
Quantum Semiconductor Devices and Technologies

ELECTRONIC MATERIALS SERIES

This series is devoted to electronic-materials subjects of active research interest and provides coverage of basic scientific concepts, as well as relating the subjects to electronic applications and providing details of the electronic systems, circuits, or devices in which the materials are used.

The Electronic Materials Series is a useful reference source for senior undergraduate and graduate-level students, as well as for research workers in industrial laboratories who wish to broaden their knowledge into a new field.

Series Editors:

Professor A. F. W. Willoughby
Dept. of Engineering Materials
University of Southampton
UK

Professor R. Hull
Dept. of Material Science & Engineering
University of Virginia
USA

Series Advisor:

Dr. Peter Capper
GEC-Marconi Infra-Red Ltd.
Southampton
UK

Titles Available:

1. *Widegap II-VI Compounds for Opto-electronic Applications*
Edited by E. Rúa
2. *High Temperature Electronics*
Edited by M. Willander and H. L. Hartnagel
3. *Narrow-gap II-VI Compounds for Optoelectronic and Electromagnetic Applications*
Edited by Peter Capper
4. *Theory of Transport Properties of Semiconductor Nanostructures*
Edited by Eckehard Schöll
5. *Physical Models of Semiconductor Quantum Devices*
Ying Fu; Magnus Willander
6. *Quantum Semiconductor Devices and Technologies*
Edited by T. P. Pearsall

Quantum Semiconductor Devices and Technologies

edited by

T. P. Pearsall

*Centre de Européen de Recherche de Fontainebleau,
Corning, S.A.
Avon, France*



SPRINGER SCIENCE+BUSINESS MEDIA, LLC

Library of Congress Cataloging-in-Publication Data

Quantum semiconductor devices and technologies / edited by T.P. Pearsall
p. cm.—(Electronic materials series ; 6)

ISBN 978-0-7923-7748-1 ISBN 978-1-4615-4451-7 (eBook)

DOI 10.1007/978-1-4615-4451-7

1. Semiconductors. 2. Quantum electronics. 3. Quantum dots.

I. Pearsall, T.P. II. Series.

TK7871.85 Q36 2000

621.3815'2—dc21

99-089632

Copyright © 2000 by Springer Science+Business Media New York

Originally published by Kluwer Academic Publishers, New York in 2000

Softcover reprint of the hardcover 1st edition 2000

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher, Springer Science+Business Media, LLC.

Printed on acid-free paper.

Contents

Preface	vii
1 Quantum nanocircuits: Chips of the future?	1
Peter Hadley and Johannes E. Mooij Delft Institute of Microelectronics and Submicron Technology Department of Applied Physics Delft University of Technology Delft, The Netherlands	
2 Self-formed quantum dot structures and their potential device applications	19
Naoki Yokoyama, Hiroshi Ishikawa, Yoshiki Sakuma, Yoshiaki Nakata, and Yoshihiro Sugiyama Fujitsu, Limited, Fujitsu Laboratories, Ltd. 10-1 Morinosato-Wakamiya Atsugi 243-0197, Japan	
3 Lithography and patterning for nanostructure fabrication	97
Guy Seeborn and Harold Craighead Department of Applied and Engineering Physics 212 Clark Hall Cornell University Ithaca, NY, 14853, USA	
4 The use of MO-VPE to produce quantum structured semiconductors	139
Werner Seifert Solid State Physics Lund University Box 118, S-22100 Lund, Sweden	

5	Growth, characterization, and application of self-assembled InGaAs quantum dots	183
	Richard P. Mirin Optoelectronics Manufacturing Group National Institute of Standards and Technology 325 Broadway Boulder Colorado, 80303, USA and Arthur C. Gossard Materials Department University of California, Santa Barbara Santa Barbara, California, 93101	
6	Structural characterization of self-organized Ge islands	233
	Anton A. Darhuber and Günther Bauer Institut für Halbleiterphysik Johannes Kepler Universität Altenbergerstrasse 69 A-4040 Linz, Austria and Pieter Schittenhelm, and Gerhard Abstreiter Walter Schottky Institut Technische Universität München Am Coulombwall D-85748 Garching, Germany	
	Index	259

Preface: Quantum Semiconductor Devices and Technologies

For the better part of a decade, an important theme in solid-state device research has been the investigation of information transfer and storage in nanostructures. A nanostructure for us is any object whose dimensions are less than 50 nm. There have been many vehicles for this discussion, both theoretical (q-bits) and experimental (quantum dots). An important question posed by experimental research is whether it is possible to have quantum dots with monodisperse features, reproducible in all characteristics, that can be made to operate at room temperature. The answer is of course yes; there are at the present moment 10³ of these dots—the atomic elements. During interactions such as bonding, the atoms display energy level changes on the order of 1 eV. This level is quite stable with regards to ambient room temperature. Larger objects such as molecules or semiconductor quantum dots will display characteristic energy level separations that are smaller, making operation at room temperature more difficult, although not impossible. Biological computing via DNA (that is, all the reactions concerned with life on earth) is an example of molecular information processing at elevated temperature.

Professor Akio Sasaki and I worked together to develop a monograph on the properties of materials and devices structured in the size regime where quantum effects are important. Although we conceived this work as a single volume, the reality of production has led us to separate publication into two volumes. This volume: **Quantum Semiconductor Devices and Technologies** treats general considerations for fabrication and operation of nanoscale devices, while Professor Sasaki will present the second volume: **Quantum Structured Semiconductors**, concentrating on specific materials systems later this year. I would like to acknowledge the contribution of Prof. Sasaki to the organization of this volume.

The fabrication of quantum nanostructure devices with reproducible performance characteristics is an important goal of this field of research. A presentation of the device structures that implement memory circuits and lasers with predictable properties is given in Chapter 2 by Yokoyama, Ishikawa, Sakuma, Nakata and Sugiyama.

The synthesis of nanostructures presents a formidable array of challenges to the current level of fabrication technologies. There is hope to overcome these difficulties by using self-organizational techniques to produce the desired structures. The nature of the challenge is reviewed in Chapter 3 by Seeborn and Craighead. The self-organizational techniques used to meet this challenge are addressed in chapter 4 on MO-CVD by Seifert, in chapter 5 on MBE of GaAs/InAs by Mirin and Gossard, and in chapter 6 on Si/Ge by Darhuber, Bauer, Schittenhelm, and Abstreiter.

The greatest difficulty that quantum nanostructure devices present however is neither related to quantum effects nor to nanotechnology, but rather to temperature. Temperature plays two roles. One, mentioned above is the relationship between the characteristic separation of energy levels relative to the temperature of operation. The other is the issue of temperature rise due to energy dissipation. Temperature rise is already a big issue in silicon integrated circuits. A motivation for pursuing quantum device circuits is to achieve a density in integration that is at least 1,000 times higher than that currently being implemented in silicon VLSI. Hadley and Mooij report on this most critical issue in chapter 1.

It is my hope that this volume helps to give some perspective for the likely directions of development of this very exciting field, in terms of both the capabilities of nanotechnologies and the limits imposed by the physics of the world we live in.

T. P. Pearsall
Avon, France

Chapter 1

Quantum nanocircuits: chips of the future?

P. Hadley and J. E. Mooij

*Delft Institute of Microelectronics and Submicron Technology (DIMES) and
Department of Applied Physics, Delft University of Technology, Delft, The
Netherlands*

1.1 Introduction

Over the years, the lateral dimensions in microelectronic circuits have been shrinking systematically by a factor of two every six years. The extrapolation of the past, formulated in Moore's law, serves as the prescription for the future as laid down in the *National Technology Roadmap for Semiconductors* [1]. This Roadmap indicates gate widths for CMOS transistors of 35 nm in the year 2012. Continuation would predict minimum feature sizes of 1 nm around 2040. Many times in the past, a breakdown of Moore's law has been predicted due to limitations in fabrication, excessive power density, or discontinuous change of physical behavior. So far, the impetus of the collective microelectronics industry has pushed aside such obstacles with remarkable ease. Nevertheless, it is hard to imagine silicon CMOS technology on the true nanometer scale. Will new quantum nanodevices take over? Many introductions to papers on quantum devices suggest that this will be the case. In this chapter, we attempt to analyze the long-term potential for microelectronics applications of quantum devices. Obviously, this analysis can only depart from the types of devices

and from the physics effects that we know of today. We will limit ourselves to electronic transport devices. We focus strongly on devices that are based on manipulation of single electrons.

Quantum devices can be made of many materials and can be based on various physical principles. However, these devices share common properties. The relevant energy levels can be estimated for their operation and also for devices that in the future would make use of perfect fabrication at the atom-by-atom level. Quantitative estimates will be given for the performance of these devices as switches and memory cells in digital applications at various temperatures. The conclusion will be that quantum devices are unlikely to replace CMOS technology in computers of the types that we know today. Future developments that involve new and different physics effects may change the picture. It may also be that quantum devices can be used in drastically different modes of operation, where quantum coherence extends over multiple elements. An extreme example is so-called *quantum computation*, where the whole computer is one coherent quantum system and information is processed in a way that has no analogy with the classical Turing-type computer. We will briefly indicate the principles of quantum computation and discuss the merits of solid-state quantum devices for this purpose. Warnings against unfounded optimism about the potential for application of quantum devices in microelectronics have been sent out repeatedly by Landauer. Titles such as “Need for Critical Assessment” [2] and “Is Quantum Mechanics Useful?” [3] speak for themselves.

1.2 General physics aspects

In quantum devices, the addition or extraction of a single electron significantly changes the energy of the system. For those devices that are fabricated lithographically with sizes down to tens of nanometers, the energy change is small compared with room temperature, and devices have to be cooled down. Future controlled atom-by-atom fabrication techniques may allow the production of quantum devices with dimensions around 1 nm. Here, as we will see, the energy changes may be up to 30 times room temperature. It is necessary to confine the electrons to a well-defined region, i.e., the molecule, the quantum dot, the cluster, or the metallic island. Electronic wave functions may only be weakly coupled to the outside world. The main effects are based on two energies: the Coulomb charging energy for one additional electron and the particle-in-box confinement energy. The general background for the physics issues discussed here can be found in [4] and [5].

1.2.1 Charging energy

In this chapter, the Coulomb charging energy will be indicated as E_c . It is equal to $e^2/2C_x$, where e is the electronic charge and C_x is the sum of the capacitances between the element considered and all other elements and conductors, including the self-capacitance to the faraway ground. A sum capacitance of 1 aF (attofarad = 10^{-18} F) yields a charging energy of 80 meV or 900 K. For a flat disk of diameter D (expressed in nanometers) surrounded by a dielectric with relative permittivity ϵ , the self-capacitance in attofarads is approximately $0.04 \epsilon D$. The contribution from the connecting leads and gates has to be added. The conclusion is that, with extreme nanofabrication, a capacitance of 0.1 aF and a charging energy of 1 eV or 10,000 K might be obtainable. This statement might be misleading without the following addition: for error-free operation at the level of digital computing, the temperature should stay below the energy be at least a factor of 30. At this time, controlled fabrication of the elements that can be used in a circuit allows for a charging energy of around 10 meV or smaller.

1.2.2 Confinement energy

The two lowest energy levels for a particle with effective mass m^* in a one-dimensional square box of width w are separated by an energy $3\hbar^2/(8em^*w^2)$. This is equivalent to about $(m_0/m^*) (1\text{ nm}/w)^2$ eV. For electrons in semiconductors, m_0/m^* can be of order 10, but for devices that are not large compared with the lattice constant, this advantage is lost as the band concept breaks down. With extreme nanofabrication, one expects that the highest obtainable confinement energy will be around 1 eV, similar to the highest charging energy.

1.2.3 Tunnel barriers

Transport in and out of the quantum elements is needed for operation, but the electronic states should not be mixed with external states. To achieve the isolation of electronic states, tunnel barriers are used. The electronic levels are well defined when the tunnel resistance is much higher than the quantum resistance $R_q = \hbar/e^2$ ($\hbar$ is Planck's constant), or 25 k Ω ; the quantum effects are smeared out or lost with more transparent tunnel barriers. This is most easily made plausible for the charging effects, where the Heisenberg uncertainty time connected to a charging energy E_c is $\hbar/(2\pi E_c) = \hbar C/(\pi e^2)$. For shorter times, no control is possible. Typical operation times are of order RC , which leads to the requirement $R \gg R_q$. For confinement, a similar reasoning, equally fundamental of character, applies. The

unavoidable consequence is that impedance levels in quantum circuits are on the order of $100\text{ k}\Omega$ or higher. Barrier heights can sometimes be made tuneable so that retention times can be long in combination with reasonable switching times.

1.2.4 Environment

The dynamical response of quantum devices is strongly determined by their electromagnetic environment. The environment includes degrees of freedom such as phonons when they couple to electromagnetic phenomena. Transitions that are otherwise forbidden can occur when additional or missing energy can be dumped into or drawn from certain environmental modes (an ohmic reservoir can be modeled by a large number of harmonic oscillators with a wide range of frequencies). When transport between quantum elements is coherent, no simple addition applies. In the design of quantum circuits with multiple elements, close attention needs to be paid to these aspects. Their overall effect may be increases of operating time, impedance, and required volume.

1.3 Physical limits of integrated circuits

As computers are made faster and more cheaply, the dimensions of the transistors that make up the integrated circuits are decreasing. This allows ever more transistors to be crowded on a computer chip. Although more transistors have been added to a chip, the price of producing one square centimeter of integrated circuit has remained constant at roughly \$100 [1]. If the economies of scale continue to operate as they have in the past, the way to continue to make computers even more cheaply is to make even smaller transistors and crowd them onto a chip more densely than in the past. This will push the circuits into the realm of mesoscopic physics, where quantum mechanics determines the behavior of the circuits.

If the price per square centimeter of an integrated circuit remains constant, then an appropriate figure of merit for comparing different technologies is the number of operations that are performed per square centimeter per second. Even if the price per square centimeter does not remain constant, it is likely that in the future, microprocessors will have an area of a few square centimeters. Larger integrated circuits will be unwieldy, and making processors smaller will limit their capabilities. Thus it is interesting to know how much computing power can be packed into a square centimeter. There are two ways to increase the computing power of an integrated circuit: either the device density can be increased or the

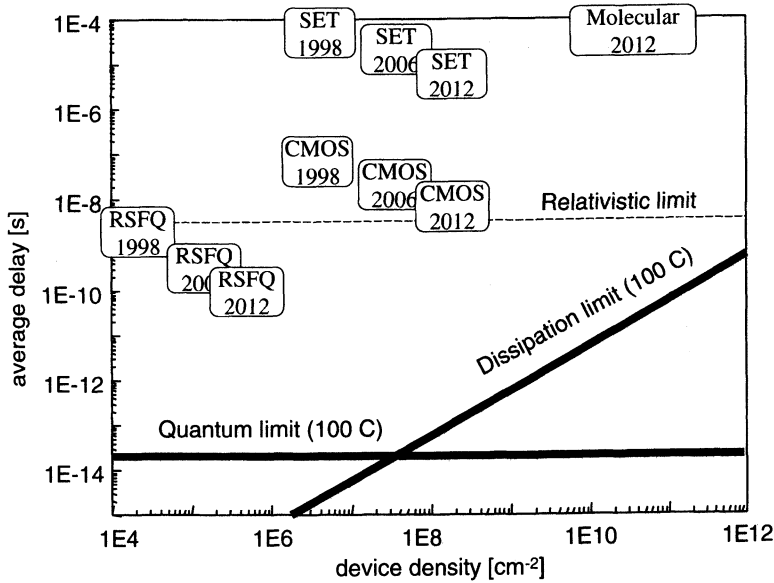


Figure 1.1: The average delay vs. device density showing the dissipation limit, the relativistic limit, and the quantum limit for room-temperature integrated circuits. The positions of CMOS circuits, SET circuits, molecular electronics, and superconducting Rapid Single Flux Quantum (RSFQ) circuits were taken from the MELARI *Roadmap for Nanotechnology* (<http://www.cordis.lu/esprit/src/melari.htm#nano>).

delay time associated with an operation can be decreased. In Fig. 1.1, the device density is plotted versus the average delay per device. The average delay per device is often much less than the clock cycle because not every device switches every clock cycle. The average delay depends on the circuit architecture. It is not possible to increase the density or decrease the average delay without bounds.

1.3.1 Dissipation limit

One of the most important limitations is formed by the dissipation of heat that is generated by the circuit. Integrated circuits that are not actively cooled typically operate at a temperature about 100°C and give off heat to the room-temperature environment. At this temperature, typical thermal fluctuations have an energy of $k_B T = 5 \times 10^{-21}$ J. In order for a bit to be stable against thermal fluctuations, the energy needed to flip a bit must be at least

10 times a typical thermal fluctuation, or 5×10^{-20} J. If quantum computing and reversible computing schemes are ignored for the moment, at least 5×10^{-20} J is dissipated for each operation that the integrated circuit performs. The power dissipated per square centimeter by an integrated circuit is therefore at least the device density times the number of operations per second per device. In real microprocessors, the dissipation is considerably more than this. The amount of thermal energy that can be transported away from the circuit is limited by the thermal conductivity of the environment of the circuit. The amount of power that can be typically transported away from the circuit is 25 W/cm^2 [1]. This dissipation limit is plotted on the density–delay diagram of Fig. 1.1. If we assume that the average cost of a square centimeter of integrated circuit remains constant, then the number of operations performed per second per unit cost can be drawn as a line in Fig. 1.1 that is parallel to the dissipation limit. The technology that comes the closest to the dissipation limit will deliver the most computational power per unit cost.

1.3.2 Quantum fluctuations

Quantum fluctuations can also disrupt circuit performance. According to Heisenberg's uncertainty principle, a quantum fluctuation of energy E can occur for a time $\hbar/E$. If the energy necessary to flip a bit is greater than 5×10^{-20} J to make the circuit stable against thermal fluctuations, then quantum fluctuations will occur for times less than 2×10^{-15} s. For the circuit to be stable against these quantum fluctuations, the switching time of a device should be at least 10 times longer than a quantum fluctuation. This means that switching times must be longer than 2×10^{-14} s. The quantum limit intersects the dissipation limit at a density of 10^7 devices/cm², which is about the current device density of CMOS. This leads to the conclusion that no dissipative, room-temperature, general-purpose computation scheme with a device density greater than CMOS circuits will be quantum noise limited. Quantum fluctuations become more relevant in cooled circuits and in high-speed signal processors.

1.3.3 RC delay

One of the factors that limits the speed of integrated circuits is the RC delay. When a particular node in a circuit is charged, the delay is the total capacitance of the node times a resistance through which the node is charged. The resistance is the sum of the output impedance of the device charging the node and the resistance of the wire leading to the node. In CMOS circuits, much attention has recently been given to the increase of the RC time due to the resistance of the wire leading to the node. This

resistance increases as all dimensions of a wire are scaled down. The formula for the resistance of a diffusive wire is

$$R = \frac{\rho l}{wt}, \quad (1.1)$$

where ρ is the resistivity, l is the length, w is the width, and t is the thickness. If l , w , and t are all decreased by a certain factor, the resistivity R increases by the same factor. An increase in R slows down the circuit. This situation gets even worse for very short wires. As the length of a wire goes to zero, the resistance does not go to zero, as Eq. (1.1) predicts. The resistance of the wire is always limited by the relationship [6]

$$R > \frac{\lambda_F^2}{wt} \left(\frac{\hbar}{e^2} \right). \quad (1.2)$$

Here λ_F is the Fermi wavelength. This means that there is a critical length below which Eq. (1.1) is no longer valid:

$$l_{crit} = \frac{\lambda_F^2}{\rho} \left(\frac{\hbar}{e^2} \right).$$

For metals, this critical length is about 10 nm. One expects an increase in the resistance when the distance between devices approaches 10 nm,

For quantum devices such as single-electron transistors and quantum dots, the dominant resistance is not the wire that connects the nodes; it is the output impedance of the device itself. The output impedance of a single-electron transistor or a quantum dot must be greater than the resistance quantum; therefore, the resistance that determines the delay is roughly 1,000 times larger in a quantum circuit than in CMOS. The way to compensate for this increase in resistance is either to decrease the capacitance of the nodes or to increase the packing density so that more computational elements are packed into a square centimeter. The capacitance of the nodes is largely determined by the physical size of the components of the circuit. In other words, for the same device density, high-resistance quantum devices are going to be slower than CMOS. This is reflected in the position of the quantum devices in Fig. 1.1. As long as a low-impedance element like a field-effect transistor can be used for logic circuits, it will be used, because the resulting circuits will be much faster.

1.3.4 Relativistic limit

Another limit that has consequences for integrated circuits is the relativistic limit that is caused by the finite speed of light. No information can be transported a distance of 1 cm in less than 0.3 ns. It is possible for inte-

grated circuits to operate in the relativistic regime, but if the average delay is to be much less than the relativistic limit, the circuit must be designed such that most of the communication between devices is local. The shorter the delay, the more local the communication must be. In Fig. 1.1, the relativistic limit was drawn for a circuit with an area of 1 cm^2 . For larger circuits, the limit will have a longer average delay time.

1.4 Single-electron transistors

Although a number of different quantum devices could be used to build circuits, the quantum device that has been most widely discussed is the single-electron transistor (SET) [7]. A single-electron transistor is a three-terminal device consisting of an input terminal (the gate) and two output terminals (the source and drain). A change in the charge on the gate modulates the current that flows from source to drain. Figure 1.2 shows the conductance through a SET as a function of bias voltage and gate voltage for a metal SET, a semiconducting SET [8], a carbon nanotube SET [9], and a superconducting SET [10]. For the metal SET and the superconducting SET, the current that flows from source to drain is a periodic function of the gate voltage with periodicity e/C_g , where C_g is the gate capacitance. For the semiconducting SET and the carbon nanotube SET, the current is not strictly a periodic function of the gate voltage because energy needed to add an electron to the island is the sum of the charging energy and the confinement energy. For the semiconducting SET and the carbon nanotube SET, the charging energy has the same order of magnitude as the confinement energy, whereas for a metal SET or a superconducting SET, the charging energy is much larger than the confinement energy.

From the current–voltage characteristics, one can determine that the input capacitance at the gate can be very small, typically $1\text{--}10\text{ aF}$, and that there is virtually no input current leakage. The output impedance of a SET is quite large, typically $100\text{ k}\Omega\text{--}100\text{ M}\Omega$ —orders of magnitude higher than in typical semiconductor devices. This impedance size causes the longer RC delay for SETs and quantum dots that was discussed above.

The data in Fig. 1.2 can also be used to determine the charge gain and the voltage gain of a SET. The charge gain is frequency dependent and determines the maximum operation frequency of the SET. When the gate charge is modulated with an amplitude Δq at a particular frequency, the current through the SET transistor is modulated with an amplitude ΔI at the same frequency. The charge that is transported through the SET transistor every half period is $\Delta IT/2$, where T is the period. The charge gain is thus $g_q = \Delta IT/(2\Delta q)$. In any circuit where the output of one SET drives the

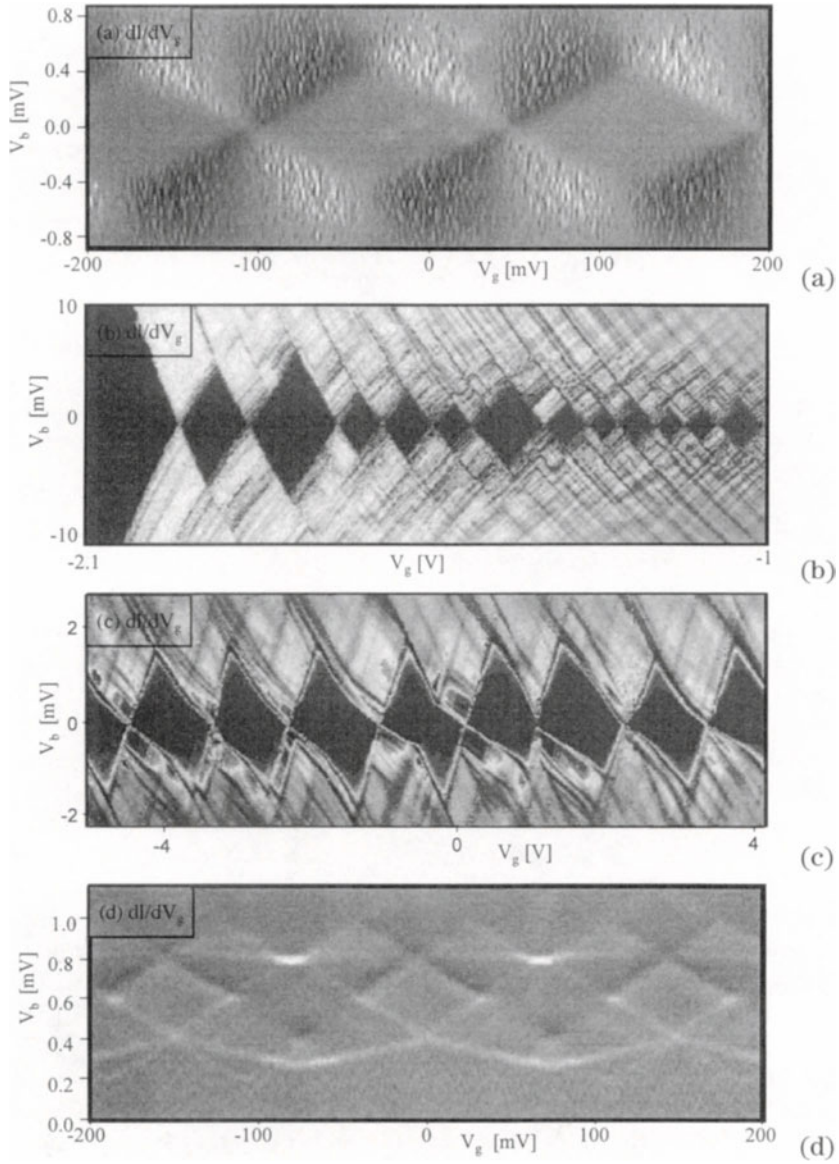


Figure 1.2: The low-temperature conductances of (a) a metal single-electron transistor (SET), (b) a semiconducting SET, (c) a carbon nanotube SET, and (d) a superconducting SET are plotted as a function of gate voltage and bias voltage. The diamond-shaped regions along the zero-bias voltage axis are regions of Coulomb blockade. In these regions, no current flows. The conductance is a periodic function of gate voltage for the metal SET and the superconducting SET, where the confinement energy is negligible. The conductance is not a periodic function of gate voltage for the semiconductor SET and the carbon nanotube SET, where the confinement energy is important.

input of the next SET, the charge gain has to be greater than 1 at the operating frequency. This is equivalent to saying that the gate must be modulated with a frequency much less than the average time between electron tunnel events e/I . Voltage gain is necessary for some applications. The voltage gain of a SET is the ratio of the modulation of the source-drain voltage to the modulation of the gate voltage at a given current bias $g_v = \Delta V_{DS}/\Delta V_g$. The voltage gain can be determined from the data in Fig. 1.2. The gain is gate voltage dependent and for small signals ranges between $g_v = -C_g/C_j$ and $g_v = C_g/(C_g + C_j)$. Voltage gain greater than 1 is only possible when the gate capacitance is greater than the junction capacitance. As an amplifier, a SET has a limited dynamic range. The gate charge modulation must be smaller than about $e/3$ for the small signal limit to apply. In the limit of large gate modulation, the voltage gain goes to zero.

The charge sensitivity of a SET is exquisite. Charge variations much smaller than the charge of an electron can be detected. A typical charge sensitivity for a SET is $10^{-4} e/\sqrt{Hz}$. At this time, the SETs with the largest bandwidth are RF SETs, which can operate up to about 1 GHz [11]. The tremendous charge sensitivity of a SET also introduces complications. A SET is so sensitive that it responds to minute changes in the charge distribution of its environment. If a crystal defect in the vicinity of a SET carries a charge e and moves slightly, this will be detected by the SET. These charged defects are called *background charges* or *offset charges* [12]. Even when the charged defects do not move, they couple charge to the SET in the same way that the gate voltage does. Thus offset charges produce a random shift along the gate voltage axis of the current-voltage characteristics shown in Fig. 1.2.

For circuits to operate properly, the charging energy should be much greater than the thermal fluctuations. At 373 K, the charging energy equals $k_B T$ for a capacitance of 2.5 aF. However, simulations show that there is hardly any modulation of the current for a SET with $C_\Sigma = 2.5$ aF at 373 K. It is necessary to have junction capacitances of about 0.1 aF before significant current modulation is achieved. Figure 1.3 shows a simulation of the current-voltage of a SET with junction capacitances of 1 aF and 0.1 aF and a modest voltage gain of 2. These capacitances are extremely small and correspond to islands only a few nanometers in size. It will be impossible to make the SETs much smaller than this because SETs or quantum dots that operate at room temperature will only be a few atoms wide. This means that there will not be much opportunity to scale room-temperature SETs smaller once they are fabricated.

The switching delay of a SET can be much longer than the RC delay explained above. This longer delay occurs at high temperatures where the current modulation of a SET decreases from the maximum low-

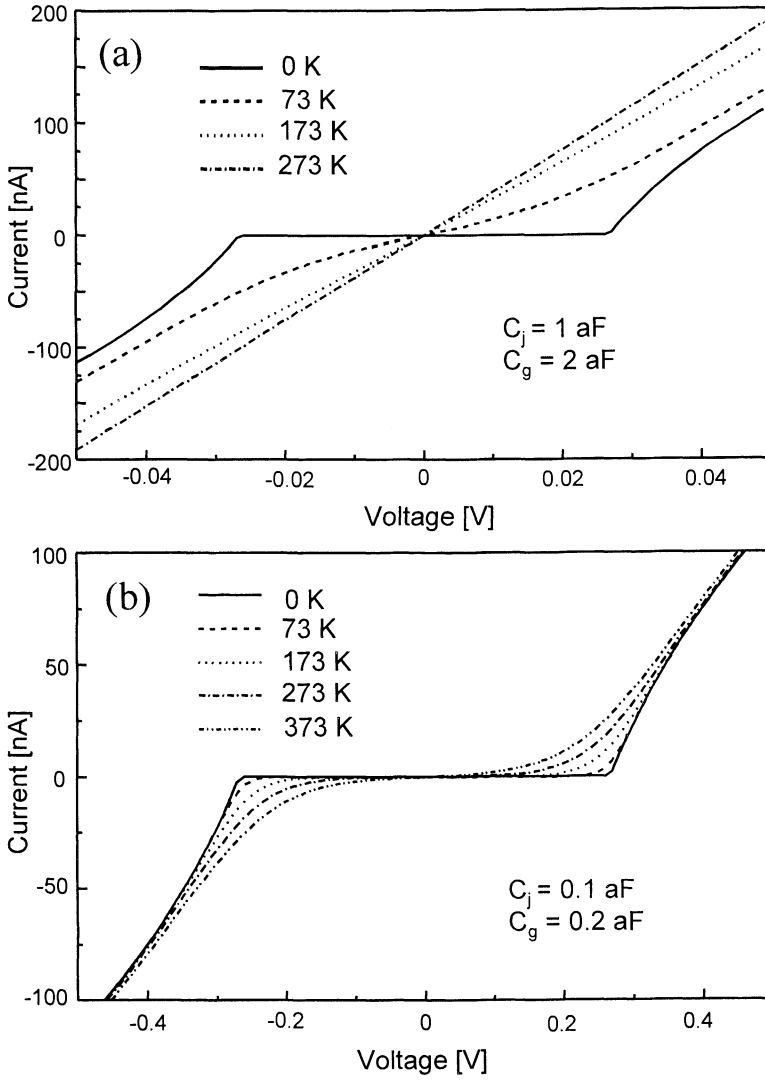


Figure 1.3: Simulations of the current-voltage characteristics of a metal single-electron transistor (SET) for various temperatures. These SETs have a voltage gain of 2. (a) Junction capacitance $C_j = 1$ aF; gate capacitance $C_g = 2$ aF; (b) $C_j = 0.1$ aF; $C_g = 0.2$ aF.

temperature value of approximately $e/(RC_x)$. This delay also occurs for multiple tunnel junctions where random offset charges reduce the current modulation to below the optimum value even at low temperature. In these cases, the minimum delay time is the inverse of the maximum frequency for which there is charge gain, $f = \Delta I/e$.

1.5 Applications

1.5.1 Logic

SETs dissipate little power and can be fabricated in sizes ranging down to a molecular scale. These characteristics make them potentially useful for logic and memory circuits. It has been suggested that large-scale integration of single-electron transistors could result in logic circuits with a high device packing density. Recently, a number of different logic schemes using single-electron-tunneling transistors have been proposed [13].

In single-electron logic (SEL) as proposed by Averin and Likharev, bits are represented by individual electrons [14]. SEL has an architecture similar to superconducting single-flux quantum logic, where individual magnetic-flux quanta are manipulated [15]. A logic scheme where bits were represented by individual electrons was also proposed by Ancona. He proposed building circuits with a regular cellular structure, with local connections as in cellular automata [16]. Tucker [17] and Yoshikawa et al. [18] suggested the use of single-electron transistors in an architecture very similar to CMOS. In this case, the bits are represented by voltage levels [19,20]. The use of multiple tunnel junctions in place of individual single-electron transistors in this sort of architecture has also been proposed [21]. Tougaw and Lent suggested that the electron configuration in a cluster of quantum dots could be used to encode the information [22]. The clusters of quantum dots are placed very close together, but they are not electrically connected. The information passes from one cluster to the next via the Coulomb interaction between the electrons. Computation proceeds locally as in a cellular automata architecture. Korotkov proposed a similar computational scheme using clusters not connected by wires [23]. Unlike Tougaw and Lent, Korotkov suggested the use of an AC driving field so that the energy needed to drive the computation could be coupled into each of the clusters directly. Ohshima and Kiehl also proposed a SET logic scheme using an AC pump (tunneling phase logic) [24]. In this scheme, the information is not coded in the configuration of some charge on a cluster but rather is coded in the relative phase between the AC pump and the periodic single-electron-tunneling oscillations. Likharev and Korotkov also

proposed reversible logic elements [25] where the amount of energy dissipated per switching event could be much less than $k_B T$. There have also been several suggestions to use single-electron-tunneling transistors to build neural nets. Goossens et al. suggested exploiting the periodic transfer characteristics of a SET transistor to build neural nets [26], while the quantum neuromorphic networks of Bandyopadhyay et al. [27] function more like a traditional Hopfield machine.

Although many logic schemes have been proposed, very few circuits have been tested experimentally. Recently, a simple SET logic circuit, which displays negative differential resistance, was demonstrated using a multilayer fabrication technology [28]. A multilayer fabrication technology is important, since wires must cross in complex circuits. The negative differential resistance device was constructed from two SETs with a voltage gain greater than 1. Voltage gain is required in many of the logic schemes mentioned above.

No logic scheme based on quantum devices is presently threatening to overthrow the semiconductor industry standard CMOS technology for logic applications. Many of these logic schemes are flawed in that they ignore the existence of offset charges. This approach is unrealistic. There is no experimental evidence that offset charges can be eliminated in any materials system. The only offset-charge-independent logic schemes use high-resistance resistors. However, the fabrication of suitable compact, low-capacitance, high-resistance resistors is a formidable technological problem. Even if this problem were to be overcome, the high resistance makes the quantum devices so much slower than CMOS that they are not competitive.

1.5.2 Quantum computation

So far, we have discussed the possible replacement of CMOS semiconductor elements by quantum devices. The roles of these two types of devices in the circuit would be identical. It turns out, however, that quantum devices provide no major advantage in this conventional setting. Much more may be gained if the intrinsic coherence of quantum states is used. The evolution in time of an isolated quantum system under the influence of a known external drive is completely deterministic. Quantum uncertainty comes in when a measurement is performed; i.e., the quantum system is made to interact with a large number of uncontrolled “classical” variables. Quantum devices in a classical matrix are continuously subjected to a measurement in this sense, resulting in severe limitations in performance due to quantum fluctuations.

Quantum systems can exist in general superpositions; for a two-level

system, the relative amplitude and the relative phase of the occupation coefficients are two independent continuous variables. In principle, these variables can be manipulated, in particular by the application of microwave pulses of the frequency that corresponds to the energy difference between the two states. Two-level quantum systems can also be coupled so that their states are entangled. There is no classical analogue to the notion of quantum entanglement.

In the last few years, the principles have been worked out for a new method of information processing that is called quantum computation [29]. In a quantum computer of this type, the bits are replaced by qubits (two-level systems as described in the previous paragraph). Qubits can be coupled by quantum gates. The whole quantum computer with N qubits lives in a superposition of its 2^N quantum states. The applied operations correspond to a unitary transformation that has to be expressed by a 2^N by 2^N matrix. For a computation, the quantum computer is first prepared in a known state and a series of operations is performed. Finally, a measurement is performed; this destroys the quantum information. During the computation, no decoherence due to coupling to external degrees of freedom is allowed.

It has been shown that a quantum computer can perform certain tasks much faster than a CMOS computer. A good example is the factorization of a large number: the time needed to factorize a number of N digits increases exponentially in a conventional computer but only with a power law in the quantum computer [30]. Quantum coherence brings in an essential advantage. To be competitive, a quantum computer needs a minimum of 10^4 or 10^5 qubits and a comparable number of operation steps, all performed within the decoherence time.

Only rudimentary quantum computers of a few qubits have been realized so far, by means of nuclear magnetic resonance and by laser manipulation of electron states in trapped single atoms. Integration to large numbers of qubits seems difficult. The use of solid-state quantum devices of the types described in this book can be considered. Ideas concentrate mostly on spin-polarized states in quantum dots and charge or fluxoid states in circuits of superconducting tunnel junctions. The suitability is mainly determined by the effective decoherence rate. Decoherence times that have been measured are on the order of 1 ns, in situations where measuring leads are connected and for systems that are susceptible to charge noise. For optimized devices, much longer times—more than 1 ms—have been calculated. Certainty can only be obtained from experiments. If solid-state quantum devices demonstrate sufficient coherence, the fact that they can be integrated into large circuits makes them excellent candidates for the development of a real quantum computer.

1.5.3 Memories

Memories are a more promising application of quantum devices than logic because density is more important for memories and speed is less important. Although great advances have been made in computer memories, there remains much room for improvement. The density of a two-dimensional memory is ultimately limited by the thickness of the address lines to the memory cells. The thinnest conductors that conduct over long distances are about 1 nm wide, resulting in a density of about 10^{13} bits/cm². Such a memory would have a linear information density about equivalent with DNA.

In single-electron memories, a bit is represented either by one electron or a few electrons on a storage node. The basic memory cell also has a charge-sensing device, either a field-effect transistor or a single-electron transistor. The charge-sensing device has to exhibit charge gain. The speed of these memories is limited by the time it takes to charge and discharge the storage node and by the maximum frequency for which the charge-sensing device exhibits charge gain. Yano et al. [31] and Stone et al. [32] used naturally occurring grains as the storage nodes and a chain of naturally occurring grains to sense the charge. There is no control of the structure of the grains, which leads to some parameter spread, but grains can be small enough that effects are observable at room temperature. Guo et al. [33] also achieved room-temperature operation using a single artificially structured polysilicon dot as the storage node and a field-effect transistor as the charge-sensing device. Likharev discussed a memory cell where the charge-sensing device is a SET; Likharev has also critically examined the optimal parameters of the tunnel barrier that is used to charge the storage node [34].

1.5.4 Metrology and sensors

Quantum devices seem certain to make a major contribution to the field of precision measurements and charge sensors. Single-electron transistors are the most sensitive charge measurement devices presently available. A SET is in many respects the dual to the SQUID, the superconducting sensor of magnetic flux. In a SQUID, a so-called flux transformer can be used to separate the sensor and the pickup coil. No such possibility exists for the SET. A long lead to the gate implies a large capacitance that is added to the total island capacity and destroys the sensitivity. Consequently, the SET has to be brought in the direct neighborhood of the potential to be measured. A natural and very interesting solution is to position the SET on the tip of a scanning probe system. This solution was first realized by Yoo et

al. [35]. When a single-electron transistor is fabricated on the tip of a scanning probe, it is possible to use the tip as an STM or AFM to determine the topology, while the SET measures potential variations. The sensitivity at room temperature should ultimately be high enough to detect single electron charges with a spatial resolution in the subnanometer scale.

Single-electron devices can also be used to make precise measurements of current. A device related to a SET is an electron pump. In an electron pump, gates are modulated to repeatedly push individual electrons through a circuit. This process produces a current with a magnitude $I = ef$, where I is the current and f is the modulation frequency. An inaccuracy of 10^{-8} has been reached by Keller et al. [36]. For metrological applications, low temperatures are not a strong disadvantage. However, nanoscale devices with much smaller capacitance will be extremely useful, since the operating frequency can be brought to a level that corresponds to more reasonable current levels (100 GHz corresponds to only 16 nA).

1.6 Conclusions

It is very unlikely that single-electron quantum devices will replace CMOS as switching elements for logic in general-purpose computation. The main disadvantage is the intrinsically high impedance level of more than 100 k Ω . If the product of device impedance and interconnect capacitance determines the response time, then with equal interconnect lengths the lower impedance device wins. Drastically new ideas would be needed for this conclusion to be changed. In the area of memories, where raw speed is less important, single-electron quantum devices may eventually be competitive.

Single-electron quantum devices can only be used for regular computation at room temperature when their dimensions are below 10 nm. Controlled fabrication for large circuits will not become available until the microelectronics *Roadmap* dictates this. Clearly, 10 nm is getting so close to the finally unavoidable end of the road that the willingness to make a switch to a completely new technology may not exist, even when the performance of quantum devices would be superior.

Nanoelectronics based on quantum devices may be relevant for applications that are complementary to standard information processing, rather than competitive. The low power level of quantum devices makes them highly suitable for fault-tolerant computation at very high densities. A local architecture will be needed to avoid long delay times due to the high resistance.

Predictions about the long-term potential of new types of devices may

well be wrong, since they are subject to the limitations of our present imagination. The field of quantum devices is young, and new ideas may provide a breakthrough. Quantum computing is an example that makes essential use of the quantum character without any classical analogue. Even if quantum devices do not replace mainstream semiconductor devices, the continued study of quantum effects that occur on a nanometer length scale is important because these effects will influence the performance of CMOS semiconductor devices when they reach the nanometer regime.

ACKNOWLEDGMENTS

This research was supported by CHARGE, Esprit project 22953, NEDO, and by Stichting voor Fundamenteel Onderzoek der Materie (FOM).

REFERENCES

- [1] *The National Technology Roadmap For Semiconductors*, Semiconductor Industry Association (1997), <http://notes.sematech.org/ntrs/PublNTRS.nsf>.
- [2] R. Landauer, *IEEE Trans. Electron Devices* **43**, 1637 (1996).
- [3] R. Landauer, *Philos. Trans. R. Soc. London, Ser. A*, **353**, 367 (1995).
- [4] *Single Charge Tunneling*, edited by H. Grabert and M. H. Devoret (Plenum Press, New York, 1992).
- [5] *Mesoscopic Electron Transport*, edited by L. L. Sohn, L. P. Kouwenhoven, and G. Schön (Kluwer Academic Publishers, Dordrecht, 1997).
- [6] Y. V. Sharvin, *Sov. Phys. JETP* **21**, 655 (1965).
- [7] K. K. Likharev, *IEEE Trans. Mag.* **23**, 1142 (1987).
- [8] L. P. Kouwenhoven, T. H. Oosterkamp, M. W. S. Danoesastro, M. Eto, D. G. Austing, T. Honda, and S. Tarucha, *Science* **278**, 1788 (1997).
- [9] Sander J. Tans, Michel H. Devoret, Remco J. A. Groeneveld, and Cees Dekker, *Nature* **394**, 761 (1998).
- [10] P. Hadley, E. Delvigne, E. H. Visscher, S. Lähteenmäki, and J. E. Mooij, *Phys. Rev. B* **58**, 15317 (1998).
- [11] R. J. Schoelkopf, P. Wahlgren, A. A. Kozhevnikov, P. Delsing, and D. E. Prober, *Science* **280**, 1238 (1998).
- [12] A. B. Zorin, F.-J. Ahlers, J. Niemeyer, T. Weimann, H. Wolf, V. A. Krupenin, and S. V. Lotkhov, *Phys. Rev. B* **53**, 13682 (1996).
- [13] A. N. Korotkov, in *Molecular Electronics*, edited by J. Jortner and M. A. Ratner (Blackwell, Oxford (1997), p. 157. K. K. Likharev, *FED J.* **6**, 5 (1995).

- [14] D. V. Averin and K. K. Likharev, in *Mesoscopic Phenomena in Solids*, edited by B. L. Altshuler, P. A. Lee, and R. A. Webb (Elsevier, Amsterdam, 1991).
- [15] K. K. Likharev and V. K. Semenov, IEEE Trans. Appl. Supercond. **1**, 3 (1991), <http://pavel.physics.sunysb.edu/RSFQ/RSFQ.html>.
- [16] M. G. Ancona, J. Appl. Phys. **79**, 526 (1996). M. G. Ancona, Superlat. Microstruct. **20** (1996).
- [17] J. R. Tucker, J. Appl. Phys. **72**, 4399 (1992).
- [18] Nobuyuki Yoshikawa, Yasuyuki Jinguu, Hiroshi Ishibashi, and Masanori Sugahara, Jpn. J. Appl. Phys. **35**, 1140 (1996).
- [19] R. H. Chen, A. N. Korotkov, and K. K. Likharev, Appl. Phys. Lett. **68**, 1954 (1996).
- [20] A. N. Korotkov, R. H. Chen, and L. K. Likharev, J. Appl. Phys. **78**, 2520 (1995).
- [21] Haroon Ahmed, J. Vac. Sci. Technology B **15**, 2101 (1997).
- [22] P. D. Tougaw and C. S. Lent, J. Appl. Phys. **75**, 1818 (1994).
- [23] A. N. Korotkov, Appl. Phys. Lett. **67**, 2412 (1995).
- [24] T. Ohshima and R. A. Kiehl, J. Appl. Phys. **80**, 912 (1996).
- [25] K. K. Likharev and A. N. Korotkov, Science **273**, 763 (1996).
- [26] M. J. Goossens, J. H. Ritskes, C. J. M. Verhoeven, and A. H. M. van Roermund, Proc. 1997 Eur. Conf. Circuit Theory and Design (ECCTD'97), 937–941 (1997).
- [27] S. Bandyopadhyay, V. P. Roychowdhury, and X. Wang, Phys. Low-Dim. Struct. **8/9**, 29–82 (1995).
- [28] C. P. Heij, D. C. Dixon, P. Hadley, and J. E. Mooij, to appear in Appl. Phys. Lett.
- [29] D. P. Divincenzo, Science **270**, 255 (1995), G. P. Berman, G. D. Doolen, R. Mainieri, and V. I. Tsifrinovich, *Introduction to Quantum Computers* (World Scientific, 1998).
- [30] P. W. Shor, in Proc. Foundations of Computer Science, IEEE Computer Society, Los Alamitos, CA, 124 (1994).
- [31] Kazuo Yano, Tomoyuki Ishii, Takashi Hashimoto, Takashi Kobayashi, Fumio Murai, and Koichi Seki, IEEE Trans. Electron Devices **41**, 1628 (1994).
- [32] N. J. Stone and H. Ahmed, APL **73**, 2134 (1998).
- [33] Lingjie Guo, Effendi Leobandung, and Stephen Y. Chou, Science **275**, 649 (1997).
- [34] K. K. Likharev and A. N. Korotkov, Proc. ISDRS (1995).
- [35] M. J. Yoo, T. A. Fulton, H. F. Hess, R. L. Willett, L. N. Dunkleberger, R. J. Chichester, L. N. Pfeiffer, and K. W. West, Science **276**, 579 (1997).
- [36] Mark W. Keller, John M. Martinis, and R. L. Kautz, Phys. Rev. Lett. **80**, 4530 (1998).

Chapter 2

Self-formed quantum dot structures and their potential device applications

**Naoki Yokoyama^{1,2}, Hiroshi Ishikawa², Yoshiki Sakuma¹,
Yoshiaki Nakata², and Yoshihiro Sugiyama¹**

*¹Fujitsu Limited; ²Fujitsu Laboratories Ltd., 10-1 Morinosato-Wakamiya, Atsugi
243-0197, Japan*

2.1 Introduction

Quantum dots (QDs) have attracted a great deal of interest during the past decade, both in quantum nanostructure physics and device engineering. This is simply because the QD structure is ultimately a scaled-down structure of electron and optical devices that we shall reach sometime in the future, and because it has possibilities for developing new principles or new concept devices. This area of research seems to have been promoted by the continuous interest and efforts of researchers to find something new to cultivate in new areas of science and technology.

The most important subject in QD research is the establishment of a method of fabricating quantum dot structures. Many challenges to fabrication have been overcome during the last two decades. The most

straightforward technique is to use high-resolution electron beam lithography for nanoscale patterning and dry or wet etching to make quantum wire or quantum dot structures [1–4]. Quantum structures fabricated through these techniques seem, however, not to be satisfactory in that they do not fully use the features of quantum dots. It appears that etching-based technologies have a drawback in that they introduce damage and/or contamination by impurities into the crystals.

Another technique is to use the regrowth of epitaxial layers, such as fractional layer growth on the step edge of a vicinal substrate [5,6], selective growth on a patterned substrate [7,8], and cleaved-edge overgrowth [9–11]. These structures have the advantages of no etching damage and no need for fine lithography but seem to have a drawback in that they have poor dot density for optical applications and have a structure that is not easy to fabricate in electron device construction.

This chapter describes recent advances in quantum dot fabrication technologies that have been developed to solve these drawbacks. It then discusses the potential device applications of these technologies, focusing on quantum dot lasers and high-density memory applications.

2.2 Tetrahedral-shaped recess (TSR) quantum dots

To use QDs in electron devices requires that the QDs be made in artificially designed positions with controllable sizes and material compositions. To meet these requirements, we have proposed and demonstrated a novel QD structure that is fabricated using tetrahedral-shaped recesses (TSRs) [12,13]. This section describes the process of fabricating InGaAs QDs in the TSRs formed in substrates and their optical characterizations to confirm the formation of quantum dots. We then discuss the growth mechanism of a QD formed at the bottom of a TSR. Further, we describe the ability to change the energy levels through control of the dot size and composition during growth.

2.2.1 TSR formation

After making circular apertures in a SiO_2 mask deposited on the semi-insulating (111)B-GaAs substrate, we etched the substrate through the apertures with an anisotropic wet chemical etchant of a 0.5% Br_2 -ethanol solution. Figure 2.1 shows a diagram of the process as well as highly resolved scanning electron microscope images of the etched surface. The etched surfaces consist of three equivalent $\{111\}$ A side facets that appear

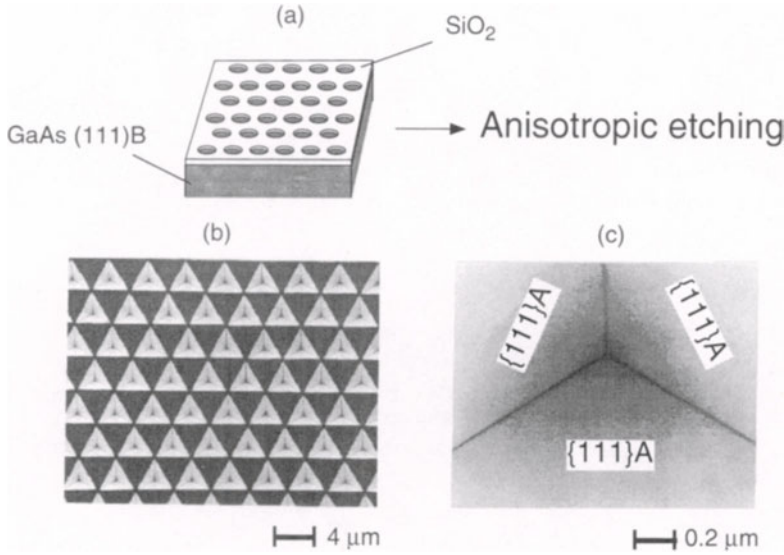


Figure 2.1: TSR forming process and SEM plan views of TSRs. (a) Initial GaAs (111)B substrates with 100 nm-thick SiO₂ patterned mask; (b) an SEM picture of TSR array; (c) a magnified SEM picture in the vicinity of a TSR bottom.

due to the extremely slow etching rate of the etchant [14], resulting in the formation of a TSR. An array of TSRs was arranged in triangular lattices in a square area, 2×2 mm, on a substrate. Typically, the side length of a surface TSR triangle was about $3 \mu\text{m}$; the distance between TSRs was $4 \mu\text{m}$. The magnified view of the etched surface, shown in Fig. 2.1c, shows that the {111}A facets were very smooth and that the shape of the bottoms were identical and acute. It should be noted that the regular shape of the recess after etching was due to the “zinc-blend-type” crystallographic characteristics, rather than the shape of the initial apertures of the SiO₂ mask.

2.2.2 Quantum dot formation and optical properties

We grew a GaAs/InGaAs/GaAs heterostructure inside the TSRs using low-pressure metal-organic chemical vapor deposition (MOCVD), as illustrated in Fig. 2.2a. Growth was performed on the TSR-formed substrates with the SiO₂ mask at a temperature of 600°C and a pressure of 50 Torr.

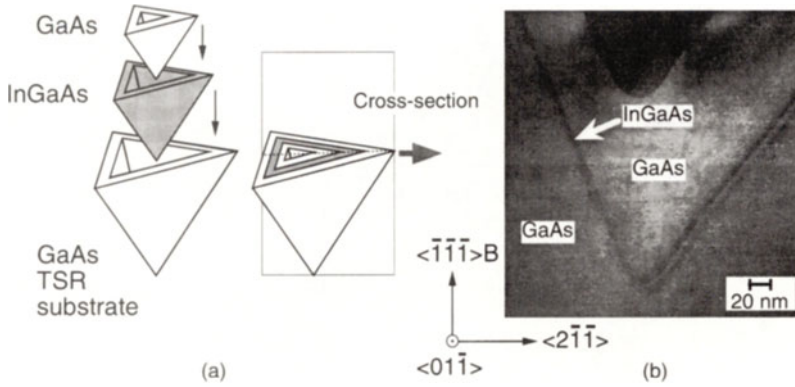


Figure 2.2: (a) Illustration of GaAs/InGaAs/GaAs growth inside a TSR; (b) a cross-sectional TEM image.

Trimethylgallium (TMGa), trimethylindium (TMIn), and AsH_3 were employed as sources. The TMGa was kept at -10°C , and the H_2 flow rate through the TMGa bubbler was 5 standard cubic centimeters per minute (sccm). The TMIn container was kept at 13.5°C . The TMIn flow rates and the growth times for the InGaAs layer were varied by fixing the flow rate of pure (100%) AsH_3 at 20 sccm. The total flow rate in the reactor was 2,000 sccm. For all samples, the thickness of the buffer GaAs was 100 nm and of the cap GaAs was 150 nm. Here, the thickness was defined toward the normal direction at the $\{111\}\text{A}$ side facets. Under these growth conditions, the GaAs and InGaAs layers had an almost uniform thickness profile on the $\{111\}\text{A}$ facets and did not grow on the SiO_2 mask because of the perfect selective area epitaxy.

Figure 2.2b is a cross-sectional transmission electron microscope (TEM) image of the sample where a 2.5 nm-thick InGaAs quantum well (QW) was grown at $\{111\}\text{A}$ sidewalls under the TMIn flow rate of 50 sccm. It is clear that the InGaAs is grown with a uniform thickness over the $\{111\}\text{A}$ sidewall of the TSR. In other words, unlike the reported quantum wires grown in the V-grooves formed on the (001) substrates [15–17], under our growth conditions we did not observe an enhancement in the growth rate in the vicinity of the bottom of the TSR.

Figure 2.3a is the PL spectrum measured at 4.2 K from the same sample shown in Fig. 2.2b. We observed two distinct emission peaks, one at 1.45 eV and another at 1.48 eV. The corresponding monochromatic cathodoluminescence (CL) images are shown in Fig. 2.3b. These plan views of CL images indicate that the emission of 1.45 eV comes from the TSR bottoms and that of 1.48 eV from the sidewalls. The results suggest that InGaAs QDs

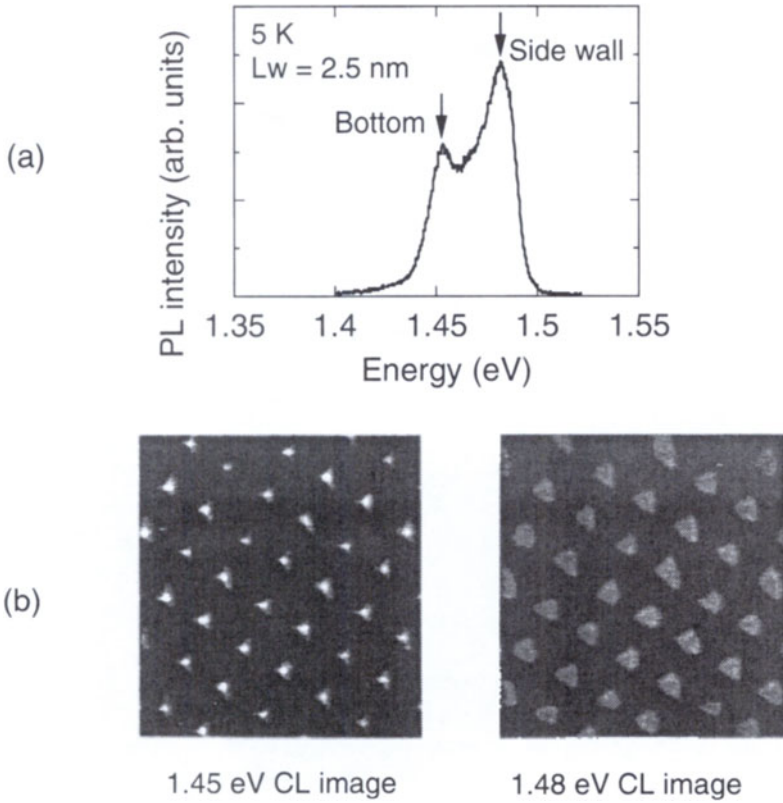


Figure 2.3: (a) PL spectrum at 5 K; (b) monochromatic CL plan views. (From Applied Physics Letters Vol. 67, Sugiyama et al., p. 257 (1995), Copyright 1995 by the American Institute of Physics)

are spontaneously formed on the bottom of every TSR, while the QWs are formed at the sidewalls.

We compared the optical property of the QD emission peaks with that of the QWs by studying the dependence of CL spectra on the excitation power of the electron beam. Here, we studied the CL from only one TSR structure by narrowing the scanning area of electron beam on the sample in the CL apparatus [18]. The results, shown in Fig. 2.4, were measured for a different sample. The sample was of InGaAs QW width, L_w , of 5 nm, which was under the TMIn flow rate of 100 sccm. The level of emission current from the electron gun was varied to change the excitation intensity under a constant acceleration voltage of 5 keV. Again, we observed two distinct peaks of QDs at 1.34 eV and QWs at 1.41 eV in the CL spectrum of the lowest

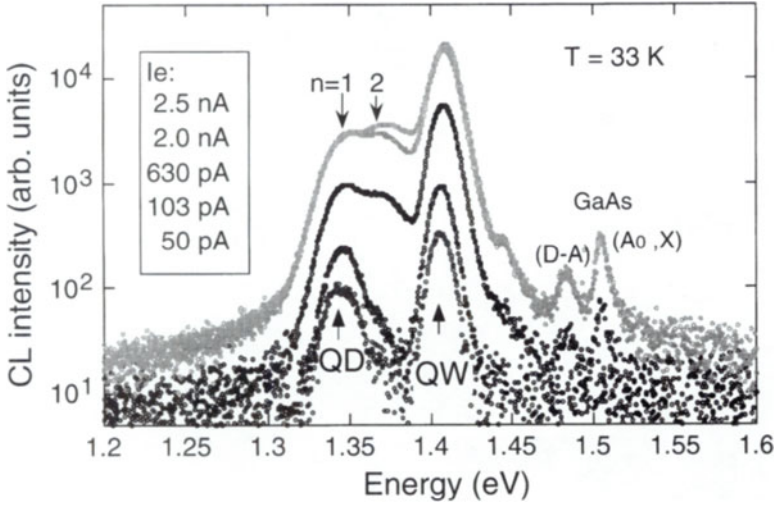


Figure 2.4: CL spectra from a single TSR measured at various electron beam intensity currents.

excitation level. With increasing excitation power, the intensity of the QD peak tends to saturate and a new peak appears at 1.38 eV. The 1.34 eV peak corresponds to the ground state emission of a QD; the peak at 1.38 eV is from the first excited state. This change in the spectra is considered to be caused by so-called *band filling*, or state-filling phenomena in the dot. Note that no band filling is observed for the QW peak, the peak position of the QW does not shift as much, and its intensity increases almost linearly with power. That the filling of the QD occurs at a much lower excitation power than do the QWs on the TSR sidewalls implies that the QD contains only a small number of energy states, suggesting a small size (or volume) of QDs.

To demonstrate more direct evidence of the zero-dimensional carrier confinement, we performed a magneto-PL study on our TSR sample under high magnetic fields. The magneto-PL at 50 K in both the Faraday ($B//z \perp E$) and Voigt ($B \perp z \perp E$) configurations was measured. B , z , and E are the magnetic field vectors, the direction along $\langle 111 \rangle B$, and the electric field vector of luminescent light. The PL spectra were gathered from the numerous TSRs within a 1 mm^2 area, using an optical fiber scope, and the dispersed light was detected by an optical multichannel analyzer (Si CCD) system combined with a single monochromator. The 488 nm line of an Ar^+ laser was used as an excitation source.

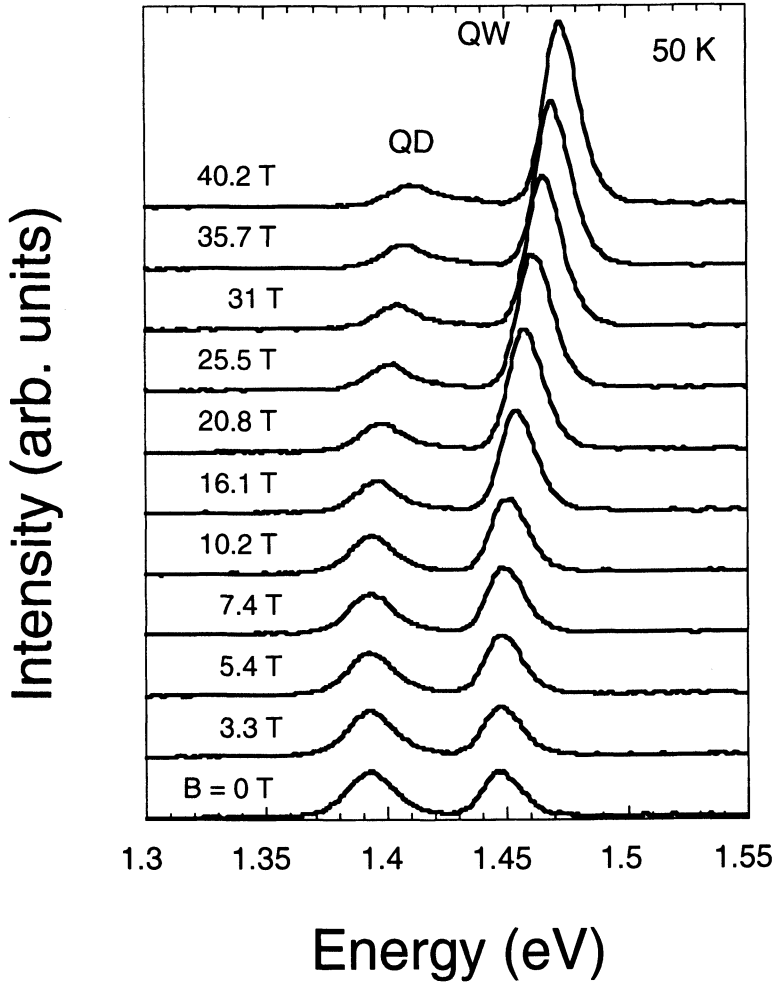


Figure 2.5: PL spectra measured at various magnetic fields in a Faraday configuration.

The evaluation sample was designed so that the width of InGaAs QW, L_w , was 5 nm at the sidewalls. From the geometrical relationship, the height of the QD region, h_d , is 15 nm. The indium content of InGaAs is represented by the H_2 flow rate through the TMIn container and was 100 sccm. Figure 2.5 shows the PL spectra under the various magnetic fields in the Faraday configuration, where B is parallel to z , i.e., the $\langle 111 \rangle_B$ direction. Both QW

and QD peaks shift diamagnetically to the higher-energy side as the intensity of the magnetic field increases. It should be noted that the intensity of the QD peak decreases with an increased magnetic field, while the intensity of the QW peak increases. The decreased intensity of the QD peak should be related to the dynamics of the photogenerated carriers during the recombination process. We believe that most of the electrons and holes emitting the QD luminescence are generated in bulk GaAs and are moved to In-rich InGaAs QDs at the bottom by way of the In-deficient QWs at the sidewalls; that is, the QW luminescence competes with QD luminescence. When a magnetic field is applied, the cyclotron motion of the carriers occurs at the QWs, decreasing the nonradiative recombination in the QWs due to the shrinkage of the excitonic wavefunction. This leads to an enhanced QW PL intensity [19] and then a decreased QD luminescence intensity.

Figure 2.6 plots the energy shift of the two emission peaks measured as functions of the applied magnetic field up to 40 T. The inset in the figure is a magnification of the data at the lower magnetic field values of less than 10 T. Compared to the energy shifts of the QW PL peak, the QD PL energy shifts are clearly smaller in both Faraday and Voigt configurations. This proves that the carriers are apparently localized in the TSR QDs. Note that no significant differences in the QD energy shifts are observed between the Faraday and Voigt configurations. This means that the carriers, or excitons, are isotropically confined in the TSR QD of this design.

In contrast, the QW energy shifts are larger in the Voigt configuration than in the Faraday, especially in the high-magnetic-field region where the energy shifts are dominated by the Landau shift. This finding is considered to be due to the two-dimensional (2D) characteristics of the exciton confined at the sidewall QWs. For QWs in general, a larger diamagnetic shift is observed when the magnetic field is perpendicular to the heterointerfaces [20,21], because the motion of the carriers is free along the QW plane but is structurally restricted perpendicular to the interfaces. Since the three {111}A facets are steeply inclined, with an angle of 70.5° from the $\langle 111 \rangle$ B direction, the perpendicular component of the magnetic field to the QW plane is larger in the Voigt configuration than in the Faraday.

To estimate the extent of the wave function of the confined carriers, we analyzed the data of the diamagnetic shifts on the basis of the hydrogen-like exciton model [22]. From the QD energy shift data in Fig. 2.6, we can deduce that β is about $12 \mu\text{eV/T}^2$ when the field is lower than 10 T, where $l_c = 8.1 \text{ nm}$. Since the exciton in the QD of the above sample is almost isotropic, the Bohr radius is estimated to be about 5.8 nm . Here, we used a GaAs dielectric constant $\epsilon = 13.1\epsilon_0$ instead of the value for the

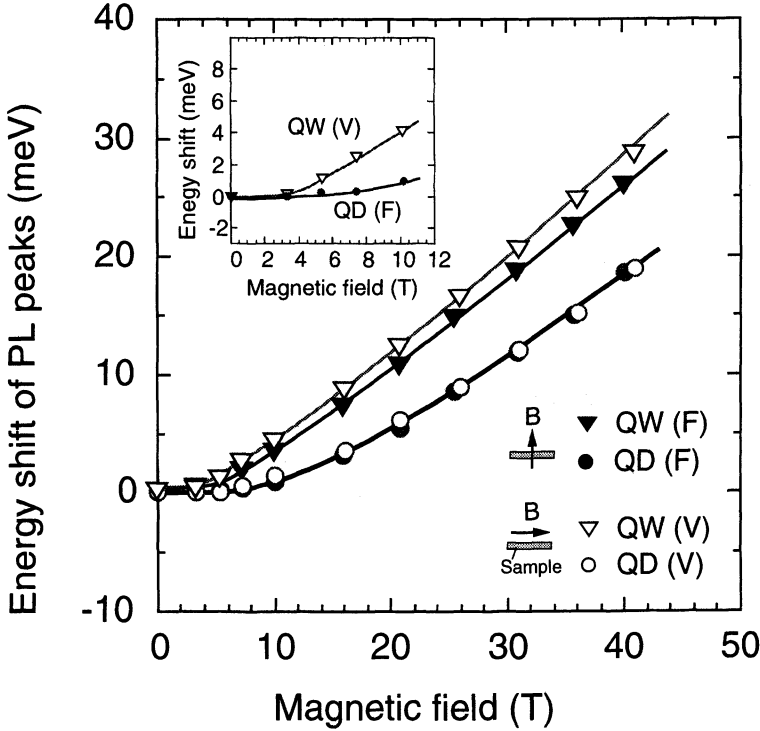


Figure 2.6: Energy shift of PL peaks as a function of applied magnetic field. F and V represent the Faraday and Voigt configurations, respectively.

InGaAs of the uncertain In content, where ϵ_0 is the dielectric constant in a vacuum. We believe that the error is within several percent because the In content of InGaAs is small. To obtain an accurate Bohr radius and the anisotropy of excitons at the sidewall QWs, we must separate the applied magnetic field into two components, namely, perpendicular and parallel to the QW plane. For simplicity, however, here we calculated the effective exciton Bohr radius a_{eff}^* from the low-field data in the Voigt configuration. The obtained effective two-dimensional (2D) exciton radius, a_{eff}^* , is 8.6 nm at the sidewall InGaAs well from the β value of $41 \mu\text{eV}/\text{T}^2$. It is reasonable that the Bohr radius of the exciton in a typical TSR QD is smaller than the effective exciton radius of the sidewall QW, and also much smaller than the calculated three-dimensional (3D) exciton radius of 11.8 nm in bulk GaAs [23]. Thus, the magneto-PL experiments confirmed that 3D carrier confinement takes place in the TSR QD.

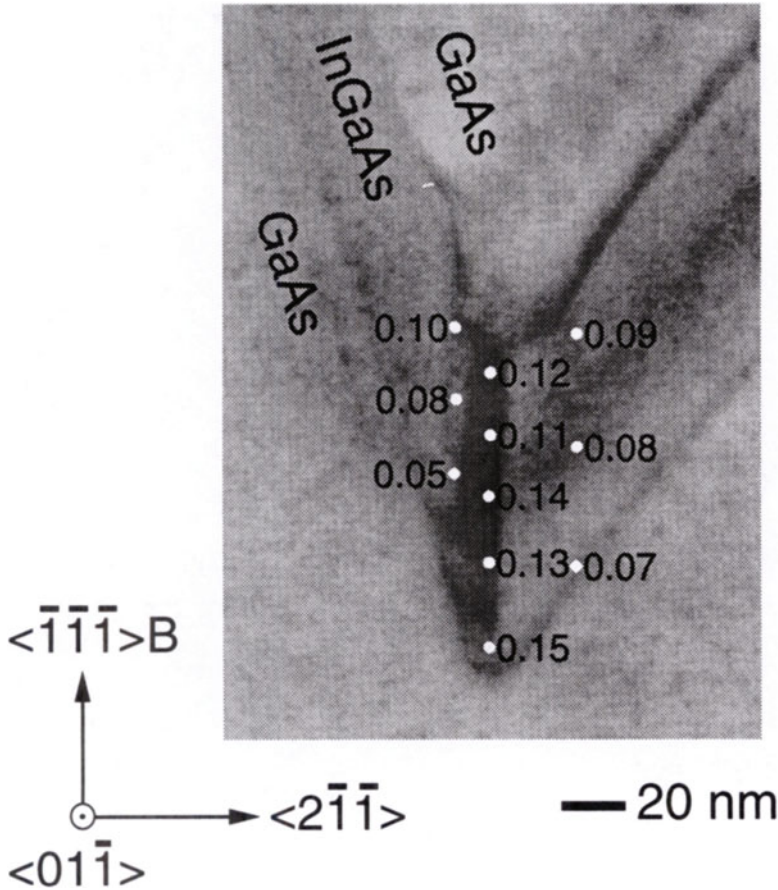


Figure 2.7: Cross-sectional TEM image of a sample with rather thick InGaAs. In contents of InGaAs near the TSR bottom are shown.

2.2.3 QD growth mechanism

To investigate the QD growth mechanism, we measured the compositions of InGaAs layers. Figure 2.7 shows the (01 $\bar{1}$) cross-sectional TEM photograph and the indium (In) composition near the bottom of TSR, where a GaAs/InGaAs/GaAs heterostructure was grown. The composition was evaluated by energy-dispersive X-ray diffraction (EDX) equipped in the TEM apparatus. The spatial resolution of the incident electron beam for EDX analysis was about 1 to 2 nm. In this sample, the L_w of the InGaAs

layer at the $\{111\}$ A facets was intended to be as thick as 50 nm, much thicker than that of the normal QD samples, to facilitate observation of the spatial distribution of the composition of the ternary alloy.

We found that a pillar-shaped nanostructure with a dark contrast grew from the TSR bottom of (111)B orientation and that the region has an In-rich composition. It is difficult to accurately estimate the absolute value of the In composition because the TSR bottom is surrounded by bulk GaAs in the TEM specimen. It should therefore be noted that the observed In compositions are relative values. Since the width, or the diameter, of the dark contrast region is seen to be 10 to 15 nm in the figure, it is small enough to accomplish the quantum confinement of carriers in the lateral direction. Therefore, the In-rich region will act as a QD when its height is within the size causing a quantum-confined effect. Thus, the growth mechanism of TSR-QD formation is not due to strain-induced 3D growth, such as the Stranski–Krastanov (SK) mode, but to the compositional modulation in the TSR. The In-rich InGaAs QD is spontaneously formed at the TSR bottom, while the In-poor InGaAs QW is formed at the $\{111\}$ A sidewall. From structural analysis, this result agrees well with the PL and CL results in Figs. 2.3 and 2.4.

Figure 2.8 is an illustration of an ideal crystallographic structure of a TSR, as viewed from the top (a) and from the $(0\bar{1}\bar{1})$ cross section (b). It is assumed that the GaAs surface is As terminated. There are three kinds of segments inside the TSR that have specific configurations in their surface bonds. The three equivalent sidewalls are $\{111\}$ A. The three valleys, crossing the two $\{111\}$ A faces, have the $\{100\}$ -like bond configuration. The bottom, crossing the three $\{111\}$ A faces, has the (111)B-like bond configuration. We speculated that on growing InGaAs in a TSR using MOCVD, the chemical composition of the InGaAs layer will be strongly affected by these local crystallographic orientations, leading to the different surface reaction kinetics.

To confirm this hypothesis, we studied the composition of InGaAs epilayers grown separately on planar (100)-, (111)A-, and (111)B-oriented GaAs substrates. The growth conditions were the same as those employed for the TSR QDs shown in Fig. 2.7. The thickness of the InGaAs growth varied from 70 nm to 100 nm. Table 2.1 summarizes the In composition of the InGaAs epitaxial layers analyzed by Auger electron microscopy (AES) and the surface morphologies. Results clearly indicate that the In composition is largest on the (111)B substrate. Although there is only a small change in the composition between the (100) and (111)A substrates, it is unclear whether the discrepancy is significant. It is obvious that the enhanced In composition at (111)B is caused by the different InGaAs growth kinetics among these orientations, because the surface reactions

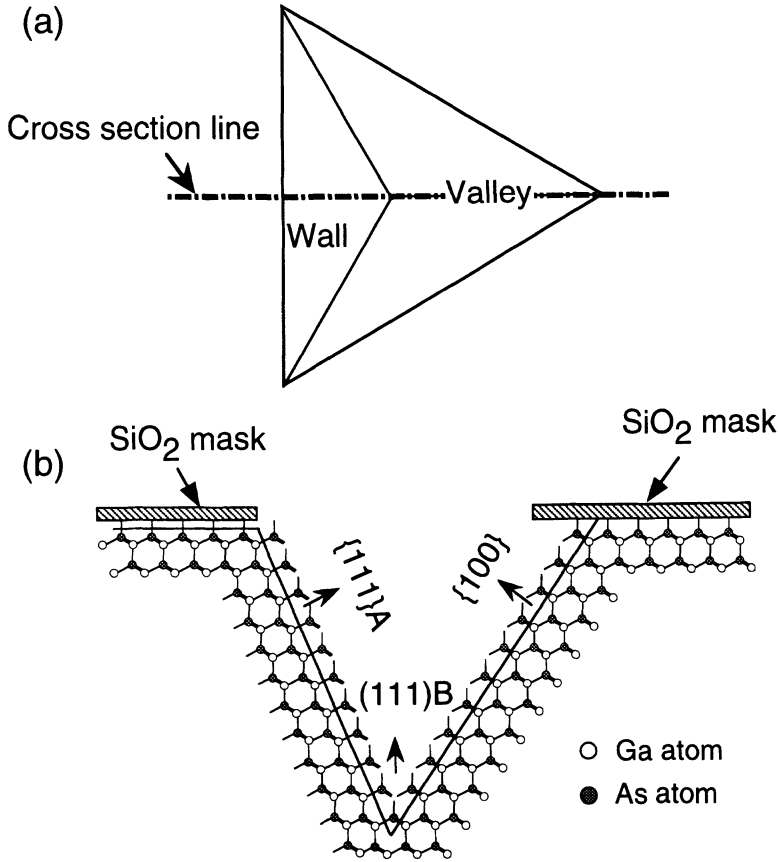
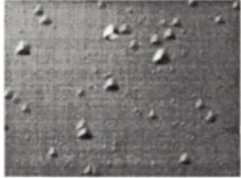


Figure 2.8: Illustration of crystallographic structure for TSR. (a) A top view; (b) a ball-and-stick model of the cross section when TSR is cleaved at the cross-section line.

play a crucial role in MOCVD. According to the similar experimental results reported in the literature [23,24], the In-rich composition of InGaAs layer on (111)B is due to the incorporation of the suppressed Ga atom into the surface, rather than of the In atom. It is also reported that the probability of In incorporation is not significantly affected by the surface orientations [23]. In other reports, the inhibition of Ga adsorption on (111)B GaAs has been observed during the GaAs homoepitaxial growth in both molecular beam epitaxy (MBE) under the As-rich growth conditions and MOCVD [25,26]. It has been interpreted that the excessively adsorbed As atoms

Table 2.1: Comparison of incomposition and surface morphology among InGaAs films grown on the substrate of different orientations.

Surface index	X_{In}	$\text{In}_x\text{Ga}_{1-x}\text{As}$ morphology
(100)	0.12*	
(111)A	0.14	
(111)B	0.36	

*Calibrated by XRD.

form chemically stable trimer structures on the GaAs (111)B, and that this specific As-rich reconstructed surface, recognized as (2×2) reconstruction in the case of MBE, deactivates the growth reactions between the surface As and the incoming Ga species [27]. This phenomena might be related to the growth of In-rich InGaAs at (111)B GaAs, as observed in our experiments.

Returning to the illustration of Fig. 2.8, we can apply the results of Table 2.1 to the self-forming mechanism of TSR QDs. As we explained earlier, the TSR bottom has a (111)B-like bond configuration. Therefore, when growing InGaAs, an In-rich region is grown spontaneously at the bottom. Unlike the ideal picture shown in Fig. 2.8, the bottom of the TSR is actually not acute in an atomic dimension, but will have a finite flat area of (111)B orientation under the growth conditions. This area will limit the lateral size of the In-rich QD. On the basis of this growth mechanism, the In-rich region becomes a QD when we grow thin InGaAs. In contrast, when

we grow a rather thick InGaAs layer, the In-rich region is automatically grown upwards and comes to be a vertical quantum dash or a wire, just as is shown in Fig. 2.7. Thus, we can control the vertical QD size through the thickness of the InGaAs growth.

2.2.4 Growth control of TSR QDs

Unlike the self-assembled QD technique we describe later, we are able to control the location of QDs by precisely positioning the mask openings by conventional photo- or electron beam lithography. To apply the QDs to some electron devices, it should be necessary for us to control the strength of zero-dimensional quantum confinement through regulating the size and the chemical composition of TSR QDs, and to develop the stacked structure of these quantum dots. In this section, we show that the TSR is a versatile method for achieving the flexible control of QDs.

Figure 2.9 shows the change of PL energies of the QD ground state and the QW at 77 K as functions of the TMIn flow rates. In this case, the nominal InGaAs thickness L_w was fixed at 5 nm. Both peaks shift to the lower energy side with increasing TMIn flow rates owing to the increased In composition in both QDs and QWs. This result indicates that we can control the energy levels of QDs by changing the gas-phase composition during InGaAs growth. The energy difference, ΔE , between QW and QD changes from 55 meV to 100 meV by doubling the TMIn flow rate from 100 sccm to 200 sccm. This means that the incorporation of the In atoms is more enhanced at the QDs than at the QWs, a phenomenon that is probably related to the different InGaAs growth kinetics between the (111)B bottom and {111}A side faces. Note that the In-rich TSR-QD is sandwiched perpendicularly by GaAs and is horizontally surrounded by the In-poorer InGaAs QWs. Therefore, the result implies that increasing the In content of the InGaAs layer enhances the zero-dimensional carrier confinement in the TSR QDs not only vertically but also laterally.

Next, we changed the thickness of the InGaAs growth, maintaining the TMIn flow rate at 100 sccm. We can control the size of the In-rich dot vertically by just changing the time of the InGaAs growth. From the geometrical relationship, the height of the TSR QD, h_d , is given by

$$h_d = \frac{L_w}{\cos\theta} = 3L_w, \quad (2.1)$$

where L_w is the width of the InGaAs QW at {111}A, and θ (the angle between the [111]A vector and the $\langle 111 \rangle$ B vector) is about 70.5° . Therefore, in the case of thick L_w , the In-rich region at the bottom will become an

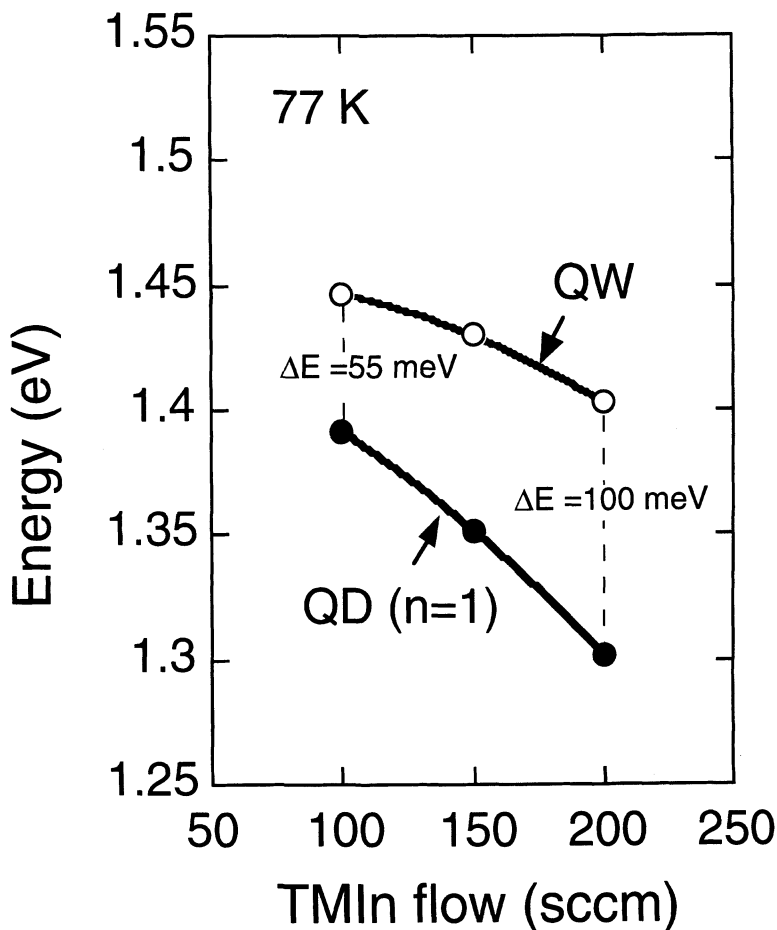


Figure 2.9: PL peak emission energies at 77 K plotted as a function of TMIn flow rates during the growth.

elongated QD—in other words, a quantum dash or a quantum wire. Figure 2.10 shows the variation of PL energy positions of the QDs' ground state and the QWs at 77 K as a function of quantum width L_w . As expected, the PL energy of the QWs changes toward the lower-energy side with increasing L_w , and the quantum size effect is observed when L_w is less than 15 nm. Both QDs and QWs show a similar PL energy dependence on L_w ($= 1/3 \hbar a$). Thus, the energy level of the QD ground state can be artificially controlled by changing the vertical length of the TSR QDs.

