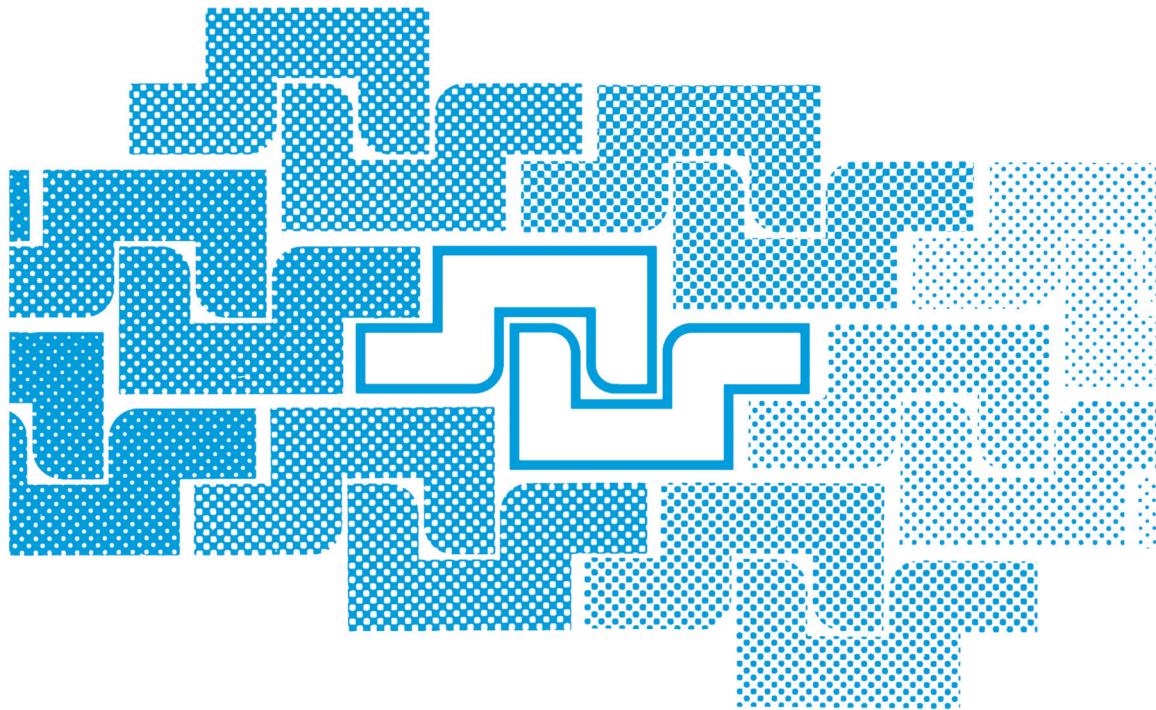




Quantum Dots: Fundamentals, Applications, and Frontiers

Edited by

Bruce A. Joyce, Pantelis C. Kelires,
Anton G. Naumovets and Dimitri D. Vvedensky

NATO Science Series

Quantum Dots: Fundamentals, Applications, and Frontiers

NATO Science Series

A Series presenting the results of scientific meetings supported under the NATO Science Programme.

The Series is published by IOS Press, Amsterdam, and Springer (formerly Kluwer Academic Publishers) in conjunction with the NATO Public Diplomacy Division.

Sub-Series

I. Life and Behavioural Sciences	IOS Press
II. Mathematics, Physics and Chemistry	Springer (formerly Kluwer Academic Publishers)
III. Computer and Systems Science	IOS Press
IV. Earth and Environmental Sciences	Springer (formerly Kluwer Academic Publishers)

The NATO Science Series continues the series of books published formerly as the NATO ASI Series.

The NATO Science Programme offers support for collaboration in civil science between scientists of countries of the Euro-Atlantic Partnership Council. The types of scientific meeting generally supported are “Advanced Study Institutes” and “Advanced Research Workshops”, and the NATO Science Series collects together the results of these meetings. The meetings are co-organized by scientists from NATO countries and scientists from NATO’s Partner countries — countries of the CIS and Central and Eastern Europe.

Advanced Study Institutes are high-level tutorial courses offering in-depth study of latest advances in a field.

Advanced Research Workshops are expert meetings aimed at critical assessment of a field, and identification of directions for future action.

As a consequence of the restructuring of the NATO Science Programme in 1999, the NATO Science Series was re-organized to the four sub-series noted above. Please consult the following web sites for information on previous volumes published in the Series.

<http://www.nato.int/science>

<http://www.springeronline.com>

<http://www.iospress.nl>



Quantum Dots: Fundamentals, Applications, and Frontiers

edited by

Bruce A. Joyce

The Blackett Laboratory,
Imperial College London, U.K.

Pantelis C. Kelires

Physics Department, University of Crete, Heraclion, Crete, Greece
and Foundation for Research and Technology-Hellas (FORTH),
Herac lion, Crete, Greece

Anton G. Naumovets

National Academy of Sciences of Ukraine,
Institute of Physics, Kiev, Ukraine

and

Dimitri D. Vvedensky

The Blackett Laboratory,
Imperial College London, U.K.



“British Association of Crystal Growth”



Springer

Published in cooperation with NATO Public Diplomacy Division

Proceedings of the NATO Advanced Research Workshop on
Quantum Dots: Fundamentals, Applications and Frontiers
Crete, Greece
20–24 July 2003

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 10 1-4020-3314-1 (PB) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN 13 978-1-4020-3314-8 (PB) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN 10 1-4020-3313-3 (HB) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN 10 1-4020-3315-X (e-book) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN 13 978-1-4020-3314-1 (HB) Springer Dordrecht, Berlin, Heidelberg, New York
ISBN 13 978-1-4020-3314-5 (e-book) Springer Dordrecht, Berlin, Heidelberg, New York

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Printed on acid-free paper

All Rights Reserved
© 2005 Springer

No part of this work may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, microfilming, recording or otherwise, without written permission from the Publisher, with the exception of any material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work.

Printed in the Netherlands.

CONTENTS

Preface	ix
List of Contributors	xi

Atomistic Processes during Quantum Dot Formation

Quantum Dots in the InAs/GaAs System: An Overview of their Formation	1
B. A. Joyce and D. D. Vvedensky	
First-Principles Study of InAs/GaAs(001) Heteroepitaxy	27
E. Penev and P. Kratzer	
Formation of Two-Dimensional Si/Ge Nanostructures Observed by STM	43
B. Voigtländer	
Diffusion, Nucleation and Growth on Metal Surfaces	55
O. Biham, I. Furman, H. Mehl and J. F. Wendelken	

The Stranski–Krastanov Transition

The Mechanism of the Stranski–Krastanov Transition	71
A. G. Cullis, D. J. Norris, T. Walther, M. A. Migliorato, and M. Hopkinson	
Off-lattice KMC Simulations of Stranski-Krastanov-Like Growth	89
M. Biehl and F. Much	
Temperature Regimes of Strain-Induced InAs Quantum Dot Formation	103
C. Heyn and A. Bolz	
Kinetic Modelling of Strained Films: Effects of Wetting and Facetting	121
D. Kandel and H. R. Eisenberg	
Ge/Si Nanostructures with Quantum Dots grown by Ion-Beam-Assisted Heteroepitaxy	135
A. V. Dvurechenskii, J. V. Smagina, V. A. Armbrister, V. A. Zinovyev, P. L. Novikov, S. A. Teys, and R. Groetzschel	

Self-Assembly of Quantum Dot Arrays

Lateral Organization of Quantum Dots on a Patterned Substrate C. Priester	145
Some Thermodynamic Aspects of Self-Assembly of Quantum Dot Arrays J. E. Prieto and I. Markov	157
The Search for Materials with Self-Assembling Properties: The Case of Si-based Nanostructures I. Goldfarb	173

Structure and Composition of Quantum Dots

X-Ray Scattering Methods for the Study of Epitaxial Self-Assembled Quantum Dots J. Stangl, T. Schüllli, A. Hesse, G. Bauer, and V. Holý	183
Carbon-Induced Ge Dots on Si(100): Interplay of Strain and Chemical Effects G. Hadjisavvas, Ph. Sonnet, and P. C. Kelires	209
Growth Information Carried by Reflection High-Energy Electron Diffraction Á. Nemcsics	221

Electrons and Holes in Quantum Dots

Efficient Calculation of Electron States in Self-Assembled Quantum Dots: Application to Auger Relaxation D. Chaney, M. Roy, and P. A. Maksym	239
Quantum Dot Molecules and Chains W. Jaskólski, M. Zieliński, A. Stróżecka, G. W. Bryant, and J. Aizpurua	257
Collective Properties of Electrons and Holes in Coupled Quantum Dots G. Goldoni, F. Troiani, M. Rontani, D. Bellucci, E. Molinari, and U. Hohenester	269

Phase Transitions in Wigner Molecules	285
J. Adamowski, B. Szafran, and S. Bednarek	

Fast Control of Quantum States in Quantum Dots: Limits due to Decoherence	301
L. Jacak, P. Machnikowski, and J. Krasnyj	

Optical Properties of Quantum Dots

Real Space <i>Ab Initio</i> Calculations of Excitation Energies in Small Silicon Quantum Dots	317
A. D. Zdetsis, C. S. Garoufalidis, and S. Grimme	

GeSi/Si(001) Structures with Self-Assembled Islands: Growth and Optical Properties	333
N. V. Vostokov, Yu. N. Drozdov, D. N. Lobanov, A. V. Novikov, M. V. Shaleev, A. N. Yablonskii, Z. F. Krasilnik, A. N. Ankudinov, M. S. Dunaevskii, A. N. Titkov, P. Lytvyn, V. U. Yukhymchuk, and M. Ya. Valakh	

Quantum Dots in High Electric Fields: Field and Photofield Emission from Ge Nanoclusters on Si(100)	353
A. A. Dadykin, A. G. Naumovets, Yu. N. Kozyrev, M. Yu. Rubezhanska, and Yu. M. Litvin	

Optical Emission Behavior of Si Quantum Dots	369
X. Zianni and A. G. Nassiopoulou	

Strain-Driven Phenomena upon Overgrowth of Quantum Dots: Activated Spinodal Decomposition and Defect Reduction	377
M. V. Maximov and N. N. Ledentsov	

Preface

The morphology that results during the growth of a material on the substrate of a different material is central to the fabrication of all quantum heterostructures. This morphology is determined by several factors, including the manner in which strain is accommodated if the materials have different lattice constants. One of the most topical manifestations of lattice misfit is the formation of coherent three-dimensional (3D) islands during the Stranski–Krastanov growth of a highly-strained system. The prototypical cases are InAs on GaAs(001) and Ge on Si(001), though other materials combinations also exhibit this phenomenon. When the 3D islands are embedded within epitaxial layers of a material that has a wider band gap, the carriers within the islands are confined by the potential barriers that surround each island, forming an array of quantum dots (QDs).

Such structures have been produced for both basic physics studies and device fabrication, including QD lasers and light-emitting diodes (LEDs) operating at the commercially important wavelengths of 1.3 μm and 1.55 μm . On a more speculative level, QD ensembles have been suggested as a possible pathway for the solid-state implementation of a quantum computer. Although some of the principles of quantum computing have been verified by other means, the practical utilization of this new computing paradigm may warrant some sort of solid state architecture. QDs are seen as possible components of such a computer, as evidenced by a number of papers appearing in the literature proposing QD-based architectures and workshops that are being organized to explore these possibilities.

In view of many developments on several fronts, a workshop to bring together experimentalists and theorists from various disciplines to stimulate interdisciplinary cross-fertilization, to review recent progress and remaining challenges, and to discuss common problems and solutions, was viewed as being particularly timely. The Scientific Affairs Division of NATO, under its Advanced Research Workshop program, and the British Association of Crystal Growth, kindly agreed to support such a meeting. Entitled “Quantum Dots: Fundamentals, Applications, and Frontiers”, this ARW was held at the Santa Marina Beach Hotel in Ammoudara, Crete between 20–24 July, 2003. The themes were the fundamentals of the growth, structure, and properties of quantum dots, their applications in devices, and their possible future uses. The main emphasis was the identification of the factors that limit the utilization of quantum dots in applications and possible strategies by which these can be overcome. This ARW highlighted the following key issues:

- The mechanism of quantum dot formation during Stranski–Krastanov growth is a complex process, with many key factors that are still not completely understood at the atomistic level, including the formation kinetics, the morphological evolution of dots (“precursors” and “pyramids”), and changes caused by capping.
- The role of atomistic and continuum modelling is becoming increasingly important in addressing these issues, with several speakers describing models that reproduce several important aspects of Stranski–Krastanov growth.
- There are theoretical and computation tools available to determine the electronic properties of quantum dots, but the dependence on shape and composition often inhibits the ability to reach unequivocal conclusions.
- Despite the experimental uncertainties about fundamental steps, the controlled growth of quantum dots is already leading to a number of promising applications, including lasers and single-photon detectors.
- Quantum computing based on a quantum dot architecture is a possibility, even if confined to low temperatures, with the scalability and large confinement energies being particularly attractive features. For detrimental effects, such as decoherence due to phonons, plausible solutions have been proposed.

The present volume contains the proceedings of this stimulating meeting. The organizers are deeply thankful to the Scientific Affairs Committee of NATO and the British Association of Crystal Growth for their sponsorship and enthusiastic endorsement of this meeting.

Bruce A. Joyce
Pantelis C. Kelires
Anton G. Naumovets
Dimitri D. Vvedensky

LIST OF CONTRIBUTORS

Janusz Adamowski

Faculty of Physics and Nuclear Techniques
AGH University of Science and Technology
al. Mickiewicza 30
30-059 Kraków Poland
email: adamowski@novell.ftj.agh.edu.pl

Günther Bauer

Institut für Halbleiter-und Festkörperphysik
Universität Linz
A-4040 Linz Austria
email: guenther.bauer@jku.at

Michael Biehl

Institut für Theoretische Physik und Astrophysik
Julius-Maximilians-Universität
Würzburg Germany
email: biehl@physik.uni-wuerzburg.de

Ofer Biham

Racah Institute of Physics
The Hebrew University
Jerusalem 91904 Israel
email: biham@phys.huji.ac.il

Tony Cullis

Department of Electronic and Electrical Engineering
University of Sheffield
Mappin Street
Sheffield S1 3JD United Kingdom
email: a.g.cullis@sheffield.ac.uk

Ilan Goldfarb

Department of Solid Mechanics, Materials and Systems
The Fleischman Faculty of Engineering
Tel Aviv University
Ramat Aviv 69978 Israel
email: ilang@eng.tau.ac.il

Guido Goldoni

INFN National Research Center for nanoStructures
and bioSystems at Surfaces (S3), and Department of Physics
University of Modena and Reggio Emilia
Via Campi 213A
I-41100 Modena Italy
email: goldoni@unimo.it

Christian Heyn

Institut für Angewandte Physik und Mikrostrukturzentrum
Universität Hamburg
Jungiusstr. 11
D-20355 Hamburg Germany
email: cheyn@physnet.uni-hamburg.de

Václav Holý

Institute of Condensed Matter Physics
Masaryk University
Brno Czech Republic
email: holy@physics.muni.cz

Włodzimierz Jaskólski

Instytut Fizyki
Uniwersytet Mikołaja Kopernika
Grudziadzka 5
87-100 Torun Poland
email: wj@nobel.phys.uni.torun.pl

Bruce Joyce (Organizer)

The Blackett Laboratory
Prince Consort Road
Imperial College London
London SW7 2AZ United Kingdom
email: b.joyce@imperial.ac.uk

Daniel Kandel

Department of Physics of Complex Systems
Weizmann Institute of Science
Rehovot 76100 Israel
email: fekandel@wisemail.weizmann.ac.il

Pantelis Kelires (Organizer)

Physics Department

University of Crete

P.O. Box 2208

710 03 Heraclion, Crete Greece

Foundation for Research and Technology-Hellas (FORTH)

P.O. Box 1527

711 10 Heraclion, Crete Greece

email: kelires@physics.uoc.gr

Zakharij Krasilnik

Institute for Physics of Microstructures

Russian Academy of Sciences

GSP-105, Nizhny Novgorod,

603600 Russia

email: zfk@ipm.sci-nnov.ru

Pawel Machnikowski

Institute of Physics

Wrocław University of Technology

Wrocław Poland

email: Pawel.Machnikowski@pwr.wroc.pl

Peter Maksym

Department of Physics and Astronomy

University of Leicester

University Road

Leicester LE1 7RH United Kingdom

email: map@leicester.ac.uk

Ivan Markov

Institute of Physical Chemistry

Bulgarian Academy of Sciences

Sofia 1113 Bulgaria

email: imarkov@ipc.bas.bg

Mikhail Maximov

Ioffe Physical-Technical Institute

Politekhnikeskaya 26

194021 St. Petersburg Russia

email: maximov.beam@mail.ioffe.ru

Anton Naumovets (Organizer)

National Academy of Sciences of Ukraine
Institute of Physics
46 Prospect Nauki
UA-03028 Kiev Ukraine
email: naumov@iop.kiev.ua

Ákos Nemcsics

Hungarian Academy of Sciences
Research Institute for Technical Physics
and Materials Science MTA-MFA
P.O. Box 49
H-1525 Budapest Hungary
email: nemcsics@mfa.kfki.hu

Pavel Novikov

Institute of Semiconductor Physics
Siberian Branch of Russian Academy of Science
Lavrentyev Prospekt 13
Novosibirsk 630090 Russia
email: novikov@isp.nsc.ru

Evgeni Penev

Fritz-Haber-Institut der Max-Planck-Gesellschaft
Faradayweg 4-6
14195 Berlin-Dahlem Germany
email: penev@fhi-berlin.mpg.de

Catherine Priester

IEMN/Dept ISEN
BP 69 59652
Villeneuve d'Ascq Cedex France
email: catherine.priester@isen.fr

Bert Voigtländer

Institut für Schichten und Grenzflächen ISG 3
Forschungszentrum Jülich
52425 Jülich Germany
email: B.Voigtlaender@fz-juelich.de

Dimitri Vvedensky (Organizer)

The Blackett Laboratory
 Prince Consort Road
 Imperial College London
 London SW7 2BZ United Kingdom
 email: d.vvedensky@imperial.ac.uk

Aristides D. Zdetsis

Department of Physics
 University of Patras
 26500 Patras Greece
 email: zdetsis@upatras.gr

Xanthippi Zianni

IMEL
 National Center for Scientific Research “Demokritos”
 153 10 Athens Greece
 email: xzianni@imel.demokritos.gr

QUANTUM DOTS IN THE InAs/GaAs SYSTEM

An Overview of their Formation

Bruce A. Joyce and Dimitri D. Vvedensky

The Blackett Laboratory, Imperial College, London SW7 2BW, United Kingdom

b.joyce@imperial.ac.uk, d.vvedensky@imperial.ac.uk

Abstract We present an overall impression of the present state of knowledge on the formation, via a modified Stranski–Krastanov growth mode, of InAs quantum dots (QDs) on GaAs substrates. We will begin with the substrate orientation and surface reconstruction specificity of QD formation, which demonstrates that QDs are the exception rather than the rule in this system, with the implication that a second process, in addition to strain relaxation, is involved in the driving force. We then discuss the formation of an alloy wetting layer, and although it may not be unique to growth on the GaAs(001) $c(4 \times 4)$ surface, it is very much more marked than on any other. This is an important effect, in that QD formation is effectively limited to the same surface reconstruction. The next stage involves this formation process and we will review the experimental evidence, including dot composition, size (volume) distribution (including scaling behaviour), and two-dimensional to three-dimensional transition effects, with some comments on possible experimental artefacts in this area. We conclude with some comments on QD shape, based mainly on reflection high energy electron diffraction (RHEED) results, but including a comparison with results from transmission electron microscopy (TEM) and scanning tunnelling microscopy (STM).

Keywords: GaAs surfaces, self-organization, quantum dots, heteroepitaxy, reconstructions, Stranski–Krastanov growth

1. Introduction

The InAs–GaAs heteroepitaxial system has a lattice mismatch of $\approx 7\%$ and under very specific conditions the growth mode follows a version of the Stranski–Krastanov (SK) mechanism, which results in the formation of coherent (dislocation-free) three-dimensional (3D) islands, after the growth of between one and two monolayers (MLs) in a two di-

mensional (2D) layer-by-layer pseudomorphic mode. The 3D islands are usually referred to as self-assembled quantum dots, or more simply QDs. The process has been extensively reviewed (see, for example, [1–4]).

During the 2D-to-3D transition the islands rapidly reach a saturation number density of between $\approx 10^9$ and $\approx 10^{11}$ cm^{-2} , depending on growth conditions, with a rather narrow size (volume) distribution, where each island contains between $\approx 10^4$ and $\approx 5 \times 10^4$ atoms, again depending on growth conditions, principally deposition rate and substrate temperature. Despite being the archetypal III-V system, however, and despite a number of suggestions that comprehensive growth models exist, there are several important questions that remain unanswered. Our aim in this paper is to discuss the issues that need to be resolved, some of which are fairly clear-cut, but with others either the experimental evidence derived from different techniques appears to be contradictory, or the theoretical models neglect important results.

The particular issues we will be discussing include:

- Substrate orientation and reconstruction specificity of QD formation
- Significance of alloying in wetting layer (WL) and QD formation
- Mechanism of the 2D-to-3D growth mode transition
- Possible scaling behaviour of 2D and 3D islands
- The shape of QDs as determined by RHEED, TEM and STM

We will confine our attention here to material grown by molecular beam epitaxy (MBE), because almost all of the fundamental results have been obtained using this method, but similar QDs can equally well be produced by metal-organic chemical vapour deposition (MOCVD).

2. Orientation and Reconstruction Specificity in QD Formation

It is generally agreed that the essential driving force for coherent QD formation, after a WL has formed, is strain relaxation, whereby the energy gain from the increase in surface area via dot formation more than compensates the increase in interfacial free energy [5]. In principle, of course, the initial misfit strain is independent of both substrate orientation and surface reconstruction, so there should be no dependence of QD formation on these parameters. It has been known for some time, however, that this is not the case [6] and as an example, Fig. 1 illustrates the relaxation mechanism which obtains in the growth of InAs

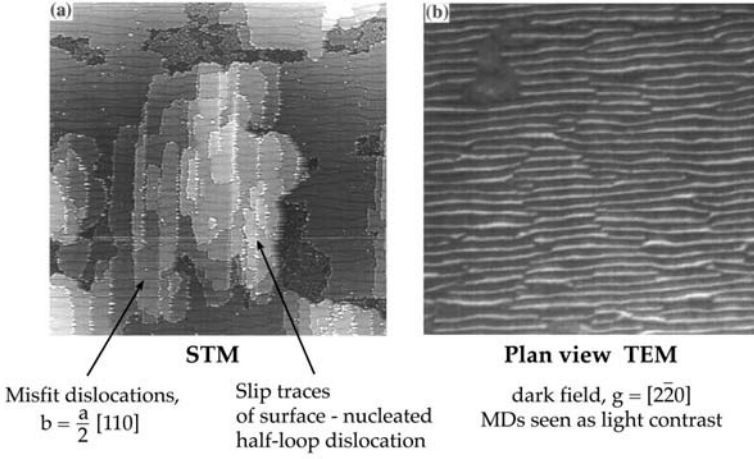


Figure 1. Misfit dislocations imaged by two different microscopies for different areas of the same 5 ML thick InAs film grown on GaAs(110) at 420°C. (a) STM and (b) TEM. Both image dimensions are 200 nm×200 nm. (a) also shows the growth mode to be 2D layer-by-layer.

on GaAs(110). The growth mode is 2D layer-by-layer and the strain is relaxed by the formation of misfit dislocations, whose geometry and surface atomic displacements are fully consistent with the crystallography of the system [7–13]. This is in fact the usual situation for this system and very similar behaviour is observed on (111)A-oriented substrates [12, 14–18], whilst on (111)B substrates growth effectively follows a Volmer-Weber (VW) mode, with the formation of comparatively large discrete islands [19].

Even on (001)-oriented substrates, QDs do not form under all conditions, but only on certain surface reconstructions, notably in the presence of excess (above stoichiometry) arsenic. Because In is significantly more volatile than Ga, growth is restricted to substrate temperatures $\leq 520^\circ\text{C}$, but typically they are in the range 450–500°C. This means that it is difficult to generate Ga-rich substrate surfaces prior to growth, but it is possible to produce cation-stable growth conditions by using a high In:As flux ratio. In this case growth is again 2D and strain is relaxed by misfit dislocations [20]. With regard to more As-rich surfaces, the $\beta 2(2 \times 4)$ is the most stable GaAs(001) surface [21], but again it is extremely difficult to maintain it for very long under typical InAs growth conditions. The sub-ML films which can be grown are, however, essentially 2D and fully strained. At longer deposition times any uncovered substrate surface tends to become $c(4 \times 4)$ reconstructed and it is this most As-rich surface phase that totally favours the formation of QDs

via some version of SK growth. These results are summarized in Table 1, which also contains a column relating to WL formation. It is clear that this occurs most readily on the GaAs(001)-c(4×4) surface and is much more difficult or even non-existent on other orientations. We will discuss this more fully in Sec. 3, but first we need to consider reasons for the specificity of QD formation.

Table 1. Summary of orientation and reconstruction effects

Substrate	Growth mode	Strain relaxatiion mechanism	Wetting layer?	QDs form?
GaAs(110)	Layer-by-layer (2D)	Misfit dislocations	No	No
GaAs(111)A (2×2)	2D	Misfit dislocations	No	No
GaAs(111)B	3D (VW?)	Not known	Not known	Large 3D islands
On GaAs(001)				
cation-stable (i.e. high In:As flux ratio)	2D	Misfit dislocations	Not known	No
(2×4) Difficult to maintain at InAs growth temperature ($< 520^\circ\text{C}$)	2D prior to conversion to c(4×4)-like at ~ 1 ML	Fully strained	Not defined	No
c(4×4)	SK-like, but 3D islands are coherent	Alloy and island formation (prior to coalescence)	Yes, strong alloying effects	Yes

One possible explanation is that the introduction of misfit dislocations is more energetically favourable than QD formation on most surfaces (all except (001)-c(4×4) of the low index orientations, which are the only ones that do not comprise more than one component). In the absence of any quantitative values this is a feasible proposition, at least as far as orientational specificity is concerned, but it is less easy to justify with reconstruction differences on the same orientation. It is also worth noting that on (110) in particular, strain relaxation via misfit dislocations is very anisotropic, being more complete along $[1\bar{1}0]$ than $[001]$ [22, 23]. This anisotropy can be explained by the orientation of $\{111\}$ slip planes with respect to the (110) surface. Only two of the four slip planes, (111) and $(11\bar{1})$, are inclined to (110) and both intersect

along $[1\bar{1}0]$. Lomer dislocations, with Burgers vector $\mathbf{b} = \frac{1}{2}a[1\bar{1}0]$, which are the most efficient in relieving strain, can be formed easily along $[001]$, so that strain relief along $[1\bar{1}0]$ is effectively complete from the early stages of growth. In the orthogonal direction Lomer dislocations cannot form, however, but the presence of two inclined $\{111\}$ planes allows the subsequent formation of 60° -type dislocations with $\mathbf{b} = \frac{1}{2}a\langle 101 \rangle$, which can then slip along these planes to reach the interface. The strain along $[001]$ is then gradually relieved via the component of $\mathbf{b}$ lying along $[001]$. No such complexity of strain relaxation via misfit dislocations occurs on (001) surfaces, but nevertheless the primary relaxation process does not involve dislocation formation.

An alternative reason might be found in the alloy WL that forms on (001) - $c(4 \times 4)$ substrates (see Sec. 3), such that the total strain is mediated over an interface region of significant thickness to provide a strain gradient, rather than across an abrupt boundary. The critical thickness for misfit dislocations is therefore not reached before QDs are nucleated as a form of growth instability. Any confirmation of this hypothesis, however, probably requires a more quantitative understanding of QD nucleation than is presently available (see Sec. 4).

The third possibility is that there is an intrinsic structural reason, within the $c(4 \times 4)$ surface, for QD formation. This is an attractive proposition and would reflect the differences in homoepitaxial nucleation of GaAs between (001) and the other low index orientations [24], but would require the $c(4 \times 4)$ surface to have properties that favour this method of strain relaxation over misfit dislocation formation. The most likely possibility is that it enhances the formation of an alloy WL, which appears to be crucial to QD nucleation. It is known to be a multi-layer structure [25] with a large effective step density which could lower significantly any steric barrier to alloying, but it is not certain that this would be sufficient. Experimentally it is clear that alloying occurs most readily on this surface [26, 27], but we need to consider in more detail whether the mechanism of the 2D-to-3D transition necessarily involves an alloy, as mentioned in the previous paragraph (see Sec. 4).

3. Alloying in the WL and QD Formation

We will concentrate on deposition on (001) - $c(4 \times 4)$ substrates, since to all intents and purposes this is the only low-index surface on which QD formation occurs unequivocally and on which there is also extensive alloying at all temperatures $\geq 350^\circ\text{C}$. We will deal first with the initial 2D WL behaviour before discussing alloying in QDs. The formation of the WL has been followed dynamically *in-situ* by RHEED and

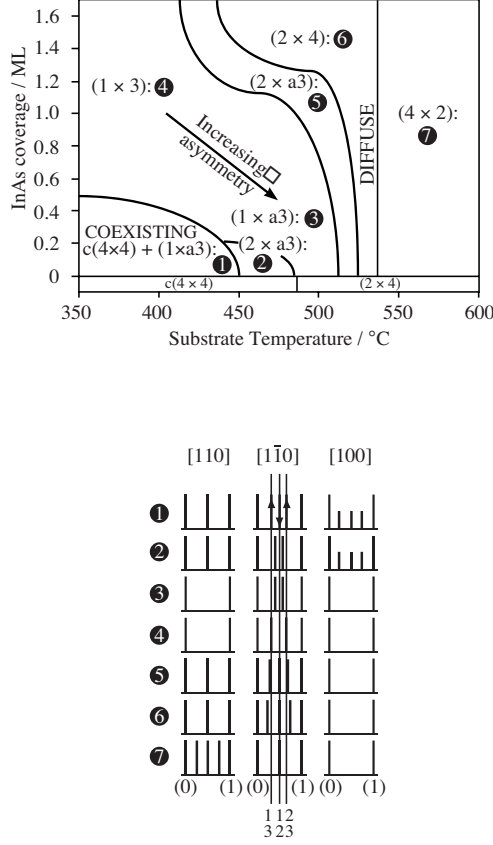


Figure 2. Schematic diagram indicating the RHEED patterns observed during WL formation of InAs on GaAs(001) as a function of nominal InAs coverage and substrate temperature. ‘a’ indicates an asymmetric pattern.

by snapshot STM images of quenched surfaces; there do not appear to be any substantial disparities between results from the two methods. The basic RHEED data are presented in Fig. 2, which shows the surface structures present as a function of substrate temperature and deposit coverage, while in Fig. 3(a,b) STM-based results illustrate how deposition of sub-ML amounts can nevertheless produce complete substrate coverage by a homogeneous layer, which confirms the presence of an (In,Ga)As alloy. The temperature dependence of the extent (rate) of alloy formation shows it to be an activated process, for which the most probable barrier is surface diffusion of In [27]. From Fig. 2 it is apparent that the alloy phase is based on a (1×3) surface unit cell, the same as that found on non-dilute bulk alloys (both components present at $>20\%$ concentration). There are small regions of the associated (2×3)

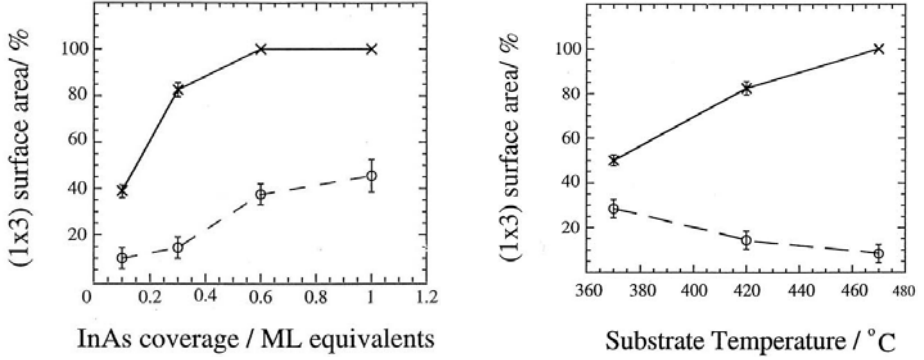


Figure 3. Percentage of surface area covered by the (1×3) reconstruction as a function of the amount of InAs deposited (in ML equivalents) at 420°C (left panel) and the substrate temperature for a nominal 0.3 ML deposition (right panel).

structure, a (2×4) phase found at high coverage and eventually a cation-stable (4×2) structure forms at high temperature, which is associated with In segregation and is probably an alloy very dilute in Ga.

Structural models of the (1×3) and (2×3) surfaces, based on an alloy composition of $\text{In}_{2/3}\text{Ga}_{1/3}\text{As}$ are characterized by continuous top-layer rows of As dimers along $[\bar{1}10]$, with two cation positions in the third layer occupied by In and one by Ga. The doubled periodicity along $[110]$ for the (2×3) reconstruction arises from an As dimer back-bonded to four third layer In atoms. These structures are illustrated in Fig. 4(a,b), but it must be emphasized that they only apply to this fixed composition alloy layer. Thermodynamic and first principles calculations show that for this composition the (2×3) structure is the most stable, but deviations from it result in incommensurate phases [28, 29, 30].

We are then left with two questions: (i) why is the $c(4 \times 4)$ structure the most susceptible of all GaAs surfaces to alloy formation, and (ii) what is the nature of the so-called incommensurate phases? With regard to the ease of alloy formation, it has been shown using a combination of RHEED and STM that InAs incorporates directly into the $c(4 \times 4)$ surface and alloyed domains form in the same layer as the remaining uncovered regions (at low coverages) [25, 31], whereas on the (2×4) , InAs islands only nucleate on top of the existing reconstruction and a uniform WL is not reached until almost a complete ML has been deposited [31]. For thicknesses ≥ 1 ML, the WL morphology is apparently independent of the starting reconstruction [31], but as we have already pointed out, it is very difficult to maintain this reconstruction for any substantial length of time under InAs deposition conditions, so it is perhaps not surprising

that differences do not persist beyond 1 ML. We can therefore look to the basic $c(4 \times 4)$ structure to account for the alloying behaviour, and two features seem to be important. One is the multi-layer structure, which means that in any defined area there is a very high step density and the other is the double layer of As in the surface, so any excess cations can be accommodated by replacing second layer As atoms with them.

We turn now to the so-called incommensurate phases, which produce asymmetric diffraction patterns, and which are clearly present under almost all conditions of substrate temperature and coverage, compared with the very limited regions where the fixed composition (2×3) phase exists. The asymmetry is made manifest by an unequal spacing between the $\frac{1}{3}$ -order diffraction features in the $[\bar{1}10]$ azimuth, even at complete ML coverage, when no substrate surface is exposed. It is a well known effect and has been reported for several semiconductor systems, including Si/GaAs, InAs, InSb and GaSb [32, 33]. It was first reported by Germer *et al.* [34] for the O_2/Ni (110) system and the analysis was extended by Houston and Park [35]. The origin is scattering from uniform sequences of different sized surface units formed from the same basic building block, which in the present case is 2 units and 3 units. The proportion of each defines the splitting of the $\frac{1}{2}$ -order rods towards the $\frac{1}{3}$ and $\frac{2}{3}$ positions, but in addition the domain sizes need to be small because of the limited coherence width of the incident electron beam. Here the domains were ≈ 45 Å and they are shown in the STM image of Fig. 5.

This model has been substantiated by generating an array of 2 and 3 block units with appropriate domain sizes which was then Fourier transformed, and led to a well defined asymmetric (1×3) “diffraction pat-

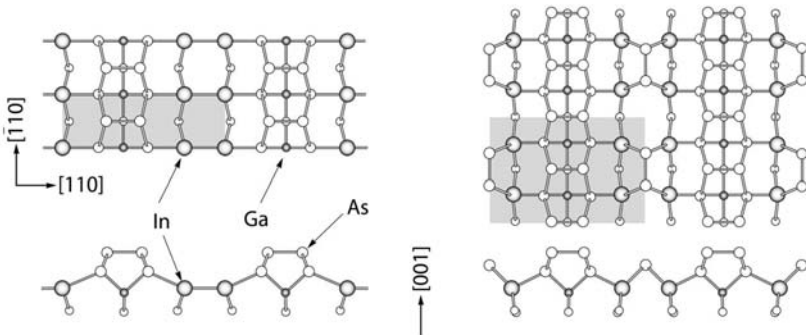


Figure 4. Structural models for the (1×3) (left panel) and (2×3) (right panel) reconstructed WL surfaces.

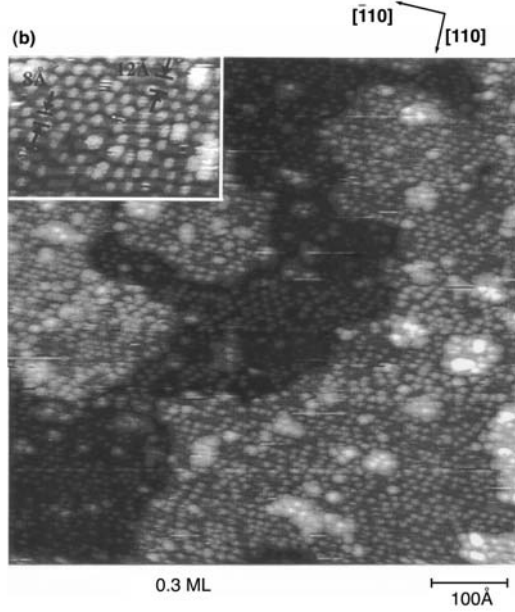


Figure 5. STM image of “domains” in the $a(1 \times 3)$ reconstructed WL surface. 0.3ML InAs deposited on a $c(4 \times 4)$ reconstructed surface at 420°C . The inset shows 2-fold and 3-fold periodicities that result in an asymmetric (1×3) RHEED pattern.

tern”. Ideally confirmation could be obtained by Fourier transforming the actual STM image, but in practice it was too noisy for the available computing power [41]. Overall, the evidence is for a reasonably well-ordered surface structure of the WL, but it is not some fixed entity with respect to either detailed structure or composition. An important point is that the RHEED data are dynamic and represent the state of the surface prior to QD formation; growth was continuous, there was no growth interruption or quenching.

In the light of the available measurements it seems reasonable to claim that the surface structure of the substrate dominates WL alloying behaviour, to the extent that it is effectively only the $\text{GaAs}(001)\text{-}c(4 \times 4)$ surface which has the appropriate form to allow this to happen. The surface structure and morphology of the alloys formed can also be determined and the conclusion here is that a range of structures is formed, and while there are certain similarities, QD formation is not restricted to any one of them, but can occur from all. An alloy WL is a precursor to QD formation, but its surface structure is not a critical factor. Irrespective of its exact form, however, it seems to be a necessary condition and in

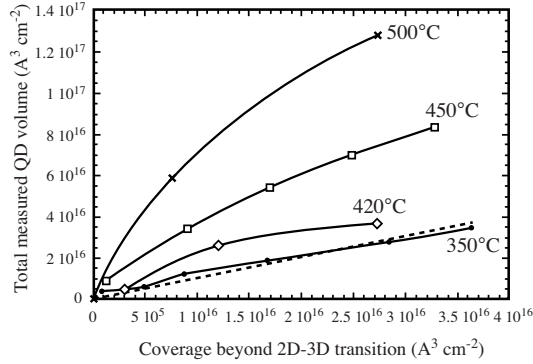


Figure 6. Total measured QD volume as a function of the amount (effective coverage) of InAs deposited beyond the 2D–3D transition at different substrate temperatures.

the next section we discuss the nature of the 2D to 3D transition and the possible role the WL plays in this.

Alloying in QD formation is in principle a much more straightforward process. There is ample evidence [37, 38, 39] to show that the amount of material in the dots at all substrate temperatures $\geq 400^\circ\text{C}$ is greater than the amount of InAs deposited beyond the 2D-to-3D transition point, and by $\approx 500^\circ\text{C}$ it is greater than the total amount of InAs deposited, as illustrated in Fig. 6. The somewhat surprising exception to this effect occurs at very low deposition rates ($\leq 0.02 \text{ ML s}^{-1}$), where the dot volume is effectively equivalent to the amount of InAs deposited once the transition has taken place [39].

Other than at these very low deposition rates, the QDs must therefore be an (In,Ga)As alloy, and furthermore, it has also been shown [40–43] in general the alloy composition is not homogeneous, but there is an In composition gradient from the apex of the dot to its base. This non-uniformity occurs mainly as a result of In surface segregation, which is dealt with more fully in the paper in these proceedings by Cullis [44], and to a lesser extent by solid state diffusion of the two cation species. Although there is no direct proof, it appears certain that at the transition point material is transferred from the alloy WL (rather than directly from the GaAs substrate) into the dot via an activated process, since the amount transferred increases with increasing temperature. Qualitatively, the process is independent of the detailed surface structure of the WL, and so far no quantitative effects have been reported. We will discuss possible mechanisms of dot formation in Sec. 4. The present anomaly is why QD alloying is not observed at very low deposition rates.

It clearly reflects a strong kinetic influence on the transition, since were the process to be close to equilibrium, alloy formation would be favoured at low rates. One possible pointer is that the number density of dots is significantly lower (by between one and two decades) and they are correspondingly larger than dots grown at more conventional rates ($\approx 0.01 \text{ ML s}^{-1}$ cf. $\approx 0.1 \text{ ML s}^{-1}$), although the transition occurs at the same effective coverage. One possibility suggested [39] relates to the magnitude of the strain field as a function of dot volume, as described by Koduvely and Zangwill [69]. This is larger for larger dots and as a consequence the detachment barrier for cations from these larger dots is reduced, although why this should have the effect of virtually eliminating the Ga component, whether it occurs in the QD or the WL, was not made clear by the authors of reference [39]. The only certainty is the extent of kinetic control involved, which must influence our thinking on the 2D-to-3D transition process, as will be evident in Sec. 4.

4. Mechanism of the 2D-to-3D Transition

The precise mechanism of the growth mode transition from the 2D WL to 3D island (QD) nucleation is probably the most controversial issue involved in the evolution of QD growth in this material system. The major problems do not arise with the initial and final states; it is quite clear that a 2D alloy WL is first formed, having variable composition, composition gradient and structure, from which comparatively large (between $\approx 10^4$ and $\approx 5 \times 10^4$ atoms) coherent (dislocation free) QDs form, in general as (In,Ga)As alloys and at number densities between $\approx 10^9$ and $\approx 10^{11} \text{ cm}^{-2}$. It is the intermediate or precursor state to fully formed dots where there is controversy. Even without theoretical considerations, the experimental results themselves are not straightforward to interpret. The difficulty arises because it is not possible to make a direct comparison between RHEED and STM results.

RHEED is the only dynamic *in-situ* technique available to observe the 2D-to-3D transition directly, but is limited in its sensitivity. There is general acceptance that the RHEED pattern changes from streaked rods produced by a 2D surface to a spot pattern formed by transmission diffraction through 3D asperities with the addition of an incremental $\approx 0.1 \text{ ML}$ ($\approx 5 \times 10^{13} \text{ atoms cm}^{-2}$), which is about the detection limit of the technique. The number of atoms in the final state of the QDs is between $\approx 5 \times 10^{13}$ and $\approx 10^{15} \text{ cm}^{-2}$, depending on size and number density. In other words, at the low end of the range, RHEED would not be able to detect the early stages of any transition, and in any case would probably not detect the formation of 2D (or shallow 3D)

precursors, which might only occupy a very small fraction (< 0.1 ML) of the surface.

Conversely, STM images have little or no sensitivity problem and are easily able to resolve the evolution of 2D precursors, even if they are present at a low number density, but unfortunately the images are not dynamic. They are snapshots taken after quenching the sample from the growth temperature and recording the image at room temperature. This could clearly lead to the condensation of any mobile surface species, so 2D or very shallow 3D precursors observed in STM images may not actually be present during growth. It is difficult to see how this can be resolved until an STM is available which is capable of producing real-time images at the growth temperature, and while this is basically possible for elemental semiconductors (see the article by Voigtländer in these proceedings), it is some way off for III-V materials.

Despite these reservations and in addition to the small incremental deposition known to be required to produce QDs, there are several other undisputed facts regarding dot formation which are summarized below:

- The QD number density decreases with increasing temperature and with decreasing In flux.
- There is a weak temperature dependence of the critical deposit thickness (amount) for QD formation, with a gradual increase for increasing temperature.
- The transition thickness is effectively independent of the In flux.
- In general, QDs are alloyed, as described in the preceding section (Sec. 3).

Whilst these observations are important and any comprehensive theory must be able to accommodate them, the most crucial factor remains the identification of the precise nature of the 2D-to-3D transition. There are two possible routes: either nucleation, or surface roughening followed by island growth and coarsening. For this system only the former has been considered, although there may be evidence that the latter should not be dismissed at this stage. At least it would be consistent with the narrow size distributions reported for QDs, since in general nucleation is a random process, but a more uniform array might be expected from co-operative formation via a nucleationless effect.

Reported results have in essence all been obtained from STM observations, since as we have indicated RHEED is not very sensitive to the presence of 2D or flat 3D precursors, despite being able to identify the rapid transition to genuine 3D entities. Heitz *et al.* [5] and Ramachandran *et al.* [47] described a so-called re-entrant 2D-to-3D morphology

change in which quasi-3D clusters first appear, then disappear, only to reappear with increasing deposition prior to the formation of 3D islands, for which the clusters act as precursors. Although they used a different terminology, Pinnington *et al.* [48] identified somewhat similar behaviour using UV light scattering, but of course they were not able to produce images, so no definitive comparison can be made. In a recent paper, Patella *et al.* [49] claimed on volumetric grounds that QDs could not be derived from quasi-3D precursors. Rather, they attributed the process to the initial formation of an intermixed $\text{In}_x\text{Ga}_{1-x}\text{As}$ ($x \approx 0.82$) layer after deposition of 1 ML, followed by a “floating” In surface population as deposition continued, from which In surface mass transport could lead to the rapid volume increase required to form QDs. AFM images of InAs deposits at ~ 0.15 ML below the critical coverage for QD formation revealed a roughened surface due to mounds of lateral dimensions $\approx 1.2 \times 0.3 \mu\text{m}$, elongated along $[1\bar{1}0]$, and about 3 nm high, possibly indicative of a growth instability, especially as such morphological details as step bunching and step-edge meandering were apparent.

Although all of these results were discussed in terms of nucleation behaviour, they could perhaps be considered as possible examples of a roughening process as the precursor to islands. In this context, it is also worth pointing out that immediately prior to the conversion of the RHEED pattern from streaks to spots it virtually disappears for a brief period, which may be indicative of surface roughening. Equally, of course, it could be the result of a mobile adatom population.

On the other hand, Krzyzewski *et al.* [50, 51] come down very heavily in favour of a process in which large 2D islands do not transform into 3D islands, but co-exist with small, irregular 3D entities, 2–4 ML high, each containing about 150 atoms. These then develop rapidly with an incremental 0.05 ML into “mature” QDs incorporating $> 1 \times 10^4$ atoms, with a number density $\geq 10^{10} \text{ cm}^{-2}$. No evidence is presented as to how this process occurs, but the values would suggest that it may be necessary for material from the WL to be included, although at the very low deposition rates used (0.017 ML s^{-1}), the same group reported that the QDs were essentially 100% InAs. Nor is it certain that the very small 3D entities can be associated unequivocally with processes occurring at the growth temperature. The possibility that they form during the quenching and cooling stages used to obtain the STM images cannot be ruled out. In favour of a nucleation based transition, however, is that the QD size distribution close to the critical coverage is large. Scaling analysis of the QD size distribution from this work implied that strain is important in the first stages of QD formation, but is not involved in their subsequent development. This latter conclusion is in agreement

with earlier work by Ebiko *et al.* [52, 53] and it is possible that elastic interactions in a dense array of islands could contribute significantly to their growth kinetics, as suggested originally by Krishnamurthy *et al.* [54]. We discuss scaling in more detail in the following section.

The role of In segregation to the surface of the WL in the transition process also needs to be considered and it has been proposed by Walther *et al.* [55] and Cullis *et al.* [56] that it controls the critical point of the 2D-to-3D transition. The argument is strongly based on nucleation concepts in that when the In adatom population becomes large enough it allows critical nuclei to form, which then grow to produce the 3D islands. They are stabilized by stress relief via the expansion of laterally unconstrained lattice planes and become effective sinks for the remaining adatoms, so the In population decreases. This is also consistent with RHEED results, since the momentary disappearance of the diffraction pattern immediately before 3D islands form could be the result of a large mobile adatom population. Segregation effects are discussed in much more detail in the article by Cullis in these proceedings.

We can only conclude that our understanding of the 2D-to-3D transition process is still in a rudimentary state, largely because we do not yet have reliable experimental techniques to evaluate surface processes on this scale in real time. Until a truly dynamic method is available we are unlikely to reach a definitive answer and at this stage we need to keep all of our mechanistic options open.

