61
electron-phonon scattering 5
electron temperature 99
electron wave 1
electron wavefront 204
electron waveguide 196
electroreflectance 104
elliptic coordinates 25
energy-dispersive x-ray spectroscopy 232, 249
ensemble average 5
envelope function 143
environmental effect 31
equi-energy line 83
equilibration length 126
ergodic hypothesis 17
evanescent mode 82
exchange effect 70
exchange interaction 67, 70
exciton 144
extended state 9, 109

Fabry–Perot interferometer 191
Fabry–Perot microcavity 161, 168
Fabry–Perot resonance 161
Fabry–Perot structure 183
Faraday configuration 97
far-infrared 96
FEL 201
Fermi energy 3, 5, 20, 22, 30, 38, 45, 54, 72, 121, 205, 206
Fermi level 24, 37, 58, 66, 90, 109, 111, 116, 133, 190
Fermi liquid 45
Fermi liquid theory 5, 47

Fermi surface 47, 90
Fermi velocity 3, 26
Fermi wavelength 3, 6, 8, 13, 36, 72, 76
Fermi wave number 3, 25, 54
Fermi wave vector 3, 90
field-effect device 193
field-effect tunneling transistor 188
filling fraction 170
fine-area photoluminescence spectrum 229
fine structure constant 109
FIR 96
FIR spectroscopy 96
first Brillouin zone 140
FIR transmission spectroscopy 96
fluctuations of orbital susceptibility 19
fluorescence 156
flux average 112
flux quantization 15, 18
flux quantum 15, 80, 88, 109
Fock–Darwin state 7, 68
Fourier transform 204
four-terminal geometry 111
four-wave mixing 105, 198
Fröhlich interaction 63
free-electron laser 201
Frenkel exciton 144
Fresnel–Fourier transform 187
Friedel oscillation 58
Friedel sum rule 39
fusion technique 174

GaAlAs 165, 176
GaAlAs/AlAs 165, 176
GaAs 22, 36, 50, 52, 66, 106, 150, 163, 174, 176, 183, 206, 228, 230, 238, 247
GaAs/AlAs 104, 228, 255
GaAs/AlGaAs 3, 61, 64, 72, 90, 97, 114, 132, 152, 171, 198, 201, 247, 255
GaInAs 176
GaInAs/AlInAs 61
GaInAs/GaAs 176, 199
GaInAs/InP 61, 207, 210
GaInAsP 176, 210
GaInAsP/InP 171, 174, 179
GaInNAs/GaAs 179
gain profile 181
galvanomagnetic effect 111
GaN 176
gas-source MBE 228, 265
Gauss's theorem 41
Ge 264
geometrical resonance 90
giant step 247
glancing-angle MBE 255
grating coupler 101
growth interruption 269
guiding center 91
guiding center coordinate 90

Hall conductivity 75
Hall field 109
Hall resistance 29
Hall resistivity 72
harmonic potential 6
Hartree–Fock calculation 67
Hartree–Fock theory 39
Hartree potential 96
hemispherical microcavity 164
HEMT 224
heterostructure 4
high electron mobility transistor 224
higher harmonics 32
high-speed device 213
Hofstadter's butterfly 8, 77
hot electron 101
hot-electron diffraction device 210

hot electron wave 205
hot plasmon 101
Hubbard model 45, 52
Hund's rule 67, 70
hysteresis 93

ideal lead 85
ideal reservoir 11
ideal wire 11
InAs 243
InAs/AlGaSb 64
incoming channel 13
inelastic scattering 5, 16, 115
inelastic scattering length 16
inelastic scattering time 115
InGaAs 66, 164, 183, 228, 243
InGaAs/GaAs 231
InGaAs quantum dot 183, 243
InP 172, 176, 228
in-plane gate 6
integer quantum Hall effect 9, 109, 120
interface roughness 106
interference 187
interference device 206
interference effect 14, 31, 80
interference switch 205
inverse Bloch oscillation 201
inverse localization length 8, 110
ion beam implantation 238
ion beam milling 238
IQHE 120

Josephson effect 34
junction network 40

Kirchhoff's diffraction theory 205
Kohn's theorem 7, 98
Kondo effect 39
Kondo temperature 39
Korringa relation 46
Kosteritz–Thouless–Berezinskii transition 42
Kramers degeneracy 10
KTB transition 42
Kubo formula 12, 62