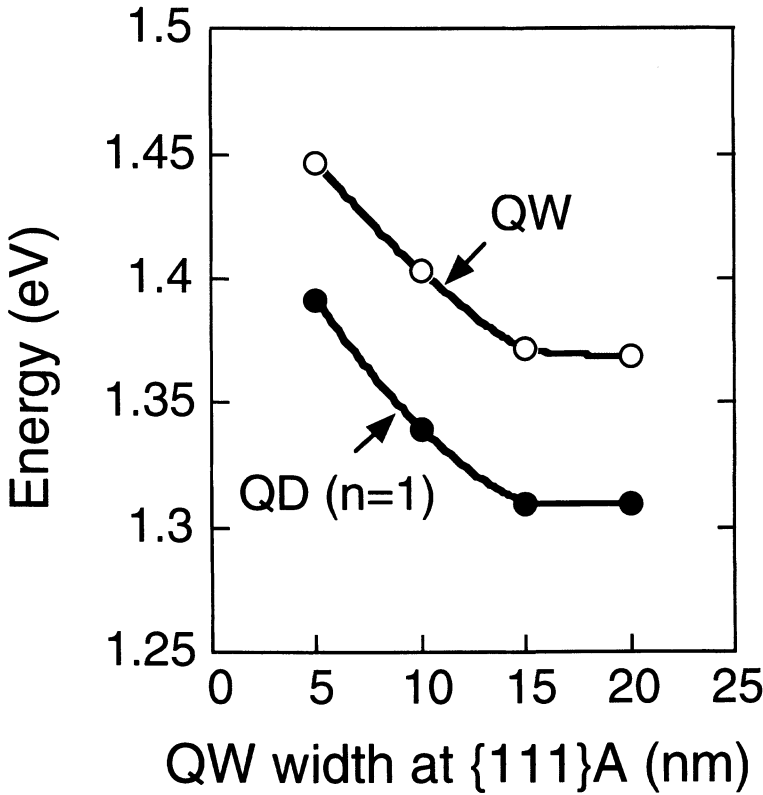


Figure 2.10: PL peak emission energies at 77K plotted as a function of InGaAs growth thickness on {111}A facets.

As stated above, when growing quantum structures inside the TSRs, we can independently control the height and indium content of the dot. This is one of the big advantages of the TSR structures, compared with the self-assembled technique governed by the SK growth mode, where there is an inseparable correlation between dot size and composition [28].

Another important feature is the capability of easy stacking of the dots. Stacking is attractive, in particular, from the standpoint of device applications, because it enables us not only to increase the density of the dots but also to create some quantum mechanical tunneling structures through the QDs [29–31]. Figure 2.11 shows a cross-sectional TEM image of double-stacked TSR QDs. We grew 20 nm-thick double InGaAs layers under a TMIn flow rate of 100 sccm on both sides of a 30 nm-thick GaAs intermediate layer so that we could easily observe the grown structures. The upper dot

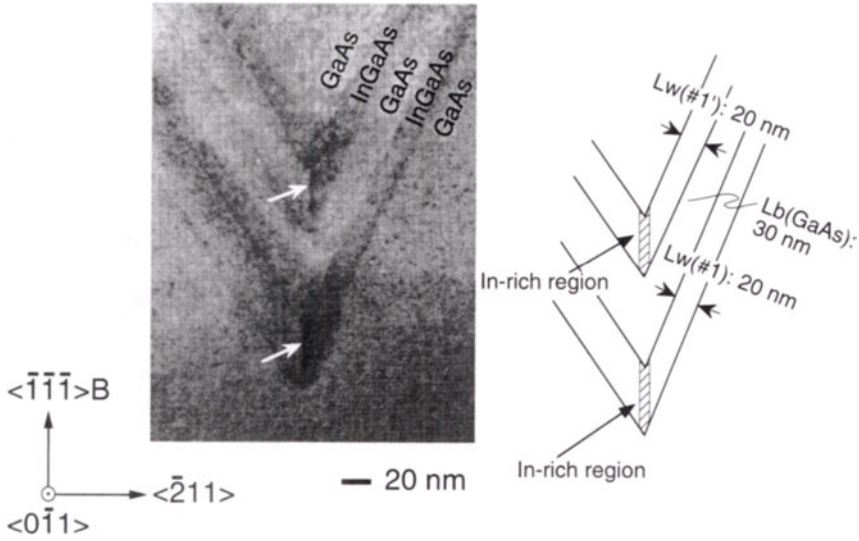


Figure 2.11: Cross-sectional TEM image of doubly stacked TSR QDs. The position of the dots is shown by white arrows.

is precisely aligned over the lower one with the same size, showing that the TSR technology is suitable for realizing such a structure.

We studied the stacked structures using conventional PL techniques. Figure 2.12a shows the 77 K PL spectra obtained from three structures: a single size of TSR QDs named A ($L_w = 5$ nm), a stack of identical dots AA, and a mixed stack of TSR QDs of two different sizes AB ($L_w = 5$ nm and 10 nm). Sample AA has almost the same spectral shape as A. In contrast, sample AB shows two distinct QD peaks. Since the position of the higher-energy dot peak coincides with that of A, the lower peak is from the larger dot. In addition, we can observe the weak QW peak in the AB structure at the energy side lower than that of A. This is due to the predominant recombination of photogenerated carriers at the 10 nm-thick QWs compared with the 5 nm-thick ones. These results clearly prove the excellent growth controllability of QDs stacked in TSR structures.

2.3 Self-assembled quantum dots in SK mode

This technique uses three-dimensional island growth on highly mismatched lattice semiconductor substrates. The growth of InAs on a GaAs

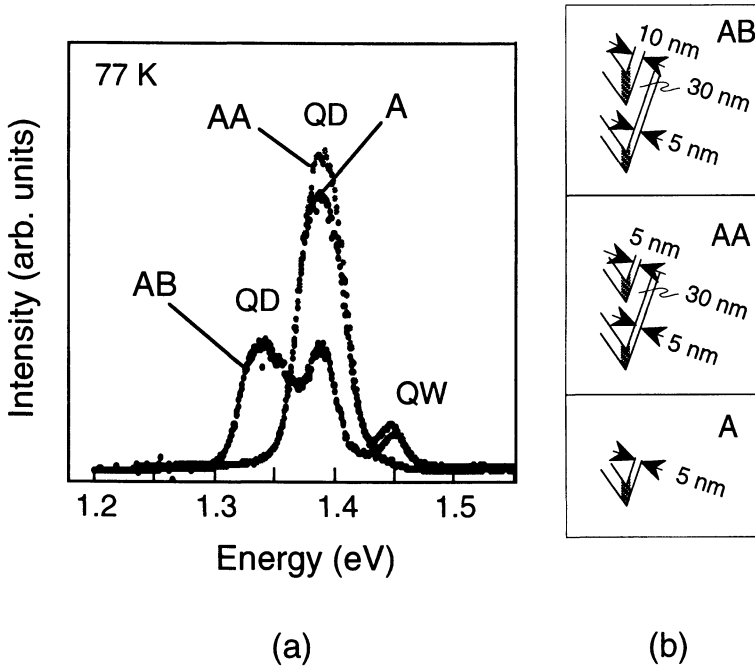


Figure 2.12: (a) PL spectra from three kinds of structures grown on TSRs; (b) Illustration of the corresponding growth structures.

substrate is a typical example, where the lattice mismatch between InAs and GaAs is about 7%; dislocation-free high-density coherent islands of InAs are self-assembled on the GaAs substrate, accompanied with a wetting layer. Typical InAs self-assembled islands have a dome or pyramid shape with a base length of about 20 nm and a height of a few nanometers. Since the exciton Bohr radius is 10 to 20 nm in an InAs–GaAs system, the island size is small enough for us to observe three-dimensional quantum confinement effect. In fact, separated photoemission spectra from discrete energy states have been observed [32,33].

Though self-assembling quantum dots are a new and exciting system, the growth process itself is not new but is already known as the Stranski–Krastanov (SK) mode growth [34]. InGaAs/GaAs islands grown via SK mode were observed and evaluated by several groups in the mid-1980s [35–38]. However, this research did not attract broad interest, at least from the standpoint of quantum dots. That the SK InAs islands on a GaAs substrate might work as quantum dots was proposed by Tabuchi et al. [39]

in 1992. Since then, many studies have identified the islands as quantum dots with 3D quantum confinement, primarily based on their optical emission properties [32,40–43]. The 3D island growth, known as a technique of self-assembled growth of quantum dots, is now the center of interest in the semiconductor material and device research field.

In addition to InGaAs quantum dots, a wide variety of quantum dots consisting of other semiconductor materials have also been fabricated through the self-assembling processes, including InGaP [44,45], CdSe [46], and GaN [47,48]. In the near future, these new material categories are expected to contribute to the improvement of many optical devices, such as many aspects of semiconductor lasers, and in many wavelengths ranging from infrared to red to blue.

This section focuses on the molecular beam epitaxial (MBE) growth of InAs islands on the GaAs substrate.

2.3.1 InAs island growth

When InAs is supplied to a GaAs substrate, three-dimensional island growth occurs followed by the two-dimensional growth of a wetting layer. In the case of MBE growth, the reflection high-energy electron diffraction (RHEED) pattern gives us important information about the surface during growth; we can observe streak reflection patterns as the layer-by-layer two-dimensional growth proceeds. The transition from streaks to spots in the RHEED pattern occurs when three-dimensional islands start to grow. After the growth, we use atomic force microscopy (AFM) to learn the shape, size, density, alignment, and distance of neighboring islands. We can get similar information from plan-view transmission electron microscopy (TEM) even when islands are covered by overgrown semiconductors. Defects such as dislocations and stacking faults are also evaluated. The cross-sectional-view TEM photograph is quite useful for observing how perpendicularly stacked quantum dots interact with each other, as will be seen in the following sections.

For the growth of InAs islands, a conventional MBE was used with metallic In, Ga, Al, and As₄ as source materials. The substrates were (001)-oriented GaAs. Before growth, the substrates were thermally cleaned at about 680°C for 1 minute under an arsenic pressure of 1.2×10^{-5} Torr. InAs islands were grown on a GaAs (100 nm)/AlGaAs (50 nm)/GaAs (400 nm) buffer layer and were covered with a GaAs (50 nm)/AlGaAs (50 nm)/GaAs (100 nm) cap layer. The substrate temperature for growth of the InAs was varied between 475 and 560°C, and was 650°C for the growth of the buffer layer. The temperatures were monitored by a pyrometer calibrated by the melting temperature of In-Al alloys. The growth rates of InAs

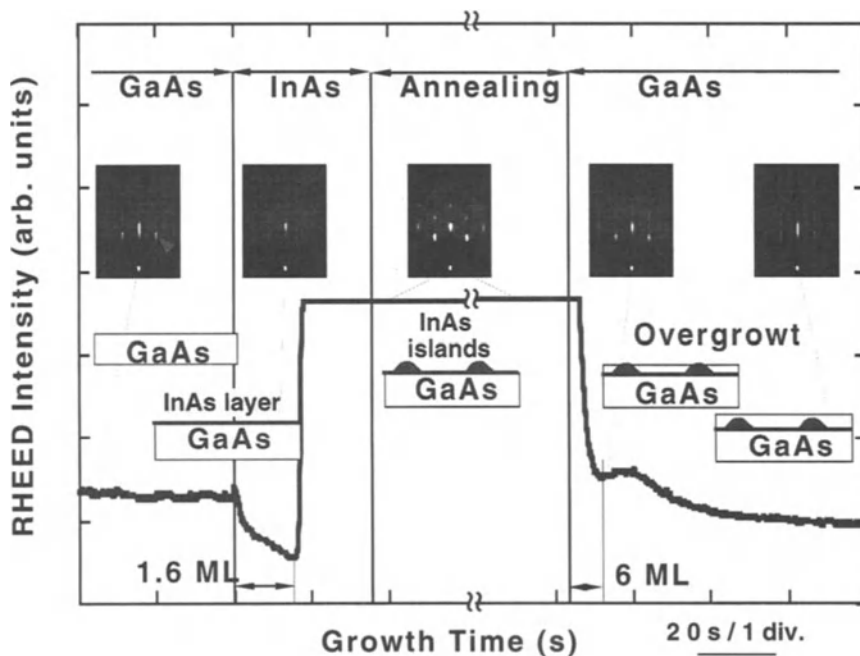


Figure 2.13: Variation of RHEED pattern and the reflection intensity for the various stages of growth of InAs islands.
(From SEMICONDUCTORS AND SEMIMETALS, Vol. 60, Nakata et al., (1999), Copyright 1999 by the Academic Press)

and GaAs were approximately 0.1 and 0.75 $\mu\text{m/h}$, i.e., it takes about seconds to grow one monolayer (ML) InAs and about 1.5 seconds to grow one ML in GaAs. The arsenic pressure used for InAs dot growth was 6×10^{-6} Torr and was 1.2×10^{-5} Torr for buffer-layer growth. Two arsenic cells were used, both set at 6×10^{-6} Torr. The arsenic pressure was changed abruptly at the interface by switching one of these cells on and off. After a 60-second annealing for the InAs islands, the GaAs layers were grown.

Figure 2.13 shows the variations in the RHEED pattern and the reflection intensities for the various stages of growth. During the growth of the buffer layer, the surface was As stabilized with 2×4 reconstruction, then transited to $c(4 \times 4)$ reconstruction while the substrate temperature was dropped to 510°C. The photograph on the left in Fig. 2.13 is the [110] azimuthal RHEED pattern obtained before InAs growth. The half-order fractional reflections were clearly observed. Just after the start of InAs growth, the fractional order reflections disappeared immediately, and the

transition from streaks to spots started as the growth of InAs reached around 1.6ML at 16 seconds, indicating a transition from two-dimensional growth to three-dimensional island growth. The reflection intensity reached an almost constant value at a 1.8ML InAs supply, and the reflection pattern did not change during the annealing afterwards, showing that the islands are stable during the interruption in growth. When GaAs supply started to overgrow InAs as a cap layer, the spot intensity rapidly decreased. The pattern returns to streaky at the 6ML GaAs growth, showing that 6ML GaAs completely covers the InAs islands and almost flattens the surface.

The island size and density were evaluated by AFM in the air. Figure 2.14 shows AFM images of the islanding surfaces when the InAs supply was 1.3, 1.6, 2.1, and 2.6ML. These images refer to different epilayers grown under the same conditions. The scanned area is $0.5 \times 0.5 \mu\text{m}^2$. The islanding growth starts at the 1.6ML supply, just as is seen in the change of the RHEED pattern. As the supply amount increases, the dot density rapidly increases and the dots begin to combine with each other. Figure 2.15 shows the in-plane diameter and the island density as a function of the nominal thickness of the InAs. As depicted in the figure, the growth proceeds from the two-dimensional growth of a wetting layer, the nucleation of the InAs islands at about a 1.6ML supply, and increase in the density to the coalescence of neighboring islands at a 2.2–2.3ML supply. Figure 2.16 shows the diameter and density of islands at a 2.1ML InAs supply as a function of the growth temperature, and the AFM images of the islands at each temperature. As the temperature increases from 470 to 560°C, the island diameter increases from 15 to 45nm, with a corresponding decrease in density. This temperature dependence is caused by the surface migration length of In atoms, which increases with an increase in temperature.

2.3.2 Multiple-layer growth and perpendicular alignment of islands

The InAs islands, buried in GaAs, work as quantum dots due to their size, which is smaller than the exciton wave function extent both in-plane and perpendicularly. When we use these quantum dots in laser diodes, multiple-layer stacking is required to increase the quantum dot density to obtain enough interaction between the confined electrons and the electromagnetic field. Figure 2.17 shows cross-sectional TEM photographs taken for multiple-stacked quantum dot layers with an intermediate GaAs layer thickness of 10nm and 20nm. For multiple-layer stacking, the growth sequence of Fig. 2.13 was repeated. The InAs islands were annealed for

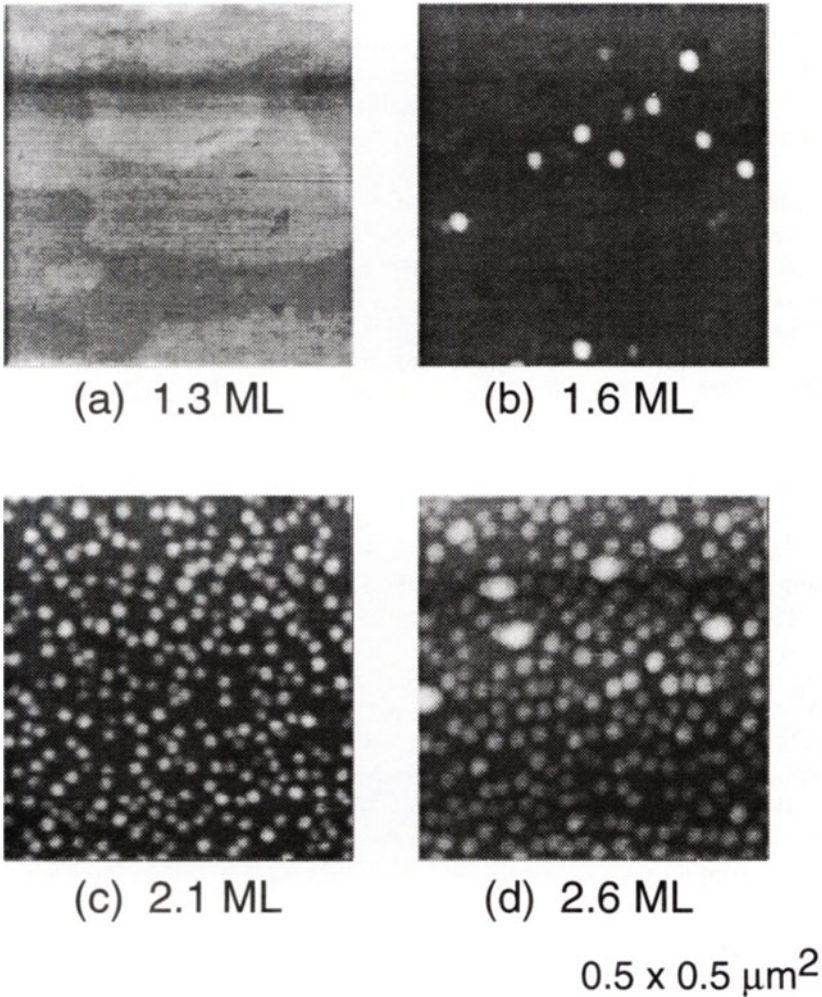


Figure 2.14: AFM images of the islanding surfaces taken for InAs supplies of 1.3, 1.6, 2.1, and 2.6 MLs.

(From SEMICONDUCTORS AND SEMIMETALS, Vol. 60, Nakata et al., (1999), Copyright 1999 by the Academic Press)

60 seconds. These intermediate layer thicknesses are thicker than the height of the InAs islands by about 3 to 5 nm, meaning that the islands are embedded in the intermediate layers.

As seen in Fig. 2.17, the island formation occurs independently between layers of the 20 nm intermediate GaAs layer structure. AFM measurements

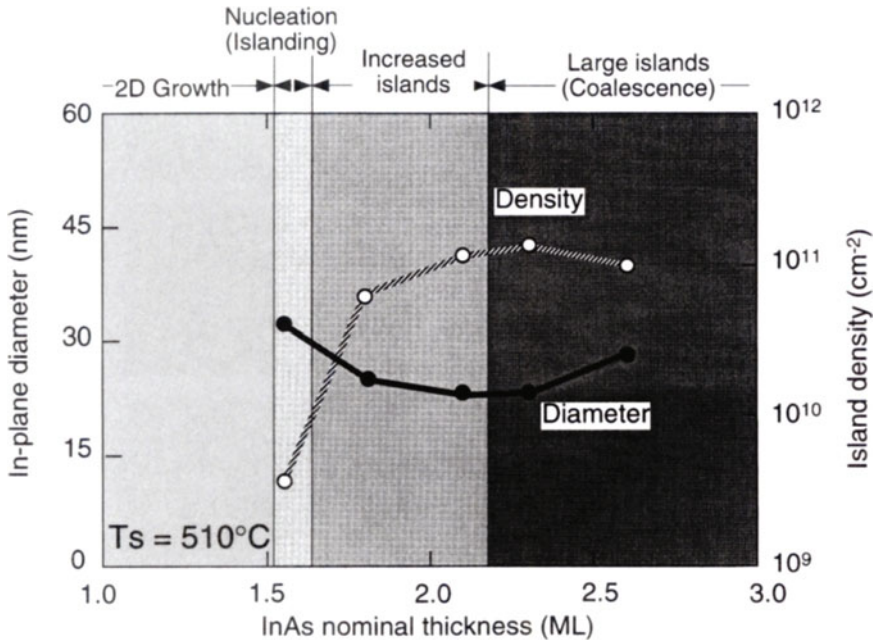


Figure 2.15: In-plane diameter and the density of islands as a function of the InAs nominal thickness.

(From SEMICONDUCTORS AND SEMIMETALS, Vol. 60, Nakata et al., (1999), Copyright 1999 by the Academic Press)

showed that the size and density of the islands were almost constant between layers. When the intermediate layer was reduced to 10 nm, the islands were found to align perpendicularly, and the size increased towards the upper layers. This perpendicular alignment of the islands is considered to be due to the strain fields induced by the islands. Xie [49] has provided an analytical description of correlated island formation in the direction of growth under strain fields.

The size and density of the islands vary as the multiple-layer stacking proceeds, as seen in Fig. 2.17a. When we grew 1.8 ML InAs islands using 10 nm-thick intermediate GaAs layers, the average diameter of the tenth-layer islands was about 45 nm, 190% larger than the first-layer islands with a 24 nm diameter. The island density decreased from $1 \times 10^{11} \text{ cm}^{-2}$ for the first-layer islands to $3 \times 10^{10} \text{ cm}^{-2}$ for the tenth-layer islands (Fig. 2.18). The same phenomena have been reported by Solomon et al. [50,51].

The increase in island size and the decrease in density can be explained as follows. The size distribution of self-assembled islands leads to spatial fluctuations of the strain fields at the GaAs surface. Supplied indium for

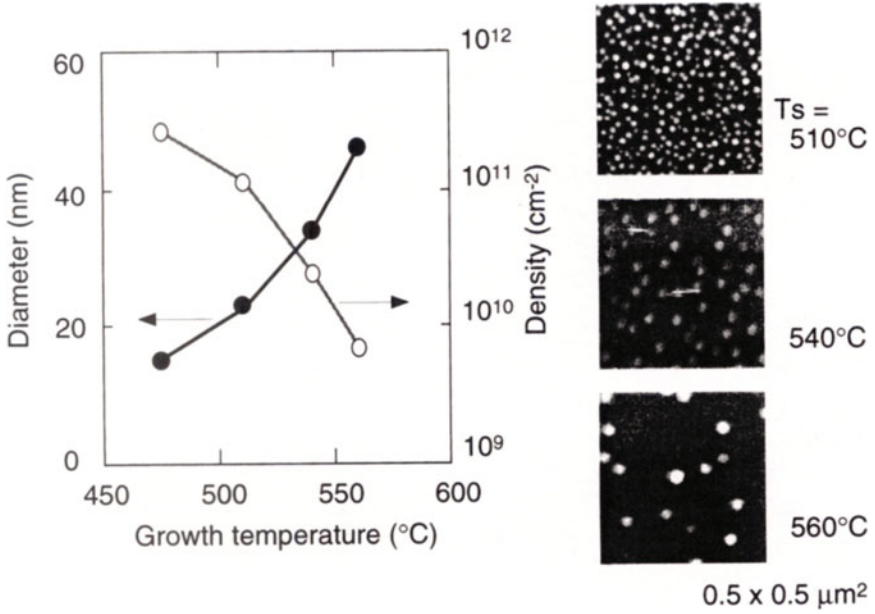


Figure 2.16: Diameter and density of islands at 2.1 ML InAs supply as a function of the growth temperature together with AFM images of islands at each temperature.

(From SEMICONDUCTORS AND SEMIMETALS, Vol. 60, Nakata et al., (1999), Copyright 1999 by the Academic Press)

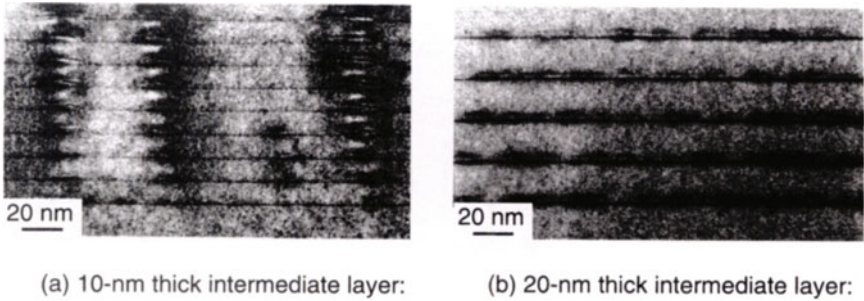


Figure 2.17: Cross-sectional TEM photographs when the intermediate GaAs layer thickness was (a) 10 nm and (b) 20 nm.

(From SEMICONDUCTORS AND SEMIMETALS, Vol. 60, Nakata et al., (1999), Copyright 1999 by the Academic Press)

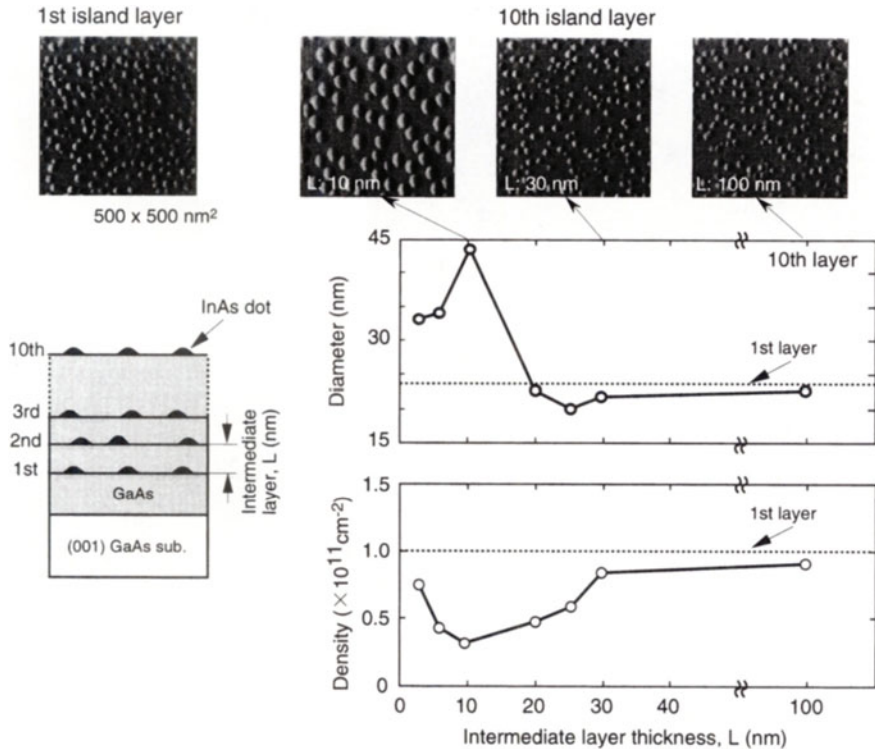


Figure 2.18: In-plane diameter and the density normalized against the first-island-layer density measured as a function of the intermediate layer thickness. AFM images for each sample are shown in the insets.

(From SEMICONDUCTORS AND SEMIMETALS, Vol. 60, Nakata et al., (1999), Copyright 1999 by the Academic Press)

the growth of the next island layer preferentially accumulates at sites with larger strain fields. Because of a smaller lattice mismatch between InAs and the surface, the island size becomes large. As a result, large islands are formed preferentially in the region of the large strain fields, preventing island formation in the region of the small strain fields. This explains the increase in island size and the decrease in density in multiple islanding growth with an intermediate layer of about 10 nm.

As shown in Fig. 2.18, as we further reduce the thickness of the intermediate layer below 10 nm, the increase in size and the decrease in density becomes less remarkable. For example, when we grew multiple layers with a 3 nm intermediate layer, the average diameter of the tenth-layer islands

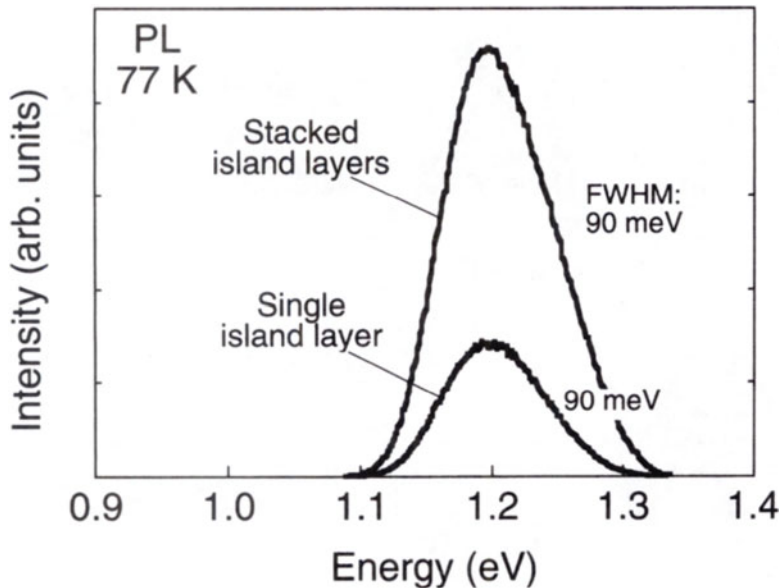


Figure 2.19: Photoluminescence spectra from single-layer and multiple-layer samples with 20 nm intermediate layers at 77 K.

(From SEMICONDUCTORS AND SEMIMETALS, Vol. 60, Nakata et al., (1999), Copyright 1999 by the Academic Press)

was about 33 nm, 140% larger than the first-layer islands with a 24 nm diameter. The island density decreases from $1 \times 10^{11} \text{ cm}^{-2}$ for the first-layer islands to $8 \times 10^{10} \text{ cm}^{-2}$ for the tenth-layer islands. This is because, due to a thinner intermediate layer, the strain field over the small size dots becomes large enough to accumulate the supplied InAs.

Photoluminescence spectra from single-layer and multiple-layer samples with 20 nm intermediate layers at 77 K are shown in Fig. 2.19. The sample was irradiated with an Ar^+ ion laser at a power of 1 mW. Luminescence was dispersed by a monochromator and detected by a cooled Ge detector. The laser spot was about $100 \mu\text{m}$ in diameter. The emission spectrum appears at around 1.2 eV, with the full width at half maximum (FWHM) of 90 meV for both samples. This large spectrum width is typically observed in self-assembled islands thus far reported. This inhomogeneous broadening is caused by the fluctuation of the quantized energies among the islands included in the measured area (10^6 to 10^7 islands). By using microprobe photoluminescence to access a limited number of islands, several researchers have observed a sharp emission spectrum with around $100 \mu\text{eV}$ [41,52–56].

The broadening of luminescence should be dominated by fluctuations in height, since the SK growth island has a height of 3 to 5 nm, much smaller than the diameter of 15 to 45 nm. This is understood by the well-known concept that a change in quantized energy is strongly influenced by a constant size fluctuation, as the size of confinement region decreases. Thus, if the height can be increased or, more accurately, controlled, the emission spectrum broadening will be greatly reduced.

2.3.3 Closely stacked InAs quantum dots

The perpendicular alignment of the islands enables the vertical coupling of the islands electrically as we reduce the intermediate GaAs thickness to the extent that the electron wave functions of neighboring wells are overlapped. This situation is just the same as the transition of multiple quantum wells to a superlattice when the barrier layer thickness is reduced. This vertical coupling enables an increase in the effective height of the quantum dots. It also enables electron tunneling between quantum dots, which might lead to novel electronic devices based on single-electron tunneling.

This section describes the growth process, the crystal structures, and the optical properties of perpendicularly stacked islands when the intermediate GaAs layer is reduced to a few nanometers close to or comparable to the height of InAs islands. Even under this condition, InAs islands are grown repeatedly. Optical diagnostics show that the stacking of the InAs islands increases the effective perpendicular size of the quantum dots due to electrical coupling, resulting in the narrowing of the spectrum FWHM to 25 meV. We call the stacked structure *closely stacked quantum dots*.

Figure 2.20 shows the growth sequence for the close stacking of InAs islands. The nominal InAs supply for island formation was fixed at about 1.8 ML, and the nominal thickness of GaAs intermediate layers was set at 2 nm and 3 nm. Prior to InAs island growth, the sample was annealed for 2 minutes. Following island growth, the sample was annealed for 1 minute. The growth rates, arsenic pressure, and growth temperature were the same as previously. Figure 2.21 shows the transition of the RHEED pattern intensity observed during the growth of InAs islands in the area indicated by the square in the inset patterns. The RHEED shows the streaky pattern for the two-dimensional growth in the early stage and varies to the spotty pattern for the three-dimensional islanding growth at a value above a critical supply amount. The SK islanding growth also occurs for the third and fifth layers. Note that the third- and fifth-layer islanding started when the growth of InAs reached about 10 ML, which is about 63% of the first-layer islanding. The reason for the smaller critical amount for the islanding is thought to be that the strain induced by the lower-layer islands accumu-

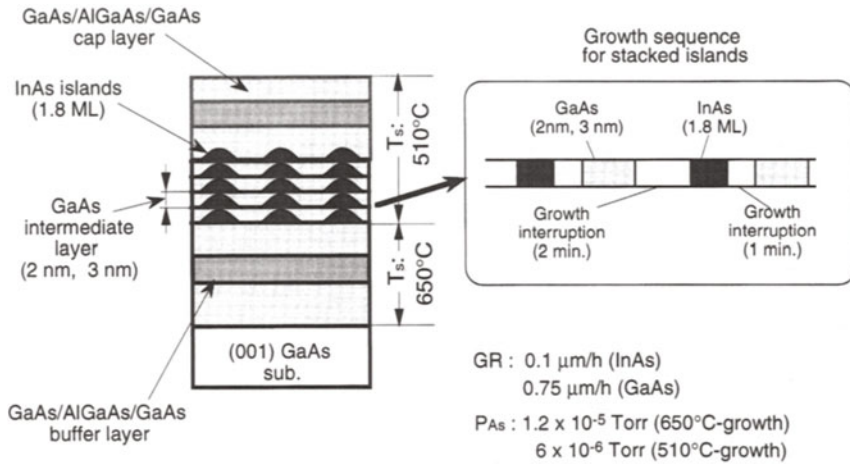


Figure 2.20: Growth sequence for the close stacking in InAs islands. The nominal InAs supply for island formation was fixed at about 1.8 ML, and the nominal thickness of GaAs intermediate layers was set at 2 and 3 nm. (From SEMICONDUCTORS AND SEMIMETALS, Vol. 60, Nakata et al., (1999), Copyright 1999 by the Academic Press)

lates InAs preferentially or that the segregation of the InAs atoms forms lower islands. The transition of the growth mode from two-dimensional growth to three-dimensional growth and the existence of wetting layers, as will be shown below in a TEM photograph, both indicated that SK growth islands were formed even on thin GaAs intermediate layers.

The island size and the density at each layer were evaluated by ex situ AFM. Figure 2.22 shows AFM images of the islanding surfaces at the (a) first, (b) third, (c) fifth, and (d) tenth layers stacked at 3 nm intervals. These images refer to different epilayers grown under the same conditions. The scanned area is $250 \times 250 \text{ nm}^2$. The upper layer islands expanded slightly as the number of stacked layers increased. Figure 2.23 shows the island-size dependence on the number of stacked layers. The average diameter of the tenth-layer islands was about 33 nm, which was 140% larger than that of the first-layer islands (about 24 nm in diameter). The island density decreases from $1 \times 10^{11} \text{ cm}^{-2}$ for the first-layer islands to $8 \times 10^{10} \text{ cm}^{-2}$ for the tenth-layer islands. The increase in size and the decrease in density can be explained by the concept of a strain field formed by the lower islands, as previously stated.

TEM photographs show the overall structural features. Figure 2.24a is a (110) cross-sectional TEM image of a five-stacked island structure grown

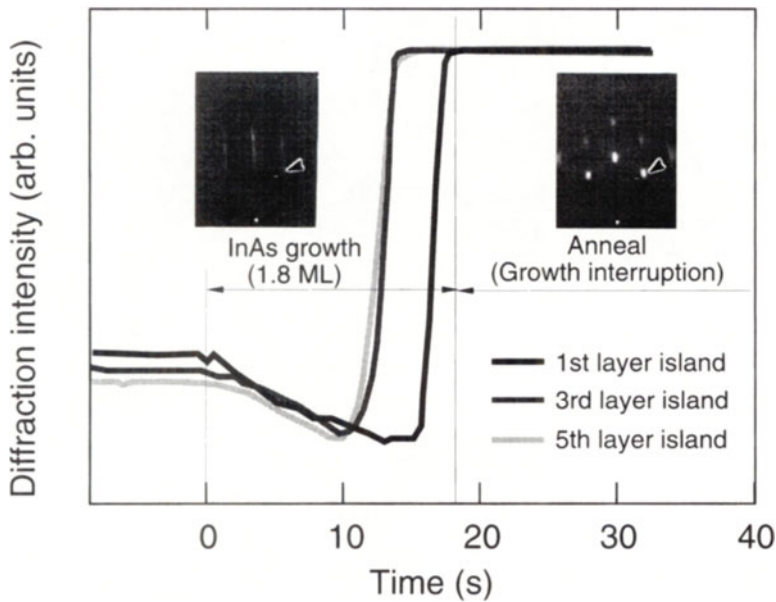


Figure 2.21: RHEED pattern intensity transition observed at the area indicated by arrowheads in the inset patterns during the growth in InAs islands.

(Reprinted from Journal of Crystal Growth, Vol. 175/176, Nakata et al., Copyright 1997, with permission for Elsevier Science)

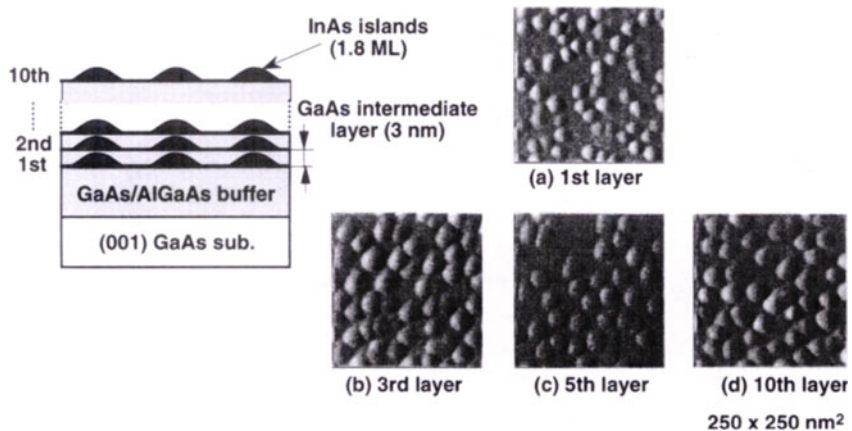


Figure 2.22: AFM images of the islanding surfaces at the (a) first, (b) the third, (c) the fifth, and (d) the tenth layers stacked at 3 nm intervals.

(Reprinted from Journal of Crystal Growth, Vol. 175/176, Nakata et al., Copyright 1997, with permission for Elsevier Science)

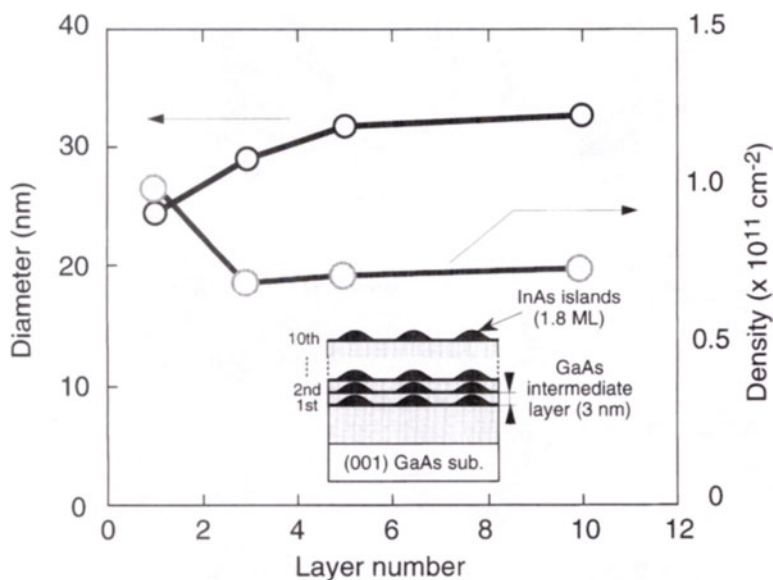


Figure 2.23: Island diameter and density plotted as functions of the number of stacked layers.

(Reprinted from Journal of Crystal Growth, Vol. 175/176, Nakata et al., Copyright 1997, with permission for Elsevier Science)

at 2 nm intervals. Each island layer was accompanied by a wetting layer, indicating that the upper-layer islands formed via SK growth, as did the first-layer islands. The upper-layer islands grew on just the lower-layer islands, aligning vertically. As a whole, these islands have a structure of about 22 nm in diameter and 13 nm in height, and appear in the image as dark megaphone-shaped strained regions. Figure 2.24b is a plane view of the TEM image obtained from a five-stacked island structure grown at 2 nm intervals. The density of islands was $8 \times 10^{10} \text{ cm}^{-2}$. This value agreed well with the AFM results. It should be noted that the size uniformity and lateral ordering are also improved compared to a single island layer. Although we attempted precise control in the vertical dimension of the islands, uniform lateral dimension was achieved in these closely stacked structures. A model to explain the formation process of this structure is illustrated in Fig. 2.25. First, the intermediate layer of GaAs overgrows away from the InAs islands. During annealing or interruption of growth after the GaAs overgrowth, InAs in the upper part of the islands above the GaAs overgrowth layers leaves the islands and regrows, forming parts of a wetting layer on GaAs. Then some of the InAs in the top regions of the remaining islands exchanges with GaAs to reduce the total energy. This

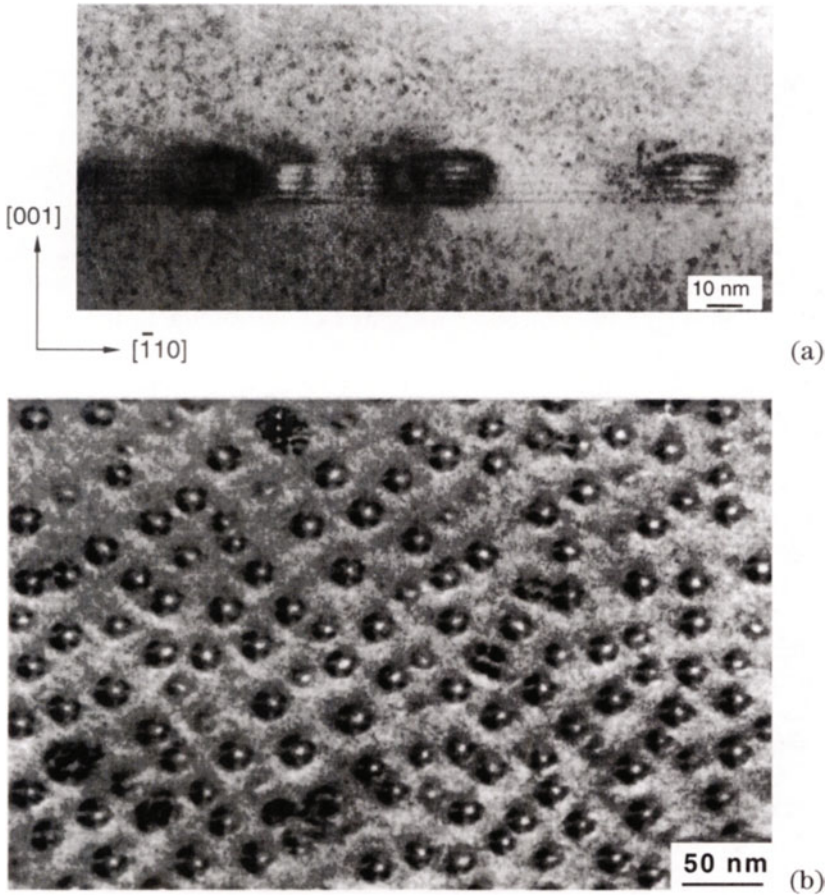


Figure 2.24: (a) (110) cross-sectional TEM image of a five-stacked island structure grown at 2 nm intervals; (b) plan view of TEM image obtained from a five-stacked island structure grown at 2 nm intervals.

(Reprinted from Journal of Crystal Growth, Vol. 175/176, Nakata et al., Copyright 1997, with permission for Elsevier Science)

leads to a decrease in the island height. This phenomenon enables us to control precisely the thickness of islands in the direction of growth and therefore the bottom diameter of the islands through the intermediate-layer thickness. The island-to-island distance of 3 to 4 ML is so thin that the electron wave functions of each island can be overlapped along the vertical direction. This finding suggests that the islands are electronically coupled and behave as a single quantum dot as a whole, which will be confirmed by the optical diagnostics as follows.

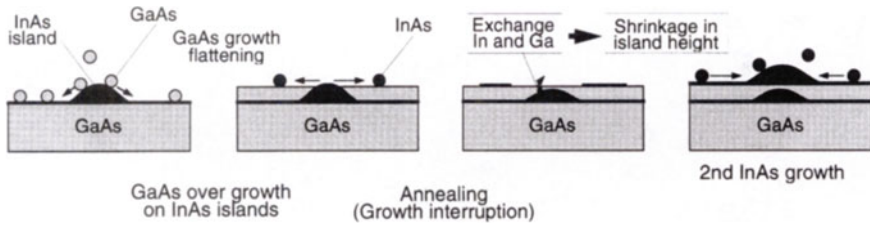


Figure 2.25: Illustration of growth process of stacked islands. The height of the InAs island after GaAs overgrowth decreases to the thickness of the interval layer during the annealing. SK growth then occurs on top of the lower island.

(From SEMICONDUCTORS AND SEMIMETALS, Vol. 60, Nakata et al., (1999), Copyright 1999 by the Academic Press)

The light-emission properties of these structures were evaluated using a photoluminescence technique with the same measurement conditions. Figure 2.26a shows the emission spectra of a single island layer at 77 K and closely stacked 3-, 5-, and 10-layer island structures grown with 3 nm-thick intermediate layers. The peak energy shifted lower as the number of stacked layers increased (about 90 meV in the five-layer islands). At the same time, the broad emission spectrum of the single island layer drastically narrowed at the five-stacked layer to an FWHM of 27 meV, about one third of that obtained for a single-layer island. For a stack of 10 layers, FWHM increases and integrated photoluminescence intensity decreases. This outcome can be caused by an increase in the total amount of strain around the stacked quantum dot structures, which induces dislocations. Figure 2.26b shows the dependence of interval layer thickness on emission spectra with the five stacked layers. As the interval layer thickness decreases, the emission spectrum shifts toward the lower energies and narrows to FWHM of 27 meV at the 3 nm interval.

Figure 2.27 shows the excitation power dependence of the photoluminescence spectra for the five-layer islands at (a) 2 nm and (b) 3 nm intervals. As the excitation power increases, both samples exhibited peaks in the higher-energy regime, which can be attributed to the higher-order quantized states. Note that energy separation between adjacent states is almost constant for both samples. The energy separation between the ground state and the first excited state emission is about 52 meV for the 2 nm-interval sample and 44 meV for the 3 nm-interval sample. The separation is smaller than for the single-layer island of about 70 meV, as seen in the electroluminescence spectrum.

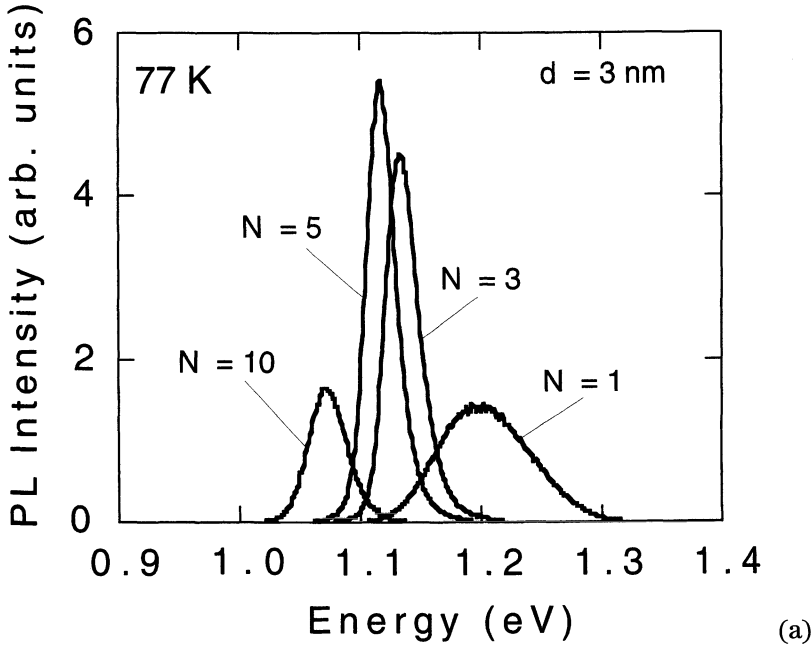
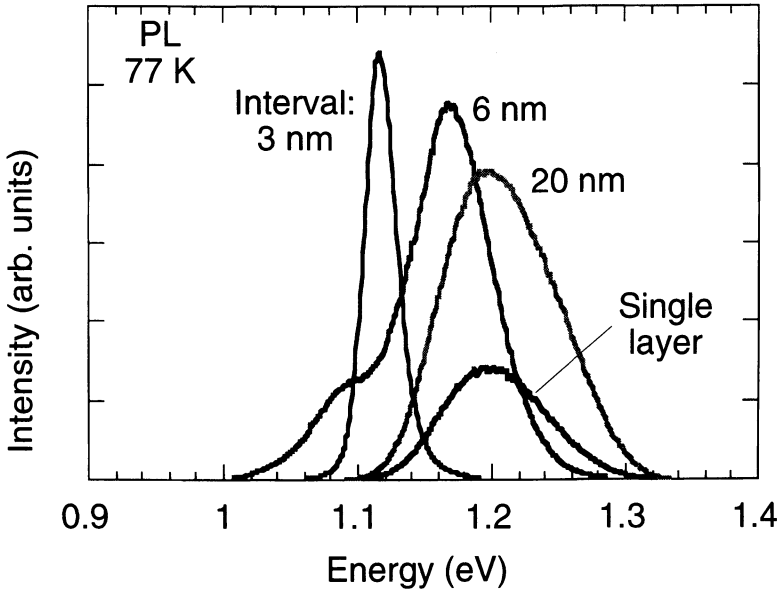


Figure 2.26: (a) Emission spectra at 77 K of a single-island layer and closely stacked 3-, 5-, and 10-layer island structures grown at 3 nm intervals; (b) emission spectra at 77 K of a single layer and closely stacked five-layer islands measured for different interval-layer thickness samples. [93] (From JJAP, Vol. 36, Pt. 2, Sugiyama et al., (1997), Copyright 1997 by the Japan Society of Applied Physics)

All the results observed above indicate that the electron wave functions between neighboring islands overlap each other and that the stacked structure can be considered to be a single quantum dot with a larger size. The evidence includes (1) the red shift of the emission spectra with an increase of the repetition number, as shown in Fig. 2.26a, (2) the red shift of the emission spectra with a decrease of the intermediate layer thickness (Fig. 2.26b), and (3) the decrease in the energy separation between the discrete energy states. Thus, we can control the quantum dot size perpendicularly by this close-stacking technique through the change in the stacking repetition number and the thickness of the interval layers.

A question that often arises when we observe luminescence spectra from quantum dot samples is whether the emission really originated from the dots and reflects the characteristics peculiar to their three-dimensional confinement. In the case of closely stacked dots, in particular, the identi-



(b)

Figure 2.26: *Continued.*

fication is not so straightforward because of their structural complexity, namely, multiple stacked structures accompanied by wetting layers. Magneto-photoluminescence, treated here, gives us not only the evidence of the three-dimensional confinement of excitons but also the extent of the exciton wave function in quantum dots. The principle can be summarized as follows. The magnetic field confines excitons in the plane perpendicular to the field and increases their energy. We can determine the magnitude of other competing confinement potentials, and thus the extent of the wave function, by evaluating the amount of energy shifts.

The samples are five-layer stacked InAs SK islands with 3 nm GaAs intermediate layers, and, for comparison, an $\text{In}_{0.18}\text{Ga}_{0.82}\text{As}/\text{GaAs}$ quantum-well sample grown on a GaAs substrate by MOCVD. The $\text{In}_{0.18}\text{Ga}_{0.82}\text{As}$ -well layer thickness was 8 nm and was subjected to a -1.3% biaxial compressive strain. Figure 2.28a shows the photoluminescence spectra of the two samples with no magnetic field applied. Figure 2.28b shows the diamagnetic shifts measured for both samples under magnetic fields perpendicular and parallel to the sample plane at 4.2 K. In the quantum well, while the shift under a perpendicular field reached up to 7.3 meV at a maximum field of 11.8 T, the shift under the parallel field is much smaller (only 1.6 meV at 11.8 T). The shift of the dots is almost independent of the field direction,

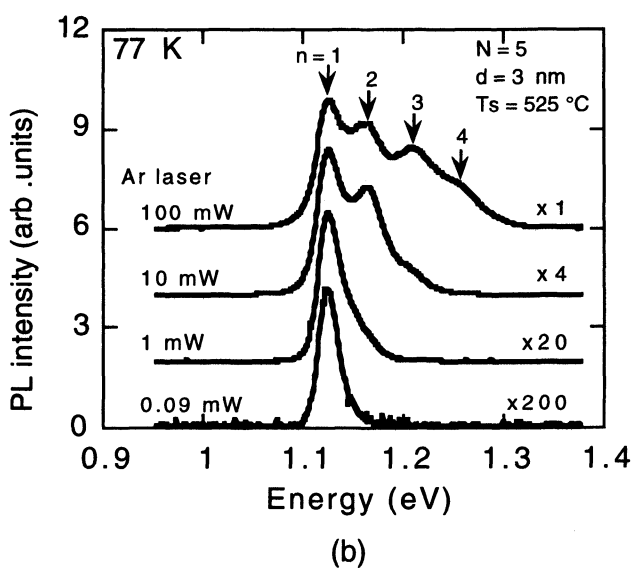
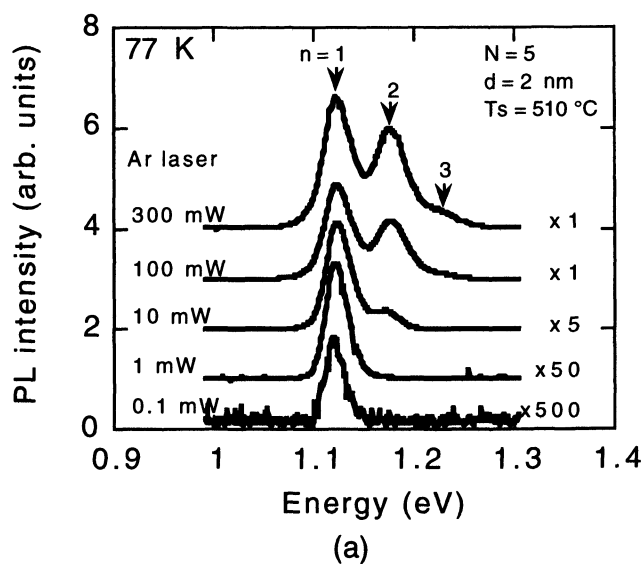


Figure 2.27: Excitation power dependence of photoluminescence spectra for the five-layer islands with (a) 2 nm and (b) 3 nm intervals.

(From JJAP, Vol. 36, Pt. 2, Sugiyama et al., (1997), Copyright 1997 by the Japan Society of Applied Physics)

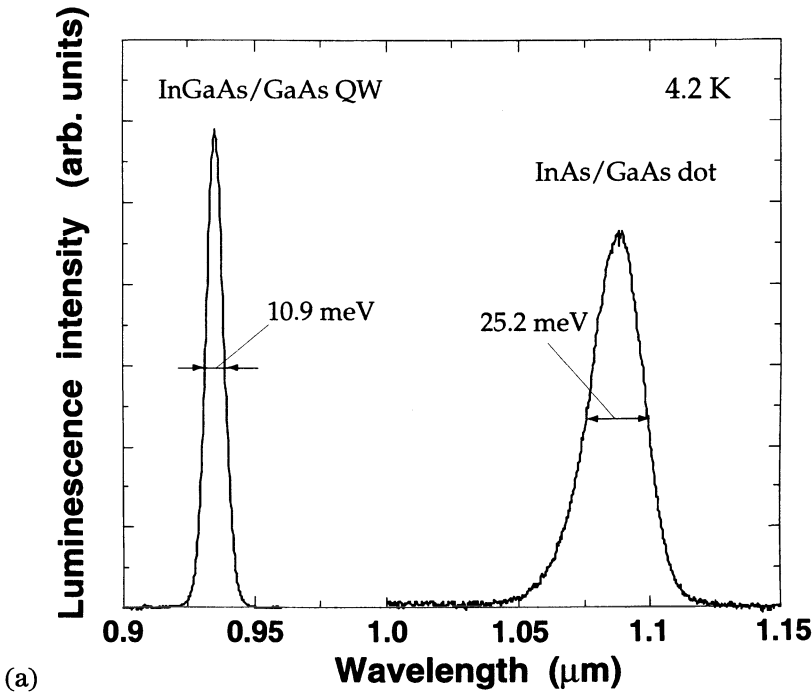
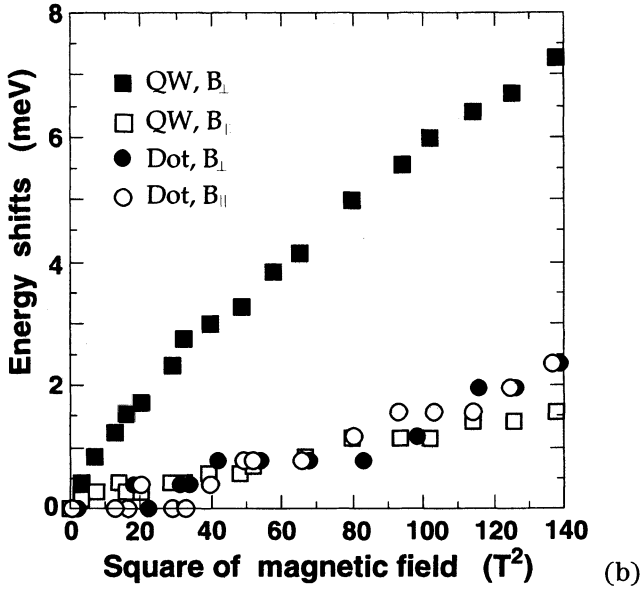


Figure 2.28: (a) Photoluminescence spectra of the two samples with no magnetic field applied; (b) diamagnetic shifts of emission spectra for both samples under magnetic fields perpendicular and parallel to the sample plane at 4.2 K.

(From Physical Review B, Vol. 55, Sugawara et al., Copyright 1997 by the American Physical Society)

and the maximum shift is 2.4 meV in either direction. These results immediately give us the following insights into the differences in quantum-confinement characteristics between the two samples. The asymmetrical shift in the quantum well is due to its one-direction quantum confinement. When the field is perpendicular to the plane, it works as a two-dimensional confinement potential for excitons in the well layer, causing large diamagnetic shifts. When the field is parallel to the plane, it affects only one direction in the quantum-well plane, since the other direction is already confined by strong potential barriers. The symmetrical, small diamagnetic shifts in the closely stacked dots show that excitons are three-dimensionally confined by an almost symmetrical confinement potential. The extent of the exciton wave function is estimated to be almost equal to the quantum-well thickness of 8 nm, since the shifts for the closely stacked

Figure 2.28: *Continued.*

dots and the quantum well under the parallel configuration are almost the same. Detailed quantitative analyses support this simple qualitative estimation [57]. This extent of the wave function, approximately 8 nm, is smaller than the dot size observed by the plane-view transmission electron microscope (Fig. 2.24b) and is definitely larger than the height of the single SK dot of about 3 nm. This indicates that the wave function of the ground-state exciton extends over two or three SK dots out of five. Such exciton localization in the entire stacked structure can be attributed to the composition and/or size inhomogeneity of each SK dot component, i.e., the potential minimum is formed by the coupling of the adjacent two or three SK dots with low band-edge energies.

2.3.4 Columnar InAs quantum dots

We have succeeded in narrowing the spectrum width using closely stacked dot structures. However, the emission efficiency was severely degraded; the photoluminescence from the closely stacked islands rapidly weakens as temperature increases and is hardly observed at room temperature. Among various growth conditions, annealing after the growth of the GaAs intermediate layer was found to affect degradation most seriously. A large number of defects and/or impurities could have been introduced during annealing, resulting in nonradiative recombination centers. In contrast, the

decrease in the spectrum width down to 25 meV was a result of annealing up to 2. Without annealing, the spectrum width was only reduced to 50–60 meV at most, presumably due to the fluctuation in the island height and space thickness.

Our next strategy to solve this problem was to make the intermediate layer even thinner than that used in the closely stacked dots so that the stacked InAs islands physically contacted each other perpendicularly to make columnar-like structures as a whole. The growth sequence is the same as stated above except that the intermediate-GaAs-layer thickness is reduced to 3 ML, and the supplied amount of InAs above the second layer is reduced. If we supply the same InAs amount as before, a large number of dislocations are introduced, and we do not obtain high emission efficiency. For example, with a 1.8 ML InAs supply for the 1 nm intermediate layer, we did not get high efficiency even at low temperatures. So we optimized the InAs supply for the stacked island formation. The InAs islands in the first layer are grown with the 1.8 ML supply as before.

Figure 2.29 shows the RHEED pattern intensity measured during the growth of the GaAs intermediate layer on the first and second InAs islands. The intermediate-layer thickness was changed from 10 to 3 ML. Immediately after the GaAs supply, the spot intensity decreased. The spot intensity recovered to that of the first islands, showing that the second islands are well fabricated. Note that the 3 ML and 5 ML GaAs is thinner than the height of the islands.

Figure 2.30 shows the RHEED pattern intensity transition when the InAs islands were repeatedly grown with an intermediate layer of 3 ML. The inset photographs are the AFM images of the topmost layer. The amount of InAs supply was 0.46 ML, 0.71 ML, and 0.81 ML. When the InAs supply was 0.46 ML, the RHEED intensity did not recover as we repeated the growth, showing that islands were not formed. When the InAs supply was 0.81 ML, the intensity did not recover at the ninth islands. The surface AFM pattern at 0.71 ML shows islands with almost same density and size as the first islands. The above results suggest that there is an optimum InAs supply amount. The TEM photograph in Fig. 2.31 shows the overall structural features. This is a cross-sectional TEM image of the structure. InAs island layers formed with an InAs supply of 0.7 ML and intermediate GaAs layers of 3 ML were grown alternately (eight repetitions) on the first island layer formed with a 1.8 ML InAs supply. The islands on each layer were seen to be in physical contact. The stacked islands, as a whole, had a columnar shape with a diameter of 17 nm and a height of about 13 nm. All island layers were accompanied by the wetting layers (contrasting horizontal shapes in the TEM image), indicating that the stacked islands were formed by the SK mode.

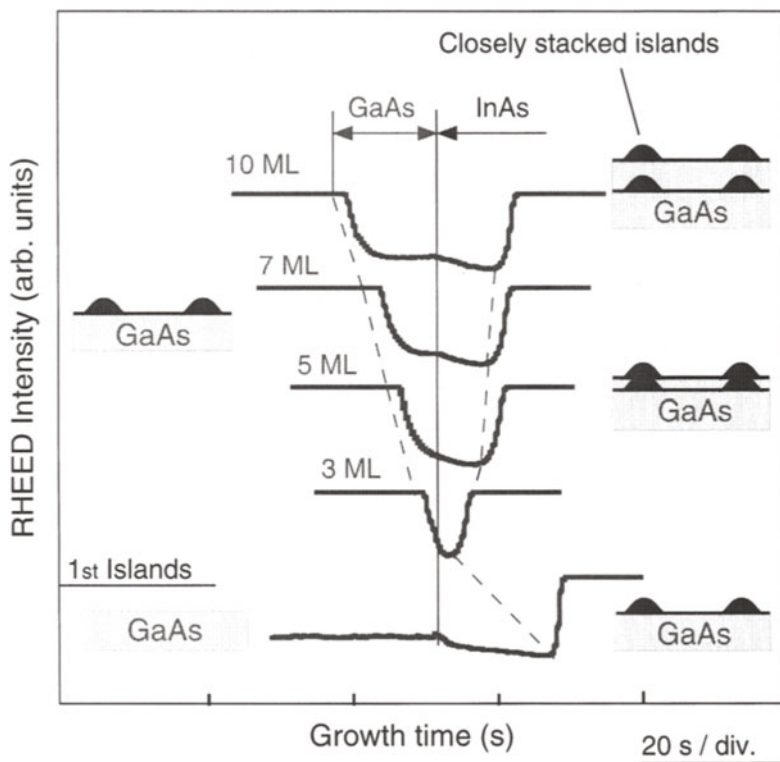


Figure 2.29: RHEED pattern intensity transition during growth of the GaAs intermediate layer on the first InAs islands and during growth of the second InAs islands.

(From IOP Conference Series: 1998 IEEE International Symposium on Compound Semiconductors, Nakata et al., Copyright 1998, with permission from IEEE)

Figure 2.32 shows the PL spectrum of the columnar-shaped quantum dots measured at 300 K. The peak wavelength of the ground level was $1.17\ \mu\text{m}$; that of the excited level was $1.10\ \mu\text{m}$. An emission peak related to the multiple wetting layers was observed at $1.01\ \mu\text{m}$. A full width at half maximum (FWHM) of the spectrum was about 42 meV. This is quite a bit smaller than for the ordinary SK-mode QDs of about 80 meV, indicating that the structural uniformity was fairly improved. The PL intensity is also rather better than for the SK-mode QDs and is three orders greater than that of the closely stacked QDs with growth interruptions.

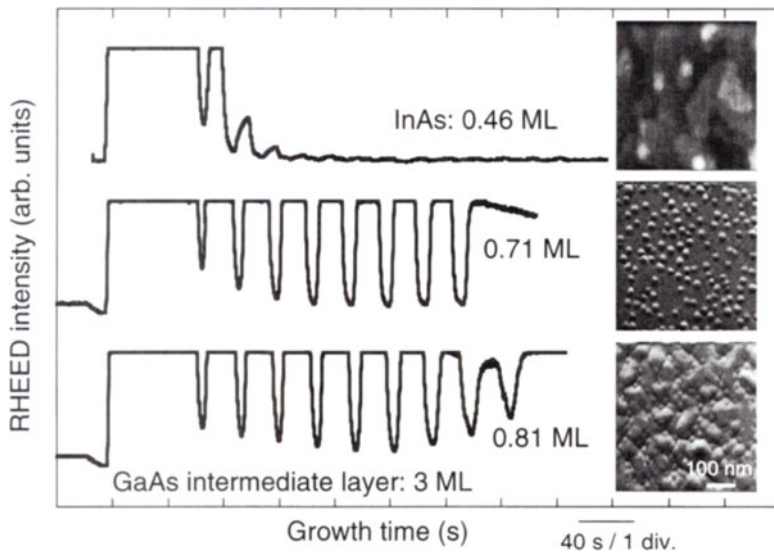


Figure 2.30: RHEED pattern intensity transition when InAs islands were repeatedly grown with an intermediate layer of 3 ML. The inset photograph shows the AFM images of the topmost layer.

(From SEMICONDUCTORS AND SEMIMETALS, Vol. 60, Nakata et al., (1999), Copyright 1999 by the Academic Press)

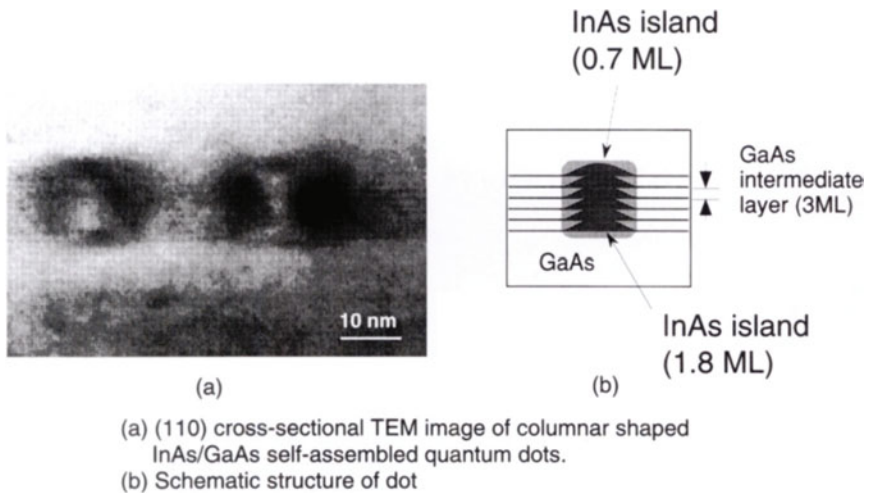


Figure 2.31: TEM photograph of the columnar-shaped InAs quantum dots together with schematics of the dots.

(From IOP Conference Series: 1998 IEEE International Symposium on Compound Semiconductors, Nakata et al., Copyright 1998, with permission from IEEE)

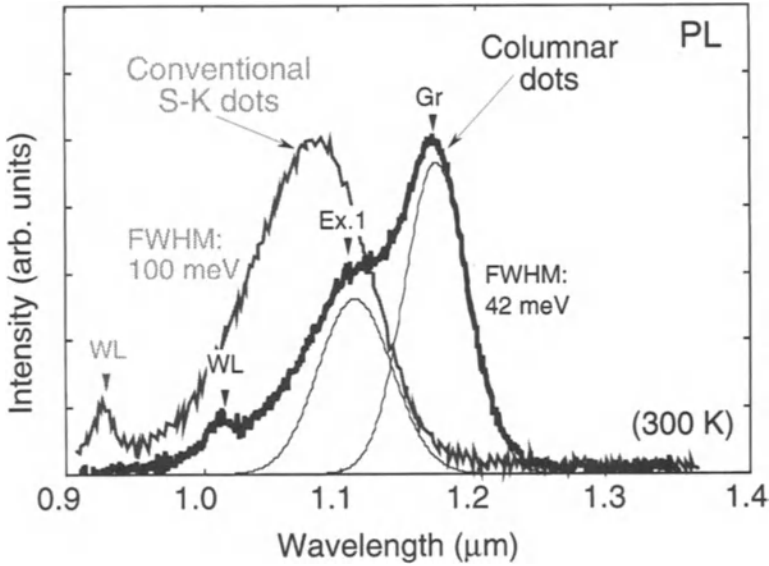


Figure 2.32: PL spectrum of the columnar-shaped quantum dots measured at 300 K, together with that of a conventional single quantum dot layer. (From IOP Conference Series: 1998 IEEE International Symposium on Compound Semiconductors, Nakata et al., Copyright 1998, with permission from IEEE)

Note that the emission from the wetting layers redshifted when compared to the ordinary SK-mode QDs. This suggests that multiple wetting layers were coupled electrically and formed a quantum level as a whole. The energy separation between the ground state and wetting layer was 168 meV, which is much smaller than the 230 meV of the ordinary SK-mode QDs. Thus, we were able to decrease the size fluctuation in the quantum dots and retain the optical quality.

Figure 2.33 plots the photoluminescence wavelength at room temperature and 77 K as a function of the stacked-layer number. As the layer number increases, the emission wavelength shifts to a longer wavelength due to the size increase. Note that the longest wavelength at room temperature is 1.24 μm at stacking 23, very close to 1.3 μm , the low-loss, low-dispersion wavelength of optical data transmission systems.

The structural and optical properties of different types of dots are summarized in Fig. 2.34.

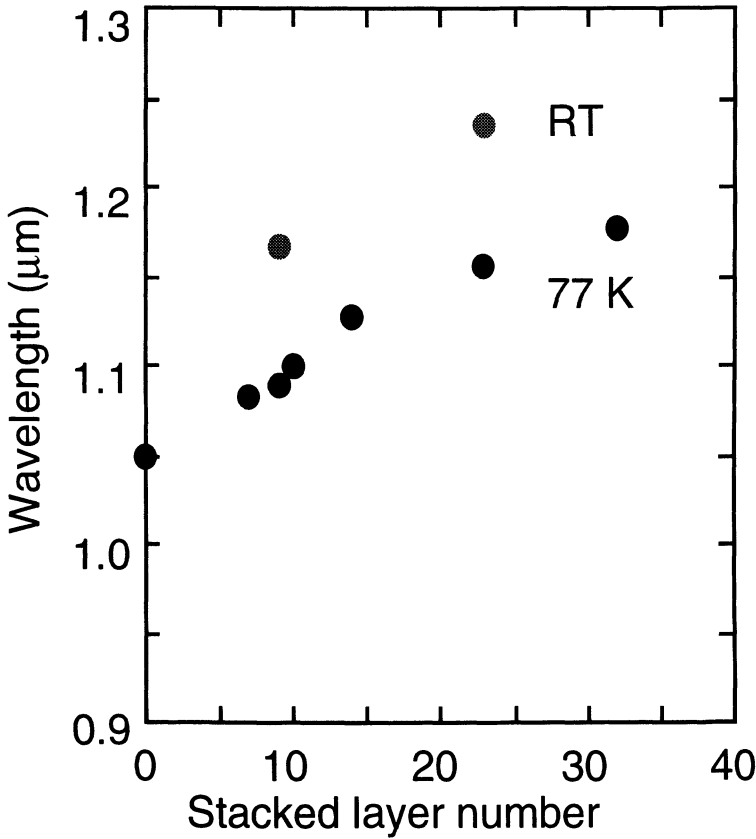


Figure 2.33: Photoluminescence wavelength at room temperature and at 77 K as a function of the stacked layer number.

(From IOP Conference Series: 1998 IEEE International Symposium on Compound Semiconductors, Nakata et al., Copyright 1998, with permission from IEEE)

2.4 Potential device applications of QDs

One of the significant features of quantum dots is the inherent discrete nature of energy levels, based on the confinement of carriers in any three-dimensional direction. This feature is expected to improve the performance of optical devices such as laser diodes, as originally proposed by Arakawa and Sakaki in 1982 [58], and is also expected to be used to develop new-principle electron and optical devices. Another feature of quantum dots is single-electron charging effects; the charging energy of an electron becomes larger than the thermal energy of ambient temperatures

	S-K island	Closely-stacked	Columnar-shaped
Schematic structure			
PL FWHM	90 meV	25 meV	40 meV
PL efficiency	Good	Poor	Good
In-plane coverage (Density)	40 % $> 1 \times 10^{11} \text{ cm}^{-2}$	40 % $> 1 \times 10^{11} \text{ cm}^{-2}$	40 % $> 1 \times 10^{11} \text{ cm}^{-2}$
Wavelength (300 K)	1.1 μm	1.2 μm	$> 1.2 \mu\text{m}$

Figure 2.34: Summary of structural and optical properties of different types of dots.
(From SEMICONDUCTORS AND SEMIMETALS, Vol. 60, Nakata et al., (1999), Copyright 1999 by the Academic Press)

due to decreased capacitance. This effect has been studied using GaAs-based nanometer structures such as those in floating-dot gate HEMTs. The feature should be more extensively studied using Si-based materials, since the single-electron charging effect becomes a real problem in nanoscale Si MOSFETs as well as a tool to break through the scaling limit of metal-oxide semiconductor (MOS) devices. This section discusses potential device applications, with a focus on the former feature of QDs. In another chapter of this book, the latter feature of quantum dot and/or nanometer structures is described in detail.

2.4.1 TSR QD memories

TSR QDs have several technological advantages with reference to the construction of electron devices. There are at least two ways of building memory devices. One is to use resonant tunneling effects through quantum dots. The memory function is achieved owing to the intrinsic voltage bistability at the node between two sets of a dot and tunnel barriers. The other is quasi-nonvolatile memory utilizing the QD as a charge-trapping floating gate, similar to a MOS FET with a nanoscale floating gate [59,60].

Transport of electrons through zero-dimensional QDs has been under extensive study recently, highlighting both the resonant tunneling

(quantum effects) and the single-electron charging phenomena (Coulomb effects). In particular, for resonant tunneling, a new theoretical prediction was proposed that a high peak-to-valley (P/V) current ratio will be attained at the negative differential resistance (NDR) region in the current-voltage (I-V) characteristics of the diode if the resonant tunneling occurs at least between two QDs [61]. This prediction is based on the expectation that inelastic phonon scattering is suppressed due to the discreteness of the atom-like density of the states in a QD, hence leading to an extremely small excess current at the NDR region. More strictly, it has been claimed that the energy relaxation of carriers is restricted to narrow energy regime where the energy spacing between the QD subbands exactly matches the LO-phonon energy of the QD material [62,63]. This is often referred to as the *phonon bottleneck problem*. It was further proposed that, if one substitutes QDs for conventional 2D QWs of the double-emitter resonant-tunneling hot-electron transistors (ME-RHET), where only one transistor is necessary in a memory cell [64], power consumption would be drastically lowered during their static random access memory (SRAM) operation [61]. The phonon bottleneck of the energy relaxation of carriers was experimentally confirmed in the I-V characteristics of magnetically confined, quasi-zero-dimensional $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}/\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ triple-barrier resonant tunneling diodes (RTDs) [65,66]. Under application of high magnetic fields perpendicular to heterointerfaces, the excess valley current was actually reduced and the P/V ratio increased up to as high as 125 at low temperatures [66]. Therefore, if similar results are achieved in a structurally confined real QD system, we could obtain SRAMs with extremely low-power consumption as well as an ultrasmall size.

Prior to making the device structures, we have to reveal electron transport through the TSR QDs. We fabricated a doubly stacked TSR-QD structure, as shown in Fig. 2.35 [67]. First, we formed TSRs in an undoped-GaAs layer that had been previously grown on a Si-doped n^+ (111)B GaAs substrate. After that, we grew Se-doped n^+ -GaAs/InGaAs buffer layers, an undoped AlGaAs/InGaAs triple-barrier resonant tunneling structure, and n^+ -GaAs/InGaAs for ohmic contact layers. The composition of both AlGaAs and InGaAs and the concentration of Se doping are controllable, but their absolute values in our experiments are unknown. The electrodes were formed by evaporating AuGe/Au layers on both the top and back of the sample. The growth and device fabrication process, including the ohmic contact to the top surface, was carried out without removing the SiO_2 mask, which had round-shaped openings and which was initially used to form the TSRs. Therefore, we were able to make a sample that allowed us to measure the electronic properties of an individual TSR structure in a self-aligned manner. Since we intended primarily to measure the transport

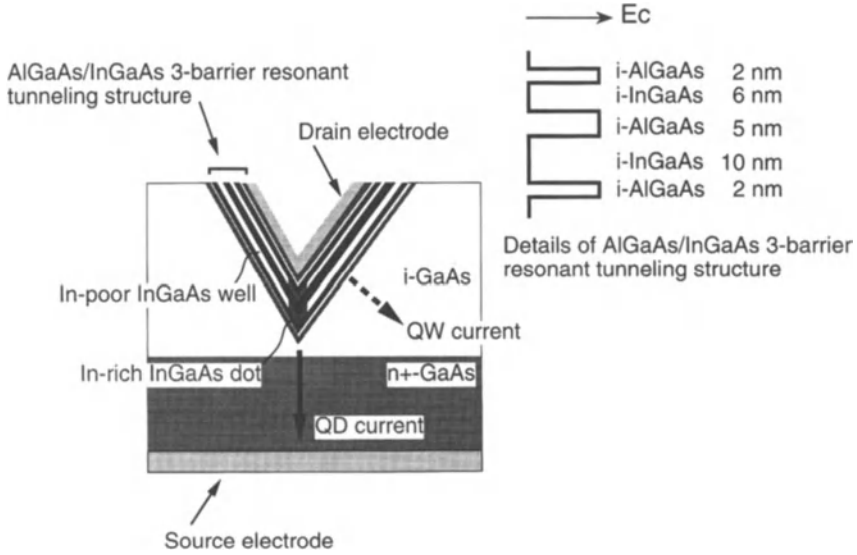


Figure 2.35: Schematic of a doubly stacked TSR-QD diode for I-V characteristics.

property through the vertically aligned QDs by suppressing the current flows through TSR side surfaces, the bottom of the TSR was initially set close to the interface between the undoped GaAs and the n^+ substrate so that the electric field might concentrate the dots.

Figure 2.36 shows the I-V characteristics at 500 mK from a structure grown on a single TSR structure. NDRs on the order of μA were observed under both positive and negative biases. Fine structures on the order of nA were observed at the lower bias region, as shown in the inset. We speculate at the present stage that the NDRs at the higher bias and the fine structures in the lower bias are due to resonant tunneling through QWs at the sidewalls and through QDs at the bottom, respectively.

Looking more closely at the lower bias region, we observed a steep rise followed by the gradual decrease in the current at several voltages in the positive region (Fig. 2.37a). According to the theoretical calculations by Bryant [68], this peculiar I-V characteristic is one of the typical features of a system including double dots, which are strongly coupled to each other in a quantum mechanical sense. In contrast, on the negative bias side, no NDRs but rather staircase patterns were observed in the I-V curve, as shown in Fig. 2.37b. Unfortunately, we have not yet obtained the expected large P/V ratio in the TSR-QD diodes. To clarify the mechanism of electron

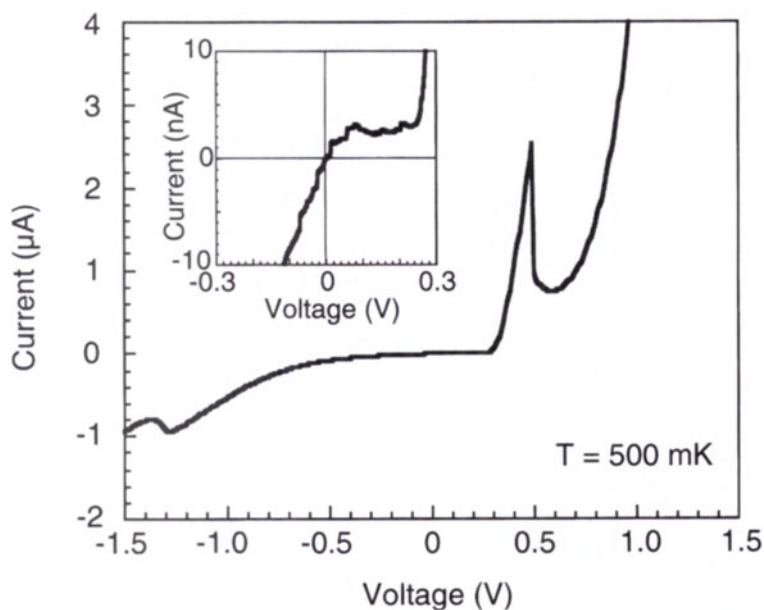


Figure 2.36: Overall I-V characteristics of the diodes at 0.5 K. The inset shows the existence of fine structure in a low-bias regime.

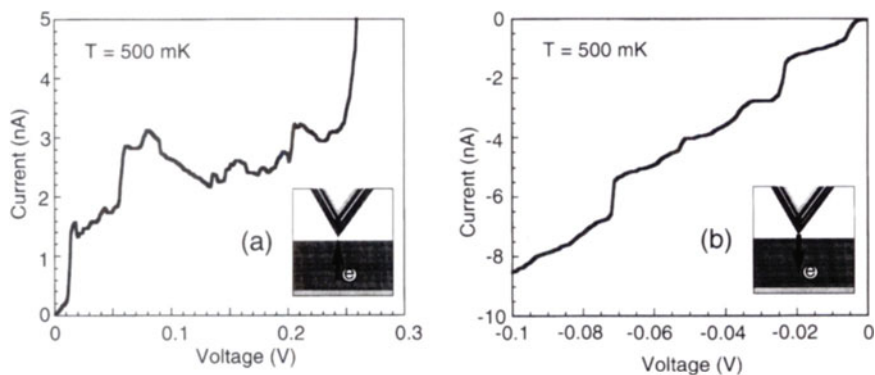


Figure 2.37: Blowup data of I-V characteristics at low bias on (a) the positive side and (b) the negative side.