5. Scaling of QD Arrays

Measurements [52, 53, 57, 58, 59] of the sizes (volumes) of QDs have shown that in the temperature range 490–540°C the size and the inter-dot separation distributions obey the scaling laws derived for *homoepitaxial growth* [60, 61, 62], where strain is not explicitly considered. The size distribution is

$$N_s = \frac{\theta}{s_{\text{av}}^2} f(s/s_{\text{av}}), \quad (1)$$

where N_s is the number density of QDs containing s atoms, θ is the fractional surface coverage, s_{av} is the average size of a QD and f is the distribution function. As shown in Fig. 7 for *dislocated* InN islands grown on GaN(0001) [59], the data closely follow the scaling function for homoepitaxial irreversible aggregation. Similar agreement was obtained for the dot separation distribution (Fig. 8), expressed as the probability of finding an island whose center is a distance r away from the center of another island,

$$N(r) = Ng(r/r_{\text{av}}), \quad (2)$$

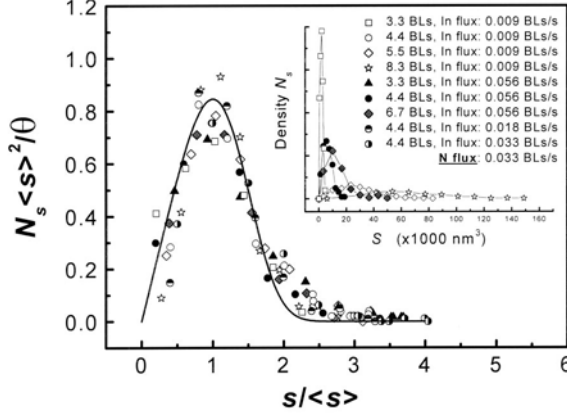


Figure 7. Scaled density of 3D InN islands on GaN(0001). The solid curve is the analytic expression of Ref. [64] with critical island size $i=1$. The inset shows the unscaled densities for selected data sets representing different coverages and deposition conditions. The lines are to guide the eye.

where N is the total number density of islands, $r_{av} \sim N^{-1/2}$ is the average inter-island separation, and g is a scaling function with the properties that $g \rightarrow 0$ as $r \rightarrow r_0$ and $g \rightarrow 1$ as $r \rightarrow \infty$, where r_0 is the average island size. This behaviour appears to indicate that the nucleation and growth of QDs is determined only by the irreversible aggregation of migrating In or Ga adatoms and not to any factors attributable to strain.

The main processes that determine the form of the scaling function are nucleation, attachment, and detachment. For irreversible homoge-

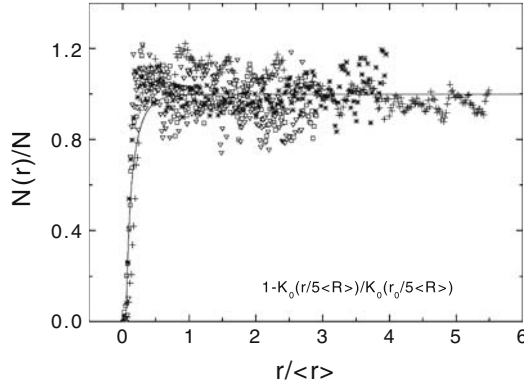


Figure 8. Pair distribution of 3D InN islands on GaN(0001). The solid curve is the indicated analytic expression from Ref. [61] with a cutoff radius $r / \langle r_0 \rangle = 0.06$.

neous nucleation, two adatoms form a critical nucleus that grows by the capture of migrating adatoms; there is no adatom detachment from islands. Any changes to this basic scenario produce discernible differences in the scaling function. In particular, strain can affect growth kinetics in several ways, e.g. elasticity calculations show that 2D islands above a certain critical size will seek to reconfigure themselves into 3D islands that relieve strain more efficiently [5], the existence of a strain- and size-dependent energy barrier to the incorporation of adatoms into a 3D island that is presumed to arise from the strain field in the neighborhood of a coherent island [66], and interactions due to substrate distortion that lead to a repulsive interaction between islands [4]. Indeed, size-dependent attachment and detachment rates have been shown [67] to produce systematic deviations from the scaling behavior noted above.

To rationalize the scaling of QDs, we note first the island-size distribution is determined by island growth rates [63, 65] which, in turn, are determined by the sizes of capture zones, defined as the region of the substrate that is closer to a given island than to any other. Thus, adatoms deposited into a capture zone are more likely to diffuse to that island than to any other [68]. On a flat substrate with no islands initially, the deposition flux creates an adatom population that, at sufficiently high densities, initiates island nucleation. Eventually a steady state is reached where no further nucleation occurs and the adatom and island densities are constant. In this regime, the capture zones remain unchanged, so island growth rates are constant. This is the scaling regime of island sizes. The apparent absence of any effect of strain is consistent with the inter-dot separations (Fig. 8), which show no evidence of the exclusion zone one would expect from the short-range repulsion between islands.

The absence of adatom detachment can be explained, perhaps, by parametrizing the island size distribution by the ratio λ of the net detachment rate from an island to the net attachment rate to an island, rather than by a nominal critical island size:

$$\lambda = \frac{R_d}{(Fk + D\sigma n)}, \quad (3)$$

where F is the flux, D the adatom diffusion constant, k the direct capture number from the incident flux, σ the diffusive capture number with an adatom density of n and R_d is the escape rate of an adatom from a 3D island. Kinetic Monte Carlo simulations of 2D homoepitaxy show that λ parametrizes a continuous family of scaling functions, but for $\lambda \approx 1$, the island size distribution fits Eq. (1), even when significant detachment occurs [69]. It must be remembered, however, that the measured islands represent a final state and it is entirely feasible that λ varies as the

islands develop. This would rationalize the fits obtained by Ebiko *et al.* [52, 53], but any effect of strain is still only implicit, in the sense that it could influence R_d , but not that it is a dominant driving force for 3D nucleation.

6. The Shape of QDs

The precise shape of QDs has important ramifications for their electronic structure and properties, so it is somewhat surprising that questions remain over such an apparently simple aspect. As with the 2D-to-3D transition, at least part of the problem derives from the interpretation of the different techniques used to measure it, but this is compounded by the fact that significant shape changes occur during encapsulation of the QDs by a higher band-gap material (usually GaAs) to provide electronic confinement. Here we will concentrate on the first issue, prior to encapsulation, since we are basically concerned with QD formation.

Results using RHEED are frequently quoted [70, 71, 72] based on the observation of chevron features extending from diffraction spots towards (occasionally away from) the shadow edge, which are attributed to faceting of the QDs. It is usually assumed that the chevrons correspond to reciprocal lattice rods normal to the facet plane, so the half-angle between the surface normal and the chevron direction determines the angle between the (001) surface plane and the facet, whose orientation can then readily be determined. On this basis the facets have been identified mainly as $\{n11\}$, where $n=3, 4$ or 5 [71, 72], although $\{136\}$ and $\{137\}$ facets have also been reported [70, 73]. This is the case for patterns obtained in the $[\bar{1}10]$ azimuth, but in the orthogonal $[110]$ azimuth, only diffraction spots are observed, apparently formed by transmission diffraction through asperities, so there can be no question of the QDs being pyramidal, a statement which has frequently been made, especially in the context of electronic structure calculations [74, 75, 76]. Patterns in both $\langle 110 \rangle$ azimuths are illustrated in Fig. 9.

There is a second serious problem with the interpretation based on a faceted shape: it is simple to show from reciprocal lattice considerations that diffraction from facets should produce not chevrons but symmetrical crosses centred on the diffraction spots. This is obviously contrary to results obtained for the InAs–GaAs system (and for several others), so we need to find an alternative explanation. The problem has been tackled by two groups using somewhat different approaches [77, 40], but each came to the same conclusion. Pashley *et al.* [77] showed by accurate calculation of the refractive deviation of an incident beam for very small grazing angles that pure refraction effects could explain not only the

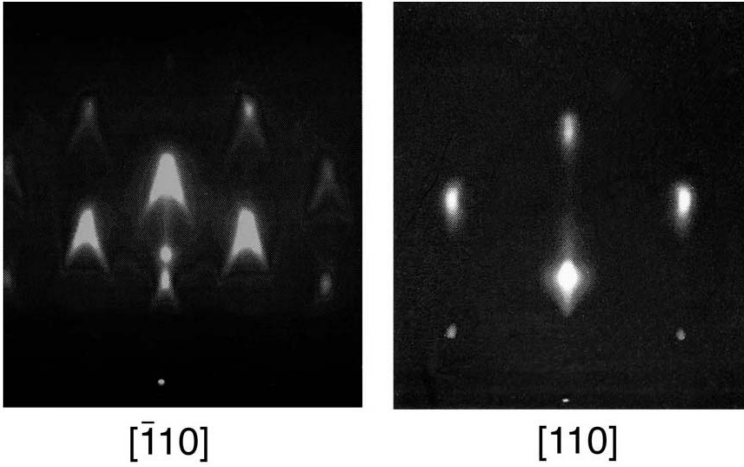


Figure 9. RHEED patterns from ≈ 2.0 ML of InAs on GaAs(001) $c(4 \times 4)$: (left) $[\bar{1}10]$ azimuth, (right) $[110]$ azimuth.

absence of a symmetrical cross, but could also account for the length and intensity distribution of the arms of the chevron (Fig. 9), i.e. the intensity falls off rapidly with distance from the centre of the diffraction spot and there is in-filling of intensity close to the peak. The QD shape required to produce these effects is lenticular (Fig. 10); the bounding faces of the QD must be curved, rather than well-developed facets, which also explains the transmission spots formed in the $[110]$ azimuth, since electrons entering and leaving zones A and C in Fig. 10 will not be refracted significantly because the grazing angles of incidence involved are too large.

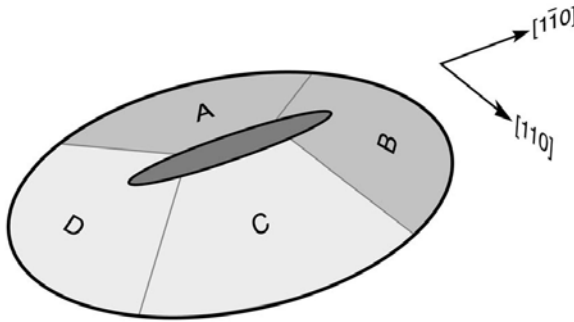


Figure 10. Division of a lenticular-shaped QD into zones for the purpose of considering the different possibilities for refractive deviations.

Hanada *et al.* [40] similarly studied refraction effects, but used kinematic theory to calculate first the points of incidence and exit for each scattering atom and then the wave vectors inside and the scattering angle at the atom for given incident and out-going wave vectors in vacuum. The results of the calculations for various possible QD shapes were compared with experimental results and they concluded that the most probable shape was asymmetric dome-shaped and the key requirement to produce a chevron was the curved surface of the dot facing the $[1\bar{1}0]$ and $[\bar{1}10]$ azimuths.

In this application of RHEED there is no question of possible insensitivity, as we are dealing with the final state array of QDs, not their nucleation, so it is only a question of interpretation. Since two different methods of calculating refraction effects give essentially identical answers, we can be reasonably certain that the QD shape shown in Fig. 10 is that present during growth, especially as virtually all published RHEED data show the same features. Finally, we note that if a genuine faceted shape does form, or is created by annealing, the RHEED pattern does indeed show clearly defined symmetric crosses. This is illustrated in Fig. 11 for a sequence of growth followed by annealing in the InP(001)–InAs system. The initial pattern is indicative of the same refraction effects as discussed above, but with annealing the pattern predicted for diffraction from a regularly-faceted structure develops. In other words, if facets are present, they are detected by RHEED.

Alternative means of shape analysis, all effectively *ex situ*, include TEM, STM and high resolution x-ray diffractometry (HRXRD). When due attention is paid to the correct interpretation of image contrast, TEM results in general agree with the conclusions reached by RHEED [79, 80]. There are several cases where regular faceted structures have

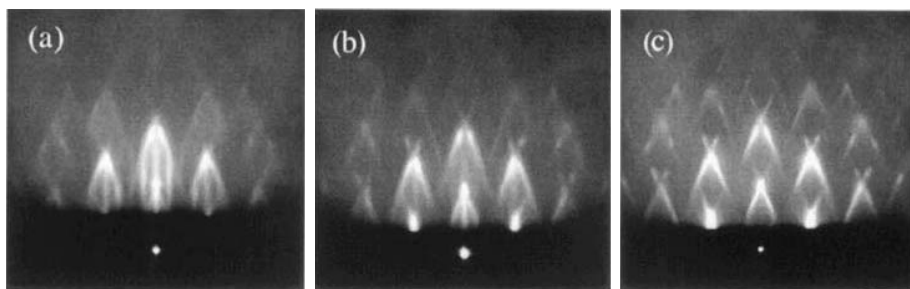


Figure 11. RHEED patterns in the $[110]$ azimuth from the growth of 2.5 ML InAs on InP(001), followed by annealing, showing the final development of symmetrical crosses from the initial chevrons: (a) 3 minute anneal at 450°C , (b) 3 minute anneal at 460°C , (c) 3 minute anneal at 480°C and cooldown at 430°C .

been reported, but in these there is often confusion between strain and atomic number (Z -dependent) contrast in the images. STM is probably not an ideal technique for precise shape analysis of features ≈ 50 Å high and there is a tendency to enhance electronically the images to show idealized regular facets. When height profiles are shown, the presence of regular facets appears much more dubious [81, 82]. HRXRD results are less definitive, but the analysis is not simple [43] (see also the article by Bauer in these proceedings). Strain and composition profiles tend to support lenticular shapes, rather than regular crystallographic facets.

Finally, it is interesting to compare experimentally determined shapes with those calculated theoretically. Pehlke *et al.* [83] used *ab initio* computations to deduce the equilibrium shape of the 3D islands and found them to be bounded by $\{110\}$ and $\{111\}$ facets and terminated by an (001) top. This is clearly not what is measured experimentally by any reliable analysis, so either the calculation has certain shortcomings, or the growth process is well removed from equilibrium. The latter is most probably true for reasons we have already discussed, so the calculation may simply be inapplicable to the experimental situation. Spencer and Tersoff [84] studied the equilibrium of a coherent strained layer which wets the substrate (SK mode) and found that discrete islands formed with a zero contact angle to the film wetting the substrate; small islands had a minimum width and hence an arbitrarily small aspect ratio, while very large islands had a shape which approached that of a sphere on a WL, but there was no faceting. Again, however, only equilibrium growth could be treated, so although the results were closer to observation, it would be unwise to conclude that accurate predictions of the real situation can be made.

References

- [1] J. M. Gérard, O. Cabrol, J. Y. Marzin, N. Lebouché, and J. M. Moison. Optical investigation of some statistic and kinetic aspects of the nucleation and growth of InAs islands on GaAs. *Mat. Sci. Eng. B* **37**: 8–16, 1996.
- [2] D. Bimberg, M. Grundmann and N. N. Ledentsov. *Quantum Dot Heterostructures*. John Wiley, 1999.
- [3] B. A. Joyce and D. D. Vvedensky. Growth modes and morphologies on GaAs surfaces, in *Thin Films: Heteroepitaxial Systems*, W.K. Lin and M.B. Santos, eds., World Scientific, Singapore, 1999, pp. 368-400.
- [4] V. A. Shchukin and D. Bimberg. Spontaneous ordering of nanostructures on crystal surfaces. *Rev. Mod. Phys.* **71**: 1125–1171, 1999.
- [5] J. Tersoff and F. K. LeGoues. Competing relaxation mechanisms in strained layers. *Phys. Rev. Lett.* **72**: 3570–3573, 1994.
- [6] B. A. Joyce, T. S. Jones and J. G. Belk. Reflection high-energy electron diffraction scanning tunneling microscopy study of InAs growth on the three low

- index orientations of GaAs: Two-dimensional *versus* three-dimensional growth and strain relaxation. *J. Vac. Sci. Technol. B* **16**: 2373–2380, 1998.
- [7] J. G. Belk, J. L. Sudijono, X. M. Zhang, J. H. Neave, T. S. Jones and B. A. Joyce. Surface contrast in two dimensionally nucleated misfit dislocations in InAs/GaAs(110) heteroepitaxy. *Phys. Rev. Lett.* **78**: 475–478, 1997.
 - [8] J. G. Belk, D. W. Pashley, C. F. McConville, J. L. Sudijono, B. A. Joyce and T. S. Jones. Surface atomic configurations due to dislocation activity in InAs/GaAs(110) heteroepitaxy. *Phys. Rev. B* **56**: 10289–10296, 1997.
 - [9] J. G. Belk, D. W. Pashley, C. F. McConville, B. A. Joyce and T. S. Jones. Surface morphology during strain relaxation in the growth of InAs on GaAs(110). *Surf. Sci.* **410**: 82–98, 1998.
 - [10] J. G. Belk, D. W. Pashley, B. A. Joyce and T. S. Jones. Dislocation displacement field at the surface of InAs thin films grown on GaAs(110). *Phys. Rev. B* **58**: 16194–16201, 1998.
 - [11] D. Maroudas, L. A. Zepeda–Ruiz and W. H. Weinberg. Kinetics of strain relaxation through misfit dislocation formation in the growth of epitaxial films on compliant substrates. *Appl. Phys. Lett.* **73**: 753–755, 1998.
 - [12] L. A. Zepeda–Ruiz, D. Maroudas and W. H. Weinberg. Theoretical study of the energetics, strain fields, and semicoherent interface structures in layer-by-layer semiconductor heteroepitaxy. *J. Appl. Phys.* **85**: 3677–3695, 1999.
 - [13] N. Oyama, E. Ohta, K. Takeda, K. Shiraishi and H. Yamaguchi. First-principles calculations on atomic and electronic structures of misfit dislocations in InAs/GaAs(110) and GaAs/InAs(110) heteroepitaxies. *J. Crystal Growth* **201**: 256–259, 1999.
 - [14] H. Yamaguchi, M. R. Fahy and B. A. Joyce. Inhibitions of three dimensional island formation in InAs films grown on GaAs(111)A surface by molecular beam epitaxy. *Appl. Phys. Lett.* **69** 776–778, 1996.
 - [15] H. Yamaguchi, J. G. Belk, X. M. Zhang, J. L. Sudijono, M. R. Fahy, T. S. Jones, D. W. Pashley and B. A. Joyce. Atomic-scale imaging of strain relaxation via misfit dislocations in highly mismatched semiconductor heteroepitaxy: InAs/GaAs(111)A. *Phys. Rev. B* **55**: 1337–1340, 1997.
 - [16] L. A. Zepeda–Ruiz, R. I. Pelzel, B. Z. Nosh, W. H. Weinberg and D. Maroudas. Deformation behavior of coherently strained InAs/GaAs(111) A heteroepitaxial systems: Theoretical calculations and experimental measurements. *J. Appl. Phys.* **90**: 2689–2698, 2001.
 - [17] B. Z. Nosh, L. A. Zepeda–Ruiz, R. I. Pelzel, W. H. Weinberg and D. Maroudas. Surface morphology in InAs/GaAs(111)A heteroepitaxy: Experimental measurements and computer simulations. *Appl. Phys. Lett.* **75**: 829–831, 1999.
 - [18] A. Ohtake, M. Ozeki and J. Nakamura. Strain relaxation in InAs/GaAs(111)A heteroepitaxy. *Phys. Rev. Lett.* **84**: 4665–4668, 2000.
 - [19] S. B. Schujman, S. R. Soss, K. Stokes and L. J. Schowalter. Self-assembled InAs islands on GaAs($\bar{1}\bar{1}\bar{1}$) substrates. *Surf. Sci.* **385**: L965–L970, 1997.
 - [20] Q. K. Xue, Y. Hasegawa, T. Ogino, H. Kiyama and T. Sakurai. In-rich 4×2 reconstruction in novel planar growth of InAs on GaAs(001). *J. Vac. Sci. Technol. B* **15**: 1270–1273, 1997.

- [21] V. P. LaBella, H. Yang, D. W. Bullock, P. M. Thibado, P. Kratzer and M. Scheffler. Atomic structure of the GaAs(001)-(2 × 4) surface resolved using scanning tunneling microscopy and first-principles theory. *Phys. Rev. Lett.* **83**: 2989–2992, 1999, and references therein.
- [22] X. Zhang, D. W. Pashley, L. Hart, J. H. Neave, P. N. Fawcett, and B. A. Joyce. the morphology and asymmetric strain relief behavior of InAs films on GaAs(110) grown by molecular-beam epitaxy. *J. Crystal Growth* **131**: 300–308, 1993.
- [23] X. Zhang, D. W. Pashley, J. H. Neave, P. N. Fawcett, J. Zhang and B. A. Joyce. Reflection high-energy electron-diffraction transmission electron-microscopy observattion of growth of InAs on GaAs(110) by molecular-beam epitaxy. *J. Crystal Growth* **132**: 331–334, 1993.
- [24] A. R. Avery, H. T. Dobbs, D. M. Holmes, B. A. Joyce and D. D. Vvedensky. Nucleation and growth of islands on GaAs surfaces. *Phys. Rev. Lett.* **79**: 3938–3941, 1997.
- [25] G. R. Bell, J. G. Belk, C. F. McConville and T. S. Jones. Species intermixing and phase transitions on the reconstructed (001) surfaces of GaAs and InAs. *Phys. Rev. B* **59**: 2947–2955, 1999.
- [26] J. G. Belk, J. L. Sudijono, D. M. Holmes, C. F. McConville, T. S. Jones and B. A. Joyce. Spatial distribution of In during the initial stages of growth of InAs on GaAs(001)-c(4 × 4). *Surf. Sci.* **365**: 735–742, 1996.
- [27] J. G. Belk, C. F. McConville, J. L. Sudijono, T. S. Jones and B. A. Joyce. Surface alloying at InAs-GaAs interfaces grown on (001) surfaces by molecular beam epitaxy. *Surf. Sci.* **387**: 213–226, 1997.
- [28] L. G. Wang, P. Kratzer, N. Moll and M. Scheffler. Size, shape, and stability of InAs quantum dots on the GaAs(001) substrate. *Phy. Rev. B* **62**: 1897–1904, 2000.
- [29] L. G. Wang, P. Kratzer and M. Scheffler. Energetics of InAs thin films and islands on the GaAs(001) substrate. *Jpn. J. Appl. Phys.* **39**(7B): 4298–4301, 2000.
- [30] M. Sauvage-Simkin, Y. Garreau, R. Pinchaux, M. B. Véron, J. P. Landesman and J. Nagle. Commensurate and incommensurate phases at reconstructed (In,Ga)As(001) surfaces – X-ray-diffraction edidence for a composition lock-in. *Phys. Rev. Lett.* **75**: 3485–3488, 1995.
- [31] T. J. Krzyzewski, P. B. Joyce, G. R. Bell and T. S. Jones. Wetting layer evolution in InAs/GaAs(001) heteroepitaxy: effects of surface reconstruction and strain. *Surf. Sci.* **517**: 8–16, 2002.
- [32] A. G. deOliveira, S. D. Parker, R. Droopad and B. A. Joyce. A generalized model for the reconstruction of (001)) surfaes of III-V compound semiconductors based on a RHEED study of InSb(001). *Surf. Sci.* **227**: 150–156, 1990.
- [33] M. R. Fahy, M. J. Ashwin, J. J. Harris, R. C. Newman and B. A. Joyce. Reflec-tion high-energy electron-diffraction study of the GaAs:Sii:GaAs system. *Appl. Phys. Lett.* **61**: 1805–1807, 1992.
- [34] L. H. Germer, J. W. May and R. J. Szostak. Thermally ordered oxygen on a nickel surface. *Surf. Sci.* **7**: 430–447, 1967.
- [35] J. E. Houston and R. L. Park. Low-energy electron-diffraction from imperfect structures. *Surf. Sci.* **21**: 209–223, 1970.

- [36] J. G. Belk. *Surface Aspects of Strain Relaxation During InAs/GaAs Heteroepitaxy*. Ph.D. Thesis, University of London, unpublished, pp. 66 *et seq.*, 1997.
- [37] B. A. Joyce, J. L. Sudijono, J. G. Belk, H. Yamaguchi, X. M. Zhang, H. T. Dobbs, A. Zangwill, D. D. Vvedensky and T. S. Jones. A scanning tunneling microscopy reflection high energy electron diffraction-rate equation study of the molecular beam epitaxial growth of InAs on GaAs(001), (110) and (111)A – Quantum dots and two-dimensional modes. *Jpn. J. Appl. Phys.* **36**(6B): 4111–4117, 1997.
- [38] P. B. Joyce, T. J. Krzyzewski, G. R. Bell, B. A. Joyce and T. S. Jones. Composition of InAs quantum dots on GaAs(001): Direct evidence for (In,Ga)As alloying. *Phys. Rev. B* **58**: R15981–R15984, 1998.
- [39] P. B. Joyce, T. J. Krzyzewski, G. R. Bell, T. S. Jones, S. Malik, D. Childs and R. Murray. Effect of growth rate on the size, composition, and optical properties of InAs/GaAs quantum dots grown by molecular-beam. *Phys. Rev. B* **62**: 10891–10895, 2000.
- [40] A. Rosenauer, U. Fischer, D. Gerthsen and A. Förster. Composition evaluation of $\text{In}_x\text{Ga}_{1-x}\text{As}$ Stranski–Krastanow-island structures by strain state analysis. *Appl. Phys. Lett.* **71**: 3868–3870, 1997.
- [41] A. Rosenauer, D. Gerthsen, D. Van Dyck, M. Arzburger, G. Böhm and G. Abstreiter. Quantification of segregation and mass transport in $\text{In}_x\text{Ga}_{1-x}\text{As}$ /GaAs Stranski–Krastanow layers. *Phys. Rev. B* **64**: art. no. 245334, 2001.
- [42] M. A. Migliorato, A. G. Cullis, M. Fearn and J. H. Jefferson. Atomistic simulation of strain relaxation in $\text{In}_x\text{Ga}_{1-x}\text{As}$ /GaAs quantum dots with nonuniform composition. *Phys. Rev. B* **65**: art. no. 115316, 2002.
- [43] I. Kegel, T. H. Metzger, A. Lorke, J. Peisl, J. Stangl, G. Bauer, N. Nordlund, W. V. Schoenfeld and P. M. Petroff. Determination of strain fields and composition of self-organized quantum dots using x-ray diffraction. *Phys. Rev. B* **63**: art.no. 035318, 2001.
- [44] A. G. Cullis. The mechanism of the Stranski–Krastanow transition, these proceedings.
- [45] H. M. Koduvely and A. Zangwill. Epitaxial growth kinetics with interacting coherent islands. *Phys. Rev. B* **60**: R2204–R2207, 1999.
- [46] R. Heitz, T. R. Ramachandran, A. Kalburge, Q. Xie, I. Mukhametzhanov, P. Chen and A. Madhukar. Observation of reentrant 2D to 3D morphology transition in highly strained epitaxy: InAs on GaAs. *Phys. Rev. Lett.* **78**: 4071–4074, 1997.
- [47] T. R. Ramachandran, A. Madhukar, I. Mukhametzhanov, R. Heitz, A. Kalburge, Q. Xie and P. Chen. Nature of Stranski–Krastanow growth of InAs on GaAs(001). *J. Vac. Sci. Technol. B* **16**: 1330–1333, 1998.
- [48] T. Pinnington, Y. Levy, J. A. Mackenzie and J. Tiedje. Real-time monitoring of InAs/GaAs quantum dot growth using ultraviolet light scattering. *Phys. Rev. B* **60**: 15901–15909, 1999.
- [49] F. Patella, S. Nufri, F. Arciprete, M. Fanfoni, E. Placidi, A. Sgarlata and A. Balzarotti. Tracing the two- to three-dimensional transition in the InAs/GaAs(001) heteroepitaxial growth. *Phys. Rev. B* **67**: art. no. 205308, 2003.

- [50] T. J. Krzyzewski, P. B. Joyce, G. R. Bell and T. S. Jones. Role of two- and three-dimensional surface structures in InAs-GaAs(001) quantum dot nucleation. *Phys. Rev. B* **66**: art. no. 121307(R), 2002.
- [51] T. J. Krzyzewski, P. B. Joyce, G. R. Bell and T. S. Jones. Understanding the growth mode transition in InAs/GaAs(001) quantum dot formation. *Surf. Sci.* **532-535**: 822–827, 2003.
- [52] Y. Ebiko, S. Muto, D. Suzuki, S. Itoh, K. Shiramine, T. Haga, Y. Nakata and N. Yokoyama. Island size scaling in InAs/GaAs self-assembled quantum dots. *Phys. Rev. Lett.* **80**: 2650–2653, 1999.
- [53] Y. Ebiko, S. Muto, D. Suzuki, S. Itoh, H. Yamakoshi, K. Shiramine, T. Haga, K. Unno and M. Ikeda. Scaling properties of InAs/GaAs self-assembled quantum dots. *Phys. Rev. B* **60**: 8234–8237, 1999.
- [54] M. Krishnamurthy, J. S. Drucker and J. A. Venables. Microstructural evolution during the heteroepitaxy of Ge on vicinal Si(001). *J. Appl. Phys.* **69**: 6461–6467, 1991.
- [55] T. Walther, A. G. Cullis, D. J. Norris and M. Hopkinson. Nature of the Stranski–Krastanow transition during epitaxy of InGaAs on GaAs. *Phys. Rev. Lett.* **86**: 2381–2384, 2001.
- [56] A. G. Cullis, D. J. Norris, T. Walther, M. A. Migliorato and M. Hopkinson. Stranski–Krastanow transition and epitaxial island growth. *Phys. Rev. B* **66**: art. no. 081305(R), 2002.
- [57] V. Bressler-Hill, S. Varma, A. Lorke, B. Z. Noshho, P. M. Petroff and W. H. Weinberg. Island scaling in strained heteroepitaxy: InAs/GaAs(001). *Phys. Rev. Lett.* **74**: 3209–3212, 1995.
- [58] T. J. Krzyzewski, P. B. Joyce, G. R. Bell and T. S. Jones. Scaling behavior in InAs/GaAs(001) quantum-dot formation. *Phys. Rev. B* **66**: art. no. 201302, 2002.
- [59] Y. G. Cao, M. H. Xie, Y. Lu, S. H. Xu, Y. F. Ng, H. S. Wu, and S. Y. Tong. Scaling of three-dimensional InN islands grown on GaN(0001) by molecular-beam epitaxy, *Phys. Rev. B* **68** art. no. 161304(R), 2003.
- [60] T. Viscek and F. Family. Dynamic scaling for aggregation of clusters. *Phys. Rev. Lett.* **52**: 1669–1672, 1984.
- [61] M. C. Bartelt and J. W. Evans. Scaling analysis of diffusion-mediated island growth in surface adsorption processes. *Phys. Rev. B* **46**: 12675–12687, 1992.
- [62] J. W. Evans and M. C. Bartelt. Nucleation and growth in metal-on-metal homoepitaxy: Rate-equations, simulations and experiments. *J. Vac. Sci. Technol. A* **12**: 1800–1808, 1994.
- [63] M. C. Bartelt and J. W. Evans. Exact island-size distributions for submonolayer deposition: Influence of correlations between island size and separation. *Phys. Rev. B* **54**: R17359–R17362, 1996.
- [64] J. G. Amar and F. Family. Critical clustersSize: Island morphology and size distribution in submonolayer epitaxial growth. *Phys. Rev. Lett.* **74**: 2066–2069, 1995.
- [65] D. D. Vvedensky. Scaling functions for island-size distributions. *Phys. Rev. B* **62**: 15435–15438, 2000.

- [66] 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*, K. Eberl, P.M. Petroff and P. Demeester, eds., Kluwer, Dordrecht (1995).
- [67] C. Ratsch, A. Zangwill, and P. Šmilauer. Scaling of heteroepitaxial densities. *Surf. Sci.* **314**: L937–L942, 1994.
- [68] P. A. Mulheran and J. A. Blackman. Capture zones and scaling in homogeneous thin-film growth. *Phys. Rev. B* **53**: 10261–10267, 1996.
- [69] H. M. Koduvely and A. Zangwill. Epitaxial growth kinetics with interacting coherent islands. *Phys. Rev. B* **60**: R2204–R2207, 1999.
- [70] H. Lee, R. Lowe-Webb, W. Yang and P. C. Sercel. Determination of the shape of self-organized InAs/GaAs quantum dots by reflection high energy electron diffraction. *Appl. Phys. Lett.* **72**: 812–814, 1998.
- [71] Y. Nabetani, T. Ishikawa, S. Noda and A. Sasaki. Initial growth stage and optical properties of a three-dimensional InAs structure on GaAs. *Appl. Phys. Lett.* **66**: 347–351, 1994.
- [72] R. P. Mirin, J. P. Ibbetson, K. Nishi, A. C. Gossard and J. E. Bowers. 1.3 μm photoluminescence from InGaAs quantum dots on GaAs. *Appl. Phys. Lett.* **67**: 3795–3797, 1995.
- [73] W. D. Yang, H. Lee, T. J. Johnson, P. C. Sercel and A. G. Norman. Electronic structure of self-organized InAs/GaAs quantum dots bounded by $\{136\}$ facets. *Phys. Rev. B* **61**: 2784–2793, 2000.
- [74] C. Pryor. Geometry and material parameter dependence of InAs/GaAs quantum dot electronic structure. *Phys. Rev. B* **60**: 2869–2874, 1999.
- [75] L.-W. Wang, J. N. Kim, and A. Zunger. Electronic structures of $[110]$ -faceted self-assembled pyramidal InAs/GaAs quantum dots. *Phys. Rev. B* **59**: 5678–5687, 1999.
- [76] T. Saito, J. N. Schulman and Y. Arakawa. Strain-energy distribution and electronic structure of InAs pyramidal quantum dots with uncovered surfaces: Tight-binding analysis. *Phys. Rev. B* **57**: 13016–13019, 1998.
- [77] D. W. Pashley, J. H. Neave and B. A. Joyce. A model for the appearance of chevrons on RHEED patterns from InAs quantum dots. *Surf. Sci.* **476**: 35–42, 2001.
- [78] T. Hanada, B.-H. Koo, H. Totsuka and T. Yao. Anisotropic shape of self-assembled InAs quantum dots: Refraction effect on spot shape of reflection high-energy electron diffraction. *Phys. Rev. B* **64**: art. no. 165307, 2001.
- [79] J. Zou, X. Z. Liao, D. J. H. Cockayne and R. Leon. Transmission electron microscopy study of $\text{In}_x\text{Ga}_{1-x}\text{As}$ quantum dots on a GaAs(001) substrate. *Phys. Rev. B* **59**: 12279–12282, 1999.
- [80] Y. Androussi, T. Benabbas and A. Lefebvre. Transmission electron microscopy analysis of the shape and size of semiconductor quantum dots. *Phil. Mag. Lett.* **79**: 201–208, 1999.
- [81] Y. Hasegawa, H. Kiyama, Q. K. Xue and T. Sakurai. Atomic structure of faceted planes of three-dimensional InAs islands on GaAs(001) studied by scanning tunneling microscope. *Appl. Phys. Lett.* **72**: 2265–2267, 1998.

- [82] J. Márquez, L. Geelhaar and K. Jacobi. Atomically resolved structure of InAs quantum dots. *Appl. Phys. Lett.* **78**: 2309–2311, 2001.
- [83] E. Pehlke, N. Moll, A. Kley and M. Scheffler. Shape and stability of quantum dots. *Appl. Phys. A* **65**: 525–534, 1997.
- [84] B. J. Spencer and J. Tersoff. Equilibrium shapes and properties of epitaxially strained islands. *Phys. Rev. Lett.* **79**: 4858–4861, 1997.

FIRST-PRINCIPLES STUDY OF InAs/GaAs(001) HETEROEPITAXY

Evgeni Penev and Peter Kratzer

Fritz-Haber-Institut der Max-Planck-Gesellschaft

Faradayweg 4–6, D-14195 Berlin-Dahlem, Germany

Abstract Density-functional theory calculations are employed to obtain important information about the morphology of III-V semiconductor surfaces and kinetics of epitaxial growth. In this way, insight into the microscopic processes governing quantum dot formation in InAs/GaAs(001) heteroepitaxy is gained. First, we investigate theoretically the atomic structure and thermodynamics of the wetting layer formed by InAs deposition on GaAs(001), including the effect of strain in our discussion. Secondly, we present results about In adatom diffusion both on the wetting layer and on the $c(4 \times 4)$ -reconstructed GaAs(001) surface. In the latter case, we demonstrate the importance of mechanical stress for the height of surface diffusion barriers. Implications for the growth of InAs quantum dots on GaAs(001) are discussed.

1. Introduction

Self-organized quantum-dot heterostructures [1, 2], or simply quantum dots (QDs), will soon leave their “teens” behind. The high technological promises they have brought to optoelectronics have triggered an enormous scientific activity. As a representative example one can take the work on the InAs/GaAs lattice-mismatched heteroepitaxial system, Fig. 1. In terms of the observed resultant morphology, the epitaxy of QDs (e.g. molecular beam epitaxy (MBE), or metal-organic chemical vapor deposition (MOCVD)) follows the Stranski-Krastanov growth mode [3]. Though our knowledge on the nature of the Stranski-Krastanov regime has considerably increased, both experimentalists and theorists are still on the way to complete understanding the intricacies of QD growth kinetics. System-specific, *microscopic* information in this respect is a must. Important information comes from direct experimental probes, like *in situ* scanning tunnelling microscopy (STM), reflection high-energy electron diffraction (RHEED), or reflectance-difference spec-

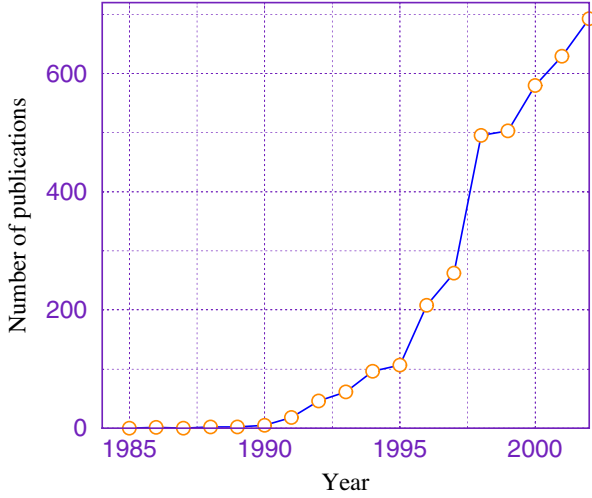


Figure 1. Number of articles published in the period 1985–2002, containing the terms ‘quantum dot(s)’, and ‘InAs’ or ‘GaAs’ in any of the title, abstract or the list of keywords according to Science Citation Index Expanded (isiknowledge.com).

troscopy (RDS). However, if one is aiming at a complete understanding of QD growth, the information accessible from experiments may not be sufficient. This is particularly true if one focuses on the characteristics of species dynamics at the surface, details of the surface atomic structure, thermodynamic quantities, etc.

This contribution addresses InAs heteroepitaxy on GaAs(001)—the III-V semiconductor system of ultimate importance for QD “self-fabrication” [4]. An extensive track record of experimental studies has established a number of anomalies in the QD self-assembly for this system [5]. It was also brought out clearly that any theory has to capture the subtleties of strain-dependent QD growth kinetics [6]. Thus the first fundamental questions to be answered are how strain affects surface morphology, and its impact on surface adatom mobility. The characteristic length and time scales for the latter lie in the microscopic range, 0.1–10 Å and 10^{-15} – 10^{-13} s, where the density-functional theory (DFT) is one of the most commonly used theoretical tools [7, 8]. By considering a few case studies we shall attempt to give an overview of an appropriate theoretical framework, building on DFT, to tackle the above mentioned problems in InAs/GaAs(001) heteroepitaxy. We feel it however compelling to note that its applicability is far more broader and has now covered problems in catalysis, biochemistry, etc. [9].

We start with a brief description of the methodology used to calculate the quantities of interest, e.g. surface energies and adatom diffusion barriers. Then the effect of strain on the equilibrium surface reconstructions of GaAs(001) and InAs(001) is discussed. Analysis of the strain dependence of indium diffusivity is considered as well. Finally, a brief outlook is provided.

2. Theoretical Framework

2.1 First-Principles Surface Thermodynamics

Understanding the properties of the InAs/GaAs(001) heteroepitaxial system naturally implies detailed knowledge of the surface properties of both subsystems. Thus, as a first step, let us consider the *equilibrium* structures of GaAs(001) and InAs(001). The key quantity determining the most stable surface reconstruction under given growth conditions—temperature T and pressure p —is the surface free energy γ , defined as

$$\gamma = \frac{1}{A} \left[G(p, T, \{N_i\}) - \sum_i \mu_i N_i \right]. \quad (1)$$

Here, A is the surface area, μ_i is the chemical potential of the i th component, N_i the corresponding number of particles, and $G(p, T, \{N_i\})$ is the Gibbs free enthalpy. For a GaAs surface Eq. (1) translates into

$$\gamma A = G(p, T, N_{\text{Ga}}, N_{\text{As}}) - \mu_{\text{Ga}} N_{\text{Ga}} - \mu_{\text{As}} N_{\text{As}}. \quad (2)$$

Exploiting the fact that in thermodynamic equilibrium

$$\mu_{\text{Ga}} + \mu_{\text{As}} = \mu_{\text{GaAs}}(p, T), \quad (3)$$

and that in MBE As-rich growth conditions are typical for the case at hand, one can eliminate μ_{Ga} , which leads to

$$\gamma A = G(p, T, N_{\text{Ga}}, N_{\text{As}}) - \mu_{\text{GaAs}}(p, T) N_{\text{Ga}} - \mu_{\text{As}} \Delta N, \quad (4)$$

where $\Delta N = N_{\text{As}} - N_{\text{Ga}}$ is the surface stoichiometry. The chemical potential μ_{As} can take on values in the range

$$\mu_{\text{As}}(\text{bulk})(p, T) - \Delta H_{\text{GaAs}}(p, T) \leq \mu_{\text{As}} \leq \mu_{\text{As}}(\text{bulk})(p, T), \quad (5)$$

ΔH_{GaAs} being the heat of formation of GaAs. It should be noted, however, that one can work out all terms in Eqs. (4) and (5) from a ground-state formalism only for $(p, T) = 0$. Nonetheless, while thermodynamic stability in general requires a discussion of the free energy, it has been

established in the literature [10, 11, 12, 13, 14, 15] that the stability of surface reconstructions is dominated by the (static) energy differences of relaxed structures. For a discussion of surface energies, it is in general adequate to neglect contributions to the free energy coming from the kinetic energy of the atoms, as well as from entropy terms. This is due to the fact that the surface free energy is the difference between two free energies, of the semi-infinite solid plus its surface on the one hand side, and of the reservoirs for the constituent atomic species (gas phase, elemental bulk phases) on the other. By taking the difference between these free energies, the kinetic and entropic terms largely cancel.

With this remark, Eqs. (4) and (5) can be transcribed in a form which provides a direct link to *ab initio* calculations:

$$\gamma A = E^{\text{tot}}(N_{\text{Ga}}, N_{\text{As}}) - N_{\text{Ga}} E_{\text{GaAs}}^{\text{tot}} - \Delta N [\mu_{\text{As}} - \mu_{\text{As(bulk)}}(0, 0)], \quad (6)$$

$$-\Delta H_{\text{GaAs}}(0, 0) \leq \mu_{\text{As}} - \mu_{\text{As(bulk)}}(0, 0) \leq 0. \quad (7)$$

In these expressions, $E^{\text{tot}}(N_{\text{Ga}}, N_{\text{As}})$ is the total energy of the system representing the GaAs surface, $E_{\text{GaAs}}^{\text{tot}}$ the total energy per Ga-As pair in bulk GaAs, $\mu_{\text{As(bulk)}}(0, 0) = E_{\text{As}}^{\text{tot}}$ the total energy per atom in bulk As metal, and

$$\Delta H_{\text{GaAs}}(0, 0) = E_{\text{GaAs}}^{\text{tot}} - E_{\text{Ga}}^{\text{tot}} - E_{\text{As}}^{\text{tot}}. \quad (8)$$

All E^{tot} contributions can now be obtained from DFT calculations. Our setup [17] employs norm-conserving pseudopotentials in conjunction with a plane-wave basis set (with a cutoff energy of 10 Ry) and the local-density approximation to the electronic exchange and correlation energy. Surfaces are represented, within the supercell approximation, by slabs of eight (or seven) atomic layers whose bottom (cation-terminated) surface is passivated [18] by pseudo-hydrogen (H^*) atoms ($Z = 1.25$). Brillouin zone (BZ) integration is carried out using a set of special $\mathbf{k}$ -points equivalent to at least 64 points in the 1×1 surface BZ. The presence of H^* in the actual calculations entails a slight modification of Eq. (6). In order to subtract the contribution of the H^* -passivated side of the slab, a similar calculation is performed for a slab with two H^* -passivated surfaces, whereby we replace $E^{\text{tot}}(N_{\text{Ga}}, N_{\text{As}}) \rightarrow E_{\text{slab}}^{\text{tot}}(N_{\text{Ga}}, N_{\text{As}}, N_{\text{H}^*}) - \frac{1}{2} E_{\text{H}^*-\text{slab}}^{\text{tot}}(N'_{\text{Ga}}, N'_{\text{As}}, N_{\text{H}^*})$. The complete energy “accounting” is exemplified for the $\beta 2(2 \times 4)$ -reconstructed GaAs(001) surface [19] in Fig. 2. Further technical provisions can be found, for examples, in Refs. [10, 20]. This scheme can be readily generalized to the case where surfaces are under strain as well as to account for a possible formation of a ternary surface alloy during the initial stages of InAs deposition on GaAs(001).

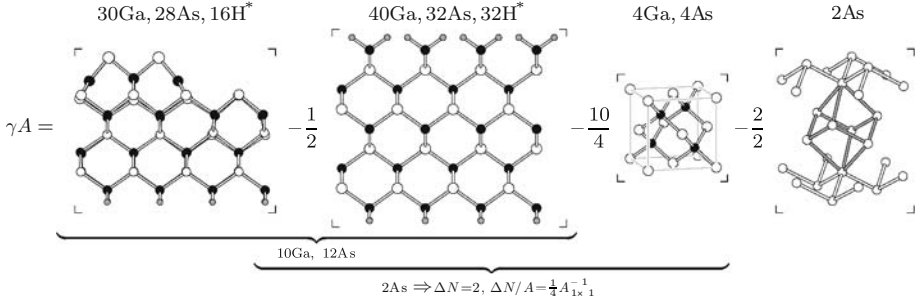


Figure 2. Pictorial representation of the procedure for calculating the surface energy of GaAs(001)- $\beta 2(2 \times 4)$. Number of atoms per species in the supercell used in the actual calculation is given above each system. Total energies of the corresponding systems are symbolically given by small corners. The As reservoir is taken to be rhombohedral bulk As metal—the so-called A7 structure [16].

2.2 Surface Diffusion

Although an equilibrium description may be applicable in some cases of III-V epitaxy [21, 22], often thermodynamic equilibrium is not established during growth. Knowledge of kinetic parameters is then crucial in finding a rationale for the observed growth morphology.

In the MBE growth of arsenide compound semiconductors it is the surface diffusion that mainly governs the incorporation of the cation species, e.g. Ga, In, Al, whereas the kinetics of arsenic incorporation is dominated by adsorption/desorption of As_2 or As_4 molecules at surface sites with enhanced local population of cations [8]—an observation dating back to the dawn of MBE experiments [23].

Typically, surface adatom mobility can be understood on the basis of the adatom’s binding specifics at the surface accessible from the relevant potential-energy surface(s) (PES). This energy “map” is defined as the minimized difference

$$E_b(X, Y) = \min_{\mathbf{R}' \subseteq \mathbf{R}} \min_Z E^{\text{tot}}(\mathbf{R}, \mathbf{R}^{\text{ad}}) - E_{\text{slab}}^{\text{tot}} - E_{\text{atom}}. \quad (9)$$

In this expression $E^{\text{tot}}(\mathbf{R}, \mathbf{R}^{\text{ad}})$ stands for the total energy of the adatom/surface system; $E_{\text{slab}}^{\text{tot}}$ and E_{atom} are those of the bare surface (modeled by a slab) and the isolated atom, respectively. Minimization is carried out with respect to the height of the adatom Z , and a given subset $\mathbf{R}'$ or eventually all coordinates $\mathbf{R}$ of the substrate atoms, $\mathbf{R}^{\text{ad}} = (X, Y, Z)$ being the adatom coordinates. Further details can be found, e.g. in Refs. [10, 20, 24].