Landau diamagnetism 17, 20
Landauer's formula 1, 11, 22, 45, 51, 111
Landau gauge 62, 90, 92
Landau level 61, 90, 112, 120, 132
Landau subband 90
laser interference pattern 90
lateral superlattice 8, 90
lattice vibration 5
linear response theory 112
local current 27
local current distribution 112
local electric field 27
localization in quantum Hall regime 110
localization length 8, 88, 117
localized state 109
LO phonon 61, 206
low-power device 213
LSI technology 187
Luttinger–Kohn Hamiltonian 151
Lyapunov exponent 79

macroscopic system 5
magic number 67
magnetic depopulation 96
magnetic-field modulation 93
magnetic flux 82
magnetic focusing 27, 74
magnetic length 82, 112, 120, 132
magnetic moment 39
magnetic unit cell 82
magnetic Weiss oscillation 93
magnetocapacitance 132
magneto-fingerprint 15
magnetophonon resonance 61

magnetoresistance 54, 64, 86, 90
Maslov index 79
MBE 188, 227, 247, 255
mean free path 3, 8, 13, 18, 27, 54, 80, 86, 90
mesoscopic physics 31
mesoscopic system 1
metal-insulator transition 8
metallic dot 6
metallic wire 6
metalorganic chemical vapor deposition 209, 227
metalorganic MBE 228
MgO/Si 176
Michaelson interferometer 96
microcavity 1, 139, 157
microcavity laser 157
micro-PL 240
microsphere 164
microwave 32, 36
miniband collapse 200
MOBILE 223
MOCVD 209, 227
MOCVD selective growth 238
model-solid theory 264
modulation bandwidth 181
molecular beam epitaxy 227
monostable-to-bistable transition logic element 223
MOSFET 4, 96, 213
MPR 61
MQB 178
MQW 255
multi-atomic step 228, 244, 247
multilayer filter 167
multiple reflection 187
multi-quantum barriers 178
multi-quantum well 255

NCS 266
NDC 218
NDR 190, 208
negative differential conductance 218
negative differential resistance 190, 208
negative magnetoresistance 9, 80
neighboring confinement structure 266
Neumann-type computing architecture 220
Ni 93
nonlocal resistance 112, 126
nonlocal transport 27
nonparabolicity 113
normal order 147
nuclear relaxation time 46

OMVPE 209, 227
optical atom 167
optical pumping 184
opto-electric effect 122
orbital current 17, 20
orbital magnetism 17
orbital susceptibility 19
organometallic vapor phase epitaxy 209, 227
orthogonal 9
outgoing channel 13

parabolic confinement 7
parabolic confinement potential 98
parquet graph method 58
peak-to-valley ratio 191
penetration depth 18
perfect conducter 11
periodic boundary condition 17
periodic electric potential 8
periodic lateral potential 90
periodic magnetic field 8
periodic orbit 72, 77
periodic orbit theory 77

persistent current 17
persistent photoconductivity 90
phase breaking time 206
phase coherence 4
phase coherence length 5, 115, 208
phase Hamiltonian 47
phase relaxation time 5
phase shift 37
phase space 72
photocurrent 104
photoluminescence 172, 184, 228, 239, 249, 261, 264
photoluminescence excitation spectrum 232
photon-assisted tunneling 201
photon density of states 170
photon DOS 170
photonic band calculation 169
photonic band gap 167
photonic crystal 167
piezoelectric potential 93
pillar type microcavity 163
pinned orbit 72
PL 172, 184, 228, 239, 249, 261, 264
planar microcavity 161, 163
plane wave 4
plasma oscillation 102
PLE 232
point contact 216
polariton 164
post type microcavity 163
power conversion efficiency 176
power-law decay 45
proximity effect 18
pyramid quantum dot 233

QCA 222
QD 181, 227, 233
quantization energy 4
quantized conductance 45
quantized Hall conductivity 110
quantized Hall resistance 111
quantized periodic orbit 79
quantum cellular automaton 222
quantum chaos 1, 79
quantum confinement effect 66
quantum dot 1, 4, 36, 66, 88, 96, 139, 227, 233, 240
quantum dot laser 181, 227
quantum-effect device 1
quantum electrodynamics 109
quantum Hall effect 1, 109
quantum Hall plateau 132
quantum interference transistor 188, 192
quantum nondemolition measurement 154
quantum point contact 13, 22
quantum well 96, 139, 195, 228, 265
quantum wire 1, 4, 6, 25, 50, 61, 72, 96, 139, 227, 238, 247, 255, 268
quantum wire laser 181
quarternion 10
quasi-Fermi level 182
quasimode radius 161
quenching of Hall effect 28, 72
QW 228, 265
QWR 96, 181, 227, 247, 255, 268