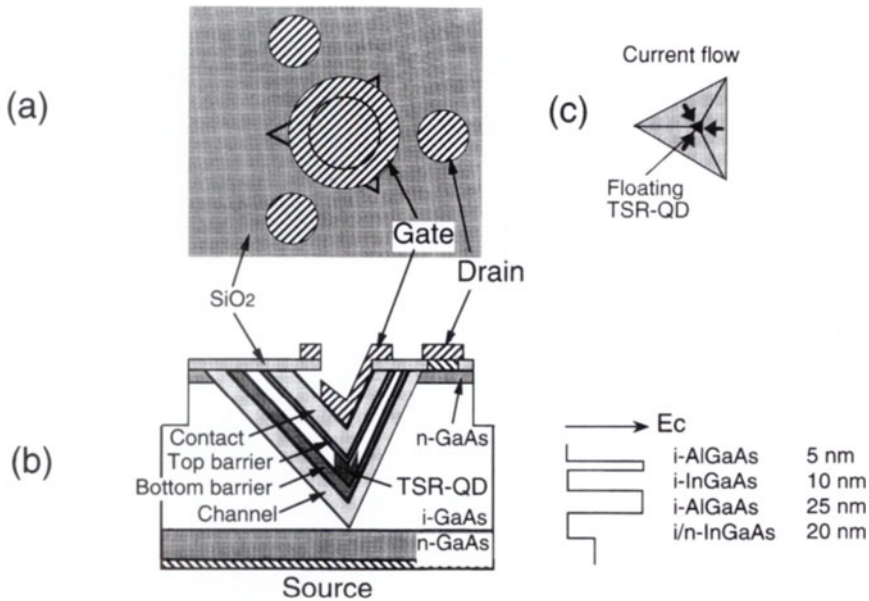


Figure 2.38: Schematics of a novel doped-channel FET with the buried gate of a floating TSR QD: (a) top view, (b) side view, and (c) relationship between current flow at sidewalls and a TSR QD.

transport through TSR QDs, we must investigate more growth structures in the TSR and optimize them.

Another opportunity to make use of TSR QDs for memory devices is the InGaAs/AlGaAs doped-channel heterostructure FET having a floating-dot gate consisting of an In-rich InGaAs TSR QD. Figure 2.38 is a schematic diagram of the device structure. Prior to TSR formation, we prepared a planar substrate by forming a thick undoped GaAs layer on a Si-doped n^+ -(111)B GaAs followed by a 75 nm-thick Se-doped n^+ -GaAs layer on the surface, using MOCVD at 800°C. After the TSR formation in the wet etching process, so that it would be buried in the undoped GaAs layer, we grew Se-doped n^+ -GaAs/InGaAs layers, an undoped AlGaAs (25 nm)/InGaAs (10 nm)/AlGaAs (5 nm) double-barrier structure, and Se-doped n^+ -GaAs/InGaAs inside the TSRs, again using MOCVD. In this structure, the buried n^+ -InGaAs layer at the sidewalls acts as a doped channel for the electrons. The source was an n^+ -substrate and the drain was surface n^+ -GaAs. The gate electrode was formed at the n^+ -InGaAs on the topmost surface of the TSR by evaporating AuGe/Au without alloying. The SiO_2 mask was not

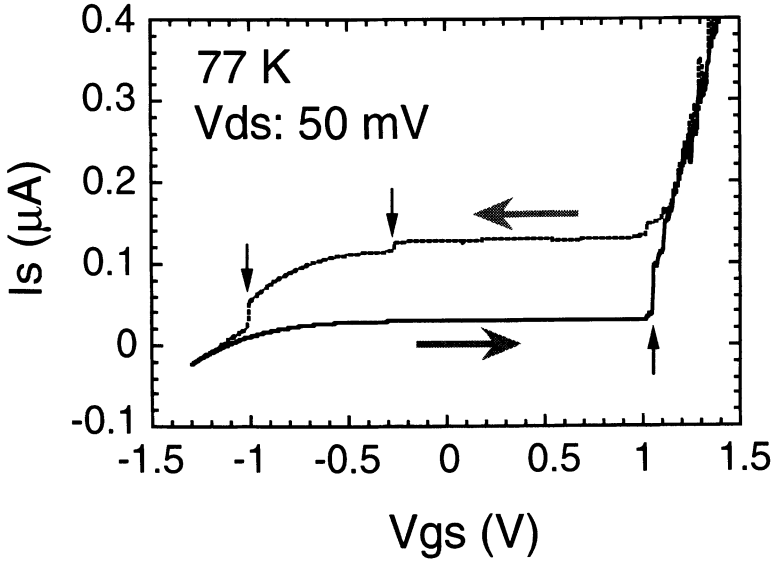


Figure 2.39: Measured V_{gs} - I_s characteristics at 77 K.

removed throughout the growth and device fabrication processes. Since In-rich InGaAs-QD self-forms at the bottom of a TSR, it is expected that we can control the trapping and detrapping of electrons from the gate electrode through the thin, top AlGaAs tunnel barrier.

Figure 2.39 plots the source current I_s variation as a function of the gate voltage V_{gs} for a single TSR structure at 77 K. Here, a constant voltage V_{ds} of 50 mV was applied between the source and drain electrodes. We found a clear hysteresis in the V_{gs} - I_s characteristics. We believe this bistability of I_s is due to either the occupied or the empty state of the electrons at the TSR QD. The lower curve, indicated by the solid line, might correspond to the characteristics of the occupied dot state; the upper curve, indicated by the broken line, might correspond to the empty dot state. When we changed V_{gs} from -1.3 V to 1.1 V, there was no significant increase in I_s . However, at the V_{gs} of 1.1 V, I_s increased in a step-like form, probably because the captured electron might be escaping from the TSR QD by tunneling through the thin AlGaAs barrier to the gate electrode. The steep increase in I_s beyond $V_{gs} = 1.1$ V might be caused by the increase in the number of injected electrons from the source to the FET channel in the vicinity of the initial TSR bottom, due to the positive gate voltage. When we decreased the V_{gs} from 1.1 V to -1.0 V, we observed a larger I_s than that

in the occupied dot state of the lower curve, because the TSR QD would remain empty. At $V_{gs} = -1.0\text{V}$, the TSR QD might be charged by electrons tunneling from the gate electrode through the AlGaAs barrier. As a result, the I_s transits to the lower curve of the occupied dot state. The reason for the small current step in the upper curve at $V_{gs} = 1.05\text{V}$ and -0.28V is not clear. The result implies that we can use the gate electrode to program the I_s states around the low V_{gs} region by charging or discharging the TSR QD.

We have observed the memory action in our novel FET structure, where the TSR QD is placed between the channel and the gate electrode. However, these are just phenomenological observations. Although this memory effect is nonvolatile within the 1-hour timescale of our experiments, an evaluation of retention time has not yet been done. Considering the operation principles of the device, we speculate that the height of the potential barrier of undoped GaAs between the apex of TSR and the n^+ substrate (source) will be strongly affected by the amount of occupied electrons in the TSR QD and that most of the FET current will be passed by way of the apex region to the n^+ substrate. If the TSR QD is charged by electrons, the injection of electrons from source to the doped channel at the TSR apex will be hindered by the increased height of the potential barrier. Therefore, the current flow through the FET channel will remain small without the large positive V_{gs} . In contrast, in the case of empty TSR QD, the barrier height will be low and the larger amount of electrons will flow from source to drain. Thus, the two types of potential profiles near the apex of the TSR will bring the memory effect to the I_s - V_{gs} characteristics. However, a great deal of work remains to be done in order to clarify a detailed mechanism of the device operation, including the memory effect.

These GaAs quantum dot devices could be expected to have a practical use when real system LSI chips using various functional devices are integrated on a Si or other material platform substrate.

2.4.2 Quantum dot lasers

2.4.2.1 Optical gain

An attractive feature of the quantum dot is its feasible large optical gain for semiconductor laser application [58]. Aside from the phonon bottleneck problem [69], the advantage of quantum dots can be illustrated by the two-level model. When the inhomogeneous broadening is much larger than the homogeneous one, which is actually the case, the optical gain can be written as

$$g(\omega) = \frac{2\sqrt{\pi}e^2\omega}{c\eta\epsilon\hbar\Delta\omega} |\langle \Psi_c | \Psi_v \rangle|^2 M^2 N (f_c - f_v) e^{-\frac{(\omega-\omega_0)^2}{\Delta\omega^2}} \quad (2.2)$$

where we assumed the Gaussian inhomogeneous broadening with the center angular wavelength of ω_0 and FWHM of $1.67\Delta\omega$. In Eq. (2.2), η is the refractive index, ϵ is the vacuum dielectric constant, e is the electron charge, M is the Kane's matrix element [70], N is the dot density, f_c and f_v are Fermi–Dirac distribution functions, and $|\langle \Psi_c | \Psi_v \rangle|^2$ is the square of the overlap integral of the ground-state conduction band and valence band envelope wave functions. Eq. (2.2) shows that the linear gain is inversely proportional to the inhomogeneous broadening and is proportional to the dot density and proportional to the second power of the overlap integral.

2.4.2.2 Self-organized quantum dots lasers

Self-assembled growth technology has progressed to achieve lasing with fairly low threshold currents [71–75]. As can be seen from Eq. (2.2), we must reduce the inhomogeneous broadening and increase the dot density to get high gain leading to the low-threshold current operation. It is also important to get large overlap integral for high gain. Here, we review our self-assembled quantum dot lasers in view of these parameters. We fabricated ridge stripe structure lasers as shown in Fig. 2.40 for the regular SK-mode grown quantum dots [73,74], and the columnar quantum dots [75].

In the regular SK type of quantum dots, lasers were fabricated with several numbers of stacks, cavity lengths, and with and without HR (high reflectivity) coating on both facets [74]. The dot density is about $1.0\text{--}1.3 \times 10^{11}/\text{cm}^2$ per layer, and the PL spectral width is about 80 meV in FWHM. Figure 2.41 shows the lasing spectrum for each device structure at 300 K. For the one-dot layer and cavity length of $300\mu\text{m}$, the lasing was at the wetting layer (Fig. 2.41a). By increasing the cavity length to $900\mu\text{m}$, the lasing shifted to the longer wavelength corresponding to the third quantum level (Fig. 2.41b). By increasing the number of stacks to three, the lasing occurred at the second sublevel (Fig. 2.41c). We achieved continuous-wave (CW) at the ground state by applying 95% HR coating on both facets (Fig. 2.41d). The threshold current density of the ground-state lasing was $1\text{ kA}/\text{cm}^2$. Increasing the number of stacks increases the effective dot density and increases the optical gain. The longer cavity length and the HR coating reduce the lasing threshold. Thus the long cavity of $900\mu\text{m}$, three-layer stacked structure and the HR coatings led to the CW lasing at the ground state at room temperature. When compared with the PL spectrum

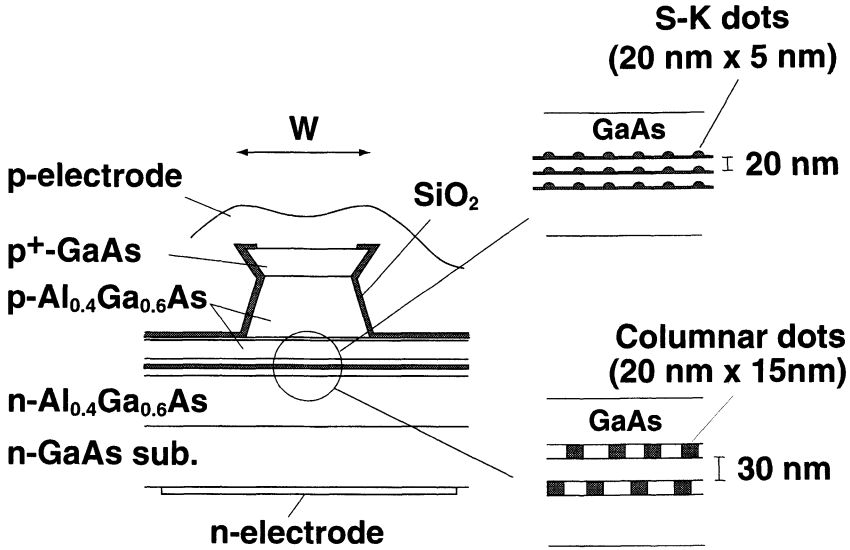


Figure 2.40: Schematic structure of a quantum dot laser. Stacked SK mode quantum dot lasers and the columnar quantum dot lasers are fabricated with this structure.

of SK dots in Fig. 2.32, we can somehow identify the discrete levels in the EL spectrum below threshold in Fig. 2.41. This is due to the smaller inhomogeneity in the size and the composition in the rather narrow stripe region of $2.5 \times 900 \mu\text{m}^2$, when compared with PL excitation area of $300 \times 300 \mu\text{m}^2$.

With the columnar quantum dots, the lowest threshold current was obtained [75]. We used two-layer stacked columnar quantum dots as an active layer. The dot density is $7 \times 10^{10}/\text{cm}^2$ per layer, and the FWHM of the PL emission is 40 meV. The I-L characteristics and lasing spectra of a sample with a cavity length of $300 \mu\text{m}$ and with HR coatings (95%) on both facets are shown in Fig. 2.42. The stripe width is $3 \mu\text{m}$. CW lasing with threshold current of 5.4 mA was achieved at 298 K. The threshold current density is $600 \text{ A}/\text{cm}^2$. The laser with cavity length of $900 \mu\text{m}$ with HR coating on both facets showed threshold current density of $320 \text{ A}/\text{cm}^2$. The lasing was at the wavelength of $1.15 \mu\text{m}$ at the low-injection current level. This wavelength coincides with the ground-state PL wavelength shown in Fig. 2.32. At higher injection current, the lasing spectrum extends to shorter wavelength region. This is a mixture of the ground-state and the second-state emissions. Figure 2.43 shows the I-L characteristics of a sample with

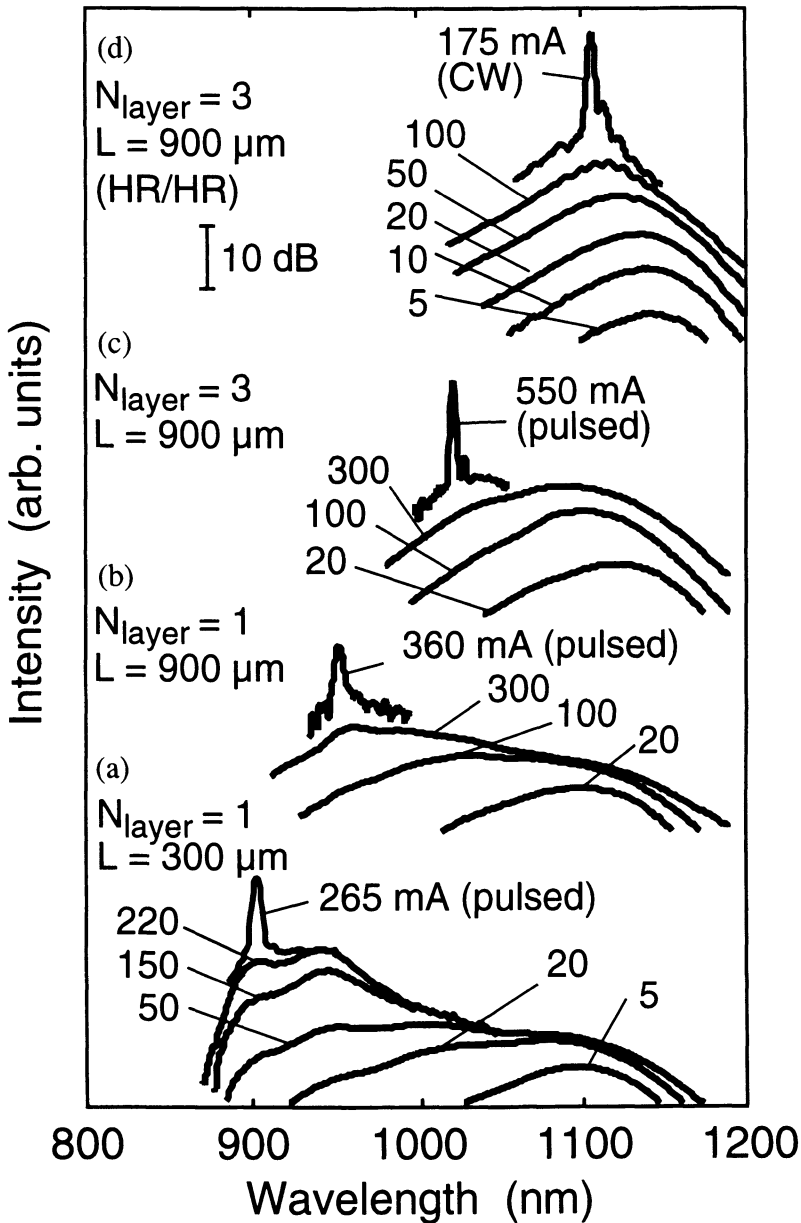


Figure 2.41: Lasing spectra of SK-mode-grown quantum dot lasers. [74]
 (From IEEE Journal of Quantum Electronics, Vol. 3, no. 2, Shoji et al., (1997), Copyright 1997, with permission from IEEE)

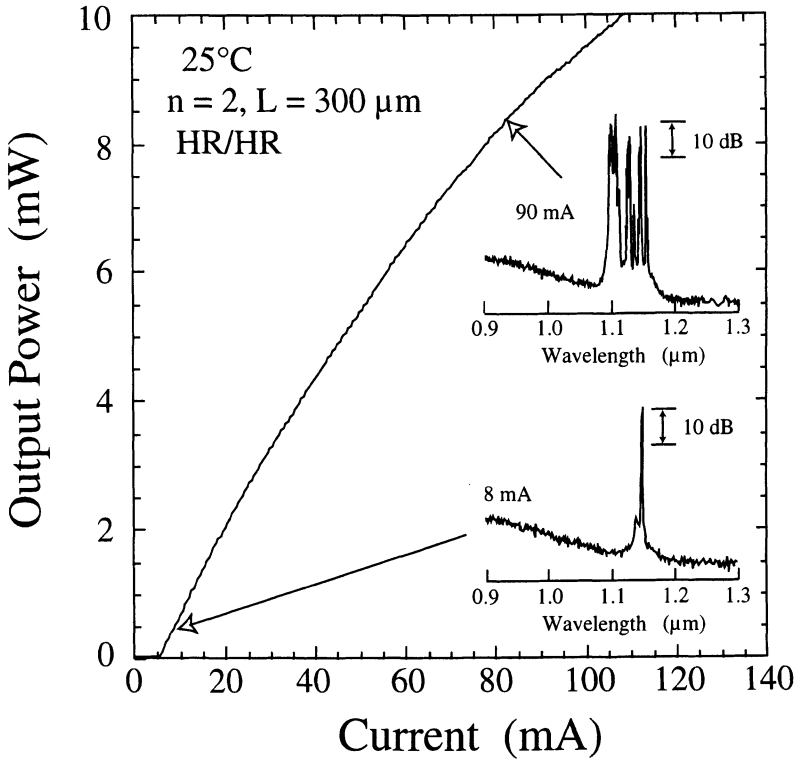


Figure 2.42: I-L characteristics and lasing spectra of a columnar dot laser. The cavity length is $300 \mu\text{m}$, and an HR coating of 95% is applied on both facets.

an as-cleaved facet and an HR-coated facet. Because of the lower reflectivity of the as-cleaved facet, high output power of 110 mW and operation up to the high temperature of 70°C are achieved. In both Fig. 2.42 and Fig. 2.43, the lasing was at the ground state at a low power level; however, lasing shifted to a mixture of ground-state and second-state emissions at higher optical output power.

The higher sublevel lasing in the regular SK-type lasers with higher optical loss (short cavity and/or no HR coating) is due to the higher optical gain at higher sublevel transition caused by the higher degeneracy of states. This means that it is of great importance to realize high dot density and a small inhomogeneous broadening to give sufficient gain to the ground state. The shift of lasing from the ground state to the mixture of

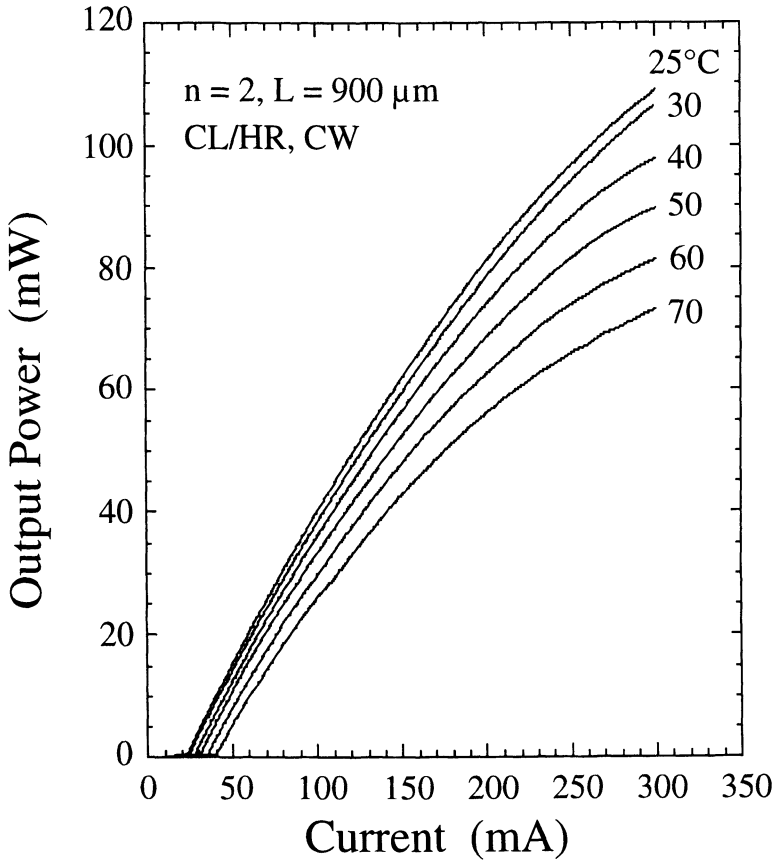


Figure 2.43: I-L characteristics of as-cleaved and HR-coated lasers with cavity length of $900 \mu\text{m}$. [75]

(From IEE Electronics letters, Vol. 34, Mukai et al., (1998), Copyright 1998, with permission from IEE)

the ground state and higher sublevel lasing was observed in the columnar quantum dot lasers as an increase in optical output power. This phenomenon may be related to the phonon bottleneck problem [69]. The slow carrier relaxation easily saturates the gain at ground state to cause the transition of lasing to the upper sublevels.

It is interesting to note that a much lower threshold current density was obtained in the columnar dot lasers than that of the regular SK dot lasers. As can be seen from Eq. (2.2), the optical gain is proportional to the ratio

of the dot density to the inhomogeneous broadening width. This value is almost the same in three-layer-stacked regular SK dot lasers and in two-layer-stacked columnar dot lasers and is about $3\text{--}5 \times 10^9/\text{cm}^2/\text{meV}$. Nevertheless, the threshold current density is much smaller in the columnar dot lasers. One of the explanations for this difference is the overlap integral of the envelope wave functions in the numerator in Eq. (2.2). In the columnar dot, the square of the overlap integral $|\langle \Psi_d | \Psi_v \rangle|^2$ is almost unity for the equivalent dot size of $20\text{ nm} \times 15\text{ nm}$. However, for the regular SK dot, this value is much smaller because of the unequal penetration of the envelope wave function of electron and hole in the thin-growth direction. This difference might have caused the difference in the linear optical gain. The other point worth mentioning is that the stack of wetting layers around the dots forms an equivalent quantum well unlike the regular SK dot lasers, as evidenced in the redshift of the wetting-layer emission in Fig. 2.32. This layer serves as a reservoir for carriers to the dot and might have enhanced the efficient supply of carriers to the quantum dots. These two factors might be responsible for the low threshold current and high power operation of columnar quantum dot lasers.

2.4.2.3 Toward higher-performance lasers

Very-low-threshold current operation comparable to the strained quantum-well lasers has been attained in the quantum dot lasers. However, there remains much to do to reach the ideal performance of the quantum dot lasers. Further improvement of fabrication technologies is essential, and an understanding of carrier dynamics is also of crucial importance to design high-performance quantum dot lasers. Here, we discuss the remaining problems in the fabrication technologies and the carrier dynamics of quantum dots.

In fabrication technologies, the following are key issues.

1. Realization of smaller inhomogeneous broadening and high dot density. We have been able to attain a low threshold current of 5.4 mA with columnar quantum dot lasers that exhibit relatively small inhomogeneous broadening and high dot density [75]. The large matrix element due to large overlap integral, and the quantum well layers formed by the stacking of wetting layers and barrier layers, may have played a role in the attainment of low threshold currents. Although we have obtained promising results, there is still room for further improvement. Inhomogeneous broadening of 27 meV was obtained in our closely stacked quantum dots. The broadening so far was 40 meV in the columnar dots. We should continue our efforts to increase the homogeneity in size and composition to realize smaller

inhomogeneous broadening. Efforts to increase the dot density must also be continued to get high optical gain. In this respect, establishment of the stacking of more layers is necessary.

2. Reduction of the nonradiative recombination. It is uncertain whether nonradiative centers exist in the columnar quantum dots. However, there are several reports on the reduction of the PL intensity at higher temperature [76,77]. This outcome is attributed to the thermal excitation of carriers from the quantum dots [76] and the subsequent nonradiative recombination around the quantum dots [77]. There is a large strain around the quantum dots. Reducing the nonradiative recombination under this high-strain field may be one of the very important steps required to reach ideal quantum dot lasers.

3. Lasing at communication wavelengths. It is of great importance to attain lasing at the communication wavelength of 1.3 and 1.55 μm in order to expand the application of quantum dot lasers. Most of the SK-type quantum dots on a GaAs substrate emit at the wavelength of 1.1–1.2 μm . Attempts to make quantum dots on an InP substrate have obtained emissions in the range of 1.4–1.7 μm [78]. Quantum dots grown by the alternate supply of the monolayer of InAs and GaAs (ALS dots), though not discussed in detail here, are of interest because of their ground-state emission wavelength of 1.3 μm [79]. These efforts have to be continued to realize high-performance lasers at communication wavelengths.

With regard to carrier dynamics, many problems remain unresolved. Because of the discrete energy levels, the energy relaxation rate of carriers to the ground state is considered to be rather slow—a situation known as the *phonon bottleneck problem* [69]. If the relaxation rate is too slow, it will cause an increase in the threshold current, the saturation of light output power, or the shift of lasing from the ground state to higher sublevels, and will limit the modulation response. Many theoretical and experimental studies have been conducted on this phenomenon. There are several models that enhance the carrier relaxation to the ground state other than LO-phonon scattering, such as Auger-like carrier scattering [80,81] and electron-hole scattering [82].

So far, the energy relaxation of carriers to the ground state of quantum dots has been evaluated by time-resolved PL measurement. An example of time-resolved PL measurement (by Mukai et al.) is shown in Fig. 2.44 [83]. The measurement was performed for quantum dots grown by ALS that have five discrete emission peaks due to quantum confinement. The PL decay was measured for each emission peak at temperatures of 20 K and 250 K. It can be seen that the decay rate is slower at lower temperatures and for the lower-state emission. This measurement gives

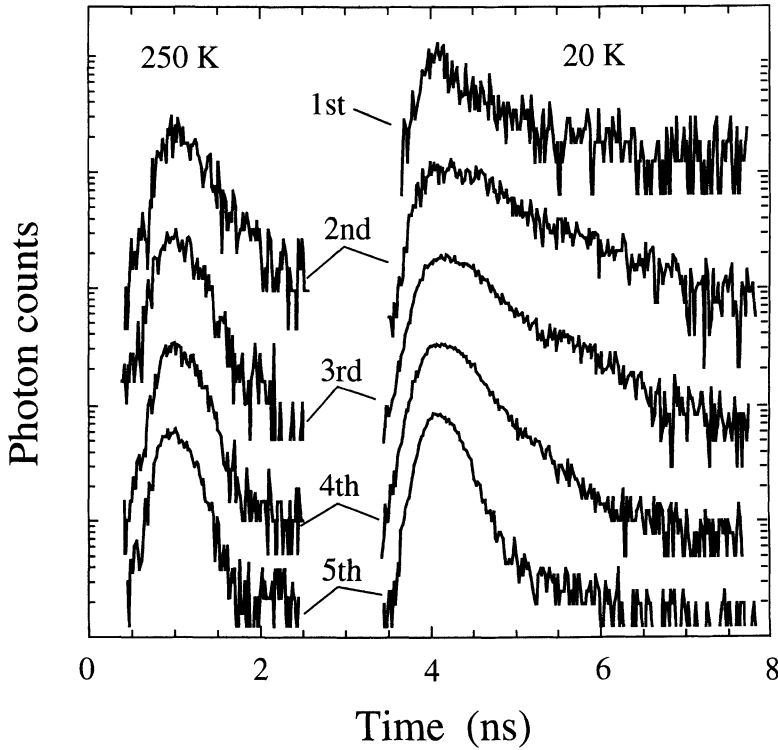


Figure 2.44: Time-resolved PL decay for each sublevel of ALS-grown quantum dots. [83]

(From Physical Review B, Vol. 54, Mukai et al., (1996), Copyright 1996 by the American Physical Society)

the energy relaxation time as 500 ps for the relaxation from the second sublevel to the ground state, and the decay as 30 ps for the fifth sublevel at 250 K.

More recently, the energy relaxation time was evaluated by measuring the modulation response of the quantum dot lasers. We can discuss the modulation response of quantum dot lasers by use of the well-developed modulation theory of quantum well lasers [84]. The maximum modulation bandwidth is given by

$$f_{\max} = \frac{1}{\sqrt{2}\pi \left(\tau_p + \frac{\tilde{\epsilon}\chi}{v_g G'} \right)} \quad (2.3)$$

where $\tilde{\epsilon}$ is the gain compression factor, χ is the factor that expresses the ratio of total carriers to that at the ground state [84,85], and G' is the differential gain. The attractive feature of the quantum dot at first sight is that we can get a large differential gain to give a large $f_{\max}$ when the inhomogeneous broadening is small. However, factors $\tilde{\epsilon}$ and χ limit the maximum modulation bandwidth. The gain compression factor $\tilde{\epsilon}$ expresses the saturation of gain for high optical power. The $\tilde{\epsilon}$ and χ depend on the energy relaxation rate of carriers. When the energy relaxation rate is slow, carriers in the ground state are depleted because of the slow supply rate of carriers from the upper sublevel or from the barrier layer for high optical power. This gives large value of $\tilde{\epsilon}$. The slow energy relaxation rate also means a larger number of carriers in the upper levels to give a larger value of χ . Despite an attractive feature of the large differential gain for small inhomogeneous broadening, the large product of $\tilde{\epsilon}\chi$ cancels out this advantage. This limits the modulation response of quantum dot lasers.

Klotzkin et al. performed a detailed study of the small signal modulation response of quantum dot lasers [86]. These authors used SK-mode-grown quantum dots and fabricated the ridge waveguide lasers with single-dot layers. The threshold current is 15–30 mA, and the lasing wavelength is about 1 μm . The 3dB modulation bandwidth was limited to 4.5 GHz. Through comparison of the small signal modulation response with the theory, the authors concluded that the modulation response is well explained by assuming the carrier capture time of 20–40 ps, which includes the carrier diffusion time in the barrier laser and the energy relaxation time in the dot. The diffusion time is on the order of a few ps. The capture time is mostly determined by the energy relaxation time.

Thus the experimental and theoretical studies suggest that the energy relaxation of carriers in the quantum dot is on the order of several tens to a hundred ps, although the complete correspondence with the relaxation mechanism is yet uncertain. The relaxation time is very large when compared with that of a few hundred fs in quantum well lasers. This limits the modulation bandwidth as studied by Klotzkin et al. [86] and is the cause of higher sublevel lasing at high operation power levels, as observed in the columnar dot lasers. What is important is to discover a new design of quantum dots to remove the limitation imposed by the long carrier energy relaxation time. One of the ideas is to use the tunnel carrier injection as proposed by Kamath et al. [87], and there may be many other new approaches. The new design of quantum dots and quantum dot lasers together with improved fabrication technologies, as discussed, would realize low-power-consumption, high-response-speed lasers for practical applications.

2.4.3 Other applications of quantum dots for optical devices: an optical intensity modulator

Other than the quantum dot laser, we can consider a variety of applications for the quantum dot, owing to its unique feature of three-dimensional carrier confinement that provides us with new optical properties. Possible applications so far discussed are, for example, use of quantum dots as a nonlinear medium [88], infrared optical detectors using subband transitions [89], and optical intensity modulators for long-distance transmission systems [90]. Here, we pick up the quantum dot intensity modulator as an interesting application and discuss its feasibility in more detail, leaving the discussion of the application to the nonlinear devices and the infrared photodetector to the references [88,89].

In long-distance and high-bit-rate communication systems, the chromatic dispersion of the optical fiber limits the transmission distance when a light source with wavelength chirping is used. The direct modulation of distributed feedback (DFB) lasers causes the wavelength chirping. At 10 Gb/s modulation, the chirping width is about 0.34 nm (at -20 dB full width of time-averaged spectrum) [91]. This causes pulse broadening and limits the transmission distance. To reduce the wavelength chirping, an external modulation scheme could be used. The MQW modulator, which uses the Stark shift of the exciton peak wavelength, has been developed. This, however, causes wavelength chirping, though much smaller than with direct modulation of lasers, due to the phase modulation of light caused by the refractive index change associated with the change in the absorption coefficient. The wavelength chirping of the intensity modulator is given by [92]:

$$\Delta\lambda(t) = -\frac{\alpha_E \lambda^2}{4\pi c} \frac{d \ln P(t)}{dt}, \quad (2.4)$$

where $P(t)$ is the time change of the modulated light power and α_E is given by

$$\alpha_E = 2k \frac{\partial \eta / \partial V}{\partial \alpha / \partial V}, \quad (2.5)$$

which is a ratio of the change of the refractive index η to the change in absorption coefficient α for the change of applied voltage to the modulator. As can be seen from Eq. (2.5), blue chirping takes place at the onset of the pulse and red chirping takes place at the end of the pulse when the α_E value is positive. This chirping causes the pulse broadening for the neg-

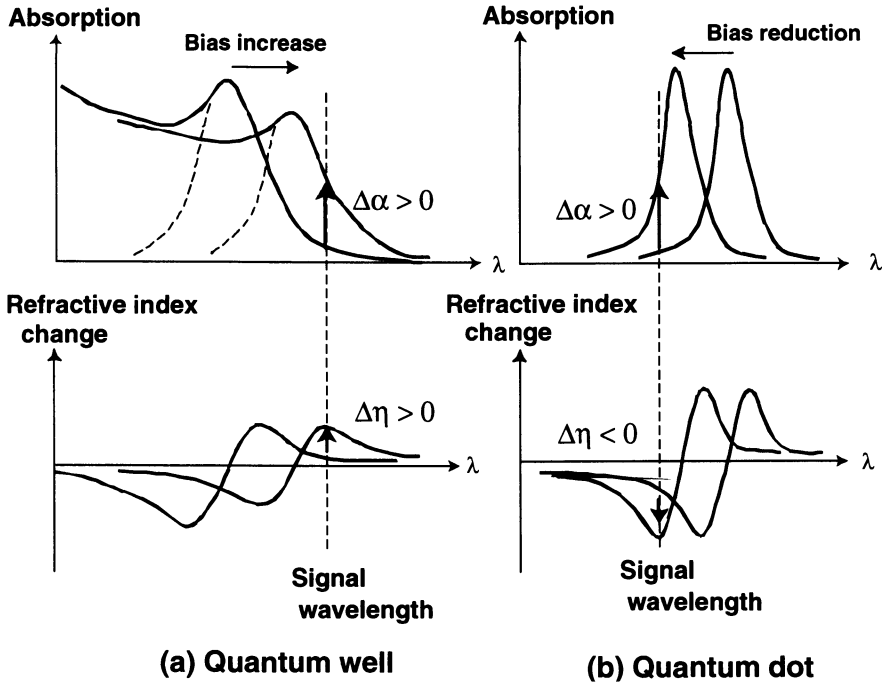


Figure 2.45: Comparison of the absorption coefficient change and refractive index change in the MQW modulator (a) and the quantum dot modulator (b). MQW refractive index change is schematically illustrated only for excitations.

ative dispersion of a fiber. The fiber dispersion is typically -17ps/nm/km at $1.55\mu\text{m}$ where the transmission loss is the lowest. If we could realize a negative α_E parameter, we can obtain pulse compression for transmission through a dispersive fiber. This improves the signal-to-noise ratio and enables long-distance transmission.

Sahara et al. proposed a modulator with a negative α_E parameter by the use of quantum dots [90]. Figure 2.45 compares the MQW intensity modulator and the quantum dot modulator. In the case of the MQW modulator, an electric field is applied to give absorption, i.e., the signal wavelength is on the red side of the exciton absorption peak. Corresponding to the shift of exciton wavelength, the refractive index increases to give a positive α_E parameter. On the other hand, in the case of quantum dots, the signal wavelength can be put at the blue side of the absorption peak owing to the discrete absorption peak of quantum dots. With the signal wavelength on the

blue side under the application of the electric field, the reduction of the electric field causes the wave length shift of the absorption peak to increases the absorption at the signal wavelength. Correspondingly, the refractive index reduces. The refractive index change is opposite to that of the quantum well in order to give a negative α_E value.

Sahara et al. assumed the amount of the shift of the absorption peak to be the same as that of the Stark shift of the MQW and calculated the change in the absorption and α_E parameter. With the assumption of the 17 meV of e^{-1} half-width of the inhomogeneous broadening, with a 0.125 dot-filling factor, they calculated the absorption change and the α_E parameter for a change in bias voltage from 130 kV/cm² to 0 V/cm². For a device length of 100 μ m and an optical confinement factor of 0.25 for the layer containing the dots and barrier, an extinction ratio of 15 dB was calculated. The α_E parameter is negative for the filed of 130 kV/cm to 80 kV. The negative α_E parameter gives blue chirping at the optical pulse end and gives pulse compression when transmitted through a negative dispersion optical fiber.

The calculation by Sahara et al. at the time was a feasibility study for the use of quantum dots. However, the idea of a negative α_E modulator uses an interesting feature of the quantum dot, i.e., we can obtain a transparent region on the shorter wavelength side (or blue side) of the absorption peak, in contrast to the quantum well or bulk, where there is no transparent region on the blue side of the band edge. In this respect, this is an example of an application that uses the unique nature of the quantum dots.

2.4.4 Self-assembled QD hole-burning memories

SK-mode self-assembled QDs show a large photoluminescence (PL) line width of as large as 80–100 meV, primarily due to the size fluctuation along the direction of growth [93]. This large inhomogeneous broadening (Γ_i) is suitable for wavelength-domain multiplication optical memory [94] or hole-burning memory. QDs can be considered to have narrow homogeneous broadening (Γ_h) due to δ -function-like density of states, which gives a large wavelength-domain multiplicity of Γ_i/Γ_h . This can lead to a high-density optical memory on a Tbit scale if the multiplicity is up to 100 or higher. This section describes preliminary observations of spectral hole burning in photocurrent spectra of InAs self-assembled QDs [95] and discusses the characteristics of spectral holes.

Figure 2.46 is a schematic drawing of the sample structure used in this experiment. Five-stacked InAs QD layers were inserted in a pin-junction

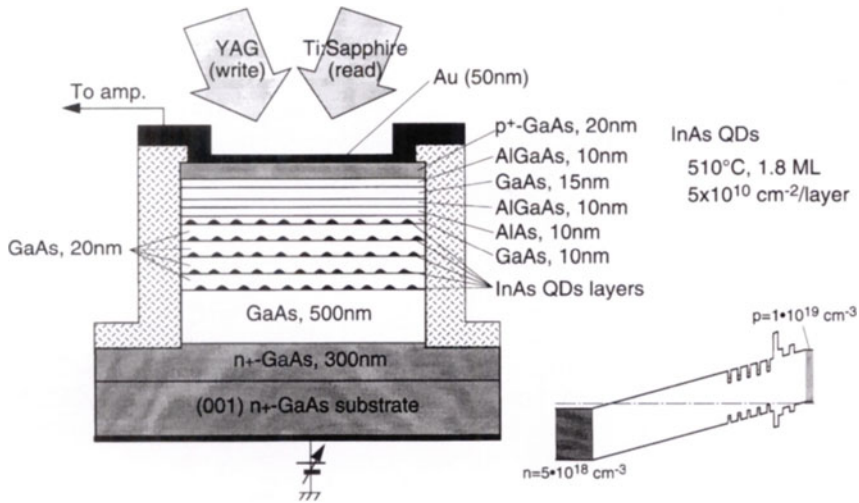


Figure 2.46: Schematic sample structure used in this experiment. The five layers of InAs self-assembled quantum dots (QDs) were embedded in a pin diode. The diameter of the sample was $500\mu\text{m}$. [95]

(From IEEE J. Select. Topics of Quantum Electron. Vol. 4, No. 5, Sugiyama et al., Copyright 1998, with permission from IEEE)

GaAs diode. The corresponding band diagram is also shown. Each QD layer formed by the nominal 1.8ML InAs has a dot density of $5 \times 10^{10}\text{ cm}^{-2}$ and is separated by 20nm GaAs interval layers to isolate the QD layers electronically. A GaAs/AlGaAs quantum well was also inserted beside the AlAs barrier layer to calibrate flat band voltage. The wafer was mesa-etched at $500\mu\text{m}$ diameter and was fabricated with nonalloy ohmic contact on top of the heavily p-doped layer. Another contact to the *n*-type region was fabricated at the bottom of the substrate. To obtain a long lifetime of the spectral hole, an electric field was applied to allow either the electrons or the holes tunnel from the QDs. The top contact on the p-GaAs was transparent, with a thin Au film (50nm) to irradiate the QDs. As a reference, a pin diode without InAs QDs was also fabricated.

The sample was continuously irradiated with a continuous-wave (CW) YAG laser at 1,064nm as a writing light source and a tunable CW Ti:Sapphire (TiS) laser as a reading light source. The band width of the TiS laser was less than 40 GHz ($165\mu\text{eV}$). The width of the YAG laser was less than 6.4 GHz ($27\mu\text{eV}$) under the reading laser power of less than $80\mu\text{W}$. Both lasers of about 2mm in diameter were aimed, unfocused, at the sample. The photocurrent corresponding to the tunable TiS laser was detected with a lock-in amplifier through a current amplifier to cut off the photocurrent

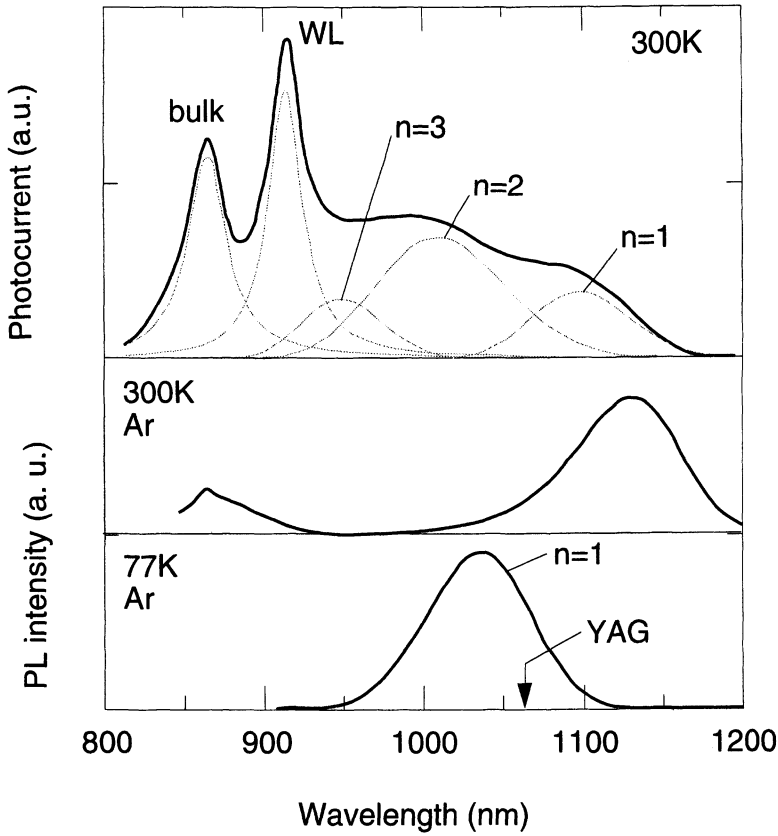


Figure 2.47: Photocurrent spectrum (top) measured with a WL lamp at 300 K with fitted Gaussian results for QD and Lorentzian results for wetting layer and GaAs. Photoluminescence (PL) spectra at 300 K (middle) and 77 K (bottom) with Ar laser excitation are also shown. [95]

(From IEEE J. Select. Topics of Quantum Electron. Vol. 4, No. 5, Sugiyama et al., Copyright 1998, with permission from IEEE)

of the CW YAG laser. During photocurrent measurement, the wavelength of the TiS laser was monitored by a wavemeter.

Prior to measuring the hole burning, the PL and the photocurrent spectra of the sample were measured. Figure 2.47 shows the photocurrent spectrum, equivalent to the absorption spectrum, obtained at 300 K with a Wl_2 lamp. Broad absorption bands related to InAs QDs fit well with three Gaussian curves. For example, the ground state was fitted with full width at half maximum (FWHM) of 82.5 meV. The higher two peaks that fit with Lorentzian curves were GaAs and the wetting layers. The peak energy of

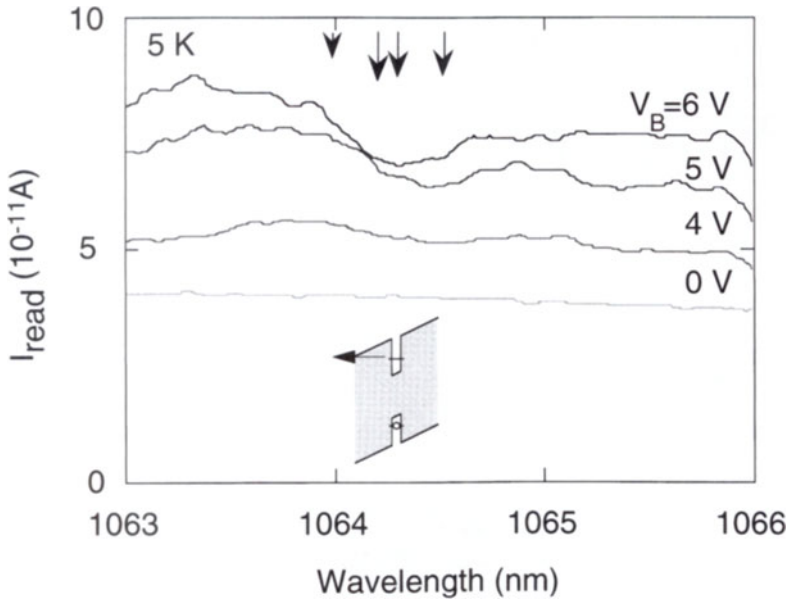


Figure 2.48: Bias voltage (V_B) dependence of photocurrent (I_{read}) spectra, corresponding to a CW Ti:Sapphire (TiS) laser as a reading light source under irradiation of a CW YAG laser as a writing light source measured at 5 K. Irradiated wavelengths of the YAG laser are indicated by the arrows. [95] (From IEEE J. Select. Topics of Quantum Electron. Vol. 4, No. 5, Sugiyama et al., Copyright 1998, with permission from IEEE)

PL at 300 K was 40 meV smaller than that of ground-state absorption. The energy difference is caused by carrier redistribution due to thermal energy [96]. PL spectra at 77 K can be considered to correspond to the ground states of QDs because the sample was excited under weak conditions. The thermal carrier distribution to the excited states is suppressed when compared to 300 K. The ground-state absorption obtained by fitting to the photocurrent spectra at 300 K agreed well with the 77 K PL spectra with consideration of the slight blue shift due to temperature change. Therefore, the photon energy of the YAG laser is estimated to be lower than the peak energy of the ground-state distribution.

Under irradiation of a YAG laser of 8 mW and a TiS laser of 32 μ W, the photocurrent response, I_{read} , to the TiS laser was measured. Figure 2.48 shows the dependence of the external bias voltage (V_B) on I_{read} at 5 K. As V_B increases, a dip appears at the wavelength of the YAG laser lines as indicated by the arrows. The depth of the dip increased as $V_B > 4$ V.

The result can be explained as follows. As V_B increases, the internal electric field increases, enhancing photogenerated carriers in QDs to tunnel from the ground state through the triangular potential barrier formed by the field. Electrons are more likely to tunnel because of their smaller effective mass. Here, the carriers are electrons; the heavy holes can be considered to stay in the QDs, as shown in the inset of Fig. 2.48. The residual holes inhibit additional optical absorption of the TiS laser under the intense irradiation of the YAG laser. As the electron tunneling rate becomes comparable with or greater than the radiative recombination rate in the QDs, spectral holes can be expected to appear in I_{read} . Such behavior was not observed in a reference sample. Therefore, the dip shows that the formation of spectral holes originating from InAs QDs. The roll-off that occurs just below 1,066 nm can be considered due to the data averaging process, because the averaged data near the edge point is strongly influenced by the shape of the noise.

Figure 2.49 shows the temperature dependence of ΔI_{read} , which is I_{read} (with YAG irradiation) $- I_{\text{read}}$ (without YAG irradiation), in the range of 5 K to 60 K under a bias voltage of 6 V. The subtraction was to decrease the background photocurrent induced by the reading light, since the lifetime of spectral holes is much shorter than the measuring time (see Section 4.2). As the temperature increases, the hole depth decreases until 40 K and vanishes to 60 K. This suggests that carriers were thermally excited, possibly, by the wetting layers. Similar excitation of carriers was reported in QD lasers [97]. However, the onset of thermally assisted tunneling cannot be excluded. Further study is needed to suppress phonon interaction and improve the operation temperature using self-assembled QDs.

Figure 2.50 shows the writing power dependence of ΔI_{read} at 5 K. The difference of emission lines of the YAG laser was from the use of a different YAG laser. The spectral hole exhibits similar features with 4 mW and 8 mW. With the intense power of 20 mW, the spectral hole was broadened with a width of more than twice 8 mW. According to Bloch's equation for two or three level systems, spectral hole width increases, and its depth saturates, maintaining a Lorentzian function shape as the writing power increases. Although this picture is for an ideal case and should be modified with other factors for a real case, the obtained spectrum for 20 mW seems not to be covered by this model because the spectral hole consists of a zero-phonon hole of around 1,064 nm as well as "antiholes" and side holes outside of the antiholes. The side holes can be considered to be due to phonon-assisted light absorption induced by the increased irradiation of the writing light. Further study is needed in this area.

To investigate the spectral hole width, we performed numerical fitting to the experimental spectra. We took into account the bandwidth of writing

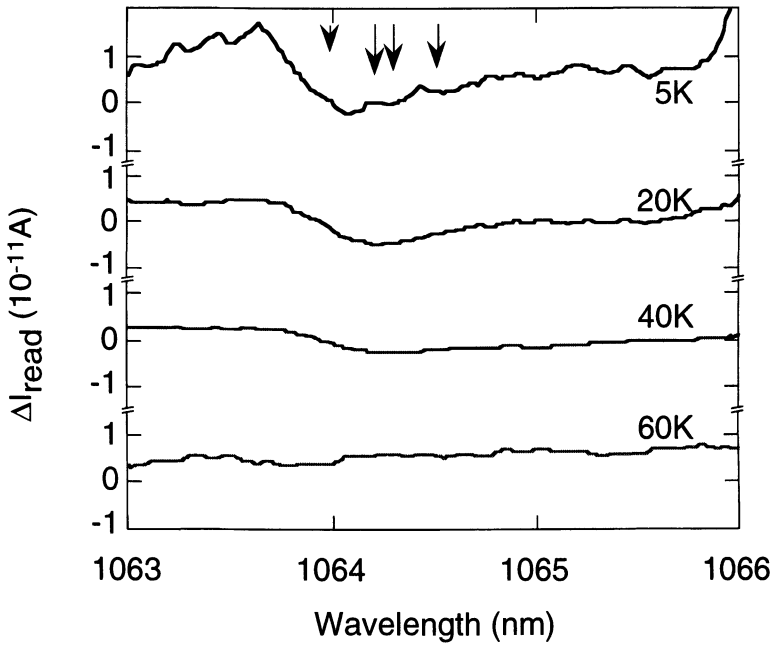


Figure 2.49: Temperature dependence of ΔI_{read} in the range from 5 K to 60 K. Irradiated wavelengths of the YAG laser are indicated by the arrows. ΔI_{read} is the difference in the reading photocurrent with and without the writing lasers. [95]

(From IEEE J. Select. Topics of Quantum Electron. Vol. 4, No. 5, Sugiyama et al., Copyright 1998, with permission from IEEE)

and reading lasers. The YAG lasers used in this experiment have a few or several emission lines 0.1 to 0.2 nm apart; each line width is less than 6.4 GHz ($27 \mu\text{eV}$). The multiple emission lines are caused by multiple transition levels in YAG crystals or in longitudinal mode limited by cavity length. In contrast, the reading laser of a CW TiS laser has more closely spaced emission lines, which can be approximated by a single Gaussian distribution with an estimated width of less than 40 GHz ($165 \mu\text{eV}$). We applied the Gaussian function for the CW TiS laser spectrum and the Lorentzian function for the absorption spectrum of the QDs for each emission line of the YAG laser, and then calculated the convolution integral between these two types of functions. In the calculation, we used an FWHM of $165 \mu\text{eV}$ for the Gaussian function. The FWHM of the Lorentzian function, Γ , was used as a fitting parameter. Figure 2.51 shows the calculated results and the experimental results of a writing power with 8 mW at 5 K. Taking into account three emission lines of writing the YAG laser as measured by a

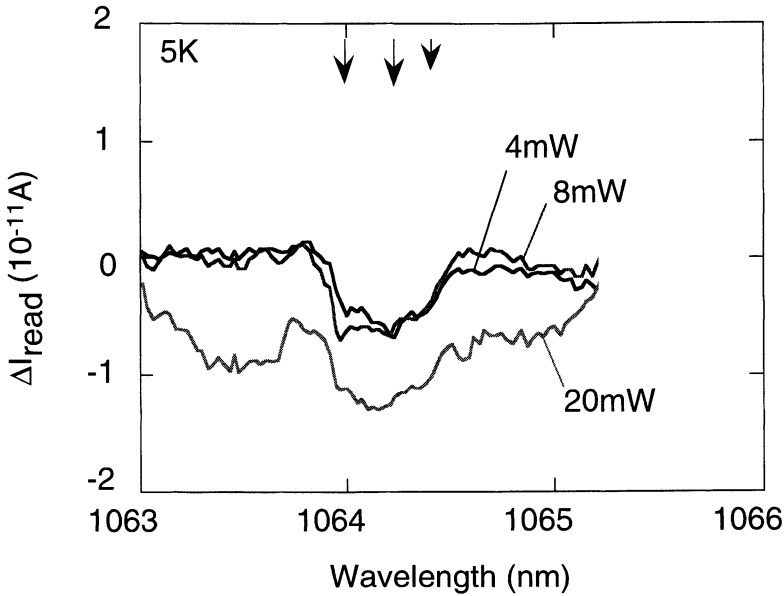


Figure 2.50: Writing power dependence of ΔI_{read} at 5K. Irradiated wavelengths of the YAG laser are indicated by the arrows. [95]
(From IEEE J. Select. Topics of Quantum Electron. Vol. 4, No. 5, Sugiyama et al., Copyright 1998, with permission from IEEE)

high-resolution monochromator, the calculated curve was a good match to the experimental result, which cannot be fitted to any single Gaussian or Lorentzian function. The obtained Γ was about $100\mu\text{eV}$. Assuming a Lorentzian distribution for both the YAG laser spectra and the absorption spectrum of the QDs, the Γ_h was estimated to be less than $80\mu\text{eV}$. This value may be overestimated because the experimental hole width is usually $2\Gamma_h$ for Lorentzian absorption spectra [98].

As is clearly seen in the experimental curve in Fig. 2.51, the spectrum seems to show “antiholes” on both sides of the writing wavelength, which is familiar in photophysical hole-burning cases. The antihole also appeared in Figs. 2.48 to 2.50. The formation mechanism and relationship of the antiholes with the spectral hole vis-à-vis energy is not yet clear. It is necessary, however, to study the origin of antiholes for the use of self-assembled QDs in hole-burning memory.

The lifetime of a spectral hole can be derived from the time evolution of the hole depth. In this case, however, the writing and reading lasers constantly irradiate the sample, and the corresponding I_{read} requires a much longer measurement time than the lifetime of the spectral hole.

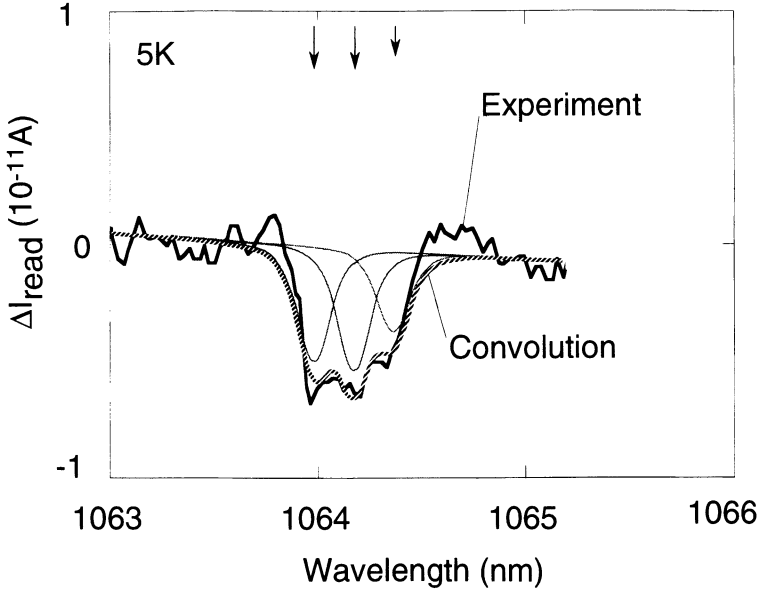


Figure 2.51: Comparison of a spectral hole with the experimental result of writing YAG laser power of 8 mW at 5 K and fitted curves assuming Gaussian for reading TiS laser and Lorentzian for absorption in QDs. [95] (From IEEE J. Select. Topics of Quantum Electron. Vol. 4, No. 5, Sugiyama et al., Copyright 1998, with permission from IEEE)

This makes direct determination of the lifetime difficult from the hole spectrum. Therefore, we have begun to consider a rate equation on a steady-state carrier.

We assume that resonantly photogenerated electrons escape much faster from QDs than do holes, so that only holes can be treated with the rate equation and that two holes can occupy the ground state in a QD due to spin degrees of freedom. In a steady state, the equation can be written as

$$\frac{\sigma n}{\tau} = c \frac{P}{S} \left(1 - \frac{\sigma}{2}\right) n, \quad (2.6)$$

where σ is the probability of hole occupancy at the ground state of QD, τ is the recombination time of the hole that equals the spectral hole lifetime, n is the number of QDs irradiated by the YAG laser, P is the incident power of the laser, S is the incident area of the laser, and c is the coefficient of proportionality for the right-hand term having a dimension of cm^2/Ws . The

left-hand term shows the recombination rate; the right-hand term shows the generation rate. Photocurrent, I_{ph} , can be also written as

$$I_{ph}(P) = qcP(1 - \sigma/2)n \quad (2.7)$$

$$= \frac{2qnS}{\tau} \left(1 - \frac{1}{1 + cP\tau/2S} \right), \quad (2.8)$$

where Eq. (2.8) is derived from Eq. (2.6). Meanwhile, the incident power of the reading laser is two orders smaller than that of the writing laser. The I_{read} that is not subtracted by I_{read} without YAG laser irradiation can be written as

$$I_{read}(P) = \left. \frac{\partial I_{ph}(P)}{\partial P} \right|_{P=P_{write}} \cdot P_{read} \quad (2.9)$$

$$= \frac{qcn}{\left(1 + \frac{cP_{write}\tau}{2S} \right)^2} P_{read}, \quad (2.10)$$

where P_{write} is the incident power of the writing laser and P_{read} is that of the reading laser. The c can be derived from Eq. (2.10) with the experimental result of $I_{read}(P_{write} = 0) = 25 \text{ pA}$. Since the laser beam was not focused on the sample surface, P_{read} of $32 \mu\text{W}$ was calibrated by the squared ratio of the sample diameter ($500 \mu\phi$) to the beam diameter ($2 \text{ mm}\phi$). The obtained c was 2.3×10^5 . Using the experimental parameters of $n = 5 \times 10^{10} \times 5 \times (27 \mu\text{eV}/82.5 \text{ meV})$, $I_{read}(P_{write} = 8 \text{ mW}) = 19 \text{ pA}$, and $S = 2 \text{ mm}\phi$, τ of $1.2 \times 10^{-6} \text{ s}$ was obtained.

Considering the practical use of self-assembled QDs for hole-burning memory, the obtained hole lifetime should be increased. However, QDs using the type I heterostructure like InAs/(Al)GaAs experience an experimental problem in that reverse bias for the pin diode to extract electrons from QDs accompanies an undesirable dark-current increase as the bias increases. This process loses some information through recombination of the residual holes with electrons, reducing the hole lifetime.

To solve this problem, we propose the use of self-assembled QDs consisting of a type II heterostructure. This structure does not need an external bias perpendicular to the hetero-interface to extract an electron from QD, which may drastically reduce dark current. Figure 2.52 shows the energy band diagram using GaSb QD or InP QD that can be grown by MBE or GSMBE or MOVPE. Photogenerated electrons tunnel to the X point of AlAs through a potential barrier and are spatially separated from holes, which reduces the recombination rate. Even a small bias parallel to the

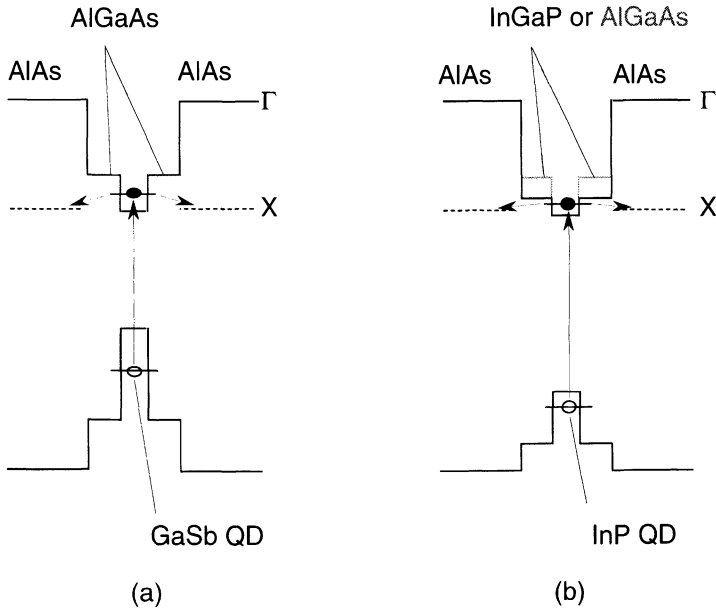


Figure 2.52: Schematic band diagram of QD using type II heterostructure: (a) GaSb/AlGaAs/AlAs system; (b) InP/InGaP(or AlGaAs)/AlAs system. The gray line in the conduction band of (b) indicates the case of AlGaAs as a tunneling barrier. [95]

(From IEEE J. Select. Topics of Quantum Electron. Vol. 4, No. 5, Sugiyama et al., Copyright 1998, with permission from IEEE)

hetero-interface may contribute to a further decrease in the recombination rate. In such a case, GaAs could be used in place of AlAs. As a consequence, the spectral hole lifetime can be increased. Using this structure, hole burning by an all-optical measurement can be also obtained.

If the spectral hole lifetime increases much more than the estimated lifetime, it becomes difficult to measure the spectral hole directly by photocurrent, and an all-optical measurement will be required. A single QD layer is less efficient in measuring spectral holes due to the small S/N ratio. The S/N ratio of hole-burning memory using QDs increases as the number of QDs increases [94]; however, the surface coverage of QDs per one layer is as low as 15% for InAs QD of $5 \times 10^{10} \text{ cm}^{-2}$. Therefore, a multiple stacked QD structure is needed to obtain a realistic S/N ratio. Although an InAs/GaAs QD is not always appropriate for the observation of persistent spectral hole-burning phenomena, this material is easy to prepare in a multiple

stacked QD structure and is useful for examining the possibility of all-optical measurement of stacked QD layers. Optical absorption spectra of self-assembled QDs has been little reported, and further investigation is necessary to study hole-burning memory.

2.5 Summary

This chapter describes recent advances in quantum dot fabrication technologies, focusing on our self-formed quantum dot technologies including TSR quantum dots and SK-mode self-assembled quantum dots. As is described in this chapter, there are many possible device applications such as quantum dot tunneling memory devices, quantum dot floating-dot gate FETs, quantum dot lasers, and quantum dot hole-burning memory devices. The quantum dot laser applications seem to be the most practicable among these applications. However, many problems remain to be solved before even this application becomes practical. The most important issue is to control the size and position of self-assembled quantum dots more precisely, with an accuracy on an atomic scale. The confinement must be enough to keep the separation energy between quantized energy levels high enough to get high-temperature characteristics. The lasing oscillation frequency should be fixed at $1.3\mu\text{m}$ or $1.5\mu\text{m}$ for optical communication. Phonon bottleneck problems should be solved by the optimization of device structures. Fortunately, there is much activity in the area of quantum dot lasers and, therefore, many breakthroughs will be made, along with the exploration of other new application areas.

At present, the most promising electrical application of quantum dots seems to be quantum dot embedded FETs. However, if we change our inspiration towards the boundary regions between electrical and other device areas, we believe there are many other interesting and useful application areas, including active image sensors, spin-electronic devices, and quantum computers.

ACKNOWLEDGMENTS

This work was partially conducted under the management of FED as a part of the MITI R & D program (Quantum Functional Device project) supported by NEDO. The authors would like to thank Prof. N. Miura and Dr. K. Uchida at ISSP of Tokyo University for the magnet-PL experiments and for valuable discussions. We also acknowledge CL measurements and useful discussions with Dr. T. Sekiguchi at IMR of Tohoku University.

The authors would like to give a hearty thank you to their collaborators, Drs. Y. Awano, T. Futatsugi, M. Sugawara, H. Shoji, K. Mukai, O. Ueda

of Fujitsu Laboratories, Ltd., and Professor S. Muto of Hokkaido University for fruitful discussions and support. We also gratefully thank Drs. H. Ishikawa and H. Imai of Fujitsu Laboratories Ltd. for their encouragement and strong support throughout this work.

REFERENCES

- [1] P. M. Petroff, A. C. Gossard, R. A. Logan, and W. Wiedmann, *Appl. Phys. Lett.* **41**, 635–638 (1982).
- [2] M. A. Reed, R. T. Bate, K. Bradshaw, W. M. Duncan, W. R. Frenseley, J. W. Lee, and H. D. Shih, *J. Vac. Sci. Technol. B* **4**, 358–360 (1986).
- [3] Y. Miyamoto, M. Cao, Y. Shingai, K. Furuya, Y. Suematsu, K. G. Ravikumar, and S. Arai, *Jpn. J. Appl. Phys.* **26**, L225–L227 (1987).
- [4] D. Gershoni, H. Temkin, G. J. Dolan, J. Dunsmuir, S. N. G. Chu, and M. B. Panish, *Appl. Phys. Lett.* **53**, 995–997 (1998).
- [5] P. M. Petroff, A. C. Gossard, and W. Wiegmann, *Appl. Phys. Lett.* **45**, 620–622 (1984).
- [6] T. Fukui and H. Saito, *Appl. Phys. Lett.* **50**, 824–826 (1987).
- [7] E. Kapon, S. Simhony, R. Bhat, and D. M. Hwang, *Appl. Phys. Lett.* **55**, 2715–2717 (1989).
- [8] E. Kapon, K. Kash, E. M. Clausen, D. M. Hwang Jr., and E. Colas, *Appl. Phys. Lett.* **60**, 477–479 (1992).
- [9] A. R. Goni, L. N. Pfeiffer, K. W. West, A. Pinczuk, H. U. Baranger, and H. L. Stormer, *Appl. Phys. Lett.* **61**, 1956–1958 (1992).
- [10] L. N. Pfeiffer, H. L. Stormer, K. W. Baldwin, K. W. West, A. R. Goni, A. Pinczuk, R. C. Ashoori, M. M. Dignam, and W. Wegscheider, *J. Crystal Growth* **127**, 849–857 (1993).
- [11] W. Wegscheider, L. N. Pfeiffer, A. Pinczuk, K. W. West, M. M. Dignam, R. Hull, and R. E. Leibenguth, *J. Crystal Growth* **150**, 285–292 (1995).
- [12] Y. Sakuma, Y. Sugiyama, S. Muto, and N. Yokoyama, *21st Int. Symp. Compound Semicon.*, San Diego (1994).
- [13] Y. Sugiyama, Y. Sakuma, S. Muto, and N. Yokoyama, *Appl. Phys. Lett.* **67**, 256–258 (1995).
- [14] Y. Tarui, Y. Komiya, and Y. Harada, *J. Electrochem. Soc.* **118**, 118–122 (1971).
- [15] E. Kapon, D. M. Hwang, and R. Bhat, *Phys. Rev. Lett.* **63**, 430–433 (1989).
- [16] S. Tsukamoto, Y. Nagamune, M. Nishioka, and Y. Arakawa, *J. Appl. Phys.* **71**, 533–535 (1992).
- [17] X. L. Wang, M. Ogura, and H. Matsuhata, *Appl. Phys. Lett.* **66**, 1506–1508 (1995).

- [18] T. Sekiguchi, Y. Sakuma, Y. Awano, and N. Yokoyama, *J. Appl. Phys.* **83**, 4944–4950 (1998).
- [19] Y. Nagamune, M. Nishioka, S. Tsukamoto, and Y. Arakawa, *Appl. Phys. Lett.* **64**, 2495–2497 (1994).
- [20] O. Akimoto and H. Hasegawa, *J. Phys. Soc. Jpn.* **22**, 181–191 (1967).
- [21] H. Q. Hou, W. Staguhn, N. Miura, Y. Segawa, S. Takeyama, Y. Aoyagi, and J. M. Zhou, *Solid State Commun.* **74**, 687–691 (1990).
- [22] Y. Nagamune, Y. Arakawa, S. Tsukamoto, M. Nishioka, S. Sasaki, and N. Miura, *Phys. Rev. Lett.* **69**, 2963–2966 (1992).
- [23] O. Kayser and J. Cryst. Growth **107**, 989–998 (1991).
- [24] B. P. Keller, S. Keller, H. Herrnberger, J. Lenzner, S. Nilsson, and W. Seifert, *J. Crystal Growth* **140**, 33 (1994).
- [25] M. Y. Yen and T. W. Haas, *Appl. Phys. Lett.* **56**, 2533–2535 (1990).
- [26] K. Kamon, M. Shimazu, K. Kimura, M. Mihara, and M. Ishii, *J. Crystal Growth* **84**, 126–132 (1987).
- [27] D. K. Biegelsen, R. D. Bringans, J. E. Northrup, and L. E. Swartz, *Phys. Rev. Lett.* **65**, 452–455 (1990).
- [28] J. M. Garcia, G. Medeiros-Ribeiro, K. Schmidt, T. Ngo, J. L. Feng, A. Lorke, J. Kotthaus, and P. M. Petroff, *Appl. Phys. Lett.* **71**, 2014–2016 (1997).
- [29] Q. Xie, A. Madhukar, P. Chen, and N. P. Kobayashi, *Phys. Rev. Lett.* **75**, 2542–2545 (1995).
- [30] G. S. Solomon, J. A. Trezza, A. F. Marshall, and J. S. Harris Jr., *Phys. Rev. Lett.* **76**, 952–955 (1996).
- [31] M. S. Miller, *Jpn. J. Appl. Phys.* **36**, 4123–4128 (1997).
- [32] K. Mukai, N. Ohtsuka, M. Sugawara, and S. Yamazaki, *Jpn. J. Appl. Phys.* **33**, L1710–L1712 (1994).
- [33] M. Grundmann, N. N. Ledentsov, O. Steir, D. Bimberg, V. M. Ustinov, P. S. Kop'ev, and Zh. I. Alferov, *Appl. Phys. Lett.* **68**, 979–981 (1996).
- [34] I. N. Stranski and L. Krastanow, *Akad. Wiss. Wien Math.-Natur. IIb* **146**, 797–810 (1937).
- [35] W. J. Schaffer, M. D. Lind, S. P. Kowalczyk, and R. W. Grant, *J. Vac. Sci. Technol. B* **1**, 688 (1983).
- [36] B. F. Lewis, T. C. Lee, F. J. Grunthaner, A. Madhukar, R. Fernandez, and F. F. Masanian, *J. Vac. Sci. Technol. B* **2**, 419–420 (1984).
- [37] F. Glas, C. Guill, P. Henoc, and F. Haussy, *Inst. Phys. Conf. Ser. No.* **87**; Section 2, 71–76 (1987).
- [38] F. Houzay, C. Guille, J. M. Moisson, P. Henoc, and F. Barth, *J. Crystal Growth* **81**, 67–72 (1987).

- [39] M. Tabuchi, S. Noda, and A. Sasaki, *Science and Technology of Mesoscopic Structures*, edited by S. Namba, C. Hamaguchi, and T. Ando, 379–384 (1992).
- [40] D. Leonard, M. Krishnamurthy, C. M. Reaves, S. P. Denbaars, and P. M. Petroff, *Appl. Phys. Lett.* **63**, 3203–3205 (1993).
- [41] J.-Y. Marzin, J.-M. Gerard, A. Israel, D. Barrier, and G. Bastard, *Phys. Rev. Lett.* **73**, 716–719 (1994).
- [42] J. Oshinowo, M. Nishioka, S. Ishida, and Y. Arakawa, *Appl. Phys. Lett.* **65**, 1421–1423 (1994).
- [43] J. M. Moison, F. Houzay, F. Barthe, L. Leprince, E. Andre, and O. Vatel, *Appl. Phys. Lett.* **64**, 196–198 (1994).
- [44] N. Carlsson, K. Georgsson, L. Montelius, L. Samuelson, W. Seifert, and R. Wallenberg, *J. Crystal Growth* **156**, 23 (1995).
- [45] C. M. Reaves, V. Bressler-Hill, S. Varma, W. H. Weinberg, and S. P. DenBaars, *Surf. Sci.* **326**, 209 (1995).
- [46] M. Arita, A. Avramescu, K. Uesugi, I. Suemune, T. Numai, H. Machida, and N. Shimoyama, *Jpn. J. Appl. Phys.* **36**, 4097 (1997).
- [47] S. Tanaka, S. Iwai, and Y. Aoyagi, *Appl. Phys. Lett.* **69**, 4096 (1996).
- [48] S. Tanaka, H. Hirayama, P. Ramvall, and Y. Aoyagi, *Proc. 2nd Int. Conf. Nitride Semiconductors*, p. 48 (1997).
- [49] Q. Xie, A. Madhukar, P. Chen, and N. P. Kobayashi, *Phys. Rev. Lett.* **75**, 2542–2545 (1995).
- [50] G. S. Solomon, J. A. Trezza, A. F. Marshall, and J. S. Harris Jr., *Phys. Rev. Lett.* **76**, 952–955 (1996).
- [51] G. S. Solomon, S. K. Komarov, J. S. Harris Jr., and Y. Yamamoto, *J. Crystal Growth* **175/176**, 707 (1997).
- [52] S. Fafard, R. Leon, D. Leonard, J. L. Martz, and P. M. Petroff, *Phys. Rev. B* **50**, 8086–8089 (1994).
- [53] N. Grundmann, N. Ledentsov, R. Heitz, L. Eeckey, J. Christen, J. Bohrer, D. Bimberg, S. S. Ruvimov, P. Werner, V. Richter, J. Heydenreich, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Hopev, and Zh. I. Alferov, *Phys. Stat. Sol. (B)* **188**, 249–258 (1995).
- [54] M. Grundmann, J. Christen, N. N. Ledentsov, J. Bohrer, D. Bimberg, S. S. Ruvimov, P. Werner, V. Richter, V. Gosele, J. Heydenreich, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Hopev, and Zh. I. Alferov, *Phys. Rev. Lett.* **20**, 4043–4086 (1995).
- [55] R. Leon, P. M. Petroff, D. Leonard, and S. Fafard, *Science* **267**, 1966–1968 (1995).
- [56] D. Hessman, P. Castrillo, M.-E. Pistol, C. Pryor, and L. Samuelson, *Appl. Phys. Lett.* **69**, 749–751 (1996).

- [57] M. Sugawara, Proc. 10th Int. Conf. On InP and Related Mater., 254–261 (1998).
- [58] Y. Arakawa and H. Sakaki, Appl. Phys. Lett. **40**, 939–941 (1982).
- [59] L. Guo, E. Leobandung, and S. Y. Chou, Appl. Phys. Lett. **70**, 850–852 (1997).
- [60] A. Nakajima, T. Futatsugi, K. Kosemura, T. Fukano, and N. Yokoyama, Appl. Phys. Lett. **70**, 1742–1744 (1997).
- [61] Y. Awano, Y. Sakuma, Y. Sugiyama, C. Wirner, M. Takatsu, T. Mori, H. Nakao, M. Yamaguchi, M. Shima, and N. Yokoyama, “Electrical and Optical Properties of Quantum Dot Structures; Magnetic-confined and Tetrahedral-Shaped Recess (TSR) Quantum Dots,” Extended Abstracts of 15th Symposium on Future Electron Devices, 46–51 (1996).
- [62] U. Bockelman and G. Bastard, Phys. Rev. B **42**, 8947–8951 (1990).
- [63] H. Benisty, C. M. Sotomayor-Torres, and C. Weisbuch, Phys. Rev. B **44**, 10945–10948 (1991).
- [64] T. Mori, S. Muto, H. Tamura, and N. Yokoyama, Jpn. J. Appl. Phys. **33**, 790–793 (1994).
- [65] C. Wirner, Y. Awano, T. Futatsugi, N. Yokoyama, T. Nakagawa, H. Bando, S. Muto, M. Ohno, and N. Miura, Jpn. J. Appl. Phys. **36**, 1958–1960 (1997).
- [66] C. Wirner, Y. Awano, N. Yokoyama, M. Ohno, N. Miura, T. Nakagawa, and H. Bando, Semicond. Sci. Technol. **12**, 998–1002 (1997).
- [67] M. Shima, Y. Sakuma, C. Wirer, T. Strutz, E. Taguchi, T. Futatsugi, Y. Awano, and N. Yokoyama, “Electron Transport Through Tetrahedral-Shaped Recess (TSR) Stacked Double Quantum Dot Structures,” *1997 IEEE International Symposium on Compound Semiconductors*, edited by M. Melloch and M. A. Reed (Institute of Physics Publishing, San Diego), (1997), pp. 539–542.
- [68] G. W. Bryant, Phys. Rev. B **44**, 3064–3069 (1991).
- [69] U. Bockelmann and G. Bastard, Phys. Rev. B **42**, 8947–8951 (1990).
- [70] E. O. Kane, J. Phys. Chem. Solids **1**, 249–261 (1957).
- [71] N. Kristaедter, N. N. Ledentsov, M. Grundmann, D. Bimberg, V. M. Ustinov, S. S. Ruvimov, M. V. Maximov, P. S. Kop’ev, Zh. I. Alferov, U. Richter, P. Werner, U. Gösele, and J. Heydenreich, Electron. Lett. **30**, 1416–1417 (1994).
- [72] H. Shoji, Y. Nakata, K. Mukai, Y. Sugiyama, M. Sugawara, N. Yokoyama, and H. Ishikawa, Electron. Lett. **32**, 2023–2024 (1996).
- [73] H. Shoji, K. Mukai, N. Ohtsuka, M. Sugawara, T. Uchida, and H. Ishikawa, IEEE Photonics Technol. Lett. **7**, 1385–1387 (1995).
- [74] H. Shoji, Y. Nakata, K. Mukai, Y. Sugiyama, M. Sugawara, N. Yokoyama, and H. Ishikawa, IEEE J. Selected Topics in Quantum Electron. **3**, 188–195 (1997).
- [75] K. Mukai, Y. Nakata, H. Shoji, M. Sugawara, K. Ohtsubo, N. Yokoyama, and H. Ishikawa, Electron. Lett. **34**, 1588–1590 (1998).

- [76] S. Fafard, S. Raymond, G. Wang, R. Leon, D. Leonard, S. Charbonneau, J. L. Merz, P. M. Petroff, and J. E. Bowers, *Surf. Sci.* **361/362**, 778–782 (1996).
- [77] K. Mukai, N. Ohtsuka, and M. Sugawara, *Appl. Phys. Lett.* **70**, 2416–2418 (1997).
- [78] S. Fafard, Z. Wasilewski, J. McCaffery, S. Raymond, and S. Charbonneau, *Appl. Phys. Lett.* **68**, 991–993 (1996).
- [79] K. Mukai, N. Ohtsuka, H. Shoji, and M. Sugawara, *Appl. Surf. Sci.* **112**, 102–109 (1997).
- [80] A. V. Uskov, F. Adler, H. Schweizer, and M. H. Pilkuhn, *J. Appl. Phys.* **81**, 7895–7899 (1997).
- [81] Al. L. Efros, V. A. Kharchenko, M. Rosen, *Solid State Comm.* **93**, 281–284 (1995).
- [82] I. Vurgaftman and J. Singh, *Appl. Phys. Lett.* **64**, 232–234 (1994).
- [83] K. Mukai, N. Otsuka, H. Shoji, and M. Sugawara, *Phys. Rev. B* **54**, 5243–5246 (1996).
- [84] R. Nagarajan, M. Ishikawa, T. Fukushima, R. S. Geels, and J. E. Bowers, *IEEE J. Quantum Electron.* **28**, 1990–2008 (1992).
- [85] H. Ishikawa and I. Suemune, *IEEE Photonic Technol. Lett.* **6**, 1315–1317 (1994).
- [86] D. Klotzkin, K. Kamath, and P. Bhattacharya, *IEEE Photonic Technol. Lett.* **9**, 1301–1303 (1997).
- [87] K. Kamath, D. Klotzkin, and P. Bhattacharya, *IEEE Lasers and Electro-Optic Society Annual Meeting*, paper Thy4, San Francisco, 1997.
- [88] H. Ishikawa, Chapter 7, *Self-assembled InGaAs/GaAs Quantum Dots, Semiconductors and Semimetals*, Vol. 60 (Academic Press, New York, 1999).
- [89] V. Ryzhii, *Semicond. Sci. Technol.* **11**, 759–765 (1996).
- [90] R. Sahara, M. Matsuda, H. Shoji, K. Morito, and H. Soda, *IEEE Photonic Technol. Lett.* **8**, 1477–1479 (1996).
- [91] K. Uomi, T. Tsuchiya, H. Nakano, M. Aoki, M. Suzuki, and N. Chninoe, *IEEE J. Quantum Electron.* **27**, 1705–1713 (1991).
- [92] F. Koyama and K. Iga, *Electron. Lett.* **21**, 1065–1066 (1985).
- [93] Y. Sugiyama, Y. Nakata, T. Futatsugi, M. Sugawara, Y. Awano, and N. Yokoyama, *Jpn. J. Appl. Phys.* **36**, L158–L161 (1997).
- [94] S. Muto, *Jpn. J. Appl. Phys.* **34**, L210–L212 (1995).
- [95] Y. Sugiyama, Y. Nakata, S. Muto, T. Futatsugi, and N. Yokoyama, *IEEE Journal of Selected Topics in Quantum Electronics*, **4**, 880–885 (1998).
- [96] Y. Sugiyama, Y. Nakata, S. Muto, K. Imamura, and N. Yokoyama, *Jpn. J. Appl. Phys.* **35**, 1320–1324 (1996).