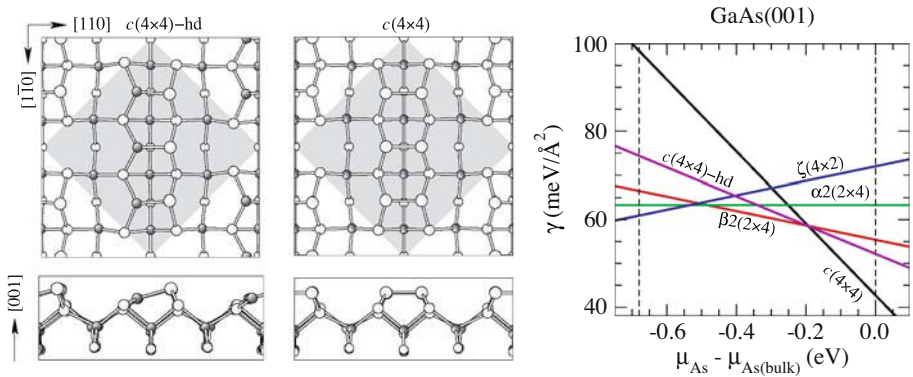


Figure 3. Equilibrium phase diagram of the GaAs(001) surface (right panel) calculated within the local-density approximation. Dashed lines mark the physically allowed range of variation of μ_{As} as determined from Eqs. (7) and (8). Structural models (topmost 4 atomic layers) of the conventional $c(4 \times 4)$ reconstruction is shown on the middle, and the $c(4 \times 4)$ -hd heterodimer model is shown on the left (Ga: shaded atoms; As: bright atoms). Shaded squares represent the surface unit cell.

3. Surface Phases under Strain

3.1 Clean GaAs(001) and InAs(001) Surfaces

Recent theoretical work [20, 25, 26, 27, 28, 29] has established that for growth conditions ranging from very As-rich to very Ga-rich the stable reconstructions of the GaAs(001) and InAs(001) surfaces follow the sequence $c(4 \times 4)$, $\beta 2(2 \times 4)$, $\alpha 2(2 \times 4)$, and $\zeta(4 \times 2)$. The calculated surface phase diagram of GaAs(001), for example, is shown in Fig. 3. However, so far only partial information is available about the influence of strain on the stability of these reconstructions (for InAs(001) see Ref. [30]). Furthermore, a very recent experiment [31] has also uncovered a new structural model for the $c(4 \times 4)$ reconstruction in the case of GaAs(001), see Fig. 3. To clarify this issue, we have performed calculations of the surface free energy for all four reconstructions for isotropically strained (001) substrates [29]. For small strain values ϵ one can write γ in the form

$$\gamma(\epsilon) = \gamma(0) + \text{Tr}(\boldsymbol{\tau})\epsilon + \mathcal{O}(\epsilon^2), \quad (10)$$

where $\boldsymbol{\tau}$ is the *intrinsic* surface stress tensor [32]. From the slope of $\gamma(\epsilon)$ at $\epsilon = 0$, we conclude that $\text{Tr}(\boldsymbol{\tau})$ is *positive* for all reconstructions of both GaAs(001) and InAs(001). In Fig. 4, we show the regions of stability of the reconstructions in the $(\mu_{\text{As}}, \epsilon)$ plane. For GaAs(001), at positive ϵ

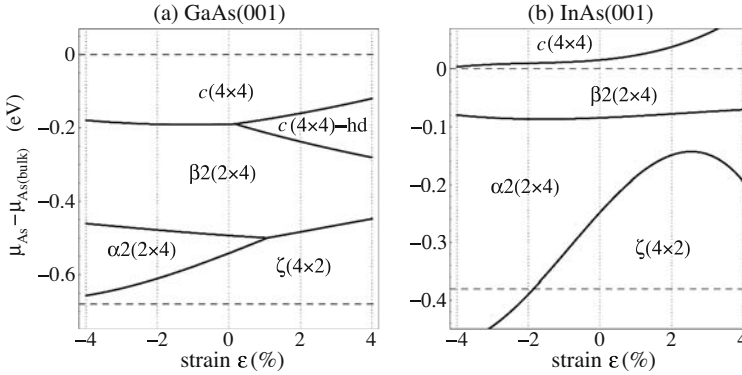


Figure 4. Diagrams of surface phases of (a) GaAs(001), and (b) InAs(001) surfaces as a function of μ_{As} and isotropic strain ϵ . Dashed lines mark the physically allowed range of variation of μ_{As} .

the “heterodimer” $c(4 \times 4)$ model, hereafter referred to as $c(4 \times 4)\text{-hd}$, appears as a stable reconstruction between the conventional $c(4 \times 4)$ and $\beta 2(2 \times 4)$ reconstructions. In MBE of InAs on GaAs, typically performed under As-rich conditions, the GaAs(001) surface displays the $c(4 \times 4)$ reconstruction, Fig. 3, which is characterized by blocks of three As dimers, or three Ga-As heterodimers, sitting on top of a complete As layer.

It is also remarkable that the stability range of $\zeta(4 \times 2)$ increases noticeably for $\epsilon > 0$, especially for InAs(001), while the $\alpha 2(2 \times 4)$ range increases for large negative strain. The latter trend is in agreement with the results of Ref. [30], although quantitative differences can be noticed. Thus, one could expect the InAs(001) surface at the heteroepitaxial strain of $\simeq -7\%$ to expose the $\alpha 2(2 \times 4)$ reconstruction over nearly the whole range of μ_{As} .

3.2 Wetting Layer

The initial delivery of InAs to the GaAs(001) substrate leads to formation of a microscopic pseudomorphic film, the so-called wetting layer (WL). Experimentally it is found that a complete layer structurally different from the substrate appears already after deposition of $1/3\text{--}2/3$ ML of InAs, depending on T . The appearance of this new structure is accompanied by a change in the RHEED pattern which gives evidence for a period of three lattice constants in the $[110]$ direction. Other probes, such as RDS [33, 34], STM and X-ray diffraction, confirm this structural transition [35, 36]. However, there is also evidence for some asymme-

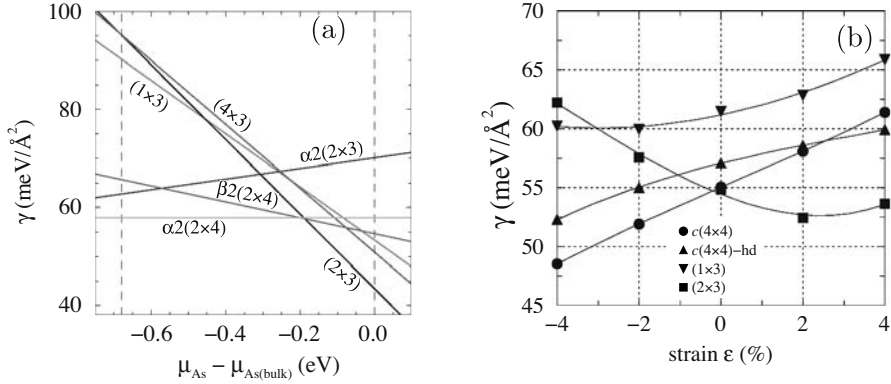


Figure 5. (a) Formation energy of $\text{In}_{2/3}\text{Ga}_{1/3}\text{As}(001)$ film with different reconstructions as a function of μ_{As} , cf. Ref. [15]. (b) Surface energy of $\text{GaAs}(001)\text{-}c(4 \times 4)$ and energy of formation of (1×3) - and (2×3) -reconstructed $\text{In}_{2/3}\text{Ga}_{1/3}\text{As}(001)$ film as functions of isotropic strain ϵ for $\mu_{\text{As}} - \mu_{\text{As(bulk)}} = -0.15$ eV. The lines are polynomial fits to the calculated points [29].

try/disorder in the diffraction patterns. A rationale for the asymmetry and the triple periodicity has been proposed a long time ago [37], based on domain formation.

Here we explore the appearance of the $\times 3$ periodicity as directly related to formation of a ternary InGaAs surface alloy with this specific surface unit cell. For InAs deposition of $2/3$ ML, a structural model with a (2×3) unit cell (see Fig. 6 (b) below) has been suggested on the basis of X-ray diffraction data [36, 38]. It consists of (continuous) rows of As dimers running in the $[\bar{1}10]$ direction, similar to the $c(4 \times 4)$ reconstruction of pure $\text{GaAs}(001)$. The indium atoms are located solely on lattice sites in the “trenches” between the chains of As dimers, avoiding the lattice sites directly below the As dimers. Thus one could speak of this film as a single ML of $\text{In}_{2/3}\text{Ga}_{1/3}\text{As}$ alloy. Alternatively, a (1×3) model has been also proposed [34].

The calculated formation energy, Fig. 5 (a), for the (2×3) structural model is found to be very close to the surface energy of the $c(4 \times 4)$ substrate, over a wide range of μ_{As} . For $-0.2 \text{ eV} < \mu_{\text{As}} - \mu_{\text{As(bulk)}} < -0.1 \text{ eV}$, the (2×3) structure is slightly more stable. We thus confirm the latter as the basic subunit of the WL. The (1×3) and other variants with different periodicity in the $[\bar{1}10]$ direction, see Fig. 5 (a), turn out to be higher in energy and can be discarded [15].

Another important feature of the pseudomorphic $\text{In}_{2/3}\text{Ga}_{1/3}\text{As}(001)$ film shows up upon considering γ as a function of the applied isotropic strain ϵ . From the slope of γ at $\epsilon = 0$ we infer that the dominant $\tau_{\alpha\beta}$ com-

ponent is *compressive*, $\text{Tr}(\boldsymbol{\tau}) < 0$, in contrast to those of pure GaAs(001) and InAs(001), as well as InGaAs(001)-(1 × 3). In Fig. 5 (b), we compare the formation energy of the WL with the surface energy of the bare GaAs(001)-c(4 × 4) substrate. The different surface stress of these surfaces renders the bare substrate more stable at compressive strain, while the (2 × 3) film is preferred for $\epsilon > 0$. We do not find conditions which render the (1 × 3) or the c(4 × 4)-hd reconstructions the most stable. Consequently, thermodynamics combined with our first-principles results imply suppression of wetting for 2/3 ML of InAs on the GaAs if the substrate is (locally) under mechanical compression, while the same amount of deposited InAs is sufficient to form a homogeneous WL on an unstrained substrate.

4. Surface Diffusion of Indium

Consider now the problem of indium migration in the process of InAs deposition on GaAs(001). In Ref. [42] we have already studied in great detail In diffusion on the GaAs(001)-c(4 × 4) surface. It is governed by the PES shown in Fig. 6 (a), from which one finds that the strongest binding site ($\mathbf{A}_1$) is located at the missing dimer position in the c(4 × 4) structure. In a first approximation, the adatom migration can be described as jumps between the $\mathbf{A}_1$ sites crossing the saddle points $\mathbf{T}_1$ leading to an isotropic diffusion characterized with energy barrier of $\Delta E = 0.65$ eV. At higher temperatures, a slight anisotropy may result upon inclusion of longer jumps across the blocks of 3 As dimers, passing the $\mathbf{T}_2$ - $\mathbf{A}_2$ - $\mathbf{T}_2$ sites, thus increasing the indium diffusion coefficient along the [110] direction. Let us note that the In/GaAs(001)-c(4 × 4) system is suitable for modeling the arrival of the first In atom to the GaAs surface in the initial stages of InAs deposition (or eventually to the GaAs capping layer in growth of multilayer QD structures), but the results cannot be taken over directly to the later stages of island growth.

As a next step we focus on In migration under those conditions where alloying in the wetting layer leads to a surface atomic arrangement with threefold periodicity along the [110] surface direction. More precisely, we have studied surface diffusion on the InGaAs(001)-(2 × 3) pseudomorphic film. While the latter may not exactly represent the typical wetting layer encountered in experiments [39, 40], we can learn from this system about the effect of alloying on the diffusivity by comparing to the case of In diffusion on GaAs(001)-c(4 × 4). It can be viewed as a particular case of surface diffusion where the adatom migrates on a homogeneously strained submonolayer film.

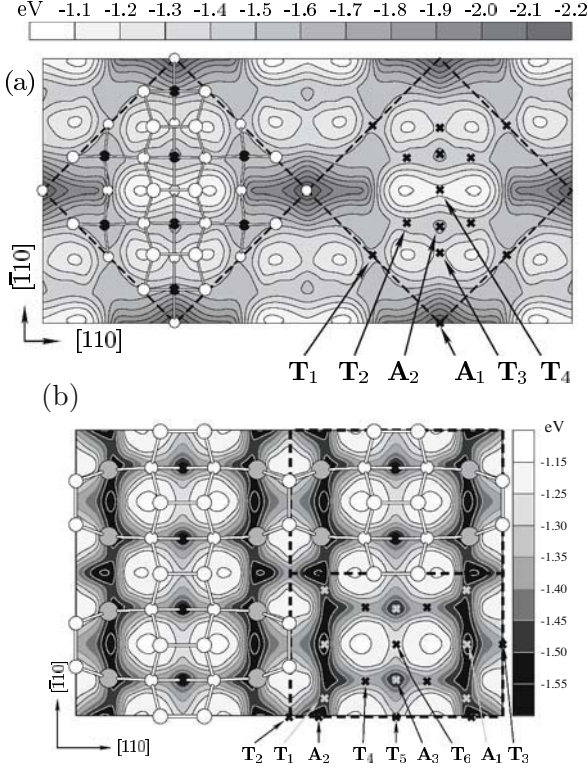


Figure 6. (a) PES for an In adatom on the GaAs(001)-c(4×4) surface, cf. Ref. [42]. Atomic positions in the clean surface unit cell are indicated for atoms in the upper four layers (As: empty circles; Ga: filled circles). (b) PES for an In adatom on the $\text{In}_{2/3}\text{Ga}_{1/3}\text{As}(001)$ -(2×3) surface. Two unit cells are indicated by dashed boxes. Overlaid on the PES plot are the topmost 3 atomic layers (In: shaded circles).

The PES for an indium adatom on the $\text{In}_{2/3}\text{Ga}_{1/3}\text{As}(001)$ -(2×3) surface is shown in Fig. 6 (b). The corrugation of this PES is remarkably small: The maximum variation of the adiabatic potential in the (001) plane is $\simeq 0.5$ eV. We find 3 symmetry-inequivalent potential minima and 6 saddle points: The energy barriers are all smaller than 0.3 eV. The adsorption site providing strongest binding, $\mathbf{A}_1$, is located between a top-layer As dimer and the As dimer bound to the third-layer In atoms. Another adsorption site, $\mathbf{A}_2$, appears next to a top-layer As dimer, but located in the gap between two As dimers bound to the third-layer In atoms. The troughs in the continuous As dimer row in $[\bar{1}10]$ direction give rise to a shallower site, $\mathbf{A}_3$. Detailed analysis of In diffusivity will be given elsewhere.

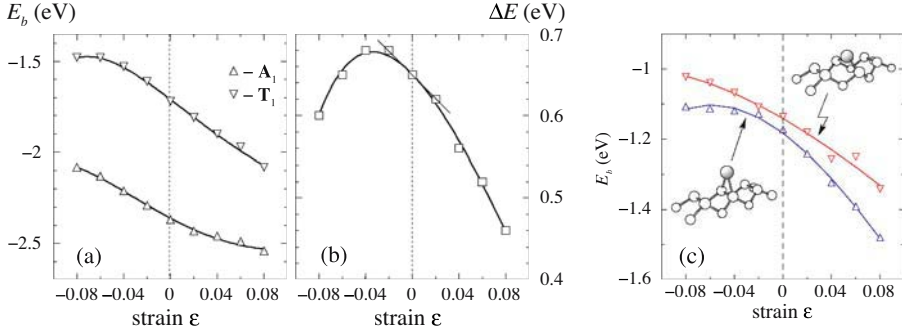


Figure 7. (a) Binding energy E_b as a function of isotropic strain ϵ for an indium adatom at the $\mathbf{A}_1$ and $\mathbf{T}_1$ sites, cf. Fig. 6 (a). (b) diffusion barrier $\Delta E \equiv E_b(\mathbf{T}_1) - E_b(\mathbf{A}_1)$ as a function of ϵ . Full curves on both panels represent least-squares polynomial fits to the calculated points [42]. (c) Binding energy of an indium adatom for the depicted bonding configurations.

Thus the typical energy scales for the potential-energy surfaces on the bare substrate and the pseudomorphic film turn out clearly different. If we define an onset temperature for diffusion by demanding that a single jump should occur at least once per second, the onset of In diffusion on the $\text{In}_{2/3}\text{Ga}_{1/3}\text{As}(001)$ film in the $[110]$ direction occurs at a temperature about 130 K lower than on the $\text{GaAs}(001)\text{-c}(4 \times 4)$ surface. For diffusion in the $[\bar{1}10]$ direction, the onset temperature is even lower by 190 K, compared to the $\text{GaAs}(001)\text{-c}(4 \times 4)$ surface. As we have not found evidence for the so-called compensation effect [41] on the attempt frequencies, one can predict a considerably higher In mobility on the $\text{In}_{2/3}\text{Ga}_{1/3}\text{As}(001)$ film as compared to the $\text{GaAs}(001)\text{-c}(4 \times 4)$ substrate.

Additionally, we have discovered [42] that In diffusivity on the $\text{c}(4 \times 4)$ -reconstructed GaAs substrate is considerably affected by the presence of mechanical strain. This situation is pertinent, e.g. to indium diffusion on a GaAs capping layer with buried QDs, acting as stressors. The strain dependence of the adatom binding at the adsorption sites $\mathbf{A}_1$ and the lowest saddle point connecting them, $\mathbf{T}_1$, is shown in Fig. 7 (a). While for $\epsilon < 0$ E_b at the adsorption site $\mathbf{A}_1$ follows approximately a linear law with a slope of -3.8 eV, the binding energy at $\mathbf{T}_1$ contains small non-linear terms in strain which do not cancel in the evaluation of the diffusion barrier $\Delta E = E_b(\mathbf{T}_1) - E_b(\mathbf{A}_1)$. Thus, ΔE for $\epsilon < 0$ is accurately described by

$$\Delta E(\epsilon) = 0.65\text{eV} + 0.03\text{eV} \times \left[1 - \left(\frac{\epsilon}{0.03} + 1 \right)^2 \right], \quad \epsilon < 0. \quad (11)$$

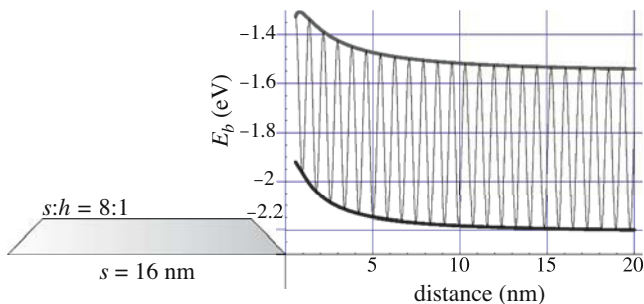


Figure 8. Migration potential (oscillating curve) for an In adatom approaching perpendicular to a very long, coherently strained InAs island (of width s and height h) on the $c(4 \times 4)$ -reconstructed GaAs(001) surface. In addition to the diffusion potential due to the atomic structure of the surface, the strain field in the substrate induced by the island gives rise to a repulsive potential that lifts both the binding energies (thick lower line) and transition state energies (thick upper line) close to the island, cf. Ref. [46]

It should be noted, however, that for small values of the strain, the diffusion barrier is to a very good accuracy linear in the strain [43, 44], $\Delta E \propto \epsilon$, as can be seen from $\Delta E(\epsilon)$ in Fig. 7 (b) in the range $\pm 2\%$. For an inhomogeneously strained sample, the pronounced strain dependence of E_b for both the adsorption site and the saddle point will introduce a position dependence of ΔE .

In order to illustrate this effect we have considered, within the flat-island approximation [45], a simple model [20, 46] for an InAs island on a GaAs(001)- $c(4 \times 4)$ substrate, Fig. 8. As can be seen from the figure, the effect of strain leads to a repulsive potential with a strength of up to 0.2 eV, that affects both the binding energy and, to a slightly smaller extent, the diffusion barriers for an In adatom that attempts to approach this island. This repulsive interaction can significantly slow down the speed of growth of strained islands. For a simple example, where two very elongated islands compete for the flux of In atoms deposited between them, we have found that strain-controlled diffusion indeed tends to equalize the size of the two islands while they are growing [42]. This effect has been proposed as one of the factors responsible for the narrowing of the island size distribution that is desirable from the point of view of the applications [47, 48].

5. Outlook

Above we have tried to develop a microscopic picture of the early stages of quantum dot formation. At the fundament of our knowl-

edge lies an understanding of the atomistic processes in epitaxial growth which are best explored through suitably conducted first-principles DFT calculations. In this contribution, we have attempted to demonstrate the application of the latter to the specific problems of surface atomic structure and adatom diffusivity in InAs/GaAs(001) heteroepitaxy. From the typically used growth temperatures and the calculated energy barriers for diffusion, we conclude that the time and length scales involved in quantum dot formation in InAs heteroepitaxy on GaAs(001) span a few orders of magnitude. The importance of a wide range of length and time scales is a very general phenomenon in epitaxial growth. It is therefore clear that the theoretical description of this phenomenon requires more than a single theoretical tool. DFT calculations must be complemented by calculations on larger scales using elasticity theory or analytical interatomic potentials. Kinetic Monte Carlo simulations are able to integrate input from various sources, and enable us to bridge the gap between the atomic scale and experimentally relevant scales. The results from our first-principles calculations constitute important input to such kinetic Monte Carlo simulations. The feasibility of this method has already been demonstrated for homoepitaxy of GaAs [15, 46]. A typical problem where its usage is indispensable is the treatment of structural disorder in the wetting layer, observed in some samples. Future research will therefore be focused on kinetic Monte Carlo simulations of In diffusion on a disordered wetting layer and possible consequences for QD growth.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft within the Sonderforschungsbereich 296. E. Penev would like to acknowledge fruitful discussions with F. Montalenti, O. Kirfel, I. Goldfarb, P. Machnikowski, and G. Goldoni during the NATO ARW on *Quantum Dots: Fundamentals, Applications, Frontiers*.

References

- [1] D. Bimberg, M. Grundmann, and N. N. Ledentsov. *Quantum dot heterostructures* (Wiley, 1998).
- [2] M. Kotrla, N. Papanicolaou, D. D. Vvedensky, and L. T. Wille, editors. *Atomistic Aspects of Epitaxial Growth* (Kluwer, Dordrecht, 2002).
- [3] I. N. Stranski and L. Krastanov. Zur Theorie der orientierten Ausscheidung von Ionenkristallen aufeinander. *Sitzungsber. Akad. Wiss. Wien, Math.-naturwiss. Kl. IIb* **146**, 797–810, 1938.
- [4] M. Sugawara, editor. *Self-assembled InGaAs/GaAs quantum dots* (Academic, New York, 1999).

- [5] B. A. Joyce and D. D. Vvedensky. Mechanisms and anomalies in the formation of InAs-GaAs(001) quantum dot structures, in *Atomistic Aspects of Epitaxial Growth*, edited by M. Kotrla and N. Papanicolaou and D. Vvedensky and L. Wille (Kluwer, Dordrecht, 2002), pp. 301–325.
- [6] T. R. Ramachandran, R. Heitz, P. Chen, and A. Madhukar. Mass transfer in Stranski-Krastanov growth of InAs on GaAs. *Applied Physics Letters* **70**: 640–642, 1997.
- [7] P. Ruggerone, C. Ratsch, and M. Scheffler. Density-functional theory of epitaxial growth of metals, in *Growth and Properties of Ultrathin Epitaxial Layers*, edited by D. A. King and D. P. Woodruff (Elsevier, Amsterdam, 1997) pp. 490–544.
- [8] P. Kratzer, C. Morgan, and M. Scheffler. Density-functional theory studies on microscopic processes of GaAs growth. *Prog. Surf. Sci.* **59**: 135–147, 1998.
- [9] For a large number of examples the reader is referred to the following URL: <http://www.fhi-berlin.mpg.de/th/paper.html>.
- [10] A. Kley. *Theoretische Untersuchungen zur Adatomdiffusion auf niederindizierten Oberflächen von GaAs* (Wissenschaft & Technik Verlag, Berlin, 1997).
- [11] C. G. Van de Walle and J. Neugebauer. First-principles surface phase diagram for hydrogen on GaN surfaces. *Phys. Rev. Lett.* **88**: art. no. 066103, 2002.
- [12] A. A. Stekolnikov, J. Furthmüller, and F. Bechstedt. Absolute surface energies of group-IV semiconductors: Dependence on orientation and reconstruction. *Phys. Rev. B* **65**: art. no. 115318, 2002.
- [13] K. Reuter and M. Scheffler. First-principles thermodynamics for oxidation catalysis: Surface phase diagrams and catalytically interesting regions. *Phys. Rev. Lett.* **90**: art. no. 046103, 2003.
- [14] F. Bechstedt. *Principles of Surface Physics* (Springer, Berlin, 2003).
- [15] P. Kratzer, E. Penev, and M. Scheffler. Understanding the growth mechanisms of GaAs and InGaAs thin films by employing first-principles calculations. *Appl. Surf. Sci.* **216**: 436–446, 2003.
- [16] R. J. Needs, R. M. Martin, and O. H. Nielsen. Total-energy calculations of the structural properties of the group-V element arsenic. *Phys. Rev. B* **33**: 3778–3784 (1986).
- [17] M. Bockstedte, A. Kley, J. Neugebauer, and M. Scheffler. Density-functional theory calculations for poly-atomic systems: Electronic structure, static and elastic properties and *ab initio* molecular dynamics. *Comp. Phys. Commun.* **107**: 187–222, 1997; URL: www.fhi-berlin.mpg.de/th/fhimd.
- [18] K. Shiraishi. A new slab model approach for electronic-structure calculation of polar semiconductor surface. *J. Phys. Soc. Jpn.* **59**: 3455–3458, 1996.
- [19] V. P. LaBella, H. Yang, D. W. Bullock, P. M. Thibado, P. Kratzer, and M. Scheffler. Atomic structure of the GaAs(001)-(2 × 4) surface resolved using scanning tunneling microscopy and first-principles theory. *Phys. Rev. Lett.* **83**: 2989–2992, 1999.
- [20] E. S. Penev. *On the theory of surface diffusion in InAs/GaAs(001) heteroepitaxy* (Technische Universität Berlin, 2002); URL: www.fhi-berlin.mpg.de/th/publications/PhD-penev-2002.pdf.

- [21] J. Tersoff, M. D. Johnson, and B. G. Orr. Adatom densities on GaAs: Evidence for a near-equilibrium growth. *Phys. Rev. Lett.* **78**: 282–285, 1997.
- [22] P. Kratzer, C. G. Morgan, and M. Scheffler. Model for nucleation in GaAs homoepitaxy derived from first principles. *Phys. Rev. B* **59**: 15246–15252, 1999.
- [23] C. T. Foxon and B. A. Joyce. Interaction kinetics of As₂ and Ga on {100} GaAs surface. *Surf. Sci.* **64**: 292–304, 1977.
- [24] M. Scheffler and P. Kratzer. *Ab initio* thermodynamics and statistical mechanics of diffusion, growth, and self-assembly of quantum dots, in *Atomistic Aspects of Epitaxial Growth*, edited by M. Kotrla and N. Papanicolaou and D. D. Vvedensky and L. T. Wille (Kluwer, Dordrecht, 2002) pp. 355–369.
- [25] S.-H. Lee, W. Moritz, and M. Scheffler. GaAs(001) surface under conditions of low As pressure: Evidence for a novel surface geometry. *Phys. Rev. Lett.* **85**, 3890–3893, 2000.
- [26] C. Ratsch, W. Barvosa-Carter, F. Grosse, J. H. G. Owen, and J. J. Zinck. Surface reconstructions for InAs(001) studied with density-functional theory and STM. *Phys. Rev. B* **62**: R7719–R7722, 2000.
- [27] W. G. Schmidt, S. Mirbt, and F. Bechstedt. Surface phase diagram of (2 × 4) and (4 × 2) reconstructions of GaAs(001). *Phys. Rev. B* **62**: 8087–8091, 2000.
- [28] W. G. Schmidt. III-V compound semiconductor (001) surfaces. *Appl. Phys. A* **75**: 89–99, 2002.
- [29] E. Penev, P. Kratzer, and M. Scheffler. (unpublished)
- [30] C. Ratsch. Strain-induced change of surface reconstructions for InAs(001). *Phys. Rev. B* **63**: art. no. 161306(R), 2001.
- [31] A. Ohtake, J. Nakamura, S. Tsukamoto, N. Koguchi, and A. Natori. New structure model for the GaAs(001)-c(4 × 4) surface. *Phys. Rev. Lett.* **89**: art. no. 206102, 2002.
- [32] W. Haiss. Surface stress of clean and adsorbate-covered solids. *Rep. Prog. Phys.* **64**: 591–648, 2001.
- [33] D. I. Westwood, Z. Sobiesierski, E. Steimetz, T. Zettler, and W. Richter. On the development of InAs on GaAs(001) as measured by reflectance anisotropy spectroscopy: Continuous and islanded films. *Appl. Surf. Sci.* **123/124**: 347–351, 1998.
- [34] T. Kita, O. Wada, T. Nakayama, and M. Murayama. Optical reflectance study of the wetting layer in (In, Ga)As self-assembled dot growth on GaAs(001). *Phys. Rev. B* **66**: art. no. 195312, 2002.
- [35] J. G. Belk, C. F. McConville, J. L. Sudijono, T. S. Jones, and B. A. Joyce. Surface alloying at InAs-GaAs interfaces grown on (001) surfaces by molecular beam epitaxy. *Surf. Sci.* **387**: 213–226, 1997.
- [36] Y. Garreau, K. Aid, M. Sauvage-Simkin, R. Pinchaux, C. F. McConville, T. S. Jones, J. L. Sudijono, and E. S. Tok. Stoichiometry and discommensuration on In_xGa_{1-x}As/GaAs(001) reconstructed surfaces: A quantitative x-ray diffuse-scattering study. *Phys. Rev. B* **58**: 16177–16185, 1998.
- [37] A. G. de Oliveira, S. D. Parker, R. Droopad, and B. A. Joyce. A generalized model for the reconstruction of {001} surfaces of III-V compound semiconductors based on a RHEED study of InSb(001). *Surf. Sci.* **227**: 150–156, 1990.

- [38] M. Sauvage-Simkin, Y. Garreau, R. Pinchaux, M. B. Véron, J. P. Landesman, and J. Nagle. Commensurate and incommensurate phases at reconstructed (In,Ga)As(001) surfaces: X-ray diffraction evidence for a composition lock-in. *Phys. Rev. Lett.* **75**: 3485–3488, 1995.
- [39] J. G. Belk, J. L. Sudijono, D. M. Holmes, C. F. McConville, T. S. Jones, and B. A. Joyce. Spatial distribution of In during the initial stages of growth of InAs on GaAs(001)-c(4×4). *Surf. Sci.* **365**: 735–742, 1996.
- [40] B. A. Joyce, D. D. Vvedensky, G. R. Bell, J. G. Belk, M. Itoh, and T. S. Jones. Nucleation and growth mechanism during MBE of III-V compounds. *Mater. Sci. Eng. B* **67**: 7–16, 1999.
- [41] D. J. Fisher. *The Meyer-Neldel Rule*. (Trans Tech Publications, Uetikon-Zürich, 2001).
- [42] E. Penev, P. Kratzer, and M. Scheffler. Effect of strain on surface diffusion in semiconductor heteroepitaxy. *Phys. Rev. B* **64**: art. no. 085401, 2001.
- [43] H. T. Dobbs, D. D. Vvedensky, A. Zangwill, J. Johansson, N. Carlsson, W. Seifert. Mean-field theory of quantum dot formation. *Phys. Rev. Lett.* **79**: 897–900, 1997.
- [44] H. T. Dobbs, A. Zangwill, and D. D. Vvedensky. Nucleation and growth of coherent quantum dots: a mean field theory, in *Surface Diffusion: Atomistic and Collective Processes*, edited by M. Tringides (Plenum, New York, 1998) pp. 263–275.
- [45] J. Tersoff and R. M. Tromp. Shape transition in growth of strained islands: Spontaneous formation of quantum wires. *Phys. Rev. Lett.* **70**: 2782–2785, 1993.
- [46] P. Kratzer, E. Penev, and M. Scheffler. First-principles studies of kinetics in epitaxial growth of III-V semiconductors. *Appl. Phys. A* **75**: 79–88, 2002.
- [47] A. Madhukar. A unified atomistic and kinetic framework for growth front morphology evolution and defect initiation in strained epitaxy. *J. Cryst. Growth* **163**: 149–164, 1996.
- [48] H. M. Koduvely and A. Zangwill. Epitaxial growth kinetics with interacting coherent islands. *Phys. Rev. B* **60**: R2204–R2207, 1999.

FORMATION OF TWO-DIMENSIONAL Si/Ge NANOSTRUCTURES OBSERVED BY STM

Bert Voigtländer

Institut für Schichten und Grenzflächen, Forschungszentrum Jülich GmbH.

52425 Jü

b.voigtlaender@fz-juelich.de

Abstract The growth of kinetically self-organized two-dimensional (2D) islands is described for Si/Si(111) epitaxy. The island size distribution for this system was measured using scanning tunneling microscopy (STM). The potential formation of thermodynamically stable strained islands of a specific size is discussed. The formation of 2D Si/Ge nanostructures at preexisting defects is studied. The step-flow growth mode is used to fabricate Si and Ge nanowires with a width of 3.5 nm and a thickness of one atomic layer (0.3 nm) by self-assembly. One atomic layer of Bi terminating the surface is used to distinguish between the elements Si and Ge. A difference in apparent height is measured in STM images for Si and Ge, respectively. Additionally, different kinds of 2D Si/Ge nanostructures, such as alternating Si and Ge nanorings having a width of 5–10 nm, were grown.

Keywords: Silicon, Germanium, nanostructures, nanowires, scanning tunneling microscopy

1. Introduction

Several approaches are used to fabricate nanostructures at surfaces. Relatively large structures are created by lithography. The greatest advantage of lithographic patterning is the very large variety of different structures that can be defined. On the other hand, lithographically defined structures are usually quite large. The resolution of optic lithography is approaching ~ 100 nm while the resolution of electron beam lithography reaches down even further. However, since writing with electron beam lithography is a sequential method it is not possible to fill a whole wafer with small structures in a reasonable time. Due to the limitations of the lithographic methods to fabricate nanostructures, new approaches

are explored for the formation of smaller nanostructures. The most extreme approach is to build nanostructures “bottom-up” atom by atom with the help of a scanning tunneling microscope at low temperatures [1]. This approach is the ultimate in terms of the size of a nanostructure. Of course this is a very slow method to build nanostructures: usually, a couple of hours are required to write some characters.

An alternative approach for the formation of small nanostructures is self-organization of atoms. Under certain conditions initially disordered atoms tend to form nanostructured clusters, crystallites, or wires at a surface. Here, three different kinds of self-organization of nanostructures at surfaces are considered. First, the formation of kinetically self-organized nanostructures formed during molecular beam epitaxy will be discussed. The island structures formed by kinetic self-organization can be quite small, though the size uniformity is a challenge. A second topic will be to show approaches for the formation of thermodynamically (meta) stable nanostructures. Here, the size uniformity should be best. A third approach for the ordered formation of nanostructures is by the nucleation of islands at defect sites which pre-exist at (or below) the surface. These defect sites can be, for instance, dislocations or step bunches.

2. Kinetically Self-Organized Islands

During molecular-beam epitaxy (MBE) growth of silicon, atoms deposited onto the surface diffuse on the surface and can be incorporated into preexisting step edges if the diffusion length is sufficient (Fig. 1). When the diffusion is slower diffusing atoms will meet and an island will nucleate on the surface. The island will grow further by attachment of diffusing atoms at the island [2] (Fig. 1). The islands that form after deposition of less than a atomic layer (monolayer ML) of Si are called two-dimensional (2D) islands, because their thickness is only one (or two) atomic layers. The islands grow epitaxially, which means that they arrange in a crystalline structure in registry with the underlying substrate atomic structure. The shape of the islands depends on the atomic structure of the underlying substrate surface. For the case of Si(111), the substrate crystal structure has a threefold rotational symmetry at the surface. Therefore, the 2D Si islands on Si(111) have triangular shape (Fig. 2).

Two-dimensional islands are the simplest example of self-organized growth of nanostructures. In the following it will be shown how the density and the size of these islands can be controlled by the kinetic parameters temperature and growth rate. First, the deposition temper-

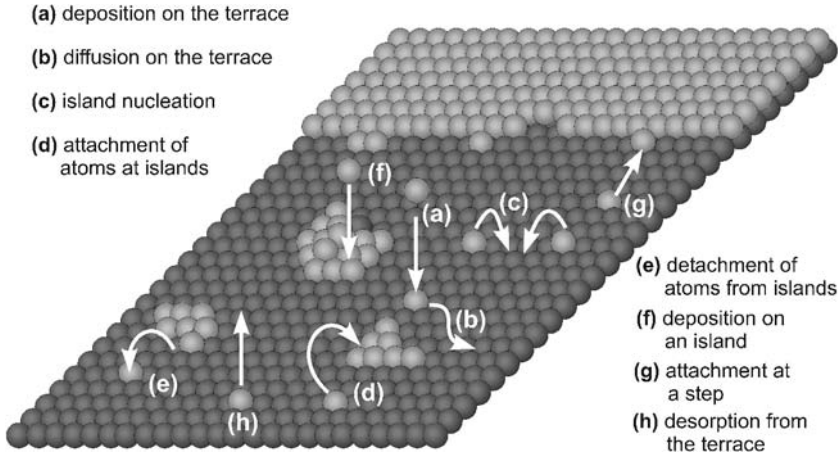


Figure 1. Schematic representation of different fundamental processes occurring during epitaxial growth leading to a self-organization of two-dimensional islands.

ature influences the island density strongly, as shown by the comparison of Fig. 2(a) and (b). When the island density is plotted as function of temperature [Fig. 3(a)], it can be seen that the island density follows an Arrhenius law: $n \sim \exp(E_{\text{act}}/kT)$, with E_{act} being an effective activation energy consisting of a diffusion energy and binding energy component. The temperature is one important parameter of growth kinetics, the deposition rate is another. It is found that the island density scales with the deposition rate in form of a power law: $n \sim F^\omega$ with a scaling exponent ω . From Fig. 3(b), the scaling exponent is determined to be $\omega = 0.75$ for Si/Si(111) homoepitaxy [3]. Combining the temperature and the rate dependence results in the following scaling law: $n \sim F^\omega \exp(E_{\text{act}}/kT)$ [4]. This shows that the island density can be controlled over a wide range by adjusting the kinetic growth parameters temperature and growth rate. The average island distance is just the square root of the inverse of the island density, $L = 1/\sqrt{n}$.

The nucleation of the islands is a random process. Despite this, the distribution of the island sizes has a peak (Fig. 4). This arises due to a saturation of the island nucleation, as will be explained in the following. In the early stage of growth islands nucleate randomly on the surface and the distance between the islands decreases. If the distance between the islands is equal to the mean distance that an adatom travels before a nucleation event happens, then the incorporation of adatoms in existing islands becomes a more probable event than the nucleation of new islands. Around each island a "capture zone" exists. Adatoms deposited in this capture zone attach to the corresponding island. Without

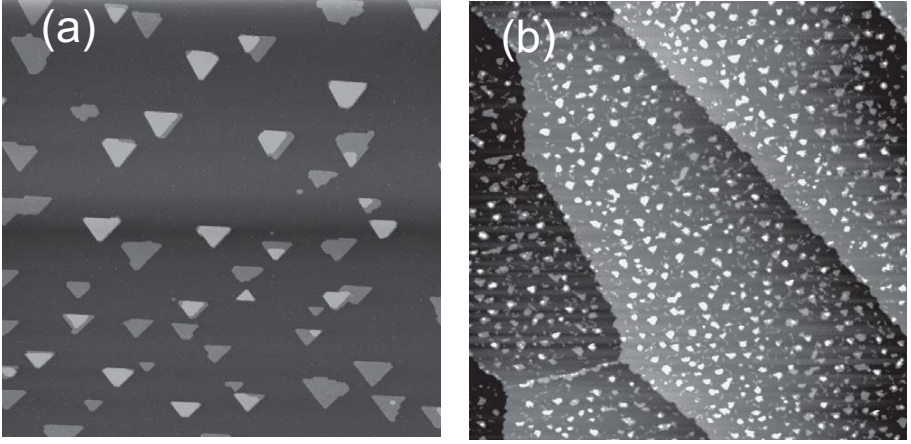


Figure 2. Scanning tunneling microscope images after the growth of 0.2 atomic layers of silicon on a Si(111) surface. The islands have triangular shape due to the symmetry of the substrate and have a height of one atomic layer (dark grey) or two atomic layers (light grey). The island density depends on temperature, as can be seen by comparison of growth at high temperatures of 500°C (a) to growth at a lower temperature 340°C (b). Both images have a size of 350 nm.

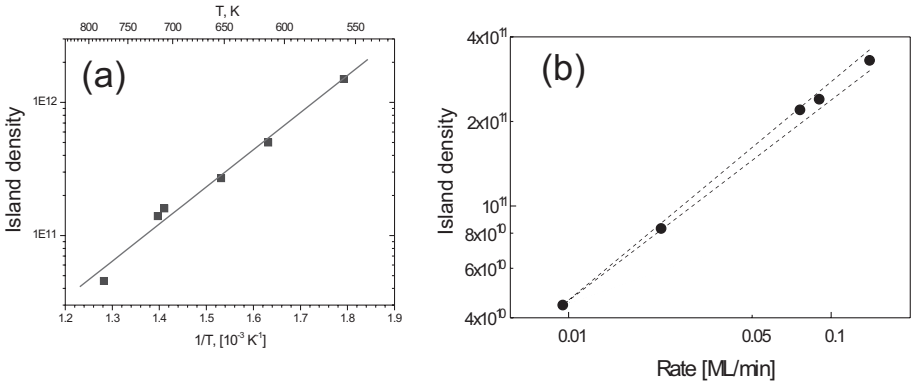


Figure 3. The island density shows an Arrhenius behavior as function of temperature. (a) As a function of deposition rate a power law is found for the island density. (b) This shows that the island density can be controlled by the kinetic parameters temperature and deposition rate. The island density is given in cm^{-2} .

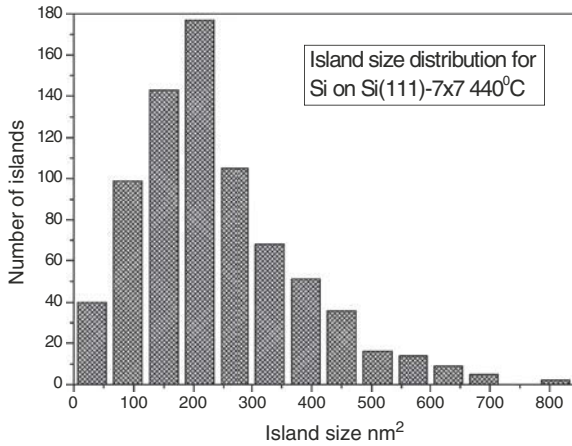


Figure 4. Island size distribution for 2D Si islands on Si(111). The width of the distribution is of the order of the average size of the islands.

this effect the distribution of island sizes would be even broader. The nucleation of further islands is suppressed beyond a certain coverage. The effect of the capture zone gives rise to two regimes in the island growth. The very early regime (nucleation regime) is characterized by the nucleation of islands (< 0.1 ML deposited). In the second regime the capture zones of existing islands overlap and the nucleation of new islands ceases. This regime is characterized by growth of existing islands to a larger size and is called the growth regime. The average island size can be controlled by the deposited amount.

In summary, the island density of two-dimensional islands can be controlled by the kinetic parameters temperature and deposition rate, while the size distribution is quite broad due to the stochastic nature of the nucleation of the islands.

3. Thermodynamically Stable Nanostructures

If nanostructure islands would be thermodynamically stable their size distribution could be narrow. A thermodynamically stable island size means that the energy (per atom) has a minimum for this stable size. For configurations with larger or smaller islands the energy (per atom) would be higher. Therefore, one has only to approach thermodynamic equilibrium to obtain a very narrow island size distribution. One way to achieve thermodynamic equilibrium is to heat a sample with different island sizes present and wait until equilibrium has established. The equilibrium configuration will be established by material transport be-

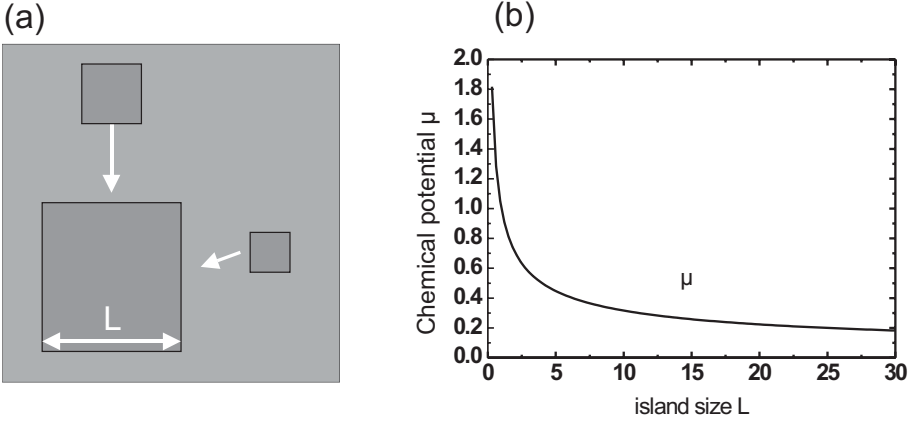


Figure 5. (a) Coarsening of a large island at the expense of small ones. (b) Chemical potential of an island.

tween the islands. Atoms will detach from islands with higher energy and attach to islands with a lower energy (per atom).

To describe material transport in a system with a variable number of atoms the chemical potential is used. The chemical potential is the change of the energy (of an island) when the number of particles changes: $\mu = dE/dN$. During equilibration atoms detach from islands where the chemical potential is highest and attach to islands with a lower chemical potential [6]. This lowers the total energy of the system, so the material transport between different islands is governed by the chemical potential. A simple example is the chemical potential of quadratic 2D islands of dimension L [Fig. 5(a)]. The energy difference between different sized islands comes from the edge energy (β is the edge energy per length). The energy of an island is $E = E_{\text{edge}} = 4L\beta$. The number of atoms N in an island depends on the dimension L as $N = L^2/\Omega$, with Ω being the area per atom. The chemical potential is

$$\mu = \frac{dE}{dN} = \frac{2\Omega\beta}{L} \sim \frac{1}{L}. \quad (1)$$

Since μ is decreasing for larger islands [Fig. 5(b)] infinite-size islands have the lowest chemical potential. This means that the stable island is infinitely large. In this case, equilibration does not result in a stable finite island size. Equilibration in this model by material transport between islands is also called coarsening because it results in the shrinkage of small islands and a growth (coarsening) of large islands.

An infinitely large stable island size is the result for homoepitaxial growth, taking only the step energy into account. The situation is different when elastic stress is taken into account, as occurs in heteroepitaxy. Here, stress is induced by the different lattice constants of the substrate material and the material of the islands. The elastic effect of strained 2D islands can be approximated by that of a surface stress domain, i.e. the surface stress at the area of the island is different from that at the rest of the surface (Fig. 5). The strain energy of a quadratic surface stress domain can be calculated using the elastic theory as $E_{\text{strain}} = 2LC'\ln L$ [5]. Adding the step edge energy results in a total energy of a strained island given by

$$E = E_{\text{edge}} + E_{\text{strain}} = 2L(2\beta - C' \ln L), \quad (2)$$

and a chemical potential

$$\mu = \Omega \left(\frac{2\beta - C'}{L} - \frac{C'}{L} \ln L \right), \quad (3)$$

which is plotted in Fig. 6. In this case the chemical potential has a minimum at $L_{\text{min}} = \exp(2\beta/C')$. This means that during coarsening the islands would approach this size. Larger islands would dissolve and smaller islands grow until all islands have size L_{min} , i.e. the lowest chemical potential. This would result in a very narrow size distribution. Unfortunately, step energies are poorly known, so that it is not possible to reliably predict the equilibrium island size. An experimental realization of a thermodynamically stable islands has not yet been confirmed.

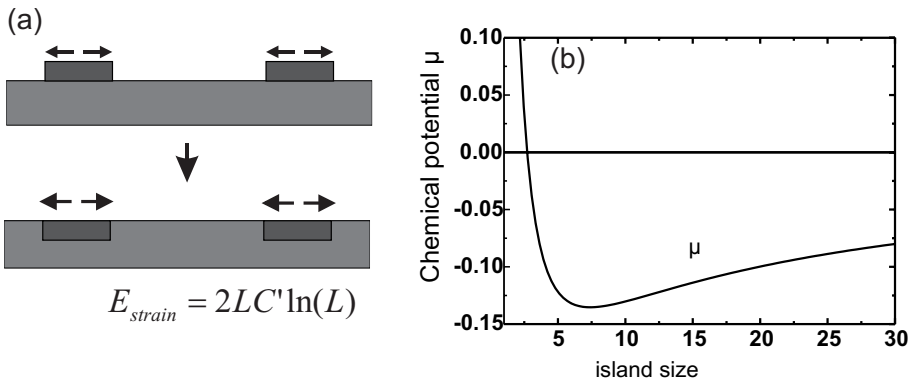


Figure 6. (a) The elastic stress induced by 2D islands with a different lattice constant than the substrate can be approximated by surface stress domains. (b) Chemical potential of an island with a energy component due to elastic strain included.