Rabi frequency 159, 164, 183
Rabi oscillation 159, 164, 183
Rabi splitting 165
randomness 19
random walk 4
reactive ion beam etching 172
reactive ion etching 238
real symmetric matrix 9
recursive Green's function technique 80, 112
reduced mass 145

reflection coefficient 57
reflection high-energy electron diffraction 247
reflection probability 11, 114
relaxation time 3, 20
renormalization group method 58
resistance standard 109
resonant absorption 61
resonant emission 61
resonant transmission 37
resonant tunneling diode 201
resonant tunneling hot electron transistor 190, 225
resonant tunneling transistor 188
RHEED 247
RHET 190, 225
RIBE 172
runaway orbit 72, 83

scattering probability 54
scattering time 3, 96
SCBA 76
Schawlow-Townes linewidth 182
Schottky gate 66
screened Coulomb interaction 59
selective area MOCVD growth 233
self-aligned quantum dot 244
self-assembled quantum wire 268
self-consistent Born approximation 76
self-limited growth 234
semiclassical regime 72
semiconductor superlattice 102
SET 31
SET oscillation 32
shallow etching 61
shallow mesa etching 6
shell structure 67
Si 188, 216, 264
SiGe 264
SiGe/Si 264
silicon-on-insulator 213, 216
SiN 216
single-electron device 216
single-electron switch 221
single-electron tunneling 1, 4, 31, 187
single-parameter scaling theory 9
single quantum well 256
SiN_x 234
SiO_2 238
SK growth mode 183, 243, 269
skipping motion 132
skipping orbit 55
S matrix 82
snake orbit 94
SOI 213, 216
specular reflection 27
spin-charge separation 45
spinodal phase separation 244
spin-orbit interaction 10
split gate 6
split-gate geometry 96
spontaneous emission 156
spontaneous emission coupling efficiency 168
spontaneous emission coupling ratio 157, 162
spontaneous emission enhancement factor 157
spontaneous emission rate 168
square lattice 80
SQW 256
stability angle 79
stable orbit 77
standing wave 157
Stark ladder 104, 199
Stark regime 105
stimulated emission 156
Stokes theorem 15
strain-induced lateral potential modulation 270

strain-induced quantum dot 270
Stranski–Krastanow growth mode 183, 233, 243, 269
Stranski–Krastanow quantum dot 227
strong localization 9, 87
superconducting ring 15, 18
superconducting tunnel junction 33
superlattice 199, 265
supersymmetric t-J model 45
surface emitting laser 176
symplectic 10

TBRTD 207
TEM 189, 231, 249, 268
terahertz spectrocsopy 96
thermal average 5
thermal diffusion length 5
thermal oxidation 216
thin film 4
Thouless energy 5
Thouless number 9
threshold current 176, 181
threshold voltage 31
THz spectroscopy 96, 198
tight-binding model 104
time-resolved THz spectroscopy 102
time-reversal symmetry 9, 16
Tomonaga–Luttinger liquid 1, 45, 52
topological winding number 109
torus geometry 110
T-QWR 255
trace formula 77
transfer resistance 28
transfer time 198
transition region 113
transmission coefficient 38
transmission electron microscope 189, 268
transmission probability 11, 24, 111
transport current 17
transport relaxation time 3, 46
traveling mode 82
traveling wave 157
triangular lattice 80
triple-barrier resonant tunneling diode 201, 207
T-shaped quantum wire 255
tunneling field-effect transistor 193
tunneling Hamiltonian 33
tunnel junction 31
two-terminal conductance 11
two-terminal geometry 111

UCF 17
Umklapp scattering 52
uncertainty principle 156, 163
uncertainty relation 9
unitary 9
universal conductance fluctuations 1, 12, 17
universality 110
unpaired electron 33
unstable orbit 79

vacuum fluctuation 156
vacuum Rabi oscillation 164
variable range hopping 88
VCSEL 176
vector potential 15, 62, 82, 90
vertical cavity surface emitting laser 176
vertical microcavity quantum dot laser 184
vertical quantum dot 66
V-groove 238, 241, 268
V-groove quantum wire 227, 240
vicinal substrate 228
VLSI 213

voltage distribution 115
voltage probe 28
voltage-tunable FIR detector 99

Wannier exciton 144
Wannier function 140
Wannier–Stark ladder 199
Wannier–Stark resonance 105
Wannier–Stark state 104
Ward identity 51
weak localization 9
Weiss oscillation 90
wet chemical etching 238
wetting layer 243
whispering-gallery mode 164

Yablonovite structure 174

Zener regime 105
Zener tunneling 104
ZnS 104
ZnSe 176