- [97] D. Bimberg, N. N. Ledentsov, M. Grundmann, N. Kirstaedter, O. G. Schmidt, M. H. Mao, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Kopév, Zh. I. Alferov, S. S. Ruvimov, U. Gösele, and J. Heydenreich, *Jpn. J. Appl. Phys.* **35**, 1311–1319 (1996).
- [98] A. Kurita, T. Kushida, T. Izumitani, and M. Matsukawa, *Opt. Lett.* **19**, 314–316 (1994).

Chapter 3

Lithography and patterning for nanostructure fabrication

G. Seebohm and H. G. Craighead

3.1 Requirements of fabrication processes

The modern electronic integrated circuit industry has been built and continues to advance on the basis of lithography and planar materials processing. A related but independent set of lithography and nanofabrication technologies continues to evolve in support of a growing effort in quantum structures and nanostructured materials studies. The demands of the nanostructure research and device work has significantly more stringent size scale requirements and a broader materials class than silicon-based electronics. One would like the ability to control material structure down to atomic dimensions, with arbitrary three-dimensional accuracy and high speed. In general, there are limits, some fundamental and some practical, to what can be done in the pursuit of “absolute” control of the structure of materials and devices.

The required feature sizes for nanostructure and quantum-based devices varies depending on the material system and relevant phenomena. In general, the interesting feature sizes are below those accessible to conventional photolithographic processes in use for integrated circuit manufacture today. Quantum structures in semiconductors may require

dimensions on the order of 10 nm, while single-electron devices may function with minimum-feature sizes on the order of 100 nm. Magnetic nanostructures may be of interest and require feature sizes in a similar range of ~ 10 nm. Biological systems provide a new set of opportunities, with interest in accessing dimensions from the size of single molecules to that of cells. In all these areas, the materials of interest could range from elemental metals to organic polymers.

Nanofabrication of materials is a multidimensional issue and involves integration of growth and analytical processes. Thin film deposition and etching methods, in particular epitaxial growth processes for semiconductors and other materials, enable tailoring the composition of materials at the atomic level in the thickness dimension. Microscopy techniques of transmission electron microscopy and scanning tunneling microscopy provide the ability to image the structure of solids with atomic resolution. It is the modification and control of material structure in the directions orthogonal to the growth direction that are, in some sense, lagging in the quest for the ideal nanostructure fabrication technology. In this chapter, we will concentrate on the methods for laterally patterning materials.

The lithographic approach to lateral nanostructure patterning is being done with a range of technologies. These involve the use of photons, charged particles, or atoms formed into a pattern by either a parallel or a serial approach. Figure 3.1 shows a schematic overview of the general lithographic approaches. They largely involve either replication of some image on a previously patterned mask or scanning of a focused beam or tip over a surface. There is always a tradeoff between feature size and throughput. Figure 3.2 shows the observed throughput of various lithographic approaches versus the feature size [1]. Existing techniques fall on a nearly universal curve of resolution versus throughput. In this chapter we will discuss the various approaches and try to develop an understanding of the possibilities and limits of these methods.

Naturally ordered and finely divided systems, such as colloidal particles and metallic island films, have been studied for decades. The term *self-assembly* has increased in usage with reference to nanostructures formed by thermodynamically stable approaches. The self-assembly approach generally enables large areas of nanostructures to be formed rapidly. However, the price for this advantage is a limitation of the geometries that can be created. Periodic systems and random or fractal structures can be readily produced at desirably small dimensions. One needs to consider how to employ these available structures rather than designing a device based on arbitrary shapes. It may well be that a combination of lithographic and self-assembly may be optimal for selected applications.

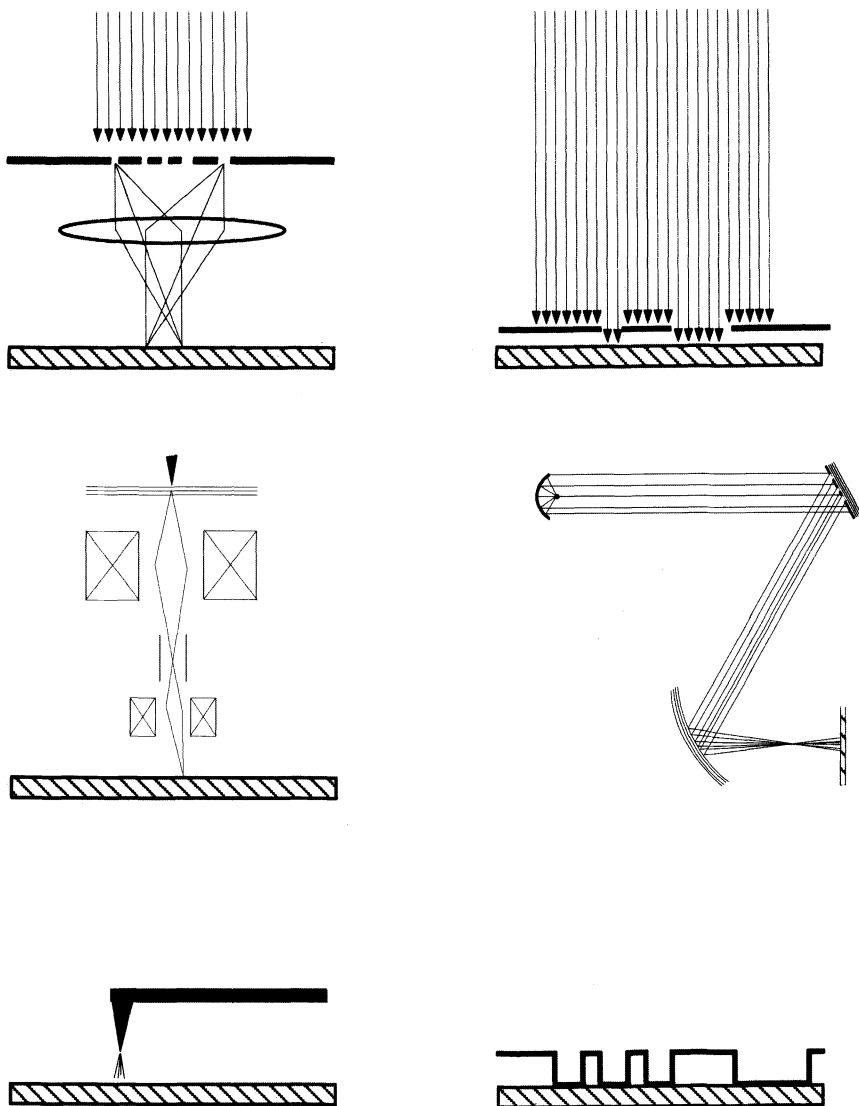


Figure 3.1: Schematic overview of the general lithographic approaches. These include projection lithography, contact lithography, scanning charged particle lithography, photolithography with reflective optics, STM lithography, and microcontact printing.

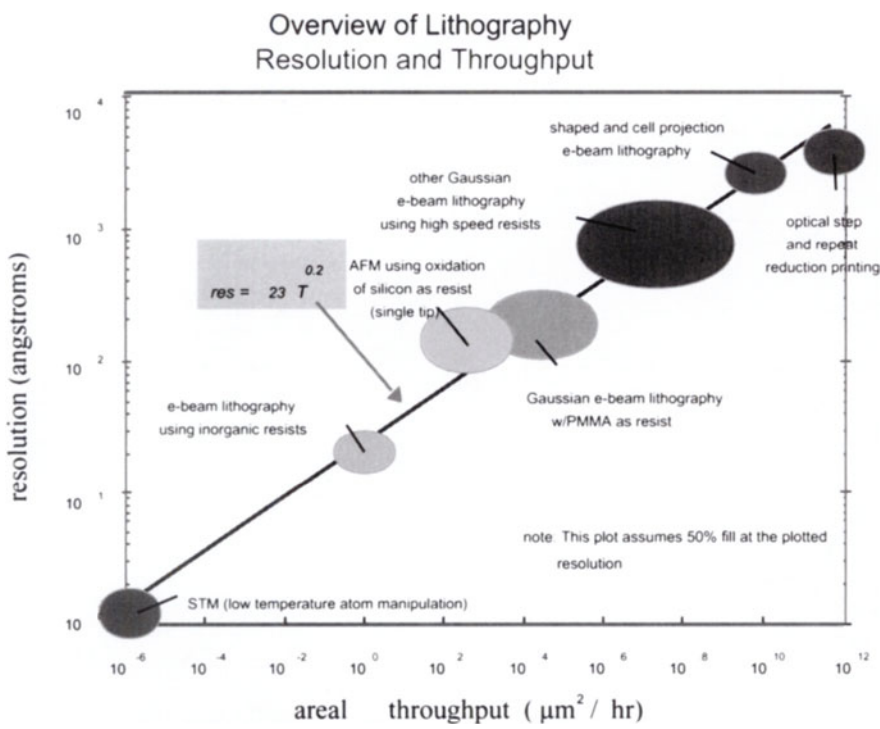


Figure 3.2: “Tennant’s Law”: Throughput plotted vs. minimum feature size for various current forms of lithography. Courtesy of D. Tennant.

Today lithography and other tools developed for integrated circuit fabrication are being used in a wide variety of fields. The field of micro-electro-mechanical systems (MEMS) is one example that uses many conventional fabrication techniques to create free-standing three-dimensional structures for uses as sensors, actuators, and transducers. Less well known are the growing number of applications of microfabrication to biophysics and biology. Microfabrication has been used to create microfluidic systems, surface patterning for cell adhesion and detection, integrated optical and electrical elements for single molecule detection, and three-dimensional structures to study cells. Figure 3.3 shows preferential adhesion of a cell on a surface [2]. The surface is patterned via microcontact printing, discussed in Section 3.5, with a self-assembled monolayer, discussed in Section 3.4.

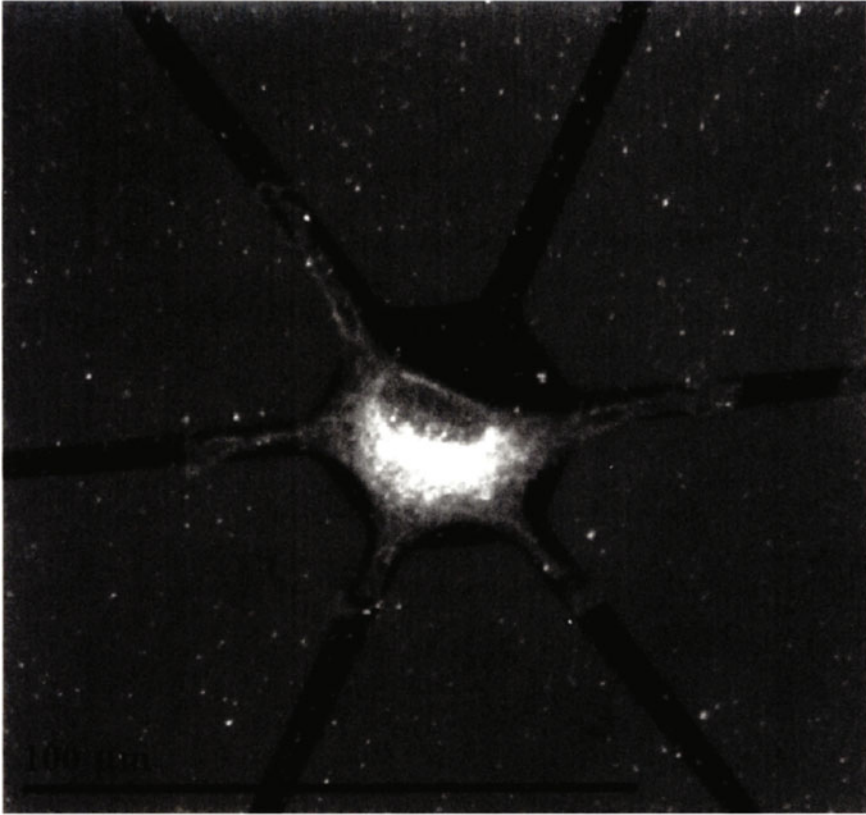


Figure 3.3: Selective attachment of cells from the central nervous system to a silicon substrate patterned with a self-assembled monolayer, DETA. The patterning was performed with microcontact printing and is used to study cell-surface interactions with the objective of producing neural nets to study how the nervous system works. Reprinted in part with permission from [2] (© 1996 Kluwer Academic Publishing).

3.2 The lithographic process in general

3.2.1 Photolithography

The fundamental task in conventional lithography is to transfer a pattern created on a computer onto a wafer. This can be done with photons, electrons, or ions. By far the most common form of industrial lithography is projection photolithography, and it illustrates the basic approach. In

projection photolithography, the design is patterned on a mask. The mask is a large-scale version of the pattern, etched in chromium on a glass substrate. This mask is illuminated, for example, with the i-line of a mercury arc lamp (365 nm), and suitable optics are used to reduce the image of the mask and focus it on the substrate. The light exposes a layer of photosensitive polymer called *resist* that has been deposited on the substrate. Exposure causes chemical changes in the resist that allow selective dissolution by developers. Thus, the pattern on the mask is duplicated in the layer of resist on the wafer. Fabrication typically continues with either a selective deposition or a selective etch process.

There are three important performance gauges for lithography systems: resolution, registration, and throughput. Resolution is the minimum feature size that can be exposed and is the most important gauge for the fabrication of quantum systems. Resolution is properly a function of both the exposure tool and the resist/developer. Registration is the accuracy with which the system can align an exposure with existing features on the substrate. Registration is of some concern for quantum systems, since the feature sizes are small. Finally, throughput is a measure of how fast a given pattern is transferred onto the substrate. This is a concern in industry, where cost is the bottom line, but is less important in a research environment. Of these three, the emphasis in this chapter is on the resolution required for quantum devices. Several good references exist about nanofabrication including registration and throughput. In particular, Campbell [3] and Muray and Brodie [4] cover all facets of conventional semiconductor processing.

The smallest feature that can be transferred with photolithography is limited by diffraction. Rayleigh's criterion is often used as a proxy for the smallest feature size. Rayleigh's criterion states that two objects are on the verge of resolvability if the central diffraction maximum of one object is at the first minimum of the other. In the far field, the smallest feature that can be resolved is proportional to $\lambda/2\text{NA}$, where λ is the wavelength and NA is the numerical aperture of the projection optics. In i-line lithography, the smallest feature size is approximately $0.35\text{ }\mu\text{m}$.

There is a tradeoff in improving the resolution in photolithography. The depth of focus (DOF) is a measure of how far up or down the substrate can be moved and still remain in focus. In photolithography, $\text{DOF} \propto \lambda/\text{NA}^2$. As the resolution is improved, the DOF is reduced. As DOF is reduced, maintaining the substrate in focus across large distances becomes more challenging, and topography on the substrate can become a problem. Limitations on resolution and DOF mean that conventional photolithography is of limited value for quantum systems and that other techniques must be used.

3.2.2 Masks

For projection photolithography systems, masks are typically made of a patterned opaque layer, commonly chromium, on a highly polished transparent substrate, usually fused silica. The pattern on the mask is 4 or 10 times larger than the final image on the wafer. This size is convenient because only $3.5\text{ }\mu\text{m}$ features on the mask are required on a 10 times projection system to generate $0.35\text{ }\mu\text{m}$ features on the substrate. Reduction allows greater tolerances for small mask defects and dust because small defects will not be resolved.

Mask technology is constantly being improved to push the diffraction limit in photolithography. The most important innovation in recent years has been the invention of phase-shifting masks [5]. Phase-shifting masks work by shifting the phase of the light from one feature so that it interferes destructively with light from another feature. The phase shift can be done by selectively adding or subtracting material with the appropriate index and thickness. This process is most easily understood in terms of a grating. In a simple grating in the region between two slits, the light from one slit will interfere constructively with light from the second. If the phase of light from the second slit is shifted by an odd multiple of π , the interference is destructive, and the energy absorbed between the two features is dramatically reduced. While the appropriate pattern for a grating is simple and obvious, application to arbitrary patterns can be very complicated. Figure 3.4 shows a one-dimensional calculation of how phase shifting can affect the resolution of two slits. Without phase shifting, the slits are not resolved. With phase shifting, the intensity goes to zero between the two slits.

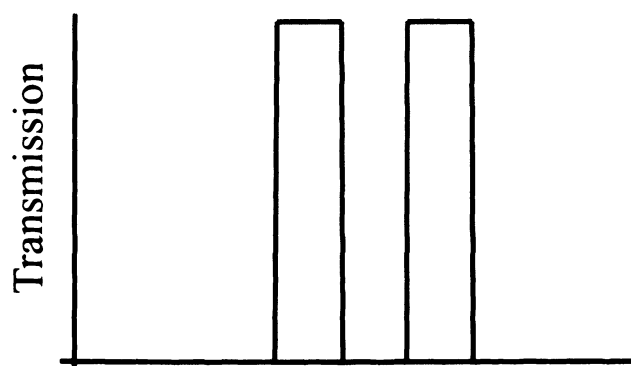
3.2.3 The future of photolithography

The future of lithography for smaller features is still uncertain. The *Semiconductor Industry Association Roadmap* is a prediction of the technology that will be needed in the future based on past historical trends [6]. The roadmap stipulates $0.25\text{ }\mu\text{m}$ resolution in lithography by 1998, $0.18\text{ }\mu\text{m}$ by 2001, $0.13\text{ }\mu\text{m}$ by 2004, $0.10\text{ }\mu\text{m}$ by 2007, and $0.07\text{ }\mu\text{m}$ in 2010. It is against this background that the debate rages about lithography for $0.18\text{ }\mu\text{m}$ and smaller features. Likely candidates include deep UV lithography, extreme UV lithography, X-ray lithography, electron beam lithography, and ion beam lithography.

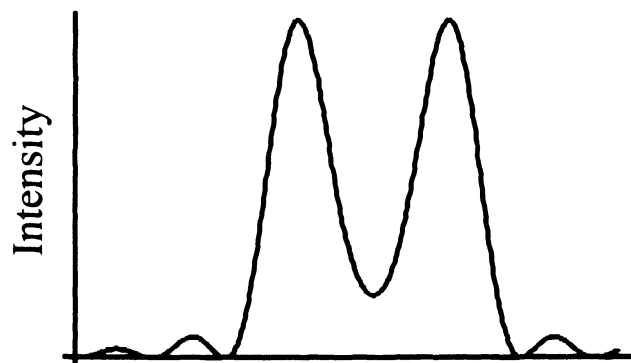
3.2.3.1 DUV

First among these is deep UV lithography (DUV). The pursuit of smaller features for denser circuits has led to the development of reliable KrF

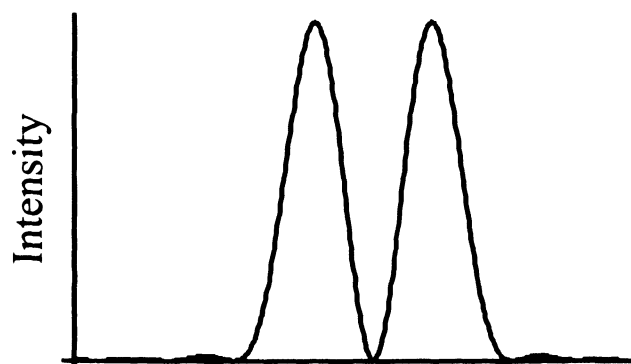
(a)



(b)



(c)



lasers for lithography at 248nm and ArF excimer lasers for lithography at 193nm. KrF sources will achieve the $0.25\mu\text{m}$ features required by industry by 1998, and several commercial vendors are already shipping 248nm systems.

Rayleigh's criterion suggests that KrF sources could be used for $0.18\mu\text{m}$ lithography if suitable high numerical aperture optics can be designed. However, the tradeoffs between resolution and depth of focus imply that the depth of focus would be intolerably short at this resolution. Instead, 193nm lithography seems a more likely candidate for $0.18\mu\text{m}$ features. F_2 excimer lasers (157nm) have also been researched as a successor to 193nm lithography [7]. However, the intensity of F_2 excimer sources is some 30 times lower than ArF sources, and further development for 157nm lithography is unlikely.

There are only a few materials for lens fabrication at 193nm. Most lens materials are either difficult to polish, like calcium fluoride, or begin to absorb near 193nm, like fused silica. The consequences of UV absorption are twofold. First, UV absorption heats the lens material, which leads to distortion. Second, UV absorption has been shown to create color centers and densify or "compact" fused silica lenses. This increases absorption and changes the index of refraction. The importance of these effects is unclear because resists are still being developed. More sensitive resists will lower the required intensity and reduce these effects. Resist schemes being developed include chemically amplified resists, multilayer resists, and top surface imaging. Features of 175nm have been demonstrated in chemically amplified resists and 150nm in top surface imaging resists [8]. These resist schemes will be discussed later.

3.2.3.2 EUV

When is an X-ray not an X-ray? When it is extremely ultraviolet light (EUV). This is the name given to sub-100nm light by the semiconductor industry. The most common sources for sub-100nm light are laser-heated plasmas, although synchrotrons can be used as well. These plasmas are created when the energy from a laser is focused onto a small spot on a thin metal

Figure 3.4: A one-dimensional demonstration of phase shift masks. The mask (a) is composed of two slits to be exposed with a 1:1 projection system. Without phase shifting (b), the image is blurred on the substrate and the slits are not resolved. If the two slits are shifted (c) in phase by π , then the intensity between the slits drops to zero and the image can be reproduced.

film such as gold. The power in the laser vaporizes the film, and the superheated metal vapor radiates light that is collected and collimated.

Unfortunately, there are no materials sufficiently transparent for production of refractive optics at these wavelengths. The problems with absorption in refractive materials can be avoided by using all reflective optics. The most common approach is multilayer mirrors. These mirrors are built up of alternating layers of high and low atomic number materials—for example, Mo and Si. At each interface, only a fraction of the incident beam is reflected, but the layer spacing is chosen so that the reflected beams interfere constructively. This approach has many attractions. First, the mask would be a patterned mirror, which would be much more robust than the membranes used in X-ray lithography. Second, the mask would not have to be in close proximity with the substrate, which greatly reduces the risk of damage. Third, both collimation and reduction are possible, which means a reduction stepper could be designed. In a reduction scheme, mask features would be several times larger than on the substrate, so mask fabrication would be easier. Such high-performance mirrors require precision deposition as well as very smooth substrates. While reflectivities greater than 63% have been achieved, desired polishing tolerances are still beyond the state of the art [9]. While EUV lithography steppers are still under development, initial results with laboratory systems are encouraging. A collaboration between Sandia and Bell Laboratories has built a 10:1 reduction Schwarzschild system that used 13.4nm radiation and five different multilayer mirrors. This system successfully resolved gratings with 100nm features [10].

3.2.3.3 X-ray lithography

X-ray lithography is being pursued as a successor to photolithography below $0.15\mu\text{m}$. X-ray lithography avoids the diffraction limitation of photolithography (a 1keV photon has a wavelength of only 1.2nm) and has higher throughput than scanning electron beam lithography. Although projection schemes exist, most X-ray lithography systems use proximity printing. In proximity printing, the mask is brought within a few microns of the wafer, and the resist is exposed through the mask with no intervening optics. As a result, proximity printing is 1:1 (i.e., the pattern on the wafer is the same size as the mask), and X-ray lithography must rely on another technique, usually electron beam lithography, to make the mask.

The masks for X-ray lithography are more expensive than for photolithography. Because no materials are sufficiently transparent to X-rays, masks are made by patterning an absorbing material, such as a heavy metal, on a very thin membrane. For example, a common combination is 300nm of gold

on a 1–2 μm -thick silicon nitride membrane. These masks are expensive to fabricate and very fragile. X-ray lithography also currently lacks an economical source. Synchrotron rings provide stable, collimated beams, but the beam is usually nonuniform and the rings cost upwards of \$15 million. Laser-heated plasmas such as those used in EUV lithography are more economical, but the beam is still nonuniform and the collimation is not as good. The advantage of X-ray over scanning electron beam lithography is the speed. X-ray lithography exposes the whole pattern at the same time (parallel) rather than one point at a time (serial). The two methods, X-ray and electron beam, are complementary, with scanning electron beam lithography used to make the masks for exposure with X-ray lithography.

The ultimate resolution of X-ray lithography is limited by penumbral blur and scattering. Diffraction is not a limitation for X-ray lithography because the wavelength of the light is close to atomic scale. Penumbral blur is a geometric effect related to imperfectly collimated sources. Even if the mask is perfectly opaque, a finite source will lead to shadowy edges. If the source size is d and the source–mask and mask–wafer distances are D and S , respectively, then from simple geometry the edge of the features on the mask will be blurred to a width $\delta \approx d S/D$. In addition, the X-ray exposure process is mediated by photoelectrons generated when the radiation is absorbed. These hot electrons travel in all directions depositing energy. As a result, the energy from an X-ray is deposited in a distribution around the incident point. In PMMA, 30 nm features have been made with X-ray lithography [11].

3.3 Particle beam lithography

Electron beam lithography (EBL) is the most common system for high-resolution lithography. EBL systems can be separated into two classes of systems: those that write features serially, called scanning systems, and those that write in parallel, such as projection systems.

EBL systems are attractive because the resolution is not limited by diffraction. Instead, the resolution is limited by scattering and proximity effects. When electrons interact with matter, they scatter either elastically or inelastically. Elastic scattering occurs primarily through scattering off nuclei and can result in large-angle scattering events. Inelastic scattering occurs primarily with electrons and is strongly peaked in the forward direction. Secondary electrons created in inelastic events are emitted primarily perpendicular to the beam by conservation of momentum. These electrons can have an important contribution to broadening the width of the exposed resist.

Broadening due to scattering comes from a combination of large-angle scattering and secondary electrons. To limit the effects of scattering, most systems use very high electron energies [12]. The range of a 100 keV electron in silicon is about $150\mu\text{m}$. However, this number is deceiving because it is the energy deposited by the electrons that is important, and this distribution is sharply peaked around the impact point. The energy deposited by electrons scattered to large angles contributes only to a very broad, very flat background. The improved beam-forming characteristics of high-energy systems make them more popular.

It is also possible to limit the effects of scattering by moving to very low energies. At 2 keV, the electron range in silicon is a mere 150 nm compared with $8\mu\text{m}$ at 20 keV. This creates its own challenges, however, because it requires very thin resist schemes such as multilayer schemes or top surface imaging. These schemes are discussed later.

3.3.1 Scanning electron systems

Scanning systems use a small beam spot to expose the pattern a single pixel at a time. Unlike all other lithography systems described so far, direct-write EBL does not use a mask. The pattern is stored in a computer and transferred through computer control of the electron beam. This makes direct-write systems very attractive for research and prototyping applications. Direct-write EBL and variable-beam-shape EBL (discussed below) are the primary tools for mask production for other forms of lithography.

EBL has the highest resolution of any conventional lithography system and has widespread use in research applications. EBL regularly achieves 20 nm resolution, as shown in Fig. 3.5, and some researchers have demonstrated sub-10 nm resolution in polymer resists [13]. EBL is regularly used in research to fabricate devices. Many examples of high-resolution features are given in Section 3.4.

All electron beam systems have three major components, namely, the source, the optics, and the stage. Electron sources available for lithography are similar to those found in scanning electron microscopes. Electron sources are usually thermionic or field emission sources. Thermionic sources emit electrons by heating a filament until the thermal energy of the electrons can overcome the work function of the metal. Field emission sources pull electrons out of a very sharp tip by applying a large electric field to overcome the work function. Field emission sources are much brighter than thermionic sources and have a much smaller spot size. In addition, the energy spread of the emitted electrons from a field emission source is much smaller, which leads to lower chromatic aberrations. On

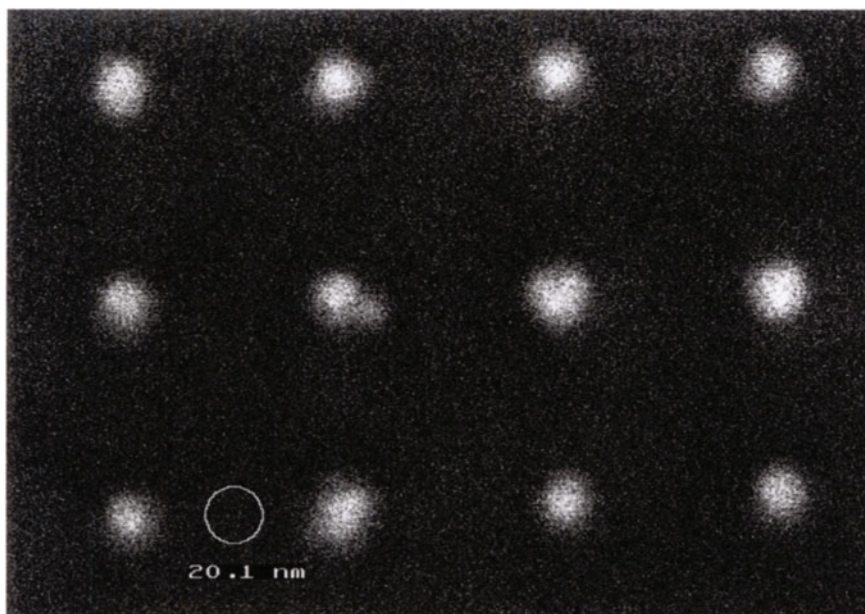


Figure 3.5: Gold dots (20 nm) fabricated with conventional electron beam lithography. The exposure dose was approximately 13 fC/dot at 100 keV energy. The resist system was 100 nm of PMMA developed in 1:3 MIBK/IPA and followed with Ti/Au liftoff.

the other hand, field emission sources are more difficult to fabricate, require better vacuum conditions, and emit nonuniformly. The most popular sources currently (no pun intended) are LaB₆ sources, which have the highest brightness of thermionic sources (10^6 A/cm³str) and require only moderate vacuum (10^{-6} torr).

The electron optics consist of a series of apertures and electric and magnetic fields to shape and focus the beam. Nearest the filament are a series of electrostatic lenses to extract and collimate the beam. Next, magnetic coils are used to align the beam with the rest of the column, as well as to correct stigmatism in the beam shape. The rest of the column usually consists of a series of apertures to limit the spot size, and magnetic lenses to focus and move the beam across the substrate. In addition, many columns have a beam blaster that can turn the beam on and off rapidly. There is a limit to how far the beam can be deflected on the substrate without introducing intolerable errors. As a result, the pattern is broken into fields, with each field small enough to be exposed by the column

without moving the substrate. The substrate must be stepped between fields for the entire pattern to be exposed. Typically, laser interferometers are used to monitor the stage position so that position errors can be compensated for by deflecting the electron beam.

Scanning systems can be subdivided into vector scan and raster scan systems. Raster scan systems raster the beam rapidly up and down while the substrate is slowly moved sideways. In the meantime, deflection plates in the column turn the beam on and off according to the pattern stored in memory. In vector scan systems, the substrate is held fixed while the beam is directed only to the areas to be exposed. The performance of the two systems is similar, although vector scan systems have a slight advantage in terms of speed. Because throughput is one of the limiting factors that prevents widescale application of EBL for device fabrication, there is a lot of effort to improve the speed. One scheme to improve throughput is to build arrays of electron microcolumns so that multiple fields can be exposed simultaneously [14]. Other schemes involve electron projection lithography, analogous to optical projection lithography.

3.3.2 Electron projection lithography

The major drawback of scanning systems is throughput. There are many alternative schemes being researched to create parallel electron beam lithography systems. These include variable-beam-shape EBL, cell-projection EBL, projection EBL, SCALPEL, and photocathode-projection EBL.

3.3.2.1 Electron-projection lithography

Variable-beam-shape EBL, cell-projection EBL, and projection EBL are all similar and use a mask or aperture to shape the beam, which is then demagnified and focused on the substrate. In projection EBL, like its optical counterpart, the pattern is contained in a mask, which is placed in the path of the beam. Electron-projection lithography is attractive because both resolution and throughput are high. The major drawback is the masks, which are complicated. The short penetration depth of electrons precludes the use of fused silica as a mask. Either an absorber on a sub-micron membrane or a stencil mask can be used. Stencil masks consist of transmission windows in 20 μm -thick silicon membranes. These masks are difficult to make, are fragile, and have serious problems with heating and distortion.

In cell-projection lithography, the mask is made up of a small number of general shapes such as squares, rectangles, and triangles. A computer is used to break the pattern up into these general shapes and to expose

each shape in the appropriate place. Throughput is increased in cell projection lithography because although each shape is exposed serially, each shape would take hundreds or thousands of pixels in a direct-write approach. This method is particularly attractive for DRAM production, where the pattern consists of a small number of shapes to be replicated many times. Lithography at $0.18\mu\text{m}$ has already been demonstrated, and $0.15\mu\text{m}$ resolution should be possible [15]. Variable-beam-shape lithography is essentially the same except that the mask is a large square aperture. The aperture is used to partially intercept the beam, resulting in a wide range of rectangles and squares.

3.3.2.2 SCALPEL

SCALPEL stands for Scattering with Angular Limitation in Projection Electron Lithography. This method uses a new type of mask for electron-projection lithography [16]. The SCALPEL technique uses a translucent mask with a thin layer of high-atomic-number absorbing material on a thin, low-atomic-number membrane. As electrons pass through the mask, they scatter, resulting in an angular distribution of electron velocities. The nature of electron interactions with matter implies that the scattering angles from the absorbing material will be much larger. An aperture can be used to distinguish between the large and small scattering angles so that only electrons scattered through small angles contribute to the exposure. SCALPEL masks are much more robust and simpler to construct than conventional electron beam masks. Conventional EBL masks use thick, freestanding absorption layers that are fragile and difficult to make. These absorption layers completely block the beam to prevent exposure. Consequently, heating and distortion are a problem. SCALPEL masks are only used to scatter and not absorb the electrons. Therefore, very little of the incident energy is absorbed by the mask, leading to lower distortions. Features as small as 80nm have been resolved in a prototype 4:1 reduction system at Bell Laboratories [17]. This system can expose a 1cm^2 field in parallel and has a depth of focus of $75\mu\text{m}$ for $0.25\mu\text{m}$ features. Masks for this system consist of 50nm Cr/W on a $0.15\mu\text{m}$ silicon nitride membrane.

3.3.2.3 Charged-particle proximity printing

There has been a recent revival in electron proximity printing [18]. This technique, sometimes called photocathode electron-projection lithography, uses a mask that is also an electron source. The mask, which is a patterned photocathode, is brought into close proximity with the substrate and illuminated with a UV source. UV photons eject electrons from the

surface of the cathode via the photoelectric effect. The ejected electrons are emitted with a random velocity but can be focused onto the substrate with a combination of electric and magnetic fields. This technique has several interesting aspects. First, it is a parallel electron beam technique and therefore has the speed of parallelism with the immunity to diffraction of electron lithography. Second, it is a 1:1 mask-based technique. This implies that submicron resolution in the pattern requires submicron resolution in the mask. Third, because photoemission is a surface phenomenon, the mask is very sensitive to surface conditions and contamination. With this technique, Saville and coworkers have demonstrated 110–550 nm features with a large depth of focus using a gold photocathode [18].

3.3.3 Ion beam systems

Ion beam lithography is another candidate for high-resolution lithography. Ion beam lithography is similar to electron beam lithography except that the beam is made of ions. This has certain advantages over electron beam lithography. The penetration depth into a substrate is much shorter for an ion than for an electron. This reduces the proximity effect and implies that a larger fraction of energy in the beam will expose the resist rather than be wasted in the substrate. However, ion beam lithography has its own set of challenges.

There are primarily three approaches to ion beam lithography being researched. These are focused ion beams, which resemble scanning electron beam lithography; proximity masked ion beam lithography; and ion-projection lithography.

Focused ion beam (FIB) lithography is a serial technique like its electron counterpart. As such, FIB lithography is slow for large-scale lithography. FIB has the advantage over electron lithography in resist sensitivity, and therefore throughput. FIB lithography does not have any advantage in resolution because the probe size for focusing ions is currently limited to ≈ 10 nm spot size. Minimum feature sizes around 20 nm have been achieved with a Ga liquid metal source used to sputter gold surfaces [19].

Proximity printing requires a mask that is transparent to ions in the regions to be exposed and opaque elsewhere. The short penetration depth that makes ion beam lithography attractive implies that masks must be made of thin, freestanding stencils. Stencil masks for ion lithography are fabricated by etching transmission windows in $1\text{ }\mu\text{m}$ freestanding membranes [20]. Because proximity printing is a 1:1 system, submicron resolution requires submicron windows in the mask, and because the mask is exposed to the same power density as the substrate, mask heating, dis-

tortion, and etching are problems. On the other hand, extremely well-collimated beams can be achieved, so penumbral blur is much less of a problem than in X-ray lithography. Mask-wafer separations larger than 1 mm are possible with a blur of only 30 nm. This is a large advantage over X-ray lithography, which requires separations on the order of $10\text{ }\mu\text{m}$. In addition, ion beam sources are compact and cheap in comparison with synchrotrons. The disadvantage of ion beam lithography lies in the stencil masks, which are more costly and more fragile than X-ray masks.

Ion beam projection lithography (IPL) will ameliorate some of the drawbacks of stencil masks. In IPL, the ion beam illuminates the mask at low energy (a few keV), reducing the damage to the mask from ion etching and milling. After the mask, the beam is accelerated and focused on the substrate with 5:1 or 10:1 demagnification. With IPL, nanometer resolution with a high depth of focus ($>10\text{ }\mu\text{m}$) can be achieved. A projection system with five-times reduction at the wafer has been built by a collaboration at IMS in Germany. This system has demonstrated a resolution of $0.2\text{ }\mu\text{m}$ over topographies as high as $0.5\text{ }\mu\text{m}$ [21]. The five-times reduction implies that critical features in the mask are much easier to fabricate, and the energy absorbed by the mask is reduced by a factor of 25. However, the optics for ion-projection lithography are not as good as those for projection electron beam lithography.

3.3.3.1 Neutral atom lithography

Neutral atom beams have recently been developed as new systems for lithography. Neutral atom beams, like ion beams, are not limited by diffraction. Unlike ion beams, neutral atom beams do not suffer space charge effects, which limit the current in charged-particle lithography systems. Sub-100 nm features have been fabricated by patterning substrates with metastable atoms. Metastable atoms are excited to an upper electronic state while still electrically neutral. Metastable atomic beams such as Ar or He have internal energies of 12 and 21 eV, respectively, which is sufficient to expose electron beam resists. Metastable beams have been used to expose alkanethiol SAMs on gold. Patterning can be achieved with either a physical mask or optical focusing. For example, 80 nm features have been demonstrated with a stencil mask in close proximity with the substrate [22]. Neutral atom beams can also be focused and collimated with laser beams. The same techniques that are used to optically cool and trap atoms can be used to collimate the beam and focus it on the substrate. Celotta and coworkers have used a standing wave pattern to directly deposit chromium in lines 38 nm wide (FWHM) on a 213 nm pitch as well as chromium dots 80 nm in diameter [23].

3.4 Resists—polymer, TSI, CAR, SAM, and inorganic resists

Equally important with the exposure tool is the resist/developer system. Two important performance gauges for a resist are sensitivity and resolution. Sensitivity is a measure of how much energy is required to expose the resist. This is the critical dose for exposure, D , measured in mJ/cm^2 [24]. Contrast is used as a proxy for resolution. Contrast is a measure of how quickly a resist switches from unexposed to completely exposed. An ideal resist will be unexposed below a critical threshold of dose, D , and completely exposed above threshold. In practice, the crossover will have a finite width. The resist will be essentially unexposed below some lower bound, D_0 , and totally exposed above some upper bound, D_1 . In between, the resist will partially dissolve in developers. A resist with poor contrast will not create sharp features.

Resist development processes have received relatively little attention in published research, although they have important effects on resolution and sensitivity. Bernstein and coworkers have reported a detailed study of developers for PMMA, the archetypal electron beam resist [25]. They demonstrated that choice of developer and developer temperature can have an exponential effect on contrast.

Some resists such as PMMA have a single ingredient, while others are made of multiple constituents. A typical photoresist contains two parts, namely, a resin and a photosensitive compound. The exposed areas of the resist are rendered either more soluble (positive tone) or less soluble (negative tone) in developer. In positive-tone photoresists, the photosensitive compound is insoluble in water until exposed. Exposure causes a chemical rearrangement into an acid that is soluble in weak bases. Some resists, such as PMMA, can be either positive or negative in tone. Low levels of electron exposure break the long polymer chains into smaller molecules that are more soluble (positive tone). Eventually, high doses crosslink the polymer fragments, rendering the exposed area less soluble (negative tone).

Lithography on substrates with preexisting topography poses its own challenges. Most resists are spin coated. In this process, the resist is deposited in solution in the middle of a substrate, which is then spun at 2–6 krpm. Spinning the substrate at high speed spreads the resist to a uniform thickness and drives off the solvent. Topography causes variations in the resist thickness that can lead to underexposure or overexposure. Immediately next to a step in the substrate, the resist will be piled up higher, and on top of the step the resist will be thinner. One solution is to

use a resist thickness twice the maximum topography on the substrate. There are three important considerations when considering the thickness of the resist. The first, as was just discussed, is topography. Second, the depth of focus of the exposure tool must be longer than the resist is deep. And third, there is a practical limit to the aspect ratio that can be achieved for very small features.

Topography is not much of a problem with X-ray or ion beam lithography. The resolution in these systems is limited by penumbral blur, which can be made very small. On the other hand, electron beam lithography is limited by scattering, which leads to larger feature sizes in thicker resists. Topography is also a problem for DUV and EUV lithography. Most organic materials strongly absorb DUV and EUV radiation so that only the top layer of a thick resist is exposed.

There are resist schemes that take advantage of this short penetration depth. Multilayer schemes typically use a thin resist spun on a thin etch mask on top of a thick planarizing layer. The upper layer is patterned with lithography and then used as a mask to etch the lower layers. These schemes can be complicated, and many different examples can be found in the literature [26]. A second resist scheme is called top surface imaging (TSI).

3.4.1 Top surface imaging

In TSI, a thick resist layer is used to cover the topography of the wafer. In the subsequent exposure, only the top fraction of the resist is exposed via electron beam or photolithography. Next, silicon is incorporated into the exposed or unexposed region in a process called *silylation*. The incorporated silicon reacts with the resist to form a hard crust. The resist is “developed” with an O_2 etch that completes the pattern transfer by preferentially removing the nonsilyated resist. TSI has been used to demonstrate ~50 nm resolution in an experimental chemically amplified resist. Figure 3.6 shows an SEM image of 67 nm resolution in an experimental resist [27]. TSI has also been used with 193 nm photolithography to demonstrate 150 nm features. Current challenges to developing TSI processes are linked to the silicon diffusion step, which is extremely non-Fickian. Consequently, the TSI process needs to be well characterized in order to control critical dimensions and edge roughness.

3.4.2 Chemically amplified resists

Photolithography at 193 nm requires significantly more complicated resist schemes because conventional photo resists are not very sensitive to DUV

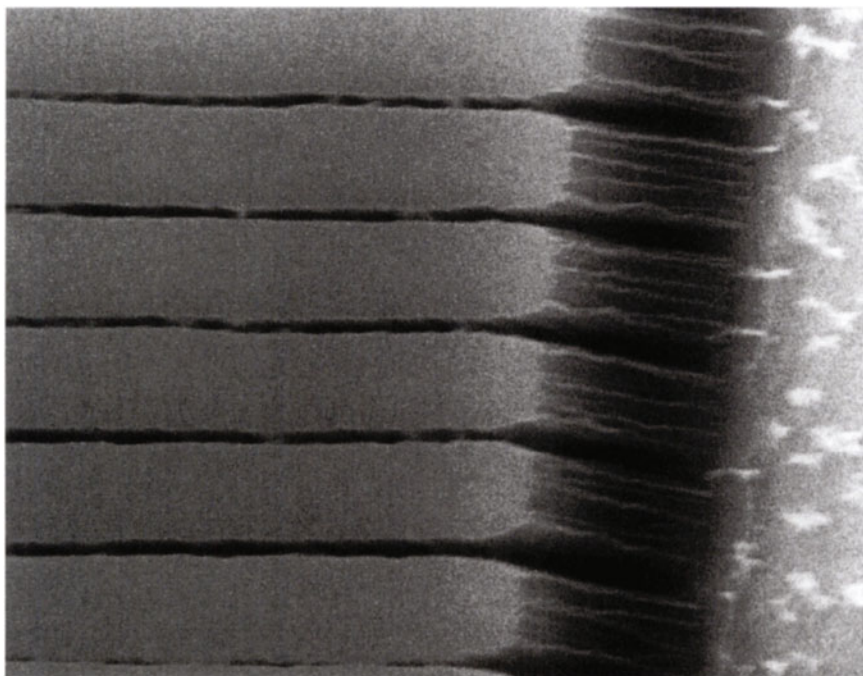


Figure 3.6: A $1.2\mu\text{m}$ TSI resist with 67 nm lines. The resist was exposed at 2 keV, silyated, and etched with an O_2 plasma. Reprinted in part with permission from [27] (© 1997 American Institute of Physics).

radiation. Special chemically amplified resists (CAR) have been developed to increase sensitivity. These resists contain a compound called a photoacid generator that decomposes when exposed to light. During post-exposure bake, the decomposition product initiates a cascade of chemical transformations, which render the resist more soluble in developer. Although these resists are very sensitive to radiation, they can be poisoned by atmospheric contamination and are very sensitive to the temperature and duration of the postexposure bake.

3.4.3 Self-assembled monolayers

Self-assembled monolayers (SAMs) are monolayers formed by the attachment of single organic chains in an ordered arrangement on a substrate surface. They take advantage of the ability of nature to form highly ordered

monolayers if the interaction energy is high enough. The monomer of a SAM consists of three parts: the foot, which binds to the surface; the body; and a head, which forms the new surface. For example, octadecyltrichlorosiloxane (OTS) has a trichlorosilane foot, SiCl_3 , which binds to silicon oxide; an alkane body $(\text{CH}_2)_{17}$; and an alkane head, CH_3 , which is hydrophobic. One of the most powerful features of SAMs is the ability to mix and match foot, body, and head components to find the right functionality. Silanes bind to silicon and silicon oxide, whereas thiol groups bind to GaAs and Au. Halides, oxygen, and conjugated bonds have been added to the body to improve the threshold for electron beam exposure. Alkyl-terminated SAMs are extremely hydrophobic and therefore form good etch masks for aqueous-based wet etches, whereas amine groups have been used for ligation sites for certain molecules as well as for the selective adhesion of biological materials.

SAMs have demonstrated high resolution, are sufficiently robust as an etch mask, and are compatible with electron beam lithography, STM lithography and microcontact printing. The most extensively studied SAMs are based on alkanethiols, such as octadecanethiol (ODT), which form SAMs on Au. Because of the high binding energy between the monomer and the substrate, exposure of a gold substrate to ODT in hexadecane results in the rapid self-assembly of the thiol as a chemisorbed monolayer. Thiol SAMs on gold are 1–2 nm thick, have a lattice structure with a periodicity ~ 0.2 nm, have a low defect density, and are relatively inert.

OTS monolayers on SiO_2 and ODT monolayers on GaAs have been exposed with electron beam lithography [28]. The monolayers act as a self-developing positive electron beam resist, with sensitivity comparable with PMMA. The exposure mechanism is suspected to be C–C and C–Si scission by secondary electrons. In both types of monolayers, 25 nm lines have been written with a 50 nm commercial electron beam system. Buffered hydrofluoric acid (6:1) was used to transfer exposed patterns 53 nm into SiO_2 . To test the ultimate resolution of SAMs as an electron resist, Lercel and coworkers exposed an alkanesilane SAM with a high-resolution SEM [29]. The FWHM of the beam was approximately 2 nm at 20 kV. Figure 3.7 shows feature sizes of 6–7 nm with this technique.

3.4.4 Inorganic resists

Inorganic resists are not widely used, although they have demonstrated the highest-resolution electron beam lithography. Figure 3.8 shows an array of holes directly etched in an inorganic resist, AlF_3 . The hole diameters are

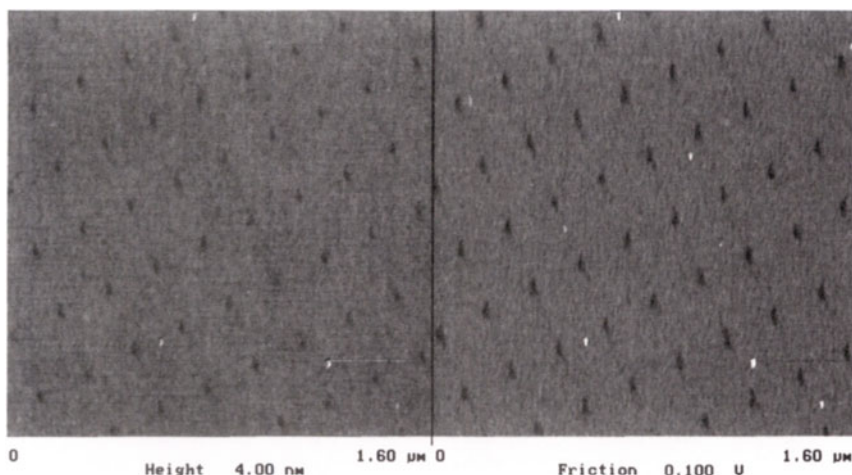


Figure 3.7: AFM image of 6–7 nm dots exposed with a Zeiss 982 SEM at 20 keV. The pattern is of the SAM itself and has not been transferred into the substrate. The image on the left is in noncontact-mode AFM and the one on the right is friction mode. Reprinted in part with permission from [29] (© 1996 American Institute of Physics).

1.5–2 nm, with a center-to-center spacing of 4 nm [30]. Electron beam exposure causes the AlF_3 to dissociate, resulting in F_2 , which evaporates, and Al, which diffuses through the resist. These resist systems have such high resolution because only the primary electron beam has sufficient energy to initiate the dissociation, and therefore the resolution is limited only by forward scattering of the beam. However, this means that the critical dose is some five orders of magnitude higher and throughput is correspondingly lower.

3.5 Soft lithography

Soft lithography is a term coined by Whitesides [31] to encompass a new group of lithography methods. The term refers to elastomeric stamps and molds, which are central to these techniques, as well as to the mechanical properties of organic materials in general. Techniques included under this umbrella are microcontact printing of SAMs, transfer molding, conformal phase-shift lithography, photopolymerization, and imprint lithography, among others.

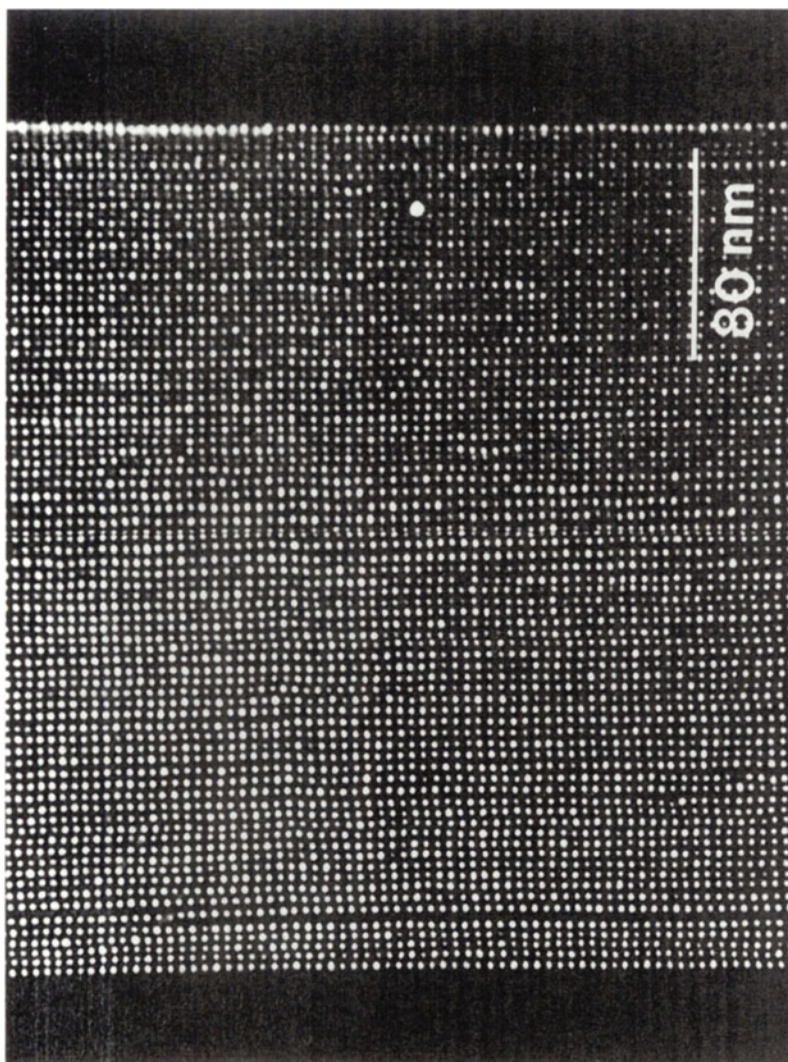


Figure 3.8: A scanning electron microscope image of an array of 1.5–2 nm features exposed in 80 nm of an inorganic resist, AlF_3 , with a 100 keV STEM. The walls of the holes are plated with aluminum. Reprinted in part with permission from [30] (© 1996 American Institute of Physics).

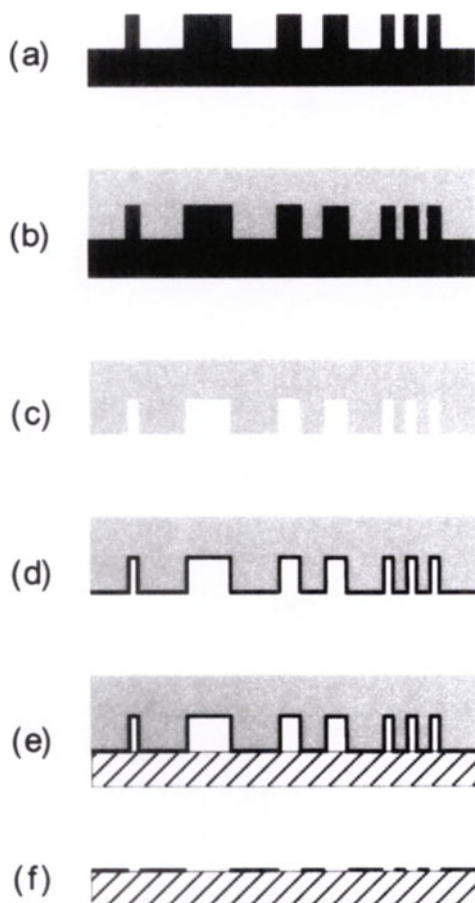


Figure 3.9: Microcontact printing starts with (a) a master that is (b) covered in prepolymer. After curing (c), the stamp is released and (d) coated with ink. The substrate is (e) stamped, and when the stamp is removed (f), only the pattern remains.

3.5.1 Microcontact printing (μ CP)

μ CP starts with a rigid master with the inverse of the desired pattern engraved on its surface. The master is used to cast an elastomeric replica, which in turn is used as a stamp to reproduce the pattern. Figure 3.9 shows a schematic of microcontact printing. Silicon, silicon oxide, and even patterned PMMA on Si have been used as masters. To create the stamp, the mask is covered with an antistick coating. An elastomeric monomer,

usually dimethyl siloxane, is poured into the mold and cured at an elevated temperature. After polymerization, the elastomer stamp can be peeled away from the master and coated with “ink,” such as an alkanethiol or alkanesilane for printing SAMs or a protein for biological experiments. The stamp is used to replicate the pattern by simply bringing it in contact with a substrate. Only ink from the embossed regions of the stamp is transferred to the substrate.

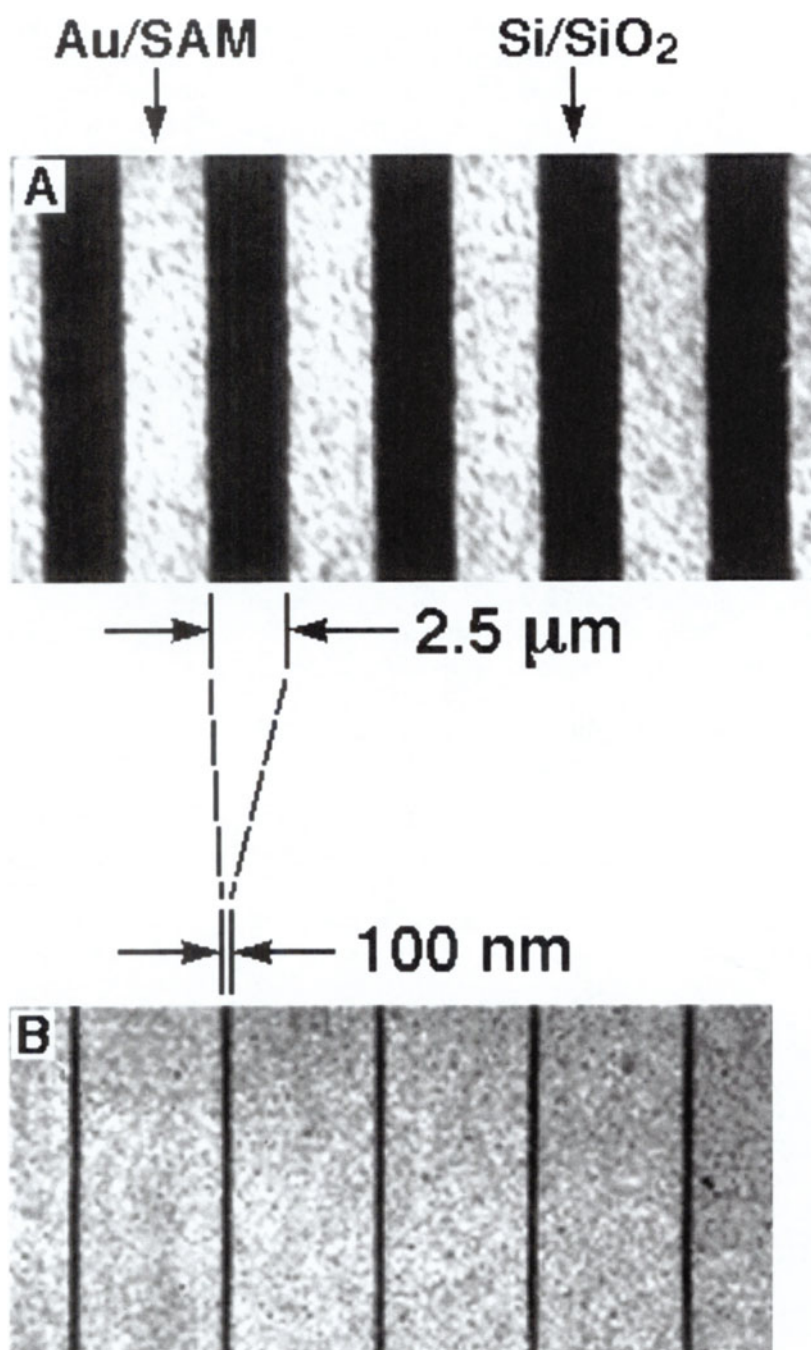
With μ CP it is possible to print a pattern with submicron features over square centimeters. μ CP of a thiol SAM on gold has reproduced 35 nm trenches and 100 nm gratings [32]. These patterns were robust enough to be transferred into the substrate with a wet CN/O_2 etch. Alkanesilane monolayers have been stamped onto Si and metal oxide surfaces. Patterns have been transferred into metal oxides with a Cl-based RIE or into Si with HF and KOH [33]. Figure 3.10 shows an image of 100 nm lines on gold fabricated with μ CP [34].

μ CP has several interesting features. First, μ CP is parallel lithography over large fields. Second, the resolution is limited by the mechanical strength of the surface relief. This implies that trenches can be formed with greater resolution than lines. Lastly, because the elastomeric stamp forms a conformal contact, long wavelength topography is not a problem. Stamps have even been used to pattern spherical and cylindrical substrates (e.g., optical fibers).

In a variation on a theme, Rogers and Whitesides [35] have reported the use of an elastomeric mask as a phase mask to produce sub-100 nm lines in photoresist with an incoherent light source. This technique, termed *conformal phase-shift lithography*, is based on the intensity modulation in the near field of light passing through a transparent substrate with surface relief. If the depth of the surface relief shifts the phase of the light by an odd multiple of π , then waves passing through the embossed regions will interfere destructively with waves passing through the engraved regions. The intensity at the edge of the embossed pattern is reduced to zero within a region of width on the order of one quarter the wavelength of the light. To implement this technique, an elastomer mask was brought into conformal contact with a substrate coated in photoresist. Exposure of the resist through the mask produces lines in positive-tone resist and trenches in negative-tone resist. In this way, 90 nm lines and 100 nm dots have been made.

3.5.2 Photopolymerization

Another method is photopolymerization. In this technique, the master is brought into contact with a bare transparent substrate, and a thin polymer



layer is drawn between the master and the substrate via capillary action. The polymer is subsequently cured with UV light and the master removed. Terris and coworkers at IBM have used a rigid master to successfully replicate 50 nm features and have promoted this method as a way of replicating ultradense CD-ROMs [36]. Current CD-ROMs have a bit density of approximately 0.1 Gbit/cm². In comparison, 50 nm features imply a bit density of 10 Gbit/cm². Whitesides and coworkers have used this technique with an elastomeric master to fabricate field effect transistors on GaAs [37].

3.5.3 Imprint lithography

The last method discussed here is imprint lithography. Like all soft lithography techniques, it begins with a rigid master formed with some other process. A thin layer of resist such as PMMA is deposited on a substrate, and the master is brought into contact and pressed down with 600–1,900 psi. While the master is in contact, the resist is heated above its glass transition temperature, allowing the polymer to reflow and conform to the master. To prevent air bubbles, the best imprint lithography is done in vacuum, and mold release agents should be added to the resist to prevent adhesion of the resist to the master. Chou and coworkers have demonstrated 25 nm features created uniformly over 1 cm² [38]. They have also demonstrated 10 nm Au/Pd liftoff using PMMA patterned with imprint lithography. Figure 3.11 shows a master and its replica with 100 nm features.

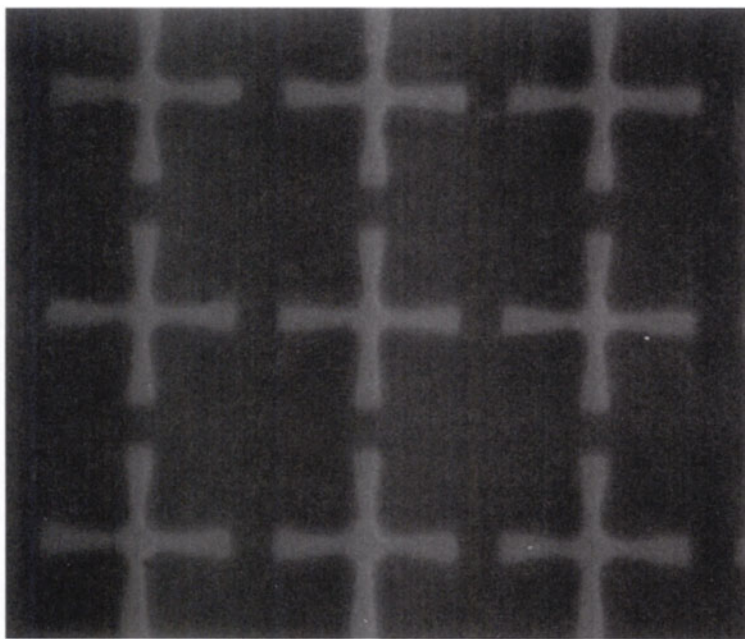
3.6 Scanned probe methods of lithography

3.6.1 STM

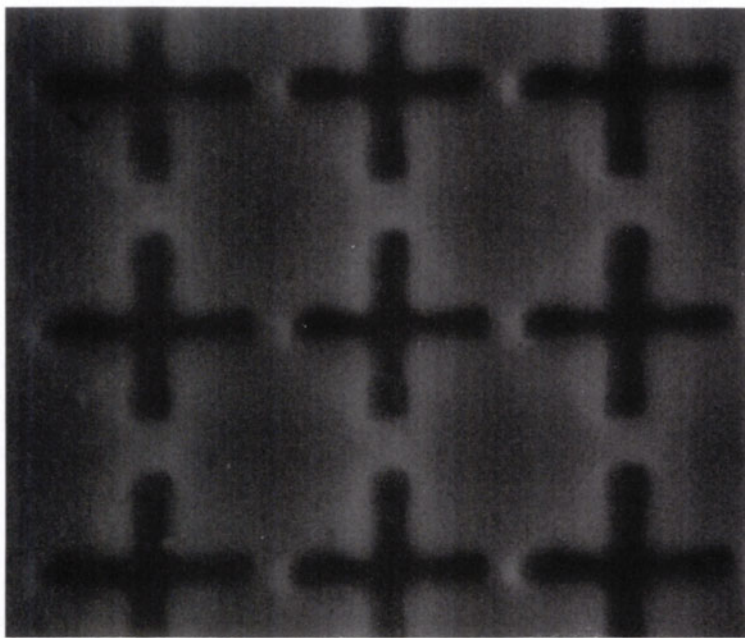
Scanning microscopies such as scanning tunneling microscopy (STM), atomic-force microscopy (AFM), or near-field scanning optical microscopy (NSOM) were originally designed to image surfaces with high resolution. Recently, these techniques have been developed to include surface modification with the same resolution. In the case of STM, this has reached the



Figure 3.10: Scanning electron micrograph of 100 nm lines formed with μ CP. The ink was an alkanethiol SAM stamped on a gold substrate. The initial pattern was 2.5 μ m wide but was allowed to spread until 100 nm features were formed. Reprinted in part with permission from [34] (© 1995 American Chemical Society).



(a)



(b)

ultimate objective of atom-by-atom control. The major limitation of scanned probe methods of lithography is throughput. Maximum writing speeds are often on the order of 1 nm/s or approximately 0.1 MHz repetition rates (assuming a 5 nm pixel), which should be compared with 25 MHz repetition rates in state-of-the-art electron beam lithography systems.

3.6.1.1 Selective oxidation

The most common STM lithography technique employs an electrically conducting tip to selectively oxidize regions of a sample. Silicon makes a good substrate for this because it can be passivated easily in HF, it oxidizes easily, and selective etches for Si versus SiO_x are well known. Cleaning a Si substrate with hydrofluoric acid (HF) removes the native oxide formed in atmosphere and terminates the surface with a monolayer of hydrogen. An STM tip or a conducting AFM tip operated in air anodically oxidizes the surface. This process resembles a conventional electrochemical oxidation, with the STM tip as the cathode and ambient humidity as the electrolyte. In this manner, the sample surface can be locally oxidized with a 10–100 nm resolution. Dagata and coworkers demonstrated 100 nm lines [39]. Features 60 nm wide and 300 nm deep have been created using a 1 nm-thick oxide as a mask for a wet chemical etch, and features 10 nm wide and 30 nm deep have been etched with an ECR-reactive ion etch [40].

Systems other than silicon have also been patterned. Titanium, chromium, and aluminum can all be locally oxidized. Aluminum can be oxidized to form Al_2O_3 , which is an effective insulator. Al wires 5–8 nm thick and 1 μm wide have been selectively oxidized to define narrow constrictions 40 nm wide by 500 nm long. This technique is especially promising for the fabrication of Coulomb blockade devices. These channels were further oxidized to form point contacts that demonstrated quantized conductance at room temperature [41]. Matsumoto and coworkers have used selective oxidation in titanium to define isolated metal islands that function as Coulomb blockade devices at room temperature [42]. Figure 3.12 shows



Figure 3.11: Master and replica with 100 nm features formed with imprint lithography. The original master (a) was fashioned using a bulk Si wafer, electron beam lithography for patterning, and a reactive ion etch to achieve a relief pattern ~400 nm deep. The replica (b) shows the inverse of the pattern formed in PMMA after compression and reflow at 170°C. Courtesy of R. Tiberio.

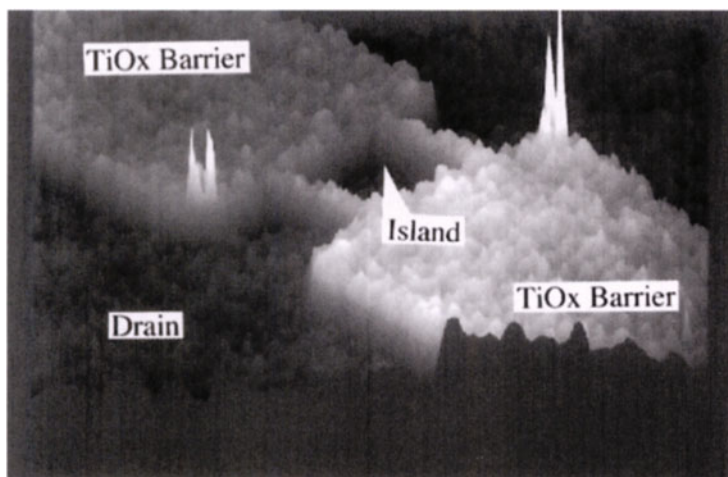


Figure 3.12: Atomic-force microscope image of a single-electron transistor fabricated with selective-oxidation STM lithography. The narrow TiO_x lines defining the island are 15–25 nm wide by 30–50 nm long. Reprinted in part with permission from [42] (© 1996 American Institute of Physics).

an AFM image of a single-electron transistor (SET) made by the nano-oxidation process. Two narrow lines of TiO_x were written to define an island between the source and the drain. The TiO_x lines are typically 15–25 nm wide, and the island is roughly 30–50 nm by 35–50 nm. The thickness of the Ti is less than 3 nm. Figure 3.13 shows drain current–voltage characteristics and conductance of the SET at room temperature.

Even though GaAs does not form as useful an oxide as silicon, selective oxidation has been done on a thin amorphous Si capping layer on GaAs and transferred into the substrate. In this scheme, silicon is effectively acting as a resist to pattern the GaAs, and 100 nm MOSFET gates have been constructed this way [43]. Spin-on-glass is another material that has been successfully patterned with an STM [44]. The patterning mechanism is thought to be decomposition of methyl groups bonded to the Si–O backbone. This results in different etch rates in BHF between patterned and unpatterned regions. Linewidths as narrow as 40 nm have been demonstrated at writing speeds of 1 mm/s. Although the minimum feature size in air is limited to about 10 nm, atomic-scale resolution has been demonstrated operating the STM in ultrahigh vacuum (UHV). Figure 3.14 shows linewidths as narrow as 2 nm on a Si (100) surface [45].

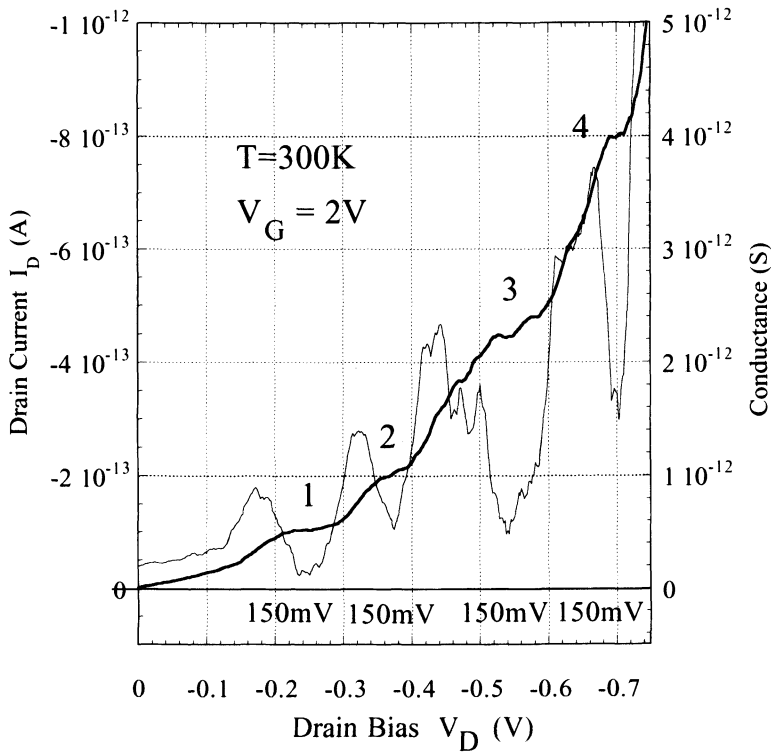


Figure 3.13: Drain current–voltage and conductance–voltage characteristics of the SET in Fig. 3.12 at room temperature. Four Coulomb staircases of $\sim 15\text{mV}$ period are observed. Reprinted in part with permission from [42] (© 1996 American Institute of Physics).

3.6.1.2 STM patterning of SAMs and organic resists

STM can be used to expose conventional electron beam resists. McCord and coworkers have patterned 20 nm lines in PMMA with an STM and have used the resulting pattern for Au/Pd liftoff [46]. Dobisz and coworkers have demonstrated 27 nm lines in SAL-601, a common negative-tone electron beam resist, with an STM [47]. Wilder and coworkers have used a modified AFM with a conducting tip to pattern a 60 nm layer of SAL-601 [48]. Their modified AFM was operated in contact mode and used the cantilever deflection to observe the topography while a negative bias between the tip and substrate caused electrons emitted from the tip to expose the resist on the substrate. In this way, 100 nm features were easily achieved and

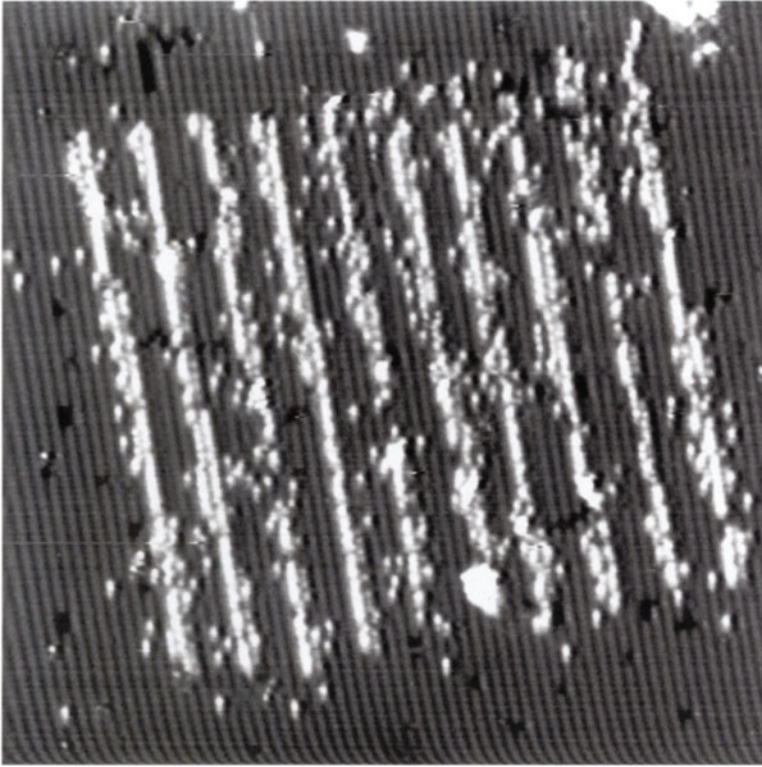


Figure 3.14: Near-atomic resolution for lithography achieved with UHV STM. The Si substrate is initially passivated with a monolayer of hydrogen. The STM patterned area appears brighter because the unpassivated Si has a higher density of states available for tunneling. Reprinted in part with permission from [45] (© 1997 IEEE).

features as small as 41 nm have been written. Because the tip follows the topography on the substrate, topography is not an issue from a lithographic standpoint. Writing speeds up to 3 mm/s have been demonstrated.

Self-assembled monolayers (SAMs) are also compatible with STM lithography. SAMs, described above, are monolayers formed by the attachment of single organic chains in an ordered arrangement on the substrate surface. Alkanethiols, such as hexadecanethiol (HDT), have a thiol foot, which binds to gold surfaces, and a hydrophobic head, which protects against aqueous-based wet etching. HDT has been used as an etch mask against a CN^-/O_2 aqueous wet chemical etch [49]. Gold exposed to 0.1 M HCN in 1 M

KOH saturated with oxygen etches at a rate of several angstroms per second, whereas regions protected with a monolayer of HDT etch 10^6 times slower. Exposure with an STM disrupts the SAM, allowing selective etch of gold underneath. Lines 10–20 nm wide have been etched into 10 nm Au with this method, with a writing speed of $200 \mu\text{m/s}$.

Marrian and coworkers have demonstrated selective electroless plating with STM lithography [50]. When a SAM contains a head that is also a ligand for a metal catalyst, a metal film can be built on the SAM. STM exposure can disrupt the ligand group so that plating does not occur. This results in a selective deposition of a metal film in the unexposed area. PEDA forms a SAM with an ethyl diamine head. Ethylene diamine can bind with a Pt catalyst that in turn acts as a nucleation site for Ni electroless plating. With this method, linewidths down to 20 nm have been demonstrated.

3.6.1.3 Single-atom manipulation

No discussion of STM nanofabrication would be complete without mention of single-atom manipulation. D. Eigler and E. Schweizer demonstrated in 1990 the use of a low-temperature STM to position Xe atoms [51]. Since that time, lines and corrals [52] of atoms have been widely publicized. The force applied by an STM tip was used to physically assemble atoms on the surface into different configurations. The potential for atomic control for fabrication of devices is clear. However throughput is severely limited with this technique.

3.6.2 AFM contact methods

AFM tips have been used to scratch metal surfaces to create weak links in superconducting films. The probe tip is simply moved across the thin metal film under high loading forces. Microbridges only 10–20 nm thick have been fabricated in this way [53]. This technique is of limited value, since even silicon nitride tips are quickly eroded with this method.

Mamin and Rugar at IBM have developed thermomechanical writing with an AFM tip. In this method, the tip of the AFM is operated in contact mode with a polymer-coated wafer. To pattern the wafer at a particular point, the tip is heated either with an external laser pulse [54] or with a pulse of current through an integrated resistor [55]. The heated tip causes the polymer coating on the substrate to soften, and the force from the tip on the softened polymer creates a depression that becomes permanent when the polymer cools. Figure 3.15 shows pit sizes fashioned this way on the order of 120 nm and 15 nm deep. This method has been promoted as a

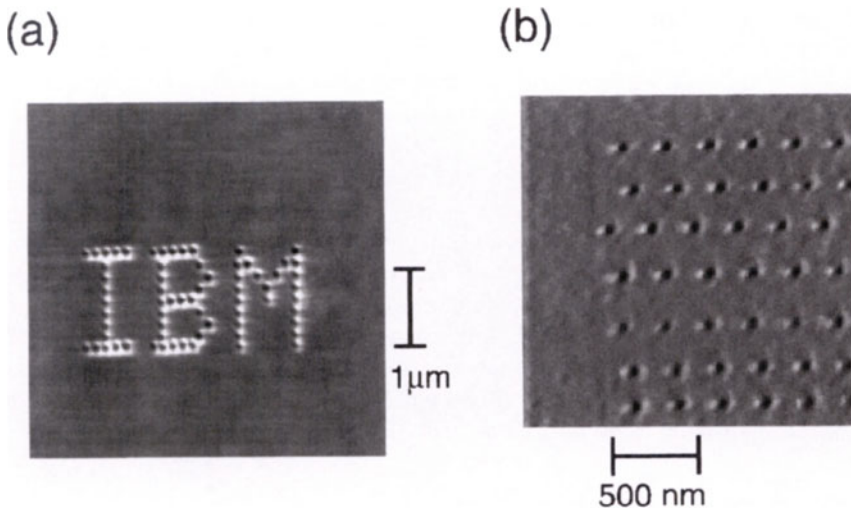


Figure 3.15: Pits formed in a polymer layer with an AFM tip. To pattern the polymer, the tip is momentarily heated, which softens the polymer and allows it to deform. The pits are on the order of 120 nm across and 15 nm deep. Reprinted in part with permission from [54] (© 1996 American Institute of Physics).

fabrication technique for ultradense CD-ROMs, although patterns created in this way could conceivably be used as a mask for further RIE or wet chemical etching.

3.6.3 Near-field optics

The final scanned probe technique that will be mentioned here employs near-field scanning optical microscopy (NSOM) to pattern media. NSOM beats the diffraction limit by scanning within a few nanometers of the surface. To date, NSOM has not been used as a lithography tool.

NSOM has been used to read and write data in magneto-optical (MO) materials. In MO applications, light from the probe heats the substrate above the Curie temperature, resulting in the formation of a domain with the magnetization aligned in the opposite direction. Betzig and a group at AT&T employed a tapered fiber optic fiber coated with 100 nm Al as an NSOM probe. The tapered fiber was drawn in a commercial puller to have a final aperture 20 nm diameter. The probe was brought within 10 nm of the surface using a conventional scanning probe feedback loop and raster

scanned across the surface. With this probe, they were able to image features with a resolution of 12 nm [56] and write 60 nm features with a bit density of 45 Gbit/in² [57]. Conventional MO technology has a density of ~1 Gbit/in².

While this technique is promising, the optical throughput in pulled fiber probes is limited. Heat-induced damage of the tip limits the maximum intensity that can be coupled into the probe to approximately 15 mW. The output efficiency is tremendously poor, on the order of 10^{-6} . A conventional optical resist has a critical dose on the order of 50 mJ/cm². With this efficiency, NSOM could only write 30 $\mu\text{m}^2/\text{s}$. In other words, assuming a 50 nm pixel size, the repetition rate would be limited to about 12 kHz, which is not practical. However, recent innovations suggest that the prospects might be improving for NSOM. Recently, two groups have reported a technique to improve the output efficiency by 10^3 [58]. The tips are formed with a wet chemical etch and have a diameter of approximately 70 nm. Further improvements in size and output efficiency might make NSOM a practical low-cost photolithography research tool.

3.7 2D crystals

Throughput remains a challenge for high-resolution electron beam lithography. In addition, feature sizes less than 20 nm or periods less than 50 nm are not easy to make. While these feature sizes are possible with scanning probe techniques, the speed is even worse. For those cases where the sub-20 nm features are periodic and require only coarse alignment, naturally occurring lattices, which exhibit self-assembly and parallel fabrication, are attractive.

Douglas and Clark at the University of Colorado [59] have used crystalline monolayers to provide relief for an angled evaporation. Their protein forms a 10 nm-thick 2D triangular lattice with a period of 22 nm. The lattice is porous with a filling factor of 30%. Evaporation of Ti/W at an angle of 40 degrees highlights the topography of the lattice. Fast-atom beam milling resulted in a 1 nm-thick Ti/W film with 15 nm diameter holes on a 22-nm period. This screen can be transferred into the substrate or charged to promote selective chemisorption in the holes in the lattice.

Park and coworkers at Princeton have demonstrated successful transfer of patterns formed in self-assembled diblock copolymers onto a substrate (Fig. 3.16) [60]. Diblock copolymers consist of two polymer chains connected with a covalent bond. With the correct choice of chain length,

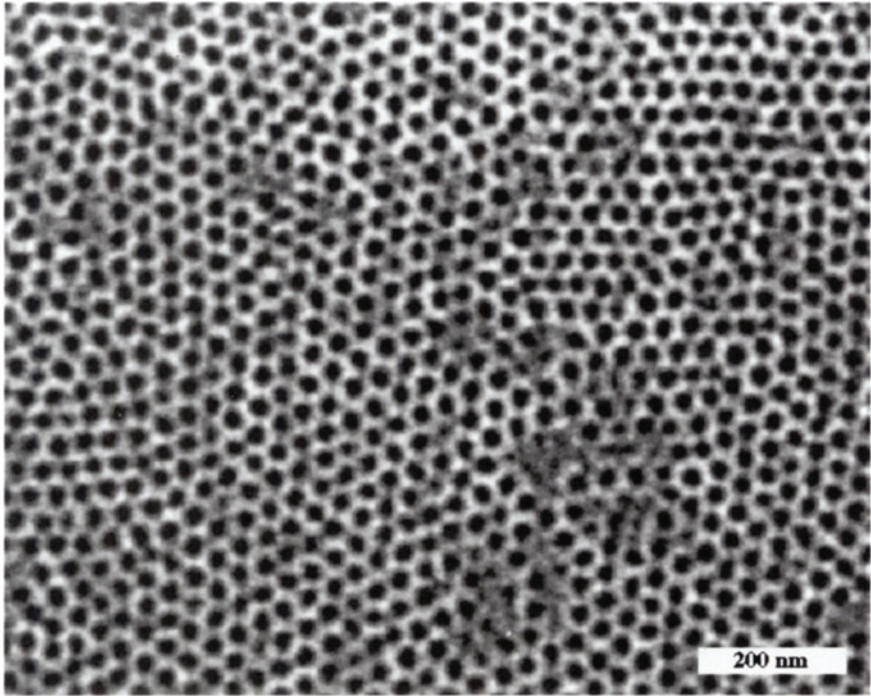


Figure 3.16: SEM image of holes etched into germanium. The original pattern was self-assembled in block copolymers and transferred into a nitride film. The pitch is approximately 40 nm. Reprinted in part with permission from [42] (© 1998 American Institute of Physics).

the polymers will spontaneously self-assemble to form an ordered lattice of micelles over a large area. Because the blocks have different chemical composition, it is possible to selectively process one block relative to the other. Hexagonal lattices with a 30 nm pitch have been transferred 15 nm deep into SiN_4 with this method. Both positive- and negative-tone processes are possible over areas as large as 45 cm^2 .