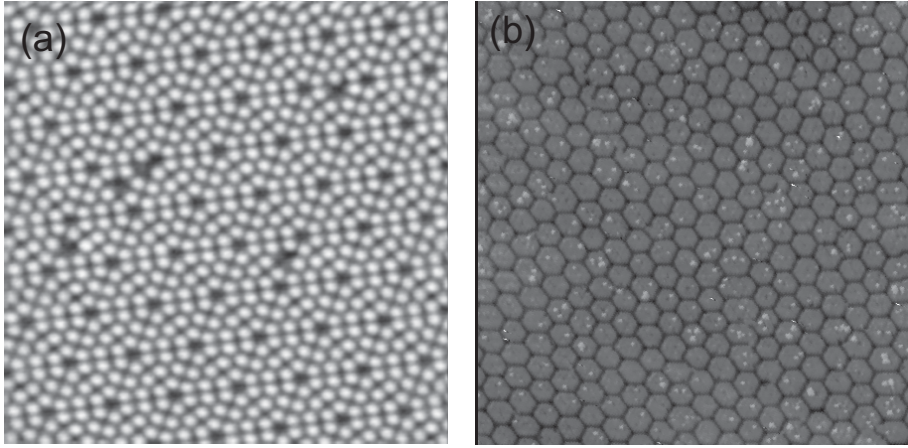


Figure 7. (a) STM image of the 7×7 reconstruction on Si(111) (periodicity 2.7 nm). (b) Hexagonal surface reconstruction on a Sb/Ge/Si(111) surface (periodicity 4 nm). Surface reconstructions are the most regular nanostructures at surfaces.

The most perfect thermodynamically stable nanostructures are surface reconstructions with a relatively large unit cell. The atomic arrangement at surfaces, especially at semiconductor surfaces often deviates strongly from the atomic structure expected for a bulk terminated surface obtained when the atomic arrangement of the surface atoms remains like the bulk structure [2]. The atoms at the surface rearrange (reconstruct) in order to reduce the number of unsaturated bonds (dangling bonds) at the surface. The most famous surface reconstruction with a large unit cell is that of the Si(111)-(7×7) surface shown in Fig. 7(a). The length of the triangular subunits is 2.7 nm. These are very small nanostructures because the main driving force for the formation of these structures is the minimization of the energy of the covalent bonds at the surface. Covalent bonding is a short range interaction compared to elastic interactions considered above. Another example for a surface reconstruction with a large periodicity is a 3 ML thick Ge layer terminated with Sb [7]. This layer forms a hexagonal arrangement with a period of 4 nm (Fig. 7(b)). Here both covalent bonding and the misfit strain induced by the larger lattice constant of the Ge are important for the formation of the hexagonal nanostructures.

In the following we compare the formation of nanostructures in equilibrium to the formation of nanostructures by growth kinetics. Equilibrium nanostructures have the advantage of a narrow size distribution around the optimum size. A disadvantage is that the size is determined

by the material parameters (strain energy and step edge energy for instance) and can not be tuned independently. The size and density of nanostructures formed under kinetic conditions can be tuned easily by variation of the growth parameters. On the other hand, the size uniformity of the islands is relatively poor.

4. Formation of Nanostructures at Pre-existing Defects

Up to now, several approaches for the self-organized growth of islands were discussed. Here, results on the formation of nanowires grown at substrate steps are described. Pre-existing step edges on the Si(111) surface act as templates for the growth of two-dimensional Ge wires at the step edges. When the diffusion of the deposited atoms is sufficient to reach the step edges, these deposited atoms are incorporated exclusively at the step edges and the growth proceeds by a homogenous advancement of the steps (step flow growth mode) [2]. If small amounts of Ge are deposited the steps advance only some nanometers and narrow Ge wires can be grown. A key issue for the controlled fabrication of nanostructures consisting of different materials is a method of characterization which can distinguish between the different materials on the nanoscale. In case of the important system Si/Ge it has been difficult to differentiate between Si and Ge due to their similar electronic structure. However, if the surface is terminated with a monolayer of Bi it is possible to distinguish between Si and Ge.

Figure 8(a) shows an STM image after repeated alternating deposition of 0.15 atomic layers of Ge and Si, respectively. Due to the step flow growth Ge and Si wires are formed at the advancing step edge [8]. Both elements can be easily distinguished by the apparent heights in the STM images. It turned out that the height measured by the STM is higher on areas consisting of Ge (light grey stripes) than on areas consisting of Si (dark grey stripes). The assignment of Ge and Si wires is evident from the order of the deposited materials (Ge, Si Ge, Si and Ge, respectively in this case). The initial step position is indicated by white arrows in the right part of Fig. 8(a). The step edge has advanced towards the left [arrowheads in Fig. 8(a)] after the growth of the nanowires. The reason for the height difference in STM between Si and Ge is the different electronic structure of a Ge-Bi bond compared to a Si-Bi bond. The apparent height of Ge areas is ~ 0.1 nm higher than the apparent height of Si wires [Fig. 8 (b)]. The width of the Si and Ge wires is ~ 3.5 nm as measured from the cross section [Fig. 8 (b)]. The nanowires are two-dimensional with a height of only one atomic layer (~ 0.3 nm).

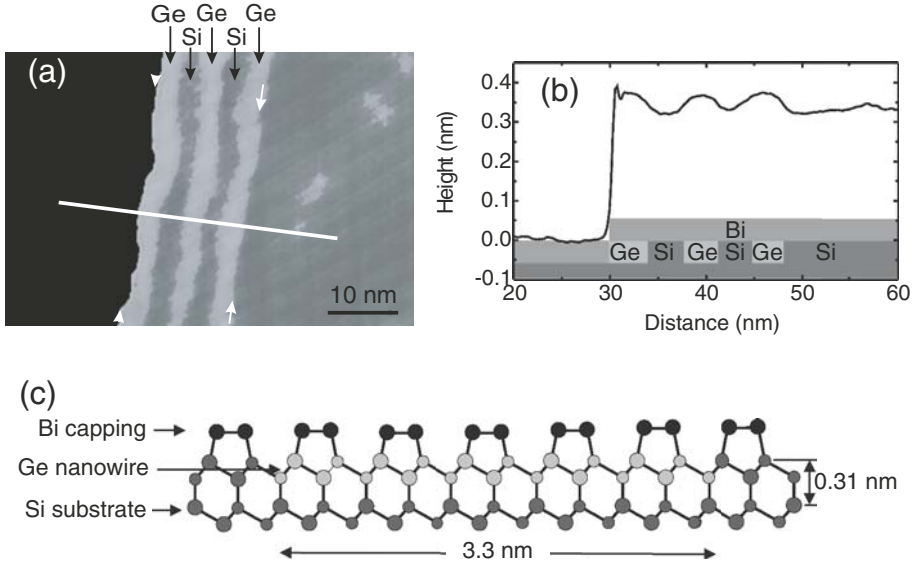


Figure 8. (a) STM image of 2D Ge/Si nanowires grown by step flow at a pre-existing step edge on a Si(111) substrate. Si wires (dark grey) and Ge wires (light grey) can be distinguished by different apparent heights. (b) The cross section along the white line in (a) shows the dimensions of the Si and Ge nanowires. The width of the wires is ~ 3.5 nm and the height is only one atomic layer (0.3 nm). (c) Atomic structure of a 3.3 nm wide Ge wire on the Si substrate capped by Bi. The cross section of the Ge wire contains only 21 Ge atoms.

Therefore, the cross section of a 3.3 nm wide Ge nanowire contains only 21 atoms [Fig. 8(c)]. The height difference arises due to an atomic layer of Bi which was deposited initially. The Bi floats always on top of the growing layer because it is less strongly bonded to the substrate than Si or Ge. The Si/Ge wires are homogeneous in width over larger distances and have a length of several thousand nanometers. Different widths of the wires can be easily achieved by different amounts of Ge and Si deposited.

5. Formation of More Complex Nanostructures by Self-Organization

One disadvantage of using self-organization for the formation of nanostructures is that only very simple kinds of nanostructures can be built by self-organization. The only examples presented so far are various kinds of islands and nanowires. In the following it will be shown that, using 2D island growth, somewhat more complex Si/Ge nanostructures, namely Si/Ge ring structures can be grown by self-assembly [Fig. 9(a)]

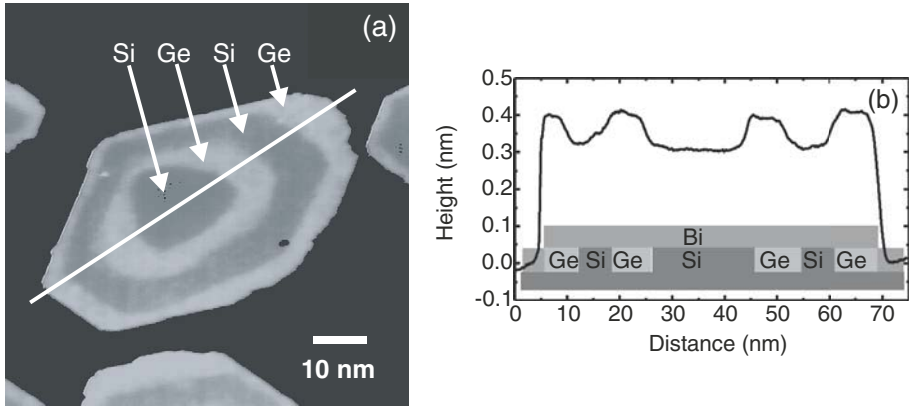


Figure 9. (a) Two-dimensional Ge/Si ring structure imaged with the STM. Ge rings are shown as light grey and Si rings are shown as dark grey. The width of the rings is 5–10 nm and the height is one atomic layer (0.3 nm). (b) Cross section along the line indicated in (a). Due to the Bi termination, the Ge rings are imaged ~ 0.09 nm higher than the Si rings. A schematic of the ring structure is shown in the inset.

[8]. In the 2D island growth mode the diffusion of deposited atoms at the surface is reduced (by decreasing the temperature) so that most of the diffusing atoms do not reach the step edges but nucleate as 2D islands or attach to existing islands. Once atomic-layer-high islands have nucleated, deposited atoms diffuse towards the island edge and are incorporated, forming a new ring of Ge or Si, respectively [Fig. 9(a)]. The measured height of the Ge rings is 0.09 nm higher than the measured height of the Si rings [Fig. 9(b)]. The width of the rings is 5–10 nm and the thickness is only one atomic layer (0.3 nm). The Si/Ge ring structure shown in Fig. 9(a) was obtained as follows. Initially Si islands form the cores (diameter 10–20 nm) of the Si/Ge ring structures. Subsequent alternating deposition of Ge and Si results in the formation of the Si/Ge ring structures around the Si core.

6. Conclusion

To fabricate semiconductor devices of dimensions beyond the limits of lithography alternative methods have to be explored. Self-organized growth of semiconductor nanostructures is one of these alternative approaches. Several examples for the formation of 2D self-organized nanostructures have been presented, including the controlled formation of different kinds of two-dimensional Si/Ge nanostructures, such as nanowires, nanowire superlattices and nanorings. The nanostructures grown have

a width down to 3.5 nm and a sub-nanometer thickness (0.3 nm), corresponding to a cross section consisting of only ~ 21 atoms.

References

- [1] D. M. Eigler and E. K. Schweizer. Positioning single atoms with a scanning tunneling microscope. *Nature* **344**: 524–526, 1990.
- [2] B. Voigtländer. Fundamental processes in Si/Si and Ge/Si epitaxy studied by scanning tunneling microscopy during growth. *Surf. Sci. Rep.* **43**: 127–254, 2001.
- [3] L. Andersohn, Th. Berke, U. Köhler, and B. Voigtländer. Nucleation behavior in molecular beam and chemical vapor deposition of silicon on Si(111)-(7 $\times$ 7). *J. Vac. Sci. Technol. A* **14**: 312–318, 1996.
- [4] J. A. Venables. Atomic processes in crystal growth. *Surf. Sci.* **299/300**: 798–817, 1994.
- [5] D. E. Jesson, B. Voigtländer, and M. Kästner. Direct observation of subcritical fluctuations during the formation of strained semiconductor islands. *Phys. Rev. Lett.* **84**: 330–333, 2000.
- [6] Feng Liu, A. H. Li, and M. G. Lagally. Self-assembly of two-dimensional islands via strain-mediated coarsening. *Phys. Rev. Lett.* **87**: art. no. 126103, 2001.
- [7] B. Voigtländer and A. Zinner. Structure of the Stranski–Krastanov layer in surfactant-mediated Sb/Ge/Si(111) epitaxy. *Surf. Sci.* **351**: L233–L238, 1996.
- [8] B. Voigtländer, M. Kawamura, N. Paul, and V. Cherepanov. Nanowires and nanorings at the atomic level *Phys. Rev. Lett.* **90**: art. no. 096102, 2003.

DIFFUSION, NUCLEATION AND GROWTH ON METAL SURFACES

Ofer Biham, Itay Furman* and Hanoch Mehl

Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel

John F. Wendelken

Condensed Matter Sciences Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee 37831-6030

Abstract Experimental, theoretical and computer simulation studies of growth of thin metal films in molecular-beam epitaxy are presented. Starting from a flat high symmetry surface such as Cu(001), a beam of Cu atoms is deposited on the surface. Within a broad range of experimental conditions, the adsorbed atoms diffuse and nucleate into islands. The process of island nucleation and growth in the sub-monolayer regime is studied using both rate equations (mean field) and kinetic Monte Carlo simulations. As additional layers grow, mounds form and the surface becomes three-dimensional in nature. The mounds effectively partition the surface into small sections with little diffusion of atoms between them. As the atomic flux continues, new islands form on the top layer of each mound. Recent studies have shown that the nucleation on the top terraces cannot be described by rate equations. We discuss the relation between the next layer nucleation and other processes in confined geometries, such as the formation of molecules on small dust grains.

1. Introduction

Thin film growth provides an excellent laboratory for statistical physics of non-equilibrium systems. The growth processes exhibit pattern formation, instabilities and dynamical phase transitions when driven away from equilibrium, as well as relaxation towards the equilibrium state when the external drive is turned off. Experimentally, molecular beam

*Present address: Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue N., Mail-stop D4-100, P.O. Box 19024, Seattle, WA 98109-1024, USA

epitaxy techniques provide excellent control on the growth conditions. Scanning probe microscopy and diffraction techniques provide detailed information about the resulting morphologies from the atomic scale to the sub-micron scale. The dynamics of surface growth is studied theoretically using a variety of techniques, ranging from a discrete description at the atomic scale to a continuum description at larger scales. This research effort is in line with the interest of the electronic and chemical industries in well characterized materials with novel electronic, optical and magnetic properties. Advances in the understanding and control of thin film growth processes is crucial for the long term effort to reduce the sizes of electronic circuits into the nano-scale.

One of the main aims of the research in this field is to predict and control the morphology of the films by tuning the material composition and growth conditions. The morphologies of thin films can be characterized by the lateral correlations, vertical roughness, and the size, shape and density of islands and nano-grains. The range of possible morphologies is rich and depends on the composition [1, 2, 3] and intensity of the flux [4, 5, 6, 7], and the composition [8, 9, 10, 11, 12], geometry [13, 14, 15] and temperature [14, 15, 16, 17, 18, 19] of the substrate.

In the case of thin film growth by molecular beam epitaxy, a beam with a well-characterized chemical composition and known flux, F , is directed towards a clean high-symmetry substrate of similar (homoepitaxy) or different (heteroepitaxy) composition. The substrate is held at a constant temperature, T , in an ultra-high vacuum chamber. The deposited atoms become adsorbed on the surface, on which they diffuse until they nucleate into new islands or stick to existing islands or steps. Several modes of growth were identified: the layer by layer growth (Frank-van der Merwe mode) in which the morphology is rather homogeneous and flat, the three dimensional growth (Volmer-Weber mode) and an intermediate growth mode (Stranski-Krastanov) that starts with a wetting layer on top of which three dimensional structures emerge.

In this paper we consider the homoepitaxial growth of thin metal films, on high symmetry surfaces such as fcc(001) that form a square substrate. In particular, we focus on the growth of Cu on Cu(001). For a broad range of flux intensities and surface temperatures, the growth takes place via the nucleation and growth of islands on the substrate. At a later stage, atoms that land on top of islands nucleate a second layer of islands and drive a growth instability that results in mound formation.

The paper is organized as follows. In Sec. 2 we consider the growth in the sub-monolayer regime. The scaling of island formation is studied using rate equations. The results of experiments and kinetic Monte Carlo simulations are shown. Multi-layer growth is presented in Sec. 3. The

nucleation of islands on the top terraces is considered and a connection to chemical reactions on small dust grains is shown. The failure of rate equations in this case is explained. The paper is summarized in Sec. 4.

2. The Submonolayer Regime

2.1 Rate Equation Analysis

Starting from a flat high-symmetry metal surface, atoms are deposited at a flux F monolayers (ML)/s. The deposited atoms adsorb on the surface, diffuse as random walkers, and nucleate new islands or aggregate into existing islands. The amount of adsorbed atoms after deposition time t is quantified by the coverage $\theta = Ft$ ML. The submonolayer regime ($\theta < 1$) can be divided into three stages. In the nucleation stage atoms diffuse on the flat surface and nucleate into two-dimensional islands. As deposition continues into the subsequent aggregation stage, nucleated islands absorb most of the newly adsorbed atoms and the nucleation of new islands is suppressed. The island density saturates and remains constant until they start to merge, thus reaching the coalescence stage. A second layer starts to grow when new islands nucleate on top of the first layer [7, 8, 19, 20, 21, 22]. The morphology of the growing film depends on the interlayer mass transport. The upward current is typically suppressed by the high energy barriers. The downward current is characterized by the Ehrlich-Schwoebel barrier E_{ES} , which impedes diffusion of adatoms down from island edges [23, 24, 25, 26]. E_{ES} is an effective kinetic parameter determined by the barriers to go down-the-step in a set of relevant local configurations [27, 28, 29, 30, 31, 32].

The nucleation of islands, in the submonolayer regime, takes place at very low coverage before significant correlations have sufficient time to emerge. Therefore, rate equations are expected to be suitable in this case. Studies of the island nucleation stage by rate equations gave rise to scaling relations for the island density as a function of the surface properties and growth conditions. The island density ρ_{I} can be expressed according to $\rho_{\text{I}} = \ell^{-2}$ where ℓ is the average distance between adjacent islands. In the simplest case, in which two atoms are sufficient to form a stable island, the rate equations take the form

$$\begin{aligned}\dot{\rho}_{\text{a}} &= F - 2A\rho_{\text{a}}^2 - A\rho_{\text{a}}\rho_{\text{I}} \\ \dot{\rho}_{\text{I}} &= A\rho_{\text{a}}^2,\end{aligned}\tag{1}$$

where ρ_{a} is the density of atoms (ML) and ρ_{I} is the density of islands. The parameter

$$A = \nu \exp\left(\frac{-E_0}{k_{\text{B}}T}\right)\tag{2}$$

is the hopping rate of atoms, where $\nu = 10^{12} \text{ s}^{-1}$ is the attempt rate, E_0 is the energy barrier for hopping and T is the surface temperature.

Scaling analysis based on the rate equations shows that the distance ℓ between adjacent islands exhibits scaling according to [33, 34, 35]

$$\ell^{-2} \sim \left(\frac{F}{\nu} \right)^\gamma \exp \left(\frac{E_{\text{eff}}}{K_B T} \right). \quad (3)$$

When this equation is used for the analysis of experimental data, the parameters γ , ν and E_{eff} are considered as fitting parameters. The energy E_{eff} is a measure of the barrier for the rate-limiting move in the island formation process and ν is the corresponding attempt rate.

The behavior described by Eq. (3) has been confirmed in a large class of systems. In particular, the exponential dependence of ℓ^{-2} on $1/T$ was reported in Refs. [4, 5, 7, 15, 16, 17, 18, 36, 37, 38] and the power-law dependence on F in Refs. [7, 5, 39, 6, 40, 41, 4]. It turns out that the exponent γ depends on microscopic details such as the nature of the diffusion process (isotropic *versus* anisotropic diffusion) [5, 15, 39, 42, 43, 44], the island structure [42, 43, 45, 46], the stability properties of small islands [33, 35, 42, 43, 45, 46, 47, 48, 36, 49] as well as their mobility [42, 43, 50, 46, 45]. The fact that small islands can be unstable or mobile turns out to be crucial, because they may split or move away unless they reach the minimal size of a stable and immobile island. These properties are quantified by the concept of critical nucleus. An island that includes i^* atoms is called critical nucleus if by adding one more atom it would become a stable (and immobile) island of size $i^* + 1$ that will continue to grow by aggregation of more atoms. In the case that islands of size i^* or less are unstable it was shown that the exponent

$$\gamma = \frac{i^*}{(i^* + 2)}, \quad (4)$$

and the effective energy barrier is

$$E_{\text{eff}} = \gamma \left(E_0 + \frac{E_{i^*}}{i^*} \right), \quad (5)$$

where E_0 is the adatom diffusion barrier and E_{i^*} is the binding energy of the critical nucleus [33, 34, 35] (note that $E_{i^*} = 0$ for $i^* = 1$).

Detailed studies of the energy landscape for diffusion of Cu atoms on Cu(001) have shown that islands of all sizes are stable up to room temperature and possibly at higher temperatures. However, islands of two atoms (dimers) are mobile and their mobility is comparable to that of single atoms [51, 52]. Therefore, the scaling of island growth of Cu on

Cu(001) is expected to be determined by the mobility of dimers, namely with $i^* = 2$. In this case, scaling analysis based on rate equations has shown that [43, 46, 45, 53].

$$\gamma = \frac{i^*}{2i^* + 1}. \quad (6)$$

A general expression for E_{eff} is not available in the case of island mobility. However, for the particular case of Cu on Cu(001), where only adatoms and dimers are mobile it was found that [42]

$$E_{\text{eff}} = \frac{(E_0 + E_2)}{10}, \quad (7)$$

where E_0 and E_2 are the barriers for diffusion of adatom and dimers, respectively. Somewhat more complicated cases appear when there is a crossover between growth regimes of different i^* [49, 54, 55] or due to the existence of “magic islands” [40, 45, 46, 56, 57].

2.2 Experiments and Monte Carlo Simulations

The growth of Cu on Cu(001) and the subsequent relaxation and coarsening were studied extensively in a series of experiments [37, 40, 41, 58, 59, 60]. Fig. 1 shows a set of STM images of the islands in the submonolayer regime at coverages of 0.08, 0.25, 0.4 and 1.0 ML, all grown at a substrate temperature of 296 K. The first image shows the early nucleation of islands: the next two images show the growth of the islands as more atoms are deposited. The third image, at 0.4 ML, shows the initial stage of island coalescence and second layer growth. At 1.0 ML multilayer growth is pronounced with the original substrate, first layer, and second layer visible. Note the irregular shape of the vacancy islands, which were observed to reshape into squares only a short time after deposition [60]. The window size in all the images is 100×100 nm.

The mean separation, ℓ , between Cu islands on Cu(001) was measured using spot-profile analysis of low-energy electron diffraction (SPA-LEED), following deposition of 0.3 ML of Cu [40, 37, 41]. This coverage was chosen to maximize the electron diffraction signal in the aggregation stage. The measurements were taken as a function of the flux F for three temperatures, $T = 213, 223$ and 263 K (Fig. 2), and as a function of the temperature T at a constant flux $F = 3.21 \times 10^{-4}$ ML/s.

Kinetic Monte Carlo (KMC) simulations provide a useful computational procedure for the study of surface diffusion and growth [61, 62, 63]. The energy barriers used in our KMC simulations were obtained via the embedded-atom method (EAM). Our simulations were performed under conditions identical to those employed experimentally in Refs. [40,

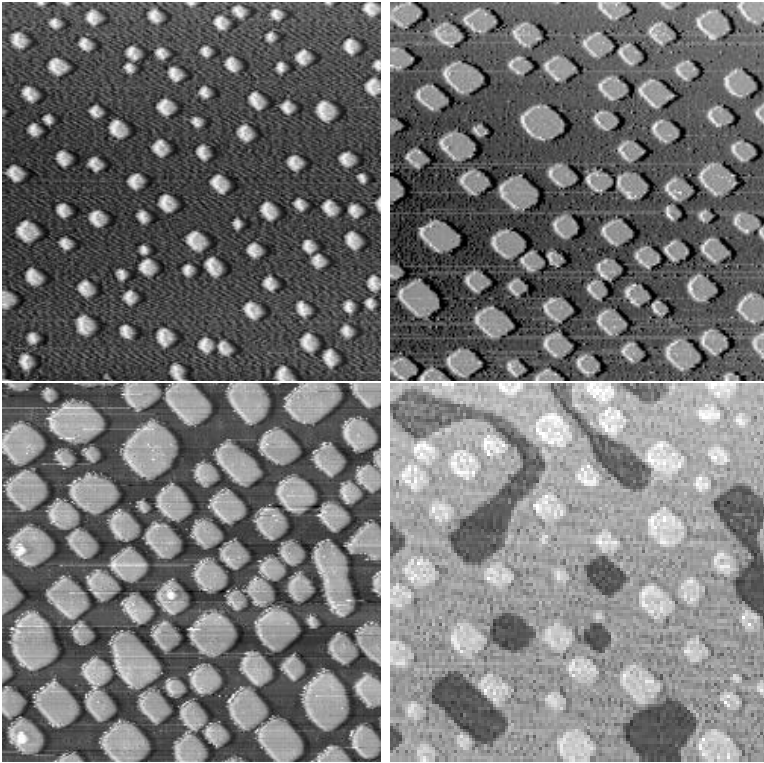


Figure 1. Submonolayer growth of Cu on Cu(001) at coverages of 0.08 ML (top left), 0.25 ML (top right), 0.4 ML (bottom left) and 1.0 ML (bottom right). The fluxes are 0.16 ML/s for the top images and 0.2 ML/s for the bottom images.

37, 41]. In our KMC simulations atoms are randomly deposited on a square-lattice substrate of 250×250 sites that corresponds to a terrace width of ~ 64 nm. These atoms attach irreversibly to the surface and hop as random walkers to unoccupied nearest-neighbor (NN) sites. Each hop involves an activation energy-barrier, E_n , that depends on the configuration of occupied and unoccupied adjacent sites in the 3×3 square around the hopping atom [64, 53, 51, 52]. The hopping-rate (in units of hops per second), A_n , for some configuration n is

$$A_n = \nu \cdot \exp(-E_n/K_{\text{B}}T), \quad (8)$$

where ν is the attempt rate common to all moves and E_n is calculated using the EAM [65, 52, 51].

We have reconstructed the experimental conditions in our KMC simulations and obtained for each experimental curve in Figs. 2 and 3 a

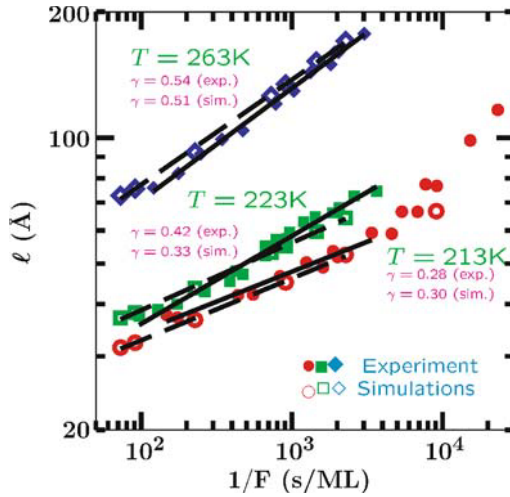


Figure 2. Comparison of experimental results (full symbols) and simulation results (empty symbols) for the island separation ℓ versus the inverse flux $1/F$, at three temperatures $T = 213$, 223 and 263 K. The coverage is $\theta = 0.3$ ML. The solid lines represent fits to Eq. (3).

corresponding simulated curve. Each simulated data point is an average over 20 runs. Fig. 2 shows the experimental results and the corresponding simulations for the island separation ℓ versus $1/F$ at $T = 213$, 223 and 263 K. All data points were taken at $\theta = 0.3$ ML. Clearly, there is good agreement between the experiment and simulation. The curves follow the power-law behavior of Eq. (3) over one and a half decades, from which the exponent γ is extracted. Based on scaling theory, one may conclude that $\gamma \simeq \frac{1}{3}$, that *dimers* and larger islands are stable and *immobile*, and that adatoms are the *only* mobile entities on the surface. However, care should be exercised when drawing such conclusions. Specifically, in our case the value $\gamma \simeq \frac{1}{3}$ is found despite significant dimer mobility. According to the scaling theory $\gamma = \frac{2}{5}$ for a system with mobile monomers and dimers. The experimental results for γ and the corresponding simulation results presented in Fig. 2 are significantly different from this value. The discrepancy is removed when we recalculate γ for $T = 263$ K at a much lower coverage of $\theta = 0.125$ ML. At this coverage, and still not too low a temperature, scaling theory is expected to apply. Indeed we find $\gamma \simeq \frac{2}{5}$, in good agreement with the scaling prediction. For the lower temperatures $T = 213$ and 223 K, the results disagree with the scaling theory, even for the lower coverage of $\theta = 0.125$ ML.

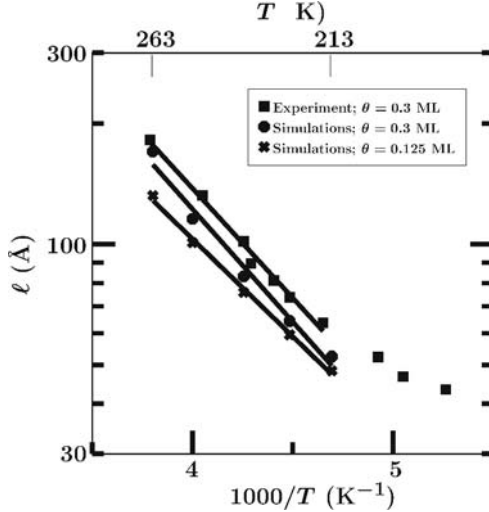


Figure 3. The island separation ℓ versus $1/T$, obtained from the experiment for $F = 3.21 \times 10^{-4}$ ML/s and simulation for $F = 4.8 \times 10^{-4}$ ML/s at coverage $\theta = 0.3$ ML. We also present the simulation results of the same runs but at an earlier stage when $\theta = 0.125$ ML. The solid lines represent fits to $\ell \sim \exp(-E_{\text{eff}}/2K_{\text{B}}T)$, with $E_{\text{eff}} = 0.108 \pm 0.005$, 0.112 ± 0.008 and 0.096 ± 0.003 eV for the experiment, and the simulations at $\theta = 0.3$ and 0.125 ML, respectively.

Unlike the energy barriers, we do not have a value for the attempt rate ν , from atomic scale calculations. Instead, the *experimental* value of ν is obtained by fitting the simulation and experimental results. This is possible since in the simulation ν sets a fundamental clock rate, while the simulation results depend only on the ratio ν/F rather than on F and ν separately. Using this property we perform the simulations for a broad range of values of ν/F and plot the island separation ℓ versus ν/F . The attempt rate ν , is then obtained as the value for which the three simulated curves in Fig. 2, overlap simultaneously the three experimental curves. It is found to be $\nu = 1.2 \times 10^{13} \text{ s}^{-1}$.

3. Multi-Layer Growth

When the islands in the first monolayer become large enough, atoms that fall on top of them start to nucleate into a second layer of islands. As the coverage increases, several atomic layers become exposed. The islands on islands develop into pyramidal mounds with a steady state angle of 5.6° [58]. Eventually, the morphology can be described as pyramid

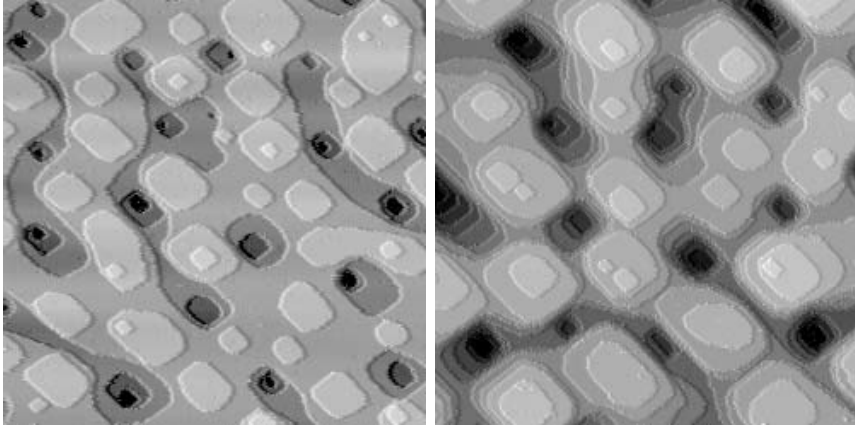


Figure 4. Typical morphologies in multi-layer growth of Cu on Cu(001) at 8 ML (left) and 40 ML (right). The mound structure and the square or rectangular shapes of the top terraces are apparent. Some new islands that nucleated on the top terraces are visible.

mounds and pyramid holes separated by meandering mid-level terraces. In Fig. 4 we show typical morphologies obtained in the growth of Cu on Cu(001) at 8 ML (left) and 40 ML (right), where $F = 0.02$ ML/s. Most of the nucleation events take place on top terraces, namely those terraces that are higher than the adjacent terraces in all directions. In the case of Cu growth on Cu(001) these terraces tend to have a square shape that resembles the submonolayer islands, as can be seen in Fig. 4. An atom that lands on a top terrace either meets another atom and nucleate into an island or hops down the step and becomes incorporated into the terrace. The residence time of a single atom on the top terrace is determined by its hopping rate, the terrace size and by the Ehrlich-Schwoebel barrier which is the additional energy barrier for moves down the step. If this barrier is zero one can assume that an atom that reaches the boundary of the terrace will hop down. As the Ehrlich-Schwoebel barrier increases the probability of an atom that reaches the edge to hop down decreases and its residence time on the terrace increases.

We will first consider the case of a vanishing Ehrlich-Schwoebel barrier. In this case atoms that reach the edge of the top terrace hop down the step. Consider an atom that is deposited into a random position on a square terrace of $S = L^2$ sites. The average number of hopping moves required for the atom to reach the edge and hop down is given by

$$s_{\text{visit}} = \alpha L^2 + \beta L, \quad (9)$$

where $\alpha = 0.14$ and $\beta = 0.56$ [66, 67, 68]. Therefore, to leading order and up to a constant factor of order one $s_{\text{visit}} \simeq S$.

The process in which atoms hop down the step and are removed from the top terrace can be considered as analogous to a desorption process in systems in which the adsorbed atoms are only weakly bounded to the surface. One such process is the adsorption and reaction of hydrogen atoms on a surface that produces hydrogen molecules. This process can be described by rate equations of the form [69, 70]

$$\dot{\rho} = F - W\rho - 2A\rho^2, \quad (10)$$

where ρ (ML) is the density of atoms on the surface. The first term on the right hand side represents the incoming flux F ML/s). The second term describes the desorption, where W (s^{-1}) is the desorption coefficient, and the third term describes the reaction that generates molecules, where A (s^{-1}) is the hopping rate of atoms on the surface. When the reaction term is suppressed, the average number of hopping moves an atom can make before it desorbs is given by

$$s_{\text{visit}} = A/W. \quad (11)$$

Here, the desorption term is determined by the binding energy of the hydrogen atoms to the surface and by the surface temperature. The rate of formation of molecules is determined by the likelihood of a hopping atom to meet another hydrogen atom on the surface before it desorbs. In order to examine the formation rate it is useful to determine the average number of vacant sites around a hydrogen atom. When the reaction term is suppressed and the system is kept under steady state conditions, the average number of vacant sites per atom is [71, 72]

$$s_{\text{vacant}} = W/F. \quad (12)$$

Clearly, a small terrace on which the number of adsorption sites is smaller than s_{vacant} has, on average, less than one atom at any given time. If the terrace size is also smaller than s_{visit} then an atom on the terrace is likely to scan the entire terrace several times, visiting the same sites again and again before it desorbs. These repetitions lower the likelihood of that atom finding other atoms and the reaction rate is reduced compared to that on a larger terrace. The analysis of the reaction process has shown that the rate equations become unsuitable once the terrace size becomes the smallest length-scale in the problem. This happens when

$$S < \min\{s_{\text{vacant}}, s_{\text{visit}}\}. \quad (13)$$

For the nucleation process on top terraces, the rate of incorporation of atoms into the terrace edge is given by $W = A/s_{\text{visit}} \simeq A/L^2$. The number of vacant sites per atom can now be obtained from Eq. (12):

$$s_{\text{vacant}} \simeq \frac{A}{F L^2}. \quad (14)$$

Since $S \simeq s_{\text{visit}}$ the limit of small terraces is obtained when $S < s_{\text{vacant}}$, namely $L < (A/F)^{1/4}$. In this limit there is on average less than one atom on a top terrace and the nucleation process is suppressed. The terrace continues to grow due to the incoming flux until it reaches the size of $L \simeq (A/F)^{1/4}$, above which the nucleation process is no longer suppressed. This is the typical terrace size at which an island of the next layer nucleates. In case of a non-zero Ehrlich-Schwoebel barrier, atoms stay longer on the terrace because they may hit the terrace edge several times before hopping down. Therefore, in this case, nucleation is expected to take place on smaller terraces. Note that in order to be suitable for the growth of Cu on Cu(001), the analysis above should be generalized, namely the dimer mobility should be taken into account.

A more complete analysis of island nucleation on top terraces was performed by Politi and Castellano [73, 74, 75]. The nucleation rate was calculated exactly and the spatial distribution of nucleation events was obtained. The connection to the recombination of hydrogen on interstellar dust grains was pointed out by Krug [76]. A promising approach to the simulation of multi-layer growth is the level-set method, which combines a continuous description within each atomic layer with a discrete resolution of the layers [77]. Its computational efficiency enables the analysis of larger systems for longer times compared to KMC methods.

4. Summary

The growth of thin metal films in molecular beam epitaxy was reviewed. Experimental results were presented for the growth of Cu on Cu(001) and compared to rate equation analysis (mean field) and to Monte Carlo simulations. In the submonolayer regime, mean field applies in the early stages of island nucleation. In multiple layer growth, the morphology is dominated by mounds and nucleation is confined to the top terraces. Deviations from mean-field results are expected, while a simple analysis of length scales provides useful predictions on island nucleation on top terraces.

Acknowledgments

The theoretical work, done at the Hebrew University was partially supported by the Adler Foundation for Space Research of the Israel Science Foundation. The experimental work was performed at Oak Ridge National Laboratory, managed by UT Battelet, LLC, for the U.S. Department of Energy under Contract No. DE-AC05-00OR22725.

References

- [1] M. C. Bartelt, A. K. Schmid, J. W. Evans and R. Q. Hwang. Island size and environment dependence of adatom capture: Cu/Co islands on Ru(0001). *Phys. Rev. Lett.* **81**: 1901–1904, 1998.
- [2] H. van der Vegt, W. Huisman, P. B. Howes, and E. Vlieg. The effect of Sb on the nucleation and growth of Ag on Ag(100). *Surf. Sci.* **330**: 101–112, 1995.
- [3] H. Brune, H. Röder, C. Boragno and K. Kern. Microscopic view of nucleation on surfaces. *Phys. Rev. Lett.* **73**: 1955–1958, 1994.
- [4] C.-M. Zhang and M. C. Bartelt and J.-M. Wen and C. J. Jenks and J. W. Evans and P. A. Thiel. Submonolayer island formation and the onset of multilayer growth during Ag/Ag(100) homoepitaxy. *Surf. Sci.* **406**: 178–193, 1998.
- [5] S. Günther, E. Kopatzki, M. C. Bartelt, J. W. Evans and R. J. Behm. Anisotropy in nucleation and growth of two-dimensional islands during homoepitaxy on “hex” reconstructed Au(100). *Phys. Rev. Lett.* **73**: 553–556, 1994.
- [6] J. A. Stroscio and D. T. Pierce. Scaling of diffusion-mediated island growth in iron-on-iron homoepitaxy. *Phys. Rev. B* **49**, 8522–8525, 1994.
- [7] H. J. Ernst, F. Fabre and J. Lapujoulade. Nucleation and diffusion of Cu adatoms on Cu(100): A helium-atom-beam scattering study. *Phys. Rev. B* **46**: 1929–1932, 1992.
- [8] E. Kopatzki, S. Günther, W. Nichtl-Pecher and R. J. Behm. Homoepitaxial growth on Ni(100) and its modification by a preadsorbed oxygen adlayer. *Surf. Sci.* **284**: 154–166, 1993.
- [9] J. E. Van Nostrand, S. J. Chey, M.-A. Hasan, D. G. Cahill and J. E. Greene. Surface morphology during multilayer epitaxial growth of Ge(001). *Phys. Rev. Lett.* **74**: 1127–1130, 1995.
- [10] J. E. Van Nostrand, S. J. Chey and D. G. Cahill. Low-temperature growth morphology of singular and vicinal Ge(001). *Phys. Rev. B* **57**: 12536–12543, 1998.
- [11] J. Tersoff, D. van der Gon and R. M. Tromp. Critical island size for layer-by-layer growth. *Phys. Rev. Lett.* **72**: 266–269, 1994.
- [12] Z. Y. Zhang and M. G. Lagally. Atomistic processes in the early stages of thin-film growth. *Science* **276**: 377–383, 1997.
- [13] W. Elliott, P. Miceli, T. Tse, and P. Stephens. Orientation dependence of homoepitaxy: An in situ X-ray scattering study of Ag. *Physica B* **221**: 65–69, 1996.
- [14] W. Elliott, P. Miceli, T. Tse, and P. Stephens. Temperature and orientation dependence of kinetic roughening during homoepitaxy: A quantitative x-ray-scattering study of Ag. *Phys. Rev. B* **54**: 17938–17942, 1996.
- [15] Y. W. Mo, J. Kleiner, M. B. Webb and M. G. Lagally. Activation energy for surface diffusion of Si on Si(001): A scanning-tunneling-microscopy study. *Phys. Rev. Lett.* **66**: 1998–2001, 1991.
- [16] J. A. Stroscio, D. T. Pierce and R. A. Dragoset. Homoepitaxial growth of iron and a real space view of reflection-high-energy-electron diffraction. *Phys. Rev. Lett.* **70**: 3615–3618, 1993.

- [17] F. Tsui, J. Wellman, C. Uher and R. Clarke. Morphology transition and layer-by-layer growth of Rh(111) *Phys. Rev. Lett.* **76**: 3164–3167, 1996.
- [18] L. Bardotti and C. R. Stoldt and C. J. Jenks and M. C. Bartelt and J. W. Evans and P. A. Thiel. High-resolution LEED profile analysis and diffusion barrier estimation for submonolayer homoepitaxy of Ag/Ag(100). *Phys. Rev. B* **57**: 12544–12549, 1998.
- [19] G. Costantini, F. Buatier de Mongeot, C. Boragno, and U. Valbusa. Temperature dependent reentrant smooth growth in Ag(001) homoepitaxy. *Surf. Sci.* **459**: L487–L492, 2000.
- [20] H. J. Ernst, F. Fabre and J. Lapujoulade. Growth of Cu on Cu(100). *Surf. Sci.* **275**: L682–L684, 1992.
- [21] H.-J. Ernst, F. Fabre, and J. Lapujoulade Observation of dynamical scaling in “spinodal decomposition” in two dimensions. *Phys. Rev. Lett.* **69**: 458–461, 1992.
- [22] J.-M. Wen, J. W. Evans, M. C. Bartelt, J. W. Burnett and P. A. Thiel. Coarsening mechanisms in a metal film: From cluster diffusion to vacancy ripening. *Phys. Rev. Lett.* **76**: 652–655, 1996.
- [23] G. Ehrlich and H. Hudda. Atomic view of surface diffusion: tungsten on tungsten. *J. Chem. Phys.* **44**: 1039–1099, 1966.
- [24] R. Schwoebel and E. Shipsey. Step motion on crystal surfaces. *J. Appl. Phys.* **37**: 3682–3696, 1966.
- [25] R. Schwoebel. Step Motion on Crystal Surfaces. II. *J. Appl. Phys.* **40**: 614–618, 1969.
- [26] S. C. Wang and G. Ehrlich. Adatom motion to lattice steps: A direct view *Phys. Rev. Lett.* **70**: 41–44, 1993.
- [27] Z.-J. Tian and T. S. Rahman. Energetics of stepped Cu surfaces. *Phys. Rev. B* **47**: 9751–9759, 1993.
- [28] J. Merikoski and T. Ala-Nissila. Diffusion processes and growth on stepped metal surfaces. *Phys. Rev. B* **52**: R8715–R8720, 1995.
- [29] J. Merikoski, I. Vattulainen, J. Heinonen and T. Ala-Nissila. Effect of kinks and concerted diffusion mechanisms on mass transport and growth on stepped metal surfaces. *Surf. Sci.* **387**: 167–182, 1997.
- [30] O. S. Trushin and K. Kokko and P. T. Salo and W. Hergert and M. Kotrla. Step roughening effect on adatom diffusion. *Phys. Rev. B* **56**: 12135–12138, 1997.
- [31] J. B. Adams, Z. Wang, and Y. Li. Modeling Cu thin film growth. *Thin Solid Films* **365**: 201–210, 2000.
- [32] Z. Y. Wang, Y. Li, and J. B. Adams. Kinetic lattice Monte Carlo simulation of facet growth rate. *Surf. Sci.* **450**: 51–63, 2000.
- [33] J. A. Venables. Rate equation approach to thin film nucleation kinetics. *Phil. Mag.* **27**: 697–738, 1973.
- [34] J. A. Venables, G. D. T. Spiller, and M. Hanbucken. Nucleation and growth of thin films. *Rep. Prog. Phys.* **47**: 399–459, 1984.
- [35] S. Stoyanov and D. Kashchiev. Thin film nucleation and growth theories: A confrontation with experiment. in *Current Topics in Materials Science*, Vol. 7, edited by E. Kaldis, (North-Holland, Amsterdam, 1981), pp. 69–141.

- [36] C. Ratsch, A. Zangwill, P. Smilauer and D. D. Vvedensky. Saturation and scaling of epitaxial island densities. *Phys. Rev. Lett.* **72**: 3194–3197, 1994.
- [37] H. Dürr, J. F. Wendelken and J.-K. Zuo. Island morphology and adatom energy barriers during homoepitaxy on Cu(001). *Surf. Sci.* **328**: L527–L532, 1995.
- [38] M. Bott, M. Hohage, M. Morgenstern, T. Michely and G. Comsa. New approach for determination of diffusion parameters of adatoms. *Phys. Rev. Lett.* **76**: 1304–1307, 1996.
- [39] T. R. Linderoth and J. J. Mortensen and K. W. Jacobsen and E. Lægsgaard, I. Stensgaard and F. Besenbacher. Homoepitaxial growth of Pt on Pt(100)-hex: Effects of strongly anisotropic diffusion and finite island sizes. *Phys. Rev. Lett.* **77**: 87–90, 1996.
- [40] J.-K. Zuo, J. F. Wendelken, H. Dürr, and C.-L. Liu. Growth and coalescence in submonolayer homoepitaxy on Cu(100) studied with high-resolution low-energy electron diffraction. *Phys. Rev. Lett.* **72**: 3064–3067, 1994.
- [41] A. K. Swan, Z.-P. Shi, J. F. Wendelken and Z. Zhang. Flux-dependent scaling behavior in Cu(100) submonolayer homoepitaxy. *Surf. Sci.* **391**: L1205–L1211, 1997.
- [42] J. Villain, A. Pimpinelli, L. Tang, and D. Wolf. Terrace sizes in molecular-beam epitaxy. *J. Phys. I (France)* **2**, 2107–2121, 1992.
- [43] J. Villain, A. Pimpinelli and D. Wolf. Layer-by-layer growth in molecular-beam epitaxy. *Comments Cond. Matter Phys.* **16**: 1–18, 1992.
- [44] M. C. Bartelt and J. W. Evans. Scaling analysis of diffusion-mediated island growth in surface adsorption processes. *Phys. Rev. B* **46**: 12675–12687, 1992.
- [45] M. Schroeder and D. E. Wolf. Magic islands and submonolayer scaling in molecular beam epitaxy. *Phys. Rev. Lett.* **74**: 2062–2065, 1995.
- [46] D. E. Wolf. Non-equilibrium ordering dynamics and pattern formation in *Scale Invariance, Interfaces and Non-Equilibrium Dynamics*, edited by M. Droz, A. McKane, J. Vannimenus, and D. E. Wolf (Plenum, New York, 1994), pp. 165–202.
- [47] J. G. Amar and F. Family. Critical cluster size: Island morphology and size distribution in submonolayer epitaxial growth. *Phys. Rev. Lett.* **74**: 2066–2069, 1995.
- [48] M. C. Bartelt, L. S. Perkins and J. W. Evans. Transitions in critical size for metal(100) homoepitaxy. *Surf. Sci.* **344**: L1193–L1199, 1995.
- [49] C. Ratsch, P. Smilauer, A. Zangwill and D. D. Vvedensky. Submonolayer epitaxy without a critical nucleus. *Surf. Sci.* **329**: L599–L604, 1995.
- [50] M. C. Bartelt, S. Günther, E. Kopatzki, R. J. Behm and J. W. Evans. Island-size distributions in submonolayer epitaxial growth: Influence of the mobility of small clusters. *Phys. Rev. B* **53**: 4099–4104, 1996.
- [51] O. Biham, I. Furman, M. Karimi, G. Vidali, R. Kennett and H. Zeng. Models for diffusion and island growth in metal monolayers. *Surf. Sci.* **400**: 29–43, 1998.
- [52] H. Mehl, O. Biham, I. Furman and M. Karimi. Models for adatom diffusion on fcc (001) metal surfaces. *Phys. Rev. B* **60**: 2106–2116, 1999.
- [53] I. Furman and O. Biham. Effects of mobility of small islands on growth in molecular-beam epitaxy. *Phys. Rev. B* **55**: 7917–7926, 1997.

- [54] M. C. Bartelt and J. W. Evans. Transition to multilayer kinetic roughening for metal (100) homoepitaxy. *Phys. Rev. Lett.* **75**: 4250–4253, 1995.
- [55] M. N. Popescu, J. G. Amar and F. Family. Self-consistent rate-equation approach to transitions in critical island size in metal (100) and metal (111) homoepitaxy. *Phys. Rev. B* **58**: 1613–1619, 1998.
- [56] G. L. Kellogg. Oscillatory behavior in the size dependence of cluster mobility on metal surfaces: Rh on Rh(100). *Phys. Rev. Lett.* **73**: 1833–1836, 1994.
- [57] Z.-P. Shi, Z. Zhang, A. K. Swan and J. F. Wendelken. Dimer shearing as a novel mechanism for cluster diffusion and dissociation on metal (100) surfaces. *Phys. Rev. Lett.* **76**: 4927–4930, 1996.
- [58] J.-K. Zuo and J.F. Wendelken. Evolution of mound morphology in reversible homoepitaxy on Cu(100). *Phys. Rev. Lett.* **78**: 2791–2794, 1997.
- [59] M. Li, J.F. Wendelken, B.-G. Liu, E.G. Wang and Z. Zhang. Decay characteristics of surface mounds with contrasting interlayer mass transport channels. *Phys. Rev. Lett.* **86**: 2345–2348, 2001.
- [60] W. W. Pai, J. F. Wendelken, C. R. Stoldt, P. A. Thiel, J. W. Evans and D.-J. Liu. Evolution of two-dimensional wormlike nanoclusters on metal surfaces. *Phys. Rev. Lett.* **86**: 3088–3091, 2001.
- [61] S. Clarke, M. R. Wilby and D. D. Vvedensky. Theory of homoepitaxy on Si(001) 1: Kinetics during growth. *Surf. Sci.* **255**: 91–110, 1991.
- [62] Y.-T. Lu and H. Metiu. Growth kinetics simulation of the Al-Ga self-organization on GaAs(100) stepped surfaces. *Surf. Sci.* **245**: 150–172, 1991.
- [63] K. A. Fichthorn and W. H. Weinberg. Theoretical foundations of dynamic Monte Carlo simulations. *J. Chem. Phys.* **95**: 1090–1096, 1991.
- [64] A. F. Voter. Classically exact overlayer dynamics: Diffusion of rhodium clusters on Rh(100). *Phys. Rev. B* **34**: 6819–6829, 1986.
- [65] M. S. Daw and M. I. Baskes. Semiempirical, quantum mechanical calculation of hydrogen embrittlement in metals. *Phys. Rev. Lett.* **50**: 1285–1288, 1983.
- [66] C. A. Walsh, J. J. Kozak. Exact algorithm for d -dimensional walks on finite and infinite lattices with traps. *Phys. Rev. Lett.* **47**: 1500–1502, 1981.
- [67] C. A. Walsh, J. J. Kozak. Exact algorithm for d -dimensional walks on finite and infinite lattices with traps. II. General formulation and application to diffusion-controlled reactions. *Phys. Rev. B* **26**: 4166–4189, 1982.
- [68] Y. Shilo and O. Biham. Sandpile models and random walkers on finite lattices. *Phys. Rev. E* **67**: art. no. 066102, 2003.
- [69] O. Biham, I. Furman, N. Katz, V. Pirronello and G. Vidali. H₂ formation on interstellar grains in different physical regimes. *Mon. Not. R. Astron. Soc.* **296**: 869–872, 1998.
- [70] N. Katz, I. Furman, O. Biham, V. Pirronello and G. Vidali. Molecular hydrogen formation on astrophysically relevant surfaces. *Astrophys. J.* **522**: 305–312, 1999.
- [71] O. Biham, I. Furman, V. Pirronello and G. Vidali. Master equation for hydrogen recombination on grain surfaces. *Astrophys. J.* **553**: 595–603, 2001.
- [72] O. Biham and A. Lipshtat. Exact results for hydrogen recombination on dust grain surfaces. *Phys. Rev. E* **66**: art. no. 056103, 2002.

- [73] C. Castellano and P. Politi. Spatiotemporal distribution of nucleation events during crystal growth. *Phys. Rev. Lett.* **87**: art. no. 056102, 2001.
- [74] P. Politi and C. Castellano. Process of irreversible nucleation in multilayer growth. I. Failure of the mean-field approach. *Phys. Rev. E* **66**: art. no. 031605, 2002.
- [75] P. Politi and C. Castellano. Process of irreversible nucleation in multilayer growth. II. Exact results in one and two dimensions. *Phys. Rev. E* **66**: art. no. 031606, 2002.
- [76] J. Krug. Reaction kinetics in confined geometries: Lonely adatoms in space. *Phys. Rev. E* **67**: art. no. 065102, 2003.
- [77] C. Ratsch, M. F. Gyure, R. E. Caflisch, F. Gibou, M. Petersen, M. Kang, J. Garcia and D. D. Vvedensky. Level-set method for island dynamics in epitaxial growth. *Phys. Rev. B* **65**: art. no. 195403, 2002.

THE MECHANISM OF THE STRANSKI–KRASTANOV TRANSITION

A. G. Cullis and D. J. Norris

Department of Electronic and Electrical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, United Kingdom

T. Walther

Institut für Anorganische Chemie, Universität Bonn, Römerstrasse 164, D-53117 Bonn, Germany

M. A. Migliorato and M. Hopkinson

Department of Electronic and Electrical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, United Kingdom

Abstract For strained-layer epitaxy, a detailed examination is carried out of the way in which strain changes due to elemental segregation within the initially-formed flat “wetting” layer can control the Stranski–Krastanov epitaxial islanding transition. Based upon these considerations, it is shown that a new segregation-based mechanism is fully compatible with transition in both the $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ and $\text{Si}_{1-x}\text{Ge}_x/\text{Si}$ systems grown over wide ranges of conditions. Quantitative segregation calculations allow critical “wetting” layer thicknesses to be derived and it is demonstrated that for the $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ system ($x = 0.25$ – 1) such calculations show good agreement with experimental measurements. The strain energy associated with the segregated surface layer is determined for the complete range of deposited In concentrations using atomistic simulations. The segregation-mediated driving force is considered to be important, also, for all other epitaxial systems which comprise chemically-similar but substantially misfitting materials and which exhibit the Stranski–Krastanov transition.

1. Introduction

The mechanisms by which epitaxial films can grow upon crystalline substrates fall into one of three categories: (i) two-dimensional (2D) layer-by-layer growth proposed by Frank and Van der Merwe [1], (ii) three-dimensional (3D) island growth proposed by Volmer and Weber [2], and (iii) 2D layer growth followed by 3D islanding, first described by Stranski and Krastanov (SK) [3]. For systems with either zero or small lattice mismatch, 2D layer-by-layer growth typically occurs, while for systems with highly mismatched and dissimilar materials, 3D island growth usually takes place. For epitaxial systems with similar materials and high lattice mismatch, the two-stage SK growth mode is common. In this latter case, a very thin, flat epitaxial layer is formed first and then a transition to 3D island growth takes place at a certain critical thickness. This growth mode has received in-depth experimental study across wide materials areas from metals to semiconducting materials (see Venables *et al.* [7] for early work on mainly metal-related deposition systems). In the semiconductors area, the transition has assumed some importance since the islands formed can be employed as quantum dots in advanced electronic devices. Accordingly, a range of semiconductor epitaxial systems exhibiting the SK growth mode have been carefully studied, and include $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ [6, 18, 19, 29, 6, 10, 20], $\text{InP}/\text{In}_x\text{Ga}_{1-x}\text{P}$ [13, 23], GaSb/GaAs [24] and SiGe/Si [11, 16, 18, 25, 26, 27]. Initially-formed SK islands are coherent [11], although they become incoherent (dislocated) [18] when island sizes increase during growth.

2. Mechanism of the Stranski–Krastanov Transition

Much work over many years has been devoted to the formulation of theoretical models based upon energy calculations and rate equations [28, 8, 12, 13, 17, 9, 27, 28] in order to attempt to explain the features of the 2D-3D SK transition. It is often concluded that 2D islands tend to transform into 3D islands when they exceed a certain critical size and such arguments have been employed to model the transition. However, despite these extensive studies, the fundamental driving force for the transition has remained unspecified [29] until recent work of Walther *et al.* [30] and Cullis *et al.* [31] has thrown new light upon the mechanism. In particular, there has been little consideration of the growth of the initial wetting layer and the factors that control the critical thickness which it must attain before the islanding transition can take place. Careful measurements [30] of SK island and “wetting” layer composition were carried out for the $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ system using electron energy

loss imaging. Based upon these measurements, it was proposed [30, 31] that segregation of elemental In to the surface of the initial flat “wetting” layer controls the point at which the transition to island growth occurs (referred to here as the WCNH mechanism). This provides a natural explanation for the general features of the transition and it is described in some detail here. The mechanism is fundamentally applicable to all materials systems exhibiting the SK transition where the growth technique does not significantly suppress segregation (see below) and is considered here with special reference to the $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ system, with some consideration also of the $\text{Si}_{1-x}\text{Ge}_x/\text{Si}$ system.

In order to produce islands by, e.g. molecular-beam epitaxy (MBE) growth of the $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ system, either an alloy or, for $x=1$, binary material (InAs) is deposited. In the WCNH mechanism proposed for the SK transition, as the initial flat epitaxial layer forms, the strain introduced by the vertical segregation of the largest atomic species (In) in the deposited material is considered to be of fundamental significance. The segregation has been simulated in the present work using the Fukatsu/Dehaese model [32, 33], which considers exchange of the Group III species between the top two layers during growth and demonstrates that the surface layer exhibits a very substantial deviation from the deposition flux concentration. (The In subsurface/surface activation and segregation energies are taken to be previously-determined values of 1.8 and 0.2 eV, respectively [33]. The assumed lattice vibration frequency is 10^{13} s^{-1} .) If a relatively dilute ($x=0.25$) alloy is deposited, Fig. 1 shows the way in which the In concentration is predicted by the theory to evolve within the growing flat layer. It is immediately evident that segregation of In to the surface enhances the surface In concentration rapidly above that of the deposition flux so that, for only ~ 1 nm of layer growth, the surface In concentration is already above 40%. It continues to increase as deposition proceeds and is estimated to attain a saturation value of 80–85% for layer thicknesses in excess of ~ 2.5 nm. (It should be noted that an alternative segregation model [34], which allows enrichment of more than one near-surface layer, gives concentrations of surface-segregated material within about 10% of those just specified. However, reliable InGaAs segregation energy values are not yet available for this model, so that it is not considered further here.)

The continuous curve in Fig. 1 shows the change in the In concentration in the surface layer during growth. Of course, it is important to determine the variation in predicted surface In concentration as a function of deposition flux composition. The curves in Fig. 2 present this quantity for deposition fluxes containing from 5% to 100% In. For each deposition condition, the surface In concentration increases progressively

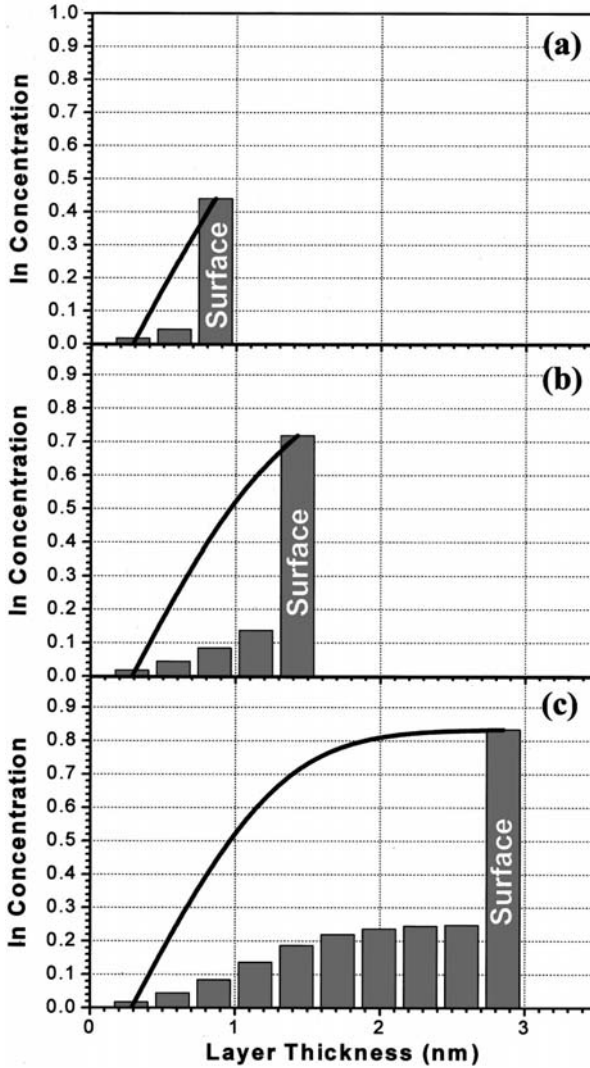


Figure 1. Composition variations in the near-surface layers driven by In segregation to the surface under the Dehaese *et al.* [33] model for a deposition flux with 25% In and different total growth thicknesses: (a) 3 monolayers (ML), (b) 5 ML and (c) 10 ML

to a saturation value which, itself, rises with increasing deposition flux concentration.

Under the WCNH mechanism, a critical surface concentration of In (and associated strain) must build up before the SK islanding transition can take place. Now, it has been shown [10] that a deposition flux of 25% In is approximately the lowest that will induce the SK transition.

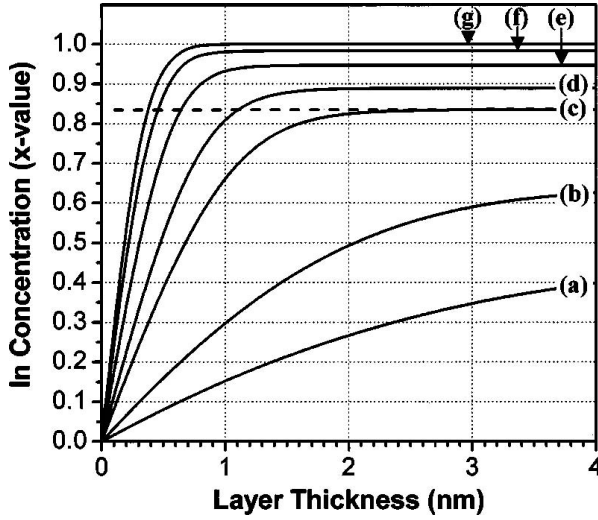


Figure 2. Composition variations in the surface ML, driven by In segregation to the surface, under the Dehaese *et al.* [33] model for deposition fluxes with (a) 5% In, (b) 10% In, (c) 25% In, (d) 35% In, (e) 55% In, (f) 80% In and (g) 100% In.

Therefore, it is possible to identify the corresponding critical surface In concentration from the associated curve in Fig. 2: the critical concentration would be predicted to be 80–85% In in the surface layer. It is then predicted that, for any particular deposition flux, the SK transition will take place after the surface In concentration rises to this critical level. It follows that the islanding transition points for layers grown over the complete range of deposition fluxes can be estimated from plots of the type given in Fig. 2, so that it is possible to estimate the critical thickness for the transition of the initial flat “wetting” layer as a function of deposited In concentration. This procedure gives the continuous “theoretical” curve in Fig. 3, which extends from 0.3 nm thickness for InAs deposition to approximately 2.5 nm thickness for a deposition flux of 25% In.

3. $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ Experimental Results

In the present work, for correlation with theory, $\text{In}_x\text{Ga}_{1-x}\text{As}$ alloy layers were grown on $(001) \pm 1^\circ$ GaAs by MBE and the thicknesses of “wetting” layers at the SK 2D-3D transition point were measured directly *ex situ* using the transmission electron microscope (TEM). Growth took place upon heat-cleaned (001) substrates exhibiting the As (2×4) reconstruction under an As overpressure and with a relatively high substrate

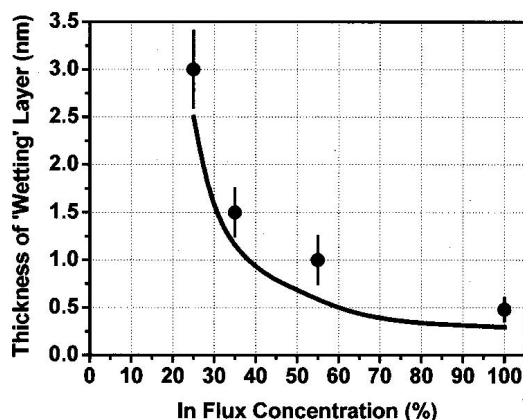


Figure 3. Variation in critical thickness of the initial flat layer for the islanding transition as a function of In concentration in the deposition flux. Measured values presented as data points and predicted values based upon the WCNH mechanism presented as continuous curve.

temperature of 540°C, chosen to encourage segregation whilst In desorption can still be considered insignificant. Growth was terminated and each substrate rapidly cooled under As flux slightly before the transition in the reflection high-energy electron-diffraction (RHEED) pattern from 2D “streaked” to 3D “spotty” was complete, since the pattern change is a little insensitive to the precise SK islanding transition point. Structural and compositional studies of these layers have been carried out using a JEOL 2010F field emission gun TEM, employing samples thinned to electron transparency in cross-sectional configuration by sequential mechanical polishing and low voltage ion beam milling.

In Fig. 4(a) the cross-sectional view of a typical island is shown and the “wetting” layer, of ~3 nm thickness, is visible as indicated by the arrows. [The nonuniform In distribution within each growth island measured by energy-filtered TEM imaging [30] is visible in Fig. 4(b).] Similar observations have been made for growth with In deposition flux concentrations of 35% and 55%, which have yielded critical “wetting” layer thicknesses of 1.5 nm and 1.0 nm, respectively. These experimental thicknesses are also plotted in Fig. 3, together with a critical “wetting” layer thickness of 0.4 nm (1.6 monolayers (ML) [20]) for InAs deposition. Other measurements of wetting layer thickness have been reported [23], but they appear to have relied wholly upon RHEED observations which, due to insensitivity, may have resulted in overestimates of the critical thickness values.

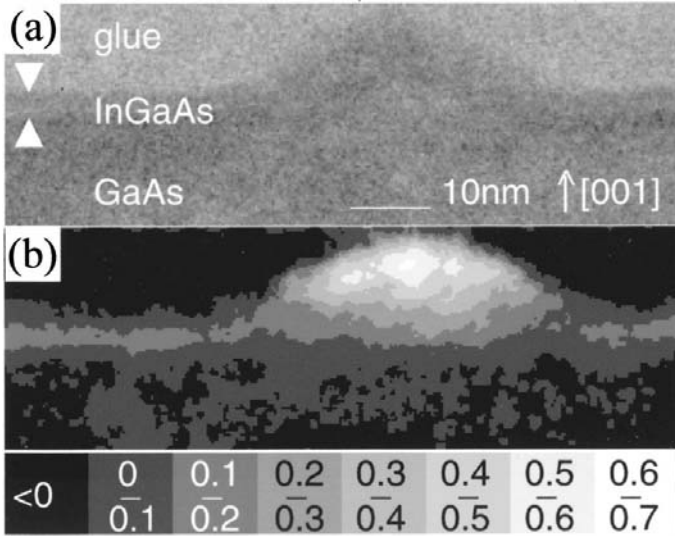


Figure 4. Cross-sectional TEM images of initial quantum dot formed for InGaAs growth upon GaAs with a deposition flux having 25% In: (a) bright-field image with arrows indicating the “wetting” layer and (b) image showing In composition variations in quantized steps (modified from [30]).

It is clearly evident from Fig. 3 that the curves derived from theory and experiment exhibit exactly the same form and are displaced from one another by no more than $\sim 0.1\text{--}0.5$ nm, depending upon In flux concentration. This small displacement of the curves may result, at least in part, from the fact that the theoretical curve does not allow for any induction period required prior to the formation of island nuclei after the “wetting” layer has achieved its critical surface In concentration. In addition, it is possible that implementation of a correctly calibrated multilayer segregation model [34] would give an even closer fit to experiment. Of particular significance, the near coincidence of the curves lends strong support for the importance of segregation in determining the SK 2D-3D transition point, as proposed in the WCNH mechanism [30, 31].

4. Atomistic Simulation of $\text{In}_x\text{Ga}_{1-x}\text{As}$ Surface Layer

When large amounts of In segregate to the growth surface of the initial flat layer, a heavily strained surface layer would be expected to be produced: indeed, the increase in surface lattice parameter has been observed by RHEED [35]. The magnitude of this strain has been estimated by atomistic simulation of the dependence of the elastic energy of the

with a rigid boundary at the bottom, open boundary at the top, and periodic boundaries in the [100] and [010] directions. The structure consisted of 10 ML of GaAs and 10 ML of $\text{In}_x\text{Ga}_{1-x}\text{As}$ with increasing In concentration, as predicted using the kinetic segregation model [31]. An As-stabilised (2×4) reconstructed $\text{In}_x\text{Ga}_{1-x}\text{As}$ (001) surface with As-As dimers [39] was reproduced by the atomic model.

The structures were relaxed using the OXON code [40], and the elastic energy of the uppermost continuous monolayer was evaluated directly from the components of the strain tensor [41]. The expression for the elastic energy is [42]

$$\int_V \left[C_{11}(e_{xx}^2 + e_{yy}^2 + e_{zz}^2) + C_{12}(e_{xx}e_{yy} + e_{xx}e_{zz} + e_{yy}e_{zz}) \right. \\ \left. + C_{44}(e_{xy}^2 + e_{xz}^2 + e_{yz}^2) \right] dV ,$$

where C_{ij} are the elastic constants and e_{ij} are the components of the strain tensor: the integrand is effectively the free energy of the system, which is a function of position and of the local composition. The element of volume dV is the volume effectively occupied by one atom in a strained unit cell [41]. The strain components were evaluated taking into consideration the local composition at every point, and not the average composition of the monolayer, in order to correctly reproduce the effect of the presence of random atomic distributions in each alloy [38].

The results of the strain calculations are shown in Fig. 5. It is clear from the figure that, as the concentration of In increases, there is a continuous increase in the elastic energy stored in the surface, and hence also the surface strain.

5. Discussion

5.1 General SK Transition Phenomena for $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$

Effects of the changing mobile atom fraction at the surface should first be considered. Based upon the WCNH model for the SK transition, as the transition point is approached the concentration of In in the surface layer approaches its maximum value driven by segregation. The large and increasing strain in the growing surface layer (see Fig. 5) makes it increasingly difficult for atoms from the deposition flux to join the surface layer, so that the number density of adatoms increases also. Indeed, at the SK transition point, the number density of adatoms is great enough to force the formation of stable 3D islands, which can then grow. The

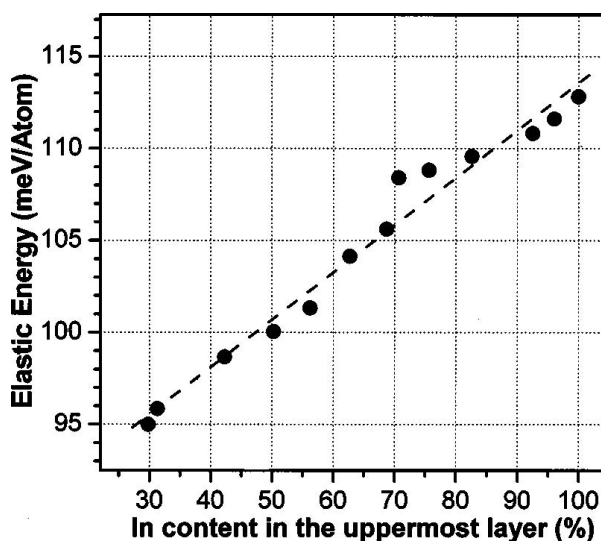


Figure 5. Variation in calculated elastic energy per atom for the uppermost continuous monolayer as a function of the local In concentration.

layer, so that the number density of adatoms increases also. Indeed, at the SK transition point, the number density of adatoms is great enough to force the formation of stable 3D islands, which can then grow. The agreement between theory and experiment for critical “wetting” layer thickness, as presented in Fig. 3, is very clear.

The islands, once formed, are stabilised by stress relief due to expansion of laterally unconstrained lattice planes [43]. The great and increasing stability of the islands would ensure that adatoms which impinge upon them have a high probability of sticking. Thus the stable islands would become very effective sinks for adatoms so that, immediately after the SK transition, the surface number density of adatoms would decrease rapidly, reaching a more stable reduced value as island capture zones overlap. The accompanying diagram in Fig. 6 qualitatively illustrates these changes. At the SK transition point, the overall surface structure is very strained and, in the case of $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$, the general disorder together with the high adatom number density gives a poor RHEED pattern: indeed, the surface structure may be viewed as held in place by the very large adatom density. When the latter decreases due to the formation of the first stable growing $\text{In}_x\text{Ga}_{1-x}\text{As}$ islands, the surface structure equilibria would be disturbed and, due to the presence of the very high surface strains, the uppermost surface layer may begin to dissociate. This process could cause step edges to retreat and vacan-

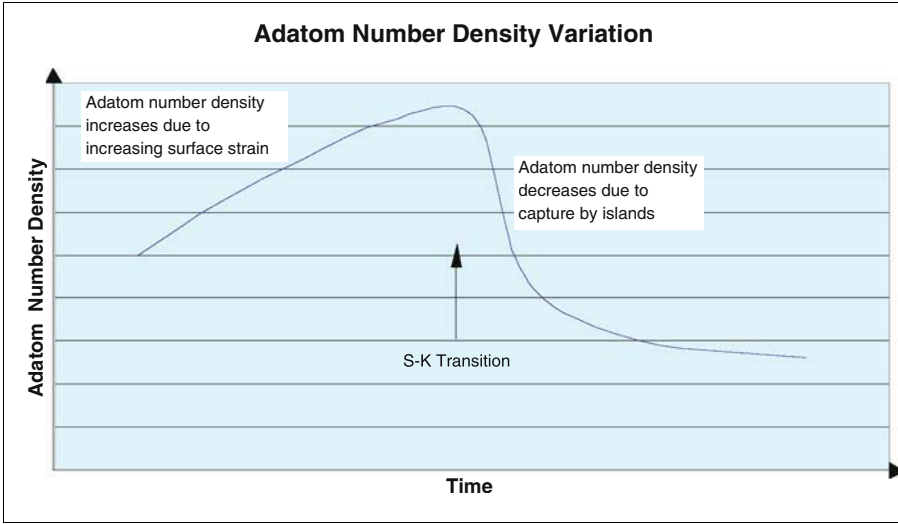


Figure 6. Theoretically-expected variation in surface adatom number density as a function of time close to the SK transition point.

cies to form within the layer itself. Such processes would increase the surface disorder and lead to a further weakening of the RHEED pattern, as is observed for the $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ system [44]. The unravelling of the surface atomic structure (possibly in a liquid-like manner) would proceed until the surface/near-surface layers with the highest In concentrations (formed originally due to segregation) are removed, whereupon a new surface structure would be established at a lower In concentration and in equilibrium with the reduced adatom population. At this point just after the SK transition, a strong RHEED pattern would be expected to return, again in agreement with observation [44], and the growing islands would give a superposed spot pattern when their volume is great enough to be detected.

The increasing surface strain energy, resulting from the increasing surface In concentration, has a pronounced effect upon layer surface growth structures. For example, for a low deposition flux In concentration of 15%, as shown in the atomic force microscope (AFM) image in Fig. 7, the layer-by-layer growth results in the formation of a very flat surface with evenly-spaced ML surface steps and few, if any, ML islands [10]. However, an increase in the deposition flux In concentration to 20% (Fig. 8) results in a strong change in surface step configurations with the formation of narrow protuberances running ahead of step fronts and the production of significant numbers of ML islands (such as at X

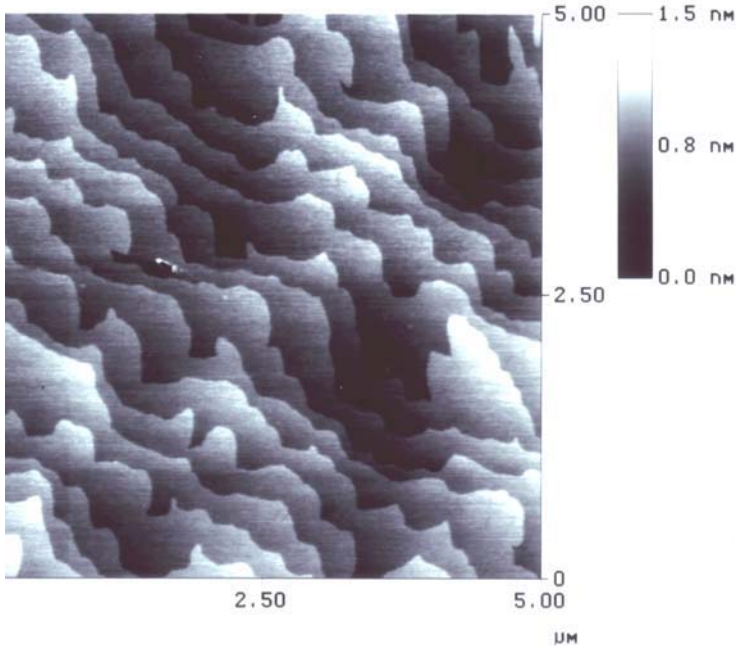


Figure 7. AFM image of $\text{In}_x\text{Ga}_{1-x}\text{As}$ layer ($x = 0.15$) 30 nm thick (after Ref. [10]).

in Fig. 8) [10]. The step front distortions would be expected to yield elastic relaxation by unidirectional expansion across the narrow protuberances. The ML islands are the expected [22] precursors of 3D islands produced by the SK transition when the deposition flux In concentration is increased by a further 5%. Similar ribbon-like wetting layer surface structures are seen in other systems (e.g. AlSb/GaAs , GaSb/GaAs and InSb/GaAs) for layers undergoing the SK islanding transition [24]. For deposition of InAs on GaAs, an initial delay in the increase of the surface In concentration to 100% [curve (g) in Fig. 2] results from the exchange of In atoms with Ga atoms in the GaAs surface. In this latter case, partial completion of the second ML is required before the surface In concentration and associated strain become high enough to trigger the SK 3D-islanding transition. Once again, before the transition occurs, layer growth processes are severely disturbed by the increasing surface strain [20].

It is interesting that, for $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$, the SK transition takes place only [39] on the As-stabilised (2×4) reconstructed (001) surface and, as shown above, requires a surface In concentration close to 100%. It is possible that the group-III-stabilised (4×2) reconstruction on (001) and the (110) and (111) crystal orientations confer altered surface ener-

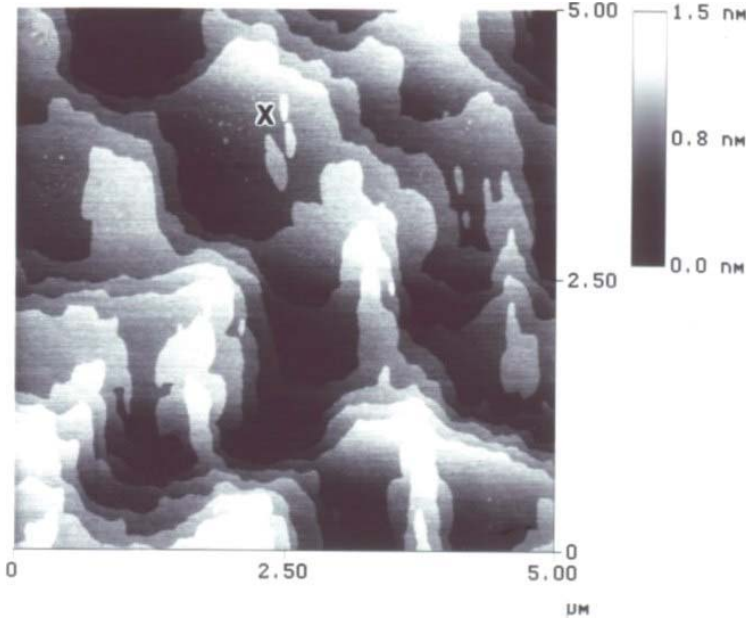


Figure 8. AFM image of $\text{In}_x\text{Ga}_{1-x}\text{As}$ layer ($x=0.2$) 30 nm thick (after Ref. [10]).

gies so that segregation build-up of even 100% In is insufficient to initiate the SK islanding transition.

5.2 The SK Transition in $\text{Si}_{1-x}\text{Ge}_x/\text{Si}$ System

Growth in the $\text{Si}_{1-x}\text{Ge}_x/\text{Si}$ epitaxial system has been studied by many workers [11, 16, 18, 25, 26, 27] and it is important to consider the nature of the islanding transition. Indeed, it has been shown [18, 26, 45] that, for the growth of Ge on Si, the initially deposited monolayers are flat and exhibit scattered pits or trenches. At a critical thickness, the SK transition occurs by the nucleation of islands (generally “hut clusters”) located at the edges of the pits, while the layer between the islands remains flat [26, 45, 46]. This appears to be an SK transition with the same characteristics as that which occurs for $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$, while the distribution of the islands produced is simply determined by the distribution of the pre-existing pits.

It has been indicated that, for $\text{Si}_{1-x}\text{Ge}_x$ alloy growth on Si, island formation can take place due to an Asaro–Tiller–Grinfeld (ATG) instability [27, 47, 48, 49] under a restricted range of growth conditions. Therefore, it is crucial to confirm the island formation mechanism for $\text{Si}_{1-x}\text{Ge}_x$ alloy layer growth. Focusing first upon MBE, deposition of the alloy on Si

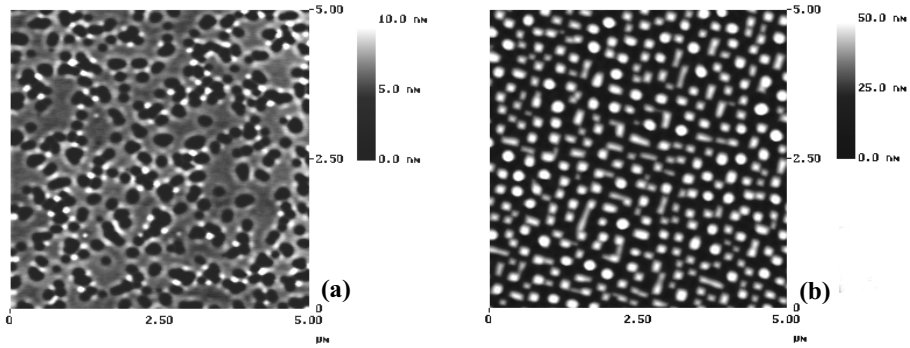


Figure 9. AFM images of $\text{Si}_{0.7}\text{Ge}_{0.3}$ alloy layers grown on (001) Si at 700°C by MBE at layer thicknesses of (a) 2.5 nm and (b) 10 nm.

has been carried out for a range of layer thicknesses, followed by studies of the growth morphology using the AFM [50]. Growth of layers with x -values of 0.3 and 0.5 have demonstrated that the islanding mechanism has fundamentally the same SK nature as for basic Ge/Si. This is illustrated in Fig. 9, which shows a layer with $x=0.3$ at mean deposit thicknesses of 2.5 nm and 10 nm. It is clear that, at 2.5 nm deposit thickness, the overall layer is very flat but exhibits an array of shallow flat-bottomed pits (the dark features). However, at this thickness, island formation has already taken place such that all the islands (the bright features) lie along pit edges. Indeed, the general growth structure seems to be of the type which yielded “quantum fortress” structures seen for lower temperature growth [51]. When the mean deposit thickness is increased to 10 nm, the islands have grown substantially larger and have increased somewhat in number density. Most importantly, there is no evidence for an ATG instability [see, e.g. Fig. 9(a)], which would induce rippling of the whole surface. Therefore, we may say that, for MBE growth and over at least the range $x=0.3$ –1, the morphological characteristics of islanding in the $\text{Si}_{1-x}\text{Ge}_x/\text{Si}$ system are as expected for a standard SK transition. In contradistinction, for growth of $\text{Si}_{1-x}\text{Ge}_x/\text{Si}$ by CVD, it has been reported [27] that an SK transition to island formation occurs for $x > 0.6$, while for $0.2 < x < 0.6$ an ATG (nucleationless) instability leads to island growth.

For CVD deposition, the presence of hydrogen surfactant would be expected [52] to at least partially suppress segregation processes: this appears to favour an ATG instability for alloy growth with the lower x -values. However, the SK transition is confirmed to occur in the standard manner (islanding on flat wetting layers) for the higher x -values during CVD growth and, as described above, for all investigated x -

values during MBE growth. We shall demonstrate elsewhere [50] that the WCNH mechanism provides a satisfying quantitative description of the $\text{Si}_{1-x}\text{Ge}_x/\text{Si}$ islanding processes by the standard SK transition mechanism.

It should be noted that ordered ripple arrays and mounds are seen upon $\text{Si}_{1-x}\text{Ge}_x/\text{Si}$ layer surfaces at greater growth thicknesses [43]. In such circumstances, it is likely that these features can form either by overlap of existing enlarged SK islands with associated undulation realignment (exactly in the manner seen for $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ layer growth [10]), or by evolution of the structures produced by the early-onset ATG instability for CVD growth in the appropriate range of alloy x -values.

5.3 The SK Transition in other Epitaxial Systems

The experimental work described here relates primarily to the MBE growth of strained layers showing the SK islanding transition in the $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ and $\text{Si}_{1-x}\text{Ge}_x/\text{Si}$ systems. Indeed, many other MBE-grown epitaxial systems of similar materials with high misfit exhibit the SK transition. It is conjectured that, in most if not all cases, the WCNH mechanism involving elemental segregation is likely to underpin the initial island formation process. However, for layer growth based upon CVD of the various types, the hydrogen surfactant generally present upon the growth surface to some degree would be expected to suppress elemental segregation to the surface [52] and lead to at least a modification of layer growth characteristics. These could possibly involve an increase in wetting layer thickness or, in the limit with sufficient suppression of segregation, it is possible that an ATG instability may occur. Further careful study of the earliest stages of islanding during CVD layer growth would be very appropriate.

6. Conclusions

The manner in which the WCNH segregation mechanism is expected to drive the SK islanding transition is considered in detail. Across the complete range of deposition fluxes for MBE growth of the (001) $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ epitaxial system, there is excellent agreement between theory and experiment regarding derivation of critical “wetting” layer thickness. This gives strong support for the segregation-based mechanism. Furthermore, the morphological characteristics of the initial growth of $\text{Si}_{1-x}\text{Ge}_x/\text{Si}$ for both MBE and ranges of CVD deposition show the same features of SK islanding and are expected to be similarly driven by elemental segregation. It is considered that, for all other

epitaxial systems exhibiting the SK transition, when elemental segregation within the “wetting” layer is not suppressed it will be a key factor determining the critical point at which islanding occurs.

Acknowledgements

The authors would like to thank Dr T. Grasby for growth of the $\text{Si}_{1-x}\text{Ge}_x/\text{Si}$ samples and EPSRC for financial support. Very helpful discussions with B. A. Joyce, A. Madhukar, I. Goldfarb and J. Tersoff are gratefully acknowledged.

References

- [1] F. C. Frank and J. H. van der Merwe. One-dimensional dislocations. I. Static theory, *Proc. Roy. Soc. London, Ser. A* **198**: 205–215, 1949.
- [2] M. Volmer and A. Weber. Keimbildung in übersättigten gebilden. *Z. Physik. Chem.* **119**: 277–301, 1926.
- [3] I. N. Stranski and L. Krastanov. Zur theorie der orientierten ausscheidung von ionenkristallen aufeinander. *Sitz. Ber. Akad. Wiss. Math.-naturwiss. Kl. Abt. IIb* **146**: 797–810, 1938.
- [4] J. A. Venables, G. D. T. Spiller and M. Hanbücken. Nucleation and growth of thin films. *Rep. Prog. Phys.* **47**: 399–459, 1984.
- [5] F. Glas, C. Guille, P. Hénoc and F. Houzay. TEM study of the molecular beam epitaxy island growth of InAs on GaAs. in *Microscopy of Semiconducting Materials 1987*, edited by A. G. Cullis and P. D. Augustus (IOP, Bristol, 1987) p. 71.
- [6] S. Guha, A. Madhukar, and K. C. 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**: 2110–2112, 1990.
- [7] C. W. Snyder, B. G. Orr, D. Kessler, and L. M. Sander. Effect of strain on surface morphology in highly strained InGaAs films. *Phys. Rev. Lett.* **66**: 3032–3035, 1991.
- [8] D. Leonard, M. Krishnamurthy, C. M. 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**: 3203–3205, 1993.
- [9] J. M. Moison, F. Houzay, F. Barthe, L. Leprince E. André, and O. Vatel. Self-organized growth of regular nanometer-scale InAs dots on GaAs. *Appl. Phys. Lett.* **64**: 196–198, 1994.
- [10] A. G. Cullis, A. J. Pidduck, and M. T. Emeny. Growth morphology evolution and dislocation introduction in the InGaAs/GaAs heteroepitaxial system. *J. Crystal Growth* **158**: 15–27, 1996.
- [11] T. R. Ramachandran, R. Heitz, N. P. Kobayashi, A. Kalburge, W. Yu, P. Chen, and A. Madhukar. Re-entrant behavior of 2D-to-3D morphology change and 3D island lateral size equalization via mass exchange in Stranski–Krastanov growth: InAs on GaAs (001). *J. Crystal Growth* **175/176**: 216–223, 199.

- [12] N. Carlsson, W. Seifert, A. Petersson, P. Castrillo, M.-E. 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**: 3093–3095, 1994.
- [13] P. M. Petroff and S. P. DenBaars. MBE and MOCVD growth and properties of self-assembling quantum-dot arrays in III-V semiconductor structures. *Superlatt. Microstruct.* **15**: 15–21, 1994.
- [14] B. R. Bennett, B. V. Shanabrook, P. M. Thibado, L. J. Whitman, and R. Magno. Stranski–Krastanov growth of InSb, GaSb, and AlSb on GaAs: Structure of the wetting layers. *J. Crystal Growth* **175/176**: 888–893, 1997.
- [15] D. J. Eaglesham and M. Cerullo. Dislocation-free Stranski–Krastanov growth of Ge on Si(100). *Phys. Rev. Lett.* **64**: 1943–1946, 1990.
- [16] Y.-W. Mo, D. E. Savage, B. S. Swartzentruber, and M. G. Lagally. Kinetic pathway in Stranski–Krastanov growth of Ge on Si(001). *Phys. Rev. Lett.* **65**: 1020–1023, 1990.
- [17] M. Tomitori, K. Watanabe, M. Kobayashi, and O. Nishikawa. STM study of the Ge growth mode on Si(001) substrates. *Appl. Surf. Sci.* **76/77**: 322–328, 1994.
- [18] I. Goldfarb and G. A. D. Briggs. The effect of mismatch strain on Stranski–Krastanov transition in epitaxial $\text{Ge}_x\text{Si}_{1-x}\text{Si}(001)$ gas-source growth. *J. Crystal Growth* **198/199**: 1032–1038, 1999.
- [19] D. E. Jesson, M. Kästner, and B. Voigtländer. Direct observation of subcritical fluctuations during the formation of strained semiconductor islands. *Phys. Rev. Lett.* **84**: 330–333, 2000.
- [20] R. M. Tromp, F. M. Ross, and M. C. Reuter. Instability-driven SiGe island growth. *Phys. Rev. Lett.* **84**: 4641–4644, 2000.
- [21] C. Ratsch and A. Zangwill. Equilibrium theory of the Stranski–Krastanov epitaxial morphology. *Surf. Sci.* **293**: 123–131, 1993.
- [22] J. Tersoff and F. K. LeGoues. Competing Relaxation Mechanisms in Strained Layers. *Phys. Rev. Lett.* **72**: 3570–3573, 1994.
- [23] V. A. Shchukin, N. N. Ledentsov, P. S. Kopev and D. Bimberg. Spontaneous ordering of arrays of coherent strained islands. *Phys. Rev. Lett.* **75**: 2968–2971, 1995.
- [24] C. Priester and M. Lannoo. Origin of self-assembled quantum dots in highly mismatched heteroepitaxy. *Phys. Rev. Lett.* **75**: 93–96, 1995.
- [25] Y. Chen and J. Washburn. Structural transition in large-lattice-mismatch heteroepitaxy. *Phys. Rev. Lett.* **77**: 4046–4049, 1996.
- [26] H. T. Dobbs, D. D. Vvedensky, and A. Zangwill. Theory of quantum dot formation in Stranski–Krastanov systems. *Appl. Surf. Sci.* **123/124**: 646–652, 1998.
- [27] J. Tersoff. Enhanced nucleation and enrichment of strained-alloy quantum dots. *Phys. Rev. Lett.* **81**: 3183–3186, 1998.
- [28] K. Nakajima. Thickness-composition diagrams of Stranski–Krastanov mode in the GaPSb/GaP and InGaAs/GaAs systems. *J. Crystal Growth* **203**: 376–386, 1999.
- [29] B. A. Joyce, D. D. Vvedensky, T. S. Jones, M. Itoh, G. R. Bell, and J. G. Belk. *In situ* studies of III-V semiconductor film growth by molecular beam epitaxy. *J. Crystal Growth* **201/202**: 106–112, 1999.

- [30] T. Walther, A. G. Cullis, D. J. Norris, and M. Hopkinson. Nature of the Stranski–Krastanov transition during epitaxy of InGaAs on GaAs. *Phys. Rev. Lett.* **86**: 2381–2384, 2001.
- [31] A. G. Cullis, D. J. Norris, T. Walther, M. A. Migliorato, and M. Hopkinson. Stranski–Krastanov transition and epitaxial island growth. *Phys. Rev. B* **66**: art. no. 081305, 2002.
- [32] S. Fukatsu, K. Fujita, H. Yaguchi, Y. Shiraki, and R. Ito. Self-limitation in the surface segregation of Ge atoms during Si molecular-beam epitaxial growth. *Appl. Phys. Lett.* **59**: 2103–2105, 1991.
- [33] O. Dehaese, X. Wallart, and F. Mollet. Kinetic model of element-III segregation during molecular-beam epitaxy of III-III'-IV semiconductor compounds. *Appl. Phys. Lett.* **66**: 52–54, 1995.
- [34] D. J. Godbey and M. G. Ancona. Concentration-dependence of Ge segregation during the growth of a SiGe buried layer. *J. Vac. Sci. Technol. B* **11**: 1392–1395, 1993.
- [35] A. Madhukar, private communication.
- [36] J. Tersoff. New empirical model for the structural properties of silicon. *Phys. Rev. Lett.* **56**: 632–635, 1986
- [37] J. Tersoff. Modeling solid-state chemistry: Interatomic potentials for multicomponent systems. *Phys. Rev. B* **39**: 5566–5568, 1989.
- [38] M. A. Migliorato, A. G. Cullis, M. Fearn and J. H. Jefferson. Atomistic simulation of strain relaxation in $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ quantum dots with nonuniform composition. *Phys. Rev. B* **65**: art. no. 115316, 2002.
- [39] B. A. Joyce, D. D. Vvedensky, G. R. Bell, J. G. Belk, M. Itoh, and T. S. Jones. Nucleation and growth mechanisms during MBE of III-V compounds. *Mats. Sci. Eng. B* **67**: 7–16, 1999.
- [40] OXON code derived by Materials Department, Oxford University.
- [41] L. D. Landau and E. M. Lifshitz. *Course of Theoretical Physics: Theory of Elasticity* (Pergamon, Oxford, 1986) p. 7.
- [42] M. A. Migliorato, A. G. Cullis, M. Fearn, and J. H. Jefferson. Atomistic simulation of $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ quantum dots with nonuniform composition. *Physica E* **13**: 1147–1150, 2002.
- [43] A. G. Cullis, D. J. Robbins, A. J. Pidduck, and P. W. Smith. The characteristics of strain-modulated surface undulations formed upon epitaxial $\text{Si}_{1-x}\text{Ge}_x$ alloy layers on Si. *J. Crystal Growth* **123**: 333–343, 1992.
- [44] B. A. Joyce, private communication.
- [45] I. Goldfarb, P. T. Hayden, J. H. G. Owen, and G. A. D. Briggs. Nucleation of “hut” pits and clusters during gas-source molecular-beam epitaxy of Ge/Si(001) in *in situ* scanning tunneling microscopy. *Phys. Rev. Lett.* **78**: 3959–3962, 1997.
- [46] I. Goldfarb, private communication.
- [47] A. Asaro and W. Tiller. Interface morphology development during stress corrosion cracking. I. Via surface diffusion. *Metall. Trans.* **3**: 1789–1796, 1972.
- [48] M. A. Grinfeld. Instability of the interface between a non-hydrostatically stressed elastic body and melts. *Dokl. Acad. Nauk. SSSR* **290**: 1358–1363, 1986 [*Sov. Phys. Dokl.* **31**: 831–834, 1986.]