There are several advantages to this approach to nanostructure fabrication. First, the processing is parallel; the lattice is assembled in one step. Second, natural lithography has a much higher resolution than can be conveniently achieved with electron beam lithography. On the other hand, because the lattice grows spontaneously everywhere, the lattice is composed of many different grains with many defects. Lastly, arbitrary structures are not possible; instead, devices must be designed around existing structures in nature.

3.8 The future of lithography

It is important to distinguish between the future of industrial lithography and lithography in a research environment. The future of industrial lithography is still uncertain but will have few surprises. The SIA Technology Working Group has put a lot of effort into listing possible technologies for sub-100nm lithography and breaking each technology down into smaller research areas to be developed. Competing technologies include EUV photolithography, X-ray lithography, electron beam lithography, SCALPEL, and ion beam lithography. Several research projects are working on the problems facing these technologies so that the semiconductor industry can maintain its historical trends in feature size and other critical parameters.

The immediate successor to conventional (248nm) lithography will be 193nm lithography along with chemically amplified or top surface imaging resists, phase-shift masks, and high NA optics, which will produce 100 nm features. After 193 nm lithography, EUV and X-ray lithography will probably be employed for 70 nm features, but eventually, for the highest-resolution features, industry will have to turn to charged particle lithography of some kind. This may be scanning EBL, projection EBL, SCAPEL, or IPL.

Moore's law predicts that the minimum feature size in industry decreases exponentially with time. It is this historical trend that has led the SIA to predict 130nm lithography will be in industrial use by 2003. However, exponential growth cannot continue indefinitely in any field, and eventually there will be a crossover from exponential growth to zero growth. This crossover in industrial lithography will probably occur at or before 20nm. This kind of resolution will probably require charged-particle lithography.

The future of lithography in research is very exciting. Most prototyping and research-scale fabrication, where throughput is not an issue, will continue with scanning electron beam lithography. With care and a little money, almost any high-resolution SEM can be converted for electron beam lithography with 25nm resolution. However, in the future we expect to see devices and circuits made with unconventional technologies. In the future we can expect to see the first devices and circuits constructed on an atom-by-atom basis. In addition, STM oxidation will provide inexpensive fabrication that will make research into quantum structures more accessible. Further in the future, there is the possibility that NSOM will develop into an inexpensive tool for in-the-lab high-resolution photolithography. NSOM has the potential to provide photolithography with a resolution better than 50nm outside of a clean room for the cost of a laser system and some electronics. This will allow research into quantum structures to take place in every university laboratory.

In addition, microcontact printing and other soft contact techniques will allow the rapid reproduction of large fields for only the cost of a master image. Microcontact printing will allow lithography by nonexperts away from a clean room environment. There is a caveat that any process design seeking to use microcontact printing or any other contact technique must be defect tolerant. Any contact between the substrate and mask always creates defects and reduces the life of the mask. On the positive side, microcontact printing allows the possibility of lithography on substrates with large long-wavelength topography, even including fibers or spheres.

3.9 Conclusions

We are closer than ever to having the ability to control material structure down to atomic dimensions. Quantum systems, magnetic nanostructures, and biological systems all require dimensions on the order of 10 nm. This is generally not possible with conventional lithographic systems, although the gap is closing rapidly. The SIA roadmap predicts that 100 nm critical dimensions will be available for industrial-scale production within 10 years. Although many conventional techniques are rapidly approaching sub-100 nm resolution, each technology has its own set of challenges that prevents wide-scale application. For example, EUV is currently limited by polishing tolerances for reflective optics; X-ray lithography is limited by mask fabrication; and electron beam reliably demonstrates sub-100 nm resolution, but the throughput is too slow.

While the limits on conventional systems are constantly being improved, a number of other techniques have already demonstrated atomic resolution. STM/AFM techniques are being developed for cheap laboratory-scale lithography with atomic resolution. Soft lithography offers parallel reproduction over square-centimeter fields with high resolution. In addition, natural lithography offers parallel reproduction of a restricted set of atomic-scale arrays. A variety of approaches exist for fabrication of quantum structures. We expect that no single technique will dominate atomic-scale lithography in the next decade and that instead different techniques will find application as appropriate for biological, magnetic, and quantum structures.

REFERENCES

- [1] D. Tennant, *Nano-Science and Technology*, edited by G. Timp (Springer Verlag, New York, 1998).

- [2] C. D. James, R. Davis, M. Meyer, A. Turner, S. Turner, G. Withers, L. Kam, G. Banker, H. G. Craighead, and M. Isaacson, *IEEE Trans. Biomedical Engineering* **47**, 17–21 (2000).
- [3] S. Campbell, *The Science and Engineering of Microelectronic Fabrication* (Oxford University Press, New York, 1996).
- [4] I. Brodie and J. Muray, *The Physics of Micro/Nano Fabrication* (Plenum Press, New York, 1992).
- [5] M. D. Levenson, N. S. Wisnawathan, and R. A. Simpson, *IEEE Trans. Elec. Dev.* **26**, 1828 (1982).
- [6] *The National Technology Roadmap for Semiconductors* (NTRS) is the Semiconductor Industry Association (SIA) 1994 report published by SEMATECH, Inc. It can be viewed at http://www.sematech.org/public/roadmap/doc/litho_toc.html. The next edition of the SIA's National Technology Roadmap is now in renewal and will be published in 1998.
- [7] T. M. Bloomstein, M. W. Horn, M. Rothschild, R. R. Kunz, S. T. Palmacci and R. B. Goodman, *J. Vac. Sci. Technol. B* **15**, 2112 (1997); H. G. Craighead, J. C. White, R. E. Howard, L. D. Jackel, R. E. Behringer, J. E. Sweeney, and R. W. Epworth, *J. Vac. Sci. Technol. B* **12**, 1186 (1983).
- [8] M. Rothschild, J. A. Burn, S. G. Cann, A. R. Forte, C. L. Keast, R. R. Kunz, S. C. Palmateer, J. H. C. Sedlacek, R. Uttaro, A. Grenville, and D. Corliss, *J. Vac. Sci. Technol. B* **14**, 4157 (1996).
- [9] R. H. Stulen, *IEEE J. Sel. Top. Quan. Elec.* **1**, 970 (1995).
- [10] G. D. Kubiak et al., *J. Vac. Sci. Technol. B* **12**, 3820 (1994).
- [11] F. Cerrina, *Proc IEEE* **84**, 644 (1997).
- [12] H. G. Craighead, R. E. Howard, L. D. Jackel, P. M. Mankiewich, *Appl. Phys. Lett.* **42**, 38 (1983).
- [13] W. Chen and H. Ahmed, *J. Vac. Sci. Technol. B* **11**, 2519 (1993).
- [14] T. H. P. Chang, et al., *J. Vac. Sci. Technol. B*, **14**, 3774 (1996).
- [15] H. Yamashita, K. Tokunaga, Y. Kojima, H. Nozue, and E. Nomura, *J. Vac. Sci. Technol. B* **13**, 2473 (1995).
- [16] H. W. P. Koops and J. Grob, in *Springer Series in Optical Sciences: X-ray Microscopy*, edited by G. Schmahl and D. Rudolph, **43**, 119 (1984); S. D. Berger and J. M. Gibson, *Appl. Phys. Lett.* **57**, 153 (1990).
- [17] L. Harriott et al., *J. Vac. Sci. Technol. B* **14**, 3825 (1996).
- [18] G. F. Saville, P. M. Platzman, G. Brandes, R. Ruel, and R. L. Willett, *J. Vac. Sci. Technol. B* **13**, 2184 (1995).
- [19] A. Kieslich, J. P. Reithmaier, and A. Forchel, *J. Vac. Sci. Technol. B* **12**, 3518 (1994).
- [20] J. C. Wolfe, S. V. Pendharkar, P. Ruchhoeft, S. Sen, M. D. Morgan, W. E. Home, R. C. Tiberio, and J. N. Randall, *J. Vac. Sci. Technol. B* **14**, 3896 (1996).

- [21] E. Hammel, A. Chalupka, J. Fegerl, R. Fischer, G. Lammer, H. Loeschner, L. Malek, R. Nowak, G. Stengl, H. Vonach, P. Wolf, W. H. Bruenger, L.-M. Buchmann, M. Torkler, E. Cekan, W. Fallman, F. Paschke, G. Stangl, F. Thalinger, I. L. Berry, L. R. Aarriott, W. Finkelstein and R. W. Hill, *J. Vac. Sci. Technol. B* **12**, 3533 (1994); W. H. Brunger, L.-M. Buchmann, and M. A. Torkler, *J. Vac. Sci. Technol. B* **12**, 3547 (1994).
- [22] K. S. Johnson, K. K. Berggren, A. Black, C. T. Black, A. P. Chu, N. H. Dekker, D. C. Ralph, J. H. Thywissen, R. Younkin, M. Tinkham, M. Prentiss, and G. M. Whitesides, *Appl. Phys. Lett.*, **69**, 2773 (1997); R. Younkin, K. K. Berggren, K. S. Johnson, M. Prentiss, D. C. Ralph and G. M. Whitesides, *Appl. Phys. Lett.* **71**, 1261 (1997).
- [23] R. J. Celotta, R. Gupta, R. E. Scholten, and J. J. McClelland, *J. Appl. Phys.* **79**, 6079 (1996).
- [24] Dose for electron beam resists is measured in $\mu\text{C}/\text{cm}^2$ for a given beam energy.
- [25] G. H. Bernstein, D. A. Hill, and W.-P. Liu, *J. Appl. Phys.* **71**, 4066 (1992).
- [26] D. E. Seeger, D. C. La Tulipe, R. R. Kunz, C. M. Garza, and M. A. Hanratty, *IBM J. Res. Dev.* **41**, 105 (1997).
- [27] C. S. Whelan, D. M. Tanenbaum, D. C. La Tulipe, M. Isaacson, and H. G. Craighead, *J. Vac. Sci. Technol. B* **15**, 2555 (1997).
- [28] M. J. Lercel, R. C. Tiberio, P. F. Chapman, H. G. Craighead, C. W. Sheen, A. N. Parikh, and D. L. Allara, *J. Vac. Sci. Technol. B* **11**, 2823 (1993).
- [29] M. J. Lercel, H. G. Craighead, A. N. Parikh, K. Seshadri, and D. L. Allara, *Appl. Phys. Lett.* **68**, 1504 (1996).
- [30] A. Muray, M. Isaacson, and I. Adesida, *Appl. Phys. Lett.* **45**, 589 (1984).
- [31] G. M. Whitesides, *1997 Conference on Solid State Sensors and Actuators*, IEEE, 1997, p 23.
- [32] H. A. Biebuyck, N. B. Larsen, E. Delsmarche, and B. Michel, *IBM J. Res. Develop.* **41**, 159 (1997).
- [33] P. M. St. John and H. G. Craighead, **68**, 1022 (1996).
- [34] D. Wang, S. Thomas, K. Wang, Y. Xia, and G. M. Whitesides, *Appl. Phys. Lett.* **70**, 1593 (1997); Y. Xia et al., *JACS* **117**, 3274 (1995).
- [35] J. A. Rogers, K. E. Paul, R. J. Jackman, and G. M. Whitesides, *Appl. Phys. Lett.* **70**, 2658 (1997).
- [36] B. D. Terris, H. J. Mamin, M. E. Best, J. A. Logan, D. Rugar, and S. A. Rishton, *Appl. Phys. Lett.* **69**, 4262 (1996).
- [37] J. Hu, R. G. Beck, T. Deng, R. M. Westervelt, K. D. Maranowski, A. C. Gossard, and G. M. Whitesides, *Appl. Phys. Lett.* **71**, 2020 (1997).
- [38] S. Chou, P. R. Krauss, and P. J. Renstrom, *J. Vac. Sci. Technol. B* **14**, 4129 (1996).

- [39] J. A. Dagata, J. Schneir, H. H. Harary, C. J. Evans, M. T. Postek, and J. Bennet, *Appl. Phys. Lett.* **56**, 2001 (1990).
- [40] E. S. Snow, P. M. Campbell, P. J. PcMarr, *Appl. Phys. Lett.* **63**, 749 (1993); E. S. Snow, W. H. Juan, S. W. Pang, and P. M. Campbell, *Appl. Phys. Lett.* **66**, 1729 (1995); E. S. Snow, P. M. Campbell, and F. K. Perkins, *Proc. IEEE* **85**, 601 (1997).
- [41] E. S. Snow, D. Park, and P. M. Campbell, *Appl. Phys. Lett.* **69**, 269 (1996).
- [42] K. Matsumoto, M. Ishii, K. Segawa, Y. Oka, B. J. Vartanian, and J. S. Harris, *Appl. Phys. Lett.* **68**, 34 (1996).
- [43] S. C. Minne, H. T. Soh, P. Flueckiger, and C. F. Quate, *Appl. Phys. Lett.* **66**, 703 (1995); E. S. Snow and P. M. Campbell, *Appl. Phys. Lett.* **64**, 1932 (1994).
- [44] S. W. Park, H. T. Soh, C. F. Quate, and S.-I. Park, *Appl. Phys. Lett.* **67**, 2415 (1995).
- [45] J. W. Lyding, T.-C. Shen, J. S. Hubacek, J. R. Tucker, and G. C. Abeln, *Appl. Phys. Lett.* **64**, 2010 (1994).
- [46] M. A. McCord and R. F. W. Pease, *J. Vac. Sci. Technol. B* **6**, 293 (1988).
- [47] E. A. Dobisz and C. R. K. Marrian, *Appl. Phys. Lett.* **58**, 2526 (1991).
- [48] K. Wilder, H. T. Soh, A. Atalar, and C. F. Quate, *J. Vac. Sci. Technol. B* **15**, 1811 (1997).
- [49] H. A. Biebuyck, N. B. Larsen, E. Delsmarche, and B. Michel, *IBM J. Res. Dev.* **41**, 159 (1997).
- [50] C. R. K. Marrian, F. K. Perkins, S. L. Brandow, T. S. Koloski, E. A. Dobisz, and J. M. Calvert, *Appl. Phys. Lett.* **64**, 390 (1994).
- [51] D. M. Eigler and E. K. Schweizer, *Nature* **344**, 524 (1990).
- [52] M. Crommic, C. Lutz, and D. M. Eigler, *Science* **262**, 218 (1993).
- [53] T. A. Jung, A. Moser, H. J. Hug, D. Brodbeck, R. Hofer, H. R. Hidber, and U. D. Schwartz, *Ultramicroscopy* **42–44**, 1446 (1992).
- [54] H. J. Mamin and D. Rugar, *Appl. Phys. Lett.* **61**, 1003 (1992); H. J. Mamin, *Appl. Phys. Lett.* **69**, 433 (1996).
- [55] B. W. Chui, T. D. Stowe, T. W. Kenny, H. J. Mamin, B. D. Terris, and D. Rugar, *Appl. Phys. Lett.* **69**, 2767 (1996).
- [56] E. Betzig, J. K. Trautman, T. D. Harris, J. S. Weiner, and R. L. Kostelak, *Science* **251**, 1468 (1991).
- [57] E. Betzig, J. K. Trautman, R. Wolfe, E. M. Gyorgy, P. L. Finn, M. H. Fryder, and C.-H. Chang, *Appl. Phys. Lett.* **61**, 142 (1992).
- [58] D. Zeisel, S. Nettesheim, B. Dutoit, and R. Zenobi, *Appl. Phys. Lett.* **68**, 2491 (1996); P. Hoffmann, B. Dutoit, and R. Salathe, *Ultramicroscopy* **61**, 165 (1995); S. J. Bukofsky and R. D. Grober, *Appl. Phys. Lett.* **71**, 2749 (1997).

- [59] K. Douglas, N. A. Clark, and K. J. Rothschild, Appl. Phys. Lett. **48**, 676 (1986); K. Douglas, N. A. Clark, and K. J. Rothschild, Appl. Phys. Lett. **56**, 692 (1990); K. Douglas, G. Devaud, and N. A. Clark, Science **257**, 642 (1992).
- [60] M. Park, C. Harrison, P. M. Chaikin, R. A. Register, and D. H. Adamson, Science **276**, 1401 (1997); C. Harrison, M. Park, P. M. Chaikin, R. A. Register, and D. H. Adamson, to be published in J. Vac. Sci. Technol. B.

Chapter 4

The use of MOVPE to produce quantum structured semiconductors

Werner Seifert

Solid State Physics, Lund University, Box 118, S-221 00 Lund, Sweden

4.1 Introduction

Metal–organic vapor-phase epitaxy (MOVPE) is today a powerful growth technique, capable of producing high-quality material of virtually all III/V and II/VI semiconductor compounds and ternary and quaternary alloys (for an overview, see, for instance, [1] and the references therein). The importance of MOVPE for mass production has been increasing. In terms of consumption of III/V wafer area, MOVPE is today rated second (20%), after liquid-phase epitaxy (LPE, 55%) and in front of the classical vapor-phase epitaxy (VPE, 18%) and the ultra-high-vacuum (UHV) technology molecular beam epitaxy (MBE, 7%) [2]. With the development of larger reactor cells, simultaneously handling up to about 10–15 2" wafers per growth session, this trend will continue. As a consequence, for the practical realization of quantized structures in future mass production technology, MOVPE will be one of the most important growth technologies. In Section

2 of this chapter, we will briefly discuss some principles and specifics of MOVPE, in particular from the perspective of growth on a scale of nanometers.

Direct growth of semiconductor heterostructures in a nanoscale, i.e., with dimensions of typically a few tens of nanometers, is still an outstanding challenge. For growth in the vertical direction, quantum wells (QWs) with thicknesses and interface abruptnesses in the range of monolayers ($1 \text{ ML} \approx 0.3 \text{ nm}$) can be realized by using MBE [3] or MOVPE [4]. Lateral structuring in order to fabricate quantum wires (QWRs) and quantum dots (QDs) is still a more difficult task. Pure lithography in combination with etching techniques in order to produce structures smaller than 100 nm is an expensive process; moreover, it reaches the limits of feasibility. As a rule, one has to use a combination of lithography, etching, and selective regrowth techniques, whereby the quantized structures will be created in smaller areas in comparison to the lithographic pattern length. This outcome can be achieved through nonplanar growth, with differences in the orientation dependence of growth rates for the different materials. Some aspects and results of this technique will be discussed briefly.

A cheaper alternative, discussed in more detail in Sections 3 and 4, is to use self-assembling processes. Self-assembling and, in a wider sense, self-organization in the lateral as well as the vertical growth direction is a natural reaction of heterocombinations in order to accommodate strain if there is a lattice misfit between the different deposited materials. Since strain is the driving force for all these self-organization processes, it is natural that the kind of combinations of materials is more important than the technology used for the deposition of the materials. Thus, QDs of (for instance) InAs, embedded into GaAs, can be produced by self-assembling with very similar results by using vacuum technologies like MBE as well as by using vapor-deposition techniques like MOVPE. However—and this will be an important point to be discussed here—there exist often different optimum conditions of the growth parameters for the different deposition techniques and specific materials, significantly affecting the results of the growth experiments. Therefore, even though we focus here more on MOVPE, a comparison of the results of the different techniques and materials will be made throughout this chapter.

When we discuss self-assembling, the reader should be aware of the temporary character of some of the explanations. In Section 3, we try to summarize the principle and to give some ideas about the underlying mechanisms. However, what counts are the facts, and it was our ambition to present this framework of understanding, as much as possible, as a coherent view of the experimental observations presented in Section 4. We

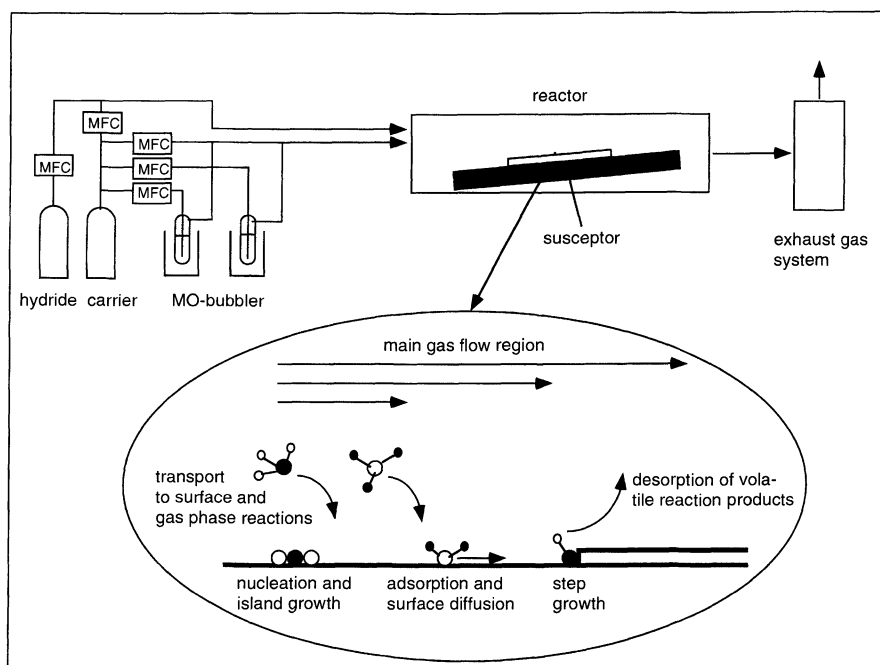


Figure 4.1: Schematics of the basic elements of a MOVPE system.

acknowledge here that the actual development is still going on that will allow a better and deeper understanding of the process of self-assembling.

4.2 Principles and specifics of MOVPE

In comparison to MBE, MOVPE has often been seen as a “black box” in which a complicated chemistry is going on, the details of which remain unclear. In fact, MOVPE technology, up to now, has not routinely had at its disposal such excellent tools for in situ monitoring of the surface during growth as RHEED offers for high-vacuum technologies. However, through the use of X-ray scattering [5] and optical reflection techniques [6], this gap has been partially filled during recent years. We will here briefly discuss some critical issues concerning the suitability of MOVPE for growth of structures in a nanosize. A schematic is depicted in Fig. 4.1.

MOVPE commonly uses hydrides like AsH_3 or PH_3 or their substitutes such as TBA or TBP (tertiary butyl arsine and -phosphine, respectively) as

the Group V precursors, and uses metal-organic compounds like TMAI, TMGa, and TMIIn (trimethylaluminum, trimethylgallium, trimethylindium) as the Group III precursors. These compounds, diluted by a carrier gas, are then transported towards a heated substrate where they react in a one-way process, simplified for GaAs as



The metal-organic compounds are placed in “bubblers” within temperature-regulated thermostates from which they are picked up by electronically controlled carrier gas flows. The amount of evaporated metal-organics is defined by the bubbler temperature (which determines the vapor pressure), the total pressure within the bubbler, and the flow of carrier gas through the bubbler. To have control over the evaporated metal-organic material is very important, since the Group III elements determine the deposition rate in MOVPE. The Group V precursor compounds are introduced in excess into the reactor cell. This is one difference between MOVPE and MBE: MOVPE is commonly working under Group V-rich conditions. This has an effect especially on the type of surface reconstruction. In the case of GaAs(001), the typical surface reconstruction is in the range of common deposition temperatures and AsH₃ pressures $c(4 \times 4)$, i.e., an additional layer of As-dimers adsorbed onto an arsenic-terminated surface [7]. In case of InP, it is the P-rich (2×1) reconstruction [8].

Modern MOVPE machines can operate under low-pressure conditions (typically in the range of 20–200 mbar), which increases the mean free path of the molecules within the reactor and improves the homogeneity of deposition. These reactors are also equipped with flow- and pressure-balanced ventilation/run pneumatic switching systems, which means that all the gas flows are continuously running. Switching valves that are close to the reactor cell direct the flows for the desired times, either toward the ventilation or toward the reactor. This technique enables very quick gas exchanges within the growth chamber ($\approx$ seconds)—a necessary condition for fabrication of abrupt semiconductor heterotransitions.

The chemistry takes place onto and close to the heated substrate. Due to the existence of a vapor phase, the reactants reach the surface by convective and finally diffusive transport. Diffusive transport within a concentration gradient boundary layer contributes to the high homogeneity of deposition onto uniform substrates and simultaneously to a high selectivity of deposition in case of growth on partially masked or patterned surfaces. Experimental as well as theoretical studies aimed at resolving the relative importance of gas phase and surface diffusion for selective growth have shown that the contributing surface diffusion remains relatively neg-

ligible [9]. This finding shows a distinct difference from MBE, where the lateral movement of reactants is only achieved by surface diffusion.

The optimum deposition conditions are to a large degree determined by the reactivity of the precursor materials and the stability of the deposited materials. Whereas at higher temperatures the growth is limited by mass transport (convective flow + vapor diffusion), at lower temperatures the limits are given by thermally activated processes (dissociation of the precursor molecules, desorption of reaction products, surface diffusion, etc.). Some precursors limit the range of temperature for depositions. PH_3 and TMGa (trimethylgallium), for instance, show increasing kinetic inertness if used for too-low temperatures. TMAI (trimethylaluminum) has strong Al–C bonds and tends therefore to incorporate C into the epilayer. These problems can be solved in part by changing the precursors to, for instance, TBP (tertiary-butyl-phosphine), TEGa (triethylgallium), and $(\text{CH}_3)_3\text{N-AlH}_3$ (trimethylaminealane), respectively, which decompose at lower temperatures. However, the deposited materials also set some limits. In order to get high-quality material and to avoid thermal decomposition, the different materials need different optimum temperatures for deposition—for example, InAs at around 500°C , InP at around 600°C , GaAs at around 650°C , and GaN at around 1000°C . At these temperatures the metal–organic compounds to a large degree decompose at the hot surface, freeing the metal atoms for incorporation into the epitaxial layer. This is a one-way process. Due to the existence of a protecting atmosphere above the substrate surface, the re-evaporation of metals away from the surface is negligible in comparison to MBE (where it is most critical for the element In).

In this chapter, we describe the use of MOVPE for producing low-dimensional structures, especially quantum dot structures by self-assembling. Self-assembling in highly mismatched systems already starts after deposition in the range of 1 to 2 ML. This is, of course, a challenge for a deposition method, which has to be calibrated for monolayer depositions by after-growth measurements. We will demonstrate that this is possible.

4.3 Quantum dot structures by self-assembling—general considerations

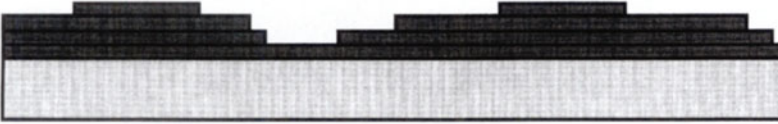
4.3.1 The principle

The growth of dislocation-free, coherent three-dimensional (3D) islands by self-assembling, most probably reported for the first time by Goldstein et al. [10] for MBE-grown dots of InAs, in GaAs, later seen by Eaglesham and

Step 1: deposition of a 2D (wetting) layer:



Step 2: growth front roughening, 3D nucleation:



Step 3: growth of coherent 3D islands:

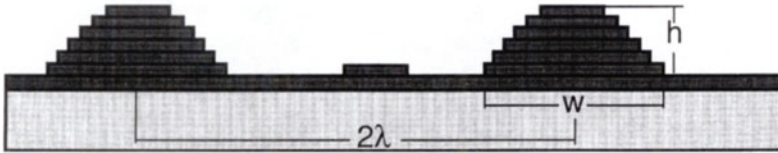


Figure 4.2: Schematics of self-assembling in the Stranski–Krastanow growth mode. The distance λ , with w , and height h are characteristic values, dependent on the deposition parameters and the materials combinations.

Cerullo [11] also for 3D islands of Ge on Si(001), is now an almost established technique for the preparation of quantum dot structures (for an overview, see, for instance, [12–16]). The term *self-assembling* characterizes a process where a strained 2D system reduces its energy by a transition into a 3D (or 2D + 3D) morphology. The principle is as follows (Fig. 4.2):

In the first step, epilayer material with a specific surface energy lower than that of the substrate grows for one or a few monolayers pseudomorphically onto the substrate surface (wetting layer). For lattice mismatched materials, an increasing elastic strain builds up with increasing thickness and the surface starts to roughen, i.e., the 2D layer-by-layer growth mechanism breaks down, and in a second step, 3D nucleation sets in. In a third step, the 3D islands develop in size by consuming the material that is mobile at the surface or that can be mobilized (for instance, by a partial decomposition of the originally grown wetting layer).

Often the term *Stranski–Krastanow growth mode* is used synonymously for this type of self-assembling (even though Stranski and Krastanow, in their original paper, discussed this growth mode for a strain-free system [17]). This growth mode stands for 2D layer-by-layer (wetting) and then 3D island growth, i.e., for a special epilayer morphology. Another morphology develops when 3D islands or clusters crystallize on a bare, unwetted substrate surface. In this case, known as *Volmer–Weber growth mode* [18] and often observed for metal deposits on semiconductor or insulator surfaces, nanosized 3D islands of rather high size homogeneity can be obtained. The term *self-assembling*, in a wider sense, should therefore include Stranski–Krastanow (SK) as well as Volmer–Weber (VW) growth.

Self-assembling is “nature’s way” to produce quantum dots. One fortunate circumstance in nature makes this possible: in order to achieve quantum confinement effects, the material forming the dots should have a lower bandgap than the surrounding barrier, and simultaneously it should have a larger lattice parameter in order to come under compressive strain. And this is precisely the case observed for most of the known SK systems, for instance, Ge/Si, InAs/GaAs, InP/GaInP, GaN/AlN, and others. One could object here that, in principle, for strain-driven processes, the sign of the strain is not important. Therefore, self-assembling should not be excluded for epilayers under tension. In fact, 3D islanding under tension has also been shown, e.g., for the materials combination PbSe(dots)/PbTe(111) [19]. However, for physical reasons, the resulting buried islands under normal circumstances should then behave as “antidots” (it is an interesting question whether systems exist in which the bandgap lowering due to tensile strain leads to their conversion into real dots).

Because of their damage-free interfaces, self-assembled and overgrown dots are optically active and promising candidates for device applications. In general, the spontaneous nature of the self-assembling produces randomly distributed 3D islands. Under certain conditions, some self-organization effects in lateral as well as in vertical direction can be realized. This opens the door for highly advanced materials engineering work on a nanoscale, using such dots as the building blocks.

4.3.2 The mechanism of self-assembling

Considering the spontaneous nature of the self-assembling process, the results of carefully optimized experiments—highly uniform and coherent 3D islands—are rather surprising. In order to explain these phenomena, we have to consider the following experimental facts:

1. Self-assembling is observed only for lattice mismatched systems, i.e., it is intimately related to the accommodation of elastic strain associated with an epitaxial lattice misfit $\Delta a/a$.
2. The results of self-assembling (surface densities and sizes of 3D islands) are affected by the growth parameters, of which the total amount of deposited material, Q ; the deposition rate, R ; and the deposition temperature, T , are the most important parameters.
3. The fact that 3D island sizes are dependent on the deposition parameters implies that there is no strain-dependent size definition for 3D islands.
4. Special issues to explain are the occurrence of bimodal size distributions in some cases (for instance, SiGe/Si, InP/GaInP/GaAs, and InAs/InP), the island faceting, the thermal stability of 3D island structures, and others.

At present, there are still controversial debates about the different models, based either strictly on thermodynamics [20–23] or more on kinetics [24–27] in order to explain the self-assembling process. The driving force of the process—the strain in the epilayer-substrate combination—is clearly defined by thermodynamics. However, kinetics becomes important as soon as activation barriers appear on the pathway between the 2D and the Stranski–Krastanow 2D + 3D morphology. These are, for instance, activation barriers for the 3D nucleation process on top of the 2D wetting layer, for diffusion of adatoms towards the 3D islands, for reorganization of material during ripening processes, etc. It is an interesting question whether the observed SK morphology is then the equilibrium or a consequence of the pathway through which it was created. Our experience in MOVPE of InP(dots)/GaInP/GaAs(001) and InAs(dots)/InP(001), as well as our critical analysis of the results of other groups using other materials combinations and other growth techniques, has led us to believe in a model that can be concisely characterized as a *nucleation and growth* model. It combines the thermodynamic aspects of self-assembling as a strain-driven process and the kinetic aspects of self-assembling as being strongly affected by the history of the process, specifically the detailed conditions during the nucleation of 3D islands.

The process is illustrated in Fig. 4.3., which shows, qualitatively, the changes in the total energy of a mismatched SK system vs. time. Epilayer material is deposited with a constant deposition rate R . The plot can be divided into three sections: period A (2D deposition), period B (2D–3D transition), and period C (ripening of islands).

Period A: At the beginning, the deposition follows a thermodynamically

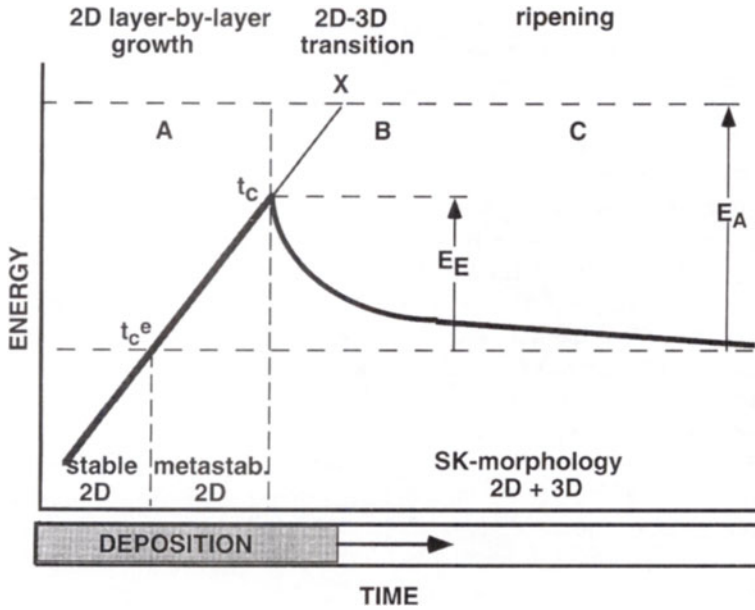


Figure 4.3: Schematics of total energy vs. time for the 2D–3D morphology transition. (A) 2D layer-by-layer growth; (B) 2D–3D transition; (C) ripening period. t_c^e is the thermodynamically defined critical wetting layer thickness, t_c is the wetting layer thickness at onset of 3D nucleation, E_A is the barrier for formation of 3D islands, E_E is the excess energy stored in the supercritically thick wetting layer, and X is the point where a pure strain-induced transition would become possible (no thermal activation).

stable 2D layer-by-layer mechanism that leads to a perfect wetting of the substrate. The accumulated elastic strain energy $E(\text{el})$ increases linearly with the deposited volume, according to Eq. (4.2):

$$E(\text{el}) = 2G \frac{(1+\nu)}{(1-\nu)} \varepsilon //^2 A t, \quad (4.2)$$

where G is the shear modulus, ν is Poisson's ratio, $\varepsilon // = (a_{\text{substr}} - a_{\text{epi}})/a_{\text{epi}}$, A is the area, and t is the layer thickness.

At point t_c^e (the thermodynamically defined critical wetting layer thickness), we leave the stable 2D growth and enter into a region where 2D growth is metastable. An excess strain energy E_E resulting from increasing thickness will be stored within the pseudomorphic 2D layer, i.e., a supercritically thick wetting layer builds up. This means that the epilayer

is potentially ready to undergo a transition towards a Stranski–Krastanow morphology, which would lower the total energy of the system by

$$\Delta E = \Delta E(\text{el}) + \Delta E(\text{surf}), \quad (4.3)$$

where $\Delta E(\text{el})$ is the (negative) change in the volume strain energy due to the more efficient lateral strain relaxation in 3D islands in comparison to 2D pseudomorphic layers, and $\Delta E(\text{surf})$ is the sum over all the (commonly positive) contributions due to the increasing surface with the formation of 3D islands. The point at which the 2D–3D transition begins depends on the height of the remaining activation barrier $E_A - E_E$ for 3D island nucleation. t_c is then the kinetically defined critical wetting layer thickness. We should mention that the existence of such a range of metastability and with it the formation of a supercritically thick wetting layer is not undisputed. The experiments confirm this, however (see, for instance, [28,29] and the results in Section 4.4 of this chapter). They show that the wetting layer at the onset of 3D nucleation ($= t_c$) is generally thicker than the remaining wetting layer after the morphology transition ($= t_c^e$).

Period B: This period can be divided into two steps: nucleation and growth.

1. Nucleation. The initiation mechanism for nucleation on top of the highly strained wetting layer is not very well understood and is still a subject of detailed theoretical investigations. Exactly at point X , the excess energy E_E of the strained 2D layer reaches the energy of the 2D–3D transition barrier E_A , and—hypothetically—a pure strain-induced 2D–3D transition would become possible. However, we have to consider that at real growth temperatures, a thermally activated nucleation sets in before X is reached. In a metastable strained 2D system with a critical thickness of n ML, any further deposition of material creates potential island nucleation sites in areas of the lateral width $w(x) \cdot w(y)$, where the thickness increases to $(n + z^*)$ ML ($z^* = 1, 2, \dots$). This is caused by the possibility that these $(n + z^*)$ ML-thick islands have to relax laterally and thus to reduce their strain potential relative to the n ML-thick surrounding layer (for further details based on experimental observations, see Section 4.4). In this picture, thickness fluctuations lead to the formation of critical nuclei, and the classical nucleation theories [30] predict for the simplest case:

$$dN/dt = \text{const } \sigma \exp[-(E_A - E_E)/kT] \quad (4.4)$$

(with dN/dt the number of critical nuclei created per time and σ the *super-saturation*, i.e., the concentration of mobile excess material at the surface). Since σ is primarily a function of the deposition rate R , the surface density ρ of 3D islands increases (for not-too-high R) linearly with R , as we will show in more detail in Section 4.4. After the 2D–3D mor-

phology transition, E_E decreases due to partial strain relaxation. This increases the effective 3D nucleation barrier $E_A - E_E$, and further nucleation will be inhibited.

2. Growth of islands. The islands grow and consume the material that is available and mobile enough to reach the islands. One peculiarity at this stage seems to be that under the high strain in the initiation period, due to the low effective nucleation barrier, more smaller (but already supercritically 3D) islands are created in the beginning than afterwards survive and grow up to larger ones (see the results discussed in Section 4.4). This would mean that immediately after the partial strain relaxation due to the formation of 3D islands, a ripening process sets in whereby only a fraction of the islands survive and grow up, consuming the material from the numerous smaller (and therefore most probably still highly strained) islands and parts of the wetting layer within a radius λ , proportional to a diffusion length $\lambda = 1/2\rho^{1/2}$ (ρ = surface density). In other words, those islands that first reach a size with an effective lateral strain relaxation grow up. Under certain conditions (low-coverage Q for simultaneously high-density ρ), smaller islands are metastable. This leads then to bimodal size distributions. What we finally see after this reorganization processes is a density ρ for the developed “stable” 3D islands, which follows very well:

$$\rho = \text{const.}(R/D)^\zeta, \quad (4.5)$$

where R is the deposition rate and $D = (2kT/h)a_o^2 \exp(-E_D/kT)$, the surface diffusion coefficient, where E_D is the activation energy of surface diffusion, k is Boltzmann’s constant, h is Planck’s constant and a_o is the lattice parameter $a_{//}$ of the substrate. ζ is found to be ≈ 1 for SK growth [31], which follows also from rate equation theory for critical nucleus sizes $i^* \gg 1$, with i = the number of atoms forming the nucleus [32]. This dependence, most clearly visible in the materials combination InP/GaInP/GaAs deposited by MOVPE, but also in MBE experiments of InAs/GaAs as well as VPE of Ge/Si [28], characterizes the self-assembling process as strongly determined by the nucleation history.

Period C: Further ripening. After the 2D–3D transition, the process has lost most of the excess energy. The mobile material is consumed, and the potential differences between slightly smaller and slightly larger islands as well as between islands and the remaining wetting layer maintain a slow further reaction, mediated by surface diffusion.

4.3.3 Sizes and shapes of self-assembled 3D islands

The strain distribution within coherent 3D islands and their surrounding is highly nonuniform. Figure 4.4 shows the calculated strain ε_{xx} in an InP island onto GaInP/GaAs(001). We see the highest compressive strain within

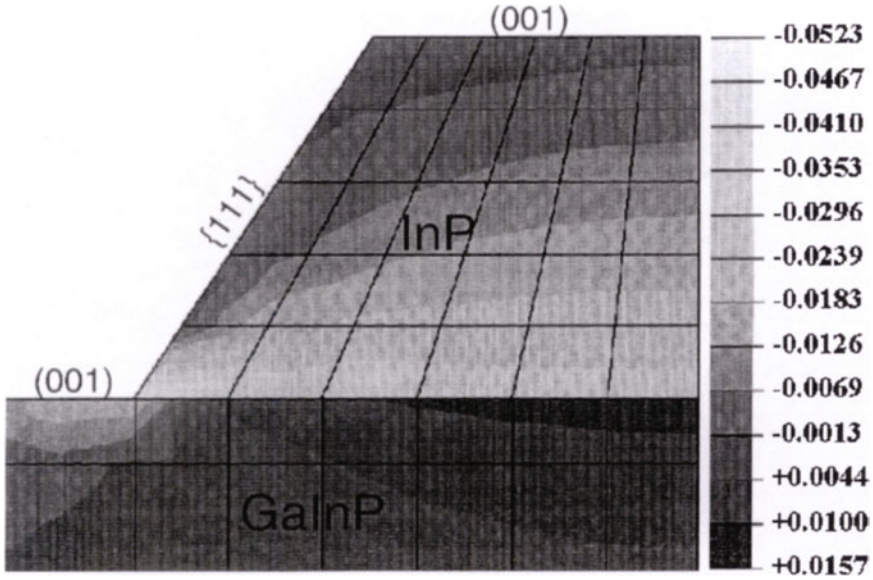


Figure 4.4: Calculated strain ε_{xx} for an InP island onto GaInP/GaAs(001) (finite element method [34]). The island has the shape of a truncated pyramid in accordance with TEM investigations (compare Fig. 4.14) and a height of about 15 nm. Note the high compressive strain at the corner substrate/island and the inversion towards tensile strain (only within the x - y plane!) at the upper (111)/(001) edge of the island.

the InP island close to the GaInP substrate, but also at the corner (001) substrate/(111) island. The compressive strain continues within the substrate outside of the island. Towards the top of the island, the strain decreases, whereby the highest relaxation is reached at the edge (111) facet/(001) top plane. There the total strain is still compressive, but the x - y components become tensile.

This nonuniform strain distribution develops during the 2D–3D transition from an originally laterally homogeneously strained 2D wetting layer. It has important implications for the growth process itself. Strain in general lowers the bond energies and therefore affects the chemical potential $\mu = \delta G / \delta n$ of the material that attaches to or is mobile at the surface. The local change of the chemical potential follows the change in the local strain energy density E along the surface [33], which can be written as (compare Eq. (4.2))

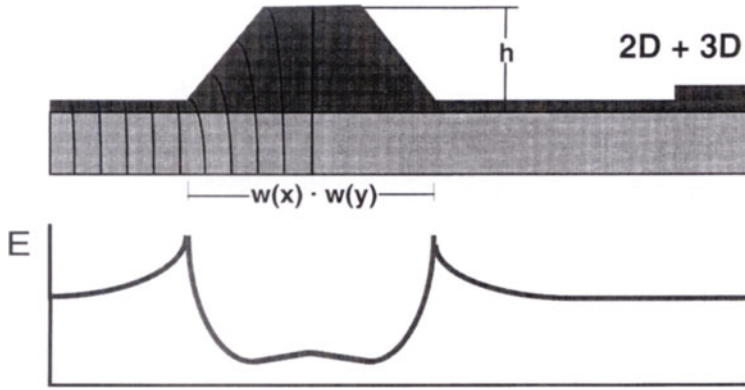


Figure 4.5: A 3D island and the local strain energy density E along the surface of the island and the wetting layer (cross section). The deformation of lattice planes due to lateral strain relaxation in the 3D island is indicated.

$$E = 2G \frac{(1+\nu)}{(1-\nu)} \varepsilon //^2. \quad (4.6)$$

A schematics of the situation is visualized in Fig. 4.5. We see the surface at the most strain-relaxed top of the 3D island as a sink in the chemical potential, where material will be attracted. The highly strained perimeter of the island develops as a peak in the chemical potential, where material will be pushed away. These local changes in the chemical potential are the driving force for reorganization and transport at the surface. The corresponding processes are decomposition (wetting layer and small strained islands), surface diffusion, and vapor-phase diffusion (in the case of MOVPE). The 3D islands grow until equilibrium is reached, i.e., until $\mu(\text{island}) = \mu(\text{wetting layer})$. One peculiarity in this process is that the increase of the chemical potential at the perimeter of the island develops and increases with the size (exactly with the aspect ratio height/width) of the island. This building up of a highly strained barrier around coherent islands is considered to act as a self-regulation and stabilization mechanism of coherent islands [22,35]. This barrier is actually an attachment barrier for material [27], favoring the growth of smaller and inhibiting the growth of larger islands. It is remarkable that the loss of coherence by insertion of a dislocation at the base of the island leads to much faster growth (see, for instance, Fig. 4.13). We will demonstrate the usefulness of this model in several parts of the following sections.

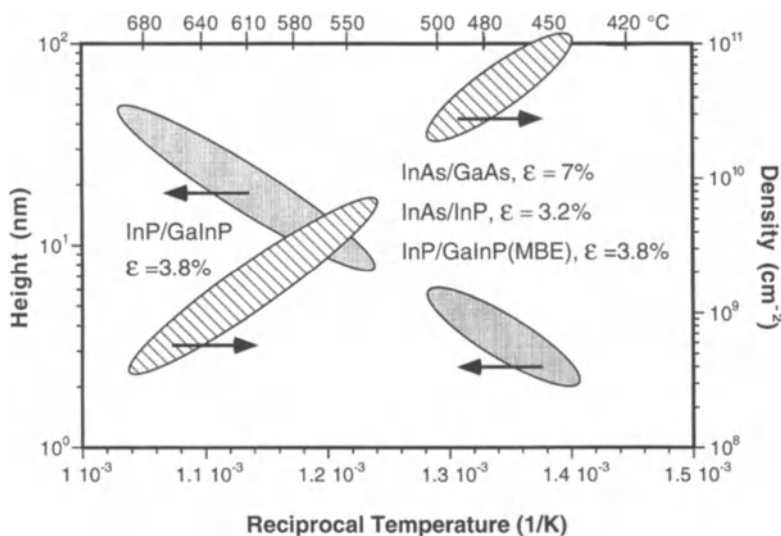


Figure 4.6: Schematic visualization of the areas where most of the experimental height and density values for 3D islands of InP on GaInP/GaAs (MOVPE), InAs on GaAs (mainly MBE or CBE), and InAs on InP (mainly MOVPE) were found. Note that dots of InP on GaInP/GaAs, grown by MBE at 470°C, show height and density values like the other materials grown at that temperature. This figure shows the inverse behavior of density and size, the almost linear dependence of this behavior on reciprocal temperature, and that the misfit in this respect is not a critical parameter.

Another question is whether there is an absolute size limit for the growth of coherent islands. The dependence of island sizes on deposition parameters, such as deposition rate and temperature, suggests that there is not. We will show in Section 4.4 that the size of the coherent islands is mainly a result of conservation of mass: the deposited material distributes over the coherent islands and the remaining wetting layer. Therefore, the sizes of the islands are found to be in the first instance inversely proportional to their surface densities. Since the density of self-assembled islands decreases with increasing temperature in accordance with Eq. (4.5), this implies simultaneously increasing island sizes with increasing deposition temperatures. This is a very general result, not specific for MOVPE growth only (see Fig. 4.6). In this picture we can understand the results of self-assembling as a balanced distribution of material over the entirety of stable 3D islands.

Table 4.1: Equilibrium surface reconstructions of InAs under As-rich conditions and their surface energies (after [36]).

Orientation	Reconstruction	Energy [meV/Å ²]
(100)	a(2 × 4)	48
(100)	c(4 × 4)	47
(111)A	(2 × 2) In vacancy	42
(110)	(1 × 1) relaxed	41
(111)B	(2 × 2) As trimer	36

Before leaving this section, a few words about the shapes of self-assembled 3D islands. The transition from a 2D towards a 3D morphology increases the surface area and reduces in this way the gain in volume elastic relaxation energy. To minimize this energy-consuming step, the islands develop facets of low specific surface energies (we assume that facets are always formed; due to strain they can be bend and their edges can be rounded). Assuming that diffusion of atoms at the island surface is faster than material exchange reactions between the islands and the wetting layer, then these facets should represent the equilibrium situation. Equilibrium shapes can be calculated using Wulff's theorem for given specific surface energies $\sigma(\text{hkl})$. These are for zincblende and diamond structures, commonly low-index planes. Theoretical calculations of equilibrium shapes [36] are in very good agreement with the experimental observations for what we name *fully developed* islands in the case of InP/GaInP, i.e., truncated pyramids, slightly elongated in $\langle 110 \rangle$, with the dominating facets $\{111\}$ B and $\{110\}$. To a lesser extent, $\{111\}$ A facets also appear in larger islands. The top facet is (001). For smaller islands, the area of $\{111\}$ facets shrinks at the expense of larger areas of $\{110\}$. The latter facets then dominate in the smaller InAs/GaAs islands [37] (see Table 4.1).

An interesting aspect of this process is shape transitions, which occur during the growth of 3D islands. These shape transitions are visible in systems where, under some growth conditions, smaller 3D islands coexist with larger 3D islands (examples are SiGe/Si [28,38,39], InP/GaInP [16,40], and InAs/InP [41]). That they can coexist is primarily dependent on the total amount Q of material that has to be distributed over a finite number of nucleated 3D islands. As mentioned in Section 4.3.2, the density of 3D islands nucleating under the high strain in the initiation period of the

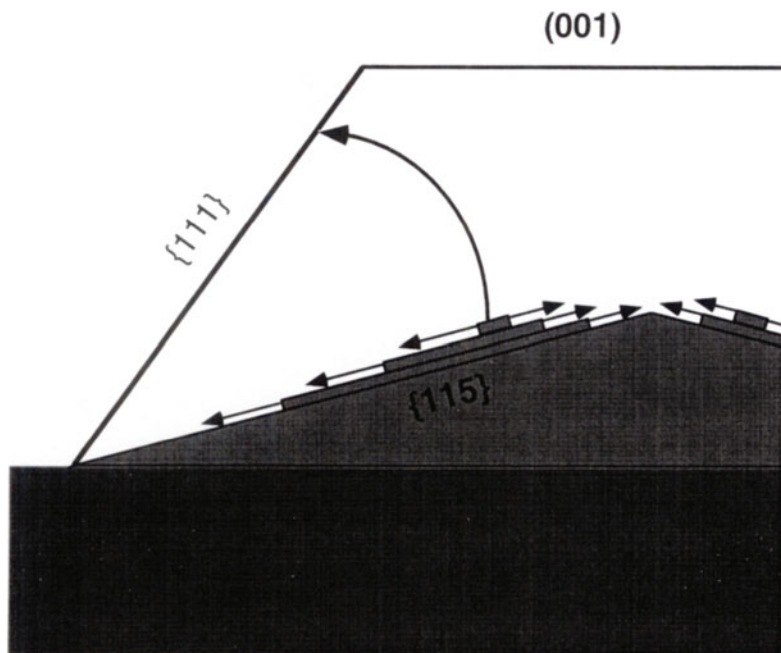


Figure 4.7: Visualization of a kinetic mechanism of a fast transition from small to fully developed islands. Nucleation preferentially starts at the most strain-relaxed apex region of small islands. This nucleation creates facets of vicinal character, which can grow very fast. Note that during this process the islands' base diameter does not increase; it could be even shrinking if material detaches from the highly strained area.

2D–3D transition is rather high. The islands formed in this period are therefore flatter, with low-angle facets like $\{10n\}$ and $\{11n\}$, where n is a number between 3 and 6. We will show in Section 4.4 that the small islands are metastable. They disappear with larger amounts of deposited material and/or during annealing periods. In the subsequent short ripening period, those islands with the most efficient strain relaxation i.e., those with highest aspect ratio, will succeed and grow up to fully developed islands. Figure 4.7 demonstrates the mechanism by which very fast growing vicinal facets can be formed by nucleation at the apex of small islands. For coherent islands of zincblende and diamond structures, the development of $\{111\}$ facets seems to be the ultimate end of ripening processes.

4.4 MOVPE-grown self-assembled quantum dot structures

4.4.1 Materials overview

The best-investigated III–V materials combination is the system $\text{Ga}_{1-x}\text{In}_x\text{As}/\text{GaAs}(001)$, from which most results come from MBE as the growth method. However, MOVPE can be used as well, with very similar results, using TMGa (trimethylgallium), TMIIn (trimethylindium), and AsH_3 as the precursors in hydrogen as the carrier gas [42,43]. In this material system, one has to be careful to use low AsH_3 pressures; otherwise, the individual dots tend to coalesce into larger clusters. This experience has already occurred with MBE [21]. Dots of InAs embedded into a GaAs matrix behave as true quantum dots, with sizes typically on the order of 15–30 nm in base diameter and 3–7 nm in height. Typical densities are on the order of 10^{10} – 10^{11} dots/cm², whereby self-organized arrays of dots can be formed. Due to the high misfit of $\approx 7\%$, these dots are highly strained; their luminescence is commonly observed in the range between 1.15 and 1.3 eV.

Very similar are dots of InAs grown on $\text{InP}(001)$ (misfit $\approx 3.2\%$). In this case, gassource MBE [41] and MOVPE [44–46] are the growth technologies used so far. Despite the lower misfit, typical dot sizes and densities are on the same order of magnitude as observed for InAs dots in combination with GaAs. The reason for this is clearly the same deposition temperature in the range of about 480–530°C, which is commonly used for growing InAs. A complication one must take care of in this system is the high reactivity of InP with AsH_3 , forming $\text{InAs}_x\text{P}_{1-x}$ interfaces at higher temperatures. Dots of InAs embedded into InP show an intense luminescence at about 0.8–0.9 eV at room temperature. Thus, this heterocombination is of particular interest since it can be regarded as a candidate for the realization of low-threshold-current quantum-dot lasers operating at the technologically interesting wavelength of 1.55 μm .

Another materials combination, widely investigated by MOVPE [40,47,48] but less so by MBE [49], is the system $\text{InP}/\text{GaInP}/\text{GaAs}(001)$ (misfit $\approx 3.8\%$), with InP forming the dots and GaInP the barrier, lattice matched to the substrate GaAs. Since optimum growth temperatures for InP in MOVPE are higher than those for InAs (best in the range between 550 and 700°C), MOVPE-grown dots of InP are generally larger. Typical base diameters are in the range of about 40–80 nm, and heights are in the range of about 12–25 nm, depending on deposition conditions. Embedded into GaInP, these dots, if fully developed, luminesce at about

1.65 eV. This heterocombination is particularly attractive for optoelectronic applications based on GaAs substrates and for basic studies of self-assembling. The growth parameters can be varied across a wide range—the deposition temperatures in MOVPE, for instance, between 550 and almost 700°C, which significantly affects surface densities, sizes, and shapes of the islands. An extension to even lower temperatures of about 470°C can be made by using MBE as the deposition technique [49]. With reference to density and size, these MBE results follow a straight extrapolation of the MOVPE results for lower temperatures.

Attempts have also been made to grow InP dots on GaP(001) (misfit $\approx 7.75\%$). In this case, however, only very poor size homogeneity has been reached so far, and no dot-related luminescence could be observed. In situ RHEED observations in chemical beam epitaxy (CBE) indicate that InP/GaP is eventually a Volmer–Weber system, i.e., no complete wetting layer of InP covering the GaP surface could be detected.

Less investigated up to now is the system GaSb/GaAs (misfit 7.8%). After MBE investigations (growth temperature 470°C), dot sizes and densities behave very similarly to InAs/GaAs [50]. MOVPE growth using the precursors TMGa (trimethylgallium) and TMSb (trimethylantimony) at 540°C results in much larger dots, with tendencies to elongate and to coalesce [51].

An active field of study is in situ-grown quantum dots in nitrides. Due to the smaller Bohr radius for electrons and holes in these materials (for GaN, $r_B \approx 3$ nm), the dots have to be very small in order to show 3D confinement effects. Unintentionally MOVPE-grown dots of $\text{Ga}_{1-x}\text{In}_x\text{N}$ instead of planar quantum wells within barriers of GaN are discussed as the reason for the observation of localized states in GaInN/GaN laser and light-emitting devices [52]. In this case, in addition to strain, the existence of an immiscibility gap for the compound GaInN may also be a driving force for dot formation. Several attempts have been undertaken to grow dots of GaInN deliberately on GaN/sapphire(0001) or on GaN/SiC(0001). One difficulty is the high dislocation density in these heterostructures. Dislocations with a screw character of an areal density of more than 10^8 dislocations/cm² act as nucleation centers and produce classical dislocated growth spirals in accordance with BCF theory, instead of coherent 3D islands [53]. Coherent GaN dots on AlGaIn have been produced after a predeposition of Si onto the AlGaIn surface. Si deposits obviously affect the nucleation process [54]. That GaN dots also form without such special treatments has been shown in MBE experiments at 680–730°C for the combination GaN/AlN (misfit 2.5%) [55].

The discussion above concerned only materials aspects. In all cases, the dots are formed by self-assembling as soon as there is a misfit in the

Table 4.2: Growth conditions for InAs dots under MOVPE.

Total pressure:	50 mbar
Substrates:	InP:Fe, (001) misoriented 0.2 °toward <011>
AsH ₃ pressure:	2.5×10^{-5} bar
Deposition rate InAs:	0.3 ML/s
Switchings:	No growth interrupt during PH ₃ /AsH ₃ -switch; 12s annealing after InAs at growth temperature (in AsH ₃ /H ₂)
Variation parameters:	Deposited amount of InAs and deposition temperature

range of about 2–8%. However, a structural aspect must also be addressed: self-assembling is dependent on the surface orientation. For example, coherent dots of InAs, easily formed on GaAs(001), do not form on GaAs(111)Ga surfaces. For this orientation, the formation of a dislocation network is the preferential strain-relaxation mechanism [56]. One can speculate about a too-low surface mobility of In on Group III-terminated polar surfaces.

In the following, we select for a more detailed discussion about self-assembling in MOVPE the materials combinations InAs/InP(001) and InP/GaInP/GaAs(001).

4.4.2 InAs on InP(001)

The specifics of InAs on InP(001) are given by the high sensitivity of an InP surface against Group V exchange reactions when switching off the PH₃ flow and switching on the AsH₃ flow. One must consider that, during switching, most of the upper P-monolayer of InP will be replaced by arsenic [4], increasing in this way the net amount of InAs for wetting or islanding at the InP surface. Low temperatures and low AsH₃ pressures help to minimize but not to completely avoid this effect.

The InAs dots, presented in Figs. 4.8. to 4.10, were grown by MOVPE under the conditions [46] given in Table 4.2.

The development of coherent InAs islands on InP with increased deposition can be seen in Fig. 4.8, which shows contact-mode AFM images together with height histograms for nominal InAs depositions of (a) 1.5, (b) 2.4, and (c) 3 ML at a growth temperature of 500°C. In this range of InAs deposition, the surface density of dots increases one order of magnitude, from 3.0×10^9 to 3.0×10^{10} dots/cm². The histogram shows a

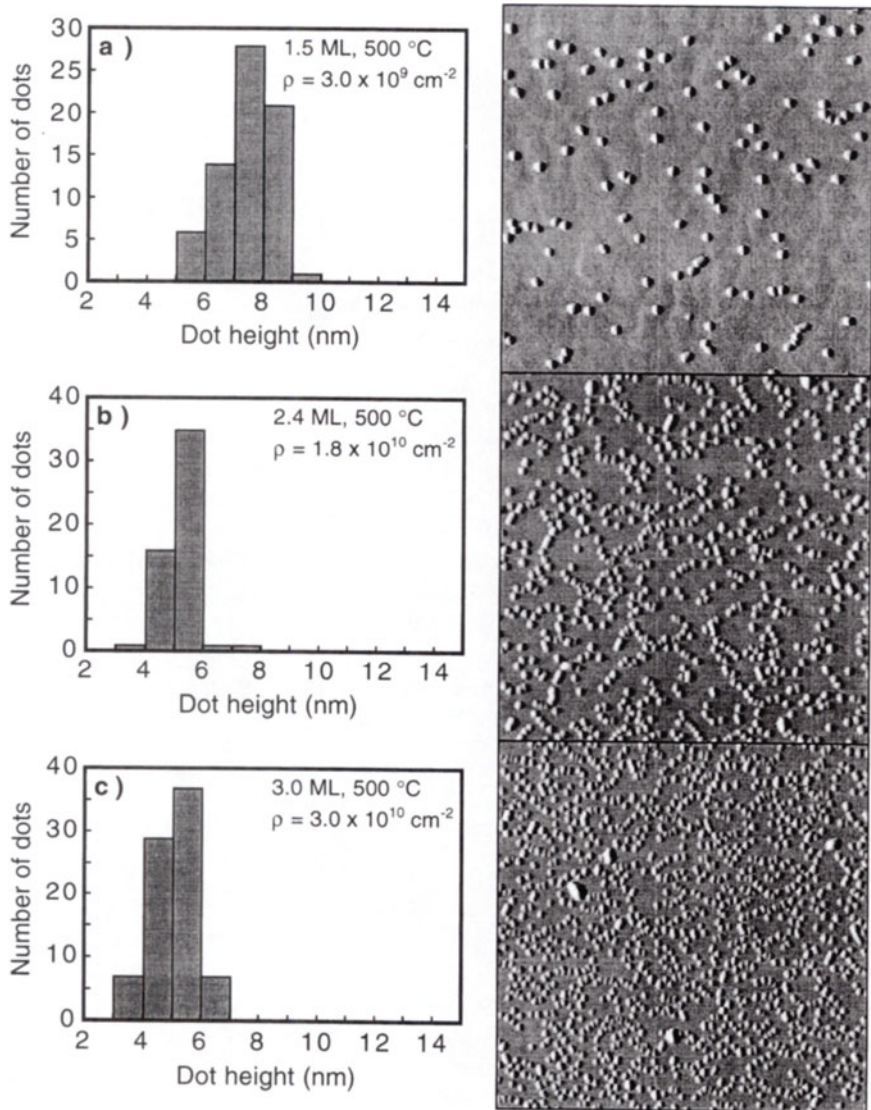


Figure 4.8: Deposition-dependent height distributions of coherent InAs dots on InP, shown together with $2 \times 2 \mu\text{m}^2$ shaded AFM images, acquired by contact-mode imaging [46].

significant result: despite the increase of the total amount of deposited InAs, together with the density increase, the dots become smaller. This inverse behavior between densities and sizes of coherent dots is a characteristic phenomenon, observable in many cases of self-assembling.

Samples grown under the same conditions as in Fig. 4.8 and then overgrown by a 100 nm InP capping layer have been investigated with PL. The resulting spectra are shown in Fig. 4.9. This material combination shows a strong dependence of the PL signal on the surface coverage. In the sample with the lowest dot density (1.5 ML InAs), only a broad doublet peak centered around 1.2 eV and corresponding to a rough InAs wetting layer of about 2 ML is visible (it cannot be ruled out that the relatively large dots luminesce at lower energies, not detected in these measurements, or that not enough dots visible in PL were formed in this particular sample). At an intermediate deposition (2.4 ML InAs), the wetting layer luminescence is still dominating, but the peak has sharpened and simultaneously shifted to a higher energy, indicating a smoothening of the wetting layer in analogy to results obtained for the InP/GaInP system (see Fig. 4.12). Besides the wetting layer peak at 1.24 eV, a broad peak appears at 0.78 eV, which is assumed to originate in the quantum dot emission. For the highest deposition (3.0 ML InAs), the intensity of the wetting layer peak decreases and the dot peak becomes dominant.

The deposition temperature can only be varied in a rather narrow range for this material. In this specific case, experiments were done at deposition temperatures of 480, 500, and 520°C for a nominal deposition of 3 ML of InAs. Histograms of the height and surface densities of coherent InAs dots on InP are shown in Fig. 4.10, together with PL spectra of corresponding capped samples. Increasing the growth temperature from 480°C to 520°C, the average height increases slightly, thus following the general trend. Considering the surface density, a minimum value was found for the sample that had the InAs deposited at 500°C. In this case, the decrease in density between 480°C and 500°C follows the general trend, whereas the increase between 500°C and 520°C is different from what has been observed, for instance, for InP dots grown on GaInP/GaAs (see Fig. 4.17). A simultaneous increase of both the size and the density of the islands on the surface must be related to a larger total amount of material in the islands. For the InAs/InP combination, the formation of additional InAs at the interface is therefore assumed to be relevant for island formation at temperatures above 500°C.

The effect on the optical properties caused by changes in the growth temperature can be seen in the PL spectra in Fig. 4.10. For the lowest growth temperature, 480°C, the dot luminescence, peaking at 0.76 eV, coexists with a broad emission between 1.1 and 1.2 eV. This high-energy

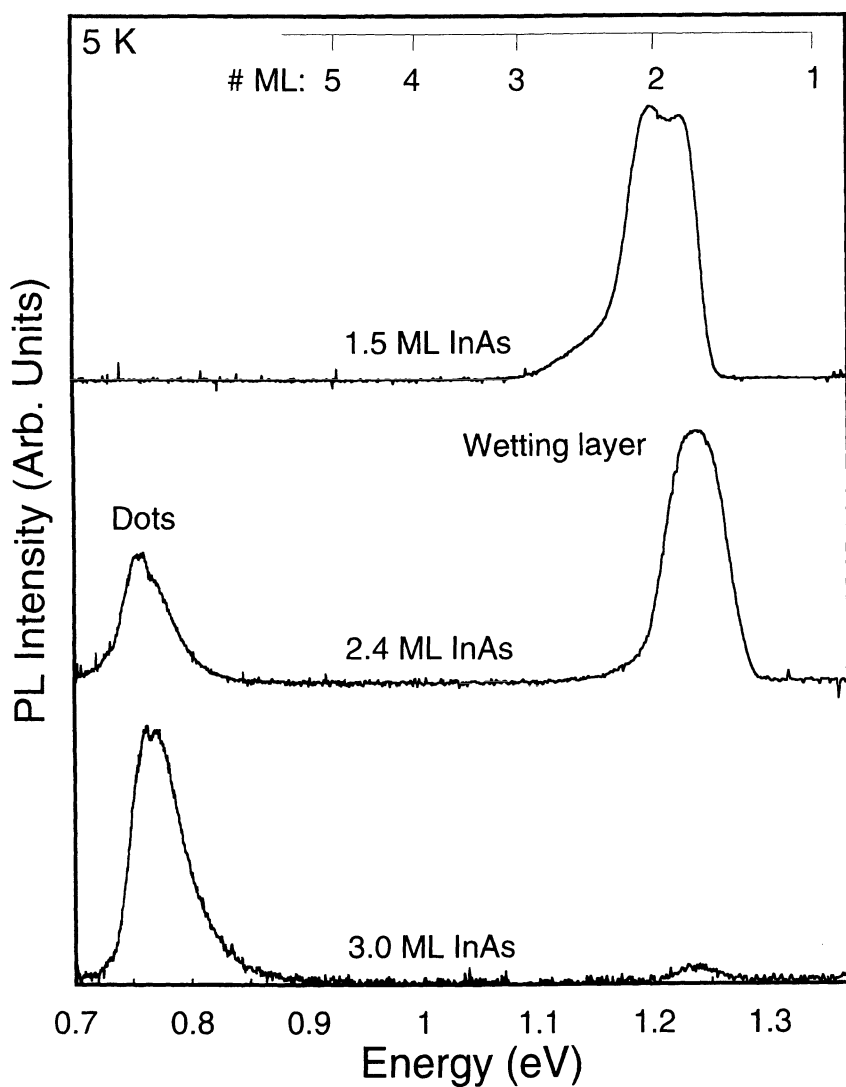


Figure 4.9: 5 K PL spectra of InAs quantum dots embedded in InP as a function of nominal InAs deposition at 500°C. The insert shows calculated emission energies for biaxially strained InAs quantum wells in InP for thicknesses up to 5 ML [46].

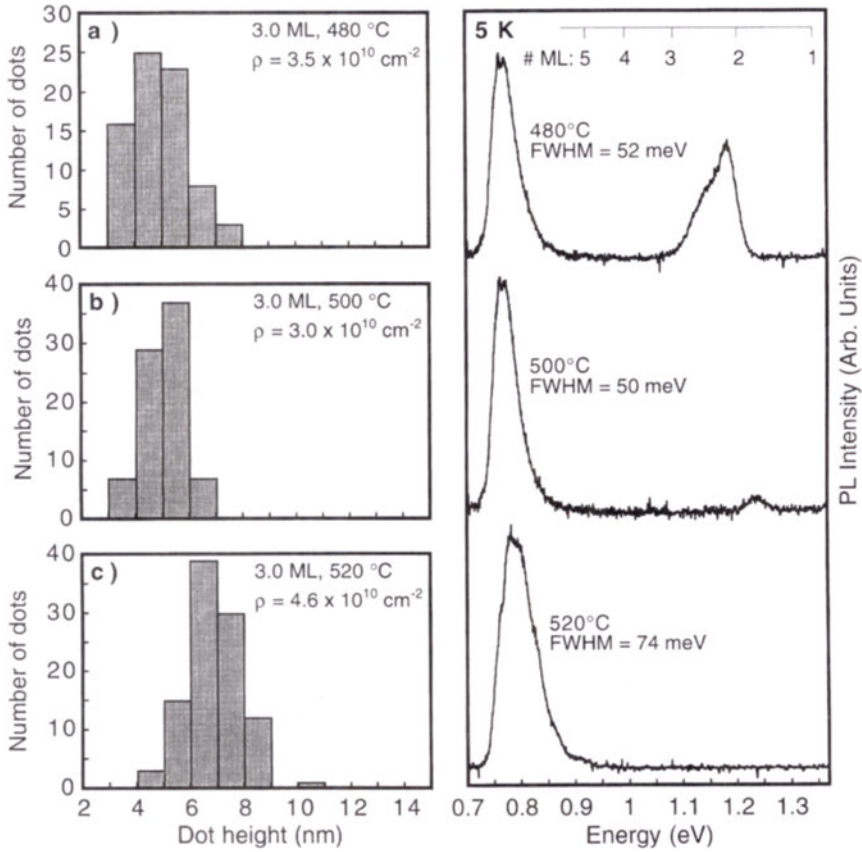


Figure 4.10: Temperature-dependent height distributions of InAs dots on InP, measured by AFM. Nominally 3.0 ML InAs was deposited on top of the InP surface. 5 K PL spectra are from embedded dots deposited under equivalent conditions and subsequently capped by 100 nm InP [46].

emission is supposed to originate from a rough InAs wetting layer. Obviously, a 12 s growth interruption at 480 °C is not enough to capture into the fully developed dots all the excess material stored in the wetting layer.

At 500 °C, only a weak peak from the wetting layer located at 1.24 eV shows up in the spectrum, most probably from a nearly perfect two-dimensional layer, since this emission occurs between the estimated energy values for 1 and 2 ML of biaxially strained InAs layers in InP. As the growth temperature was further increased, to 520 °C, the dot luminescence broadened, mainly towards the high-energy side. The blueshift is unex-

Table 4.3: Conditions for the MOVPE experiments in Section 4.4.3.

Total pressure:	100 mbar
Substrates:	GaAs:Si, (001) exact
PH ₃ pressure:	1.5×10^{-3} bar
Deposition rate:	Variable, in standard runs InP 0.5–0.6 ML/s, GaInP 1.1–1.2 ML/s
Switchings:	Before deposition of InP in most of the samples, a ≈ 2 ML-thick GaP layer was deposited (the effect is a slight delay of InP dot formation [57]) For InP depositions directly on GaAs, a ≈ 2 ML-thick GaP layer was always deposited in order to suppress As carryover effects (GaP-stabilized GaAs)
—Annealings:	Variable, in standard runs 12 s after InP deposition at growth temperature (in PH ₃ /H ₂)
—Variables:	Deposited amount Q of InP, deposition rate R , deposition temperature T , annealing times t

pected, since the increase in average dot size, as seen by the AFM, should lead to a decrease in emission energy due to a weakening of the quantization in the larger dots. It can, however, be understood in terms of a mechanism whereby the phosphorus, replaced by arsenic at the lower heterointerface, is subsequently recaptured into the dots, thus forming dots composed of an InAs_xP_{1-x} alloy, rather than pure InAs. For practical applications, the energy of this quantum dot luminescence can be expected to be tunable to some degree either by alloying of the dot material (InAsP, GaInAs) or by modifications of the barrier (Al alloys).

4.4.3 InP on GaInP/GaAs(001)

4.4.3.1 Phenomenology of the 2D–3D transition

Despite the fact that, in comparison with RHEED, MOVPE does not have such excellent tools for in situ observations, if we use PL in combination with AFM and TEM as characterization tools we can very easily follow the different steps of the 2D–3D transition. Unless other data are given below, the MOVPE experiments presented in the subsequent sections were done under the conditions outlined in Table 4.3 [47,57].

Figure 4.11 shows PL spectra of InP/GaInP structures containing dif-

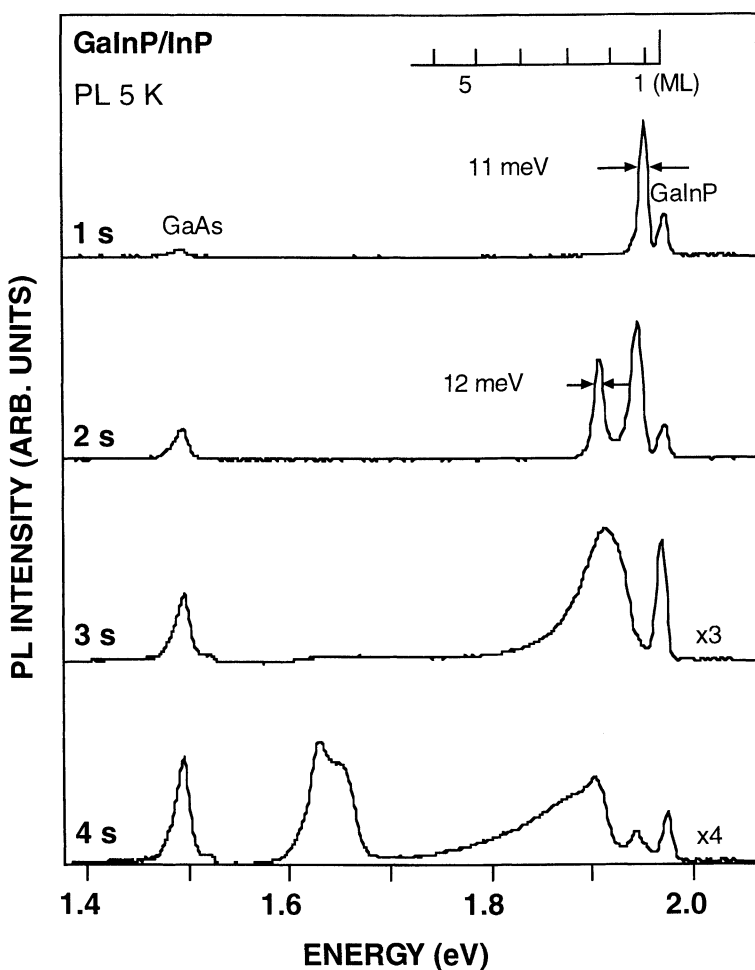


Figure 4.11: The development of PL spectra with increasing InP-on-GaInP deposition time. The deposition rate is about 0.6 ML/s. InP is embedded between 300 nm GaInP barrier layers (MOVPE, 580°C). The insert shows calculated energy values for biaxially strained InP quantum wells. The peak at 1.64 eV represents the luminescence from fully developed coherent InP islands; the continuous emission between 1.7 and 1.85 originates from smaller islands [47].

ferent quantities of InP embedded in GaInP (lattice matched to GaAs) and grown at 580°C by low-pressure MOVPE. In agreement with the scenario for SK growth, discussed in Section 4.3.1, we see the luminescence of 1 ML- and 2 ML-thick 2D quantum wells for deposition of 0.6 and 1.2 ML InP. With more InP, the strain increases and the growth front roughens, observable as the broadening of the quantum well luminescence for 1.8 ML InP. With deposition of even more InP, luminescence from 3D islands becomes visible. In this case, two types of coherent islands are formed: smaller islands with the luminescence in the range between 1.70 and 1.85 eV, and fully developed islands, responsible for the doublet peak at around 1.64 eV. We will come back to this case of bimodality when we discuss the AFM and TEM results.

These samples above were grown continuously, i.e., there was no growth interruption after InP deposition. Very important for the result of self-assembling is, however, the time period immediately after the onset of 3D nucleation. In Fig. 4.12, the development of the PL spectra is shown for different annealing times, namely, 0 s, 3 s, and 12 s after InP deposition before overgrowth of the structure by a GaInP top layer. The deposited amount of 1.8 ML of InP is slightly above the (kinetically defined) critical wetting layer thickness for onset of nucleation. We see that without annealing, i.e., for continuous growth, the luminescence at about 1.64 eV, arising from fully developed dots, is very weak. It increases for 3 s annealing time, while growth interruption of more than 12 s had no further effect on the development of the characteristic dot luminescence peak. This means that 3D island formation can be suppressed to some degree by continuous deposition of material and that coherent islands develop during growth interruption very quickly within a few seconds only. Also remarkable are the changes in the range of wetting-layer luminescence. The broad peak, corresponding to emissions from a rough wetting layer in the sample grown without an interruption, splits up after annealing into well-resolved peaks corresponding to emission from more extended areas of 1 ML- and 2 ML-thick quantum wells.

These experiments illustrate that 3D islands develop at 580°C within a time span of only a few seconds. This happens obviously by reorganization of material deposited primarily two-dimensionally at the surface, since the islands in this case grow without further materials supply. This outcome supports the model developed in Section 4.3.2, specifically the assumption of the existence of a range of metastability for a supercritically thick wetting layer.

A further interesting point coming out from these experiments is the extremely high growth rate of the 3D islands. The fully developed islands, responsible for the light emission at about 1.64 eV, have typically heights

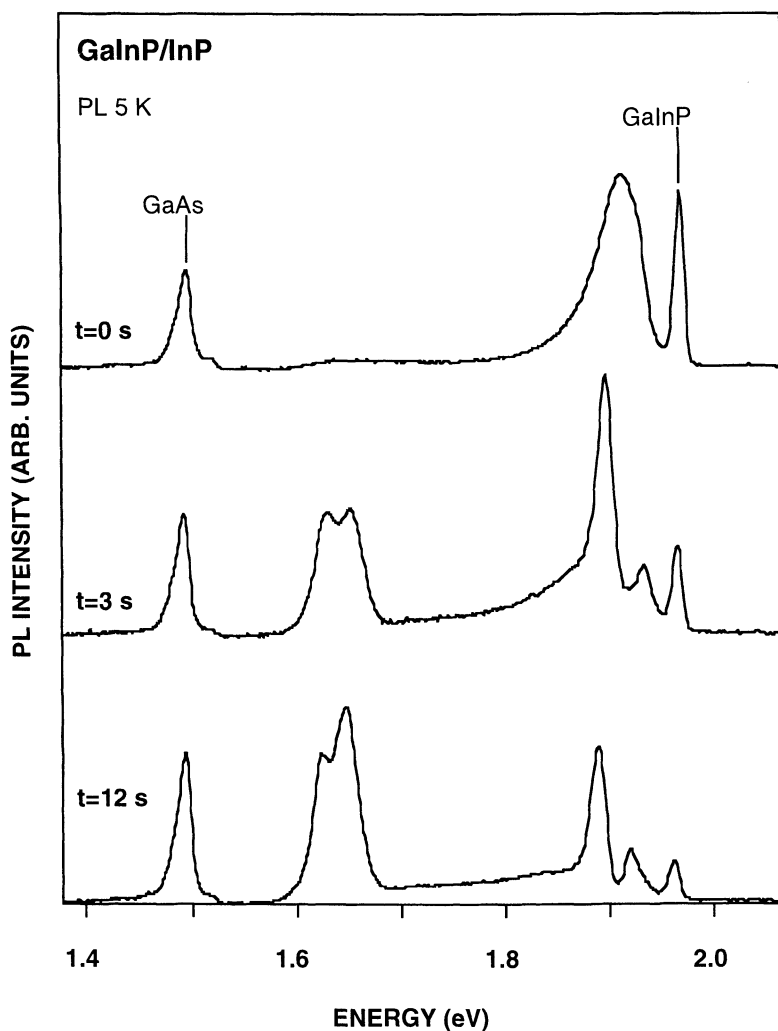


Figure 4.12: Development of 3D island luminescence at ≈ 1.64 eV for different annealing times 0 s, 3 s, and 12 s after about 1.8 ML InP deposition before capping by GaInP. The other conditions are as in Fig. 4.11 [47].

of more than 10 nm, i.e., they must grow up to this heights in a few seconds (compare with typical $\langle 001 \rangle$ growth rates of 0.3 nm/s for InP in MOVPE). This fact also speaks for essential contributions of reorganization effects to growth of the islands.

Which kind of reorganization can be expected? It seems very likely that above a certain thickness, parts of the wetting layer become “active,” i.e., after the formation of 3D islands they push off the accumulated excess material and thin down to a thickness t_c^e , which is more or less in equilibrium with the 3D islands. Hand in hand with this mobilization of material from the wetting layer, materials transfer from smaller towards larger islands, i.e., we see Ostwald ripening processes during the early stages of island formation. For the InP/GaInP system, as for $\text{Si}_{1-x}\text{Ge}_x/\text{Si}$ [38,39], we observe a bimodal size distribution of coherent islands [40]. This distribution is dependent not only on the deposited material but also on the deposition rate, the deposition temperature, and annealing times. For a low amount of deposition, low temperature and no annealing dominate small flat islands with rather high density on the order of 10^{10} cm^{-2} . When embedded into GaInP, these islands are highly confined in the growth direction ($h < 6\text{ nm}$) and give rise to photoluminescence emissions in a range between 1.7 and 1.85 eV (see Fig. 4.11). The disappearance of these smaller islands with increasing amount of deposited material or by annealing of the sample characterizes the islands as preliminary stages to the “fully developed” larger islands with steeper facets—in this material, typically facets of type $\{111\}\text{B}$ and $\{110\}$. With even more material, noncoherent islands are formed with larger size but principally the same shape. An AFM picture of a surface with the three different types of 3D islands is shown in Fig. 4.13. There are almost no islands with intermediate height values between the small and the fully developed islands. This result is confirmed by a careful height/diameter (aspect ratio) analysis of InP/GaAs islands, grown under conditions similar to those of the sample in Fig. 4.13. This analysis [58] shows a critical diameter of about 40 nm for the small islands, whereas with more deposition the height of the islands suddenly reaches the value of more than 10–12 nm, corresponding to the size and aspect ratio of the fully developed islands. A further interesting aspect of Fig. 4.13 is the sporadic random distribution of coherent islands, which shows that in first approximation their sizes are unaffected by the island environment—i.e., environment does not play a role, whether the islands develop in isolation in an island-free area or whether they develop in the neighborhood of noncoherent islands. This finding will help determine whether coherent islands have their own size-regulation mechanism.