- [49] B. J. Spencer, P. W. Voorhees, and J. Tersoff. Enhanced instability of strained alloy films due to compositional stresses. *Phys. Rev. Lett.* **84**: 2449–2452, 2000.
- [50] D. J. Norris, A. G. Cullis, T. Grasby and E. H. C. Parker, in preparation.
- [51] J. L. Gray, R. Hull, and J. A. Floro. Control of surface morphology through variation of growth rate in SiGe/Si(100) epitaxial films: Nucleation of “quantum fortresses”. *Appl. Phys. Lett.* **81**: 2445–2447, 2002.
- [52] M. Schowalter, M. Melzer, A. Rosenauer, D. Gerthsen, R. Krebs, J.-P. Reithmaier, A. Forchel, M. Arzberger, M. Bichler, G. Abstreiter, M. Grau, M. C. Amann, R. Sellin, and D. Bimberg, Segregation in III-V semiconductor heterostructures studied by transmission electron microscopy. in *Microscopy of Semiconducting Materials 2003*, edited by A. G. Cullis and P. A. Midgley (Institute of Physics, Bristol, 2003) 2004.

OFF-LATTICE KMC SIMULATIONS OF STRANSKI-KRASTANOV-LIKE GROWTH

Michael Biehl and Florian Much

*Institut für Theoretische Physik und Astrophysik and Sonderforschungsbereich 410,
Am Hubland, D-97074 Würzburg*
biehl@physik.uni-wuerzburg.de

Abstract We investigate strained heteroepitaxial crystal growth in the framework of a simplifying $(1+1)$ -dimensional model by use of off-lattice kinetic Monte Carlo simulations. Our modified Lennard–Jones system displays the so-called Stranski–Krastanov growth mode: initial pseudomorphic growth ends by the sudden appearance of strain induced multilayer islands upon a persisting wetting layer.

1. Introduction

In addition to its technological relevance, epitaxial crystal growth is highly attractive from a theoretical point of view. It offers many challenging open questions and provides a workshop in which to develop novel methods for the modeling and simulation of non-equilibrium systems, in general. In particular, strained heteroepitaxial crystal growth attracts significant interest as a promising technique for the production of, for instance, high quality semiconductor films. Recent overviews of experimental and theoretical investigations can be found in, e.g. [1, 2, 3]. A particularly attractive aspect is the possibility to exploit self-organizing phenomena for the fabrication of nanostructured surfaces by means of molecular-beam epitaxy (MBE) or similar techniques.

In many cases, adsorbate and substrate materials crystallize in the same lattice but with different bulk lattice constants. Frequently, one observes that the adsorbate grows *layer by layer*, initially, with the lateral spacing of atoms adapted to the substrate. The misfit induces compressive or tensile strain in this *pseudomorphic* film and, eventually, misfit dislocations will appear. These relax the strain and the adsorbate grows with its natural lattice constant far from the substrate, eventually.

Dislocations will clearly dominate strain relaxation in sufficiently thick films and for large misfits. In material combinations with relatively small misfit an alternative effect governs the initial growth of thin films: Instead of growing layer by layer, the adsorbate aggregates in three-dimensional (3D) structures. The term *3D islands* is commonly used to indicate that these structures are spatially separated. The situation is clearly different from the emergence of *mounds* due to the Ehrlich–Schwoebel (ES) instability [1, 2], for instance.

At least two different growth scenarios display 3D island formation: In *Volmer–Weber* growth, such structures appear immediately upon the substrate when depositing the mismatched material. The situation resembles the formation of non-wetting droplets of liquid on a surface. This growth mode is frequently observed in systems where adsorbate and substrate are fundamentally different, an example being Pb on a graphite substrate [1].

In the following we concentrate on the so-called *Stranski–Krastanov* (SK) growth mode, where 3D islands are found upon a pseudomorphic wetting layer (WL) of adsorbate material [1, 2, 3]. Two prominent prototype SK systems are Ge/Si and InAs/GaAs where, as in almost all cases discussed in the literature, the adsorbate is under compression in the WL. In order to avoid conflicts with more detailed definitions and interpretations of the SK growth mode in the literature we will resort to the term *SK-like growth* in the following. It summarizes the following sequence of phenomena during the deposition of a few monolayers (ML) of material:

- 1 The layer by layer growth of a pseudomorphic adsorbate WL up to a *kinetic thickness* h_{WL}^* .
- 2 The sudden appearance of 3D islands, marking the so-called 2D-3D or SK transition
- 3 Further growth of the 3D islands, fed by additional deposition and by incorporation of surrounding WL atoms
- 4 The observation of separated 3D islands of similar shapes and sizes, on top of a WL with reduced *stationary thickness* h_{WL} .

Admittedly, the above *operative definition* already reflects our interpretation of SK growth to a certain extent. Besides these basic processes a variety of phenomena can play important roles in the SK scenario, including the interdiffusion of materials and the segregation of compound adsorbates. These effects are certainly highly relevant in many cases, see [4, 5] and other contributions to this volume. However, SK-like growth is

observed in a variety of material systems which may or may not display these specific features. For instance, intermixing or segregation should be irrelevant in the somewhat *exotic* case of large organic molecules like PTCDA deposited on a metal substrate, e.g. Ag(111). Nevertheless, this system displays SK-like growth in excellent accordance with the above operative definition [6].

Despite the extensive investigation of SK growth, a complete detailed theoretical picture is still lacking, apparently. This concerns in particular the nature of the 2D-3D transition. One problem clearly lies in the richness of the phenomenon. On the other hand, the very diversity of SK-like systems gives rise to the hope that this growth scenario might be governed by a few basic universal mechanisms. Accordingly, it should be possible to capture and identify these essential features in relatively simple prototype systems.

This hope motivates the investigation of simplifying models without aiming at the reproduction of material specific details. Some of the key questions in this context are: under which conditions does a WL emerge and persist? How does its thickness before and after the SK transition depend on the growth conditions? Which microscopic processes trigger and control the sudden formation of 3D islands? How do the island size and their spatial arrangement depend on the parameters of the system?

Following earlier investigations of related phenomena, e.g. [8–11], we choose a classical pair potential ansatz to represent the interactions between atoms in our model. Here, we restrict our studies to the fairly simple case of a modified Lennard–Jones (LJ) system in $1 + 1$ spatial dimensions, i.e. growth on a one-dimensional substrate surface. As we interpret our model as a cross-section of the physical $(2 + 1)$ -dimensional case, we still use the common term 2D-3D transition for the formation of multilayer from monolayer islands.

We investigate our model by means of kinetic Monte Carlo (KMC) simulation, see e.g. [13] for an introduction and overview. This concept has proven useful in the study of non-equilibrium dynamical systems in general and, in particular, in the context of epitaxial growth, see e.g. [1, 2]. Most frequently, pre-defined lattices are used for the representation of the crystal. So-called solid-on-solid (SOS) models which neglect lattice defects, bulk vacancies, or dislocations have been very successful in the investigation of various relevant phenomena, including scenarios of kinetic roughening or mound formation due to instabilities [1, 2]. There is, however, no obvious way of including mismatch and strain effects in a lattice gas model. A potential route is to introduce additional elastic interactions between neighboring atoms in an effective fashion. In fact, such models of heterosystems have been studied in some

of the earliest KMC simulations of epitaxial growth [14, 15], see [16] for a recent example of a so-called *ball and spring* model. In alternative approaches the strain field of a given configuration is evaluated using elasticity theory as, for instance, in [17].

In order to account for strain effects more faithfully, including potential deviations from a perfect lattice structure, it is essential to allow for continuous particle positions. Given at least an approximation for the interatomic potentials, a molecular dynamics (MD) type of simulation [18] would be clearly most realistic and desirable, see [19] for one example in the context of heteroepitaxial growth. However, this method suffers generally from the restriction to short physical times on the order of 10^{-6} s or less. MBE relevant time scales of seconds or minutes do not seem feasible currently even when applying sophisticated acceleration techniques [20].

Here, we put forward an *off-lattice* KMC method which has been introduced in [7, 8, 11] and apply it in the context of the SK scenario. Some of the results have been published previously in less detail [12]. The paper is organized as follows. In the next section we outline the model and simulation method. Before analysing the actual SK-like growth in Sec. 4 we present some basic results concerning various diffusion scenarios in Sec. 3. In the last section we summarize and discuss open questions and potential extensions of our work.

2. Model and Method

In our off-lattice model we consider pairwise interactions given by LJ potentials of the form [18]

$$U_{ij}(U_0, \sigma) = 4U_0 \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right], \quad (1)$$

where the relative distance r_{ij} of particles i and j can vary continuously. As a widely used approximation we cut off interactions for $r_{ij} > 3\sigma$.

The choice of the parameters $\{U_0, \sigma\}$ in Eq. (1) characterizes the different material properties in our model: interactions between two substrate (adsorbate) particles are specified by the sets $\{U_s, \sigma_s \equiv 1\}$ and $\{U_a, \sigma_a\}$, respectively. Instead of a third independent set we set $U_{as} = \sqrt{U_s U_a}$, $\sigma_{as} = \frac{1}{2}(\sigma_s + \sigma_a)$ for the inter-species interaction.

As the lattice constant of a monatomic LJ crystal is proportional to σ , the relative misfit is given by $\epsilon = (\sigma_a - \sigma_s) / \sigma_s$. Here we consider only cases with $\sigma_a > \sigma_s$, i.e. positive misfits ϵ of the order of a few percent. For the simulation of SK-like growth we set $U_s > U_{as} > U_a$. In such systems, the formation of a WL and potential layer-by-layer growth should be

favorable, in principle. If not otherwise specified, we have set the misfit to $\epsilon = 4\%$, a typical value for SK systems, and used the LJ prefactors $U_s = 1.0$ eV, $U_a = 0.74$ eV, $U_{as} \approx 0.86$ eV.

Growth takes place on a substrate represented by six atomic layers, with the bottom layer fixed and periodic horizontal boundary conditions. In the following we mostly refer to systems with $L = 800$ particles per substrate layer, additional simulations with $L = 400$ or 600 revealed no significant L -dependence of the results presented here. The deposition of single adsorbate particles is performed with a rate $R_d = L F$, where F is the deposition flux. As we interpret the substrate lattice constant as our unit of length, flux and deposition rate (measured in ML/s) assume the same numerical values. The rates of all other significant changes of the configuration are given by Arrhenius laws. We consider only hopping diffusion events at the surface and neglect bulk diffusion, exchange processes or other concerted moves. Furthermore, diffusion is restricted to adsorbate particles at the surface whereas jumps of substrate particles onto the surface are not considered. As we will demonstrate in a forthcoming publication, these simplifications are justified for small misfits ϵ in the LJ system because the corresponding rates are extremely low, see also [7, 8].

The rate R_i for a particular event i is taken to be of the form

$$R_i = \nu_0 \exp\left(-\frac{E_i}{k_B T}\right), \quad (2)$$

where T is the simulation temperature and k_B the Boltzmann constant. For simplicity, we assume that the attempt frequency ν_0 is the same for all diffusion events. In order to relate to physical units we use $\nu_0 = 10^{12} \text{ s}^{-1}$ wherever numerical results are given. The activation barriers E_i are calculated *on-line* given the actual configuration of the system. This can be done by a minimal energy path saddle point calculation [9, 18]. Here, we use a *frozen crystal* approximation, which speeds up the calculation of barriers significantly (see [21] for an application in the context of strained surfaces). Note that the calculations are particularly simple in $1 + 1$ dimensions: the path between neighboring local minima of the potential energy is uniquely determined and the transition state corresponds to the separating local maximum.

An important modification concerns interlayer diffusion. LJ systems in $1 + 1$ dimensions display a strong additional barrier, which hinders such moves at terrace edges [1, 2]. This so-called Ehrlich–Schwoebel (ES) effect is by far less pronounced in $(2 + 1)$ -dimensional systems, because interlayer moves follow a path through an energy saddle point rather than the pronounced maximum at the island edge. In our investi-

gation of the SK-like scenario we remove the ES barrier for all interlayer diffusion events *by hand*. One motivation is the above mentioned overestimation in one dimension. More importantly, we wish to investigate strain-induced island formation without interference of the ES instability. Note that the latter leads to the formation of mounds even in homoepitaxy [1, 2].

The rates for deposition and diffusion are used in a rejection-free KMC simulation [13]. Using a binary search tree technique, one of the possible events j is drawn with probability R_j / R , where $R = (R_d + \sum_i R_i)$ is the total rate of all potential changes. Time is advanced by a random interval τ according to the Poisson distribution $P(\tau) = R e^{-R\tau}$ [13]. In order to avoid the artificial accumulation of strain due to inaccuracies of the method, the potential energy should be taken to the nearest local minimum by variation of all particle positions in the system. This relaxation process affects both, adsorbate and substrate atoms. In order to reduce the computational effort, we restrict the variation to particles within a radius $3\sigma_s$ around the location of the latest event, in general. The global minimization procedure is performed only after a distinct number of steps. It is important to note neither procedure leads to a substantial rearrangement. Significant changes of the activation energies due to relaxation signal the necessity to perform the global procedure more frequently.

3. Diffusion Processes

Before analysing the SK-like scenario, we compare the barriers for hopping diffusion in various settings on the surface. The investigation of systems like Ge/Si(001) reveal a very complicated scenario due to anisotropies and the influence of surface reconstructions [22, 23]. For Ge on Ge(111) the barrier for hopping diffusion is higher on the surface of a compressed crystal, whereas diffusion is faster on relaxed Ge [23].

Lennard–Jones or similar models do not, in general, reproduce this feature. Schroeder and Wolf [21] consider $(2+1)$ -dimensional single species LJ systems and evaluate the diffusion barrier for a single adatom on surfaces in various lattice types. Among other results they find that mechanical compression of the crystal lowers the barrier for surface diffusion. However, in the mismatched two species system, it is more important to compare diffusion on (a) the substrate, (b) the WL and (c) the surface of partially relaxed islands.

The strong adsorbate-substrate interaction ($U_{as} > U_a$) favors the formation of a WL, but it also yields deep energy minima and a relatively high diffusion barrier for adsorbate particles on the substrate. This ef-

fect is much weaker for particles on a complete wetting monolayer of adsorbate and, hence, the corresponding diffusion barrier is significantly lower. The faster diffusion further stabilizes the WL, as deposited particles will reach and fill in gaps easily. In principle, the trend extends to the following layers. However, due to the short range nature of the LJ potential the influence of the substrate essentially vanishes on WL of three or more monolayers. With the example choice $\epsilon = 4\%$ and $U_s = 1.0$ eV, $U_a = 0.74$ eV, $U_{as} \approx 0.86$ eV, we find an activation barrier of $E_a^0 \approx 0.57$ eV for adsorbate diffusion on the substrate and $E_a^1 \approx E_a^2 \approx 0.47$ eV for diffusion on the first and second adsorbate layer, respectively.

For the SK scenario the diffusion on islands of finite extension is particularly relevant. Figure 1 shows the barriers for diffusion hops on islands of various heights located upon a wetting monolayer. We wish to point at two important features: (i) Diffusion on top of islands is, in general, slower than on the WL and the difference increases with the island height. In our model, this is an effect of the partial relaxation or over-relaxation in the island top layer. (ii) Depending on the lateral

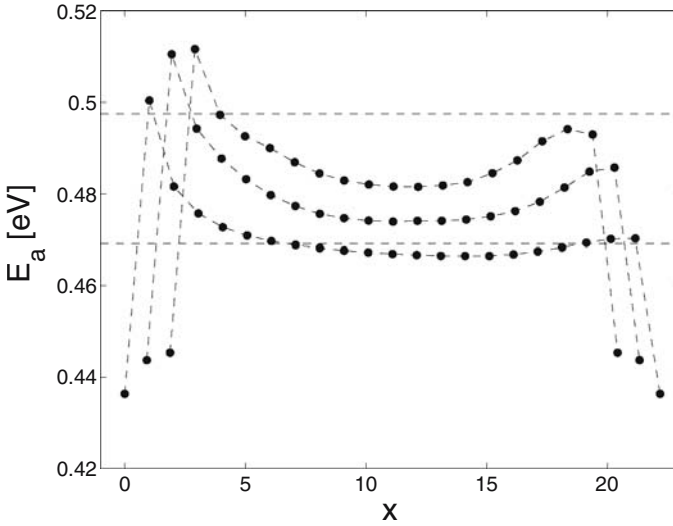


Figure 1. Diffusion barriers as obtained in our model for a single adatom on a flat symmetric multilayer island with 24 base particles and height 1, 3, 5 layers (bottom to top curves). Symbols represent the activation energies for hops from the particle position x to the left neighbor site. The island is placed on top of a single WL, interaction parameters are given in the text of section 3. The leftmost barriers correspond to downward jumps at the island edge with suppression of the ES effect. Horizontal lines mark the barrier for adatom diffusion on the WL (lower line) and on perfectly relaxed adsorbate material (upper line).

island size and its height, there is a more or less pronounced diffusion bias towards the island center, reflecting the spatially inhomogeneous relaxation. A similar effect has been observed in $(2+1)$ -dimensional LJ systems [21]. Note that (ii) has to be distinguished from the diffusion bias imposed by the ES effect, which would be present even in homoepitaxy and with particle positions restricted to a perfect undisturbed lattice. Clearly, (i) and (ii) favor the formation of islands upon islands and hence play an important role in SK-like growth. They concern adatoms which are deposited directly onto the islands as well as particles that hop upward at edges, potentially. As we will argue in the following section, upward diffusion moves play the more important role in the 2D-3D transition of our model.

4. SK-Like Growth Scenario

In our investigation we follow a scenario which is frequently studied in experiments [24, 25]. In each simulation run a total of 4 ML adsorbate material is deposited at rates in the range $0.5 \text{ ML/s} \leq R_d \leq 9.0 \text{ ML/s}$. After deposition is complete, a relaxation period with $R_d = 0$ of about 10^7 diffusion steps follows, corresponding to a physical time on the order of 0.3 s. As we have demonstrated in [11], strain relaxation through dislocations is not expected for, say, $\epsilon = 4\%$ within the first few adsorbate layers. Indeed misfit dislocations were observed in none of the simulations presented here. Results have been obtained on average over at least 15 independent simulation runs for each data point.

In our simulations we observe the complete scenario of SK-like growth as described in the introduction. Illustrating mpeg movies of the simulations are available upon request or directly at our web pages [26]. A section of a simulated crystal after deposition and island formation is shown in Figure 2. Ultimately, the formation of islands is driven by the relaxation of strain. As shown in the figure, material within the 3D island and at its surface can assume a lattice constant close to that of bulk adsorbate. On the contrary, particles in the WL are *forced* to adapt the substrate structure.

During deposition, monolayer islands located on the WL undergo a rapid transition to bilayer islands at a well-defined thickness h_{WL}^* . For the systematic determination of h_{WL}^* we follow [24] and fit the density ρ of 3D islands as $\rho = \rho_0 (h - h_{\text{WL}}^*)^\alpha$, finding comparable values of the exponent α . Figure 3 displays the results for two different substrate temperatures T and various deposition rates R_d . The increase of h_{WL}^* with decreasing T agrees qualitatively with several experimental findings, see [25] as one example. Heyn [4] discusses the effect of adsorbate/substrate

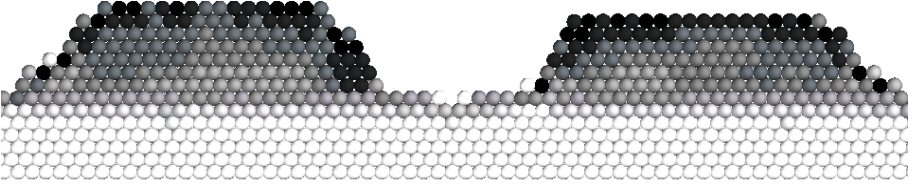


Figure 2. A section of a simulated crystal as obtained for $R_d = 7.0$ ML/s and $T = 500$ K. Islands are located on a stationary WL with $h_c \approx 1$, the six bottom layers represent the substrate. The darker a particle is displayed, the larger is the average distance from its nearest neighbors.

interdiffusion on the T -dependence of h_{WL}^* in the InGaAs system. Reassuringly, our result is consistent with his findings without intermixing.

A key observation is the significant increase of h_{WL}^* with increasing deposition flux. Qualitatively the same flux dependence is reported for InP/GaAs heteroepitaxial growth in [25]. This behavior leads to the conclusion that the emergence of islands upon islands, i.e. the SK transition, is mainly due to particles performing upward hops onto existing monolayer islands. If, on the contrary, the formation of second or third layer nuclei by freshly deposited adatoms was the dominant process, one

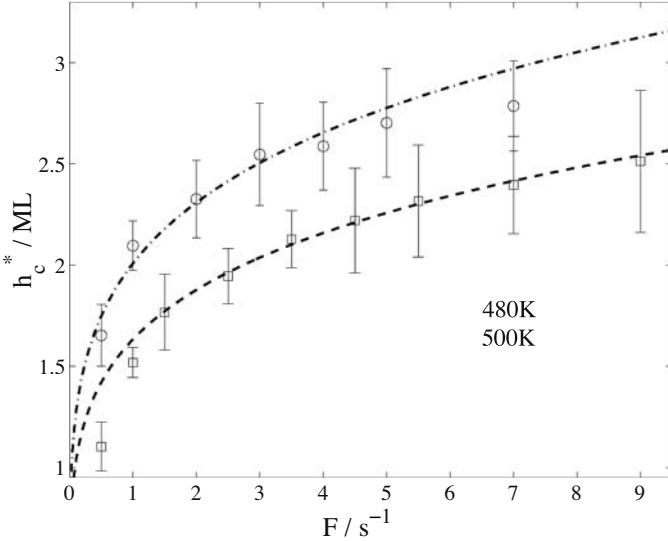


Figure 3. Kinetic critical WL thickness h_{WL}^* versus the deposition flux for two different substrate temperatures. Both curves correspond to Eq. (3) with h_0 and γ obtained from the data of $T = 500$ K only.

would expect more frequent nucleation and an earlier 2D-3D transition at higher growth rates. Note again that we have suppressed the ES effect explicitly which would also lead to more frequent mound formation at higher deposition rates [1, 2].

In our model, upward diffusion is the limiting effect and sets the characteristic time for the SK transition. The frequency of these processes strongly depends on the temperature and it has to be compared with the deposition rate. A high incoming flux will fill layers before particles can perform upward hops, and hence it will delay the 2D-3D transition, cf. Fig. 3. These considerations, together with the arguments of [25], suggest a functional dependence of h_{WL}^* on F and T of the form

$$h_{\text{WL}}^* = h_0 \left(\frac{F}{R_{\text{up}}} \right)^\gamma, \quad (3)$$

where R_{up} is the Arrhenius rate for upward hops. In our simulations we find a typical value which corresponds to an activation barrier $E_{\text{up}} \approx 1.0$ eV close to the transition. Using this value we have performed a non-linear fit according to Eq. (3) based on the data for $T = 500$ K. Inserting the obtained parameters $h_0 \approx 3.0$ ML and $\gamma \approx 0.2$ in Eq. (3) yields good agreement for $T = 480$ K, as well. Note that the value of γ is expected to vary with the material system in the range $0 < \gamma < 0.5$ according to [25]. Our assumption that upward diffusion is more important than nucleation of deposited adatoms on top of islands is further supported by the observation that multilayer islands tend to form on a WL of, say, two monolayers thickness even without particle deposition.

After the 2D-3D transition, islands grow by incorporating newly deposited material, but also by consumption of the surrounding WL. Very large islands are observed to split by means of upward diffusion events onto their top layer [26]. The migration of WL particles towards and onto the islands can also extend into the relaxation period after deposition ends. Eventually, a stationary WL thickness $h_{\text{WL}} \approx 1$ is observed in our example scenario with $U_{\text{as}} \approx 0.86$ eV. By increasing the strength of the adsorbate/substrate interaction we can achieve, e.g. $h_{\text{WL}} \approx 2$ for $U_{\text{as}} \approx 2.7$ eV, but it is difficult to stabilize a greater stationary WL thickness due to the short range nature of the LJ potential.

Finally, we discuss the properties of 3D islands emerging in SK growth. As an example, their base length b is measured as the number of particles in the island bottom layer. The results discussed in the following were obtained at the end of the relaxation period with $R_d = 0$. Whereas mean values do not change significantly, fluctuations are observed to decrease with time in this phase. We observe that, for fixed temperature, the island size decreases with increasing deposition flux, as observed in several

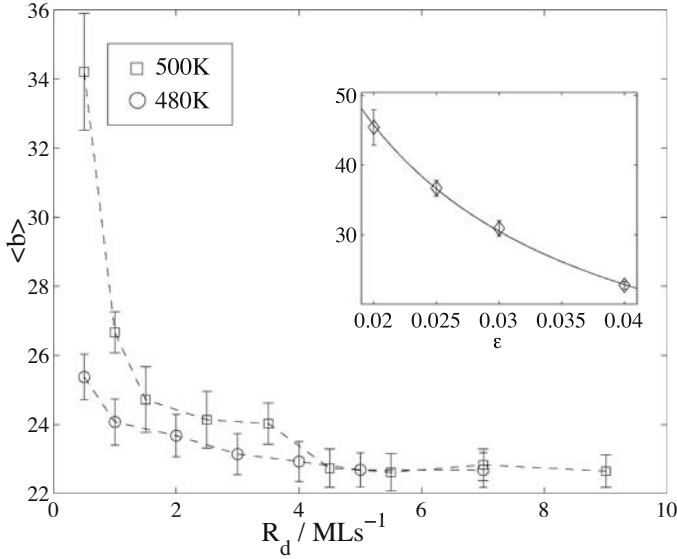


Figure 4. Average base size $\langle b \rangle$ of multilayer islands as a function of R_d at $T = 500$ K, together with standard error bars. The inset shows the result for $R_d = 4.5$ ML/s and different misfit parameters, and the solid line corresponds to $\langle b \rangle = 0.91/\epsilon$.

experimental studies [25]. However, $\langle b \rangle$ becomes constant and independent of T for large enough deposition rate, cf. Fig. 4. Corresponding behavior is found for the island density and their lateral distance, hinting at a considerable degree of spatial ordering [12]. The saturation behavior further demonstrates the importance of upward hops *versus* aggregation of deposited particles on islands. The latter process would yield a continuous increase of the island density with R_d . We find a narrower distribution of island sizes with increasing R_d [12]. The island size distribution in the saturation regime will be studied in greater detail in forthcoming investigations.

We conclude our discussion by noting that, in the saturation regime of high growth rates, the typical island size follows a simple power law: $\langle b \rangle \propto 1/\epsilon$, cf. Fig. 4. Very far from equilibrium, the only relevant length scale in the system is the relative periodicity $1/\epsilon$ of the adsorbate and substrate lattices. This characteristic length was already found to dominate in the formation of misfit dislocations for larger values of ϵ [11].

5. Conclusion and Outlook

Despite its conceptual simplicity and the small number of free parameters, our model reproduces various phenomena of heteroepitaxial

growth. We believe that, with a proper choice of interaction parameters, our model should be capable of reproducing all three prototype growth modes: layer-by-layer growth (for very small misfits), Volmer–Weber (for $U_{as} < U_a$) and, as demonstrated here, the Stranski–Krastanov mode. Our work provides a fairly detailed and plausible picture of the latter. The key features are:

- (i) The strong adsorbate/substrate interaction favors the WL formation and results in a relatively slow diffusion of adatoms on the substrate. Diffusion on the WL is significantly faster and the corresponding barrier decreases with the WL thickness.
- (ii) Strain relaxation leads to a pronounced bias towards the island center on top of finite mono- and multilayer islands located on the WL. In addition, diffusion is slower on top of the partially relaxed islands than on the WL, in our model.

Whereas (a) favors the formation and persistence of the WL, (b) clearly destabilizes layer by layer growth. We find that the microscopic process which triggers the transition is upward diffusion of adatoms from the WL and at island edges. The corresponding barriers decrease with the WL thickness analogous to (a). As a result of the competing effects, the 2D-3D transition occurs at a critical thickness which depends upon T and R_d as suggested in Eq. (3).

We hypothesize that strain effects induce spatial modulations with a characteristic length scale ϵ^{-1} and thus control the island size far from equilibrium. We find a corresponding saturation regime for large enough deposition fluxes. The precise mechanism of the size selection will be the subject of a forthcoming project. A related open question concerns the crossover from kinetically controlled island sizes to the equilibrium behavior which should be achieved in the limit of very small R_d . Several arguments [2, 17] suggest that the typical island size close to equilibrium should be of order ϵ^{-2} .

Further investigations will concern the effects of intermixing and segregation which have been excluded from our model, so far. To this end we will consider the co-deposition of both species and allow for exchange diffusion at the substrate/adsorbate interface. In order to test the potential universality of our results, we will introduce different types of interaction potentials in our model. Ultimately, we plan to extend our model to the relevant case of $2+1$ dimensions and to more realistic empirical potentials for semiconductor materials, e.g. [21].

Acknowledgments

F. Much has been supported by the Deutsche Forschungsgemeinschaft. M. Biehl thanks the organizers and all participants of the NATO ARW on Quantum Dots: Fundamentals, Applications, and Frontiers for pleasant and useful discussions. We also thank B. Voigtländer for communicating the results of [23] prior to publication.

References

- [1] A. Pimpinelli and J. Villain, *Physics of Crystal Growth*, Cambridge University Press (1998).
- [2] P. Politi, G. Grenet, A. Marty, A. Ponchet, and J. Villain. Instabilities in crystal growth by atomic or molecular beams. *Phys. Rep.* **324**: 271–404, 2000.
- [3] W. K. Liu and M. B. Santos (eds.), *Thin Films: Heteroepitaxial systems*, World Scientific (1999).
- [4] C. Heyn. Critical coverage for strain-induced formation of InAs quantum dots, *Phys. Rev. B* **64**: art. no. 165306, 2001.
- [5] A. G. Cullis, D. J. Norris, T. Walther, M. A. Migliorato, and M. Hopkinson. Stranski–Krastanov transition and epitaxial island growth. *Phys. Rev. B* **66**: art. no. 081305(R), 2002.
- [6] L. Chkoda, M. Schneider, V. Shklover, L. Kilian, M. Sokolowski, C. Heske, and E. Umbach. Temperature-dependent morphology and structure of ordered 3,4,9,10-perylene-tetracarboxylicacid-dianhydride (PCTDA) thin films on Ag(111). *Chem. Phys. Lett.* **371**: 548–552, 2003.
- [7] A. C. Schindler, *Theoretical aspects of growth in one and two dimensional strained crystal surfaces*, dissertation, Duisburg (1999).
- [8] A. C. Schindler, D. D. Vvedensky, M. F. Gyure, G. D. Simms, R. E. Caflisch and C. Connell. Atomistic and continuum elastic effects in heteroepitaxial systems, in: *Atomistic Aspects of Epitaxial Growth*, M. Kotrla, N. I. Papanicolaou, D. D. Vvedensky, L. T. Wille (eds.), Kluwer (2002), pp. 337–353.
- [9] H. Spjut and D. A. Faux. Computer simulation of strain-induced diffusion enhancement of Si adatoms on the Si(001) surface. *Surf. Sci.* **306**: 233–239, 1994.
- [10] J. Kew, M. R. Wilby, and D. D. Vvedensky. Continuous-space Monte Carlo simulations of epitaxial growth. *J. Crystal Growth* **127**: 508–512, 1993.
- [11] F. Much, M. Ahr, M. Biehl, and W. Kinzel. Kinetic Monte Carlo simulations of dislocations in heteroepitaxial growth. *Europhys. Lett.* **56**: 791–796, 2001.
- [12] M. Biehl and F. Much. Simulation of wetting-layer and island formation in heteroepitaxial growth. *Europhys. Lett.* **63**: 14–20, 2003.
- [13] M. E. J. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics*, Oxford University Press (1999).
- [14] A. Madhukar. Far from equilibrium vapor phase growth of lattice matched III-V compound semiconductor interfaces: some basic concepts and Monte-Carlo computer simulations. *Surf. Sci.* **132**: 344–374, 1983.

- [15] S. V. Ghaisas and A. Madhukar, *Role of surface molecular reactions in influencing the growth mechanism and the nature of nonequilibrium surfaces: a Monte Carlo study of molecular-beam epitaxy*, *Phys. Rev. Lett.* **56** (1986) 1066.
- [16] K. E. Khor and S. Das Sarma. Quantum dot self-assembly in growth of strained-layer thin films: a kinetic Monte Carlo study. *Phys. Rev. B* **62**: 16657–16664, 2000.
- [17] M. Meixner, E. Schöll, V. A. Shchukin, and D. Bimberg. Self-assembled quantum dots: crossover from kinetically controlled to thermodynamically limited growth *Phys. Rev. Lett.* **87**: art. no. 236101, 2001.
- [18] F. Jensen, *Introduction to Computational Chemistry*, Wiley (1999).
- [19] L. Dong, J. Schnitker, R. W. Smith, D. J. Sroloviy. Stress relaxation and misfit dislocation nucleation in the growth of misfitting films: A molecular dynamics simulation. *J. Appl. Phys.* **83**: 217–227, 1997.
- [20] A. F. Voter, F. Montalenti, and T. C. Germann. Extending the time scale in atomistic simulations of materials. *Ann. Rev. Mater. Res.* **32**: 321–346, 2002.
- [21] M. Schroeder and D. E. Wolf. Diffusion on strained surfaces. *Surf. Sci.* **375**: 129–140, 1997.
- [22] V. Cherepanov and B. Voigtländer. Influence of strain on diffusion at Ge(111) surfaces. *Appl. Phys. Lett.* **81**: 4745–4747, 2002.
- [23] V. Cherepanov and B. Voigtländer. Influence of material, surface reconstruction, and strain on diffusion at the Ge(111) surface. *Phys. Rev. B* **69**: art. no. 125331, 2004
- [24] D. Leonard, K. Pond, and P. M. Petroff. Critical layer thickness for self-assembled InAs islands on GaAs. *Phys. Rev. B* **50**: 11687–11692, 1994.
- [25] J. Johansson and W. Seifert. Kinetics of self-assembled island formation: Part I: Island density. *J. Crystal Growth* **234**: 132–138, 2002; Part II: Island size, **234**: 139–144, 2002.
- [26] mpeg movies and other illustrations are available from the web pages [{~much}](http://www.physik.uni-wuerzburg.de/~biehl).
- [27] J. Tersoff. New empirical approach for the structure and energy of covalent systems. *Phys. Rev. B* **37**: 6991–7000, 1988.

TEMPERATURE REGIMES OF STRAIN-INDUCED InAs QUANTUM DOT FORMATION

Christian Heyn and Arne Bolz

Institut für Angewandte Physik, Jungiusstr. 11, 20355 Hamburg, Germany

heyne@physnet.uni-hamburg.de

Abstract We study the mechanisms of strain-induced InAs quantum dot (QD) formation using theoretical and experimental techniques with focus on the influence of the growth parameters such as temperature and flux. The QDs are grown using solid-source molecular-beam epitaxy on GaAs and AlAs substrates and investigated with *in situ* electron diffraction, x-ray diffraction techniques, and atomic force microscopy. The experimental data of the critical time up to quantum dot formation and of the QD structural properties are modelled in terms of a kinetic rate-equations-based growth model. We distinguish three main temperature regimes. At low temperatures ($T \leq 420^\circ\text{C}$), QD formation is assumed to be mainly controlled by kinetic migration of adatoms on the surface. At higher temperatures, the additional process of intermixing of the QDs with substrate material is observed, which crucially modifies the QD formation process. Due to this intermixing, the strain energy inside the dots is significantly reduced and, accordingly, the driving force for QD formation. As a consequence, the critical time for QD formation increases. At $T > 520^\circ\text{C}$ for InAs on GaAs and $T > 540^\circ\text{C}$ for InAs on AlAs, desorption of In from the QDs becomes important and yields a further delay of QD formation.

Keywords: Quantum dots, molecular beam epitaxy, InAs, GaAs, AlAs, electron diffraction, x-ray diffraction, rate-equations, strain, temperature, kinetics, intermixing, desorption

1. Introduction

Self-assembling mechanisms that generate crystalline quantum-size structures without any lithographic steps are a fascinating aspect of solid-state physics. A prominent example are strain-induced InAs quantum dots (QDs) grown on GaAs or AlAs in the so-called Stranski-Krastanov (SK) mode [1, 2, 3, 4]. One main advantage of these co-

herently strained SK islands is their small size distribution [3] when compared with islands grown in Volmer-Weber mode. The size uniformity is very important for the fabrication of, e.g. quantum dot lasers, where a QD ensemble is used as the optically active medium. Since the electronic properties of these QDs are crucially determined by their shape and composition, a detailed understanding of the underlying self-assembling mechanism is essential for controlled fabrication of QDs with well-defined properties.

The driving force for the self-assembling mechanism is compressive strain inside the deposited film which is due to the lattice mismatch between substrate and deposit. During SK growth, first a thin wetting layer (WL) is formed. At a critical coverage θ_C , the accommodated strain energy initiates the transition from a flat two-dimensional (2D) surface morphology into three-dimensional (3D) QD-like islands. The formation, growth, and properties of the QDs can be controlled via the process parameters such as temperature T and beam fluxes. To get more insight in the underlying mechanisms, we focus in this work on the analysis of how the influence of growth parameters can be modelled in terms of a kinetic rate-equations-based growth model.

The influence of the growth temperature is quite complex and is the major topic of this work. A very important temperature dependent process is intermixing of the QDs with substrate material [5, 6, 7], which crucially modifies the QD formation process as well as their electronic properties. But the mechanism for intermixing is not yet clear. According to the experiments of Joyce *et al.* we assume negligible intermixing at $T \leq 420^\circ\text{C}$ [5], which allows us to study kinetics of QD formation without any influence of the unknown composition. At $T > 520^\circ\text{C}$, desorption of In comes into play and yields a further significant modification of the QD formation process [8]. In this work we distinguish three major temperature regimes that are denoted according to the prevailing process: the migration regime ($T \leq 420^\circ\text{C}$), the intermixing regime ($420^\circ\text{C} < T < 520^\circ\text{C}$), and the desorption regime ($T > 520^\circ\text{C}$).

2. Experimental Setup

The QDs are fabricated in a solid-source MBE system (Riber 32p) on (001) GaAs substrates. After thermal oxide desorption, a GaAs buffer is grown at $T = 600^\circ\text{C}$. For our samples with AlAs as starting surface, we grow an additional 20 nm thick AlAs layer at $T = 600^\circ\text{C}$. Afterwards the temperature is reduced for InAs deposition to $T = 415, \dots, 550^\circ\text{C}$. Quantum dots are grown with growth rates of 0.04 monolayers/s (ML/s) and 0.1 ML/s.

The experimental methods applied in this work to study growth and structural properties of the quantum dots can be distinguished between *in situ* and *ex situ* methods. As a true *in situ* technique, reflection high-energy electron-diffraction (RHEED) allows direct observation of the dynamic behavior during QD formation. However, the real space representation of the reciprocal space diffraction pattern is not completely clear at present. This is demonstrated, e.g. by the so-called chevrons attached to the transmission diffraction spots of InAs QDs in the $[\bar{1}10]$ azimuth, which origin is still debated [9]. We use a 12 keV RHEED system with a CCD camera and a PC for image processing [10].

A typical RHEED pattern of a $c(4\times 4)$ reconstructed GaAs surface at $T=500^\circ\text{C}$ is shown in Fig. 1(a). Immediately after InAs deposi-

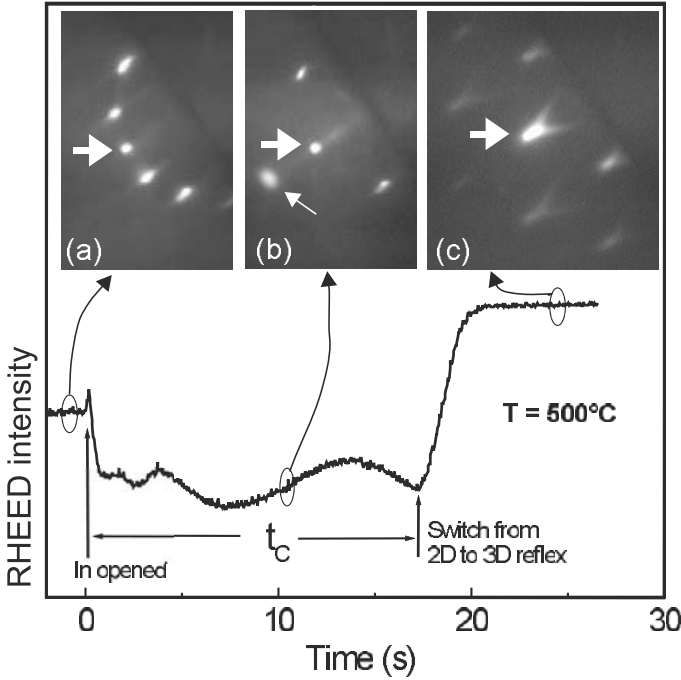


Figure 1. Time evolution of the intensity of 2D and 3D growth related reflexes together with corresponding RHEED patterns of GaAs and InAs surfaces. (a) flat GaAs surface at 500°C in the $[\bar{1}10]$ azimuth, (b) after deposition of 1.0 ML InAs, the small arrow marks a spot caused by thermal radiation of the In cell which is no RHEED spot, (c) transmission diffraction and appearance of chevrons after deposition of 2.0 ML InAs. The large arrows in (a) ... (c) indicate the respective 2D and 3D reflexes used for the measurement of the time dependent intensity. The 3D reflex appears at a critical time t_c .

tion is started, the surface reconstruction changes [Fig. 1(b)]. As for the GaAs buffer surface, the RHEED pattern at low InAs coverages is caused by diffraction from flat two-dimensional surface features. When the coverage θ exceeds a critical value θ_C , the RHEED pattern changes qualitatively. Now the peaks represent the reciprocal InAs lattice and originate from transmission diffraction through three-dimensional quantum dot-like islands [Fig. 1(c)]. We note that after transition into bulk-like reflection typical intensity tails (chevrons) [11] are attached to the RHEED spots that are directly associated to dot formation.

The critical time t_C for the 2D-to-3D transition is precisely determined from the intensity evolution of the brightest part of the RHEED pattern [10]. The location of the intensity maximum is automatically determined by a specially designed tracking software during data recording. Before t_C the intensity maximum is centered in the 2D related RHEED feature, while at larger time the maximum is found in a new developed spot that is close to the former one and related to the 3D morphology. An example for the intensity evolution of a 2D spot for $t < t_C$ and of a 3D spot for $t > t_C$ is shown in Fig. 1. The corresponding RHEED spots, marked by arrows, are shown in Figs. 1(a–c). We would like to note that t_C so determined represents the instant when the RHEED intensity from the 2D surface features is equal to that of the 3D morphology. Obviously, this instant is close to but not identical to the first occurrence of 3D features on the surface. Therefore, there might be a slight difference between values of the critical coverage taken by RHEED and by microscopic techniques.

The *ex situ* measurements are done after growth with atomic force microscopy (AFM) and height-resolved grazing incidence x-ray diffraction (GIXRD) at the Hamburg synchrotron radiation laboratory (HASYLAB). With the AFM we measure the density and height of the QDs and estimate their diameter. The GIXRD data are analyzed according to the method of Kegel *et al.* [6] in order to determine the Ga and Al content in the different layers of the QDs. To avoid oxidation of the highly reactive Al-compounds, the QD samples grown on AlAs substrates are transferred via an UHV-chamber to the synchrotron beam-line.

3. Growth model for strain-induced 3D islands

Many approaches to model the SK growth mechanisms presented in the literature mainly employ equilibrium arguments [12, 13, 14, 15] or kinetics of adatoms on the surface [16, 17, 18, 19, 20, 21, 22]. The mean-field growth model presented in this paper represents a kinetic approach and is an extension of our earlier rate-equation-based model

[20]. One major improvement in comparison to the older model is the consideration of layer-dependent strain and composition. This allows us to calculate, e.g. the height-resolved Ga content in the QDs as a function of the growth conditions.

We assume that after deposition has started the wetting layer (WL) is formed directly on top of the GaAs or AlAs substrate. Characteristic for this layer is a chemical attraction to the substrate which hinders upward migration of In atoms from the WL into higher levels. Once the wetting layer coverage θ_W is large enough, atoms will impinge in the active layer on top of the WL and nucleate there one monolayer (ML) high 2D islands with rate R_N . These islands grow laterally by attachment of diffusing adatoms with rate R_A and vertically by both direct hits from the vapor beam as well as by atoms hopping from island edges on top of the islands with rate R_T . As long as the islands are small, the influence of the strain is negligible and $R_T \simeq 0$. During this 2D growth regime, the behavior is close to layer-by-layer growth as, e.g. during GaAs homoepitaxy. When the island diameter becomes larger, the strain induced by the lattice mismatch between substrate and deposit reduces the lateral binding energy E_N between adatoms and island edges. Regarding the mass transport on top of the islands, the activity of atoms at island perimeter sites is the essential parameter. Since E_N is assumed to be equivalent to the energy barrier for upward migration of atoms from island edges to the top of the islands, this effect enhances R_T . This results in a strong increase of the island height and thus in their transformation into three-dimensional quantum dot-like islands. In this picture, the central quantity for transformation of the initially 2D growth islands into 3D QDs is the rate R_T at which atoms from the edge of the island base layer migrate on top of the island.

3.1 Strain energy

We assume that QDs are truncated pyramids with an angle of 26° between the side-facets and the substrate, as reported in Ref. [23]. An important quantity is the number s of atoms in the QD base layer, since this value controls the capture of diffusing adatoms by the island as well as the upward migration rate R_T . For InGaAs QDs, one gets $R_T = R_{T,\text{In}} + R_{T,\text{Ga}}$. We apply an Arrhenius-type law to calculate these rates as $R_{T,i} = \nu_0 x_i \sqrt{s} \exp(-E_N/k_B T)$, with $i = \text{In}$ or Ga , and x_i is the content of species i in the QD base layer. The lateral binding energy $E_N = E_0 - \kappa_a E_{\text{strain}}$ between atoms and the island edge is reduced by the strain energy E_{strain} in the QD base layer. E_0 and κ_a are constants, $\nu_0 = (2k_B T)/h$ is taken as vibrational frequency [24], T is the temperature,

k_B is Boltzmann's constant, and h is Planck's constant. We assume identical values of E_N for In and Ga. In order to specify E_{strain} of a QD, we apply a simple spring approximation, where the spring energy is proportional to the square of the compressed length.

For the topmost compressed layer $l = l_{\text{top}}$ of the QD this approach yields a strain energy $E_{\text{top}} = \kappa_b s_{\text{top}} \delta_{\text{top}}^2$, with constant κ_b , the number s_l of atoms, the lattice mismatch $\delta_l = (a_l - a_{l-1})/a_{l-1}$, and the In-content $x_{\text{In},l}$ dependent lattice constant $a_l = x_{\text{In},l}(a_{\text{InAs}} - a_{\text{GaAs}})$ of the top layer. The strain energy of the layer directly below the top layer has two components: the strain energy of the layer itself plus the additional strain energy from the layer above $E_{\text{top-1}} = \kappa_b s_{\text{top-1}} \delta_{\text{top-1}}^2 + c_U E_{\text{top}}$, where c_U describes the coupling of the strain energy between different layers. For the strain energy in the base layer $l=1$ one gets: $E_{\text{strain}} = \kappa_b s_1 \delta_1^2 + \kappa_b \sum_{l>1} c_U^{l-1} s_l \delta_l^2$. This allows the calculation of the binding energy

$$E_N = E_0 - c_B(s_1 \delta_1^2 + \sum_{l>1} c_U^{l-1} s_l \delta_l^2), \quad (1)$$

where $c_B = \kappa_a \kappa_b$. E_N can be used to calculate the upward migration rate R_T via the Arrhenius ansatz.

3.2 Rate Equations

In order to describe the dynamics of QD formation, we use a mean-field growth model based on kinetic rate equations. The central quantities calculated with the rate model are the densities of mobile Indium n_{In} and Gallium n_{Ga} adatoms (monomers), the island density n_I , the lateral QD size characterized by the average number s of atoms in the QD base layer, as well as the In and Ga coverage in the different layers of the QD. As a simplification in comparison to our earlier rate models [20, 25], the present model does not consider the island size distribution but calculates the average island size s , which tremendously reduces the computational time.