A more detailed view of the shape of fully developed islands is given in the high-resolution TEM picture of Fig. 4.14a and 4.14b. The figures show the InP islands as truncated pyramids with $\{111\}\text{B}$ and $\{110\}$ side facets, slightly elongated in $\langle 110 \rangle$.

The truncated pyramid shape with $\{111\}\text{B}$ and $\{110\}$ facets and almost no $\{111\}\text{A}$ facets is in accordance with the higher surface energies of $\{111\}\text{A}$

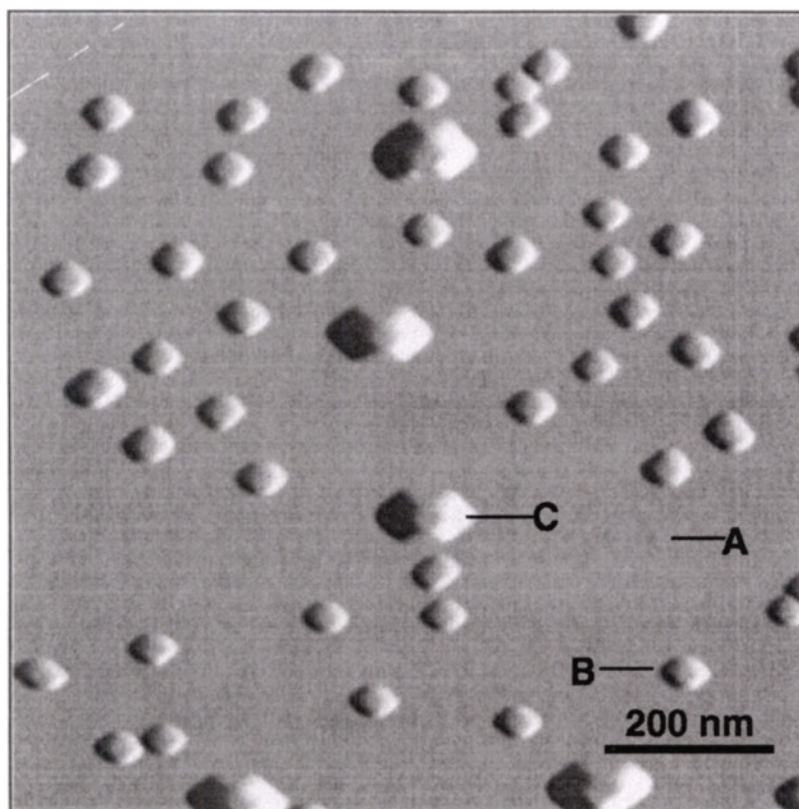


Figure 4.13: High-resolution AFM image of 3.5 ML InP on GaInP (MOVPE 580°C), annealed for 12 s before cooling down. The small flat islands (A) almost disappeared; coherent “fully developed” islands (B) with an average height of 15 nm dominate. A few noncoherent islands (C, average height 41 nm) have also formed [16].

planes (see Table 4.1; here we assume that the gradation of the surface energy values in Table 4.1 is valid for III–Vs in general). There is a trend for smaller fully developed islands to be less elongated and to show larger {110} areas at the expense of {111} [59]. One can imagine that fully developed islands, once they have formed, are rather stable, simply for kinetic reasons: further uptake of material without nucleation atop the low-energy facets can only be realized at the top of the island. There are, however, limits given by the increasing positive energy contributions when edges meet to form a pyramid tip.

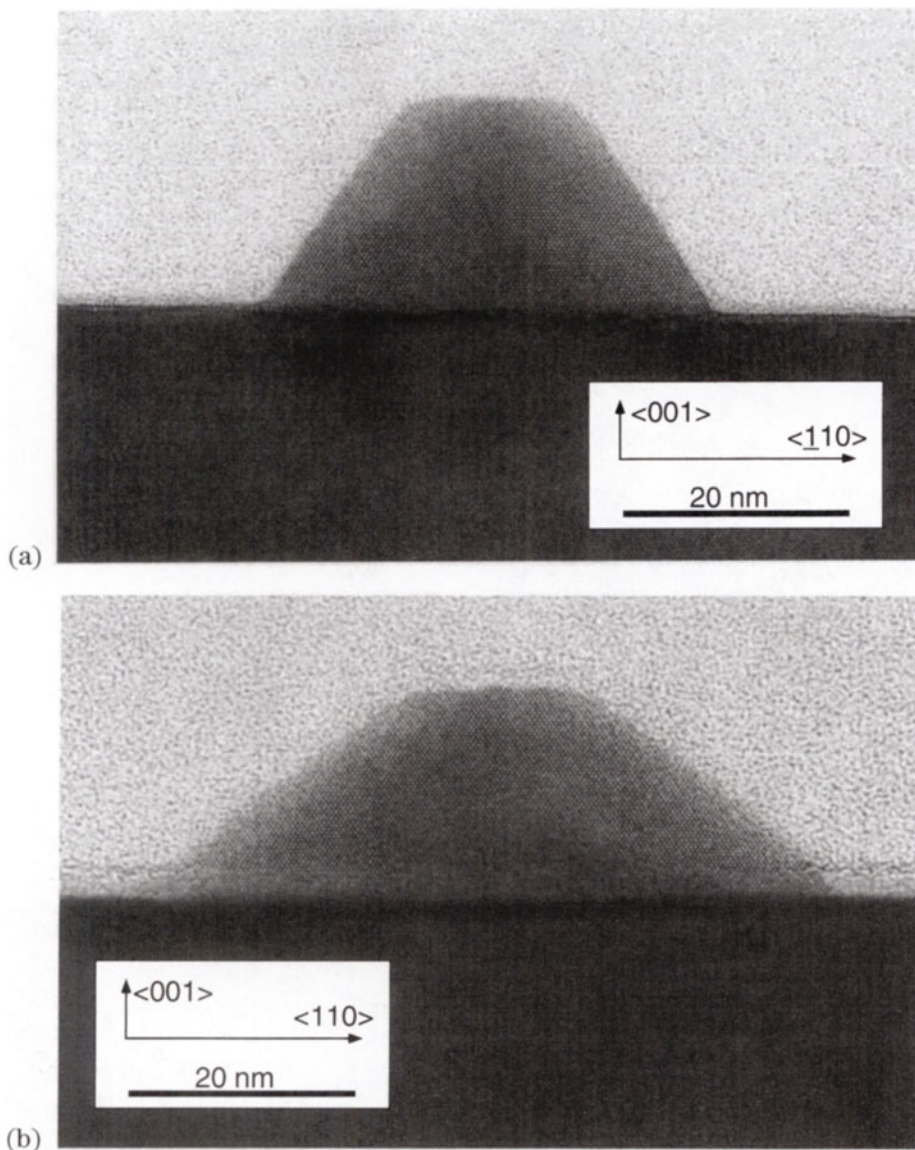


Figure 4.14a: High-resolution TEM of freestanding InP/GaInP islands, e-beam in direction $\langle 110 \rangle$. Note the bending of the $(1\bar{1}1)_B$ lattice planes close to the substrate due to the compressive strain [59].

Figure 4.14b: High-resolution TEM of freestanding InP/GaInP islands, e-beam in direction $\langle 110 \rangle$. The slightly diffuse edge of the island defines the projection of the intersection of two $\{110\}$ facets [59].

4.4.3.2 Manipulations of densities and sizes of 3D islands

A necessary condition for applications of self-assembling is to have control over densities and sizes of quantum dots. As already outlined in Section 4.3, these quantities are not independent from each other. We will show below the experimental results that uphold this conclusion and that were the experimental basis for the general proportionality expressed in Eq. (4.5). For this purpose, InP dots were deposited directly on GaAs(001), the surface of which was P-stabilized by deposition of 2 ML GaP in order to reduce As carryover effects. This direct deposition on GaAs has no effect on densities and sizes of InP in comparison to depositions on GaInP. The advantage is that one can vary more easily the deposition temperature without affecting the quality of the substrate (GaInP shows temperature-dependent ordering effects). First we will consider the changes in surface densities of InP 3D islands as a function of deposition temperature T and deposition rate R . The deposited amount Q was constant at 3.5 ML. The AFM images for two selected temperatures and three selected deposition rates are shown in Fig. 4.15. We see the highest densities (and smallest sizes) in the field of lower temperature and highest deposition rate and the lowest densities (and largest sizes) in the field of higher temperature and lowest deposition rate.

If we consider that the 3D islands are a result of a nucleation process, then the average distance 2λ between coherent islands should show some relation to characteristic length scales of nucleation processes on surfaces, which for deposition/diffusion processes often are written as a power law [60]:

$$\lambda \propto (D/R)^\gamma. \quad (4.7)$$

D is the diffusion constant for surface diffusion, given by $D = (2kT/h)a^2 \exp(-E_D/kT)$, where k is Boltzmann's constant, T is the absolute temperature, h is Planck's constant, a is the lattice parameter, and E_D is the energy barrier for surface diffusion. R is the deposition rate. The exponent γ is determined by the actual process (we will use it here as an empirical parameter).

For λ we use half of the average island separation, $\lambda = 1/(2\rho^{1/2})$ (Fig. 4.2), with ρ as the density of the islands. We should note that λ for our case is not equal but proportional to the surface diffusion length [61]. By holding T constant and plotting $\ln \lambda$ versus $\ln R$, γ was estimated to be approximately 1/2. E_D was then extracted from a plot of $\ln(\lambda/T^\gamma)$ versus $1/T$, for R constant, and estimated to be 1.2 eV. With these values, we obtain an interesting dependence when ρ is expressed as a function of R/D . Since $\gamma = 1/2$ and $\rho = 1/4\lambda^2$, we obtain the direct proportionality $\rho \propto R/D$, i.e., Eq. (4.5). The results are presented in Fig. 4.16.

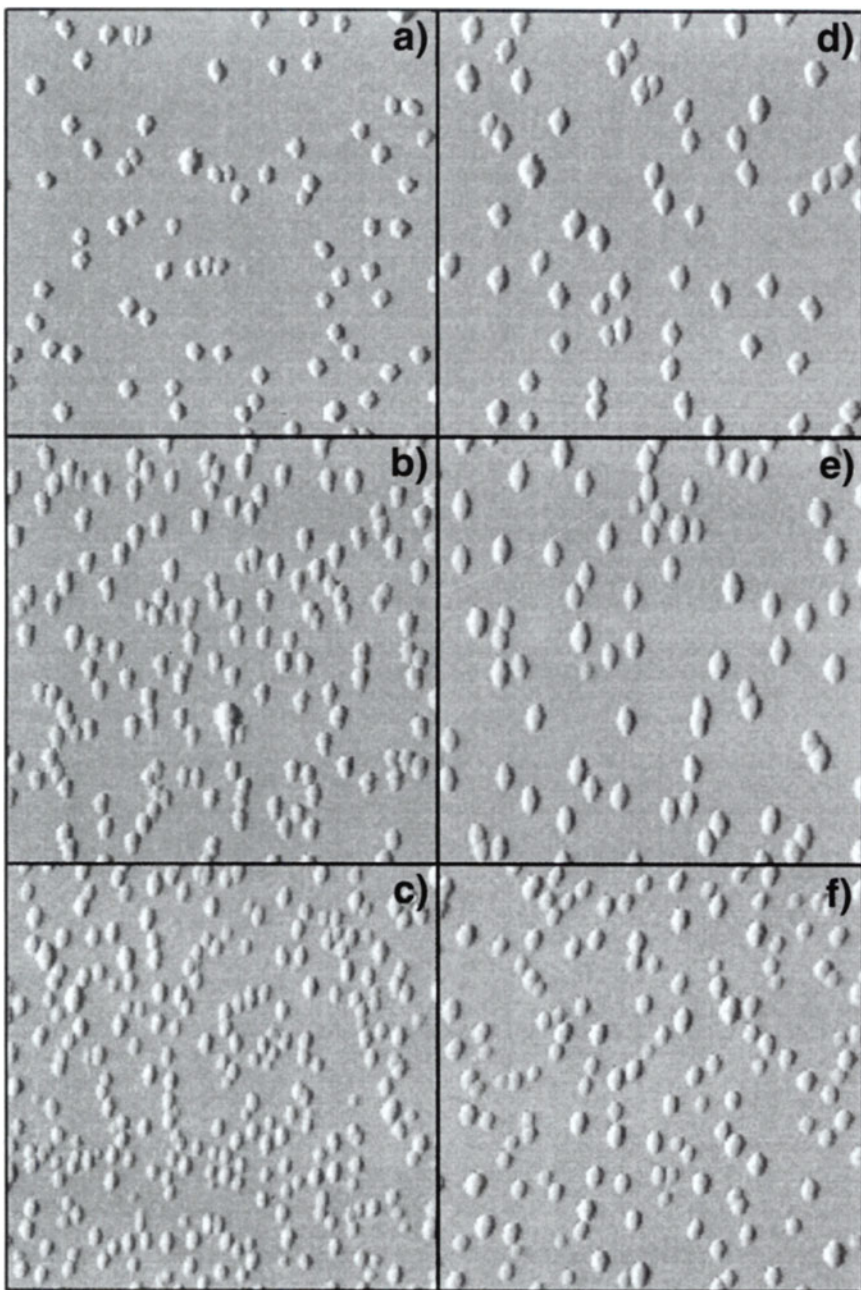


Figure 4.15: AFM images, $2 \times 2 \mu\text{m}^2$, of InP islands grown on GaAs. The temperature was 610°C (a,b, and c) and 640°C (d,e, and f). The deposition rates were 0.5 (a,d), 1.2 (b,e), and 3.5 (c,f) ML/s. The total amount of InP was 3.5ML in all cases. All images are viewed in shadow mode [31].

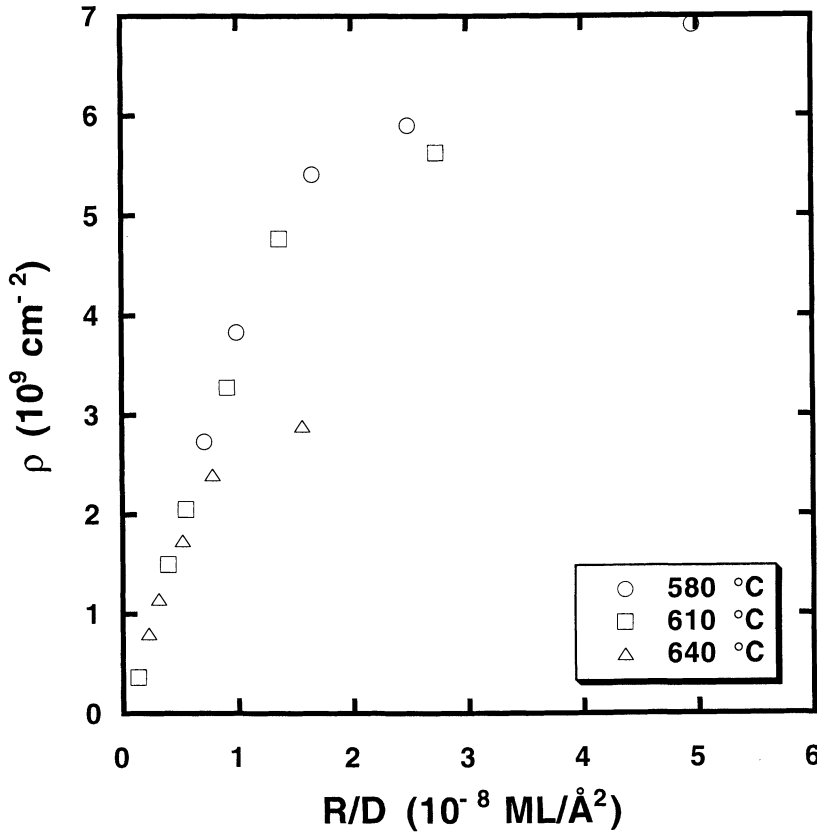


Figure 4.16: The density of InP 3D islands as a function of R/D . Note that for low R , all densities fall together into one line, showing a direct linear proportionality $\rho \propto R/D$ [31].

We see that this proportionality is strongly valid, at least for not-too-high deposition rates. This result is closely related to the results observed for deposition of Ag on Si and W on Si [32,62]. The rate equation theory, developed in [32], predicts for the maximum 3D cluster density a proportionality $\rho_x \propto (R/\nu)^p \exp(E/kT)$. In this formula, ν is the atomic vibration frequency ($\approx 2kT/h$), $p = i/(i + 2.5)$, and $E = (E_i + iE_D)/(i + 2.5)$, with E_i the barrier to form a critical nucleus and E_D the activation barrier for surface diffusion. Note that this proportionality becomes equal to $\rho \propto R/D$ as soon as the critical size i^* to form a 3D nucleus becomes $i^* \gg 1$. In other words, the observed proportionality is nothing else than an expression for the

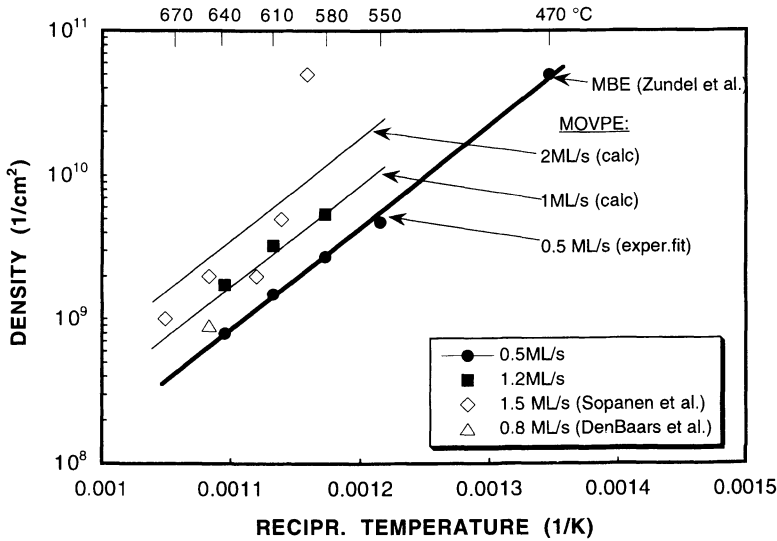


Figure 4.17: A plot of density vs. reciprocal temperature for 3D islands of InP on GaAs(001) or GaInP/GaAs(001). Nominal deposition was 3–3.5 ML of InP. The $R = 0.5 \text{ ML/s}$ results were used to fit the proportionality factor, resulting in $\rho = 7.325 R \exp(E_D/kT)$ and $E_D = 1.2 \text{ eV}$. Using this equation, the lines for 1 and 2 ML/s were calculated. The values from the literature are from DenBaars et al. [40], Sopanen et al. [48], and Zundel et al. [63].

density of stable 3D islands surviving at the surface after the initial strain-induced nucleation process.

At present, we cannot explicitly express the proportionality factor in Eq. (4.5), but we can fit the numerical value by using experimental data. This is done for InP 3D islands, nominal 3.5 ML, deposited with a rate of 0.5 ML/s. The result is shown in Fig. 4.17. An extrapolation of this 0.5 ML/s fit towards lower temperatures covers also the density of InP islands, deposited at 470°C by MBE [63]. And if we now vary the deposition rate, R , we can approximately describe also the found island densities for different deposition temperatures, including values published by other authors.

The sizes of the fully developed coherent islands vary inversely with the densities, as shown in Fig. 4.18. We should mention that instead of the correct volume values we used the heights as representative size values. The reason is that most of the “sizes” are measured by AFM, which gives correct values for the height but commonly too-large values for the lateral

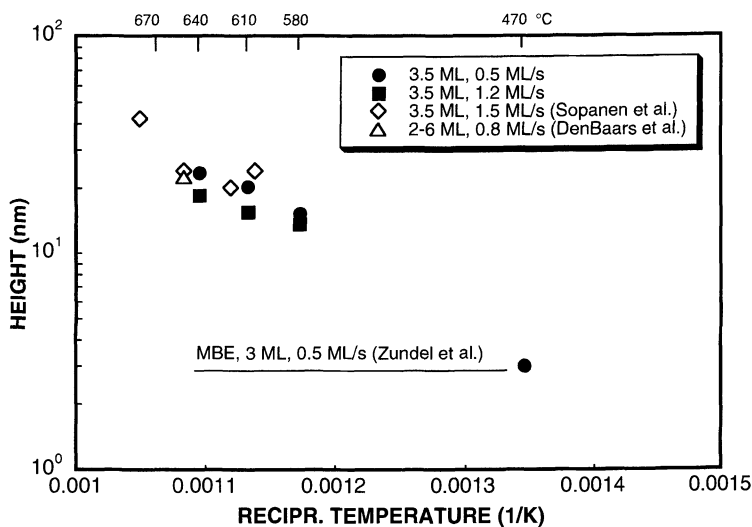


Figure 4.18: A plot of height vs. reciprocal temperature for InP islands (data as in Fig. 4.17).

widths due to tip convolution effects. These results confirm that densities and sizes are related to each other. In order to produce high densities of small quantum dots, one has to use low temperatures and high deposition rates; in the opposite case, one produces low densities and large quantum dots.

There are also ways to manipulate dot sizes independent of dot densities. InP dots can be annealed for a longer time, whereby they lose height while their base diameter remains almost unaffected [64]. In this way also, larger InP “dots” can be reduced in size to heights below 10 nm (Fig. 4.19).

The effect is visible in PL measurements: the typical 1.65 eV peak originating from larger InP dots embedded into GaInP shifts to higher energies after more than 10 minutes annealing at 600°C. It is interesting that during this annealing process the size homogeneity does not degrade. The stable base diameter is most probably due to interdiffusion processes at the highly strained island base. In-diffusion of Ga, for instance, could reduce the strain and simultaneously stabilize the island base by alloying.

4.4.3.3 Size homogeneities of MOVPE-grown InP dots

Using self-assembling for fabrication of quantum dot structures, one has to accept that the result will always be an ensemble of dots with a certain

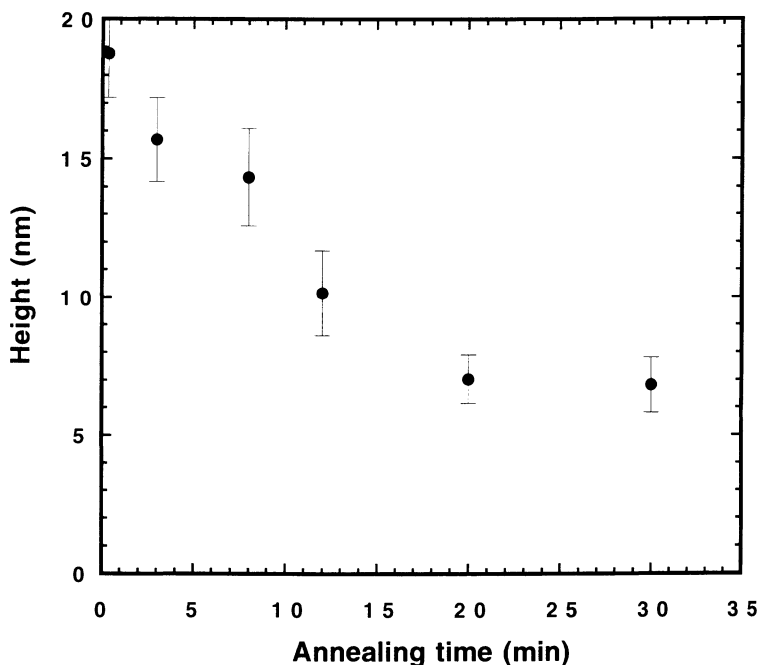


Figure 4.19: Effect of annealing in PH_3/H_2 on the height of coherent InP islands. $T = 600^\circ\text{C}$ [64].

size distribution. However, there are some rules to reduce this disadvantageous effect to a minimum. Besides finding the optimum for the deposited amount, deposition temperature, and deposition rate, one important rule is to use the right timing for the process. Right timing means to initiate the 2D–3D transition in such a way that, in as short as possible a time interval, the necessary density of nucleation sites will be created and the islands can grow. This outcome requires in general uninterrupted and not too slow deposition of the total amount Q . It seems that in this respect MOVPE, with large contributions from vapor transport, is more sensitive than MBE, where surface diffusion dominates as the lateral transport mechanism. We will demonstrate this, using the results of an experiment [65] in which three different samples of self-assembled InP islands were grown, and the deposited amount Q was (A) 2.5 ML, (B) 3.5 ML, and (C) 2.5 + interrupt 12 s + 1 ML. After deposition, the samples were annealed for 12 s under PH_3 at deposition temperature and then cooled down.

Figure 4.20 shows the AFM images and the height histograms for InP islands. Deposition of 2.5 ML of InP results in nearly all coherent SK islands

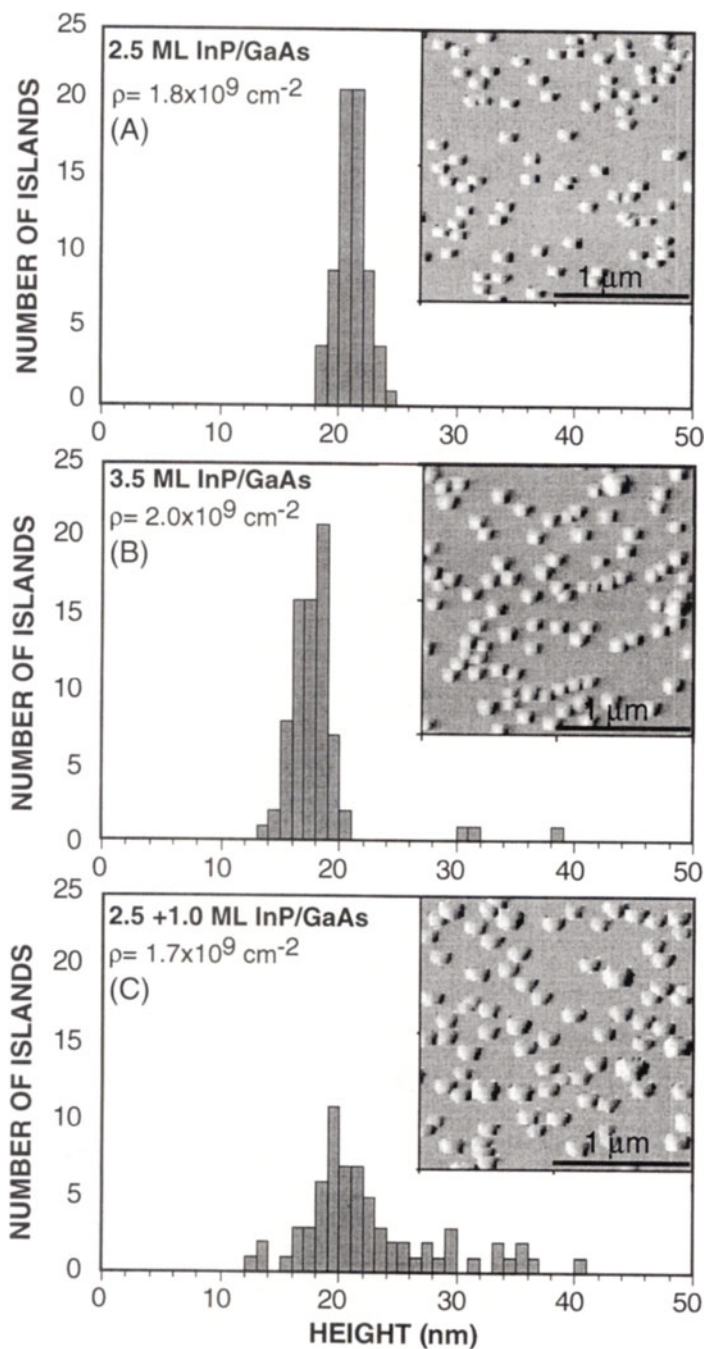


Figure 4.20: AFM images and histograms of height distributions of InP islands grown continuously (A) 2.5MLInP; (B) 3.5MLInP; and discontinuously with a 12s growth interrupt (C) 2.5 + interrupt + 1 MLInP. Deposition temperature: 600°C.

of rather high size homogeneity, with an average height of 21 nm. An increased deposition of 3.5 ML of InP produces coherent islands with a slightly higher surface density and a slightly lower average height of 18 nm. Their smaller size in comparison to sample A is an effect of the slightly higher density and the competition in growth due to the appearance of a few noncoherent islands (islands with heights of 30–32 nm and 38 nm). The size homogeneity of the coherent islands is still high. In contrast, the same amount of 3.5 ML of InP, deposited in portions of 2.5 and 1 ML, with 12 s annealing in between, results in a rather inhomogeneous size distribution (C). This means that the additional 1 ML of InP, deposited in a second portion, was not equally distributed over all the islands formed in the first step from the 2.5 ML deposition. In other words, the size-regulation mechanism, responsible for the high homogeneity of sizes if one deposits the material of 3.5 ML at once, is not furthermore effective if one deposits the same amount of material in portions.

What makes the difference? As long as 3D islands grow by the mechanism of reorganization of material, they will preferentially be provided with material via surface diffusion. Their growth should then be affected by the existence of a strain-induced barrier E_B around coherent islands. For this case (see also Fig. 4.5), the net flux of material towards the islands, j , is given by

$$j = k_{\text{diff}}(\mu_{\text{wl}} - \mu_{\text{isl}}) \quad (4.8)$$

where $\mu_{\text{wl}} - \mu_{\text{isl}}$ is the difference of the chemical potentials of mobile surface atoms at the wetting layer and the island, respectively, and the rate constant $k_{\text{diff}} = k' \exp(-E_B/kT)$. The barrier E_B is size regulating, since it increases with increasing island height and favors the smaller islands to grow faster. If, however, 3D islands grow by direct transport of material from a supersaturated vapor phase, the flux of material towards the island is given by

$$j = k_g(\mu_g - \mu_{\text{isl}}) \quad (4.9)$$

with μ_g the chemical potential of atoms in the vapor and k_g the mass transfer coefficient of vapor-phase diffusion. In the vertical direction above the islands, there is no size-dependent diffusion barrier; k_g is independent of island sizes. The differences in mass transport mainly arise due to differences in the driving force for vapor-phase diffusion, i.e., the difference of the chemical potentials in Eq. (4.9). The variable is the local chemical potential μ at the surface, which decreases in the sequence $\mu_{\text{wl}} > \mu_{\text{small isl}} > \mu_{\text{larger isl}}$ due to the more efficient strain relaxation. Thermodynamically, the larger (but still not fully developed!) islands are the absolute sinks in the chemical surface potential. Consequently, transport of material from

the vapor phase towards already existing islands should preferentially be directed towards the larger islands.

We should also learn from these experiments that any subsequent deposition onto a surface that is already covered by islands will result in bad size homogeneities. This is in fact an experience that one can have in vapor-deposition systems.

4.5 Summary

The aim of this chapter has been to make the reader familiar with MOVPE and its use in fabricating quantum structured semiconductors. We have mainly focused on self-assembling as one of the most promising methods to produce quantum dots in situ. We have demonstrated that by understanding self-assembling as a strain-driven process of nucleation and growth, the dependence of sizes and densities of quantum dots on deposition parameters finds its natural explanation. The main parameters affecting sizes and densities are the deposited amount of the material, Q ; the deposition temperature, T ; and the deposition rate, R . Then, in a first approximation, the density, ρ , of stable three-dimensional islands follows a power-law $\rho \propto (R/D)^\zeta$ (with D as the surface diffusion coefficient and ζ found to be ≈ 1) that characterizes self-assembling primarily as a kinetic process of nucleation and growth. The sizes depend on the deposited amount that has to be distributed over the number of three-dimensional islands. Therefore, we always found inverse behavior of density and sizes. This interdependence of density and sizes to some degree limits the flexibility of the method. The limitation can be overcome by, for instance, annealing procedures, whereby the dots shrink in height.

For dots of InP on GaInP/GaAs, we also discussed some peculiarities—for instance, the high density of smaller and flat three-dimensional islands in the initiation period, which comes out as a consequence of a lower nucleation barrier on top of a highly strained, i.e., supercritically thick, wetting layer. Therefore, after partial strain relaxation due to formation of three-dimensional islands, no further nucleation, but ripening under formation of fully developed and stable coherent islands takes place. For too-low deposited amounts of material, and at lower temperatures, the smaller 3D islands are metastable, and as a consequence, bimodal size distributions can be observed.

There are a few other interesting aspects of self-assembling that unfortunately could not be included in this chapter. These include especially the phenomena of lateral and vertical self-organization. In this field, empirical results exist already. Lateral alignment of dots by growth on patterned sur-

faces, as well as stacking of dots in the vertical direction, has been demonstrated. Most of the work has been done in MBE, with in a few cases also using MOVPE [66] and CBE [67]. Further progress in this field will open the doors for completely new quantum structured semiconductors. We consider this as an interesting task for future work.

ACKNOWLEDGEMENTS

I am grateful to many colleagues and co-workers at Solid State Physics and Inorganic Chemistry at Lund's University, especially to Niclas Carlsson and Jonas Johansson for their contributions to the presented experimental results and the fruitful discussions with them, as well as to Reine Wallenberg, Christina Georgsson, and Anna Carlsson for the TEM data. I also acknowledge the support and permanent interest of L. Samuelson and enlightening discussions with D. Vvedensky and A.-L. Barabási.

This work was performed within the Nanometer Structure Consortium in Lund and was supported by grants from the Swedish National Board for Industrial and Technical Development (NUTEK), the Swedish Natural Science Research Council (NFR), and the Swedish Research Council for Engineering Sciences (TFR).

REFERENCES

- [1] G. B. Stringfellow, *Organometallic Vapor-Phase Epitaxy: Theory and Practice* (Academic Press, Inc., London, 1989).
- [2] R. L. Moon, information given at the 9th International Conference on Metal Organic Vapor Phase Epitaxy 1998, La Jolla, CA.
- [3] K. Ploog, *Surface Interface Anal.* **12**, 279 (1988).
- [4] L. Samuelson and W. Seifert, in *Handbook of Crystal Growth*, Vol. 3, edited by D. T. J. Hull (Elsevier, Amsterdam, 1994), pp. 745–783.
- [5] P. H. Fuoss, D. W. Kisker, G. B. Stephenson, and S. Brennan, *Mater. Sci. Eng. B* **30**, 99 (1995).
- [6] W. Richter and J. T. Zettler, *Appl. Surface Sci.* **100/101**, 465 (1996).
- [7] K. Ploska, J.-T. Zettler, W. Richter, J. Jönsson, F. Reinhardt, J. Rumberg, M. Pristovsek, M. Zorn, D. Westwood, R. H. Williams, *J. Crystal Growth* **145**, 44 (1994).
- [8] M. Zorn, T. Trepk, J.-T. Zettler, B. Junno, C. Meyne, K. Knorr, T. Wethkamp, M. Klein, M. Miller, W. Richter, and L. Samuelson, *Appl. Phys. A* **65**, 333 (1997).
- [9] K. F. Jensen, in *Handbook of Crystal Growth*, Vol. 3, edited by D. T. J. Hurle (Elsevier, Amsterdam, 1994), pp. 541–599.

- [10] L. Goldstein, F. Glas, J. Y. Marzin, M. N. Charasse, and G. L. Roux, *Appl. Phys. Lett.* **47**, 1099 (1985).
- [11] D. J. Eaglesham and M. Cerullo, *Phys. Rev. Lett.* **64**, 1943 (1990).
- [12] P. M. Petroff and S. P. DenBaars, *Superlat. Microstruct.* **15**, 15 (1994).
- [13] J.-M. Gerard, in *Confined Electrons and Photons*, Series B340, edited by E. Burstein and C. Weisbuch (Plenum Press, New York, 1995), p. 357.
- [14] A. Madhukar, P. Chen, Q. Xie, A. Konkar, T. R. Ramachandran, N. P. Kobayashi, and R. Viswanathan, in *Low Dimensional Structures Prepared by Epitaxial Growth or Regrowth on Patterned Substrates*, edited by K. Eberl et al. (Kluwer Academic Publishers, Boston, 1995).
- [15] R. Nötzel, *Semicond. Sci. Technol.* **11**, 1365 (1996).
- [16] W. Seifert, N. Carlsson, M. Miller, M.-E. Pistol, L. Samuelson, and L. R. Wallenberg, *Prog. Crystal Growth Characterist.* **33**, 423 (1996).
- [17] I. N. Stranski and L. Krastanow, *Sitz. Ber. Akad. Wiss., Math.-naturwiss. Kl. Abt. IIb* **146**, 797 (1938).
- [18] M. Volmer and A. Weber, *Z. Physik. Chem.* **119**, 277 (1926).
- [19] M. Pinczolits, G. Springholz, and G. Bauer, *Appl. Phys. Lett.* **73**, 250 (1998).
- [20] C. Ratsch and A. Zangwill, *Surf. Sci.* **293**, 123 (1993).
- [21] V. A. Shchukin, N. N. Ledentsov, P. S. Kop'ev, and D. Bimberg, *Phys. Rev. Lett.* **75**, 2968 (1995).
- [22] A.-L. Barabasi, *Appl. Phys. Lett.* **70**, 2565 (1997).
- [23] I. Daruka and A. L. Barabasi, *Appl. Phys. Lett.* **72**, 2102 (1998).
- [24] B. G. Orr, D. Kessler, C. W. Snyder, and L. Sander, *Europhys. Lett.* **19**, 33 (1992).
- [25] D. E. Jesson, K. M. Chen, and S. J. Pennycook, *MRS Bull.* **21**, 31 (1996).
- [26] D. E. Jesson, G. Chen, K. M. Chen, and S. J. Pennycook, *Phys. Rev. Lett.* **80**, 5156 (1998).
- [27] H. T. Dobbs, D. D. Vvedensky, A. Zangwill, J. Johansson, N. Carlsson, and W. Seifert, *Phys. Rev. Lett.* **79**, 897 (1997).
- [28] G. Abstreiter, P. Schittenhelm, C. Engel, E. Silveira, A. Zrenner, D. Meertens, and W. Jäger, *Semicond. Sci. Technol.* **11**, 1521 (1996).
- [29] D. I. Westwood, Z. Sobierski, C. C. Matthai, E. Steimetz, T. Zettler, and W. Richter, *J. Vac. Sci. Technol. B* **16**, 2358 (1998).
- [30] M. Zinke-Allmang, L. C. Feldman, and M. H. Grabow, *Surf. Sci. Rep.* **16**, 377–463 (1992).
- [31] J. Johansson, N. Carlsson, and W. Seifert, *Physica E* **2**, 667 (1997).
- [32] J. A. Venables, *Surface Sci.* **299/300**, 798 (1994).
- [33] D. J. Srolovitz, *Acta Metall.* **37**, 621 (1989).

- [34] A. Carlsson, L. P. Wallenberg, C. Persson, and W. Seifert, *Surface Science* **406**, 48 (1998).
- [35] Y. Chen and J. Washburn, *Phys. Rev. Lett.* **77**, 4046 (1996).
- [36] E. Pehlke, N. Moll, A. Kley, and M. Scheffler, *Appl. Phys. A* **65**, 525 (1997).
- [37] S. Ruvimov, P. Werner, K. Scheerschmidt, U. Gösele, J. Heydenreich, U. Richter, N. N. Ledentsov, M. Grundmann, D. Bimberg, V. M. Ustinov, A. Y. Egorov, P. S. Kop'ev, and Z. I. Alferov, *Phys. Rev. B* **51**, 14766 (1995).
- [38] G. Medeiros-Ribeiro, A. M. Bratkovski, T. I. Kamins, D. A. A. Ohlberg, and R. S. Williams, *Science* **279**, 353 (1998).
- [39] F. M. Ross, J. Tersoff, and R. M. Tromp, *Phys. Rev. Lett.* **80**, 984 (1998).
- [40] S. P. DenBaars, C. M. Reaves, V. Bressler-Hill, S. Varma, H. Weinberg, and P. M. Petroff, *J. Crystal Growth* **145**, 721 (1994).
- [41] A. Ponchet, A. L. Corre, H. L'Haridon, B. Lambert, and S. Salaün, *Appl. Phys. Lett.* **67**, 1850 (1995).
- [42] J. Oshinowo, M. Nishioka, S. Ishida, and Y. Arakawa, *Appl. Phys. Lett.* **65**, 1421 (1994).
- [43] F. Heinrichsdorff, A. Krost, M. Grundmann, D. Bimberg, A. Kosogov, and P. Werner, *Appl. Phys. Lett.* **68**, 3284 (1996).
- [44] H. Marchand, P. Desjardins, S. Guillon, J.-E. Paultre, Z. Bougrioua, R. Y.-F. Yip, and R. A. Masut, *Appl. Phys. Lett.* **71**, 527 (1997).
- [45] M. Taskinen, M. Sopanen, H. Lipsanen, J. Tulkki, T. Tuomi, and J. Ahopelto, *Surface Sci.* **376**, 60 (1997).
- [46] N. Carlsson, T. Junno, L. Montelius, M.-E. Pistol, L. Samuelson, and W. Seifert, *J. Crystal Growth* **191**, 347 (1998).
- [47] N. Carlsson, W. Seifert, A. Petersson, P. Castrillo, M.-E. Pistol, and L. Samuelson, *Appl. Phys. Lett.* **65**, 3093 (1994).
- [48] M. Sopanen, H. Lipsanen, and J. Ahopelto, *Appl. Phys. Lett.* **67**, 3768 (1995).
- [49] A. Kurtenbach, K. Eberl, and T. Shitara, *Appl. Phys. Lett.* **66**, 361 (1995).
- [50] F. Hatami, N. N. Ledentsov, M. Grundmann, J. Böhrer, F. Heinrichsdorff, M. Beer, D. Bimberg, S. Ruvimov, P. Werner, U. Gösele, J. Heydenreich, U. Richter, S. V. Ivanov, B. Y. Meltser, P. S. Kop'ev, and Z. I. Alferov, *Appl. Phys. Lett.* **67**, 656 (1995).
- [51] B. M. Kinder and E. M. Goldys, *Appl. Phys. Lett.* **73**, 1233 (1998).
- [52] S. Nakamura and G. Fasol, *The Blue Laser Diode* (Springer Verlag, Berlin Heidelberg New York, 1997).
- [53] S. Keller, B. P. Keller, M. S. Minsky, J. E. Bowers, U. K. Mishra, S. P. DenBaars, and W. Seifert, *J. Crystal Growth* **189/190**, 29 (1998).
- [54] S. Tanaka, S. Iwai, and Y. Aoyagi, *Appl. Phys. Lett.* **69**, 4096 (1996).

- [55] F. Widmann, B. Daudin, G. Feuillet, Y. Samson, J. E. Rouvière, and N. Pelekanos, *J. Appl. Phys.* **83**, 7618 (1998).
- [56] H. Yamaguchi, M. R. Fahy, and B. A. Joyce, *Appl. Phys. Lett.* **69**, 776 (1996).
- [57] N. Carlsson, K. Georgsson, L. Montelius, L. Samuelson, W. Seifert, and R. Wallenberg, *J. Crystal Growth* **156**, 23 (1995).
- [58] M. Berti, A. V. Drigo, A. Giuliani, M. Mazzer, A. Camporese, G. Rossetto, and G. Torzo, *J. Appl. Phys.* **80**, 1931 (1996).
- [59] K. Georgsson, N. Carlsson, L. Samuelson, W. Seifert, and L. R. Wallenberg, *Appl. Phys. Lett.* **67**, 2981 (1995).
- [60] A. L. Barabasi and H. E. Stanley, *Fractal Concepts in Surface Growth* (University Press, Cambridge, 1995).
- [61] M. Hata, T. Isu, A. Watanabe, Y. Kajikawa, and Y. Katayama, *J. Crystal Growth* **114**, 203 (1991).
- [62] J. A. Venables, G. D. T. Spiller, and M. Hanbücken, *Rep. Prog. Phys.* **47**, 399 (1984).
- [63] M. K. Zundel, P. Specht, K. Eberl, N. Y. Jin-Phillipp, and F. Phillipp, *Appl. Phys. Lett.* **71**, 2972 (1997).
- [64] J. Johansson, W. Seifert, V. Zwiller, T. Junno, and L. Samuelson, *Appl. Surface Science* **134**, 47 (1998).
- [65] W. Seifert, J. Johansson, N. Carlsson, A. Gustafsson, and J.-O. Malm, *J. Crystal Growth*, **197**, 19 (1999).
- [66] W. Seifert, N. Carlsson, A. Petersson, L.-E. Wernersson, and L. Samuelson, *Appl. Phys. Lett.* **68**, 1684 (1996).
- [67] S. Jeppesen, M. Miller, D. Hessman, B. Kowalski, I. Maximov, and L. Samuelson, *Appl. Phys. Lett.* **68**, 2228 (1996).

Chapter 5

Growth, characterization, and applications of self-assembled InGaAs quantum dots

Richard P. Mirin¹ and Arthur C. Gossard²

¹Optoelectronic Manufacturing Group, National Institute of Standards and Technology, 325 Broadway, Boulder, CO 80303; ²Materials Dept., University of California, Santa Barbara, Santa Barbara, CA 93101

5.1 Introduction

This chapter focuses on the growth, characterization, and applications of self-assembled $\text{In}_x\text{Ga}_{1-x}\text{As}$ quantum dots (QDs) on GaAs substrates. These QDs are formed by the so-called Stranski–Krastanow (SK) transition of the highly lattice-mismatched $\text{In}_x\text{Ga}_{1-x}\text{As}$ film. The $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ system is the prototypical system in III–V growth for self-assembly of QDs, and consequently it is the most heavily studied. The lattice constant of InAs is about 7.2% larger than that of GaAs; this is the largest mismatch among AlGaInAsP alloys to GaAs or InP substrates. From a practical standpoint, the large mismatch to a common substrate means that this material system is readily accessible to most laboratories with epitaxial growth capabilities such as molecular beam epitaxy (MBE) and organometallic

vapor-phase epitaxy (OMVPE). Note that since the formation of these self-assembled QDs is driven by a reduction of the strain energy, many other III–V semiconductors can conceivably be used to form QDs on a variety of different substrates. Many of these have been demonstrated, such as GaSb/GaAs [1], InP/Ga_{0.5}In_{0.5}P [2], and InAs/InP [3], but we will not discuss them further.

In_xGa_{1-x}As with an In mole fraction of greater than about 0.2 [4] changes from layer-by-layer (two-dimensional) growth to island (three-dimensional) growth. These experimental results are consistent with theoretical predictions [5] about the competing mechanisms for strain relaxation. These predictions suggest that at low misfit values, the surface relaxes by forming dislocations, before any surface roughening occurs. At large misfit values, the surface roughens because of a net decrease in the surface energy caused by a partial relaxation of the strain at the edges of the islands. Furthermore, the island edges provide high stress points at which dislocations can later nucleate [6]. Experiments verify this phenomenon as well [7].

The exact In mole fraction at which the 2D-to-3D transition occurs depends strongly on growth conditions. This transition from two-dimensional growth to three-dimensional growth can be easily seen during MBE growth using the reflection high-energy electron diffraction (RHEED) pattern, which has a streaky structure when the surface is flat (two-dimensional) and a spotty structure when the islands appear (three-dimensional). The In_xGa_{1-x}As islands (quantum dots) that form are quantum sized (smaller than the de Broglie wavelength of the electrons and holes) and coherently strained (no dislocations) to the underlying GaAs substrate. The two-dimensional layer of In_xGa_{1-x}As that is grown immediately before the island formation is known as the Stranski–Krastanow wetting layer. Once the layer of In_xGa_{1-x}As reaches some thickness (the exact thickness depends on composition, substrate temperature, and As flux), the film will undergo a transition from flat (two-dimensional) to islanded (three-dimensional). This increases the surface area of the film, but allows for partial relief of the strain energy by bowing at the edges of the islands. Continued deposition beyond the 2D-to-3D transition is possible, with a rapid increase in island density, followed by saturation of the island density and a subsequent increase in island size [8]. Beyond a given island size that depends strongly on several factors (In mole fraction, growth temperature, As flux, and species), dislocations form and the dislocated islands expand rapidly, consuming smaller islands. Finally, a heavily dislocated but continuous film whose thickness is spatially modulated forms.

QDs formed from III–V semiconductors are interesting to physicists and engineers because the QDs provide three-dimensional confinement of

Table 5.1: Electrical and optical characteristics of various heterostructures.

Name	Dimension	$\rho(k)$	$\rho(E)$
Bulk	3	$\frac{k^2}{\pi^2}$	$\frac{\sqrt{E}}{2\pi^2} \left[\frac{2m^*}{\hbar^2} \right]^{3/2}$
Quantum well	2	$\frac{k}{\pi L_z}$	$\frac{m^*}{\pi \hbar^2 L_z}$
Quantum wire	1	$\frac{2}{\pi L_x L_y}$	$\frac{2}{\pi L_x L_y} \frac{1}{\sqrt{E}} \left[\frac{2m^*}{\hbar^2} \right]^{1/2}$
Quantum dot	0	$\frac{2}{L_x L_y L_z} \delta(k)$	Single dot: $2\delta(E)$ Array: $D\rho_A (2\pi(\Delta E)^2)^{-1/4} \exp \left[-\left(\frac{E - E_Q}{2\Delta E} \right)^2 \right]$

D : Degeneracy of the state; ρ_A : areal density of quantum dots; ΔE : energy spread of the quantum dots; E_Q : peak energy.

electrons and holes. QDs have dramatically different electrical and optical properties than bulk, quantum-well (QW) (one-dimensional confinement), or quantum-wire (two-dimensional confinement) heterostructures. The electrical and optical characteristics of all these heterostructures are determined by the density of states (DOS) associated with the given dimensionality (see Table 5.1). As shown in Fig. 5.1, the bulk DOS is proportional to energy to the one-half power. A QW DOS is a staircase in energy, with another upward step taken at each subband. The quantum wire DOS has a peak at each subband, and then decreases with energy to the negative one-half power.

Unlike these other heterostructures, a QD has a discrete DOS, rather than a continuous (in energy) DOS. In this way, a QD is very similar to an atom. At each allowed eigenvalue of the Hamiltonian that describes the potential of the quantum dot, there is a delta function in energy that corresponds to an allowed state in the dot, as illustrated in Fig. 5.1. The amplitude of this delta function is equal to the degeneracy of the given state. For an array of quantum dots, such as will be discussed here, the amplitude of the delta function must be multiplied by the areal density of quantum dots. This allows us to compare the DOS for all these heterostructures.

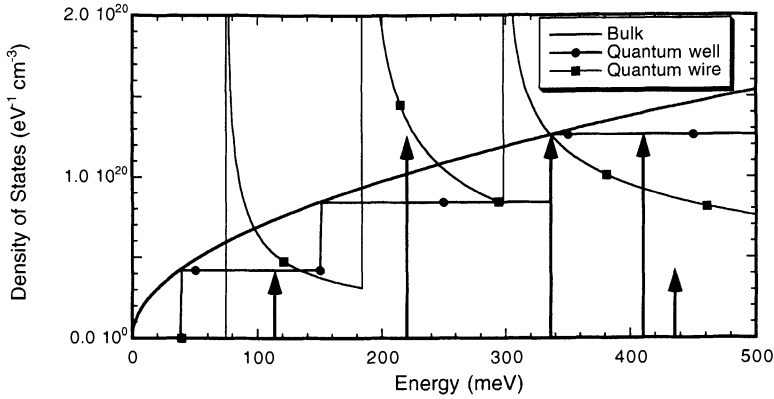


Figure 5.1: Density of states for bulk (three-dimensional), quantum-well (two-dimensional), quantum-wire (one-dimensional), and quantum-dot (zero-dimensional) structures. See table 5.1. An effective mass (m^*) of $0.1 m_e$ and a confining length of 10 nm in each dimension is assumed. The amplitude of the delta functions for the various states of the quantum dots correspond to the areal density of an arrays of dots multiplied by the degeneracy of the given state.

The discrete DOS of an array of QDs led to a seminal paper by Arakawa and Sakaki in 1982 [9] that predicts improved characteristic temperatures in semiconductor lasers with QD active regions. This prediction, along with predictions of decreased threshold currents and increased differential gain [10], has led to a strong interest in techniques to form uniform arrays of quantum dots for optoelectronics applications. However, only with the recent advances in self-assembled quantum dot growth, which will be described hereafter, have QD lasers been fabricated that have demonstrated better performance than QW lasers, such as higher characteristic temperature and larger differential gain.

In the following sections, we will discuss the important experimental and theoretical results that have led to a strong interest in self-assembled quantum dots. We will first discuss the growth and structural characterization of QDs. Then we will proceed to describe the optical and electrical characterization of the QDs and to show how these properties can be correlated with the morphological characterization. Applications of QDs, with an emphasis on QD lasers, will be discussed next. Finally, some current challenges and future directions will be examined.

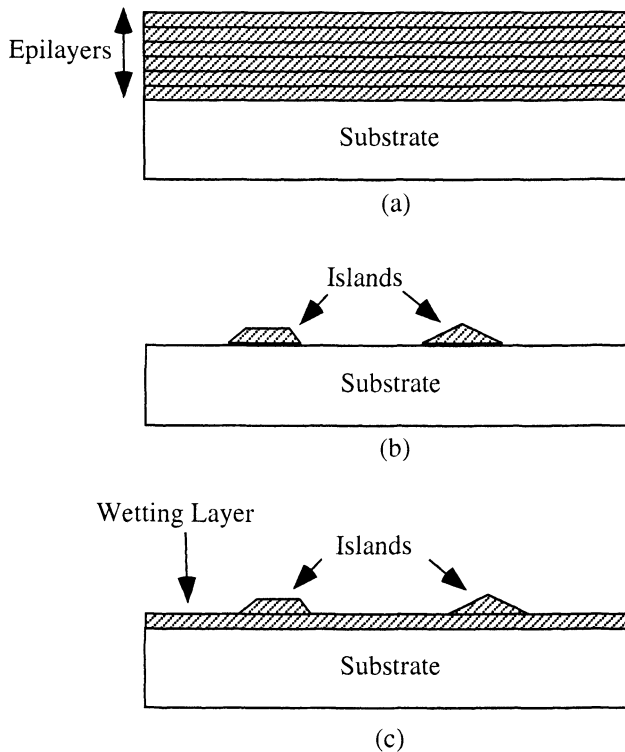


Figure 5.2: (a) Frank-van der Merwe (layer-by-layer) growth; (b) Stranski-Krastanow (layer-plus-island) growth; (c) Volmer-Weber (island) growth.

5.2 Crystal growth and structural characterization

5.2.1 Background and history

Crystal growth can occur by three different methods (Fig. 5.2): (1) layer-by-layer growth (Frank-van der Merwe), (2) layer-plus-island growth (Stranski-Krastanow), or (3) island growth (Volmer-Weber). The particular growth method that one material exhibits when deposited on another material is determined by the surface energies and the kinetics. For III-V semiconductors grown by MBE or OMVPE, the crystal grower almost

always prefers layer-by-layer growth because this gives smooth interfaces and uniform quantum wells. However, not all III–V semiconductors can be grown heteroepitaxially in a layer-by-layer fashion. The lattice constant a_0 of the most common III–V compounds (we limit ourselves to compounds with Group III elements Al, Ga, and In, and Group V elements As, P, and Sb) varies from about 5.45 Å for GaP to 6.48 Å for InSb. Therefore, layers that are strained due to lattice mismatch can be grown. The lattice mismatch between GaAs ($a_0 = 5.65$ Å) and InAs ($a_0 = 6.05$ Å) is about 7.2%. Because of this large mismatch, only about the first two monolayers of InAs that are deposited on GaAs will form planar layers. Beyond this thickness, the InAs forms islands on the surface.

The work on QDs has proceeded only due to the large amount of work that went into characterizing the early stages of growth (≤ 10 monolayers) of InAs and $\text{In}_x\text{Ga}_{1-x}\text{As}$ on GaAs. This initial work was generally focused on morphological considerations, although some photoluminescence (PL) results were reported early on and were associated with the formation of InAs clusters [11]. Much of this early work was focused on either extending the range of two-dimensional growth or investigating various models that predict the formation of dislocations rather than islanding, something that was experimentally shown to be incorrect.

Schaffer et al. [12] were the first to study the nucleation process of InAs on GaAs. These authors were interested in improving the properties of bulk films of InAs grown by MBE on (100) GaAs. They observed that the RHEED pattern developed spots after two monolayers of InAs were deposited under As-rich conditions. However, since these authors were interested in obtaining smooth, thick films of InAs on GaAs, they did not examine the properties of the islands that formed at the early stages of growth.

In 1985, Goldstein et al. [11] studied the properties of the InAs islands (clusters) using scanning transmission electron microscopy (STEM), X-ray diffraction, and PL at 77 K. They also observed the change in the RHEED pattern to a spotty pattern with the deposition of more than two monolayers of InAs. They measured the size of the InAs islands with STEM and measured PL at 77 K that was associated with those clusters. Furthermore, they observed vertical alignment of the islands when a superlattice of InAs/GaAs was grown. Finally, they suggested that these clusters “are thus proved to be of interest to study low-dimensional (<2) objects showing good optical properties.” Glas et al. [13] used STEM to investigate the islands that form in the early stages of InAs growth. Figure 5.3 shows a plan view STEM image of a (nominally) two-monolayer-thick InAs film. The authors estimate the island density at around 10^{11} cm^{-2} . The islands are clearly aligned along $\langle 100 \rangle$ directions.

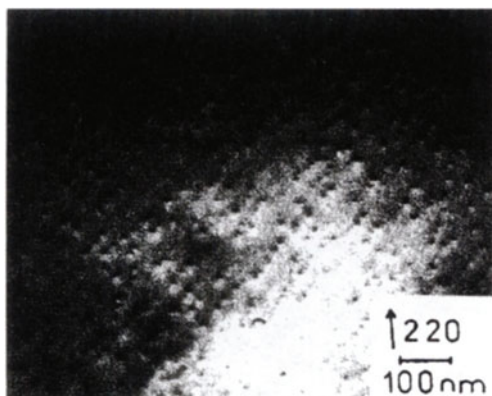


Figure 5.3: Plan-view TEM image of a nominally two-monolayer-thick InAs film on (100) GaAs, grown by MBE. The islands can clearly be seen on the image, and they are aligned along $\langle 100 \rangle$ directions [13].

Guha et al. [7] used high-resolution transmission electron microscopy (HR-TEM) to investigate the early stages of growth of $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ on GaAs (100). They grew seven monolayers of $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ at a substrate temperature of 520°C using MBE. Dark field (400) TEMs show that the substrate lying underneath the islands is deformed due to the strain induced by the islands. HR-TEMs show that these islands are defect free up to about 8.0 nm in height and 25.0 nm in width (Fig. 5.4). Beyond these dimensions, stacking faults can be seen toward the edges of the larger islands, consistent with theoretical predictions [6]. This may put a limit on the size of the islands that can be grown coherently. Note, however, that Guha et al. give no indication in their paper about the change of the RHEED pattern from streaky to spotty, so we have no way of knowing how far beyond the 2D-to-3D transition these islands are measured. Nonetheless, this is direct evidence that it is possible to obtain defect-free islands.

Snyder et al. [14] used RHEED, scanning tunneling microscopy (STM), and TEM to investigate the MBE growth of $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ on GaAs (100). At a growth temperature of 520°C , they observe the 2D-to-3D transition in the RHEED pattern after deposition of about five monolayers of $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$. An STM image of the surface after deposition of 12 monolayers of $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ shows that the morphology is islanded, with coalescence of islands occurring (Fig. 5.5). Thus, it is extremely difficult to discern whether defects have been generated in these islands already (note that almost twice as much material has been deposited than at the point at which Guha et al. [7] saw the onset of defect formation).

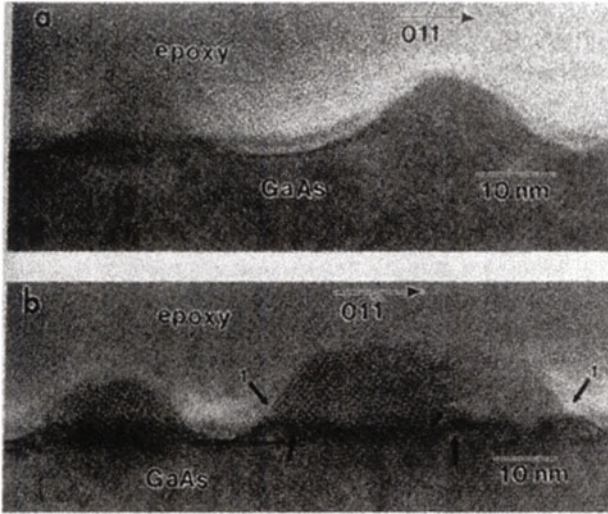


Figure 5.4: High-resolution TEM images of (a) two coherent islands and (b) a coherent island (left) and a defective island (right). The arrows point to defects [7].

5.2.2 Single layers of quantum dots on (100) GaAs

Goldstein et al. [11] were the first to associate PL with the InAs islands. They suggested that islanded growth might be a technique for making reduced-dimensional structures. Further suggestions that the islands could be used as QDs came in 1991 from Tabuchi et al. [15], who used TEM and PL to show that these clusters (or islands) could be well controlled and could emit light at 77 K. These authors observed no dislocations when two monolayers of InAs were deposited. The PL emission wavelength increased, and the full width at half maximum (FWHM) increased, as the film went from two-dimensional to three-dimensional, as was reported by Goldstein et al. [11]. Despite this early work that indicated that InAs grown on GaAs could form nanoscale clusters that emitted light, interest in using this technique to form arrays of quantum dots did not appear until Leonard et al. [16] using plan-view TEM, and Moison et al. [17], using atomic force microscopy (AFM), demonstrated that *uniformly sized* arrays of quantum dots can be grown.

The transition from a flat surface, without QDs, to a surface that contains QDs is very abrupt. Moison et al. [17] demonstrated that the RHEED transition from streaky to spotty occurs after 1.75 ± 0.1 monolayers of InAs

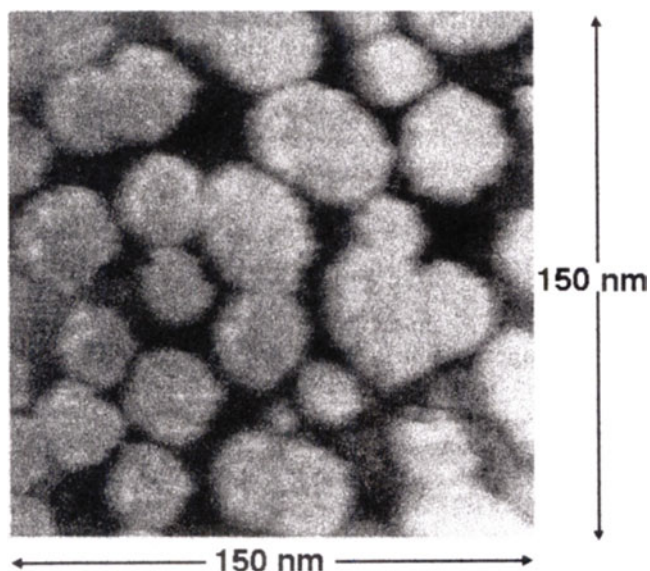


Figure 5.5: A $150\text{ nm} \times 150\text{ nm}$ STM image after deposition of 12 monolayers of $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ at 520°C . The height range is approximately 8 nm . The RHEED pattern changes from streaky to spotty after about 5 monolayers of $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ [14].

have been deposited at a substrate temperature of 500°C . AFM images taken after deposition of 1.7 monolayers of InAs show a flat surface, with atomic steps visible on the surface. Between 1.75 monolayers and about 3.0 monolayers, quantum-sized islands are observed (Fig. 5.6). The islands increase in size (height and base diameter) as more InAs is deposited. However, beyond about three monolayers of InAs, coalescence occurs and the island density decreases dramatically. Histograms taken at an InAs coverage of 2.3 monolayers show a dot height of $3.9\text{ nm} \pm 0.75\text{ nm}$ and a half-base diameter of $14.8\text{ nm} \pm 1.3\text{ nm}$. These dimensions are appropriate for lateral confinement of electrons and holes and can thus be called quantum dots.

Leonard et al. [16] showed that the RHEED pattern during growth of $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ transformed from streaky to spotty after approximately four monolayers had been deposited. Plan-view TEM images showed an approximate island diameter of $27.8\text{ nm} \pm 2.8\text{ nm}$, and no dislocations were observed. Low-temperature PL measurements show that the dots give a broad peak ($\text{FWHM} \approx 60\text{ meV}$) centered at about 1.2 eV . The broad peak can be accounted for by the size distribution of the islands. The integrated

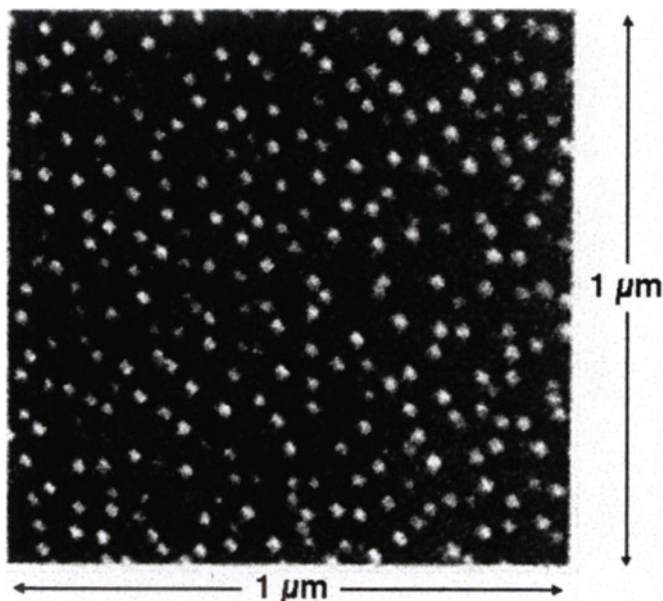


Figure 5.6: A $1\,\mu\text{m} \times 1\,\mu\text{m}$ AFM image after deposition of 2.3 monolayers of InAs at 500°C [17].

PL intensity from the QDs is comparable to that of reference QW, which indicates the high quality of the QD material.

Most of the early work done on $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ QDs has used MBE rather than OMVPE. There are some early reports of OMVPE growth of SK QDs [18], but MBE has the big advantage of using the in situ RHEED transition as a measure of when the QDs form. Steimetz et al. [19] have used reflectance anisotropy spectroscopy (RAS) and spectroscopic ellipsometry (SE) during OMVPE growth to monitor the formation of islands during InAs growth. By combining these in situ techniques with ex situ scanning electron microscopy (SEM), they found that large cluster formation (rather than small island formation) was favored when (1) the temperature was increased, (2) the deposition rate was decreased, and (3) the AsH_3 partial pressure was increased.

Both OMVPE and MBE are extremely sensitive to the total amount of InAs that is deposited. Leonard et al. [8], using MBE, found that the density of coherent islands changed from $8 \times 10^6\text{ cm}^{-2}$ to $2.5 \times 10^8\text{ cm}^{-2}$ with an increase of only about 0.02 monolayers of InAs. Incoherent islands are not observed until about 2.5 monolayers of InAs have been deposited, which is about 20%–30% more than the 2D-to-3D transition. OMVPE growth has similar differences in the density of coherent islands for small changes in

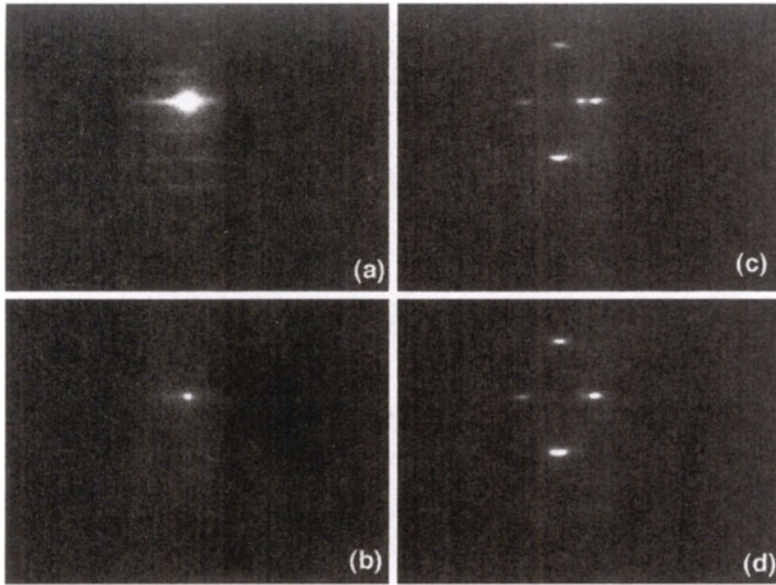


Figure 5.7: Typical RHEED pattern with the electron beam on the [011] azimuth for deposition of (a) 0 monolayers, (b) 1.8 monolayers, (c) 2.0 monolayers, and (d) 2.7 monolayers of InAs.

InAs thickness, but the appearance of incoherent islands occurs much sooner. Gieger et al. [20] grew InAs on (100) GaAs using low-pressure OMVPE. They found that a difference in InAs thickness of only about 0.1 monolayers can mean the difference between having an areal density of about 10^9 cm^{-2} coherent islands and having 10^{10} cm^{-2} coherent islands along with 10^9 cm^{-2} incoherent islands. Heinrichsdorff et al. [21] also investigated the effects of OMVPE growth conditions (growth temperature, V/III ratio, composition, and growth interruption time) on the formation and properties of InGaAs QDs. They found that small changes in In mole fraction (from 71% to 78%) could lead to the formation of large, incoherent clusters. Also, smooth, flat islands with lateral dimensions of 100–200 nm can be found between well-defined QDs under certain growth conditions. Thus, careful control over all growth parameters must be maintained in order to successfully grow QDs with OMVPE.

The exact shape of the QDs has been controversial. It is important to know as much as possible about the shape of these QDs so that modeling of the band structure can proceed. RHEED, TEM, and AFM have all been used to infer information about the shape of the QDs. An example of the typical RHEED pattern at various levels of coverage is shown in Figs. 5.7

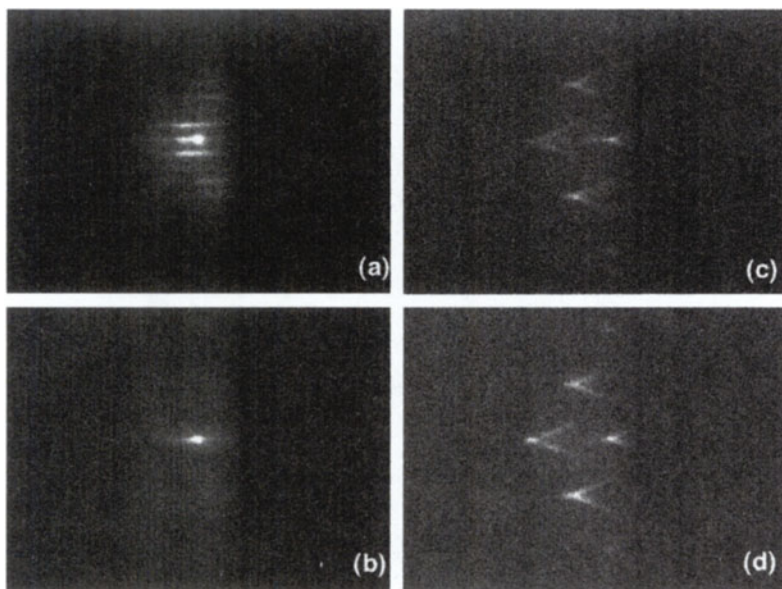
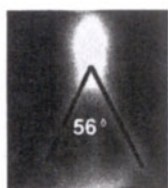
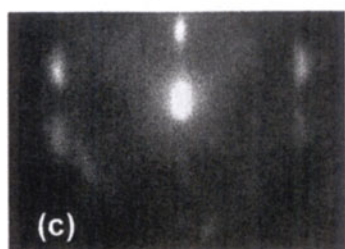
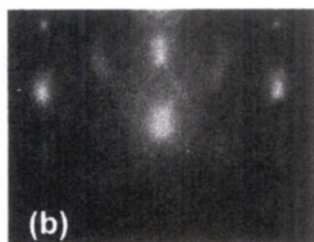
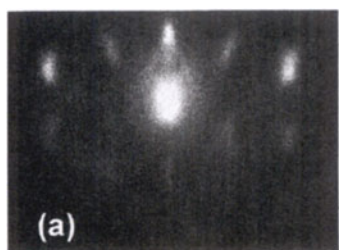


Figure 5.8: Typical RHEED pattern with the electron beam on the $[0\bar{1}1]$ azimuth for deposition of (a) 0 monolayers, (b) 1.8 monolayers, (c) 2.0 monolayers, and (d) 2.7 monolayers of InAs.

and 5.8. As first reported by Nabetani et al. [22] the observation of well-defined chevrons along the $[0\bar{1}1]$ azimuth indicates diffraction from sharp facet planes. The included angle of the chevron is about 50° , which indicates that the facet planes have an angle of about 25° with respect to the $[001]$ azimuth. This is a (311) plane. The spot pattern when the RHEED beam is along the $[011]$ azimuth suggests that no well-defined facet is present along this direction. More recently, Lee et al. [23] have given a new interpretation of the RHEED results. They point out that there should be crosses, rather than chevrons, if the islands have (311) facets along the $[0\bar{1}1]$. They looked at the RHEED pattern in other azimuths, particularly $[3\bar{1}0]$ and $[1\bar{3}0]$ (Fig. 5.9). From these observations, along with those along $[0\bar{1}1]$ and $[011]$, they conclude that the QDs are composed of $\{136\}$ planes, and that the islands have C_{2v} symmetry (Fig. 5.10). These results should lead to better models of QD shape and correspondingly better band structure calculations.

AFM results have been complicated by the fact that the image that is obtained is a convolution of the AFM tip and the island. Unless the tip is much sharper than the island, it will be difficult to get a very accurate



(d)

Figure 5.9: RHEED patterns along (a) $[3\bar{1}0]$, (b) $[1\bar{3}0]$, and (c) $[310]$ azimuths after deposition of 1.68 monolayers of InAs. Panel (d) presents a closeup of the facet-related streaks from panel (b) [23].

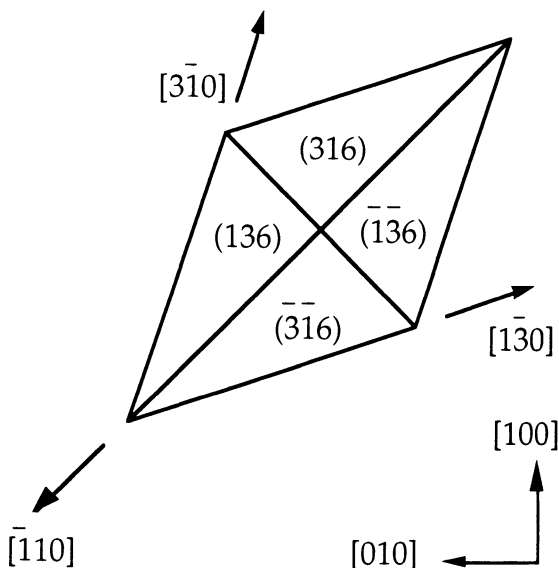


Figure 5.10: Proposed structure for the InAs quantum dots, based on the RHEED patterns shown in Fig. 9 [23].

measurement of the island shape. Moison et al. [17] describe the shape of the dots as rounded pyramids. The AFM images suggest a (410) facet plane, as opposed to the (511) plane that their TEM images suggest. Leonard et al. [8] report a radially symmetric, planoconvex (lens-like) shape of their QDs. However, this may be due to an inability of the AFM to measure the sharp tip of a QD.

TEM has also been used to ascertain the shape of the QDs. Depending upon the diffraction conditions that are used, TEM can be sensitive to variations in both strain and chemical composition. Grundmann et al. [24] have used plan-view TEM under diffraction conditions that are far from the exact Bragg condition, which should be sensitive to the morphology of the QDs, rather than the strain (Fig. 5.11). They conclude that the QDs are pyramidal shaped with a square base (C_{4v} symmetry) whose sides lie along $\langle 100 \rangle$ directions. They also see alignment of the QDs along the $\langle 100 \rangle$ directions, as has been previously reported [13] and as is shown in Fig. 5.3. Further examination of TEM images will likely be needed, especially considering the recent RHEED interpretation of Lee et al. [23].

There have been several reports of in-plane alignment of QDs [8, 13, 18, 24–26]. Glas et al. [13] and Grundmann et al. [24] showed that InAs QDs generally line up along $\langle 100 \rangle$ directions on (001) substrates. The reasons

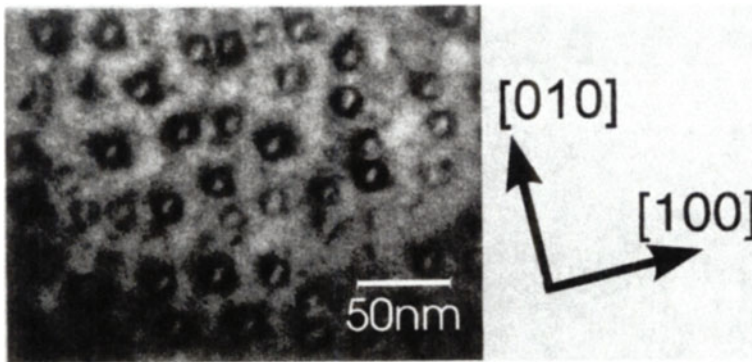


Figure 5.11: Plan-view TEM image using off-Bragg conditions in order to enhance the contrast due to morphological differences, rather than contrast due to strain fields or compositional differences [24].

for this alignment are not yet understood, but interaction between QDs via the strain field is one likely candidate. Leonard et al. [8] and Kitamura et al. [27] observed that $\text{In}_x\text{Ga}_{1-x}\text{As}$ QDs can be preferentially nucleated at step edges (Fig. 5.12). This suggests that In adatoms preferentially incorporate at the step edges due to the high density of kinks at the step edges. Mirin [25] found that aligned InGaAs could be formed on top of GaAs terraces with multiple steps at the edges of the terraces (Fig. 5.13). Indium adatoms are unable to migrate down the step edges (Schwoebel barrier) and instead are reflected at the step edge. This leads to island nucleation along the top edge of the terraces.

5.2.3 Overgrowth of self-assembled quantum dots

The quantum dots that are formed by the Stranski–Krastanow transition are highly nonplanar and have partial strain relaxation at the edges of the dot. Accordingly, the QDs must be buried in AlGaAs or some other large bandgap material in order to form a top barrier for electrons and holes. The growth of this top barrier is critical for obtaining highly efficient radiative recombination in the QDs. If the overgrowth quality is poor (interface states form at the QD–top barrier interface), then carrier trapping in the surface states can drastically reduce the efficiency of the QDs. Furthermore, the thickness of this overgrowth layer is important for the growth of QD superlattices. The QDs must be buried by the top barrier, but the spacer thickness between layers can be varied, and this variable can be adjusted to give electronic coupling between the layers if desired;

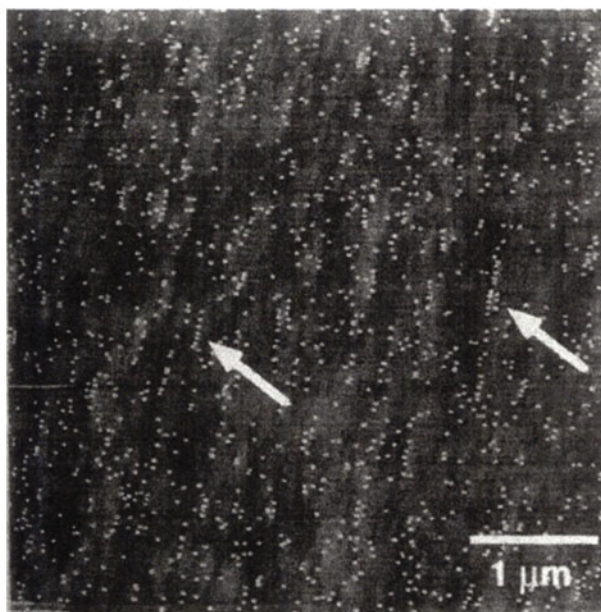


Figure 5.12: A $5\mu\text{m} \times 5\mu\text{m}$ AFM image of InAs quantum dots that shows preferential nucleation at step edges [8].

the growth of QD superlattices will be discussed further in the next section.

Xie et al. [28] have compared cap layer growth using migration enhanced epitaxy (MEE) and conventional MBE. Using low temperature photoluminescence, they show an order-of-magnitude improvement in integrated PL intensity for a sample with 2.0 monolayers of InAs grown at 500°C , capped with GaAs grown at 400°C by MEE, as compared to samples with identical InAs QDs but GaAs cap layers grown by conventional MBE at 480°C and 400°C .

One potential problem of using MEE to grow the GaAs cap layer is the strong intermixing of the InAs with the GaAs cap layer due to segregation of the In atoms. The segregation is enhanced on the high Miller index surfaces that form the sides of the quantum dots [29]. Mirin et al. [30] have used a different approach to optimize the GaAs overgrowth. They grow InGaAs QDs using alternating MBE at a substrate temperature of about 515°C . Next they deposit varying amounts of GaAs on top of the QDs at the growth temperature of 515°C using conventional MBE, and then raise the temperature to 570°C before continuing the GaAs growth. The PL inten-

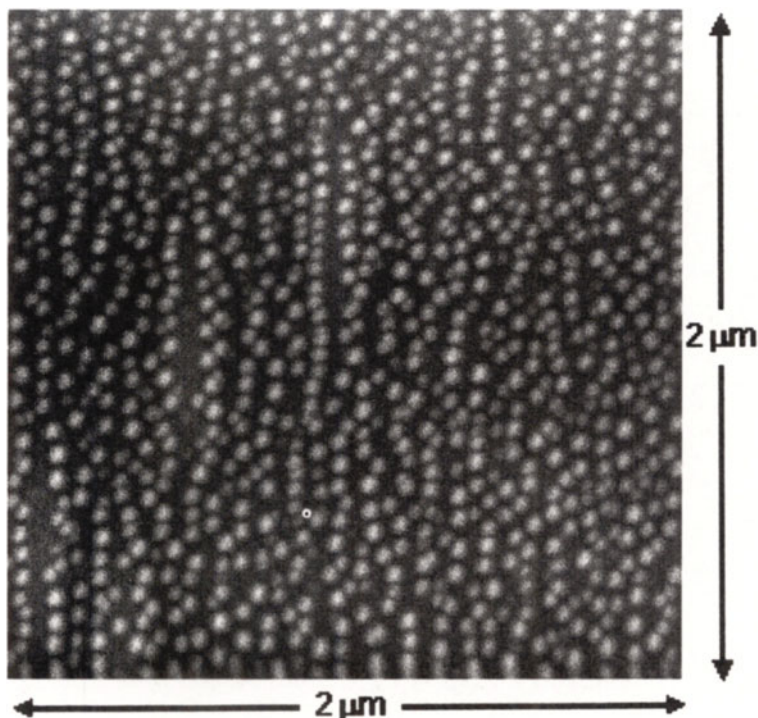


Figure 5.13: A $2\mu\text{m} \times 2\mu\text{m} \times 15\text{nm}$ AFM image of InGaAs quantum dots aligned along the tops of GaAs terraces that are formed during the growth of the buffer layer.

sity and peak emission wavelength vary as the amount of GaAs deposited at 515°C is varied. Some intermediate thickness of GaAs gives the best luminescence results. If too thin a layer is deposited, the peak emission wavelength blueshifts and the intensity is decreased; if too thick a layer is deposited, the intensity is substantially decreased. The variation in peak emission wavelength is due to In segregation and/or re-evaporation from the QDs. Both of these outcomes lead to a higher emission energy. This technique minimizes the thickness of GaAs grown at low temperatures, which can lead to defects, while maintaining the shape and In composition of the QDs, which leads to a stable emission wavelength.

The overgrowth and burial of the QDs has also been investigated using AFM [31] (Fig. 5.14). They find that when the islands have been only partially buried (2.0nm of GaAs deposited), there is a large dip at the island center where very little GaAs has covered the InAs island. As the

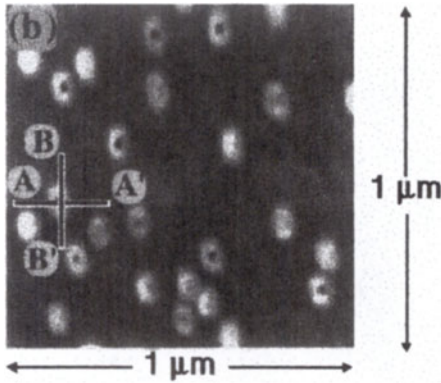


Figure 5.14: A $1\mu\text{m} \times 1\mu\text{m}$ AFM image of InAs quantum dots (areal density about $5 \times 10^9\text{cm}^{-2}$) capped with 2.0 nm of GaAs. The dark spots at the center of some of the islands correspond to holes in the overgrowth [31].

coverage increases to about 5.0 nm, the dip fills in, leaving a depression of about two monolayers. This AFM study has important implications for the growth of stacked layers of QDs and the calculation of electron and hole wave functions in stacks of QDs.

5.2.4 Quantum dot superlattices

In order to increase the spatial density of QDs, several layers of $\text{In}_x\text{Ga}_{1-x}\text{As}$ QDs can be grown with GaAs barriers in between the layers (quantum dot superlattices [QDSLs]). Several groups have demonstrated that growing multiple layers of InGaAs QDs can lead to a spatial correlation between subsequent layers [11,13,25,32–35]. Improved uniformity (both island size and spacing) within a given layer of a multilayer stack has been theoretically predicted [36] and demonstrated [37]. QDSLs have been particularly important for achieving high-performance QD lasers; this will be discussed further in a later section.

Goldstein et al. [11] first observed the vertical alignment of the InAs islands using STEM. They point out that this “memory effect” indicates that the GaAs is nonuniformly strained. Yao et al. [32] were the first to investigate the spacer thickness between multiple layers of islands. They found that in order to obtain vertical alignment, the GaAs spacer had to be thin enough for the strain field from the underlying QDs to propagate up to where the subsequent layer of QDs was formed. Xie et al. [33] found very similar results using InAs QDs formed by depositing two monolayers of

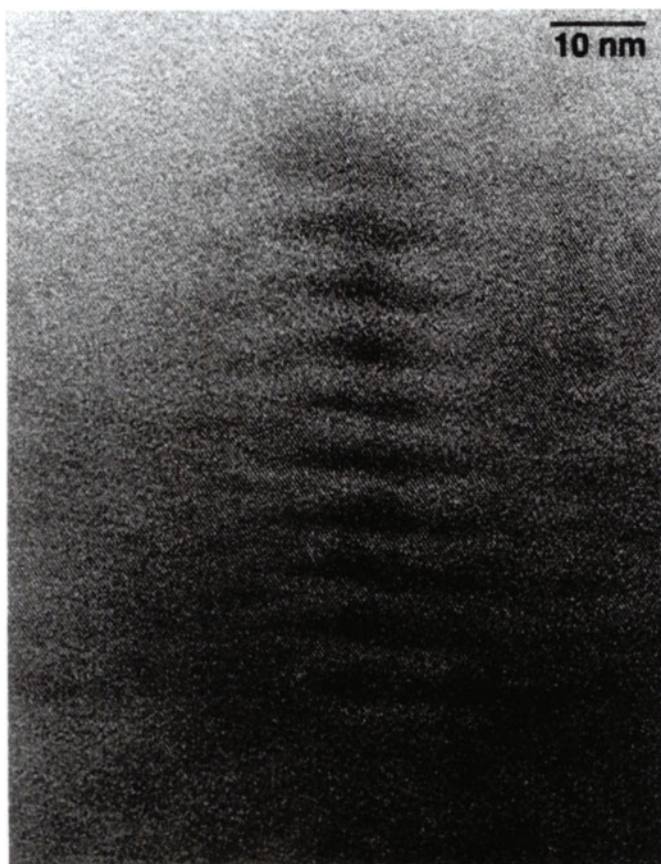


Figure 5.15: A TEM cross section image shows an aligned column of quantum dots [35].

InAs. Their systematic study of vertical alignment and spacer thickness compares very well with a phenomenological growth model that includes surface diffusion during growth. This model demonstrates that the strain field due to the underlying islands creates a local potential minimum directly over the island, thus providing a preferential site for subsequent island formation.

Solomon et al. [35] demonstrated vertical alignment of InAs islands that are formed from three monolayers of InAs with 20 monolayers of GaAs in between each layer of islands. They demonstrated that up to ten layers of islands can be grown without forming dislocations (Fig. 5.15). Furthermore, they demonstrated that the 8K photoluminescence linewidth

monotonically decreases from about 70 meV to 55 meV by adding layers of QDs, implying a more uniform size distribution. Solomon et al. [37] also performed AFM measurements on 1-, 10-, and 20-layer stacks of InAs QDs. The results indicate that the QD density decreases, but the mean QD height and in-plane diameter both increase when the number of layers is increased.

These results agree well with the model of Tersoff et al. [36], in which the nucleation of islands is the prime consideration [5]. The strain from the underlying layers gives a local minimum in the mismatch, and this local minimum leads to a reduced nucleation barrier. Since the island nucleation rate depends exponentially on the nucleation barrier, a local minimum in the nucleation barrier leads to a much higher probability of island nucleation. Calculations show that, depending upon initial island spacing, an equilibrium distribution of island spacing and sizes can be achieved in as few as five layers. These simulations also suggest that, when the initial islands are closely spaced (meaning that the strain fields do not overlap) in the plane of the islands, some of the columns will terminate and the density of islands will decrease. This has also been verified experimentally [37].

One point that is generally neglected in the above papers is the layer-to-layer uniformity and its effect on device performance. During the growth of these QD superlattices, the first QD layer requires more monolayers than the subsequent layers, as indicated by the change in the RHEED pattern from streaks to spots [25,38]. This is consistent with the model of Tersoff et al. [36], where the nucleation rate controls the formation of a layer of islands. The local variations due to the underlying islands reduce the nucleation barriers, increasing the island nucleation rate and leading to the observation of the RHEED transition with fewer number of deposited monolayers.

5.2.5 Growth on non-(100) substrates

Although (100) GaAs is by far the most commonly available substrate and thus of the most interest for QD formation, there has been interest in investigating other orientations of substrates. It is expected that there will be some differences in the growth of QDs grown on orientations other than (100) because the atomic arrangement on the surface, and hence the bond strengths, is different. This will affect the growth, since these differences in surface structure affect surface diffusion length, sticking coefficient, step density termination, etc. Particularly noteworthy are the results obtained by Nötzel et al. [39]. They used OMVPE to grow $\text{In}_{0.25}\text{Ga}_{0.75}\text{As}$ and $\text{In}_{0.4}\text{Ga}_{0.6}\text{As}$ on $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$ buffer layers on (311)B GaAs. They found that

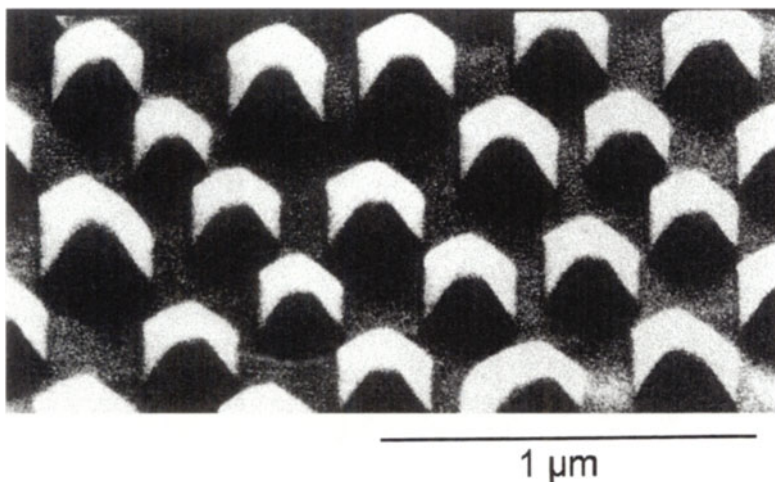


Figure 5.16: Scanning electron micrograph of ordered array of $\text{Al}_{0.5}\text{Ga}_{0.5}\text{As}$ microcrystals that contain buried $\text{In}_{0.2}\text{Ga}_{0.8}\text{As}$ quantum disks [39].

a growth interruption after the growth of the originally flat InGaAs films leads to a spontaneous rearrangement of the InGaAs film into an array of islands buried by AlGaAs. These islands, called *quantum disks* by the authors, have lateral diameters ranging from 30 to 200 nm. The islands form remarkably well-ordered arrays, as seen in the scanning electron micrograph in Fig. 5.16. Room-temperature PL linewidths as low as 6 meV have been measured from these arrays of QDs, and room-temperature lasing has been achieved with threshold current densities that are slightly lower than similar lasers grown on (100) substrates [40]. However, the differential efficiency from the (311)B lasers is much worse than that from the (100) lasers, and the output power saturates more readily, so it is not necessarily advantageous to use the (311)B substrate for laser growth.

Although the OMVPE QDs on (311)B GaAs are not due to an SK transition, SK growth of InAs and $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ QDs on (311)B GaAs by MBE has been demonstrated [41]. The thickness required for the 2D-to-3D transition to occur was shown to be only slightly larger for a (311)B substrate than for a (100) substrate. AFM measurements show that the uniformity of the QDs on the (311)B is much better than that of the (100) QDs; the standard deviation of the height is 30% larger for the (100) oriented QDs and the standard deviation of the diameter, 50%. The best PL linewidths from the (311)B QDs are 35 meV at 2 K and 41 meV at 300 K, as opposed to 61 meV and 102 meV from the QDs on (100) GaAs.

Gonzalez-Borrero et al. [42] reported on growth of QDs on several different orientations, including (100), (211)A, (111)A/B, (511)A/B, and (711)A/B. All orientations except (511)A and (711)A exhibited QD formation. Optical properties of the orientations on which QDs did form showed similar results as the more common (100) orientation. Further study is needed to determine if there are other advantages to using other than (100)-oriented substrates for QD formation.

5.2.6 Summary

Arrays of self-assembled $\text{In}_x\text{Ga}_{1-x}\text{As}$ QDs with In mole fractions ranging from about 30% to 100% have been grown using both MBE and OMVPE. The thickness at which the transition from two-dimensional growth to three-dimensional growth occurs depends on several factors, including In mole fraction (strain), growth temperature, As flux, and substrate orientation. As the In mole fraction increases, the growth becomes especially sensitive to the growth conditions. Multiple layers of quantum dots (QDSLs) can be grown, which is important for device applications. These layers align vertically because of the strain field from the underlying QDs. The limitations on dot-to-dot uniformity and layer-to-layer uniformity need further study. Recent theoretical modeling of self-assembled island growth [43–45] shows promise for predicting methods to improve uniformity and control areal density.

The InGaAs quantum dot arrays that are currently being grown are far from perfect. The biggest and most critical challenge that needs to be faced is the size uniformity. Although the size uniformity now is around $\pm 10\%$, better uniformity should lead to bigger improvements in device performance. There has not however, been any substantial improvement in the uniformity of these arrays since the work was first begun, which suggests that there may be some fundamental limitations that are preventing the desired improvements. This is one area that is ripe for new avenues of research.

Another difficult challenge is control of the areal dot density and position. It would be desirable to be able to control the dot density, size, and energies of the quantized levels, but this has so far proven difficult. Most of the reports of dot densities are between $3 \times 10^{10} \text{ cm}^{-2}$ and $8 \times 10^{10} \text{ cm}^{-2}$, with some reports around $1\text{--}2 \times 10^{11} \text{ cm}^{-2}$. Higher dot densities would be desirable for laser applications, and lower densities would be desirable for studying the properties of individual quantum dots. Control of the in-plane position of the QDs is also desirable for engineering complex single-electron structures and arrays of such structures. This may be possible using growth on patterned substrates, but no systematic technique has yet

been developed. Better understanding of the growth mechanisms involved in the QD formation are needed for researchers to obtain better control over these properties.

5.3 Optical characterization of quantum dots

5.3.1 Photoluminescence, photoluminescence excitation spectroscopy, and polarization-dependent photoluminescence

The optical properties of self-assembled QDs are obviously important for development of improved devices and/or devices with new functionality, based upon QDs. It is important to measure properties of both arrays of QDs, as a measure of size uniformity and possible use in current generations of devices, as well as measuring properties of individual QDs, since the individual characteristics will set limitations on future devices. In this section, experimental results on both arrays of QDs and individual QDs will be discussed.

The arrays of QDs typically exhibit low-temperature PL linewidths that range from about 30 to 100 meV [11,15,16,46]. These large linewidths are due to variations in the size, shape, and alloy composition of the QDs. The size distribution of these arrays of QDs is typically $\pm 10\%$. This leads to an inhomogeneously broadened linewidth that is dominated by the size distribution; since the height of the QDs is typically only about 20% of the lateral dimension, the height fluctuations lead to much larger potential fluctuations than do the variations in lateral dimension. The best indication that the linewidth is inhomogeneously broadened due to the size distribution is that the linewidth does not monotonically decrease when the temperature is decreased, as the linewidth of a QW does. In fact, there are reports that the linewidth will decrease slightly when the temperature is decreased from room temperature, but will actually increase again when the temperature is lowered even further [25,47–49]. This might be caused by dot-dot tunneling that is preferred at higher temperatures or by excited-state emission that overlaps with the ground-state emission.

Once arrays of quantum dots had been demonstrated and preliminary PL experiments showed promising results, the proof that these QDs were truly zero-dimensional structures could be obtained. Photoluminescence excitation (PLE) and resonant PL experiments by Fafard et al. [50] clearly demonstrated the zero-dimensional nature of the QDs. An $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ QD array with a low-temperature PL FWHM of 50 meV was studied. The PL

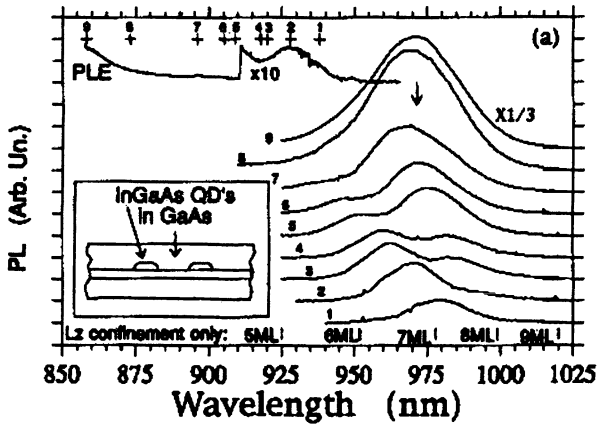


Figure 5.17: The PL spectra obtained when the pump laser wavelength is varied. The peak emission wavelength is seen to change as the pump laser excites carriers selectively into various (uncoupled) quantum dots. The PLE spectra in the upper left indicates the laser excitation wavelength for the various PL spectra [50].

peak wavelength was around 972 nm when the QDs were excited by the 514.5 nm line of an Ar^+ laser. By varying the energy of the exciting laser around the peak energy of the first excited state in the QDs, the peak emission wavelength varied from about 960 nm to about 980 nm (Fig. 5.17). This resonant excitation into the excited states of the QDs indicates that the carriers excited in a given QD are mostly confined to that QD and are unable to diffuse to a lower energy state, as would typically happen in a QW. Further evidence is seen in the PLE spectrum measured on these QDs (Fig. 5.18). As the detection wavelength is varied over the peak of the QD ground state emission, the PLE spectra show a shift in the position of the first excited state. This shift again indicates that only a certain fraction of the QDs, specifically those QDs that have their ground state at the detection wavelength *and* have their first excited state at the excitation energy, are detected with this measurement.

As noted above, there is evidence (RHEED, AFM, and TEM) that the islands have an anisotropic shape. Nabetani et al. [22] measured the polarization dependence of the PL at 10 K. They observed that the PL emission with the electric field parallel to the $[110]$ azimuth was approximately twice that when the electric field was parallel to the $[1\bar{1}0]$. This is a good indication of the anisotropic shape and/or strain of the QDs, and it will have important consequences for the polarization properties of QD vertical cavity lasers [51].

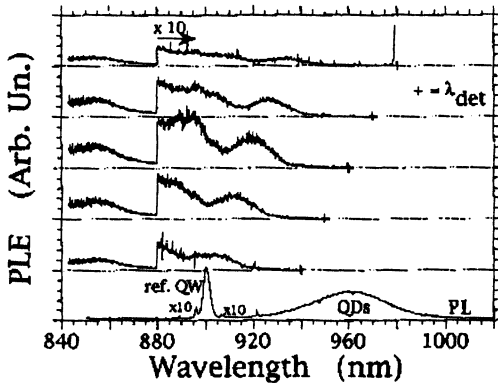


Figure 5.18: PLE spectra obtained when the detection wavelength is varied along the peak of the emission spectra. The change in PLE between the PLE spectra shows that carriers that are excited in a given quantum dot tend to recombine in that dot, rather than diffusing to the lowest available energy state in the system (which is often in a different dot) [50].

5.3.2 Single QD photoluminescence

The properties of individual QDs are important because they set limitations on improvements in device performance. Most of the PL results on arrays of QDs have linewidths of 40 meV or more. This linewidth does not vary much as the temperature changes because the linewidth is limited by the inhomogeneous broadening due to the size distribution of the QDs. Because of the high density of QDs ($\geq 10^9 \text{ cm}^{-2}$), a typical PL experiment with a spot diameter of $100 \mu\text{m}$ excites 10^5 QDs or more. Therefore, in order to investigate smaller numbers of QDs, special techniques must be used.

Marzin et al. [52] used e-beam lithography to define square mesas with dimensions from 100 nm to 5,000 nm in order to isolate some number of QDs. They used a microscope objective to focus the exciting laser light down to a spot diameter of about $2 \mu\text{m}$. As the mesa size was reduced, the PL spectra showed more evidence of discrete transitions. When the mesa size was 500 nm, individual QD emission lines were detected with linewidths less than $100 \mu\text{eV}$ (limited by the experimental setup) (Fig. 5.19). Fafard et al. [53] also etched mesas to isolate finite ensembles of QDs. Their results were similar to those of Marzin et al. [52]; as the size of the mesa decreased (and thus the number of QDs contributing to the spectrum decreased), the spectrum became better resolved—that is, the spikes due to individual QDs became apparent, rather than the

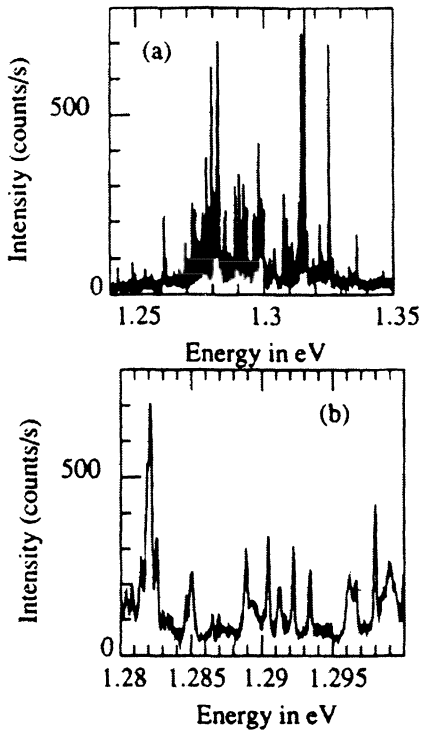


Figure 5.19: (a) 10 K PL spectrum of a 500 nm square mesa that contains QDs formed by depositing 2.2 monolayers of InAs at 520°C. (b) A blowup of part of the spectrum in (a) shows the very sharp emission lines from individual quantum dots [52].

broad Gaussian peak that appeared when large numbers of QDs were sampled.

Grundmann et al. [25] used spatially resolved cathodoluminescence (CL) to excite small numbers of QDs. They also observed sharp lines (as narrow as $150\text{ }\mu\text{eV}$, instrument limited) in the spectra (Fig. 5.20). There was no measurable difference in the linewidth as the temperature was increased from 5 K up to 50 K. This is further evidence that the ensemble linewidths (30–100 meV) that are typically measured at low temperatures are due to inhomogeneous broadening from the size distribution of the array of QDs. Furthermore, this linewidth puts an upper limit on the amount of improvement that can be obtained by using extremely uniform arrays of QDs in devices.

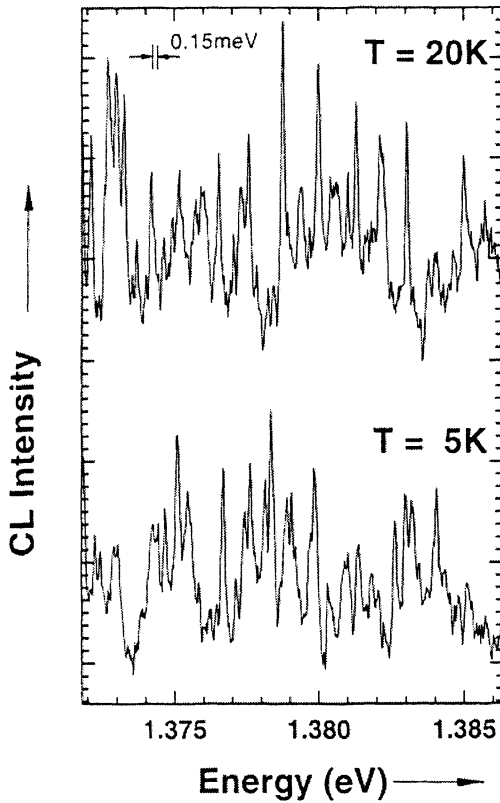


Figure 5.20: Low-temperature cathodoluminescence measurements on arrays of QDs using a tightly focused electron beam spot show the emission lines from individual dots [25].

5.3.3 Time-resolved photoluminescence

Time-resolved PL measurements on self-assembled QDs are particularly interesting because of the prediction of a phonon bottleneck [54,55] in QDs due to the discrete density of states in the QDs. This prediction states that when an electron is captured by a QD into an excited state of the QD, that electron will have a difficult time cooling to the ground state and then recombining radiatively because there are no intermediate energy states in the QD that are exactly one longitudinal optical (LO) phonon energy away. This is unlike a QW or quantum wire, where the density of states is a continuum and hot electrons can easily cool by phonon emission into the continuum.

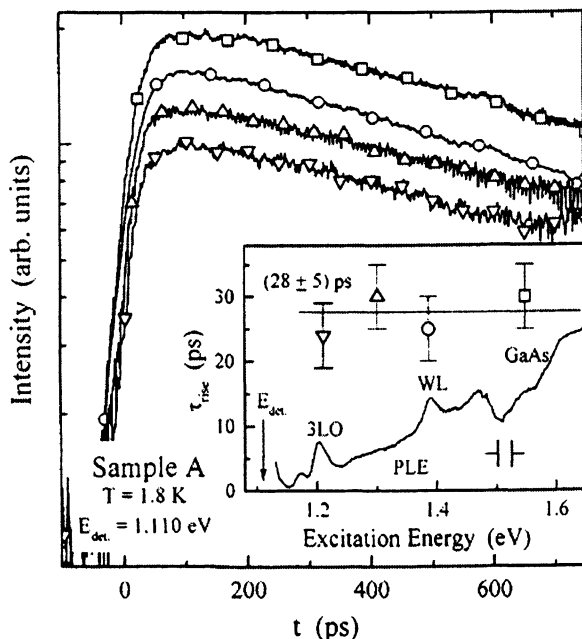


Figure 5.21: Time-resolved PL measurements on InAs QDs with various excitation energies of the pump laser. The decay times are all 28 ± 5 ps, which indicates that only the last relaxation process is resolved [58].

Wang et al. [56] compared the PL decay time for a QD and a QW. At 1.4 K, they found that the QD luminescence has a decay time of 882 ps, as compared to 336 ps for the QW luminescence. However, despite this increased decay time, the integrated PL intensity was similar for the QDs and the QW. If a phonon bottleneck was the cause of this reduced decay time, then the nonradiative lifetime should be even longer than this. Kamath et al. [57] also performed time-resolved experiments on QDs. For a sample with six layers of QDs, they found a decay time for the ground state of 700 ps and a decay time for the first excited state of 250 ps at a temperature of 13 K. The rapid decay of the excited state emission is believed to be due to relaxation to the ground state and re-emission to the barriers.

Heitz et al. [58] used PLE and time-resolved PL to investigate multiphonon processes for carrier relaxation in InAs/GaAs QDs. As shown in Fig. 5.21, they found that excitation at various energies still leads to the same PL rise time of 28 ± 5 ps. This implies that only the energy relaxation process that corresponds to the lowest energy excitation is resolved. This

lowest-energy relaxation process is shown to be almost resonant with the absorption into the first excited hole state in the QD. Therefore, the first excited hole state is a relaxation bottleneck; however, because this relaxation process is still faster than the nonradiative or radiative recombination processes, it is not observed in PL experiments.

5.3.4 Summary

The optical properties of arrays of QDs and individual QDs have been studied using several different techniques. The arrays of QDs exhibit large inhomogeneous linewidths (30–100 meV) because of the size distribution of the QDs. However, when spectroscopic techniques with high spatial resolution are used, individual QD emission lines can be resolved that are limited by the instrument resolution to about $100\mu\text{eV}$ wide. This linewidth sets an upper limit on possible device improvements. The emission from the QDs is shown to be polarized in the plane of the QDs because of the anisotropic shape of the QDs. Time-resolved PL measurements show that although there may be a phonon bottleneck that limits carrier relaxation, the nonradiative processes are still much slower than the carrier relaxation times, indicating the high quality of the interfaces of the QDs with the surrounding matrix.

Several remaining optical properties of QDs have not yet been measured. For example, although the wave function overlap between electrons and holes is expected to be very large because of the spatial confinement, there is little or no information available about excitonic properties of QDs. Also, little is known about the nonlinear optical properties of QDs, except for some intensity-dependent PL measurements that give saturation intensities. Until many more of these unknown properties are measured, we will not have a good understanding of the limitations of QDs.

5.4 Electrical characterization of quantum dots

5.4.1 Capacitance spectroscopy

Capacitance spectroscopy (C–V measurements) is a powerful way to investigate the electron and hole states in the self-assembled QDs. A particular advantage of capacitance spectroscopy over other spectroscopic techniques is the ability to separately investigate the electron and hole levels in the QD, rather than to see a convolution of the two, as in PL or PLE measurements. A typical structure for performing C–V measurements

consists of a layer of QDs sandwiched between a tunneling barrier to a back ohmic contact and a spacer layer to a Schottky contact [59]. A bias is applied between the two contacts, which allows the Fermi level to be swept through the range of quantized levels in the dots. At low temperatures, the capacitance is directly proportional to the DOS at the Fermi level:

$$c = \frac{dQ}{dV} = q \frac{dn}{d\mu} \frac{d\mu}{dV} = q \frac{dn}{d\mu} \rho(E_F), \quad (5.1)$$

where $\rho(E_F)$ is the DOS at the Fermi level, n is the electron concentration, μ is the chemical potential, and q is the electronic charge. Thus, the C–V curve can be used to determine the energies of the confined states in the QD, as well as the states in the wetting layer.

Drexler et al. [59] performed the first capacitance and magnetocapacitance spectroscopy on self-assembled $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ QDs. They investigated the electron states at 4.2 K and found an energy-level spacing between the ground state and the first excited state of 41 meV. This finding is significant because it is more than twice the Coulomb charging energy of 18 meV. Therefore, these self-assembled QDs are suitable for investigating electron–electron interactions in QDs, experiments that are extremely difficult to interpret in other QD systems in which the Coulomb charging energy is comparable to the energy level spacing. Medeiros-Ribeiro et al. [60] also performed C–V measurements on InAs QDs. They investigated both *n*-type and *p*-type samples at 4.2 K in order to find the electron and hole states. As shown in Fig. 5.22, two electron states and one hole state are clearly distinguishable in the C–V plot. The capacitance dramatically increases when the Fermi level is swept into the two-dimensional wetting layer that lies beneath the QDs because of the large increase in the DOS at this energy. The energy difference between the electron and hole ground states is 1.3 eV, which is close to the PL transition peak of 1.27 eV.

Brunkov et al. [61] performed capacitance spectroscopy on *n*-type samples at temperatures ranging from 77 K to 400 K. They proposed a model, based on Poisson's equation, that allows them to extract several parameters from the C–V data, including the areal dot density, the electron energy, the degree of filling of the electron states, and the distance of the QD layer from the Schottky barrier. The C–V measurements can then be correlated with other measurements; that is, the distance of the QD layer from the Schottky barrier should be known according the sample design, and the areal density can be correlated with plan-view TEM images or AFM images. Good agreement is obtained between the C–V measurements and the other measurements.

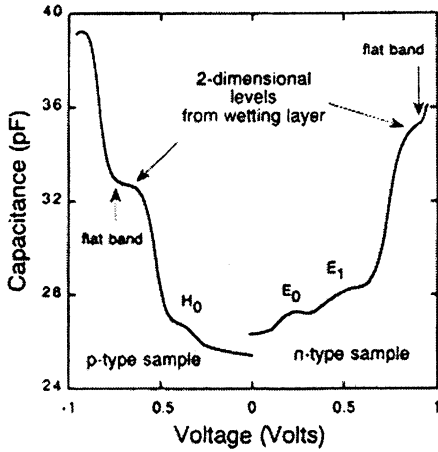


Figure 5.22: Capacitance–voltage characteristics for *p*- (left side) and *n*- (right side) type samples of InAs quantum dots. The formation of the wetting layer two-dimensional electron/hole gas is shown together with the estimated flatband voltage. The 2DEG is not completely formed in the *n*-type sample due to gate leakage [60].

5.4.2 Transport

Sakaki et al. [62] have investigated transport in a two-dimensional electron gas (2DEG) that is modulated by a layer of InAs quantum dots parallel to the plane of the 2DEG. The QDs are grown after the 2DEG and are placed at various distances from the 2DEG. As the distance from the channel to the QDs is decreased, the mobility decreases and the electron concentration increases. This finding indicates that electrons can be scattered and/or localized by the QD potential, especially when there is a strong parallel conduction path. One potential use for such a structure is a field-effect transistor with a programmable threshold voltage [63]. Horiguchi et al. [64] have studied parallel transport through a layer of self-assembled InAs QDs using a split-gate configuration. They observe peaks in the *I*–*V* characteristics separated by about 42 meV. Modulation was observed up to about 20 K.

5.4.3 Summary

Very few electrical properties of InAs QDs have been measured. Capacitance spectroscopy has proved to be a useful tool for investigating the quantum levels of both electrons and holes in QDs. Parallel transport

through arrays of QDs is poorer than a 2DEG because of scattering by the QDs. Further experiments need to be done to investigate the transport properties of arrays of QDs and single QDs.

5.5 Quantum dot lasers

5.5.1 Quantum dot edge-emitting lasers

In addition to the fundamental physics that can be studied in zero-dimensional systems, several applications can be improved by using QDs. The most widely studied application is the QD laser, that is, a semiconductor diode laser that uses QDs as the active region. Since the first report of a QD laser with self-assembled QDs in 1994 [65], several other groups have reported room-temperature operation of QD lasers [66–71]. To date, the best reported threshold current densities are only slightly worse than the best reports from QW lasers, even though QW lasers have been studied for almost 15 years.

The first proposal for a “multidimensional quantum well” laser was by Arakawa and Sakaki in 1982 [9]. In this seminal paper, the authors showed that the threshold current in QD lasers should be much less sensitive to temperature than that of bulk or quantum well lasers, that is, a QD laser could have an infinite value of characteristic temperature T_0 ($I_{thr} = I_0 \exp(T/T_0)$). This is a substantial improvement over the expected $T \ln(T)$ increase in threshold current of a quantum well laser. This result can be understood by looking at the expression for gain [72].

$$g(h\nu_0) = K \int \rho_r(E_{21})(f_2 - f_1)L(h\nu_0 - E_{21})|M_T(E_{21})|^2 dE_{21}, \quad (5.2)$$

where $g(h\nu_0)$ is the gain at energy $h\nu_0$, K is a constant, ρ_r is the reduced DOS at the interband energy E_{21} , L is a lineshape broadening function, M_T is the matrix element for the transition, and $f_2(f_1)$ is the Fermi function for the conduction (valence) band. The Fermi functions have the important temperature dependence of the carrier distribution in them:

$$f_2 = \frac{1}{\exp[(E_2 - E_{fc})/k_B T] + 1} \quad (5.3a)$$

and

$$f_1 = \frac{1}{\exp[(E_1 - E_{fv})/k_B T] + 1}, \quad (5.3b)$$

where E_{fc} (E_{fv}) is the quasi-Fermi level in the conduction (valence) band, k_B is Boltzmann's constant, and T is temperature. At a given temperature

T , if the intersubband energy difference ΔE_{10} between the first excited state and the ground state is several times $k_B T$, then all the carriers will occupy the ground-state energy level. Therefore, assuming that the losses are independent of temperature, the carrier density at threshold will remain constant for all temperatures such that $\Delta E_{10} \gg k_B T$.

Additional improvements in the performance of QD lasers versus QW lasers, including reduced threshold current (almost one order of magnitude improvement compared to a quantum well) and higher differential gain, were discussed by Asada et al. [10] in 1986. These improvements can also be easily understood by examining Eq. (5.2). If all the carriers go into the single energy state from which lasing occurs, rather than a distribution of energy states (including many states that do not contribute gain at the lasing energy), then the threshold current will be lowered. Similarly, the differential gain dg/dn will be higher if *all* injected electrons contribute to the lasing mode. The theory in these papers has led many researchers to attempt to fabricate QD lasers by various techniques. However, these other techniques, such as electron beam lithography, combined with etch-and-regrowth, or magnetic field confinement, are complicated and not practical enough to generate other than academic interest.

The self-assembled $\text{In}_x\text{Ga}_{1-x}\text{As}$ QDs have provided a fairly simple technique for investigating QD lasers. The crystal growth can be performed in one step, which is clearly beneficial compared to etch-and-regrow schemes for fabricating QDs. The interfaces between the QDs and the surrounding matrix are buried in situ, which is advantageous for eliminating defects. The laser structure is identical to a QW laser, except for the substitution of one or more layers of QDs for the QWs. Processing of the QD lasers proceeds the same as a QW laser. Therefore, QD lasers that perform better than QW lasers can be implemented without significant changes in manufacturing processes.

The first QD laser using InGaAs QDs was demonstrated in 1994 by Kirstaedter et al. [65]. This laser had only a single layer of QDs, but it still lased from the ground state at 77 K. Despite the somewhat high threshold currents of this initial device (950 A/cm^2 at room temperature, 120 A/cm^2 at 77 K), it demonstrated a high characteristic temperature of 350 K from 50–120 K (Fig. 5.23). This was the first demonstration of the improved properties of QD lasers using self-assembled QDs as the active region.

Since then, several other groups have also demonstrated room-temperature operation of QD lasers [66–68,70,73] and vertical cavity QD lasers [51,69,71,74,75]. Ground-state lasing at room temperature has been demonstrated using multiple layers of QDs. These lasers have threshold current densities comparable to the best QW lasers at a similar wavelength,

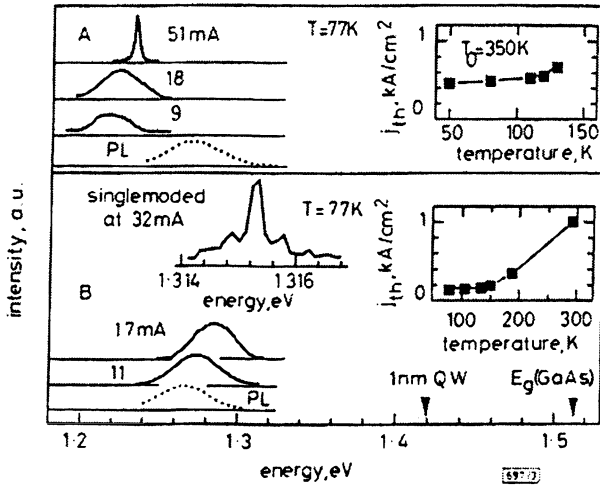


Figure 5.23: A single layer of $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ quantum dots emits from the ground state at low temperatures. The characteristic temperature from 50 K to 120 K is 350 K [65].

and furthermore, the QD lasers have even better characteristic temperatures than the QW lasers.

Room-temperature operation of a QD laser emitting from the ground state with low threshold current density and high internal quantum efficiency has been achieved [76]. Room-temperature threshold current densities as low as 63 A/cm^2 have been achieved by using three layers of InAs QDs inserted in an AlGaAs matrix. At 77 K, threshold current densities as low as 18 A/cm^2 have been obtained. The barrier height for electrons and holes is increased by using an AlGaAs barrier rather than a GaAs barrier, which decreases the spread of carriers into the barrier states. This allows the range of high characteristic temperature to be extended to 300 K. Characteristic temperatures as high as 350 K have been measured from 77 to 300 K.

One interesting characteristic that QD lasers emitting from excited states have is the temperature dependence on lasing wavelength (Fig. 5.24). In a QW laser, the lasing wavelength decreases as the temperature of the device is decreased because of the increase in bandgap. However, in a QD laser that lases from excited states, the lasing wavelength is much less sensitive to temperature than that of a QW laser. The QD active region also has a bandgap that increases with decreasing temperature, but the increase in bandgap is partially offset by the decrease in state filling

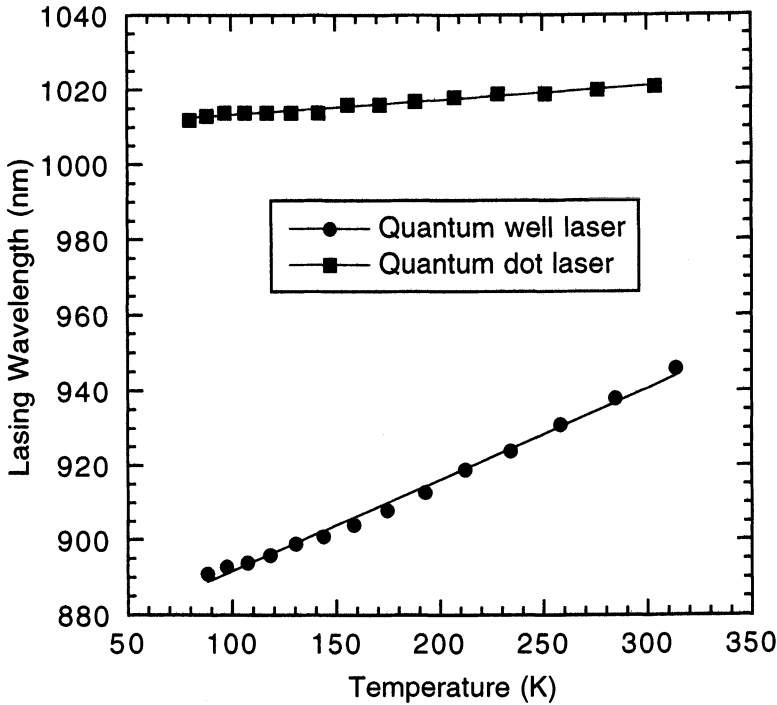


Figure 5.24: Temperature dependence of lasing wavelength for InGaAs QD and QW lasers. The QD laser lases from excited states. The lasing wavelength of the QD laser is much less sensitive to temperature than the QW laser because the reduction in state filling of the QDs as the temperature decreases partially compensates for the increase in bandgap [68].

(decreased J_{th}) as the device is cooled. This leads to a laser whose lasing wavelength is only weakly dependent on temperature, as seen in Fig. 5.24.

The other interesting property of QD lasers that lase from excited states is that the lasing wavelength depends strongly on laser length (Fig. 5.25). This property is different from QW lasers, where the lasing wavelength is practically independent of laser length because of the step-like DOS. The length dependence of the lasing wavelength concurs with the subthreshold electroluminescence from the QDs. For single-layer QD lasers below threshold, the ground state saturates and excited states dominate the EL. The position of the EL peak energy increases as the current density increases, until threshold is finally obtained. Thus, the length dependence on lasing wavelength becomes clear: as the laser length decreases, the distributed mirror loss α_m ($\alpha_m = \text{Length}^{-1} \cdot \ln(\text{Reflectance}^{-1})$) increases and the

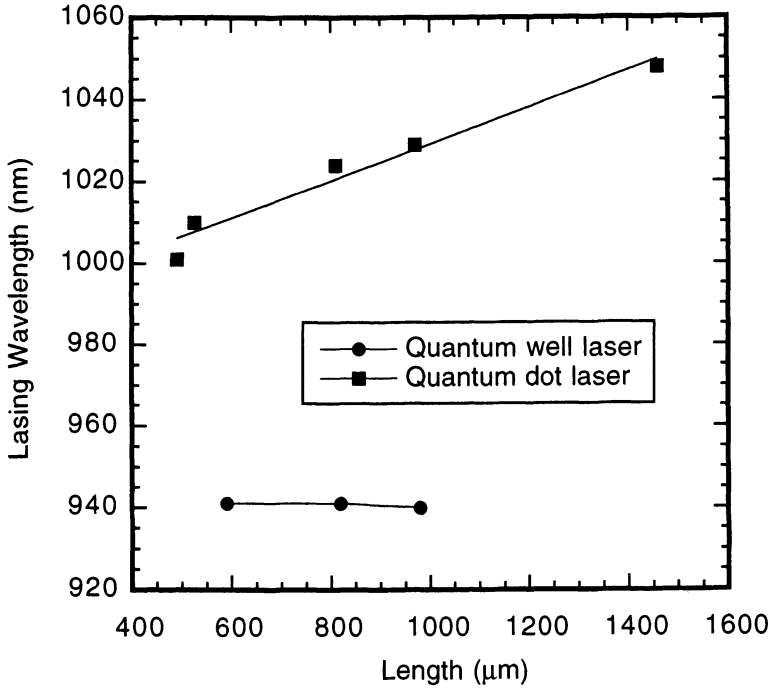


Figure 5.25: Length dependence of lasing wavelength for InGaAs QD and QW lasers. The QD laser's lasing wavelength is strongly dependent on length, and therefore distributed mirror loss, because of the continuum DOS that arises from the size distribution of the QDs [68].

distributed gain required to reach threshold increases. Thus, the threshold current density increases as the length decreases, and the lasing energy increases due to state filling.

Zhukov et al. [77] have reported that *negative* values of characteristic temperature can be obtained at low temperatures (<150 K). As shown in Fig. 5.26, when an AlGaAs spacer layer is used between the QDs, so that the continuum states are energetically even higher than if a GaAs spacer is used, then below about 150 K the threshold current density starts to increase again. This outcome is completely unexpected in the case of an equilibrium carrier distribution. However, since the QDs described by Zhukov et al. [77] are widely spaced, it is possible that at low temperatures the thermal excitation between dots is eliminated, and thus a nonequilibrium carrier distribution is obtained. Thus, the system cannot be described by a constant Fermi level, and the observed behavior becomes explicable.

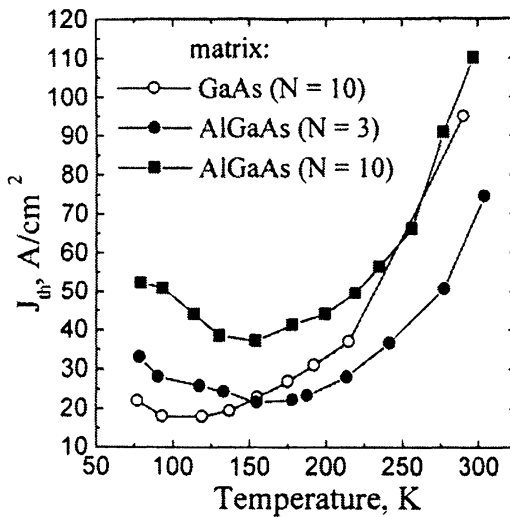


Figure 5.26: Temperature dependence of threshold current density for vertically coupled quantum dot lasers with different number of quantum dot layers and different spacer layer compositions. Note the negative characteristic temperatures seen at low temperatures [77].

5.5.2 Quantum dot VCSELs

QD active regions have also been used to make VCSELs [51,69,74,78,79]. The first QD VCSEL had a high threshold current of about 30 mA [69]. However, rapid improvements in performance have been obtained by using state-of-the-art oxide aperture, as is now common for quantum well VCSELs. Lott et al. [74] have used vertically coupled QDs (three QD layers) to achieve room-temperature CW lasing with threshold currents around 200 μ A. They obtain 12% efficiency and output power around 175 μ W from lasers with a threshold current of 270 μ A. Huffaker et al. [78] have used a single QD layer to achieve room-temperature CW operation with a threshold current of 235 μ A and a differential slope efficiency of 17% and output power greater than 100 μ W.

These results are not as good as the best oxide-confined QW lasers [79] where threshold currents of 36 μ A and slope efficiencies of 23% have been measured. Nonetheless, the QD results can be expected to improve as growth improvements are developed and the QDs are used in optimized designs. In addition, the QD VCSEL has demonstrated an important characteristic—stable polarization—that can only be obtained in QW VCSELs by using additional processing steps [51]. Because the self-assembled QDs

have an anisotropic shape [22,23,47], it is expected that one in-plane polarization will be favored, as has been seen in polarization-dependent PL measurements [22]. Saito et al. [51] have demonstrated 18 dB of polarization suppression at 1 mW output power. If this polarization ration can be maintained in the low-threshold current oxide-aperture VCSELs, then a strong case can be made for immediate used of QD VCSELs in systems that require highly polarized lasing emission.

5.5.3 High-speed quantum dot lasers

Investigations into the high-speed modulation of in-plane QD lasers are under way. These investigations are still in their early stages, but are interesting because of the various mechanisms that may enhance or impede the modulation bandwidth of these lasers. Because of the DOS of an array of QDs, the differential gain, and thus the modulation bandwidth, is expected to be higher than QW or bulk lasers [10]. However, it is possible that slow relaxation of carriers due to a phonon bottleneck may limit the high-speed performance of QD lasers [55].

Kirstaedter et al. [80] performed the first measurements of gain and differential gain on an InAs QD laser. They measured a $25\text{ }\mu\text{m}$ -wide stripe laser that had only a single layer of QDs. This device lased from the ground state at 77 K, the temperature at which all their measurements were made. At a current density of 80 A/cm^2 , a material gain of $6.8 \times 10^4\text{ cm}^{-1}$ is reported. A differential gain of $2 \times 10^{-12}\text{ cm}^2$ at a current density of 20 A/cm^2 is also reported. This value of differential gain is much higher than quantum well lasers and suggests that high modulation bandwidths will be possible unless carrier transport effects (capture and cooling) limit the performance.

Kamath et al. [81] were the first to report on small-signal bandwidth and differential gain at room temperature. Their laser was a single-mode ridge waveguide ($3\text{ }\mu\text{m}$ wide) with a single-layer QD active region consisting of 10 monolayers (nominally) of $\text{In}_{0.4}\text{Ga}_{0.6}\text{As}$. Note that these lasers were lasing from an excited state with a spontaneous lifetime of 250 ps. Under pulsed operating conditions, they measured a small-signal, 3 dB bandwidth of 7.5 GHz. From the modulation response, they were able to extract a differential gain of $1.7 \times 10^{-14}\text{ cm}^2$. This result is about two orders of magnitude smaller than the value reported by Kirstaedter et al. [80]. However, this difference is at least in part due to the difference in temperature between the two experiments, as well as the assumptions of each group in extracting this value. This same group also used high-frequency electrical impedance measurements to characterize these lasers [82]. Their results indicate that carrier capture times are substantially longer for QD

lasers as compared to QW lasers, which will limit the overall modulation bandwidth of these lasers.

Mao et al. [83] reported measurements on the relaxation oscillation frequency from three stacked InAs QDs sheets. They measured the relaxation oscillations at various power levels from this laser. At a power level of 9.2 mW per facet, the relaxation oscillation frequency saturated at 5.3 GHz, which can be extrapolated to a 3 dB bandwidth of 8.2 GHz, neglecting any carrier transport effects. Gain saturation is speculated to limit the response in these lasers. Further investigations into the carrier transport limitations of QD lasers will be necessary before any more conclusive statements can be made regarding their high-speed performance.

5.5.4 Summary

Quantum dot lasers have demonstrated extremely promising results already, with threshold currents of edge-emitting lasers approaching the best quantum well lasers. The threshold current density is insensitive to temperature when the QD lasers emit from the ground state, as predicted by theory; this is likely to reduce packaging costs for these lasers as compared to QW lasers. The modulation bandwidths of QD lasers do not yet approach the best QW lasers. It is still unclear if there are fundamental limitations, such as a carrier relaxation bottleneck, that will prevent QD lasers from achieving the improved modulation bandwidths that have been predicted. Unusual characteristics from QD lasers that lase from excited states have been observed, including a lasing wavelength that is relatively insensitive to temperature and a lasing wavelength that depends strongly on laser length. These can be explained by state-filling effects in the higher-lying states of the QDs.

QD lasers must meet several challenges before they become technologically and commercially important. In order to compete directly with the established technology (quantum well lasers), QD lasers must offer either improved performance or new functionality. Improved performance includes lower threshold currents, higher modulation bandwidths, higher characteristic temperature, and higher output powers. The reduced threshold current and increased characteristic temperature are likely to be achieved once the size uniformity is improved. Higher modulation bandwidth will require further understanding of the carrier relaxation processes in the QDs and engineering of the QDs to enhance the relaxation rates. New functionality, such as the demonstration of polarization control in VCSELs, has already been demonstrated by QD lasers. With improved uniformity, other new devices such as single-longitudinal-mode Fabry-Perot lasers might be practical.

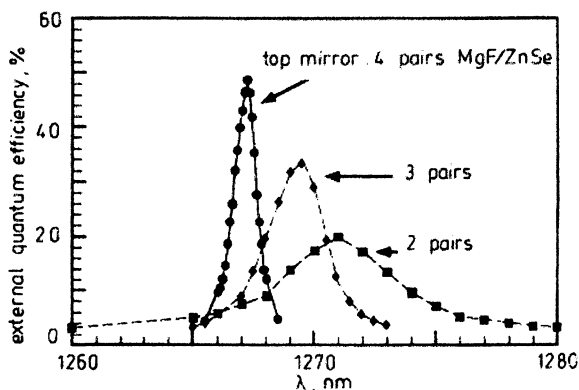


Figure 5.27: External quantum efficiency of a QD photodetector with varying number of pairs in the MgF/ZnSe top Bragg reflector. The spectral width is only 1.2 nm for the four-pair top mirror [85].

5.6 Other applications

QDs may be able to solve problems that cannot be solved with quantum wells. One example is the ability to reach the technologically important wavelength of $1.3\text{ }\mu\text{m}$ on GaAs [47,84] by growing QDs well beyond the 2D-to-3D transition. Although this technology has not yet been used to form a $1.3\text{ }\mu\text{m}$ laser on GaAs, a resonant cavity photodetector near $1.3\text{ }\mu\text{m}$ has been demonstrated [85]. This device demonstrates the possibility of additional functionality using quantum dots. It has an external quantum efficiency of 49% with a spectral bandwidth of 1.2 nm (Fig. 5.27), the narrowest spectral response of any normal-incidence photodetector. The quantum dot absorbing region has a relatively low absorbance, which is an advantage in obtaining narrow spectral bandwidths.

Huffaker et al. [75] have demonstrated the longest wavelength of a laser that uses InGaAs QDs on GaAs. They obtained an 1,154 nm VCSEL with a room temperature CW threshold of $502\text{ }\mu\text{A}$ for a $10\text{ }\mu\text{m}$ -diameter structure. The subthreshold electroluminescence from this device had a spectral linewidth of about 64 meV, which is not unusual for these arrays of self-assembled QDs. However, the longer wavelength emission ($1.3\text{ }\mu\text{m}$) involves larger QDs, so fluctuations in QD size become less significant. Linewidths as low as 28 meV have been reported [47] for long-wavelength-emitting QDs, which suggests that further improvements in threshold current density may be expected.

Another possible application for self-assembled QDs is optical memory [86,87]. In this case, it is desirable to have a large inhomogeneous linewidth

of the system, such as that provided by the size fluctuations of the InGaAs QDs. The proposed device structure relies on spectral hole burning, a nonlinear optical effect. This is the first proposed device based on the nonlinear optical properties of self-assembled QDs.

The anisotropic shape of the QDs leads to interesting optical properties. Polarization control in VCSELs was discussed earlier. Another example where this anisotropy leads to new functionality is with normal-incidence infrared detectors [88,89]. Normal-incidence infrared absorption (at wavelengths greater than about $10\text{ }\mu\text{m}$, due to intersubband transitions) is extremely weak in quantum wells on GaAs due to the opposing symmetry of the (adjacent) initial and final states. However, the anisotropy of the QDs leads to states that are unlikely to be purely symmetric or anti-symmetric, thus allowing transitions between adjacent initial and final states.

5.7 Conclusions and future work

The study of self-assembled InGaAs QDs is still in its infancy, but the future looks very promising. Threshold current densities of QD lasers are already approaching those of the best QW lasers, and the temperature performance of the QD lasers is superior to that of the QW lasers. New and unexpected results have been demonstrated by QD devices, such as polarization control in VCSELs and narrow spectral bandwidths in resonant cavity photodetectors. These devices need further development before they will replace the current generation of QW devices, but this may just be a matter of development time.

Many measurements remain to be made on QDs. Many of the properties of arrays of QDs have been measured, but only a few measurements on individual QDs have been made to date. Making measurements on individual QDs is both an opportunity and a challenge, mainly because of the difficulties associated with obtaining isolated QDs. Better understanding of the growth will lead to control over QD position and density, which will facilitate experiments with individual QDs.

Two areas have had relatively little investigation: possible electronic applications of self-assembled QDs and applications based on nonlinear optical properties of QDs. Devices such as single-electron storage devices and single-electron transistors may be possible, even at room temperature, if the Coulomb blockade effects can be utilized. These devices may require control over the positioning of the QDs, so further improvements in growth will be needed before they are implemented. Optical nonlinearities in QDs, particularly those based on excitonic effects, are expected to be enhanced

because of the large electron-hole wave function overlap due to the spatial carrier confinement. These properties have not yet been measured, and new devices based on these strong nonlinearities can be expected.

In summary, the study of self-assembled QDs of $\text{In}_x\text{Ga}_{1-x}\text{As}$ on GaAs has progressed rapidly since the technique for fabricating uniform arrays of quantum dots was discovered. The physics of these structures is beginning to be understood, and subsequent improvements in device performance can be expected. Despite the preliminary state of research on the growth and physics of these self-assembled quantum dots, devices that demonstrate performance comparable to or better than that of quantum well devices have already been fabricated. Further improvements in device performance are expected as the growth technique is improved to obtain even better uniformity from these arrays of QDs. We can expect other new applications of self-assembled QDs to arise as both the physics and technology of these structures becomes better understood.

ACKNOWLEDGMENTS

We thank our many colleagues who participated in this work, particularly John Bowers, John English, Evelyn Hu, James Ibbetson, Devin Leonard, Gilberto Medeiros-Riberiro, Kenichi Nishi, Pierre Petroff, Tom Reynolds, and Klaus Schmidt. Critical reviews of this manuscript by David Christensen (NIST) and Kenichi Nishi (NEC) are greatly appreciated. Financial support for this work at UC Santa Barbara was provided by the NSF's Science and Technology Center for Quantized Electronic Structures (QUEST). We also thank the many contributors of figures that are included here.

REFERENCES

- [1] F. Hatami, N. Ledentsov, H. Grundmann, J. Böhrer, F. Heinrichsdorff, M. Beer, D. Bimberg, S. Ruvimov, P. Werner, U. Gösele, J. Heydenreich, U. Richter, S. Ivanov, B. Meltser, P. Kop'ev, and Z. Alferov, "Radiative recombination in type-II GaSb/GaAs quantum dots," *Appl. Phys. Lett.* **67**(5), 656–658 (1995).
- [2] N. Carlsson, W. Seifert, A. Petersson, P. Castrillo, M. Pistol, and L. Samuelson, "Study of the two-dimensional–three-dimensional growth mode transition in metalorganic vapor phase epitaxy of GaInP/InP quantum-sized structures," *Appl. Phys. Lett.* **65**(24), 3093–3095 (1994).
- [3] S. Fafard, Z. Wasilewski, J. McCaffrey, S. Raymond, and S. Charbonneau, "InAs self-assembled quantum dots on InP by molecular beam epitaxy," *Appl. Phys. Lett.* **68**(7), 991–993 (1996).

- [4] A. Marti Ceschin and J. Massies, "Strain-induced 2D-3D growth mode transition in molecular beam epitaxy of $\text{In}_x\text{Ga}_{1-x}\text{As}$ on GaAs (001)," *J. Crystal Growth* **114**, 693–699 (1991).
- [5] J. Tersoff and F. LeGoues, "Competing relaxation mechanisms in strained layers," *Phys. Rev. Lett.* **72**(22), 3570–3573 (1994).
- [6] D. E. Jesson, S. J. Pennycook, J.-M. Baribeau, and D. C. Houghton, "Direct imaging of surface cusp evolution during strained-layer epitaxy and implications for strain relaxation," *Phys. Rev. Lett.* **71**(11), 1744–1747 (1993).
- [7] S. Guha, A. Madhukar, and K. Rajkumar, "Onset of incoherency and defect introduction in the initial stages of molecular beam epitaxial growth of highly strained $\text{In}_x\text{Ga}_{1-x}\text{As}$ on GaAs (100)," *Appl. Phys. Lett.* **57**(20), 2110–2112 (1990).
- [8] D. Leonard, K. Pond, and P. Petroff, "Critical layer thickness for self-assembled InAs islands on GaAs," *Phys. Rev. B* **50**(16), 11687–11692 (1994).
- [9] Y. Arakawa and H. Sakaki, "Multidimensional quantum well laser and temperature dependence of its threshold current," *Appl. Phys. Lett.* **40**(11), 939–941 (1982).
- [10] M. Asada, Y. Miyamoto, and Y. Suematsu, "Gain and the threshold of three-dimensional quantum-box lasers," *IEEE J. Quantum Electron.* **QE-22**(9), 1915–1921 (1986).
- [11] L. Goldstein, F. Glas, J. Marzin, M. Charasse, and G. LeRoux, "Growth by molecular beam epitaxy and characterization of InAs/GaAs strained-layer superlattices," *Appl. Phys. Lett.* **47**(10), 1099–1101 (1985).
- [12] W. Schaffer, M. Lind, S. Kowalczyk, and R. Grant, "Nucleation and strain relaxation at the InAs/GaAs(100) heterojunction," *J. Vac. Sci. Technol. B* **1**(3), 688–695 (1983).
- [13] F. Glas, C. Guille, P. Hénoc, and F. Houzay, *TEM Study of the Molecular Beam Epitaxy Island Growth of InAs on GaAs*. Microscopy of Semiconducting Materials (Institute of Physics, Oxford, England, 1987).
- [14] C. Snyder, J. Mansfield, and B. Orr, "Kinetically controlled critical thickness for coherent islanding and thick highly strained pseudomorphic films of $\text{In}_x\text{Ga}_{1-x}\text{As}$ on GaAs(100)," *Phys. Rev. B* **46**(15), 9551–9554 (1992).
- [15] M. Tabuchi, S. Noda, and A. Sasaki, "Mesoscopic structure in lattice-mismatched heteroepitaxial interface layers," in *Science and Technology of Mesoscopic Structures*, edited by S. Namba, C. Hamaguchi, and T. Ando (Tokyo, Springer-Verlag, 1992).
- [16] D. Leonard, M. Krishnamurthy, C. R. Reaves, S. P. DenBaars, and P. M. Petroff, "Direct formation of quantum-sized dots from uniform coherent islands of InGaAs on GaAs surfaces," *Appl. Phys. Lett.* **63**(23), 3203–3205 (1993).
- [17] J. M. Moison, F. Houzay, F. Barthe, L. Leprince, E. Andre, and O. Vatel, "Self-organized growth of regular nanometer-scale InAs dots on GaAs," *Appl. Phys. Lett.* **64**(2), 196–198 (1994).

- [18] J. Oshinowo, M. Nishioka, S. Ishida, and Y. Arakawa, "Highly uniform InGaAs/GaAs quantum dots (~15 nm) by metalorganic chemical vapor deposition," *Appl. Phys. Lett.* **65**(11), 1421–1423 (1994).
- [19] E. Steimetz, F. Schienle, J. Zettler, and W. Richter, "Stranski–Krastanow formation of InAs quantum dots monitored during growth by reflectance anisotropy spectroscopy and spectroscopic ellipsometry," *J. Crystal Growth* **170**, 208–214 (1997).
- [20] M. Geiger, A. Bauknecht, F. Adler, H. Schweizer, and F. Scholz, "Observation of the 2D–3D growth mode transition in the InAs/GaAs system," *J. Crystal Growth* **170**, 558–562 (1997).
- [21] F. Heinrichsdorff, A. Krost, M. Grundmann, D. Bimberg, F. Bertram, J. Christen, A. Kosogov, and P. Werner, "Self organization phenomena of InGaAs/GaAs quantum dots grown by metalorganic chemical vapour deposition," *J. Crystal Growth* **170**, 568–573 (1997).
- [22] Y. Nabetani, T. Ishikawa, S. Noda, and A. Sasaki, "Initial growth stage and optical properties of a three-dimensional InAs structure on GaAs," *J. Appl. Phys.* **76**(1), 347–351 (1994).
- [23] H. Lee, R. Lowe-Webb, W. Yang, and P. Sercel, "Determination of the shape of self-organized InAs/GaAs quantum dots by reflection high energy electron diffraction," *Appl. Phys. Lett.* **72**(7), 812–814 (1998).
- [24] M. Grundmann, N. Ledentsov, R. Heitz, L. Eeckey, J. Christen, J. Böhrer, D. Bimberg, S. Ruvimov, P. Werner, U. Richter, J. Heydenreich, V. Ustinov, A. Egorov, A. Zhukov, P. Kopev, and Zh. Alferov, "InAs/GaAs quantum dots radiative recombination from zero-dimensional states," *Phys. Stat. Sol. (B)* **188**, 249–258 (1995).
- [25] R. Mirin, "Alternating molecular beam epitaxy and characterization of InGaAs quantum dots and quantum dot lasers," *Electrical and Computer Engineering*. (University of California, Santa Barbara, 1997), p. 223.
- [26] K. Nishi, T. Anan, A. Gomyo, S. Kohmoto, and S. Sugou "Spontaneous lateral alignment of $\text{In}_{0.25}\text{Ga}_{0.75}\text{As}$ self-assembled quantum dots on (311)B GaAs grown by gas source molecular beam epitaxy," *Appl. Phys. Lett.* **70**(26), 3579–3581 (1997).
- [27] M. Kitamura, M. Nishioka, J. Oshinowo, and Y. Arakawa, "*In-situ* fabrication of self-aligned InGaAs quantum dots on GaAs multiautomic steps by metalorganic chemical vapor deposition," *Appl. Phys. Lett.* **66**(26), 3663–3665 (1995).
- [28] Q. Xie, P. Chen, A. Kalburge, T. Ramachandran, A. Nayfonov, A. Konkar, and A. Madhukar, "Realization of optically active strained InAs island quantum boxes on GaAs (100) via molecular beam epitaxy and the role of island induced strain fields," *J. Crystal Growth* **150**, 357–363 (1995).
- [29] M. Ilg, M. I. Alonso, A. Lehmann, K. H. Ploog, and M. Hohenstein, "Investigation of InAs submonolayer and monolayer structures on GaAs (100) and (311) substrates," *J. Appl. Phys.* **74**(12), 7188–7197 (1993).

- [30] R. P. Mirin, J. P. Ibbetson, J. E. Bowers, and A. C. Gossard, "Overgrowth of InGaAs quantum dots formed by alternating molecular beam epitaxy," *J. Crystal Growth* **175/176**, 696–701 (1997).
- [31] J. Garcia, G. Medeiros-Ribeiro, K. Schmidt, T. Ngo, J. Feng, A. Lorke, J. Kotthaus, and P. Petroff, "Intermixing and shape changes during the formation of InAs self-assembled quantum dots," *Appl. Phys. Lett.* **71**(14), 2014–2016 (1997).
- [32] J. Y. Yao, T. G. Andersson, and G. L. Dunlop, "The interfacial morphology of strained epitaxial $\text{In}_{(x)}\text{Ga}_{(1-x)}\text{As}/\text{GaAs}$," *J. Appl. Phys.* **69**(4), 2224–2230 (1991).
- [33] Q. Xie, A. Madhukar, P. Chen, and N. P. Kobayashi, "Vertically self-organized InAs quantum box islands on GaAs (100)," *Phys. Rev. Lett.* **75**(13), 2542–2545 (1995).
- [34] O. Schmidt, N. Kirstaedter, N. Ledentsov, M. Mao, D. Bimberg, V. Ustinov, A. Egorov, A. Zhukov, M. Maximov, P. Kop'ev, and Z. Alferov, "Prevention of gain saturation by multi-layer quantum dot lasers," *Electron. Lett.* **32**(14), 1302–1304 (1996).
- [35] G. S. Solomon, J. A. Trezza, A. F. Marshall, and J. S. Harris, "Vertically aligned and electronically coupled growth induced InAs islands in GaAs," *Phys. Rev. Lett.* **76**(6), 952–955 (1996).
- [36] J. Tersoff, C. Teichert, and M. Lagally, "Self-organization in growth of quantum dot superlattices," *Phys. Rev. Lett.* **76**(10), 1675–1678 (1996).
- [37] G. Solomon, S. Komarov, J. Harris, and Y. Yamamoto, "Increased size uniformity through vertical quantum dot columns," *J. Crystal Growth* **175/176**, 707–712 (1997).
- [39] R. Nötzel, J. Temmyo, H. Kamada, T. Furuta, and T. Tamamura, "Strong photoluminescence emission at room temperature of strained InGaAs quantum disks (200–30 nm diameter) self-organized on GaAs (311)B substrates," *Appl. Phys. Lett.* **65**(4), 457–459 (1994).
- [40] J. Temmyo, E. Kuramochi, M. Sugo, T. Nishiya, R. Nötzel, and T. Tamamura, "Strained InGaAs quantum disk laser with nanoscale active region fabricated with self-organisation on GaAs (311)B Substrate," *Electron. Lett.* **31**(3), 209–211 (1995).
- [41] K. Nishi, R. Mirin, D. Leonard, G. Medeiros-Ribeiro, P. Petroff, and A. Gossard, "Structural and optical characterization of InAs/GaAs self-assembled quantum dots grown on (311)B GaAs," *J. Appl. Phys.* **80**(6), 3466–3470 (1996).
- [42] P. Gonzalez-Borrero, E. Marega, Jr., D. Lubyshev, E. Petitprez, and P. Basmaji, "Molecular-beam epitaxy of self-assembled InAs quantum dots on non-(100) oriented GaAs," *J. Crystal Growth* **175/176**, 765–770 (1997).
- [43] A.-L. Barabási, "Self-assembled island formation in heteroepitaxial growth," *Appl. Phys. Lett.* **70**(19), 2565–2567 (1997).

- [44] H. T. Dobbs, D. D. Vvedensky, A. Zangwill, J. Johansson, N. Carlsson, and W. Seifert, "Mean-field theory of quantum dot formation," *Phys. Rev. Lett.* **79**(5), 897–900 (1997).
- [45] K. E. Khor and S. Das Sarma, "Surface morphology and quantum dot self-assembly in growth of strained-layer semiconducting films," *J. Vac. Sci. Technol. B* **15**(4), 1051–1055 (1997).
- [46] M. Grundmann, N. Ledentsov, O. Stier, J. Böhrer, D. Bimberg, V. Ustinov, P. Kop'ev, and Zh. Alferov, "Nature of optical transitions in self-organized InAs/GaAs quantum dots," *Phys. Rev. B* **53**(16): R10509–R10511 (1996).
- [47] R. Mirin, J. Ibbetson, K. Nishi, A. Gossard, and J. Bowers, "1.3 μm photoluminescence from InGaAs quantum dots on GaAs," *Appl. Phys. Lett.* **67**(25), 3795–3797 (1995).
- [48] D. Lubyshev, P. González-Borrero, E. Marega, Jr., E. Petitprez, N. La Scala, Jr., and P. Basmaji, "Exciton localization and temperature stability in self-organized InAs quantum dots," *Appl. Phys. Lett.* **68**(2), 205–207 (1996).
- [49] H. Lee, W. Yang, and P. Sercel, "Temperature and excitation dependence of photoluminescence line shape in InAs/GaAs quantum-dot structures," *Phys. Rev. B* **55**(15), 9757–9762 (1997).
- [50] S. Fafard, D. Leonard, J. L. Merz, and P. M. Petroff, "Selective excitation of the photoluminescence and the energy levels of ultrasmall InGaAs/GaAs quantum dots," *Appl. Phys. Lett.* **65**(11), 1388–1390 (1994).
- [51] H. Saito, K. Nishi, S. Sugou, and Y. Sugimoto, "Controlling polarization of quantum-dot surface-emitting lasers by using structurally anisotropic self-assembled quantum dots," *Appl. Phys. Lett.* **71**(5), 590–592 (1997).
- [52] J.-Y. Marzin, J.-M. Gerard, A. Izrael, and D. Barrier, "Photoluminescence of single InAs quantum dots obtained by self-organized growth on GaAs," *Phys. Rev. Lett.* **73**(5), 716–719 (1994).
- [53] S. Fafard, R. Leon, D. Leonard, J. Merz, and P. Petroff, "Visible photoluminescence from N-dot ensembles and the linewidth of ultrasmall AlInAs/AlGaAs quantum dots," *Phys. Rev. B* **50**(11), 8086–8089 (1994).
- [54] U. Bockelmann and G. Bastard, "Phonon scattering and energy relaxation in two-, one-, and zero-dimensional electron gases," *Phys. Rev. B* **42**(14), 8947–8951 (1990).
- [55] H. Benisty, C. M. Sotomayor-Torres, and C. Weisbuch, "Intrinsic mechanism for the poor luminescence properties of quantum-box systems," *Phys. Rev. B* **44**(19), 10945–10948 (1991).
- [56] G. Wang, S. Fafard, D. Leonard, J. Bowers, J. Merz, and P. Petroff, "Time-resolved optical characterization of InGaAs/GaAs quantum dots," *Appl. Phys. Lett.* **64**(21), 2815–2817 (1994).
- [57] K. Kamath, N. Chervela, K. Linder, T. Sosnowski, H.-T. Jiang, T. Norris, J. Singh, and P. Bhattacharya, "Photoluminescence and time-resolved photolumines-

- cence characteristics of InGaAs/GaAs self-organized single- and multiple-layer quantum dot laser structures," Appl. Phys. Lett. **71**(7), 927–929 (1997).
- [58] R. Heitz, M. Veit, N. Ledentsov, A. Hoffmann, D. Bimberg, V. Ustinov, P. Kop'ev, and Zh. Alferov, "Energy relaxation by multiphonon processes in InAs/GaAs quantum dots," Phys. Rev. B **56**(16), 10435–10445 (1997).
 - [59] H. Drexler, D. Leonard, W. Hansen, J. Kotthaus, and P. Petroff, "Spectroscopy of quantum levels in charge-tunable InGaAs quantum dots," Phys. Rev. Lett. **73**(16), 2252–2255 (1994).
 - [60] G. Medeiros-Ribeiro, D. Leonard, and P. Petroff, "Electron and hole energy levels in InAs self-assembled quantum dots," Appl. Phys. Lett. **66**(14), 1767–1769 (1995).
 - [61] P. Brunkov, S. Konnikov, V. Ustinov, A. Zhukov, A. Egorov, M. Maksimov, N. Ledentsov, and P. Kop'ev, "Capacitance spectroscopy of electron energy levels in InAs quantum dots in a GaAs matrix," Semiconductors **30**(5), 492–496 (1996).
 - [62] H. Sakaki, G. Yusa, T. Someya, Y. Ohno, T. Noda, H. Akiyama, Y. Kadoya, and H. Noge, "Transport properties of two-dimensional electron gas in AlGaAs/GaAs selectively doped heterojunctions with embedded InAs quantum dots," Appl. Phys. Lett. **67**(23), 3444–3446 (1995).
 - [63] G. Yusa and H. Sakaki, "GaAs/*n*-AlGaAs field-effect transistor with embedded InAs quantum traps and its programmable threshold characteristics," Electron. Lett. **32**(5), 491–493 (1996).
 - [64] N. Horiguchi, T. Futatsugi, Y. Nakata, and N. Yokoyama, "Electron transport properties through InAs self-assembled quantum dots in modulation doped structures," Appl. Phys. Lett. **70**(17), 2294–2296 (1997).
 - [65] N. Kirstaedter, N. N. Ledentsov, M. Grundmann, D. Bimberg, V. M. Ustinov, S. S. Ruvimov, M. V. Maximov, P. S. Kop'ev, Zh. I. Alferov, U. Richter, P. Werner, U. Gösele, and J. Heydenreich, "Low threshold, large T_0 injection laser emission from (InGa)As quantum dots," Electron. Lett. **30**(17), 1416–1417 (1994).
 - [66] D. Bimberg, N. Ledentsov, M. Grundmann, N. Kirstaedter, O. Schmidt, M. Mao, V. Ustinov, A. Egorov, A. Zhukov, P. Kopev, Zh. Alferov, S. Ruvimov, U. Gosele, and J. Heydenreich, "InAs-GaAs quantum pyramid lasers—*in-situ* growth, radiative lifetimes, and polarization properties," Jpn. J. Appl. Phys. **35**, 1311–1319 (1996).
 - [67] K. Kamath, P. Bhattacharya, T. Sosnowski, T. Norris, and J. Phillips, "Room temperature operation of $\text{In}_{0.4}\text{Ga}_{0.6}\text{As}/\text{GaAs}$ self-organized quantum dot lasers," Electron. Lett. **32**(15), 1374–1375 (1996).
 - [68] R. Mirin, A. Gossard, and J. Bowers, "Room temperature lasing from InGaAs quantum dots," Electron. Lett. **32**(18), 1732–1733 (1996).
 - [69] H. Saito, K. Nishi, I. Ogura, S. Sugou, and Y. Sugimoto, "Room-temperature lasing operation of a quantum-dot vertical-cavity surface-emitting laser," Appl. Phys. Lett. **69**(21), 3141–3142 (1996).

- [70] H. Shoji, Y. Nakata, K. Mukai, Y. Sugiyama, M. Sugawara, N. Yokoyama, and H. Ishikawa, "Room temperature CW operation at the ground state of self-formed quantum dot lasers with multi-stacked dot layer," *Electron. Lett.* **32**(21), 2023–2024 (1996).
- [71] D. Huffaker, L. Graham, and D. Deppe, "Low-threshold continuous-wave operation of an oxide-confined vertical cavity surface emitting laser based on a quantum dot active region and half-wave cavity," *Electron. Lett.* **33**(14), (1997).
- [72] L. A. Coldren and S. Corzine, *Diode Lasers and Photonic Integrated Circuits* (Wiley-Interscience, New York, 1995).
- [73] H. Shoji, K. Mukai, N. Ohtsuka, M. Sugawara, T. Uchida, and H. Ishikawa, "Lasing at three-dimensionally quantum confined sublevel of self-organized In_{0.5}Ga_{0.5}As quantum dots by current injection," *IEEE Photon. Technol. Lett.* **7**(12), 1385–1387 (1995).
- [74] J. Lott, N. Ledentsov, V. Ustinov, A. Egorov, A. Zhukov, P. Kop'ev, Zh. Alferov, and D. Bimberg, "Vertical cavity lasers based on vertically coupled quantum dots," *Electron. Lett.* **33**(13), 1150–1151 (1997).
- [75] D. Huffaker, H. Deng, and D. Deppe, "1.15 μ m Wavelength oxide-confined quantum dot vertical-cavity surface-emitting laser," *IEEE Photon. Technol. Lett.* **10**(2), 185–187 (1998).
- [76] V. M. Ustinov, A. Yu. Egorov, A. R. Kovsh, A. E. Zhukov, M. V. Maximov, A. F. Tsatsul'nikov, N. Yu. Gordeev, S. V. Zaitsev, Yu. M. Shernyakov, N. A. Bert, P. S. Kop'ev, Zh. I. Alferov, N. N. Ledentsov, J. Böhrer, D. Bimberg, A. O. Kosogov, P. Werner, and U. Gösele, "Low-threshold injection lasers based on vertically-coupled quantum dots," *J. Crystal Growth* **175/176**, 689–695 (1997).
- [77] A. Zhukov, V. Ustinov, A. Egorov, A. Kovsh, A. Tsatsulnikov, N. Ledentsov, S. Zaitsev, N. Gordeev, P. Kopev, and Zh. Alferov, "Negative characteristic temperature of InGaAs quantum dot injection laser," *Jpn. J. Appl. Phys.* **36**(6B), 4216–4218 (1997).
- [78] D. Huffaker, O. Baklenov, L. Graham, B. Streetman, and D. Deppe, "Quantum dot vertical-cavity surface-emitting laser with a dielectric aperture," *Appl. Phys. Lett.* **70**(18), 2356–2358 (1997).
- [79] D. Huffaker and D. Deppe, "Improved performance of oxide-confined vertical-cavity surface-emitting lasers using a tunnel injection active region," *Appl. Phys. Lett.* **71**(11), 1449–1451 (1997).
- [80] N. Kirstaedter, D. Schmidt, N. Ledentsov, D. Bimberg, V. Ustinov, A. Egorov, A. Zhukov, M. Maximov, P. Kop'ev, and Zh. Alferov, "Gain and differential gain of single layer InAs/GaAs quantum dot injection lasers," *Appl. Phys. Lett.* **69**(9), 1226–1228 (1994).
- [81] K. Kamath, J. Phillips, H. Jiang, J. Singh, and P. Bhattacharya, "Small-signal modulation and differential gain of single-mode, self-organized

- In_{0.4}Ga_{0.6}As/GaAs quantum dot lasers,” Appl. Phys. Lett. **70**(22), 2952–2953 (1997).
- [82] D. Klotzkin, K. Kamath, and P. Bhattacharya, “Quantum capture times at room temperature in high-speed In_{0.4}Ga_{0.6}As-GaAs self-organized quantum-dot lasers,” IEEE Photon. Technol. Lett. **9**(10), 1301–1303 (1997).
- [83] M. Mao, F. Heinrichsdorff, A. Krost, and D. Bimberg, “Study of high frequency response of self-organized stacked quantum dot lasers at room temperature,” Electron. Lett. **33**(19): 1641–1642 (1997).
- [84] K. Mukai, N. Ohtsuka, M. Sugawara, and S. Yamazaki, “Self-formed In_{0.5}Ga_{0.5}As quantum dots on GaAs substrates emitting at 1.3 μ m,” Jpn. J. Appl. Phys. **33**(12A), L1710–L1712 (1994).
- [85] J. Campbell, D. Huffaker, H. Deng, and D. Deppe, “Quantum dot resonant cavity photodiode with operation near 1.3 μ m wavelength,” Electron. Lett. **33**(15), 1337–1338 (1997).
- [86] S. Muto, “On a possibility of wavelength-domain-multiplication memory using quantum boxes,” Jpn. J. Appl. Phys. **34**(2B), L210–L212 (1995).
- [87] Y. Sugiyama, Y. Nakata, S. Muto, N. Horiguchi, T. Futatsugi, Y. Awano, and N. Yokoyama, “Observation of spectral hole burning in photocurrent spectrum of InAs self-assembled quantum dots embedded in *pin* diode,” Electron. Lett. **33**(19), 1655–1657 (1997).
- [88] D. Pan, Y. P. Zeng, M. Y. Kong, J. Wu, Y. Q. Zhu, C. H. Zhang, J. M. Li, and C. Y. Wang, “Normal incident infrared absorption from InGaAs/GaAs quantum dot superlattice,” Electron. Lett. **32**(18), 1726–1727 (1996).
- [89] D. Pan, E. Towe, and S. Kennerly, “Strong normal-incidence infrared absorption and photocurrent spectra from highly uniform (In,Ga)As/GaAs quantum dot structures,” Electron. Lett. **34**(10), 1019–1020 (1998).

Chapter 6

Structural characterization of self-organized Ge islands

A. A. Darhuber¹, G. Bauer¹, P. Schittenhelm², and G. Abstreiter²

¹Institut für Halbleiterphysik, Johannes Kepler Universität, Altenbergerstr. 69, A-4040 Linz, Austria; ²Walter Schottky Institut, Technische Universität München Am Coulombwall, D-85748 Garching, Germany

6.1 Introduction

Coherent island formation in the SiGe material system has been studied for a number of years [1,2]. The lattice mismatch of 4.2% allows the fabrication of islands with diameters ranging from about 100 to 2,500 Å.

For III–V material systems, the goal of Stranski–Krastanow growth has been the fabrication of optoelectronic devices with superior properties as compared with quantum well devices. In GaAs quantum-dot lasers operating up to 60°C were reported to exhibit a high gain and temperature stability [3]. The interest in self-assembled Ge-rich dots on silicon was driven by the perspective of an integration of quantum-functional devices with conventional silicon technology.

The first device structures containing Ge islands in active regions have extended the sensitivity of Si-based infrared photodetectors in the 1.5 μm range. An unexpected and dramatic increase of the luminescence

efficiency of this indirect bandgap material has been observed recently in Ge dots grown on fractional monolayers of carbon deposited on silicon substrates [4]. These recent achievements are a big step towards silicon optoelectronic devices based on the properties of nanostructured semiconductor materials.

6.2 Growth of self-organized Ge dots

Self-assembled (Si) Ge islands have been fabricated using low-pressure chemical vapor deposition (LPCVD) [5] and gas [6] and solid [7] source molecular beam epitaxy (MBE). Some kinetic trends in dot evolution are common to all fabrication methods, e.g., the diameter of coherent islands increases with increasing growth temperature and decreases with increasing Ge concentration and increasing growth rate [8,9]. A growth interruption after the onset of Stranski–Krastanow growth also leads to the lateral merging of originally smaller dots and the formation of larger islands [10].

6.3 Structural characterization

6.3.1 Description of the problem

The structural effects occurring during formation and overgrowth of self-assembled quantum dots are quite complex. The driving force for the formation of islands is the reduction of strain energy by elastic relaxation, i.e., without nucleation of dislocations. The surface diffusion, which leads to the self-organization, may be anisotropic and influenced by the strain fields of the islands. Since the growth takes place at high temperatures and low growth rates, diffusion of Ge from the wetting layer into the Si substrate may reduce the Ge concentration at the surface, which is counteracted by segregation. During overgrowth, both vertical and lateral segregation and diffusion of Ge occur, leading to an inhomogeneous three-dimensional distribution of the Ge concentration. The complexity of the Ge distribution manifests itself in the strain fields, which are also highly inhomogeneous and extend far into the surrounding Si matrix. These long-range strain fields are the origin of both the vertical stacking and the preferred lateral arrangement along the “soft” crystal directions with minimal Young’s modulus (see Fig. 6.1). Minima occur along the [100] directions, i.e., 130.2 GPa for Si and 103.4 GPa for Ge, respectively, and maxima along the [110] directions, with 169.2 GPa for Si and 137.2 GPa for Ge, respectively.

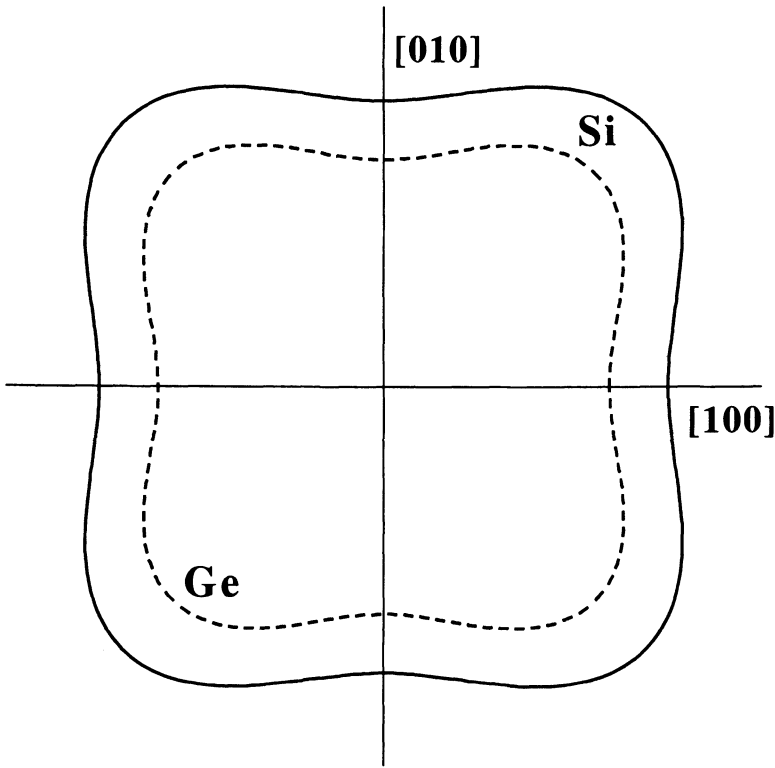


Figure 6.1: Polar plot of Young's modulus in the (001) plane for Si and Ge.