Surface activity is described in terms of thermally activated rates. The most important rates are the nucleation rate $R_{N,i,j} = n_i n_j \sigma_1 D$ at which islands are generated by collisions between diffusing adatoms of species i and j , the rate $R_{A,i} = n_i n_I \sigma_s D$ at which adatoms of species i attach to existing islands, the upwards migration rate $R_{T,i}$ of species i , and the desorption rate R_D of In atoms, where the σ are capture numbers [26], $D = \nu_0 \exp(-E_S/k_B T)$ represents the surface diffusion coefficient, and E_S the energy barrier for surface diffusion of a free adatom. Desorption of Ga or Al is assumed to be negligible. The material supply is provided by the beam fluxes F_{In} and F_{Ga} as well as by the rate R_X at which

substrate material intermixes into the QDs. R_X is assumed to increase only the WL coverage and the monomer density, whereas atoms from the beam fluxes impinge on the higher levels of the QDs, too.

The following set of coupled rate-equations describes the InAs on GaAs system, but InAs on AlAs can also be calculated by simply changing the indices from Ga to Al. The species resolved wetting layer coverage $\theta_W = \theta_{W,\text{In}} + \theta_{W,\text{Ga}}$ follows:

$$\dot{\theta}_{W,\text{In}} = F_{\text{In}}(1 - \theta_W) \quad (2)$$

$$\dot{\theta}_{W,\text{Ga}} = (F_{\text{Ga}} + R_X)(1 - \theta_W) \quad (3)$$

The monomer densities in the active layer on top of the WL are described by:

$$\dot{n}_{\text{In}} = F_{\text{In}}(\theta_W - \theta_B) - 2R_{\text{N,In,In}} - R_{\text{N,In,Ga}} - R_{\text{A,In}} \quad (4)$$

$$\dot{n}_{\text{Ga}} = F_{\text{Ga}}(\theta_W - \theta_B) + R_X\theta_W - 2R_{\text{N,Ga,Ga}} - R_{\text{N,In,Ga}} - R_{\text{A,Ga}} \quad (5)$$

Islands are nucleated in the active layer on top of the WL due to collisions between diffusing adatoms with density:

$$\dot{n}_I = R_{\text{N,In,In}} + R_{\text{N,Ga,Ga}} + R_{\text{N,In,Ga}} \quad (6)$$

The QD base layer ($l = 1$) coverage $\theta_B = \theta_{B,\text{In}} + \theta_{B,\text{Ga}}$ follows:

$$\dot{\theta}_{B,\text{In}} = 2R_{\text{N,In,In}} + R_{\text{N,In,Ga}} + R_{\text{A,In}} - R_{\text{T,In}} - R_D \quad (7)$$

$$\dot{\theta}_{B,\text{Ga}} = 2R_{\text{N,Ga,Ga}} + R_{\text{N,In,Ga}} + R_{\text{A,Ga}} - R_{\text{T,Ga}} \quad (8)$$

yielding the average QD size $s = \theta_B/n_I$.

The coverage of the layer ($l = 2$) directly above the base layer obeys:

$$\dot{\theta}_{2,\text{In}} = F_{\text{In}}(\theta_{l-1} - \theta_l) + R_{\text{T,In}} - \xi_{2,\text{In}} - R_D \quad (9)$$

$$\dot{\theta}_{2,\text{Ga}} = F_{\text{Ga}}(\theta_{l-1} - \theta_l) + R_{\text{T,Ga}} - \xi_{2,\text{Ga}} \quad (10)$$

For all higher layers ($l > 2$) the coverage follows:

$$\dot{\theta}_{l,\text{In}} = F_{\text{In}}(\theta_{l-1} - \theta_l) + \xi_{l-1,\text{In}} - \xi_{l,\text{In}} - R_D \quad (11)$$

$$\dot{\theta}_{l,\text{Ga}} = F_{\text{Ga}}(\theta_{l-1} - \theta_l) + \xi_{l-1,\text{Ga}} - \xi_{l,\text{Ga}} \quad (12)$$

$R_{\text{T,In}}$, $R_{\text{T,Ga}}$, ξ_{In} , and ξ_{Ga} describe the mass transfer between the different layers of the QD. Strain-induced migration with R_T is assumed from the island base layer $l = 1$ into the layer $l = 2$ directly above, only.

This, together with direct hits from the beam flux, increases the number of atoms in layer $l=2$. On the other hand, the number of sites in layer $l=2$ is limited due to the pyramid shape and controlled by the base layer size s . Excessive atoms in layer $l=2$ migrate one layer higher. The same mechanism yields an upward migration of atoms from all higher layers, too, and is described by ξ_{In} and ξ_{Ga} . The QD height is calculated as sum of the layer filling levels $h = \sum_l \text{atoms/sites}$.

The above set of coupled rate-equations is calculated self-consistently with respect to the surface diffusion length [26]. Central model parameters are the surface diffusion barrier E_S and the strain related parameters E_0 , c_U , c_B . Additional process parameters are T , F_{In} , and F_{Ga} . Parameters R_X , R_D related to intermixing and desorption will be discussed in the respective sections.

In order to compare calculation results of the critical time for the 2D-to-3D transition with our RHEED data, in the growth model, islands are regarded as 3D once their height exceeds 4 ML. This selection prevents counting islands that are not formed strain-driven but by direct hits from the vapor beam. Independent of the parameter set, the density of 3D islands increases abruptly at a certain time t . This is in good accordance with our RHEED observations which at the critical time t_C abruptly change from 2D to 3D patterns. In the following, we thus assume that the time at which the calculated 3D island density increases, can be identified with t_C .

4. The Migration Regime ($T \leq 420^\circ\text{C}$)

The most relevant quantity in context with surface morphology evolution is the coverage and regarding the 2D-to-3D transition especially the critical coverage θ_C . On the other hand, RHEED measurements yield the critical time t_C which is related to the coverage via

$$\theta_C = (F_{\text{In}} + F_{\text{Ga}})t_C + \theta_X(T) - \theta_D(T), \quad (13)$$

with the temperature dependent coverage θ_X due to intermixing of substrate material into the deposit, and the amount of material θ_D removed by desorption. Based on the experiments of Joyce *et al.* [5], we assume intermixing to be negligible at $T \leq 420^\circ\text{C}$: $R_X \simeq 0$, $\theta_X \simeq 0$. Desorption can be neglected as well [8] resulting in: $R_D \simeq 0$ and $\theta_D \simeq 0$. Therefore, in this regime the critical coverage can be directly calculated from experimental values of t_C according to Eq. (13), and the average QD composition is known from the beam flux ratio.

In a series of RHEED experiments at $T=420^\circ\text{C}$, the influence of the lattice mismatch $\delta = (1 - x_{\text{Ga}})(a_{\text{InAs}} - a_{\text{GaAs}})/a_{\text{GaAs}}$ is studied for $\text{In}_{1-x}\text{Ga}_x\text{As}$ deposition (Fig. 2). Variations of $F_{\text{Ga}} = 0.0, \dots, 0.23 \text{ ML/s}$

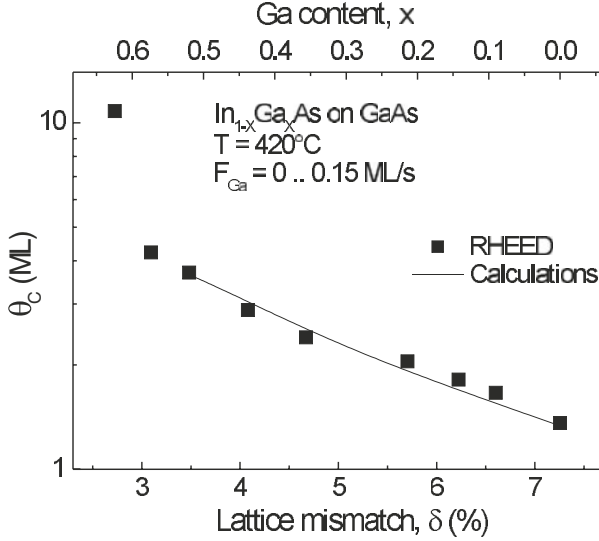


Figure 2. Dependence of measured and calculated θ_C on the Ga content x_{Ga} during $In_{1-x}Ga_xAs$ deposition. The In flux of 0.093 ML/s is kept constant, and only the Ga flux is varied. Assuming homogeneous conditions, the lattice mismatch is calculated from x_{Ga} .

at constant $F_{In} = 0.093$ ML/s yield $x_{Ga} = 0, \dots, 0.71$. In the RHEED data, up to a Ga content $x_{Ga} = 0.62$, θ_C increases significantly with decreasing δ , which is obviously due to the reduced strain inside the growing film. For larger values of x_{Ga} , no 3D growth related features are observable in the RHEED spots not even for very long deposition times. From this arises a minimal lattice mismatch of $\delta = 0.027$ to gain strain-induced 3D island formation.

In order to parameterize our growth model, we chose the surface diffusion energy barrier $E_S = 0.7$ eV and the lateral binding energy $E_0 = 2.2$ eV according to Ref. [20]. In case of pure InAs deposition, c_U is not important and the only free parameter $c_B = 0.0864$ is chosen for best agreement with the experimental value of $\theta_C = 1.36$ ML. In the next step $c_U = 0.792$ is chosen to fit experimental $\theta_C = 3.7$ ML at $x_{Ga} = 0.52$. As is demonstrated in Fig. 2, the calculated values of θ_C show very good quantitative reproduction of the experimental data. This paramount result establishes the applied spring approach as a suited approximation to describe the influence of strain during InGaAs quantum dot formation. Furthermore, the minor importance of intermixing and desorption in this regime is confirmed. At higher Ga contents, the RHEED measurements show intensity oscillations during the 2D growth regime that are associ-

ated with layer-by-layer growth. The wetting layer thickness determined from the number of oscillations lies beyond one monolayer and is thus not compatible with the geometrical assumptions of our model.

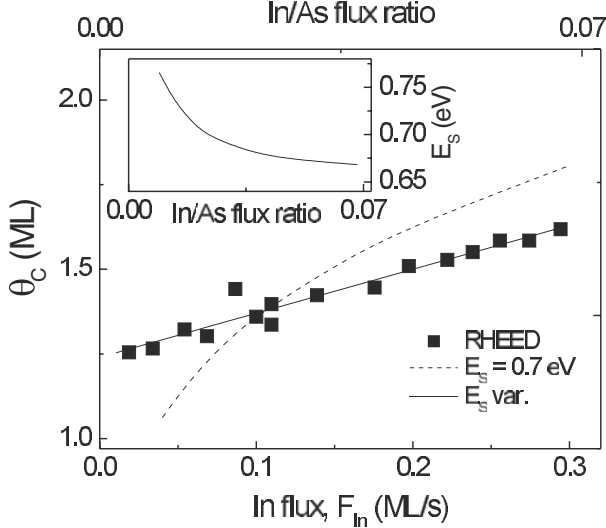


Figure 3. Critical coverage θ_C for 2D to 3D transition as function of In flux and corresponding In/As flux ratio. Plotted are RHEED measurements (symbols) together with calculation results with constant (dashed line) and varied (solid line) E_S . The inset shows the influence of In/As flux ratio on surface diffusion barrier E_S as is described in text.

RHEED measurements of θ_C as function of F_{In} are depicted in Fig. 3 together with calculation results. The RHEED data feature an almost linear increase of θ_C with F_{In} . This behavior can be explained in terms of the nucleation rate, where the calculated island density follows $n_I \sim (D/F)^{-1/3}$ [20]. Therefore, an increase of F_{In} yields a higher density of smaller islands. The critical lateral island size for formation of 3D islands is obtained later and, thus, θ_C is increased. The calculations show an increasing θ_C as well. However, if a constant E_S is assumed, the calculated values marked by the dashed line deviate significantly from experimental data, in particular at low F_{In} . The deviation is explained with the experimental procedure. In the experiments, only the In cell temperature is varied to set the flux, whereas the arsenic pressure is kept constant. This results in a change of the In/As flux ratio as well, which is not considered in the growth model. For GaAs homoepitaxy, a high Ga/In ratio is expected to enhance the surface diffusion coefficient D [27]. A similar effect in InAs epitaxy would counteract the

flux in the measurements via (D/F) and reduces its influence on θ_C . In order to give a quantitative characterization of this effect, we have adjusted the surface diffusion barrier in some calculations to reproduce the experimental θ_C vs. F_{In} data (solid line in Fig. 3). That way determined values of E_S are plotted in the inset of Fig. 3 as function of In/As flux ratio. We find a weak modification of the surface diffusion barrier from $E_S = 0.77$ eV at $F_{\text{In}}/F_{\text{As}} = 0.009$ down to $E_S = 0.67$ eV at $F_{\text{In}}/F_{\text{As}} = 0.068$.

5. The Intermixing Regime ($420^\circ < T < 520^\circ\text{C}$)

In this regime, desorption is still negligible [8], but intermixing becomes important. For deposition of pure InAs, Eq. (13) simplifies to $\theta_C = F_{\text{In}}t_C + \theta_X(T)$. Due to the *a priori* unknown θ_X , a direct calculation of θ_C from RHEED data is not possible. Furthermore, Eq. (13) demonstrates the complex correlation between temperature, intermixing, and the time until the 2D-to-3D transition. An increase of T enhances the surface kinetics which results in a decrease of θ_C . On the other hand, due to intermixing the lattice mismatch δ is reduced and θ_C is increased. As established by our RHEED experiments, the latter process dominates and t_C increases. As a third contribution, the additional coverage θ_X given by the intermixed substrate material reduces t_C as well.

Figure 4 shows RHEED measurements of t_C as function of the growth temperature at $F_{\text{In}} = 0.036$ ML/s. Up to about $T = 500^\circ\text{C}$ we observe an almost linear increase of t_C . At $T > 520^\circ\text{C}$, t_C increases rapidly, which we attribute to desorption (see next section). We find a very similar behavior of t_C for AlAs as the substrate material, but the onset of desorption is shifted towards higher temperatures of about 540°C . This indicates a stronger chemical binding in comparison to the InAs/GaAs system.

In a first step, we have calculated temperature dependent t_C without considering intermixing and desorption. Due to the small F_{In} in the experiments, $E_S = 0.77$ eV is chosen. The values of the strain-related model parameters are set as described in the previous section. We find a decrease of calculated t_C with increasing T (dotted line in Fig. 4) which reflects the Arrhenius-type nature of the growth model. However, the qualitative discrepancy between experiments and the model results indicates that the influence of intermixing is significantly stronger than the effect of temperature dependent surface migration.

In order to include intermixing, we incorporate a rate R_X at which substrate material mixes into the deposit. In particular, R_X increases

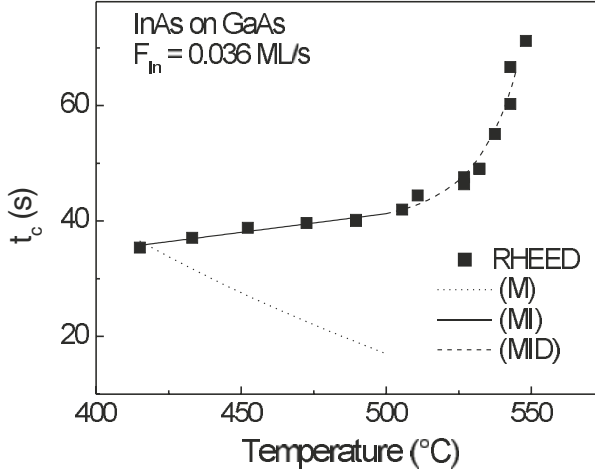


Figure 4. Influence of the growth temperature on t_c . Squares mark measured values and lines calculation results. (M) denotes calculations considering surface migration but no intermixing and desorption, (MI) those that include intermixing, and (MID) calculations incorporating both intermixing and desorption.

the wetting layer coverage θ_{Ga} and the monomer density n_{Ga} according to Eqs. (3) and (5). As an approximation, R_X is assumed to be independent of the deposition time. Now we use R_X as input-parameter for the model and find an increase of the calculated critical time t_c with R_X . This allows us to adjust $R_X(T)$ for $T \leq 500^\circ\text{C}$ in order to achieve agreement with the experimental t_c (Fig. 4). Fig. 5 shows an Arrhenius plot of R_X . From these data we calculate the average Ga content $x_{\text{Ga}} = R_X / (R_X + F_{\text{In}})$ in the QDs as plotted in Fig. 6. In a second series of experiments and calculations, x_{Ga} was determined for higher In flux $F_{\text{In}} = 0.1 \text{ ML/s}$ as well. As is visible in Fig. 6, both graphs show a similar behavior with $x_{\text{Ga}} = 0.45, \dots, 0.5$ at $T = 500^\circ\text{C}$.

At $T > 510^\circ\text{C}$, the above analysis becomes unreliable due to the beginning influence of desorption. To extrapolate values of x_{Ga} even for higher temperatures, we apply an empirical fit $x_{\text{Ga}} \approx 0.0183(T - 420)^{0.75}$. The fit represents a good approximation of the Ga content calculated with the model (dashed line in Fig. 6) as well as of $R_X = x_{\text{Ga}} F_{\text{In}} / (1 - x_{\text{Ga}})$ (dashed line in Fig. 5). These extrapolated values of R_X will be used as input-parameters for calculations at $T > 500^\circ\text{C}$ described in the next section.

In addition we have performed height-resolved grazing-incidence x-ray diffraction (GIXRD) experiments using the method introduced by Kegel *et al.* [6]. Figure 7 establishes that for both material systems InAs on

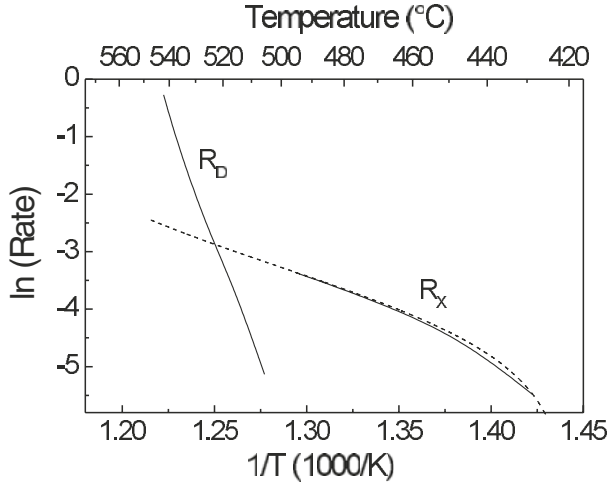


Figure 5. Arrhenius plot of the intermixing rate R_X and the desorption rate R_D (solid lines) and a fit of R_X , as described in the text (dashed line).

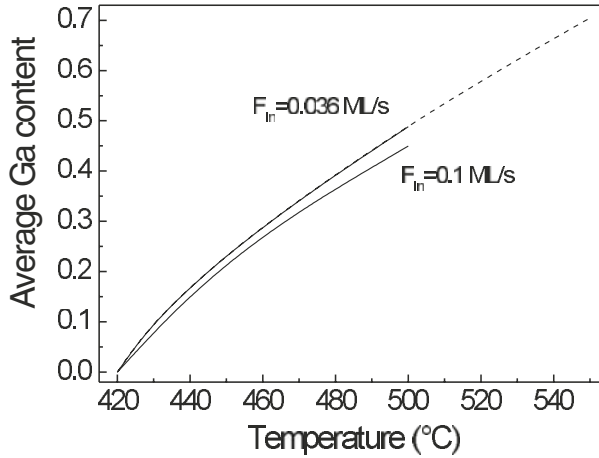


Figure 6. Average Ga content in the QDs as function of temperature for $F_{\text{In}} = 0.036$ and 0.1 ML/s (solid line) and a fit of x_{Ga} , as described in the text (dashed line).

GaAs as well as InAs on AlAs a significant amount of substrate material is incorporated especially in the lower layers of the quantum dots. These results demonstrate the crucial significance of intermixing and the related modification of the strain status in the QDs. Calculations reveal a qualitatively similar layer-dependence of the Ga content, but with

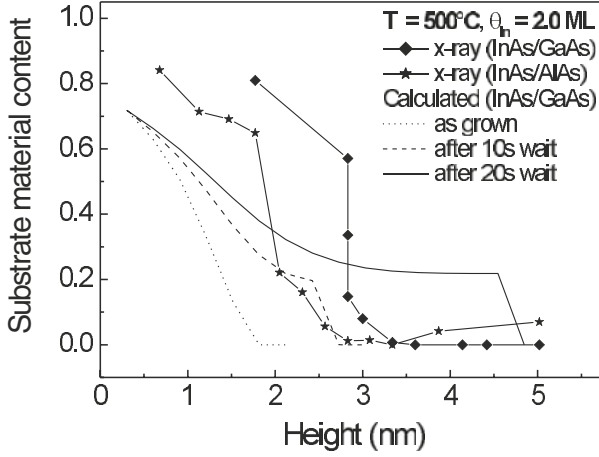


Figure 7. Height-resolved substrate material concentration in the QDs measured with GIXRD and calculated with our growth model.

slightly smaller values than in the GIXRD data. Further calculations are performed for 10 s and 20 s growth interruption ($F_{\text{In}} = 0$), which represents the fact that after MBE growth the QD samples need a certain time to cool-down before x-ray measurements. We find a significant change of the QD shape and in particular an increase of their height establishing the non-equilibrium nature of the QD formation process [20].

6. The Desorption Regime ($T > 520^\circ\text{C}$)

The desorption regime is characterized by an almost abrupt increase of t_C (Fig. 4) that cannot be explained with intermixing. We assume that only In atoms desorb at the temperatures studied here, whereas Ga atoms stick on the surface [8]. Desorption of In results in two effects: first the QD volume shrinks and second the QDs become more Ga rich. The latter effect yields a strain reduction with consequences as discussed above.

We incorporate desorption in terms of a rate R_D , at which In atoms are removed from the outermost layers of the QDs [Eqs. (7), (9), and (11)]. Since, at these high temperatures, the intermixing rate cannot be determined by the method described in the previous section, we use the extrapolated values of R_X in Fig. 5 for the calculations. Similar to the procedure in the intermixing regime, we adjust R_D as input-parameter for best agreement between calculated and experimental t_C

(Fig. 4). In Fig. 5 values of R_D so determined are plotted. From the nearly linear behavior of $\ln(R_D)$ vs. $1/T$ one might expect a simple Arrhenius-type desorption process. We note, that in an earlier study [8] of InAs QD desorption after growth cessation, we have observed layer-by-layer desorption starting from the top of the QDs, a stabilization of the QDs by an As flux, and evidence for a precursor state of the incoming As molecules. All these findings indicate a rather complex desorption mechanism.

7. Conclusions

The mechanisms of strain-induced InAs quantum dot formation are very intricate as they reflect the interaction between several basic processes on the surfaces. In this work the central goal is to find out, which kind of processes are important to explain the relevant features of our experiments. Among these processes the most important are surface diffusion, nucleation of initially 2D islands, lateral island growth by attachment of mobile adatoms, vertical island growth by migration of atoms from the edge on islands top, intermixing with substrate material, and desorption. Obviously, a complete picture of self-assembled QD formation would require the detailed knowledge of all these processes, their rates, energy barriers, and dependencies. We introduce a spring approximation for the strain energy inside the QDs as function of QD size and height-resolved composition. This approach allows us to calculate the strain-driven upward migration rate that is the central quantity for the transformation of the initially 2D islands into 3D QDs. Our growth model employing the spring approach quantitatively reproduces the influence of lattice mismatch on the experimental critical time t_C at low temperatures, where intermixing with substrate material and desorption are negligible. We find that the increase of t_C at higher temperatures can be explained by intermixing and desorption of In. To determine the respective rates, measurements and calculation results are compared. The quantitative agreement between model results and the experimental temperature dependence of t_C demonstrates that our growth model includes all processes relevant for strain-induced QD formation.

Acknowledgments

The authors would like to thank S. Mendach, R. L. Johnson, and W. Hansen for very valuable discussions, T. Maltezopoulos for AFM measurements, and the Deutsche Forschungsgemeinschaft for support under SFB 508 and GrK "Nanostrukturierte Festkörper".

References

- [1] D. Leonard, M. Krishnamurthy, S. Fafard, J.L. Merz, and P.M. Petroff. Molecular-beam epitaxy growth of quantum dots from strained coherent uniform islands of InGaAs on GaAs. *J. Vac. Sci. Technol. B* **12**: 1063–1066 1994.
- [2] A. Madhukar, Q. Xie, P. Chen, and A. Konkar. Nature of strained InAs three-dimensional island formation and distribution on GaAs(100). *Appl. Phys. Lett.* **64**: 2727–2729, 1994.
- [3] J. M. Moison, F. Houzay, F. Barthe, L. Lepronce, E. Andre, and O. Vatel. Self-organized growth of regular nanometer-scale InAs dots on GaAs. *Appl. Phys. Lett* **64**: 196–198, 1994.
- [4] D. Bimberg, M. Grundmann, and N. N. Ledentsov. *Quantum Dot Heterostructures*. John Wiley, 1999.
- [5] P. B. Joyce, T. J. Krzyzewski, G. R. Bell, B. A. Joyce, and T. S. Jones. Composition of InAs quantum dots on GaAs(001): Direct evidence for (InGa)As alloying. *Phys. Rev. B* **58**: R15981–R15984, 1998.
- [6] I. Kegel, T. H. Metzger, A. Lorke, J. Peisl, J. Stangl, G. Bauer, J. M. Garcia, and P. M. Petroff. Nanometer-scale resolution of strain and interdiffusion in self-assembled InAs/GaAs quantum dots. *Phys. Rev. Lett.* **85**: 1694–1697, 2000.
- [7] K. Zhang, Ch. Heyn, W. Hansen, Th. Schmidt, and J. Falta, Strain status of self-assembled InAs quantum dots. *Appl. Phys. Lett.* **77**: 1295–1297, 2000.
- [8] Ch. Heyn Stability of InAs quantum dots. *Phys. Rev. B* **66**: art. no. 075307, 2002.
- [9] D. W. Pashley, J. H. Neave, B. A. Joyce. A model for the appearance of chevrons on RHEED patterns from InAs quantum dots. *Surf. Sci.* **476**: 35–42, 2001.
- [10] Ch. Heyn, D. Endler, K. Zhang, and W. Hansen. Formation and dissolution of InAs quantum dots on GaAs. *J. Crystal Growth* **210**: 421–428, 2000).
- [11] H. Lee, R. Lowe-Webb, W. D. Yang, and P. C. Sercel, Determination of the shape of self-organized InAs/GaAs quantum dots by reflection high energy electron diffraction. *Appl. Phys. Lett* **72**: 812–814, 1998.
- [12] J. Tersoff and F. K. LeGoues. Competing relaxation mechanisms in strained layers. *Phys. Rev. Lett.* **72**: 3570–3573, 1994.
- [13] C. Priester and M. Lannoo. Origin of self-assembled quantum dots in highly mismatched heteroepitaxy. *Phys. Rev. Lett.* **75**: 93–96, 1995.
- [14] V. A. Shchukin, N. N. Ledentsov, P. S. Kop'ev, and D. Bimberg. Spontaneous ordering of arrays of coherent strained islands. *Phys. Rev. Lett.* **75**: 2968–2971, 1995.
- [15] N. Moll, M. Scheffler, and E. Pehlke. Influence of surface stress on the equilibrium shape of strained quantum dots. *Phys. Rev. B* **58**: 4566–4571, 1998.
- [16] H. T. Dobbs, D. D. Vvedensky, A. Zangwill, J. Johansson, N. Carlson, and W. Seifert. Mean-field theory of quantum dot formation. *Phys. Rev. Lett.* **79**: 897–900, 1997.
- [17] B. A. Joyce, J. L. Sudijono, J. L. Belk, H. Yamaguchi, X. M. Zhang, H. T. Dobbs, A. Zangwill, D. D. Vvedensky, and T. S. Jones. A scanning tunneling microscopy – reflection high energy electron diffraction – rate equation study of

the molecular beam epitaxial growth of InAs on GaAs(001), (110) and (111)A-quantum dots and two-dimensional modes. *Jpn. J. Appl. Phys.* **36**: 4111–4117, 1997.

- [18] H. M. Koduvely and A. Zangwill. Epitaxial growth kinetics with interacting coherent islands. *Phys. Rev. B* **60**: R2204–R2207, 1999.
- [19] Ch. Heyn and C. Dumat. Formation and size evolution of self-assembled quantum dots. *J. Crystal Growth* **227/228**: 990–994, 2001.
- [20] Ch. Heyn. Critical coverage for strain-induced formation of InAs quantum dots. *Phys. Rev. B* **64**: art. no. 165306, 2001.
- [21] M. Meixner, E. Schöll, V. A. Shchukin, and D. Bimberg. Self-assembled quantum dots: Crossover from kinetically controlled to thermodynamically limited growth. *Phys. Rev. Lett.* **87**: art. no. 236101, 2001.
- [22] F. Much and M. Biehl. Simulation of wetting-layer and island formation in heteroepitaxial growth. *Europhys. Lett.* **63**: 14–20, 2003.
- [23] J. Marquez, L. Geelhaar, and K. Jacobi. Atomically resolved structure of InAs quantum dots. *Appl. Phys. Lett.* **78**: 2309–2311, 2001.
- [24] T. Shitara, D. D. Vvedensky, M. R. Wilby, J. Zhang, J. H. Neave, and B. A. Joyce. Morphological model of reflection high-energy electron-diffraction intensity oscillations during epitaxial growth on GaAs(001). *Appl. Phys. Lett.* **60**: 1504–1506, 1992.
- [25] Ch. Heyn and W. Hansen. Ga/In-intermixing and segregation during InAs quantum dot formation. *J. Crystal Growth* **251**: 140–144, 2003.
- [26] G. S. Bales and D. C. Chrzan. Dynamics of irreversible island growth during submonolayer epitaxy. *Phys. Rev. B* **50**: 6057–6067, 1994.
- [27] Ch. Heyn and M. Harsdorff. Simulation of GaAs growth and surface recovery with respect to gallium and arsenic surface kinetics. *Phys. Rev. B* **55**: 7034–7038, 1997; Ch. Heyn, T. Franke, and R. Anton. Correlation between island-formation kinetics, surface roughening, and RHEED oscillation damping during GaAs homoepitaxy. *Phys. Rev. B* **56**: 13483–13489, 1997.

KINETIC MODELLING OF STRAINED FILMS: EFFECTS OF WETTING AND FACETTING

Daniel Kandel and Helen R. Eisenberg

*Department of Physics of Complex Systems, Weizmann Institute of Science,
Rehovot 76100, Israel*

Abstract The existence of a wetting layer in strained films is not well understood, despite extensive studies of the stability of strained films. In this paper we show that the dependence of the reference state free energy on film thickness leads to a finite thickness wetting layer, which decreases with increasing lattice mismatch strain. We also show that anisotropic surface tension gives rise to a metastable enlarged wetting layer.

1. Introduction

The growth of epitaxially strained thin films, in which there is lattice mismatch between the substrate and the film, has attracted significant attention in recent years. Under suitable conditions, the strain in the flat film relaxes by the formation of dislocation-free islands on the film surface via surface diffusion. These coherent islands can self-organize to create periodic arrays which can be utilized to create quantum dot structures, to which this Workshop is devoted.

It has been observed experimentally that dislocation-free flat films of less than a certain thickness (the critical wetting layer) are stable to surface perturbations, while thicker films are unstable [1, 2, 3, 4, 5, 6, 7, 8, 9]. The thickness of the wetting layer is substance dependent and decreases with increasing lattice mismatch strain [5, 6, 7, 8, 9], $\varepsilon = (a_s - a_f)/a_f$, where a_s and a_f are the substrate and film lattice constants. Despite considerable efforts, the physics of the critical wetting layer is poorly understood. The present work addresses this important question.

We show that the variation of nonlinear elastic free energy with film thickness can give rise to a wetting layer, which decreases with increasing lattice mismatch strain. We explain how and at what depth a flat film becomes unstable to perturbations of varying amplitude for films with

both isotropic and anisotropic surface tension. We study the evolution of these small perturbations and observe island facetting. We show that anisotropic surface tension gives rise to a metastable enlarged wetting layer. The perturbation amplitude needed to destabilize this wetting layer decreases with increasing lattice mismatch. The results of this study appear in Ref. [10, 11].

2. Problem Formulation

We model the evolution of a thin film on a substrate using continuum theory. The lattice mismatch between the film and the substrate creates a strain in the film, ε . Both the substrate and the film are assumed to be elastically isotropic with the same elastic constants. The surface of the solid is at $y = h(x, t)$ and the film is in the $y > 0$ region with the film-substrate interface at $y = 0$. The system is modeled to be invariant in the z -direction, and all quantities are calculated for a section of unit width in that direction. This is consistent with plane strain where the solid extends infinitely in the z -direction and hence all strains in this direction vanish, i.e. $e_{xz} = e_{yz} = e_{zz} = 0$. We assume there is no material mixing between the substrate and the film.

We assume that surface diffusion is the dominant mass transport mechanism. Gradients in the chemical potential produce a drift of surface atoms with an average velocity v given by the Nernst-Einstein relation

$$v = -\frac{D_s}{k_B T} \frac{\partial \mu}{\partial s}, \quad (1)$$

where D_s is the surface diffusion coefficient, s is the arc length, T is the temperature, k_B is the Boltzmann constant and μ is the chemical potential at the surface, i.e. the increase in free energy when an atom is added to the solid surface at the point of interest. Taking the divergence of the surface current produced by the atom drift gives an expression for the height evolution [12],

$$\frac{\partial h}{\partial t} = \frac{D_s \eta \Omega}{k_B T} \frac{\partial}{\partial x} \frac{\partial \mu}{\partial s}, \quad (2)$$

where η is the number of atoms per unit area on the solid surface and Ω is the atomic volume.

In the continuum approximation $\mu = \Omega \frac{\delta F}{\delta h}$, where F is the free energy of the solid and $\delta F / \delta h$ is the functional derivative of F . The free energy is composed of elastic and surface terms:

$$F = F_{el} + \int dx \gamma \sqrt{1 + (\partial h / \partial x)^2}, \quad (3)$$

where γ is the surface tension and F_{el} is the elastic free energy including any elastic contributions to the surface tension.

The elastic free energy can be written in terms of the elastic free energy density, f_v , as $F_{el} = \int dx dy f_v$. f_v is expanded as a power series in the strain:

$$f_v = f_v^{(0)} + \sigma_{ij}^{(0)} e_{ij} + \frac{1}{2} c_{ijkl} e_{ij} e_{kl} + \dots, \quad (4)$$

where $f_v^{(0)}$ is the free energy density in a zero strain reference state, $\sigma_{ij}^{(0)}$ is the stress in the reference state and c_{ijkl} are the elastic coefficients of the material. Thus, $F_{el} = F_{el}^{(0)} + \delta F_{el}$, where $F_{el}^{(0)}$ is the reference state elastic free energy and the linear elasticity correction is

$$\delta F_{el} = \int dx \int_{-\infty}^{h(x)} dy \left[\sigma_{ij}^{(0)} e_{ij} + \frac{1}{2} c_{ijkl} e_{ij} e_{kl} \right]. \quad (5)$$

In linear elasticity theory, deformations are assumed to be small and so terms of third order and higher are neglected. The stress-strain relationship is given by $\sigma_{ij} = \partial f_v / \partial e_{ij}$, which under linear elasticity gives Hooke's law:

$$\sigma_{ij} = \sigma_{ij}^{(0)} + c_{ijkl} e_{kl}. \quad (6)$$

Our system has periodic boundary conditions in the x -direction and is infinite in the negative y -direction. We assume that the forces on the upper surface due to surface tension (as given by Marchenko and Parshin [13]) are negligible in comparison to the forces due to the mismatch stress. This assumption is fulfilled as long as $\gamma/R \ll M\varepsilon$, where R is the radius of curvature of the surface and M is the plane strain modulus. For typical values of γ , M and ε , this condition is satisfied when R is larger than the lattice constant. As typical surface features have length scales of the order of 100 nm this assumption is valid. Hence the boundary conditions are given by

$$\begin{aligned} \sigma_{ij} n_j &= 0 & \text{at} & & y &= h(x), \\ \sigma_{ij} &\rightarrow 0 & \text{when} & & y &\rightarrow -\infty. \end{aligned} \quad (7)$$

For each value of x , our reference state corresponds locally to a *flat* film of thickness $h(x)$ constrained to have the lateral lattice constant of the substrate; i.e.

$$F_{el}^{(0)} = \int dx \int_{-\infty}^{h(x)} dy f_v^{(0)}(h(x), y), \quad (8)$$

where $f_v^{(0)}(h(x), y)$ is the elastic free energy density of a flat film of thickness $h(x)$ with the substrate lateral lattice constant. We calculate the correction to the elastic free energy of the perturbed state, δF_{el} , using linear elasticity theory.

When looking at the stability of a strained flat film of thickness C , the obvious first choice for a reference state is that of a flat film of thickness C constrained to have the lateral lattice constant of the substrate. For later calculations we must fully define the reference state and hence need to know its stress $\sigma_{ij}^{(0)}(C, y)$ and free energy density $f_v^{(0)}(C, y)$. One simple approach to calculate these quantities would be to use linear elasticity with the unstressed film as a reference state. In linear elasticity a flat film of any thickness constrained to have the substrate lateral lattice constant and free to move in the y -direction is in equilibrium and has the elastic free energy density of an infinitely strained film. Hence, such a calculation does not predict any C or y dependence in $\sigma_{ij}^{(0)}$ and $f_v^{(0)}$, except for a step function at the film-substrate interface. For example, in the case of plane strain, where the mismatch strain is uniaxial (i.e. $e_{xz} = e_{yz} = e_{zz} = 0$, $e_{xy} = 0$, $e_{xx} = \varepsilon$), linear elasticity gives $\sigma_{ij}^{(0)} = M\varepsilon$ and $f_v^{(0)} = \frac{1}{2}M\varepsilon^2$, where M is the plane strain modulus. Therefore, variation of the elastic free energy and stress of a flat film with film height is a nonlinear phenomenon, and a model outside of linear elasticity theory must be used to calculate them. As will be shown in Sec. 3, small variations in the reference free energy density with film thickness are crucial in predicting wetting layer thickness, and a reference free energy density which has no variation with film thickness will lead to thin films that have no wetting layer.

The disadvantage of our choice of the reference state is that the dependence of h on x leads to lateral variations of the reference state. As a result, the reference stress does not satisfy the condition of mechanical equilibrium. However, the needed corrections vanish in the limit $a/\lambda \rightarrow 0$, where a is the length scale over which stress varies in the y -direction and λ is the lateral length of typical surface structures. This is because in this limit there are no lateral variations in the reference stress. As typical experimental islands have $\lambda \sim 100nm$, and as a is of the order of the lattice constant (see below), the corrections to the reference stress are small and have been ignored.

Though linear elasticity cannot be used to calculate properties associated with the reference state, it can still be used to find the correction to the elastic free energy of the perturbed state, δF_{el} . For convenience we work in terms of the reference elastic free energy per unit length in

the x -direction,

$$f_{el}^{(0)}(h(x)) \equiv \int_{-\infty}^{h(x)} dy f_v^{(0)}(h(x), y), \quad (9)$$

instead of the free energy per unit volume. With this definition and the linear elasticity expression for δF_{el} , $\delta F/\delta h$ takes the form (see [10, 11])

$$\frac{\delta F}{\delta h} = \tilde{\gamma}\kappa + \frac{df_{el}^{(0)}}{dh} + \left[\frac{1}{2} S_{ijkl} \sigma_{ij} \sigma_{kl} - \frac{1}{2} S_{ijkl} \sigma_{ij}^{(0)} \sigma_{kl}^{(0)} \right] \Big|_{y=h(x)}, \quad (10)$$

where κ is surface curvature, θ is the angle between the normal to the surface and the y -direction, and $\tilde{\gamma}(\theta) = \gamma(\theta) + \partial^2 \gamma / \partial \theta^2$ is the surface stiffness. We have used the inverted Hooke's law $e_{ij} = S_{ijkl}(\sigma_{kl} - \sigma_{kl}^{(0)})$, where S_{ijkl} are the compliance coefficients of the material.

As the above equation gives $\delta F/\delta h$ at the solid surface, all variables in the equation are also given at the surface. In particular $\sigma_{ij}^{(0)}(h, y=h)$ is taken as the stress at the surface of a flat solid of height $h(x)$ and hence must vanish when $h \leq 0$, since then the film is absent. $df_{el}^{(0)}(h)/dh$ is determined by calculating how the reference elastic free energy of the solid changes as monolayers are added to the solid surface. When $h \leq 0$, $df_{el}^{(0)}/dh = 0$, as the substrate is completely relaxed. In principle, Eq. (10) should also contain derivatives of γ with respect to h . However, we believe that the variation of surface tension with h away from a step dependence is due to elastic effects. Since we included all elastic contributions in the zero-strain elastic free energy, we modeled γ as a step function, taking the value of the substrate surface tension for $h \leq 0$ and the film surface tension for $h > 0$. Thus all partial derivatives of γ with respect to surface height vanish and were omitted from Eq. (10).

Equations (2) and (10) form a complete model of surface evolution. In order to solve this model, the chemical potential (given by Eq. (10)) for a given surface must be found, and this involves solving the linear elasticity problem. For an isotropic solid under plane strain, this can be reduced to finding the stress function, W , which satisfies the boundary conditions (7) and the biharmonic equation (see, e.g. Timoshenko [14] or Mikhlin [15]):

$$\Delta^2 W = \Delta(\Delta W) = \frac{\partial^4 W}{\partial x^4} + 2 \frac{\partial^4 W}{\partial x^2 \partial y^2} + \frac{\partial^4 W}{\partial y^4} = 0, \quad (11)$$

with

$$\sigma_{xx} = \frac{\partial^2 W}{\partial y^2}, \quad \sigma_{xy} = -\frac{\partial^2 W}{\partial x \partial y}, \quad \sigma_{yy} = \frac{\partial^2 W}{\partial x^2}. \quad (12)$$

In order to model the early evolution of faceted islands, and to study the effect of an anisotropic form of surface tension on the wetting layer, we used the cusped form of surface tension given by Bonzel and Preuss [16], which shows facetting in a free crystal:

$$\gamma(\theta) = \gamma_0 [1 + \beta |\sin(\pi\theta/(2\theta_0))|] , \quad (13)$$

where $\beta \approx 0.05$ and θ_0 is the angle of maximum γ . The value of γ_0 was taken as 1 J/m^2 in the substrate and about 75% of that in the film (as is the case for Ge/Si). This ensures a wetting layer of at least one monolayer. We considered a crystal which facets at $0^\circ, \pm 45^\circ$ and $\pm 90^\circ$ with $\theta_0 = \pi/8$. The cusp gives rise to $\tilde{\gamma} = \infty$. However, a slight miscut of the low-index surface along the z -direction leads to a rounding of the cusp, which can be described by

$$\gamma(\theta) = \gamma_0 \left(1 + \beta \sqrt{\sin^2(\frac{\pi}{2\theta_0}\theta) + G^{-2}} \right) , \quad (14)$$

where, for example, $G = 500$ corresponds to a miscut angle, $\Delta\theta \approx 0.1^\circ$. All the results mentioned in this paper relate to surfaces with a very small miscut angle in the z -direction.

3. Linear Stability Analysis

In this section we carry out a linear stability analysis of Eq. (2) against a sinusoidal perturbation of wavenumber k , similar to that carried out in Ref. [17, 18, 19] for an infinitely thick stressed film. We thus look for a height profile of the form, $h(x, t) = C + \delta(t) \sin kx$, which solves Eq. (2) to first order in δ . To calculate the linear elastic energy we find the solutions of (11), which satisfy the boundary conditions (7). $\sigma_{xy}^{(0)}$ vanishes because the film is hydrostatically strained, and $\sigma_{yy}^{(0)} = 0$ since in the reference state the force on the surface vanishes. Hence the only non-zero component of the reference stress is $\sigma_{xx}^{(0)}(h, y)$.

To first order in δ the stresses at the surface are (see Ref. [10, 11]):

$$\begin{aligned} \sigma_{xx} &= -2\delta k \sigma_{xx}^{(0)}(C, C) \sin(kx) + \sigma_{xx}^{(0)}(C, C) + \delta \sin(kx) d\sigma_{xx}^{(0)}/dh|_{h=C} \\ \sigma_{yy} &= 0 \\ \sigma_{xy} &= \delta k \sigma_{xx}^{(0)}(C, C) \cos(kx). \end{aligned}$$

Using the above stresses in Eq. (10), we obtain the expression

$$\frac{\delta F}{\delta h} = \delta \sin(kx) \left[\frac{d^2 f_{el}^{(0)}}{dh^2} - 2k \frac{(\sigma_{xx}^{(0)}(C, C))^2}{M} + \tilde{\gamma}_0 k^2 \right] + \frac{df_{el}^{(0)}}{dh}(C, C), \quad (15)$$

where $\tilde{\gamma}_0 = \tilde{\gamma}(\theta = 0)$. Combining the above equation with the evolution equation (2) gives the following equation for $\delta(t)$:

$$\frac{d\delta}{dt} = Kk^2 \left[-k^2\tilde{\gamma}_0 + 2k \frac{(\sigma_{xx}^{(0)}(C, C))^2}{M} - \frac{d^2 f_{el}^{(0)}}{dh^2} \right]_{h=C} \delta, \quad (16)$$

where $K = \frac{D_s \eta \Omega^2}{k_B T}$. Each term in the brackets in this equation has a simple physical significance. The first term is a surface tension term. Surface tension acts to reduce surface curvature, κ , and so this term is negative, thereby reducing the perturbation amplitude, and is linear in $\kappa \sim k^2$. The second term in this equation is a mismatch stress term. Regions of high stress have large chemical potential, and so atoms tend to detach from these regions and attach to regions of small chemical potential. In a mismatch stressed solid, valleys are regions of high stress, hence material moves from the valleys to the hills of a perturbed surface increasing perturbation amplitude. The contribution of this term is proportional to the density of valleys, which is linear in k . The last term is a reference state term. If $d^2 f_{el}^{(0)}/dh^2 > 0$ it costs more energy to add a monolayer to a flat film than to remove a monolayer, and hence it costs energy to perturb a film. Thus, positive $d^2 f_{el}^{(0)}/dh^2$ stabilizes a flat thin film, whereas negative $d^2 f_{el}^{(0)}/dh^2$ leads to an instability. Obviously this reference state term is present even if the film is flat and hence is independent of k .

Equation (16) implies that the flat film is stable at all perturbation wavelengths as long as

$$\frac{[\sigma_{xx}^{(0)}(C, C)]^4}{M^2} \leq \tilde{\gamma}_0 \left. \frac{d^2 f_{el}^{(0)}}{dh^2} \right|_{h=C}, \quad (17)$$

and the equality holds at the critical wetting layer thickness. $\tilde{\gamma}_0$ is positive if $\theta = 0$ is a surface seen in the equilibrium free crystal [20]. As mentioned earlier, $\tilde{\gamma}_0 \rightarrow \infty$ at a perfect facet and is large and positive on a surface with a small miscut, as is the case for most of the materials used in epitaxial films. Therefore, a linearly stable wetting layer of finite thickness can exist only if $d^2 f_{el}^{(0)}/dh^2 > 0$. Note that for the wetting layer to have a finite rather than an infinite thickness, $d^2 f_{el}^{(0)}/dh^2$ must decrease to a value less than the left-hand side of Eq. (17) as the thickness of the film increases. $\sigma_{xx}^{(0)}(C, C)$ depends linearly on the lattice mismatch ε , and hence the left-hand side of (17) is proportional to ε^4 , while the right-hand side of (17) is proportional to ε^2 due to the dependence of $f_{el}^{(0)}$ on lattice mismatch. Therefore, if $d^2 f_{el}^{(0)}/dh^2 > 0$, the

thickness of the wetting layer increases with decreasing lattice mismatch and diverges in the limit $\varepsilon \rightarrow 0$.

The maximum thickness of a flat film which is stable to infinitesimal perturbations is given by (17) when the equality holds. A film slightly thicker is unstable to perturbations of wavelength

$$\lambda = \frac{2\pi M\tilde{\gamma}_0}{\left[\sigma_{xx}^{(0)}(C, C)\right]^2}. \quad (18)$$

For films which are nearly perfect facets, these wavelengths are larger than the typical sample size and so practically such perturbations will never occur. However as will be explained in Sec. 5, the film can be nonlinearly unstable to smaller wavelength perturbations of a non-zero amplitude at physically reasonable wavelengths. Hence the inequality in (17) is only useful in predicting the stability of films with large miscut angles or above the roughening temperature. At small miscut angles the stability of the film to large perturbations will predict its maximum thickness. This issue is discussed in more detail in Sec. 5.

4. Nonlinear Elastic Free Energy of a Flat Film

The dependence of the nonlinear elastic free energy of a flat film $f_{el}^{(0)}(h)$ on film thickness, h , is vital in the determination of both the wetting layer thickness and thin film evolution. As a result of the sharp interface between the substrate and the film, we expect $\sigma_{ij}^{(0)}$ to behave as a step function of y with small corrections due to elastic relaxation. If we ignore these small corrections, the resulting free energy $f_{el}^{(0)}(h)$ is proportional to film thickness, and its second derivative vanishes. Hence according to Eq. (17), the thickness of the critical wetting layer vanishes. The correction due to elastic relaxation is therefore extremely important. As discussed earlier, this correction vanishes within linear elasticity theory.