The partial elastic relaxation of the dots is, however, reduced after overgrowth and depends on the thickness of the Si cap layer. An analytical calculation of the strain fields in and around Ge islands is not possible. For qualitative behavior, we have studied a simple model in which we have replaced the islands by cylinders in an isotropic matrix.

The corresponding lateral strain relaxation ϵ_{xx} and the strain in the growth direction ϵ_{zz} in the center of the cylinder with a radius of $750 \text{ \AA}$, a height of $70 \text{ \AA}$, and an assumed Ge content of 70% are plotted as a function of the cap-layer thickness in Fig. 6.2. The chosen dimensions agree with the size of Ge dots grown at 670°C . As can be seen, there is little change if the cap-layer thickness exceeds $1,000 \text{ \AA}$, but below this value, the influence of the surface on the strain state has to be taken into account. If the dot diameter decreases, this limiting cap-layer thickness is correspondingly smaller than $1,000 \text{ \AA}$.

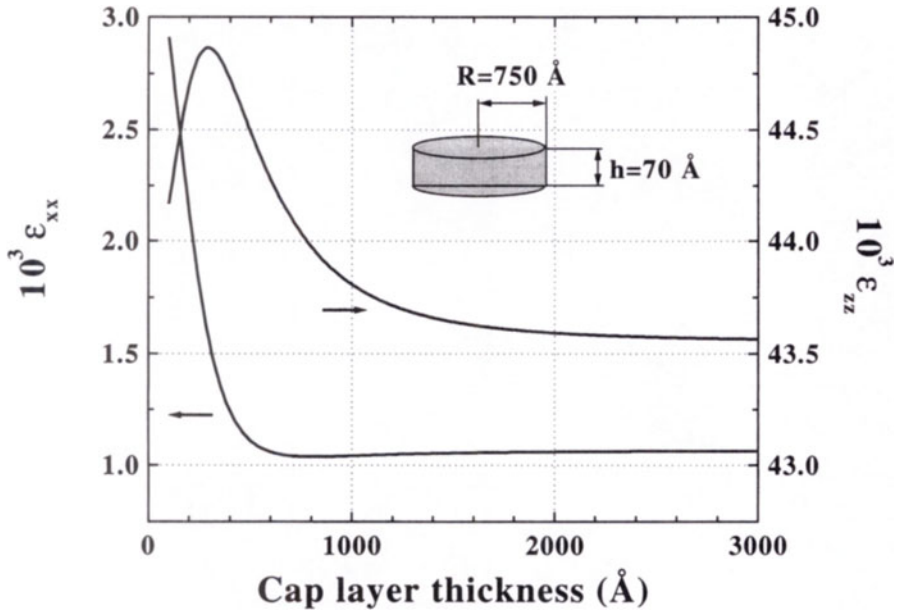


Figure 6.2: The calculated dependence of the strains ϵ_{xx} and ϵ_{zz} in the center of a cylindrical GeSi dot on the depth of the dot center below the free surface for a dot radius of $R = 750 \text{ \AA}$, a dot height of $h = 70 \text{ \AA}$, and a Ge concentration $x = 0.7$. Isotropic elastic constants were assumed identical for the dot lattice and the surrounding Si matrix.

Figure 6.3 shows a contour plot of ϵ_{xx} for a homogeneous, cuboidal dot within an elastically isotropic matrix. The material on the left- and right-hand sides of the island is compressively strained and ϵ_{zz} is negative, whereas above and below the island, the sign of ϵ_{zz} is positive, which corresponds to an expansion of the crystal lattice. It is this local increase of the lattice constant that favors the nucleation of islands in multilayers directly above existing ones and leads to the “stacking” of islands. Tersoff et al. [11] showed that the strain fields also enhance the lateral ordering of islands in a multilayer [12]. If in one layer two islands are rather close to each other, or if one island is much larger than another one in the vicinity of the first, their associated strain fields overlap and only one broader potential minimum occurs in the next layer, if the Si spacer layer is sufficiently thick. This effect can be seen in Fig. 6.4, where two AFM images of a single SiGe dot layer and the 25th layer in a dot superlattice are compared. As can be seen, the dot ensemble after 25 periods is significantly

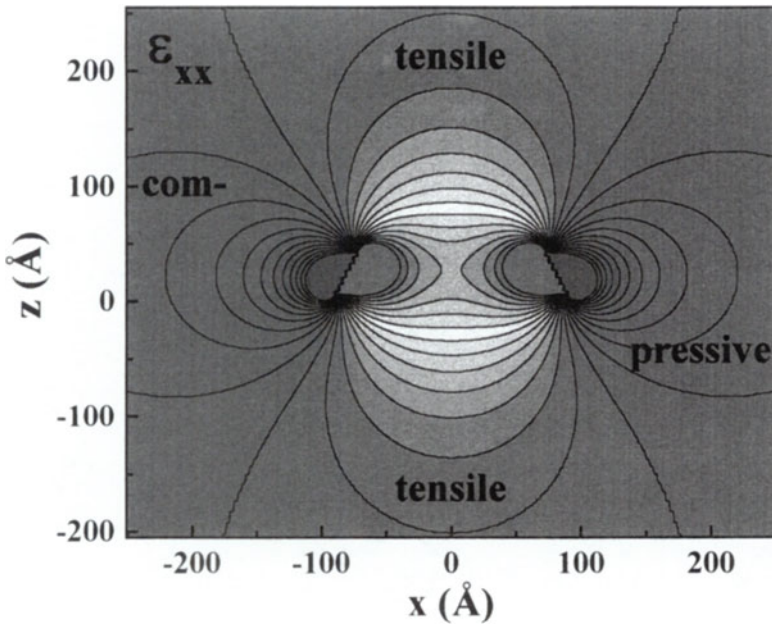


Figure 6.3: Contour map of the in-plane strain ϵ_{xx} in the central xz -plane of a homogeneous cuboidal dot oriented along the x -, y -, and z -coordinate axes in an isotropic matrix.

more homogeneous with reference to the positions, sizes, and shapes of the dots [13]. Whereas the tendency for vertical ordering of dots in multilayers is quite strong, the tendency for lateral ordering in the Si/SiGe system is much less pronounced. Recently, Holý et al. [14] have shown that the ordering of the dots in a two-dimensional square array is actually driven by the elastic anisotropy of the Si host crystal. The preferential nucleation is induced by the elastic deformation field produced by the dots lying at the buried interface in the multilayer stack.

The two-dimensional lateral ordering of self-assembled Ge islands can be improved considerably if the islands are grown on vicinal (001) Si substrates with regular, periodic ripples on the surfaces. This can be realized through the growth of a Si/SiGe multilayer that leads to a regular terrace pattern formed through step-bunching [15,16], which is used as template for further island growth [17–19].

The electronic and optical properties of self-organized Ge dots are influenced both by the local variation of the Ge concentration and by the inhomogeneous strain fields. For a comprehensive structural characterization,

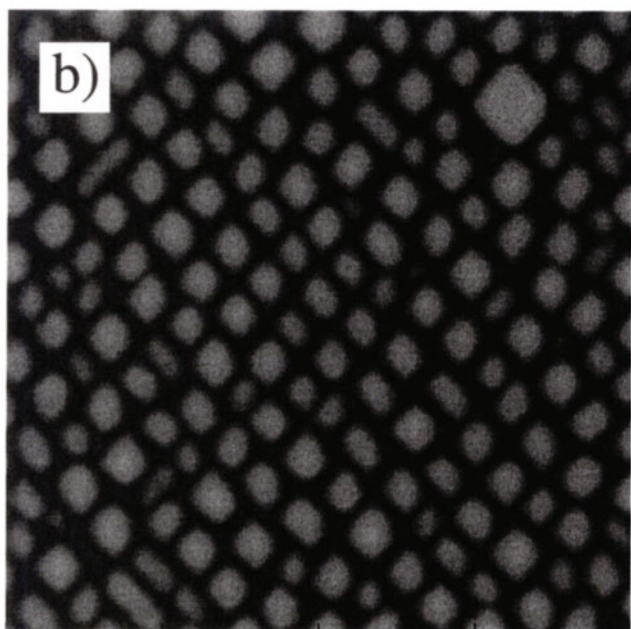
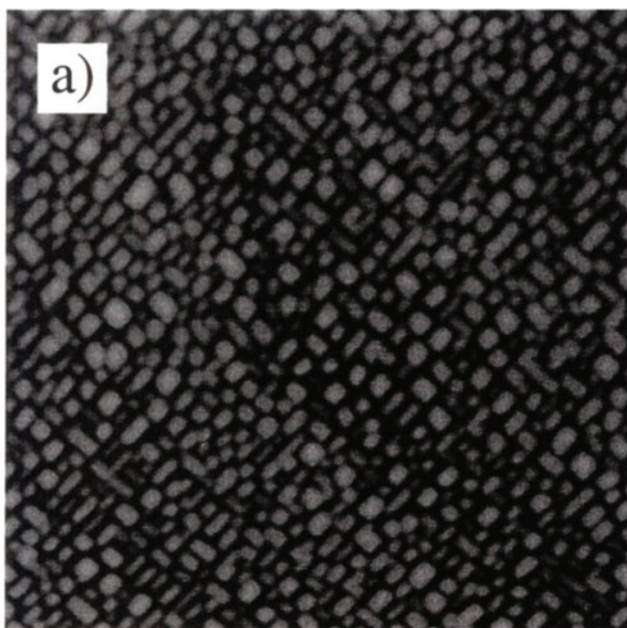


Figure 6.4: AFM images of (a) the first and (b) the 25th layer of a SiGe dot superlattice (taken from [13], courtesy of C. Teichert).

the experimental assessment of the position dependence of both the Ge concentration and the lattice parameter variations is desirable. Since the structures are very small and the parameters can change over distances equal to a few times the lattice constant, this is an extremely difficult task. Only a combination of several independent characterization methods is likely to have success. In the following, we describe several methods that have been applied in order to study different aspects of Ge-dot ensembles.

6.3.2 Atomic-force and scanning tunneling microscopy

Atomic-force microscopy is the standard characterization method of structures on semiconductor surfaces. It yields the two-dimensional surface profiles of dot ensembles and hence parameters like the density, the diameter and height, the shape and facets of uncapped self-organized dots. Moreover, the extent of ordering of the dots can easily be investigated. However, the investigations are more or less restricted to uncapped dots, the shape of which may change during overgrowth with a cap layer, which is necessary to make use of any optical or electronic properties of these structures.

The surface modulation produced by Ge dots decreases rapidly after overgrowth with silicon, as was found by Schittenhelm [10]. The height of dots evolving after the deposition of nominally 6 ML Ge was determined by AFM as 116 Å. After growth of 30 Å Si, the surface modulation was only about 4 nm, while after deposition of 100 Å silicon, it was about 15 Å. This is an indication for the occurrence of lateral segregation and the diffusion of Ge during overgrowth with silicon.

Theiss et al. have studied the onset of plastic relaxation of Ge islands on Si[111] by scanning tunneling microscopy (STM) [20]. The in-plane lattice constant at the surface and on top of the islands has been measured as a function of the island height. The authors identified three stages of different relaxation behavior: for heights below 50 ML, the in-plane lattice constant is rapidly increasing to a value even above the Ge bulk value; in the region between 50 and 80 ML, the relaxation at the surface decreases again to about 2% below the bulk value; and for island heights above 80 ML, the surface lattice spacing approaches the bulk value again.

Voigtländer and Zinner have investigated the growth of Ge on Si[111] in situ with an STM and recorded the time evolution of the Stranski-Krastanow growth process [21]. Thus the complete evolution of the morphology can be followed during the island growth. With increasing coverage, the aspect ratio of a single Ge island first increases and after passing

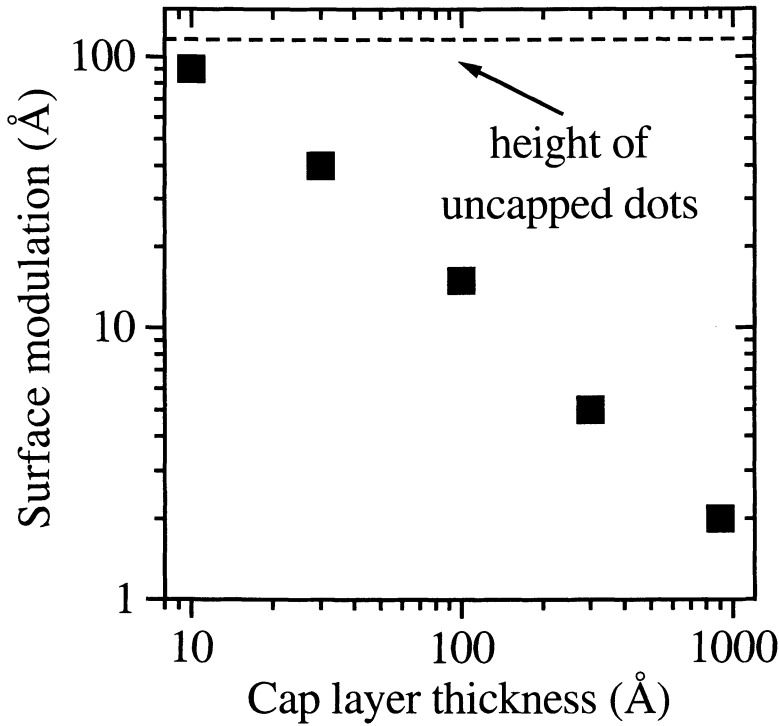


Figure 6.5: Surface modulation versus deposited Si cap-layer thickness (taken from [10]).

through a maximum, decreases again. This inversion of the aspect ratio was identified as the transition from coherent to dislocated 3D islands, i.e., when a critical height for dislocation formation is reached, the coherent islands transform into dislocated islands. Consequently, in the partially relaxed islands, the average strain is diminished, and thus the driving force for high aspect ratio decreases with further coverage.

6.3.3 Transmission electron microscopy

Transmission electron microscopy (TEM) provides image contrast both for material composition and strain and is therefore sensitive both to the (position-dependent) composition and the elastic relaxation of the islands. Moreover, cross-section TEM allows for the study of the vertical positional correlation of the Ge islands. Figure 6.6 shows the perfect vertical stacking of 19 individual Ge dot layers with a Si spacer layer thickness of about

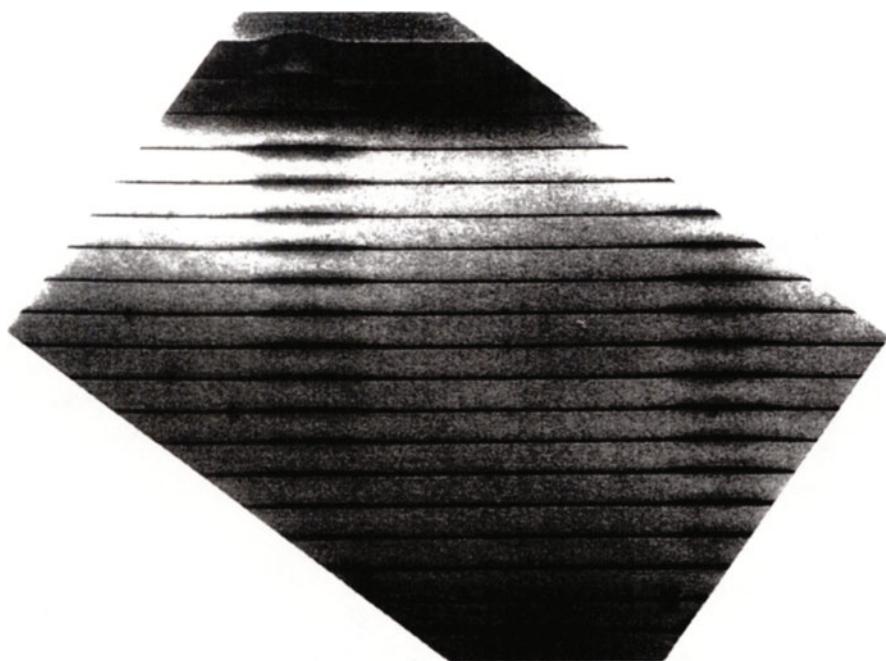


Figure 6.6: Cross-section TEM of a 19-period Si/Ge dot multilayer.

400 Å and nominally deposited 5.5 ML of Ge in each of the periods. The dot height increases slightly towards the sample surface, whereas the wetting-layer thickness decreases at the same time. The height of the uncapped dots at the sample surface is approximately twice as large as in the last overgrown dot layer, which is consistent with the AFM observations presented in the last subsection.

Plan-view TEM allows the investigation of lateral ordering, even of capped dot layers. The influence of process steps like annealing on the shape and composition of the islands can be studied [22]. Mateeva et al. performed cross-section TEM investigations of Ge-dot multilayers. A vertical merging of two smaller islands into a single larger one in subsequent layers and the changes of the dot shape after overgrowth have been observed [23].

Ruvimov and Scheerschmidt have investigated the high-resolution TEM contrast of pyramidal-shaped quantum dots and have shown that the overgrown pyramidal islands always seem to be truncated due to the strong strain fields and the lower concentration at the top of the dots [24].

6.3.4 Photoluminescence

Photoluminescence (PL) is sensitive to the band alignment, which depends strongly on both the local composition and the strain status of SiGe. Figure 6.7 shows PL spectra of self-assembled Ge-dot layers and the corresponding dependence on the deposited number of Ge monolayers. The peaks around 1 eV stem from the wetting layer. They are shifting towards lower energies from 4 to 5 ML because of decreasing quantum confinement in the two-dimensional layer [25]. However, after the formation of the first dots, which occurs at about 5.33 ML, the wetting-layer peaks exhibit a blueshift because of the thinning of the wetting layer. The dot peaks at about 0.9 eV become predominant for nominal Ge-layer thicknesses of 5.5 ML and above. The redshift between the curves for 6 and 12 MLs of Ge is probably related to incoherent strain relaxation due to dislocations. In Fig. 6.8, the peak positions of both 3D islands and the wetting layer are depicted, which allows for a precise determination of the thickness of the two-dimensional wetting layer.

The physical nature of the carrier recombination processes in Ge dots has been studied recently by Fukatsu et al. [26]. Phononless radiative recombination across the indirect bandgap was observed. Three-dimensional confinement of electrons in an electronic dot is found to effectively remove the momentum conservation rule for interband transitions.

6.3.5 Raman scattering

As mentioned before, the position of the PL peaks is sensitive both to the strain status and the composition of the dots. Persans et al. have tried to separate strain and compositional effects in Ge dots on Si[111] by combining electroreflectance spectroscopy and Raman scattering [27]. The Raman peak positions of both the Si-Ge and the Ge-Ge modes depend on the strain status and on the composition of the SiGe layers. For cubic and biaxially strained SiGe, the dependence of the peak positions on composition has been determined by Alonso and Winer [28] and Cerdeira et al. [29]. In the case of partially relaxed Ge islands, the strain distribution is strongly inhomogeneous and position dependent, and the shear terms can be comparable to the normal strains. It is therefore difficult to relate the peak positions to the strain and composition inside the dots.

6.3.6 X-ray diffraction and reflectivity

X-ray scattering is sensitive to the lattice constants of crystalline layers and also to local changes in the strain state. Moreover, lateral and vertical

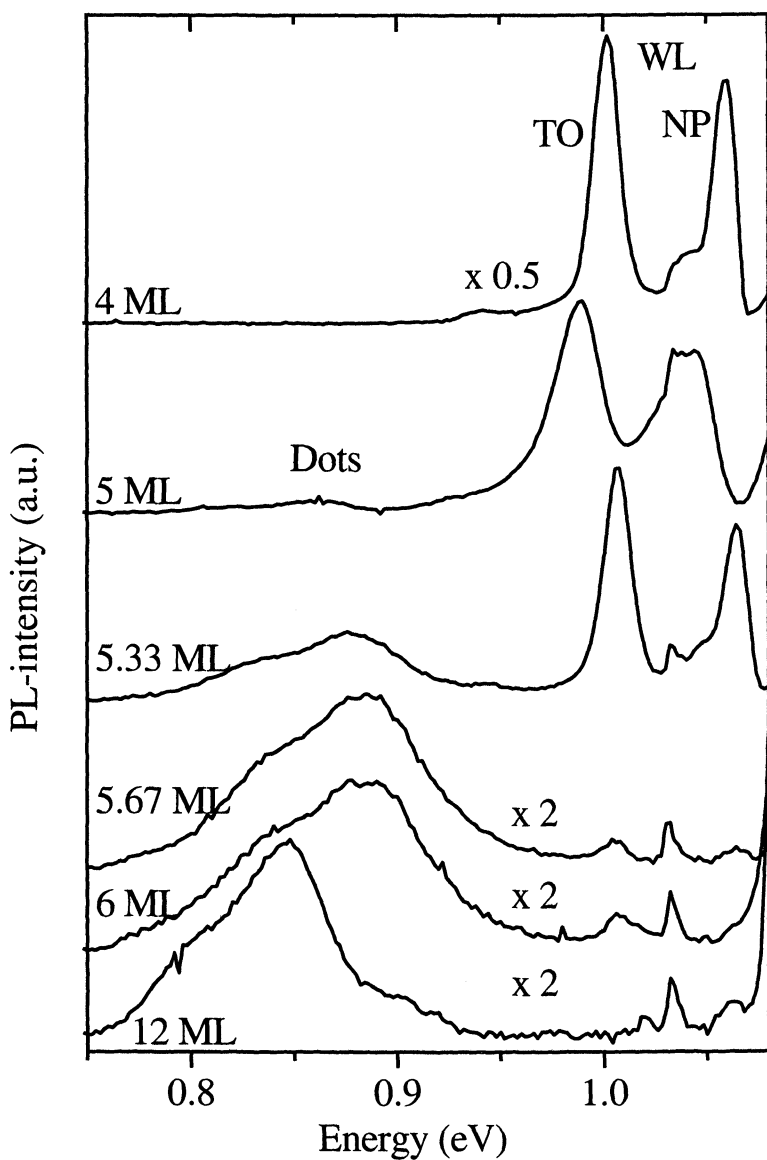


Figure 6.7: Photoluminescence spectra of Ge-dot layers as a function of the nominal Ge coverage (taken from [10]).

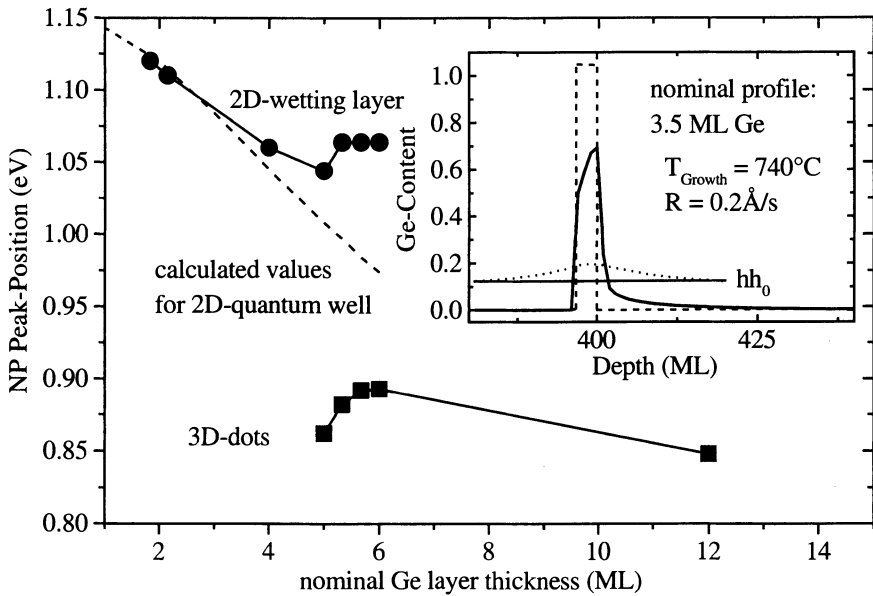


Figure 6.8: Photoluminescence peak positions as a function of the nominal Ge coverage (taken from [10]).

correlations of positions and shapes of inhomogeneities can be studied with a high statistical reliability because of the large illuminated areas. With the method of reciprocal space mapping, a two- or even three-dimensional distribution of scattered intensity in the reciprocal space is measured, which is influenced by the size, shape, composition, and lateral and vertical arrangement of the dots. Since the acquired information is not only the energetical position of a single optical transition but also a multi-dimensional distribution, the potential for the assessment of the modulation of the Ge concentration and the strain fields is given.

Conventional X-ray-diffraction reciprocal space maps of self-assembled Ge dots were first studied by Baribeau [30]. Williams et al. have studied the height, the relaxation, and the annealing behavior of uncovered Ge islands grown on Si[001] using in situ grazing incidence diffraction (GID) [31]. More recently, Steinfort et al. have employed GID along with a phenomenological model for the determination of the height dependence of the elastic relaxation in uncovered Ge-hut clusters grown on Si[001] substrates at a temperature of 430°C [32].

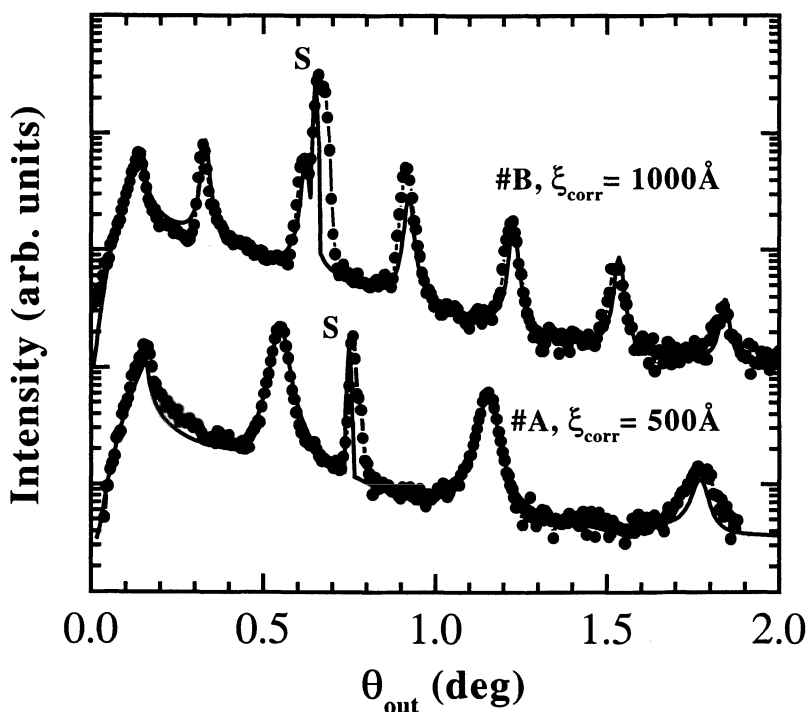


Figure 6.9: Measured (dotted lines) and simulated (full lines) 2θ -scans of Ge-dot multilayers with 10nm (lower curve) and 20nm (upper curve) Si spacer layers. The peaks labeled “S” are at positions where the trajectory of the scan crosses the rod of specular reflection.

6.3.6.1 X-ray reflectivity

X-ray reflectivity is sensitive only to variations of the electron density and hence to modulations of the composition; it is not obscured by the strain distribution in the investigated structures. It can be used for the determination of layer thicknesses, Ge concentrations, and the rms values of the interface roughness. Moreover, resonant diffuse scattering (RDS) [33] allows for the study of the vertical and lateral correlation properties of the interface profiles. Figure 6.9 presents measurements (dotted line) and simulations (full line) of the diffuse scattering of Ge-dot multilayers formed after repeated deposition of 5.5ML Ge at a growth temperature of 670°C. The lower curve refers to a sample with 100 Å Si spacer layers between the dot layers, whereas the upper curve corresponds to a sample with 20nm spacers. A series of peaks is visible, which would be absent if the interface

profiles were totally uncorrelated. The smaller the vertical correlation length ξ_{corr} , the broader are these RDS maxima. The numerical values for ξ_{corr} are $1,000 \pm 200 \text{ \AA}$ for the $200 \text{ \AA}$ spacer sample and $500 \pm 200 \text{ \AA}$ for the $100 \text{ \AA}$ spacer sample.

The measured data indicate a clear tendency towards an increase of ξ_{corr} with increasing thickness of the Si spacer layers. The value of ξ_{corr} not only reflects the positional correlation of the dots but is also affected by a vertical inhomogeneity of the dot sizes. If, for instance, the positions of the dot centers were fully vertically correlated ($\xi_{\text{corr}} \rightarrow \infty$) but the sizes of the dots increased towards the free surface, then a smaller value of ξ_{corr} would be determined from the measured diffuse maxima. Cross-section TEM revealed a good vertical correlation of the dot positions in all samples and a distinct gradient of the dot sizes in the samples with 100 and $200 \text{ \AA}$ Si spacer layers. Consequently, as for the latter, the value of ξ_{corr} is rather affected by the size gradient.

6.3.6.2 X-ray diffraction

Double-crystal diffractometry is the standard characterization method for epitaxially grown semiconductor layers. It yields reliable values for the layer thicknesses and compositions of laterally homogeneous systems. However, for the characterization of laterally inhomogeneous structures such as etched or self-organized quantum dots, which exhibit an essentially three-dimensional distribution of diffracted intensity in the reciprocal space, it is inappropriate, since only a one-dimensional projection of the intensity distribution perpendicular to the diffraction plane and along the Ewald sphere is measured.

Two-dimensional reciprocal space maps (RSMs) of a 10-period Si/Si_{0.5}Ge_{0.5} dot multilayer grown by MBE at 600°C and a growth rate of $0.13 \text{ \AA/s}$ are shown in Fig. 6.10. In order to discriminate between features due to the relative arrangement and features due to the strain relaxation of the dots, we have acquired several asymmetric RSMs. In Fig. 6.10a, S denotes the Si substrate peak, A the analyzer streak—an instrumental artefact, and SL_0 and SL_1 the zeroth and first-order superlattice satellites, respectively. The presence of the dots manifests itself in the occurrence of broad peaks of diffuse scattering D_0 and D_1 on the left- and right-hand sides of SL_0 and SL_1 . The spacing between D_0 and SL_0 is identical in the (113) and (224) RSMs, whereas it increases by approximately a factor of $\sqrt{2}$ in the (404) RSM. Therefore, the peak D_0 is attributed to a lateral (short-range) ordering of the dots in a square array with main axes along the $\langle 100 \rangle$ directions, which are directions of minimal stiffness. A “dot-lattice constant” of $1,400 \text{ \AA}$ was deduced.

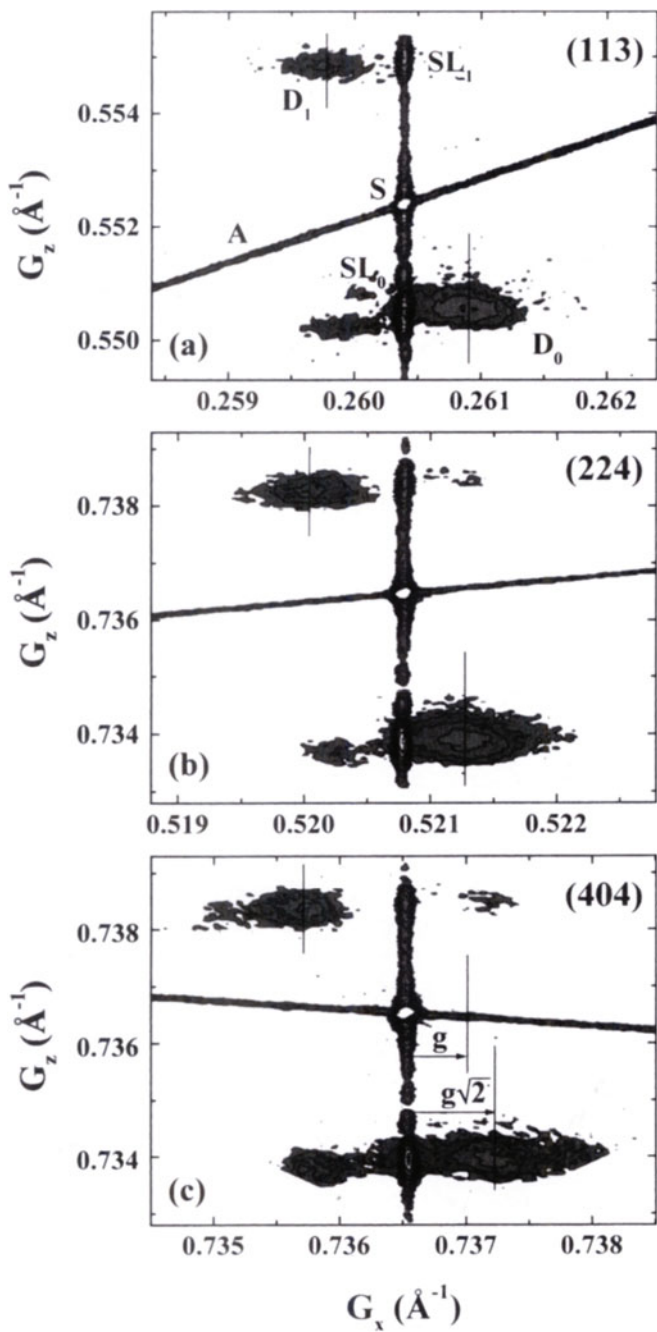


Figure 6.10: Asymmetric reciprocal space maps of a Si/Si_{0.5}Ge_{0.5} dot multilayer. See text for discussion.

The spacing between the peaks labeled D_1 and SL_1 increases with increasing in-plane component $\sqrt{h^2 + k^2}$ of the scattering vector (hkl) and thus is influenced by the elastic relaxation of the dots and the strained region above and below them. In order to extract quantitative information from the reciprocal space maps, a model for the simulation of the X-ray intensity distribution has been developed [34]. The input parameters for this model are the island shape and size and the Ge content, which is assumed to be constant. The resulting quantities are the average relaxation in the dots, the average dot spacing, and the dispersion of the positional distribution of the dots, which is a measure of the range of lateral order in the dot array, as well as the vertical correlation length.

In Fig. 6.11, measured and simulated RSMs of a 19-period Si/Ge dot multilayer grown by MBE at a temperature of 670°C and a deposition rate of 0.075 Å/s are shown. The nominal Ge coverage was 5.5 ML; the thickness of the Si spacer layers was about 300 Å. The diameter of the islands was approximately 1,500 Å and their height about 7 nm. Because of the small ratio of height to width, the extent of elastic relaxation is expected to be small. A value of $\Delta\epsilon_{xx} = (1.4 \pm 0.1) \cdot 10^{-3}$ was deduced from the simulations, which is in good agreement with the estimated value predicted by the isotropic cylinder model presented above.

6.3.6.3 Grazing incidence diffraction

Conventional X-ray diffraction is a coplanar technique, where the normal to the diffracting plane and the incident and the diffracted wavevectors share a common plane. In grazing incidence diffraction, the diffracting lattice planes are perpendicular to the sample surface and the scattered wave vectors are probed in an exit angle range from $\alpha_f = 0$ to a few degrees with respect to the sample surface [35,36]. Via variation of the incidence angle α_i and the exit angle α_f , the penetration (or information) depth can be varied from a few angstroms to several thousand angstroms.

Reciprocal space maps close to the reciprocal lattice vector (220) have been measured for different values of α_i using synchrotron radiation [37]. For $\alpha_i = 0.15^\circ$, a planar cut through the intensity distribution is shown in Fig. 6.12a. At this incident angle, only the top layer of the superlattice contributes to the scattering signal if the intensity is integrated over an exit angle of $\alpha_f \leq \alpha_c$. The Bragg peak at $q_r = 0$ stems from the pseudomorphic part of the region close to the surface, and the intensity distribution around $q_r = -0.025 \text{ \AA}^{-1}$ is due to the lattice parameter distribution of the relaxed Ge islands. The elastic relaxation manifests itself in a peak shift in radial direction q_r towards the origin of reciprocal space and is expected to be largest for the topmost layer, since the surface of the sample is free. Thus, the

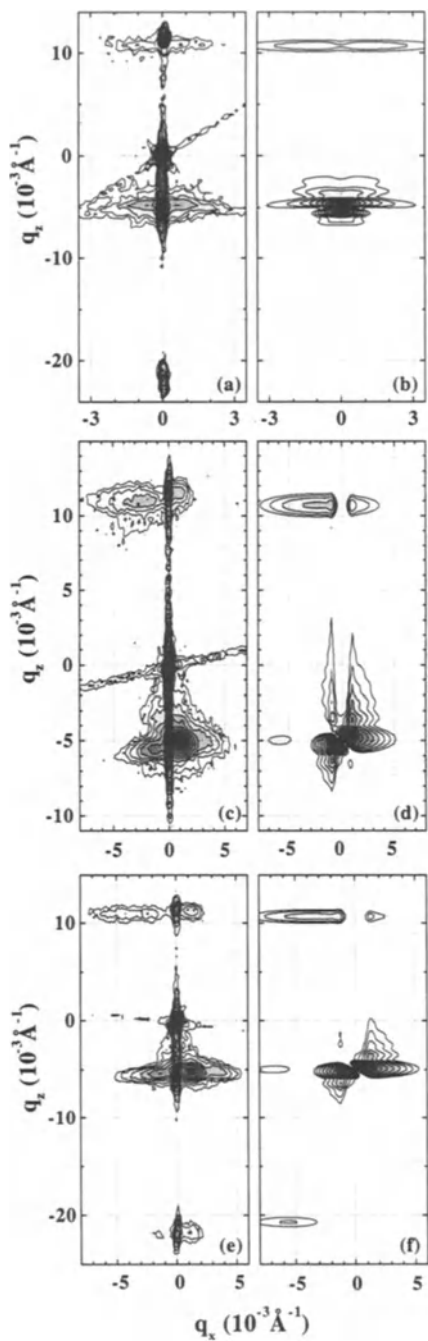


Figure 6.11: Measured (left-hand side) and simulated (right-hand side) reciprocal space maps of a 19-period Si/Ge dot multilayer.

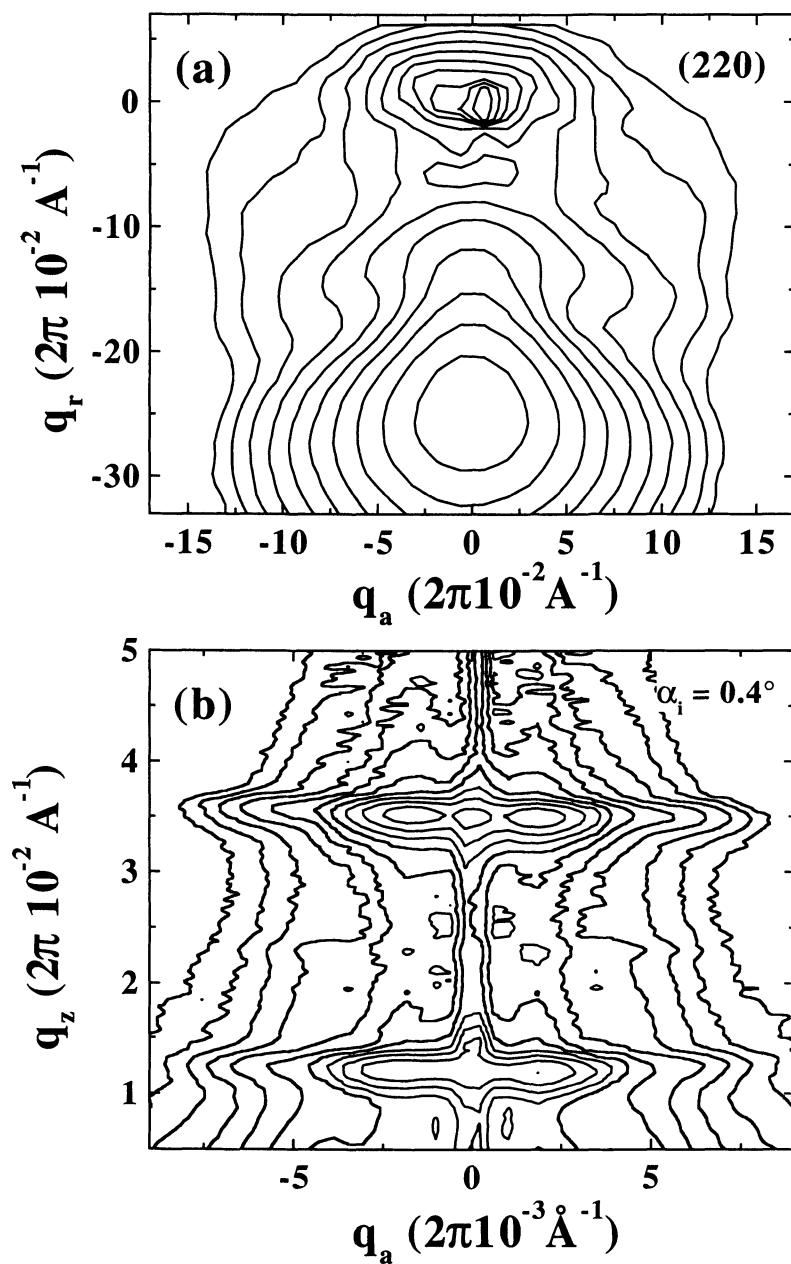


Figure 6.12: Horizontal (a) and vertical (b) grazing-incidence reciprocal space maps of an uncapped 19 period Si/Ge dot multilayer.

mentioned shift should increase with decreasing penetration depth of the X-rays. This behavior is indeed observed in lateral cuts corresponding to different information depths. In order to resolve the satellite maxima expected for laterally correlated dots, high-resolution scans perpendicular to the reciprocal lattice vector have been carried out while simultaneously recording the intensity distribution in growth direction along q_z . The RSM for an incidence angle $\alpha_i = 0.4^\circ$ is shown in Fig. 6.12b. We clearly find the dot-induced peaks on both sides of the (220) rod at $q_a = \pm 0.0018 \text{ \AA}^{-1}$, which corresponds to a lateral correlation length $d_{110} = 3,460 \pm 220 \text{ \AA}$. From the measurement close to the (004) reflection, we find a correlation length of $d_{001} = 4,900 \pm 200 \text{ \AA}$, again indicating that the dots form roughly a two-dimensional square lattice oriented along the [100] direction. If we consider the q_z dependence of the dot correlation peaks at $q_a = \pm 0.0018 \text{ \AA}^{-1}$ in Fig. 6.12b, strong maxima appear at the same q_z positions as the superlattice peaks. This means that the lateral location of the Ge islands in the individual layers is highly replicated during the growth process, at least in the upper 10 bilayers probed with the chosen penetration depth.

Using GID, one can visualize the reciprocal lattice of the dots. The statistical parameters of the dot distribution, such as the lateral and vertical ordering, mean dot distance and the root mean square deviations can be determined for a comparatively large sample area [38]. Such information is indeed required for any possible applications of Ge dots embedded in a matrix and complements the information obtainable from transmission electron microscopy.

Recently, grazing-incidence small-angle X-ray scattering (GISAXS), which combines small-angle X-ray scattering and total external reflection, has been used to discriminate between the shape of uncovered Ge islands and their depth-dependent strain relaxation, i.e., the strain relaxation as a function of the distance from their top [39,40]. This new technique is a powerful new tool for the determination of in-plane electron density fluctuations in the nanometer range and allows for studies of the recently observed shape transitions of Ge nanocrystals [41]. The strain data from GISAXS experiments on uncovered dots can be easily compared to results of finite element calculations [42].

6.4 Future concepts

6.4.1 Infrared detectors

The self-organization of Ge dots provides a possibility to overcome the limit of the critical thickness and to fabricate structures with a Ge concentration higher than possible in two-dimensional systems without the

nucleation of misfit dislocations or structural defects. This decreases the bandgap locally; hence, the sensitivity of SiGe-based interband infrared detectors can be extended to longer wavelengths. However, this desirable compositional effect is counteracted by the influence of elastic relaxation on the bandgap. Since the latter increases with increasing relaxation, it is advantageous for applications such as infrared detectors to fabricate large and flat dots that do not relax significantly. Vertical stacking of dot layers provides both for an increased volume for absorption and for a further reduction of the effective bandgap via vertical coupling of the dot energy levels. It is thus possible to reduce the effective bandgap of essentially strained SiGe structures to about 0.55 eV, which is well below the bandgap of cubic Ge [43].

Figure 6.13 shows a comparison of photocurrent spectra of a (two-dimensional) three-period SiGe multi-quantum well and a three-period Ge-dot multilayer with a nominal coverage of 12 ML Ge. The photocurrent increases substantially, i.e., by a factor of 10, in the wavelength range between 0.8 and 1.2 eV in the npn structure with a Ge-dot layer in the p-base region, as shown by Schittenhelm et al. [44]. The onset of the photocurrent is shifted to smaller energies, which indicates a smaller bandgap in the sample with the dot layers. It should be pointed out that this infrared photodetector based on Ge dots works with high external efficiency. For the detection of radiation with longer wavelengths, intersub-band transitions could be used as well.

6.4.2 Enhanced light emission from Si-based heterostructures

Brunner et al. demonstrated that the introduction of carbon can significantly enhance the luminescence efficiency of the indirect bandgap SiGe material system [45]. SiGe provides confinement mainly for holes, whereas the band offset in $\text{Si}_{1-y}\text{C}_y$ occurs mainly in the conduction band. Therefore, the growth of neighboring confinement $\text{Si}_{1-x}\text{Ge}_x/\text{Si}_{1-y}\text{C}_y$ or $\text{Si}_{1-x-y}\text{Ge}_x\text{C}_y/\text{Si}_{1-y}\text{C}_y$ structures increases the wave function overlap of the carriers and hence the luminescence intensity. Recently, it was shown by the same group that a fractional monolayer of carbon deposited prior to Ge epitaxy changes the self-organization of islands dramatically [4,46]. The critical thickness for Ge-dot formation is reduced, and the dots themselves are much smaller since carbon reduces the surface diffusion lengths. Typical diameters and heights are 10 Å and 1 Å, respectively, for a growth temperature of 550°C. Moreover, the carbon layer provides for electron confinement very close to the Ge islands, where holes are captured. In this fashion, the strongest luminescence ever reported in the SiGeC material system has been achieved at an energy of 0.99 eV (see Fig. 6.14), which is very close

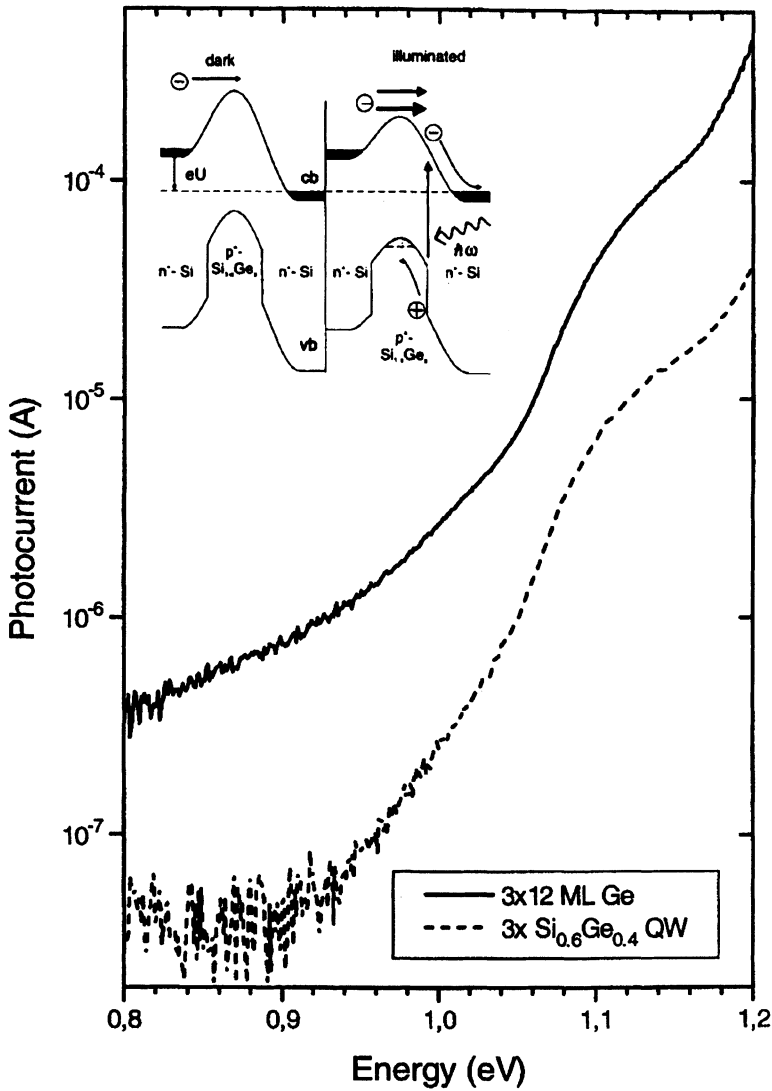


Figure 6.13: Photocurrent spectra of a SiGe quantum well and Ge-dot phototransistors.

to the technologically relevant region around $1.3\mu\text{m}$. However, as shown by Schmidt et al. [47], in the carbon-induced Ge-dot multilayers no vertical alignment of the dots is observed due to a kinetically limited islanding process.

Figure 6.14: Photoluminescence spectra from a $\text{Si}_{0.84}\text{C}_{0.16}/\text{Si}_{0.988}\text{C}_{0.012}$ double quantum well (DQW) structure with $d_{\text{SiGe}} = 1.2 d_{\text{SiC}} = 27 \text{ \AA}$, a single layer of 0.2 ML carbon-induced 2.2 ML Ge quantum dots, and a 50-period 0.2 ML C/2.4 ML Ge sample. The Si spacer layer thickness in the multilayer sample is $96 \text{ \AA}$. Inset: Schematic band alignment of neighboring confinement $\text{Si}_{1-y}\text{C}_y/\text{Si}_{1-x}\text{Ge}_x$ quantum wells and for C-induced Ge dots. Gradual variation of the alloy composition and the strain of the C-induced dots is assumed. The dots cause an additional strong lateral carrier localization (taken from [46]; courtesy of K. Eberl).

6.5 Conclusions

We have presented a survey of methods used for the characterization of self-organized Ge dots ranging from imaging techniques and optical methods to X-ray scattering. Each method addresses a particular structural aspect of the Ge islands, and only a combination is likely to allow reliable conclusions. The ultimate goal, namely, the determination of the position dependence of the composition and the local strain fields in buried dots, has not been reached yet. Attempts to obtain locally resolved information on the chemical composition with scanning electron microscopy using energy-dispersive X-ray analysis are in progress [49]. A spatial resolution of about 3 nm could be achieved so far, and through the use of EELS this value could be even reduced further. Other attempts to reveal the depth-dependent chemical composition of buried islands based on the analysis of digitized high-resolution TEM images were published recently [50].

ACKNOWLEDGMENTS

The X-ray diffraction work was carried out in collaboration with V. Holý, J. Stangl, T. H. Metzger, I. Kegel, and U. Pietsch, and we thank them for their most valuable contributions. We thank A. O. Kosogov and P. Werner for the TEM work, and are grateful to C. Teichert (Montan-Universität Leoben) and K. Eberl (Max-Planck-Institut für Festkörperforschung) for providing Figs. 6.4 and 6.14.

REFERENCES

- [1] D. J. Eaglesham and M. Cerullo, Phys. Rev. Lett. **64**, 1943 (1990).
- [2] Y.-W. Mo, D. E. Savage, B. S. Swartzentruber, and M. G. Lagally, Phys. Rev. Lett. **65**, 1020 (1990).

- [3] N. N. Ledentsov, Proc. 8th Int. Conf. on the Physics of Narrow Gap Semiconductors (Shanghai, 1997), World Scientific (Singapore, 1998).
- [4] O. G. Schmidt, C. Lange, K. Eberl, O. Kienzle, and F. Ernst, Appl. Phys. Lett. **71**, 2340 (1997).
- [5] R. Apetz, L. Vescan, A. Hartmann, C. Dieker, and H. Lüth, Appl. Phys. Lett. **66**, 445 (1995).
- [6] See, e.g., H. Sunamura, N. Usami, Y. Shiraki, and S. Fukatsu, Appl. Phys. Lett. **66**, 3024 (1995).
- [7] See, e.g., P. Schittenhelm, M. Gail, J. Brunner, J. F. Nützel, and G. Abstreiter, Appl. Phys. Lett. **67**, 1292 (1995).
- [8] P. Schittenhelm, G. Abstreiter, A. Darhuber, G. Bauer, P. Werner, and A. Kosogov, Thin Solid Films **294**, 291 (1997).
- [9] C. K. Seal, D. Samara, and S. K. Banerjee, Appl. Phys. Lett. **71**, 3564 (1997).
- [10] P. Schittenhelm, Ph.D. thesis, Walter Schottky Institut, Technical University Munich (1997).
- [11] J. Tersoff, C. Teichert, and M. G. Lagally, Phys. Rev. Lett. **76**, 1675 (1996); C. Teichert, M. G. Lagally, L. J. Peticolas, J. C. Bean, and J. Tersoff, Phys. Rev. B **53**, 16334 (1996).
- [12] Q. Xie, A. Madhukar, P. Chen, and N. P. Kobayashi, Phys. Rev. Lett. **75**, 2542 (1995).
- [13] C. Teichert, Y. H. Phang, L. J. Peticolas, J. C. Bean, and M. G. Lagally, *Surface Diffusion: Atomistic and Collective Processes*, edited by M. C. Tringides, NATO ASI Series (Plenum Press, New York, 1997).
- [14] V. Holý, J. Stangl, S. Zerlauth, G. Bauer, N. Derowski, D. Lübbert, and U. Pietsch, J. Phys. D: Appl. Phys. **32**, A234 (1999).
- [15] V. Holý, A. A. Darhuber, J. Stangl, G. Bauer, J. Nützel, and G. Abstreiter, Phys. Rev. B **57**, 12435 (1998).
- [16] Feng Liu, J. Tersoff, and M. G. Lagally, Phys. Rev. Lett. **80**, 1268 (1998).
- [17] J. Zhu, K. Brunner, and G. Abstreiter, Appl. Phys. Lett. **73**, 620 (1998).
- [18] C. Teichert, J. Tersoff, and M. G. Lagally, in *Morphological Organization in Epitaxial Growth and Removal*, edited by Z. Zhang and M. G. Lagally (World Scientific, Singapore, 1998, in press).
- [19] F. Liu and M. G. Lagally, Surf. Sci. **386**, 169 (1997).
- [20] S. K. Theiss, D. M. Chen, and J. A. Golovchenko, Appl. Phys. Lett. **66**, 448 (1995).
- [21] B. Voigtländer and A. Zinner, Appl. Phys. Lett. **63**, 3055 (1993).
- [22] A. O. Kosogov, P. Werner, U. Gösele, N. N. Ledentsov, D. Bimberg, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, P. S. Kop'ev, N. A. Bert, and Zh. I. Alferov, Appl. Phys. Lett. **69**, 3072 (1996).

- [23] E. Mateeva, P. Sutter, J. C. Bean, and M. G. Lagally, *Appl. Phys. Lett.* **71**, 3233 (1997).
- [24] S. Ruvimov and K. Scheerschmidt, *Phys. Stat. Sol. (a)* **150**, 471 (1995).
- [25] G. Abstreiter, P. Schittenhelm, C. Engel, E. Silveira, A. Zrenner, D. Meertens, and W. Jäger, *Semicond. Sci. Technol.* **11**, 1521 (1996).
- [26] S. Fukatsu, H. Sunamura, Y. Shiraki, and S. Komiyama, *Appl. Phys. Lett.* **71**, 258 (1997).
- [27] P. D. Persans, P. W. Deelman, K. L. Stokes, L. J. Schowalter, A. Byrne, and T. Thundat, *Appl. Phys. Lett.* **70**, 472 (1997).
- [28] M. I. Alonso and K. Winer, *Phys. Rev. B* **39**, 10056 (1989).
- [29] F. Cerdeira, A. Pinczuk, J. C. Bean, B. Batlogg and B. A. Wilson, *Appl. Phys. Lett.* **45**, 1138 (1984).
- [30] J.-M. Baribeau, *J. Cryst. Growth* **157**, 52 (1995).
- [31] A. A. Williams, J. M. C. Thornton, J. E. Macdonald, R. G. van Silfhout, J. F. van der Veen, M. S. Finney, A. D. Johnson, and C. Norris, *Phys. Rev. B* **43**, 5001 (1991).
- [32] A. J. Steinfert, P. M. Scholte, A. Ettema, F. Tuinstra, M. Nielsen, E. Landemark, D.-M. Smilgies, R. Feidenhans'l, G. Falkenberg, L. Seehofer, and R. L. Johnson, *Phys. Rev. Lett.* **77**, 2009 (1996).
- [33] V. Holý and T. Baumbach, *Phys. Rev. B* **49**, 10668 (1994).
- [34] A. A. Darhuber, P. Schittenhelm, V. Holý, J. Stangl, G. Bauer, and G. Abstreiter, *Phys. Rev. B* **55**, 15652 (1997).
- [35] U. Pietsch, H. Metzger, S. Rugel, B. Jenichen, and I. K. Robinson, *J. Appl. Phys.* **74**, 2381 (1993).
- [36] G. T. Baumbach, S. Tixier, U. Pietsch, and V. Holý, *Phys. Rev. B* **51**, 16848 (1995).
- [37] A. A. Darhuber, V. Holý, P. Schittenhelm, J. Stangl, I. Kegel, Z. Kovats, T. H. Metzger, G. Bauer, G. Abstreiter, and G. Grübel, *Physica E* **2**, 789 (1998).
- [38] V. Holý, A. A. Darhuber, J. Stangl, S. Zerlauth, F. Schäffler, G. Bauer, N. Darowski, D. Lübbert, U. Pietsch, and I. Vavra, *Phys. Rev. B* **58**, 7934 (1998).
- [39] I. Kegel, T. H. Metzger, P. Fratzl, J. Peisl, A. Lorke, J. M. Garcia, and P. M. Petroff, *Emophys. Lett.* **45**, 222 (1999).
- [40] M. Schmidbauer, T. Wiebach, H. Raidt, M. Hanke, R. Köhler, and H. Wawra, *Phys. Rev. B* **58**, 10523 (1998).
- [41] G. Medeiros-Ribeiro, A. M. Bratkovski, T. I. Kamins, D. A. A. Ohlberg, and R. S. Williams, *Science* **279**, 353 (1998).
- [42] S. Christiansen, M. Albrecht, H. P. Strunk, and H. J. Maier, *Appl. Phys. Lett.* **64**, 3617 (1994).

- [43] K. Eberl and W. Wegscheider, *Handbook of Semiconductors*, Vol. 3a, edited by P. S. Moss (Elsevier Science, Amsterdam 1994), p. 595.
- [44] P. Schittenhelm, C. Engel, F. Findeis, G. Abstreiter, A. A. Darhuber, G. Bauer, A. O. Kosogov, and P. Werner, *J. Vac. Sci. Technol. B* **16**, 1575 (1998).
- [45] K. Brunner, W. Winter, and K. Eberl, *Appl. Phys. Lett.* **69**, 1279 (1996).
- [46] K. Eberl, O. G. Schmidt, S. Schieker, N. Y. Jin-Phillip, and F. Phillipp, *Solid-State Electron.* **42**, 1593 (1998).
- [47] O. G. Schmidt, S. Schieker, K. Eberl, O. Kienzle, and F. Ernst, *Appl. Phys. Lett.* **73**, 659 (1998).
- [48] J. Bevk, A. Ourmazd, L. C. Feldman, T. P. Pearsall, J. M. Bonar, B. A. Davidson, and J. P. Mannaerts, *Appl. Phys. Lett.* **50**, 760 (1987).
- [49] O. Kienzle, private communication.
- [50] M. Strassburg, V. Kutzer, U. W. Pohl, A. Hoffmann, I. Broser, N. Ledentsov, D. Bimberg, A. Rosenauer, U. Fischer, D. Gerthsen, I. L. Krestnikov, M. V. Maximov, P. S. Kop'ev, and Zh. I. Alferov, *Appl. Phys. Lett.* **72**, 942 (1998).

Index

Page numbers in *italics* indicate figures. Page numbers followed by “t” indicate tables.

- AFM. *See* Atomic force microscopy
- Alkanethiols, in lithography, 128
- Applications of quantum nanocircuits, 12–16
 - logic, 12–13
 - metrology, sensors, 15–16
 - quantum computation, 13–14
- Atomic-force, scanning tunneling microscopy, self-organized Ge islands, 239–240
- Atomic force microscopy
 - lithography, 123
 - self-formed quantum dot structures, 37
- Average delay, *vs.* device density, quantum nanocircuits, 5
- Background charges, quantum nanocircuits, 10
- Band filling, self-formed quantum dot structures, 24
- Bottleneck, phonon, self-formed quantum dot structures, 62, 74
- CAR resist, 114–118, *116*
- Charging energy, quantum nanocircuits, 3
- Chemically amplified resists, 115–116, *117*
- Closely stacked InAs quantum dots, self-formed quantum dot structures, 45–55, *46–55*
- CMOS semiconductor elements, replacement of, by quantum devices, 1–13
- Columnar InAs quantum dots, self-formed quantum dot structures, 55–59, *57–61*
- Communication wavelengths, self-formed quantum dot structures, 74
- Confinement energy, quantum nanocircuits, 3
- Conformal phase-shift lithography, 121
- Crystal growth, self-assembled InGaAs quantum dots, 187–205
- Crystals, in lithography, for nanostructure fabrication, 131–132, *132*
- Deep UV lithography, 103–105
 - for nanostructure fabrication, 103
 - photolithography, 103–105
- Density, device, *vs.* average delay, quantum nanocircuits, 5
- DFB lasers. *See* Distributed feedback lasers
- Dissipation limit, integrated circuits, 5–6
- Distributed feedback laser, self-formed quantum dot structures, 77
- DUV. *See* Deep UV lithography
- EBL. *See* Electron beam lithography
- Electrical characterization of quantum dots, self-assembled InGaAs quantum dots, 211–214
 - capacitance spectroscopy, 211–213, *213*
 - transport, 213
- Electron beam lithography, for nanostructure fabrication, 107
- Electron projection lithography, 110–111
- Epitaxial growth, molecular beam, self-formed quantum dot structures, 37
- EUV. *See* Extremely ultraviolet light

- Extremely ultraviolet light,
photolithography, 105–106
- Fabrication technologies, key issues, 73
- Feedback lasers, distributed, self-formed
quantum dot structures, 77
- Gate widths, quantum nanocircuits, 1
- Ge islands, self-organized
atomic-force, scanning tunneling
microscopy, 239–240
diffraction, grazing incidence, 248–251,
250
infrared detectors, 251–252, 253
light emission, enhanced, from Si-based
heterostructures, 252–255, 254–255
low-pressure chemical vapor
deposition, 234
molecular beam epitaxy, 234
photoluminescence, 242, 243, 244
Raman scattering, 242
self-organized Ge dots, growth of, 234
small-angle x-ray scattering, grazing
incidence, 251
Stranski-Krastanow growth, 233, 234
structural characterization, 233–258
transmission electron microscopy,
240–241, 241
x-ray diffraction, 242–251, 247, 249
x-ray reflectivity, 245, 245–246
- Grazing incidence
diffraction, self-organized Ge islands,
248–251, 250
small-angle x-ray scattering, self-
organized Ge islands, 251
- HDT. *See* Hexadecanethiol
- Heisenberg's uncertainty principle, 6
- Hexadecanethiol
lithography, 128
scanned probe methods of lithography,
128
- Hole-burning, self-assembled quantum
dot, 79–89, 80–82, 84–86, 88
- Imprint lithography, 123, 124–125
- InAs
island growth, self-formed quantum dot
structures, 37–39, 38, 40, 41
- MOVPE, 157t, 157–162, 158, 160, 161
- Incomposition, surface morphology,
among InGaAs films, 31
- Infrared detectors, self-organized Ge
islands, 251–252, 253
- InGaAs quantum dots, self-assembled,
183–231
capacitance spectroscopy, electrical
characterization of quantum dots,
211–213, 213
crystal growth, 187–205
electrical characterization of quantum
dots, 211–214
growth on non-(100) substrates,
202–204, 203
heterostructures, electrical, optical
characteristics of, 185
high-speed quantum dot lasers,
quantum dot lasers, 220–221
longitudinal optical phonon energy, 209
molecular beam epitaxy, 183
optical characterization of quantum
dots, 205–211
organometallic vapor-phase epitaxy, 184
overgrowth of self-assembled quantum
dots, 197–200, 200
photoluminescence, 205–207, 206
polarization-dependent
photoluminescence, 205–207, 206
quantum disks, 203
quantum dot edge-emitting lasers,
214–219, 216–219
quantum dot lasers, 214–222
quantum dot VCSELs, 219–220
reflectance anisotropy spectroscopy,
192
reflection high-energy electron
diffraction, 184
scanning electron microscopy, 192
single layers, quantum dots on GaAs,
190–197, 192–199
single quantum dot photoluminescence,
207–209, 208, 209
Stranski-Krastanow transition, 183
superlattices, quantum dot, 200–202,
201
time-resolved photoluminescence,
209–211, 211
transport, electrical characterization of
quantum dots, 213
- Inhomogeneous broadening, smaller, high
dot density, self-formed quantum dot
structures, 73

- Inorganic resists, 114–118, *116*, *119*
 InP, MOVPE
 on GaInP/GaAs, 162–177
 size homogeneities of MOVPE-grown dots, 173–177, *175*
 Integrated circuits, physical limits, 4–8, *5*
 dissipation limit, 5–6
 quantum fluctuations, 6
 RC delay, 6–7
 relativistic limit, 7–8
- Lasers
 distributed feedback, 77
 edge-emitting quantum dot, 214–219, *216–219*
 high-speed quantum dot, 220–221
 quantum dot, 214–222
 self-formed quantum dot structures, 67–77
 self-organized quantum dots, 68–73, *69–72*
- Lateral dimensions, microelectronic circuits, quantum nanocircuits, 1
- Light emission, enhanced, from Si-based heterostructures, 252–255, *254–255*
- Lithography for nanostructure fabrication, 97–138
 2D crystals, 131–132, *132*
 AFM contact methods, 129–130, *130*
 alkanethiols, such as hexadecanethiol, 128
 CAR, 114–118, *116*
 charged-particle proximity printing, 111–112
 chemically amplified resists, 115–116, *117*
 conformal phase-shift lithography, 121
 deep UV lithography, 103–105
 electron beam lithography, 107
 electron projection lithography, 110–111
 extremely ultraviolet light, 105–106
 fabrication processes, 97–101, *99–101*
 focused ion beam, 112
 future of, 133–134
 imprint lithography, 123, *124–125*
 inorganic resists, 114–118, *116*, *119*
 ion beam projection lithography, 113
 ion beam systems, 112–113
 lithographic process, 101–107
 magneto-optical materials, 130
 masks, 103, *104–105*
 micro-electro-mechanical systems, 100
 microcontact printing, *120*, 120–121
 near-field optics, 130–131
 neutral atom lithography, 113
 octadecyltrichlorosiloxane, 117
 particle beam lithography, 107–113
 photolithography, 101–102
 photopolymerization, 121–123, *122–123*
 polymer, 114–118
 process, 101–107
 Rayleigh's criterion, 102, 105
 resist, photosensitive polymer, 102
 resists, 114–118
 SAM, 114–118, *116*
 SCALPEL, 111–112
 scanned probe methods of lithography, 123–131
 scanning electron systems, 108–110, *109*
 scanning tunneling microscopy, 123–129
 selective oxidation, 125–127, *126–128*
 self-assembled monolayers, 116–117, *118*, 128
 self-assembly, 98
 silylation, 115
 single-atom manipulation, 129
 soft lithography, 118–123
 Tennant's law, 100
 top surface imaging, 115, *116*
 TSI, 114–118, *116*
 x-ray lithography, 106–107
- Logic, quantum nanocircuits, 12–13
- Low-pressure chemical vapor deposition, self-organized Ge islands, 234
- Low-temperature conductances, quantum nanocircuits, 9
- LPCVD. *See* Low-pressure chemical vapor deposition
- Magneto-optical materials, lithography, 130
- Magneto-photoluminescence, self-formed quantum dot structures, 51
- Masks, in lithographic process, 103, *104–105*
- MBE. *See* Molecular beam epitaxy
- MEMS. *See* Micro-electro-mechanical systems
- Metal-organic vapor-phase epitaxy. *See* MOVPE
- Metal-oxide semiconductors, self-formed quantum dot structures, 60

- Metrology, sensors, quantum nanocircuits, 15–16
- Micro-electro-mechanical systems, lithography, 100
- Microcontact printing, 120, 120–121
- Molecular beam epitaxy
self-assembled InGaAs quantum dots, 183
self-formed quantum dot structures, 30, 37
self-organized Ge islands, 234
- MOS. *See* Metal-oxide semiconductors
- MOVPE, quantum structured semiconductors, 139–181
2D–3D transition, phenomenology of, 162t, 162–168, 163, 165, 167, 168
3D islands, 149–154, 150–152, 153t, 154, 169–173, 170–174
InAs, 157, 157t, 157–162, 158, 160, 161
InP, 162–177, 175
nucleation, growth model, 146
principles of, 141, 141–143
quantum dot structures, by self-assembling, 143–154
self-assembling, 144, 145–149, 147
Stranski-Krastanow growth mode, 144, 145
surface energies, conditions and, 153
tertiary-butyl-phosphine, 143
triethylgallium, 143
trimethylaluminum, 143
trimethylaminealane, 143
vapor-phase epitaxy, 139
Volmer-Weber growth mode, 145
- Multiple-layer growth, perpendicular alignment of islands, self-formed quantum dot structures, 39–45, 42–44
- mwCP. *See* Microcontact printing
- National Technology Roadmap for Semiconductors*, 1
- NDR. *See* Negative differential resistance
- Near-field optics, lithography, 130–131
- Negative differential resistance region, self-formed quantum dot structures, 61
- Nonradiative recombination, reduction of, self-formed quantum dot structures, 74
- NSOM. *See* Near-field scanning optical microscopy
- Nucleation, MOVPE, growth model, 146
- Octadecyltrichlorosiloxane, 117
- Offset charges, quantum nanocircuits, 10
- OMVPE. *See* Organometallic vapor-phase epitaxy
- Optical characterization, of quantum dots, self-assembled InGaAs quantum dots, 205–211
- Optical gain, self-formed quantum dot structures, 67–68
- Optical intensity modulator, self-formed quantum dot structures, 77–79, 79
- Optical phonon energy, longitudinal, self-assembled InGaAs quantum dots, 209
- Organometallic vapor-phase epitaxy, self-assembled InGaAs quantum dots, 184
- OTS. *See* Octadecyltrichlorosiloxane
- Particle beam lithography, 107–113
charged-particle proximity printing, 111–112
electron projection lithography, 110–111
focused ion beam, 112
ion beam projection lithography, 113
ion beam systems, 112–113
neutral atom lithography, 113
SCALPEL, 111–112
scanning electron systems, 108–110, 109
- Patterning, for nanostructure fabrication, 97–138
2D crystals, 131–132, 132
AFM contact methods, 129–130, 130
alkanethiols, such as hexadecanethiol, 128
atomic-force microscopy, 123
CAR, 114–118, 116
charged-particle proximity printing, 111–112
chemically amplified resists, 115–116, 117
conformal phase-shift lithography, 121
deep UV lithography, 103–105
electron beam lithography, 107
electron projection lithography, 110–111
extremely ultraviolet light, 105–106
fabrication processes, 97–101, 99–101
focused ion beam, 112
future of lithography, 133–134
future of photolithography, 103–107
imprint lithography, 123, 124–125
inorganic resists, 114–118, 116, 119
ion beam projection lithography, 113
ion beam systems, 112–113

- lithographic process, 101–107
- magneto-optical materials, 130
- masks, 103, 104–105
- micro-electro-mechanical systems, 100
- microcontact printing, 120, 120–121
- near-field optics, 130–131
- neutral atom lithography, 113
- octadecyltrichlorosiloxane, 117
- particle beam lithography, 107–113
- photolithography, 101–102
- photopolymerization, 121–123, 122–123
- polymer, 114–118
- Rayleigh's criterion, 102, 105
- resist, photosensitive polymer, 102
- resists, 114–118
- SAM, 114–118, 116
- SCALPEL, 111–112
- scanned probe methods of lithography, 123–131
- scanning electron systems, 108–110, 109
- scanning tunneling microscopy, 123–129
- selective oxidation, 125–127, 126, 127, 128
- self-assembled monolayers, 116–117, 118, 128
- silylation, 115
- single-atom manipulation, 129
- soft lithography, 118–123
- Tennant's law, 100
- top surface imaging, 115, 116
- TSI, 114–118, 115, 116
- x-ray lithography, 106–107
- Phonon bottleneck, self-formed quantum dot structures, 62, 74
- Photolithography, 101–102
 - deep UV lithography, 103–105
 - extremely ultraviolet light, 105–106
 - future of, 103–107
 - Rayleigh's criterion, 105
 - x-ray lithography, 106–107
- Photoluminescence
 - self-assembled InGaAs quantum dots, 205–207, 206
 - self-organized Ge islands, 242, 243, 244
 - single quantum dot, self-assembled InGaAs quantum dots, 207–209, 208, 209
- Photopolymerization, 121–123, 122–123
- Photosensitive polymer, resist, 102
- Physics of quantum nanocircuits, 2–4
 - charging energy, 3
 - confinement energy, 3
 - environment, 4
 - tunnel barriers, 3–4
- Polymer, 114–118
 - chemically amplified resists, 115–116, 117
 - inorganic resists, 117–118, 119
 - octadecyltrichlorosiloxane, 117
 - photosensitive, resist, 102
 - self-assembled monolayers, 116–117, 118
 - silylation, 115
 - top surface imaging, 115, 116
- Quantum computation, 2
- Quantum disks, self-assembled InGaAs
 - quantum dots, 203
- Quantum dot, 19–95
 - MOVPE, 143–154
 - superlattices, 200–202, 201
- Quantum dot lasers, 67–77
 - high-speed quantum dot lasers, 220–221
 - quantum dot edge-emitting lasers, 214–219, 216–219
 - quantum dot VCSELs, 219–220
 - self-assembled InGaAs quantum dots, 214–222
- Quantum fluctuations, integrated circuits, 6
- Quantum nanocircuits, 1–18
 - applications, 12–16
 - average delay, *vs.* device density, 5
 - background charges, 10
 - charging energy, 3
 - CMOS semiconductor elements,
 - replacement of, by quantum devices, 1–13
 - confinement energy, 3
 - dissipation limit, 5–6
 - environment, 4
 - gate widths, 1
 - Heisenberg's uncertainty principle, 6
 - lateral dimensions, microelectronic circuits, 1
 - logic, 12–13
 - low-temperature conductances, 9
 - metrology, sensors, 15–16
 - offset charges, 10
 - physical limits, integrated circuits, 4–8, 5
 - physics of, 2–4
 - quantum computation, 2, 13–14

Quantum nanocircuits (*Continued*)

- quantum fluctuations, 6
 - RC delay, 6–7
 - relativistic limit, 7–8
 - single-electron, 15
 - single-electron transistors, 8–12, 9, 11
 - transistors, decreasing dimensions, 4
 - tunnel barriers, 3–4
- Raman scattering, self-organized Ge islands, 242
- RAS. *See* Reflectance anisotropy spectroscopy
- Rayleigh's criterion
- lithographic process, 102
 - photolithography, 105
- RC delay, integrated circuits, 6–7
- Recombination, nonradiative, reduction of, self-formed quantum dot structures, 74
- Reflectance anisotropy spectroscopy, self-assembled InGaAs quantum dots, 192
- Reflection high-energy electron diffraction, 37
- self-assembled InGaAs quantum dots, 184
- Relativistic limit, integrated circuits, 7–8
- Resist, 114–118
- CAR, 114–118, 116
 - chemically amplified resists, 115–116, 117
 - inorganic resists, 114–118, 116, 119
 - octadecyltrichlorosiloxane, 117
 - photosensitive polymer, 102
 - polymer, 114–118
 - SAM, 114–118, 116
 - self-assembled monolayers, 116–117, 118
 - silylation, 115
 - top surface imaging, 115, 116
 - TSL, 114–118, 116
- RHEED. *See* Reflection high-energy electron diffraction
- SAM, 114–118, 116
- SCALPEL, particle beam lithography, 111–112
- Scanned probe methods of lithography, 123–131
- AFM contact methods, 129–130, 130
 - alkanethiols, such as hexadecanethiol, 128

- atomic-force microscopy, 123
 - magneto-optical materials, 130
 - near-field scanning optical microscopy, 123, 130
 - scanning tunneling microscopy, 123–129
 - selective oxidation, 125–127, 126–128
 - self-assembled monolayers, 128
 - single-atom manipulation, 129
- Scanning electron microscopy, 192
- Scanning tunneling microscopy, 127–129
- atomic-force, self-organized Ge islands, 239–240
 - lithography, 123–129
- Selective oxidation, lithography, 125–127, 126–128
- Self-assembled 3D islands, MOVPE, sizes, shapes, 149–154, 150–152, 153t, 154
- Self-assembled InGaAs quantum dots, 183–231
- capacitance spectroscopy, 211–213, 213
 - crystal growth, 187–205
 - electrical characterization of quantum dots, 211–214
 - growth on non-(100) substrates, 202–204, 203
 - heterostructures, electrical, optical characteristics of, 185
 - high-speed quantum dot lasers, 220–221
 - longitudinal optical phonon energy, 209
 - molecular beam epitaxy, 183
 - optical characterization of quantum dots, 205–211
 - organometallic vapor-phase epitaxy, 184
 - overgrowth of self-assembled quantum dots, 197–200, 200
 - photoluminescence, 205–207, 206
 - polarization-dependent photoluminescence, 205–207, 206
 - quantum disks, 203
 - quantum dot edge-emitting lasers, 214–219, 216–219
 - quantum dot lasers, 214–222
 - quantum dot superlattices, 200–202, 201
 - quantum dot VCSELs, 219–220
 - reflectance anisotropy spectroscopy, 192
 - reflection high-energy electron diffraction, 184
 - scanning electron microscopy, 192
 - single layers, quantum dots on GaAs, 190–197, 192–199
 - single quantum dot photoluminescence, 207–209, 208, 209

- Stranski-Krastanow transition, 183
- time-resolved photoluminescence, 209–211, 211
- Self-assembled quantum dots
 - hole-burning, 79–89, 80–82, 84–86, 88
 - in SK mode, 35–59
- Self-formed quantum dot structures, 19–95
 - atomic force microscopy, 37
 - band filling, 24
 - closely stacked InAs quantum dots, 45–55, 46–55
 - columnar InAs quantum dots, 55–59, 57–61
 - communication wavelengths, lasing at, 74
 - distributed feedback lasers, 77
 - fabrication technologies, key issues, 73
 - growth control of tetrahedral-shaped recess quantum dots, 32–35, 33–36
 - growth process of stacked islands, 50
 - higher-performance lasers, 73–76, 75
 - hole-burning, 79–89, 80–82, 84–86, 88
 - InAs island growth, 37–39, 38, 40, 41
 - incomposition, surface morphology, among InGaAs films, 31
 - inhomogeneous broadening, smaller, high dot density, 73
 - lasers, 68–73, 69–72
 - magneto-photoluminescence, 51
 - metal-oxide semiconductors, 60
 - molecular beam epitaxy, 30, 37
 - multiple-layer growth, perpendicular alignment of islands, 39–45, 42–44
 - negative differential resistance region, 61
 - nonradiative recombination, reduction of, 74
 - optical gain, 67–68
 - optical intensity modulator, 77–79, 79
 - phonon bottleneck, 62, 74
 - potential device applications, of quantum dots, 59–89
 - quantum dot formation, optical properties, 21–28, 22–26
 - quantum dot growth mechanism, 28, 28–32, 30, 31t
 - quantum dot lasers, 67–77
 - reflection high-energy electron diffraction, 37
 - self-organized quantum dots lasers, 68–73, 69–72
 - in SK mode, 35–59
 - Stranski-Krastanov mode growth, 36
 - tetrahedral-shaped recess formation, 20–21, 21
 - tetrahedral-shaped recess quantum dots, 20–35, 61–67, 63, 64, 65, 66
 - transmission electron microscopy, 37
 - trimethylgallium, 22
 - trimethylindium, 22
 - Self-organized Ge islands
 - atomic-force, scanning tunneling microscopy, 239–240
 - grazing incidence, 248–251, 250
 - infrared detectors, 251–252, 253
 - light emission, enhanced, from Si-based heterostructures, 252–255, 254–255
 - low-pressure chemical vapor deposition, 234
 - molecular beam epitaxy, 234
 - photoluminescence, 242, 243, 244
 - Raman scattering, 242
 - self-organized Ge dots, growth of, 234
 - small-angle x-ray scattering, 251
 - Stranski-Krastanow growth, 233, 234
 - structural characterization, 233–258
 - transmission electron microscopy, 240–241, 241
 - x-ray diffraction, 246–248, 247, 249
 - reflectivity, 242–251
 - x-ray reflectivity, 245, 245–246
 - Self-organized quantum dot lasers, 68–73, 69–72
 - SEM. *See* Scanning electron microscopy
 - Semi-conductor Industry Association Roadman*, 103
 - Sensors, metrology, quantum nanocircuits, 15–16
 - Silylation, polymer, 115
 - Single-atom manipulation, lithography, 129
 - Single-electron
 - quantum nanocircuits, 15
 - transistors, 8–12, 9, 11
 - Small-angle x-ray scattering, grazing-incidence, self-organized Ge islands, 251
 - Soft lithography, 118–123
 - conformal phase-shift lithography, 121
 - imprint lithography, 123, 124–125
 - microcontact printing, 120, 120–121
 - photopolymerization, 121–123, 122–123
 - Stacked islands, growth process of, self-formed quantum dot structures, 50
 - STM. *See* Scanning tunneling microscopy

- Stranski-Krastanow growth mode, 36, 145, 233, 234
 self-assembling in, 144
- Surface morphology, incomposition, among InGaAs films, 31
- TBP. *See* Tertiary-butyl-phosphine
- TEGa. *See* Triethylgallium
- TEM. *See* Transmission electron microscopy
- Temperature, low, conductances, quantum nanocircuits, 9
- Tennant's law, lithography, 100
- Tertiary-butyl-phosphine, MOVPE, 143
- Tetrahedral-shaped recess quantum dots, 20–35, 61–67, 63, 64, 65, 66
 growth control of tetrahedral-shaped recess quantum dots, 32–35, 33, 34, 35, 36
 quantum dot formation, optical properties, 21–28, 22, 23, 24, 25, 26
 quantum dot growth mechanism, 28, 28–32, 30, 31t
 tetrahedral-shaped recess formation, 20–21, 21
- Three-dimensional islands
 MOVPE, densities, sizes of, manipulations of, 169–173, 170, 171, 172, 173, 174
 self-assembled, MOVPE, sizes, shapes, 149–154, 150, 151, 152, 153t, 154
- Time-resolved photoluminescence, self-assembled InGaAs quantum dots, 209–211, 211
- TMAI. *See* Trimethylaluminum
- TMGa. *See* Trimethylgallium
- TMIn. *See* Trimethylindium
- Top surface imaging, 115, 116
- Transistors, decreasing dimensions, 4
- Transmission electron microscopy, 37
 self-organized Ge islands, 240–241, 241
- Triethylgallium, MOVPE, 143
- Trimethylaluminum, MOVPE, 143
- Trimethylaminealane, MOVPE, 143
- Trimethylgallium, 22
- Trimethylindium, 22
- TSI, 114–118, 115, 116
- Tunnel barriers, quantum nanocircuits, 3–4
- Two-dimensional-3D transition, MOVPE, phenomenology of, 162t, 162–168, 163, 165, 167, 168
- Uncertainty principle, Heisenberg's, 6
- Vapor-phase epitaxy, MOVPE, 139
- Volmer-Weber growth mode, MOVPE, 145
- VPE. *See* Vapor-phase epitaxy
- X-ray diffraction, self-organized Ge islands, 242–251, 246–248, 247, 249
- X-ray lithography, 106–107
- X-ray reflectivity, self-organized Ge islands, 245, 245–246