The nonlinear elastic free energy of the reference state, $f_{el}^{(0)}(h)$, is calculated for a solid with a flat surface of height h . Hence in order to calculate $df_{el}^{(0)}/dh$, we determine $f_{el}^{(0)}$ for flat solids of heights $h + \frac{1}{2}\delta$ and $h - \frac{1}{2}\delta$, and use the estimate

$$\frac{df_{el}^{(0)}}{dh} = \frac{1}{\delta} [f_{el}^{(0)}(h + \tfrac{1}{2}\delta) - f_{el}^{(0)}(h - \tfrac{1}{2}\delta)]. \quad (19)$$

Ideally, first principles, substance specific calculations should be performed in order to evaluate $\sigma_{xx}^{(0)}(h, h)$ and $f_{el}^{(0)}(h)$, and we are currently

carrying out such calculations. However, the qualitative general behavior of $f_{el}^{(0)}(h)$ can be obtained from much simpler models. We used two-dimensional networks of balls and springs of varying lattice-type and spring constants. In these calculations δ is one monolayer, and $f_{el}^{(0)}(h)$ is calculated at film thicknesses of integer numbers of monolayers from no film up to 10 monolayers of film. Values of $f_{el}^{(0)}(h)$ for film heights of fractional monolayers are interpolated from the values calculated at integer monolayer heights.

We studied ball-and-spring models with natural spring length which has a step variation over the interface, and the balls are placed on a lattice with the substrate lattice constant. Thus balls in the film are connected by springs which have length larger than their natural length by a factor of $1 + e$, where e is the homogeneous strain in the film. The network was then allowed to relax, with the film free to move in the y -direction, and periodic boundary conditions being applied in the x -direction to ensure that the system boundaries in this direction were fixed to the natural substrate length.

We calculated the mismatch stress $\sigma_{xx}^{(0)}$ within the relaxed film and at the film surface. We also calculated the dependence of the mismatch surface stress $\sigma_{xx}^{(0)}(h, h)$ and the nonlinear elastic free energy $f_{el}^{(0)}(h)$ on film thickness, h . We carried out these calculations for various two dimensional networks of balls and springs with different spring constants. Simulations showed that the relaxation in the y -direction depended on film thickness and on the depth of the atom in the lattice but was independent of x .

Note that the springs in a simple square lattice can relax completely in the y -direction. Therefore, for such a lattice the relaxation is independent of spring depth within the film or film thickness. Only when diagonal bonds, such as those in a fcc lattice were present did the springs show depth dependent relaxation. The inability of the springs to completely relax due to the presence of diagonal bonds was a necessary condition for $f_{el}^{(0)}(h)$ to vary nonlinearly with film height. In such incompletely relaxed films the nonlinear dependence of $f_{el}^{(0)}$ on h arises from the elastic relaxation at the surface and its coupling to the relaxation at the film-substrate interface. A similar effect should occur in real systems due to surface reconstruction, for example.

A typical behavior of $df_{el}^{(0)}/dh$ is shown in Fig. 1, where it is seen that $f_{el}^{(0)}(h)$ indeed depends on the thickness h . Moreover, the model predicts that $d^2 f_{el}^{(0)}/dh^2 > 0$ and decreases with increasing film thickness, and therefore according to the inequality (17) and the discussion following it,

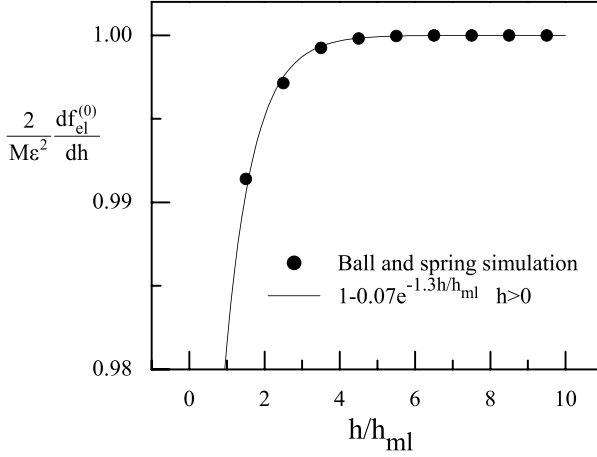


Figure 1. Variation with film thickness of the elastic free energy of a relaxed ball-and-spring system, $df_{el}^{(0)}/dh$, as a function of film thickness h . The free energy is normalized to the infinite film linear elastic energy density, $\frac{1}{2}M\epsilon^2$. h_{ml} is the thickness of one monolayer.

there should be a linearly stable wetting layer, whose thickness is finite and increases with decreasing lattice mismatch.

While the detailed dependence of $df_{el}^{(0)}(h)/dh$ on film thickness close to the substrate-film interface varied between different networks, it showed the same general behavior. In all systems $d^2f_{el}^{(0)}(h)/dh^2$ showed an exponential decay with a decay length of about a monolayer from the interface. The dimensionless quantity $(2/(M\epsilon^2))(df_{el}^{(0)}/dh)$ was independent of lattice mismatch sign and magnitude. $df_{el}^{(0)}/dh$ increased with film thickness and asymptotically approaches the elastic free energy density of an infinite film, $M\epsilon^2/2$, as expected.

The results of both Tersoff [21] and Wang [22, 23] show similar behavior suggesting that the fundamental form of $f_{el}^{(0)}(h)$ which gives rise to the wetting layer is due to universal elastic effects and that the specifics of the system only change the details of how thick that wetting layer will be. In previous works [24, 25, 26] on the physics of the wetting layer it was assumed that the reference state energy variation is a smooth function of h , mainly in order to avoid non-analyticities at the interface. In contrast, our reference state energy variation behaves as a step function of the surface height with a small but important correction.

For the calculations used later in this paper we used the function

$$\frac{df_{el}^{(0)}(h)}{dh} = \frac{M\varepsilon^2}{2} [1 - 0.05 \exp(-h/h_{ml})] \quad (20)$$

for $h > 0$, and

$$\frac{df_{el}^{(0)}(h)}{dh} = 0 \quad (21)$$

for $h \leq 0$. h_{ml} is the thickness of one monolayer. The reference state stress was taken to be $\sigma_{xx}^{(0)}(h, h) = M\varepsilon$ for $h > 0$. The thickness dependence of $\sigma_{xx}^{(0)}(h, h)$ only slightly alters the wetting layer thickness predicted from Eq. (17), and hence was ignored.

5. The Stability of Thin Films

According to Eq. (17), anisotropic surface tension greatly enlarges the linearly stable wetting layer thickness. Does this conclusion survive beyond linear stability analysis? When a linearly stable flat film is perturbed strongly, so that the surface orientation in some regions is far from the $\theta = 0$ direction, the local surface stiffness in these regions is much smaller than that for $\theta = 0$. This tends to destabilize the linearly stable film. We carried out numerical simulations (using the procedure described in Refs. [10, 11]) that showed that films thinner than the linear wetting layer were unstable to perturbations greater than a certain critical amplitude (Fig. 2): hence, such films are *metastable*. When large perturbations were applied, faceted islands developed in the film, which underwent ripening at later stages of the evolution.

We carried out simulations for films perturbed by random perturbations and by perturbations of a single wavelength. The critical perturbation amplitude, δ_c , depends on the wavelength of the perturbation, λ , taking its minimal value

$$\delta_c^m = \min_{\lambda} \delta_c(\lambda) \quad (22)$$

at $\lambda/l_0 \sim 10 - 50$, where $l_0 = 2\gamma_0/M\varepsilon^2$. δ_c^m is plotted as a function of lattice mismatch in Fig. 3. The linear wetting layer thickness for $G = 500$, $M = 1.5 \times 10^{11}$ N/m² and $h_{ml} = 5\text{\AA}$ is also shown for comparison.

When the lattice mismatch is small, δ_c^m is much larger than the linear wetting layer thickness. Hence, flat films thinner than the linear critical thickness are stable at small lattice mismatch. As the linear critical thickness at small lattice mismatch is very large, we expect the film to first become unstable to misfit dislocations. This is indeed seen in experiments [7, 8].

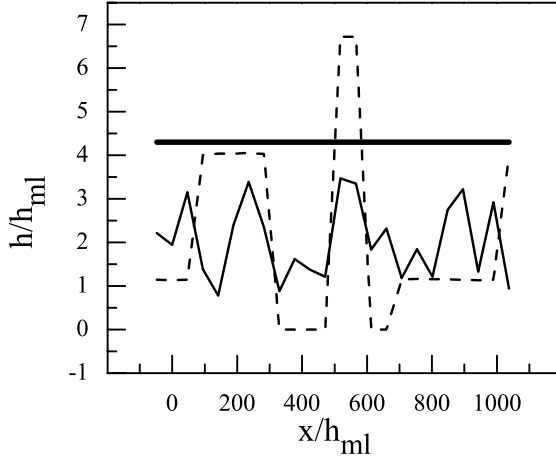


Figure 2. Evolution of a randomly perturbed film, in which perturbations were larger than the critical perturbation amplitude. Lattice mismatch in this film is 4%. The initial film surface is shown as a thin solid line. The dashed line shows the film surface at a later time. The linear wetting layer thickness is shown as a thick solid line.

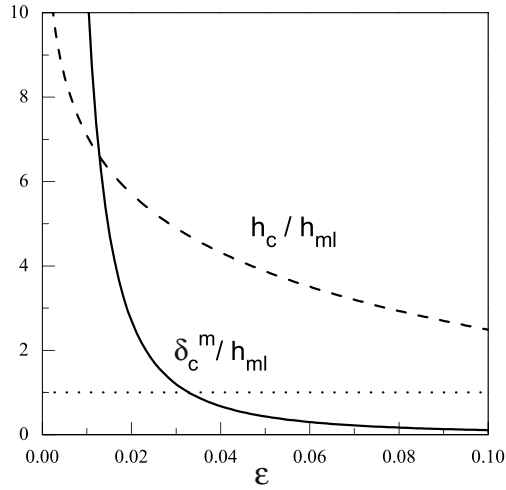


Figure 3. Variation of the minimal critical perturbation amplitude, δ_c^m , and linear wetting layer thickness, h_c , with lattice mismatch, ε . The minimal critical perturbation amplitude, δ_c^m/h_{ml} is represented by the solid line. The linear wetting layer thickness, h_c/h_{ml} , is represented by the dashed line. The dotted line shows the size of one monolayer, h_{ml} , for comparison.

At intermediate lattice mismatch, δ_c^m is of the order of a few monolayers. Hence we expect such a film to become unstable as perturbations of this amplitude are physically likely. In this regime films should develop growing perturbations at wavelengths given by $\lambda/l_0 \sim 10\text{--}50$. This corresponds to wavelengths of a few hundred nanometers. This typical wavelength decreases as lattice mismatch increases, agreeing with experiment [5, 7]. As ε increases, δ_c^m decreases in this regime from about 10 monolayers to approximately one monolayer, and we expect the thickness of the film needed to support such perturbations to correspondingly decrease. Such a trend is seen in experiments [5, 6, 7, 8].

For very large mismatch, a perturbation smaller than a monolayer is sufficient in order to destabilize the linearly stable wetting layer. Therefore, in practice, the wetting layer is a single monolayer in this case.

References

- [1] Y.-W. Mo, D. E. Savage, B. S. Swartzentruber and M. G. Lagally. Kinetic pathway in Stranski-Krastanov growth of Ge on Si(001). *Phys. Rev. Lett.* **65**: 1020–1023, 1990.
- [2] J. Massies and N. Grandjean. Oscillation of the lattice relaxation in layer-by-layer epitaxial growth of highly strained materials. *Phys. Rev. Lett.* **71**: 1411–1414, 1993.
- [3] T. R. Ramachandran, R. Heitz, P. Chen and A. Madhukar. Mass transfer in Stranski-Krastanow growth of InAs on GaAs. *Appl. Phys. Lett.* **70**: 640–642, 1997.
- [4] T. I. Kamins, E. C. Carr, R. S. Williams and S. J. Rosner. Deposition of three-dimensional Ge islands on Si(001) by chemical vapor deposition at atmospheric and reduced pressures. *J. Appl. Phys.* **81**: 211–219, 1997.
- [5] J. A. Floro, E. Chason, R. D. Twisten, R. Q. Hwang and L. B. Freund. SiGe coherent islanding and stress relaxation in the high mobility regime. *Phys. Rev. Lett.* **79**: 3946–3949, 1997.
- [6] J. A. Floro, E. Chason, L. B. Freund, R. D. Twisten, R. Q. Hwang and G. A. Lucadamo. Evolution of coherent islands in $\text{Si}_{1-x}\text{Ge}_x/\text{Si}(001)$. *Phys. Rev. B* **59**: 1990–1998, 1999.
- [7] R. M. Tromp, F. M. Ross and M. C. Reuter. Instability-driven SiGe island growth. *Phys. Rev. Lett.* **84**: 4641–4644, 2000.
- [8] D. D. Perovic, B. Bahierathan, H. Lafontaine, D. C. Houghton and D. W. McComb. Kinetic critical thickness for surface wave instability *vs.* misfit dislocation formation in $\text{Ge}_x\text{Si}_{1-x}/\text{Si}(100)$ heterostructures. *Physica A* **239**: 11–17, 1997.
- [9] H. J. Osten, H. P. Zeindl and E. Bugiel. Considerations about the critical thickness for pseudomorphic $\text{Si}_{1-x}\text{Ge}_x$ growth on Si(001), *J. Cryst. Growth* **143**: 194–199, 1994.
- [10] H. R. Eisenberg and D. Kandel. Wetting layer thickness and early evolution of epitaxially strained thin films. *Phys. Rev. Lett.* **85**: 1286–1289, 2000.

- [11] H. R. Eisenberg and D. Kandel. Origin and properties of the wetting layer and early evolution of epitaxially strained thin films. *Phys. Rev. B* **66**: art. no. 155429, 2002.
- [12] W. W. Mullins. Theory of thermal grooving. *J. Appl. Phys.* **28**: 333–339, 1957.
- [13] V. I. Marchenko and A. Ya. Parshin. Elastic properties of crystal surfaces. *Sov. Phys. JETP* **52**: 129–131, 1980.
- [14] S. Timoshenko and J. N. Goodier. *Theory of Elasticity* (McGraw-Hill, 1951).
- [15] S. G. Mikhlin. *Integral Equations* (Pergamon, New York, 1957).
- [16] H. P. Bonzel and E. Preuss. Morphology of periodic surface profiles below the roughening temperature: Aspects of continuum theory. *Surf. Sci.* **336**: 209–224, 1995.
- [17] R. J. Asaro and W. A. Tiller. Surface morphology development during stress corrosion cracking: Part I: via surface diffusion. *Metall. Trans.* **3**: 1789–1796, 1972.
- [18] M. A. Grinfeld. Instability of the separation boundary between a nonhydrostatically stressed elastic body and a melt. *Sov. Phys. Dokl.* **31**: 831–835, 1986.
- [19] D. J. Srolovitz. On the stability of surfaces of stressed solids. *Acta. Metall.* **37**: 621–625, 1989.
- [20] C. Herring. Some theorems on the free energies of crystal surfaces. *Phys. Rev.* **82**: 87–93, 1951.
- [21] J. Tersoff. Stress-induced layer-by-layer growth of Ge on Si(100). *Phys. Rev. B* **43**: 9377–9380, 1991.
- [22] L. G. Wang, P. Kratzer, M. Scheffler and N. Moll. Formation and stability of self-assembled coherent islands in highly mismatched heteroepitaxy. *Phys. Rev. Lett.* **82**: 4042–4045, 1999.
- [23] L. G. Wang, P. Kratzer, N. Moll and M. Scheffler. Size, shape and stability of InAs quantum dots on the GaAs(001) substrate. *Phys. Rev. B* **62**: 1897–1904, 2000.
- [24] C.-H. Chiu and H. Gao. A numerical study of stress controlled surface diffusion during epitaxial film growth. *Mater. Res. Soc. Symp. Proc.* **356**: 33–44, 1995.
- [25] B. J. Spencer. Asymptotic derivation of the glued-wetting-layer model and contact-angle condition for Stranski-Krastanow islands. *Phys. Rev. B* **59**: 2011–2017, 1999.
- [26] R. V. Kukta and L.B. Freund. Minimum energy configuration of epitaxial material clusters on a lattice-mismatched substrate. *J. Mech. Phys. Solids.* **45**: 1835–1860, 1997.

Ge/Si NANOSTRUCTURES WITH QUANTUM DOTS GROWN BY ION-BEAM-ASSISTED HETEROEPITAXY

A. V. Dvurechenskii, J. V. Smagina, V. A. Armbrister, V. A. Zinovyev,
P. L. Novikov, S. A. Teyss

*Institute of Semiconductor Physics, SB RAS
prospekt Lavrent'eva 13, 630090 Novosibirsk, Russia*

R. Groetzschel

*Institute of Ion Beam Physics and Material Research,
Forschungszentrum Rossendorf, D-01314, Dresden, Germany*

Abstract Scanning tunneling microscopy (STM) experiments were performed to study growth modes induced by hyperthermal Ge ion action during molecular-beam epitaxy (MBE) of Ge on Si(100). Continuous and pulsed ion-beams were used. STM studies have shown that ion-beam action during heteroepitaxy leads to decrease in critical film thickness for transition from two-dimensional (2D) to three-dimensional (3D) growth modes, enhancement of 3D island density and narrowing of size distribution, as compared with conventional MBE experiments. The crystal perfection of Ge/Si structures with Ge islands embedded in Si was analyzed by the Rutherford backscattering/channeling technique (RBS) and transmission electron microscopy (TEM). The studies of Si/Ge/Si(100) structures indicated defect-free Ge dots and Si layers for the initial stage of heteroepitaxy (5 monolayers of Ge) in pulsed ion beam action growth mode at 350°C. Continuous ion-beam irradiation was found to induce dislocations around Ge dots. The results of kinetic Monte Carlo (KMC) simulations have shown that two mechanisms of ion-beam action can be responsible for stimulation of 2D-3D transition: (i) surface defect generation by ion impacts, and (ii) the enhancement of surface diffusion.

1. Introduction

Self-assembled Ge islands on Si(100) have been intensively investigated as the basis of future electronic and optical devices [1]. At present,

it is commonly accepted that the energy gain caused by the strain relaxation in island apexes is the key factor in the transition from a two-dimensional (2D) to three-dimensional (3D) island growth. The 3D islands are formed due to the morphological instability of strained films in systems with a large (more than 2%) lattice mismatch between a film and substrate, among which Ge/Si (4%) and InAs/GaAs (7%) are most familiar. The conventional way to control island formation (size, shape and density) is the variation of growth conditions by the alteration of substrate temperature and molecular flux. However, to establish a method to achieve sufficiently uniform island sizes with a regular spatial distribution remains a critical issue. This should be solved since well-defined sizes with little dispersion are generally required for any practical applications.

The new facility to tune island dimensions and their surface densities is expected to be the use of ion-beams with energy exceeding the energy in the molecular beam, but less than the energy of defect generation in the bulk of the growing layer (and substrate). The results of our recent study indicate that irradiation with low-energy Ge^+ ions during Ge/Si(111) heteroepitaxy stimulates the nucleation of 3D Ge islands and reduces the critical thickness at which the 2D–3D transition occurs [2, 3]. In this work, we present the results of an investigation of size ordering in an ensemble of Ge nanoislands formed by heteroepitaxy under low-energy ion-beam irradiation and crystal perfection of Ge/Si structures with quantum dots embedded in Si. In order to clarify the effect of the ion irradiation on 3D island nucleation we have carried out the simulation of ion assisted growth of Ge films on Si with the kinetic Monte Carlo (KMC) method.

2. Experimental

The experiments were carried out in an ultrahigh-vacuum chamber of a molecular-beam epitaxy (MBE) setup equipped with an electron-beam evaporator for Si and effusion cell (boron nitride crucible) for Ge. A system of ionization and acceleration of Ge^+ ions provided a degree of ionization of Ge molecular beam from 0.1 to 0.5%. A pulsed accelerating voltage supply unit generated ion-current pulses with a duration of 0.5–1 s and an ion energy of 50–200 eV. The angle of incidence of the molecular and ion beams on the substrate was 55° to the surface normal. The analytical section of the chamber included reflection high-energy (20 keV) electron-diffraction (RHEED).

Heteroepitaxy was carried out at substrate temperatures varied in the range of 300–500° C. The rate of Ge deposition varied from 0.05 to 0.1

monolayer (ML) per second. Three types of Ge/Si(100) heterostructures were investigated: grown by (i) conventional MBE of Ge on Si, (ii) MBE with single pulsed Ge^+ ion action for each Ge monolayer completed at layer-by-layer growth mode, and (iii) MBE under continuous irradiation by the Ge^+ ion beam. The surface morphology was studied *ex situ* with scanning tunneling microscopy (STM). Structures were analyzed by Rutherford backscattering/channeling technique (RBS) with 1.2 MeV He^+ ions and transmission electron microscopy (TEM). For that, a 150 nm thick cap layer of Si was grown at 500°C by conventional MBE (with no irradiation) over the Ge layer.

3. Results and Discussion

Examples of STM patterns observed for structures with 5 monolayers of Ge grown on Si(100) at 350°C are shown in Fig. 1. An average island size obtained by conventional MBE was 22 nm and their dispersion (full width at half maximum) was 3.5 nm. Those for experiments with pulsed irradiation of Ge^+ ion beam were 6.5 ± 0.7 nm. In the case of continuous irradiation with a Ge^+ ion-beam, the average size of islands diminished (18 nm), but their dispersion increased (5.4 nm) in comparison with those obtained from conventional MBE. The surface density of Ge nanoislands for the structures of second type was $6.8 \times 10^{11} \text{ cm}^{-2}$, which is approximately seven times higher than that for the structures of first type ($\sim 10^{11} \text{ cm}^{-2}$). The density of Ge nanoislands in experiments under continuous ion irradiation with Ge^+ was $2 \times 10^{11} \text{ cm}^{-2}$.

A decrease in the full width at half maximum of the size distribution function is evidence for size ordering in an ensemble of Ge nanoclusters. We found that the ordering process is caused by pulsed ion-beam actions at each Ge monolayer completed in layer-by-layer growth mode. RHEED was used for *in situ* control of the stressed state of Ge/Si(100) surface. In addition, the starting point of hut and dome clusters formation was observed due to specific reflexes produced in RHEED images by (105) and (113) facets. Figure 2 shows the evolution of the Ge lattice constant during the conventional and ion-assisted MBE. The arrows in this figure separate the stages of 2D growth, the growth of hut and dome clusters, respectively. One can see that the ion-beam action results in an earlier 2D-3D transition as well as the formation of dome clusters. The effect was found to be dependent on the energy of ions (Fig. 3). Under 200 eV Ge^+ irradiation the hut clusters are formed at 1 ML and the dome clusters at 2 ML earlier as compared to 100 eV Ge^+ irradiation.

The treatment of RBS spectra permitted us to calculate the backscattering yield for Ge layers embedded in Si and for Si layers. The backscat-

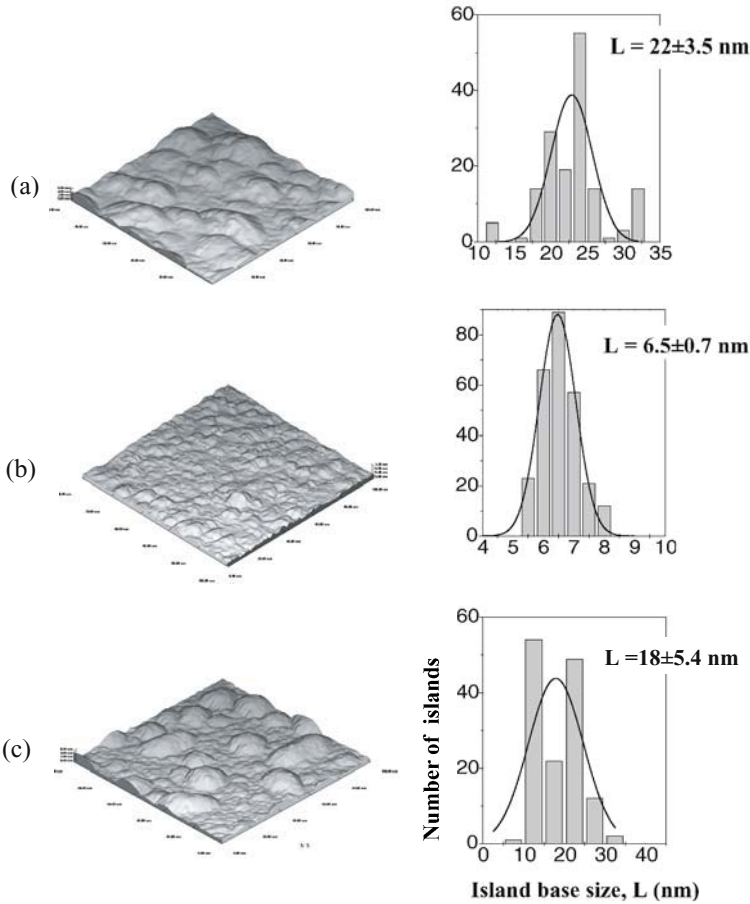


Figure 1. STM images of $100 \times 100 \text{ nm}^2$ surface area and size distribution of 3D islands for three types of Ge/Si(100) heterostructures: (a) conventional MBE of Ge on Si; (b) MBE with pulsed irradiation by 100 eV Ge^+ ions; (c) MBE with continuous 100 eV Ge^+ ion-irradiation. The rate of Ge deposition was 0.1 ML/s. Substrate temperature was 350°C .

tering yield for perfect crystal of Si(100) is about 3%. The backscattering yield from Ge layers turns out to be sensitive to the growth conditions (Fig. 4). The perfect structure of the 2.5% backscattering yield was found in the pulsed irradiation mode over the range of 1 to 5 Ge ML deposited at 350°C . For the lower temperature of 300°C , the yield exceeded 5% in similar structures. This increase in backscattering has also been observed for even more thick Ge layers at higher growth temperatures (400°C – 500°C). The enlarged yield was found also in structures formed with continuous beam irradiation at temperatures in the range

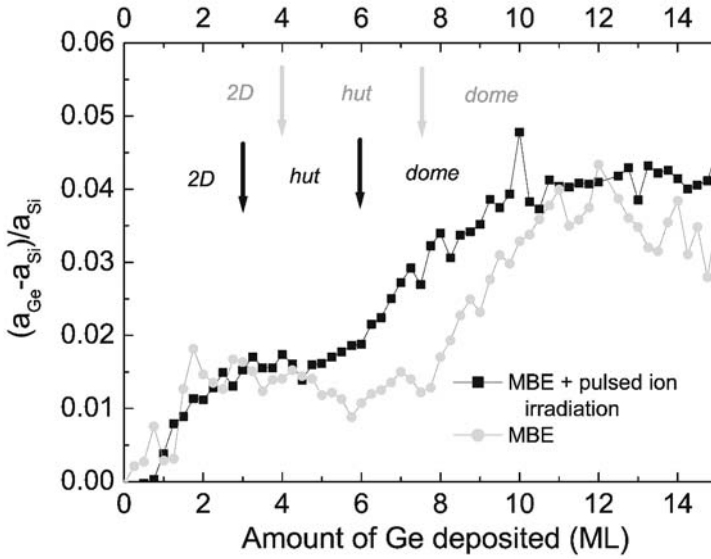


Figure 2. Change in Ge lattice constant during conventional and ion-assisted MBE of Ge on Si(100). The arrows indicate the appearance of hut and dome clusters registered by RHEED.

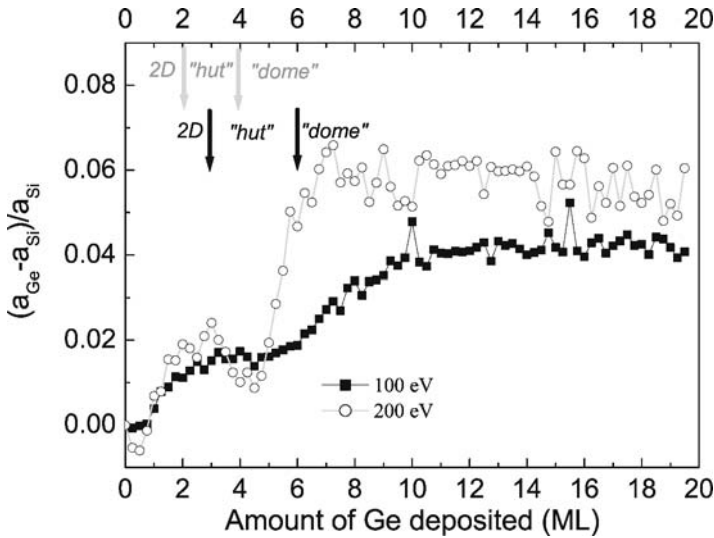


Figure 3. Change in Ge lattice constant during ion-assisted MBE of Ge on Si(100) for different Ge^+ energies. The arrows indicate the appearance of hut and dome clusters registered by RHEED.

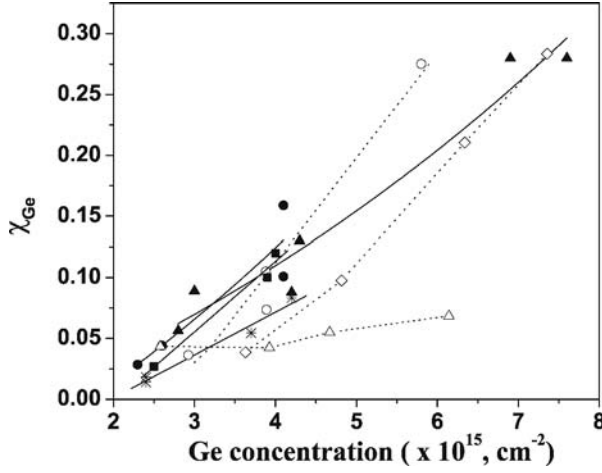


Figure 4. Backscattering yield from Ge embedded into Si layers as dependent on Ge layer thickness. Conventional MBE: $\triangle$ -300°C , $\diamond$ -400°C , $\circ$ -500°C ; MBE with continuous ion beam: $\blacksquare$ -300°C , $\bullet$ -350°C ; MBE with pulsed ion beam: $\blacktriangle$ -300°C , $*$ -350°C .

$300\text{--}350^{\circ}\text{C}$. The yield from the Si matrix (Fig. 5) slightly depends on the growth conditions and corresponds to perfect Si structure.

Transmission electron microscopy (TEM) studies indicated defect-free Ge dots and Si layers for the initial stage of heteroepitaxy (5 Ge ML) in pulsed action growth mode at 350°C . Continuous beam irradiation was found to induce dislocations around Ge dots. The size ordering of the Ge islands during heteroepitaxy with pulsed Ge^+ ion beam irradiation is, most likely, caused by the following factors: a synchronization of island nucleation by pulsed ion-beam and ion-induced enhancement of surface diffusion. The latter facilitates adatom exchange between islands. The increase in backscattering yield can be attributed to altering the elastic deformation inside the Ge islands due to ion-assisted change of their size and density and/or to generation and accumulation of point defects in bulk region of Ge.

4. Modeling

The process of 2D-3D transition under ion irradiation was simulated by the kinetic Monte Carlo method. The main elementary processes included in the model were atom deposition, diffusion hops and ion impact. At the first step we have modeled the pure heteroepitaxy of Ge/Si(111) without ion irradiation. The diffusion activation energy was assumed to depend on the bonding environment and elastic energy associated

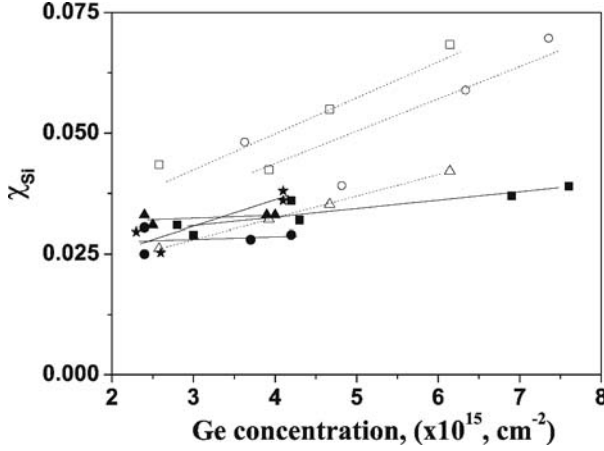


Figure 5. Backscattering yield from Si cap layer of Si/Ge/Si heterostructure as dependent on Ge layer thickness. Conventional MBE: Δ – 300°C, $\square$ – 400°C, $\circ$ – 500°C; MBE with continuous ion beam: $\blacktriangle$ – 300°C, $\star$ – 350°C; MBE with pulsed ion beam: $\blacksquare$ – 300°C, $\bullet$ – 350°C.

with strain, $E = E_{\text{bond}} - E_{\text{strain}}$, where $E_{\text{bond}} = n_1 E_1 + n_2 E_2$, (E_1 is the nearest-neighbour binding energy, E_2 is the next nearest-neighbour binding energy, n_1 is the number of the nearest neighbours, n_2 is the number of the next nearest neighbours); E_{strain} is the strain energy per atom, calculated using the Keating potential [4]. It follows from those calculations that the strain energy is maximal near the island edge and depends on the island size. We took E_{strain} into account only for atoms on the island edge. The simulation of growth within above assumptions results in the 2D-3D transition as soon as the critical thickness of Ge layer is achieved. The main features of the simulation model are presented in detail elsewhere [5, 6].

At the second step we included the low energy ion beam irradiation in the model. The ion beam was assumed to be responsible for the following processes: (a) sputtering of the material, (b) generation of additional adatoms and surface vacancy clusters, and (c) ion-assisted enhancement of surface diffusion. According to molecular dynamics simulation of low-energy interaction with Si surface [7], qualitatively correct for Ge, an ion impact in conditions similar to those in our experiment, produces a cluster of 10 vacancies, 9 excited adatoms and one atom sputtered [8]. In our simulations we used a magnitude of surface diffusion coefficient 10 times higher than that for case without ion-irradiation, which agrees with recent experimental measurements [9]. The simulation has shown that the growth can occur in two regimes: 2D layer-by-layer growth,

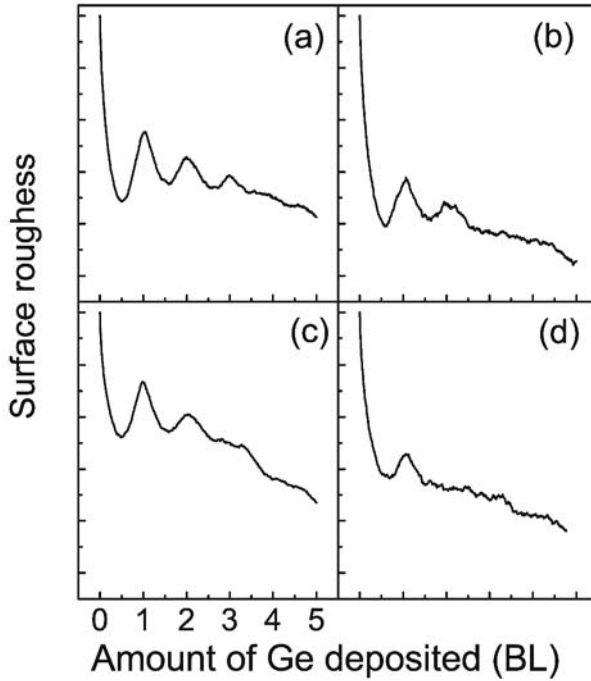


Figure 6. The calculated surface roughness of the Ge/Si(111) structure obtained by KMC simulations as dependent on the amount of Ge deposited. The model included: (a) no ion beam effect, (b) generation of adatoms by the ion beam, (c) enhancement of surface diffusion, (d): both (b) and (c).

when the oscillations of surface roughness is observed, and 3D growth, when oscillations disappear (Fig. 6). The 2D-3D transition is confirmed also by images of the simulated surface. For the case when the main ion-assisted process is the generation of additional adatoms and surface vacancy clusters, the 2D-3D transition occurs earlier [Fig. 6(b)], than in the case of the usual heteroepitaxy [Fig. 6(a)]. The number of oscillations is reduced to 2. The density of 3D islands is higher than in the case of usual epitaxy taken at same amount of Ge deposited [3.4 biatomic layers (BLs), corresponding to the beginning of 2D-3D transition].

For the case when the main ion-assisted process is the enhancement of surface diffusion, we obtained that the transition occurred at the same critical thickness as in the first case [Fig. 6(c)]. But the size and density of islands are different. The islands become larger and higher and the density decreases. Additionally, the surface roughness is lower in comparison with the case when only the surface defect generation by the ion beam was taken into account.

The ion-assisted enhancement of surface diffusion leads to an increase in the average size of 2D islands. As a consequence, the strain energy becomes higher, which promotes the hops of atoms from an edge to the upper layer. This leads to nucleation of 3D islands at the earlier stage of growth. The facilitation of the 2D-3D transition by the defect generation mechanism is interpreted as the result of ion impacts producing additional adatoms, which can hop on the top of existing 2D islands and nucleate 3D islands. Thus, both mechanisms promote the transition to 3D growth and the simulations including both mechanisms simultaneously have shown a stronger effect on 2D-3D transition [Fig. 6(d)]. In this case, the critical thickness is decreased up to 1 BL.

5. Summary and Conclusions

Our experimental results demonstrate that Ge/Si(100) heteroepitaxy with a pulsed low-energy ion-beam action enables the creation of defect-free 3D Ge islands with small sizes and high density. It provides a narrower size distribution of islands in comparison with conventional MBE. This is promising for potential applications in the technology of nanostructures. The results of KMC modeling have shown that both the generation of adatoms by ion-beams and enhancement of surface diffusion promote the transition from 2D to 3D growth mode.

Acknowledgments.

We acknowledge the support of the Russian Fund of Fundamental Research (grants 02-02-16020, 01-02-16830), INTAS (projects 2001-0615, 01-2257) and the Federal Scientific and Technical Program (project 40.012.1.1.1153).

References

- [1] K. Brunner. Si/Ge nanostructures, *Rep. Prog. Phys.* **65**: 27–72, 2002.
- [2] A. V. Dvurechenskii, V. A. Zinovyev, V. A. Kudryavtsev, and J. V. Smagina. Effects of low-energy ion irradiation on Ge/Si heteroepitaxy from molecular beam. *JETP Letters* **72**: 131–133, 2000.
- [3] A. V. Dvurechenskii, V. A. Zinoviev, and Zh. V. Smagina. Self-organization of an ensemble of Ge nanoclusters upon pulsed irradiation with low-energy ions during heteroepitaxy on Si. *JETP Letters* **74**: 267–269, 2001.
- [4] P. N. Keating. Effect of invariance requirements on the elastic strain energy of crystals with applications to the diamond structure. *Phys. Rev.* **145**: 637–645, 1966.
- [5] A. V. Dvurechenskii, V. A. Zinovyev, V. A. Kudryavtsev, Zh. V. Smagina, P. L. Novikov, and S. A. Tey. Ion-beam assisted surface islanding during Ge MBE on Si. *Phys. Low-Dim. Struct.* **1/2**: 303–314, 2002.

- [6] K. E. Khor and S. Das Sarma. Quantum dot self-assembly in growth of strained-layer thin films: A kinetic Monte Carlo study. *Phys. Rev. B.* **62**: 16657–16664, 2000.
- [7] V. A. Zinovyev, L. N. Aleksandrov, V. A. Dvurechenskii, K.-H. Heinig, D. Stock. Modelling of layer-by-layer sputtering of Si(111) surfaces under irradiation with low-energy ions. *Thin Solid Films* **241**: 167–170, 1994.
- [8] J. A. Floro, B. K. Kellerman, E. Chason, S. T. Picraux, D. K. Brice, and K. M. Horn. Surface defect production on Ge(001) during low-energy ion bombardment. *J. Appl. Phys.* **77**: 2351–2357, 1995.
- [9] R. Ditchfield and E. G. Seebauer. Semiconductor surface diffusion: Effects of low-energy ion bombardment. *Phys. Rev. B.* **63**: art. no. 125317, 2001.

LATERAL ORGANIZATION OF QUANTUM DOTS ON A PATTERNED SUBSTRATE

Catherine Priester

IEMN, dept ISEN, UMR CNRS 8520, Villeneuve d'Ascq, France

catherine.priester@isen.fr

Abstract The work reported here focuses on the role of nanopatterning in strained heteroepitaxy. This study makes use of an atomistic description. Two types of prepatterning are considered:

(i) A perfectly periodic strain field induced by a buried array of twist interface dislocations in a twist-bonded bilayer substrate. Network periodicity is controlled by the misorientation angle between the substrate and the surface bonded layer. For a thin enough surface bonded layer (a few tens of nm) the strain field variations appear to be strong enough to laterally organize quantum dots when a strained layer is grown (Ge deposited on a Si/Si twist bonded sample). The mechanism is similar to what happens in vertical alignment in multi-quantum dots layers.

(ii) Nanomesas: a quite regular array of nanomesas can be obtained by using stress-selective etching of the surface of twist bonded samples. A preliminary study of strained growth on nanomesas compared to strained growth on ideally flat substrates is also reported. It is shown how and why the elastic relaxation at the edges of the mesas can delay or even inhibit the 2D-3D transition. However, related to the design parameters of these nanomesas, one still gets 3D quantum dots (whose shapes are quite different from the usual self-assembled quantum dots shapes) which are very well laterally organized and calibrated.

Keywords: Self-assembled quantum dots, lateral organization, mismatched heteroepitaxy, twist boundary

1. Introduction

Intensive investigations have been devoted to the fabrication of quantum dots with controlled parameters. Self assembly appears to be an attractive path; however, size and shape calibration requirements are not easy to fulfil. To overcome the difficulty, prepatterning is a very efficient tool for reaching a high quality lateral organization of similar quantum

dots. Indeed, if a perfect regular lateral organization is reached, better size and shape calibration will follow.

In this work we report on two types of substrate surface patterning, both of which result from the use of a substrate with a thin twist-bonded overlayer. Section 2 is devoted to the description, from an atomistic point of view, of the strain field at the surface of a thin Si layer twist-bonded on a Si substrate, resulting from the regular array of dislocations buried at the bonded interface. We study how this strain field depends on design parameters, and how its strength influences lateral organization of the Ge islands that nucleate when Ge is deposited on such a surface. The period of this strain field has to lie in the range of a few tens of nanometers for an optimal efficiency. In the third section, the organizing mechanism we study is bound to morphological considerations: the patterning is no longer strain induced, but made of nanomesas which result from strain sensitive chemical etching of the substrates used in Sec. 2. The early steps of a theoretical study of strained heteroepitaxy on a regular array of nanomesas are therefore displayed in Sec. 3.

2. Lateral Organization due to the Strain Field from a Buried Array of Dislocations

The use of an inhomogeneous strain field for steering lateral organization has been previously proposed. One of the most famous examples is the strain inhomogeneity from interacting buried strained dots [17]. Bourret has also demonstrated, using continuum elasticity, that the surface strain field modulation from a buried incoherent interface is strong enough to organize quantum dots [2]. On the other hand, Fournel *et al.* [3] have combined a silicon bonding technique with layer transfer to provide a new type of substrate with a thin silicon layer on a thick silicon substrate with a tilt and twist-bonded interface. The angle between the two crystal orientations is performed with an accuracy of $\pm 0.005^\circ$ [4].

In this work we focus on systems with an array of pure twist interface dislocations (TWIDs). Making use of an atomistic description we quantify their ability to laterally organize quantum dots. The lateral period is tuneable from a distance of few angstroms up to more than $200\text{ }\mu\text{m}$ by controlling the misalignment angle between the two crystals. The distance between two neighboring TWIDS is directly related to the twist angle, and in Table 1 are reported the most interesting twist boundaries, together with the inter-dislocation distance. Because of computer time limitations we have calculated here only systems corresponding to the largest angles and the smallest periods (i.e. corresponding to the first three lines of Table 1).

Table 1. Several small angle twisted bonded interfaces.

Twist Boundary	Angle	d (nm)
Σ_{265}	4.98°	4.78
Σ_{613}	3.27°	6.72
Σ_{925}	2.66°	8.26
Σ_{1301}	2.25°	9.79
Σ_{2381}	1.66°	13.25
Σ_{4901}	1.16°	19.01

Whereas such systems with rather small twist angles ($0.8\text{--}5^\circ$) are attractive for lateral organization, systems with larger twist angles (a few tens of degrees) appear to provide interesting possibilities for compliance purposes. These latter cases have been previously described from the atomistic point of view by making use of Keating's potential [5, 6]. Such a formalism is used here for describing the systems we are interested in. They are composed of one (001) diamond semi-infinite Si substrate plus a thin (001) diamond Si layer, rotated with respect to each other on an [001] axis by an angle θ . We only consider periodic systems which correspond to coincidence angles (these are well known and studied as grain boundaries systems [7, 8, 9]).

As can be seen in Fig. 1(a) for the Σ_{265} bonding interface, new bonds linking neighbouring atoms across a twisted interface can be established first in accordance with the sp^3 bond criterion and second without any dangling bonds. "Distorted lines" clearly appear that correspond to the TWIDs recently observed by Rouvière *et al.* [10]. These interface dislocations induce a strain field that spreads into the whole sample. Figures 1(b, c, d) display this strain field at three distances in the bonded layer. This figure clearly bears evidence of the strong decrease of the strain field amplitude as one goes higher in the bonded layer. As previously shown by Bourret [2], this decay is somewhat smoother for smaller twist angles. Such a decay would let us think that, for realistic samples (twist bonded layer thicker than a few nanometers), there would be no chance of efficient lateral organization of dots. However, owing to the strain field induced by the partially strained Ge dots, this is not the case.

The interaction between the two deformation fields (one due to interface dislocations and the other to the partial elastic relaxation of dots) can strongly enhance the effect of a rather weak dislocation-induced surface-deformation field. Such an enhancement mechanism has been pointed out for vertical organization in multi-quantum dot layers [11]. For quantifying this phenomenon, we have chosen the five substrate samples listed in Table 2. The first three differ by the twist angle, but not by the bonded layer thickness, and the last three by the layer thickness,

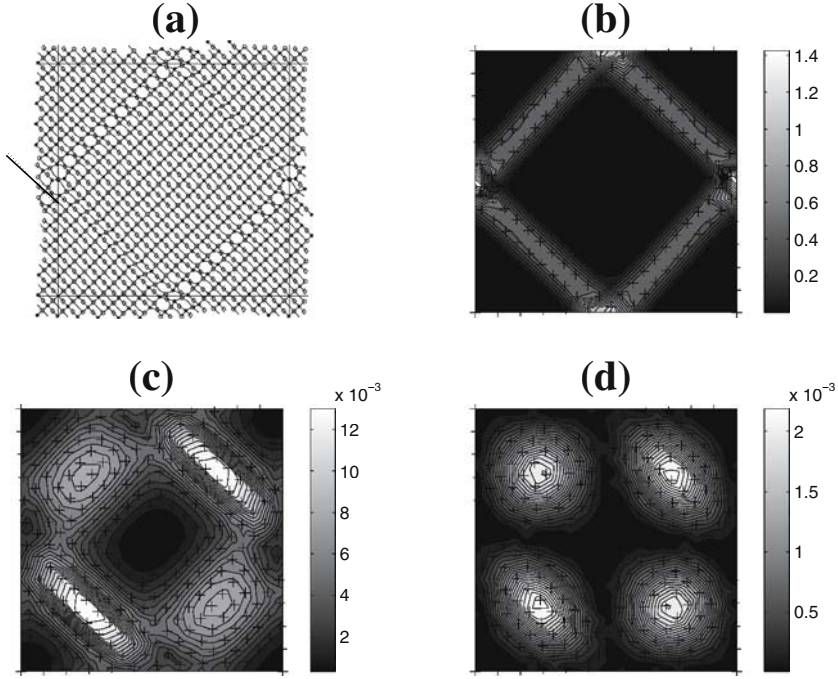


Figure 1. The $\Sigma_{265}(\theta = 4.98^\circ)$ system. (a) Top view of the bonded interface. Tags are circles and points for Si-atoms in the host substrate, and stars and squares for Si-atoms in the bonded layer. The sp^3 bonds are shown by lines. The [110] directions for the host substrate and for the bonded layer are indicated by dotted and solid lines, respectively. (b, c, d) Local strain energy maps at (b) the upper interface plane, (c) 8 atomic planes (i.e. 1.2 nm) upper (d) 16 atomic planes (i.e. 2.4 nm) upper.

but not by the twist angle. At the surface of these samples plus two Ge wetting monolayers, we move a “probe” that is an island with a shape intermediate between the 2D platelets one gets during 2D growth mode and the 3D huts that appear just after the 2D-3D transition: two- or four-atomic planes high, nonfacetted, and roughly circular. As a result of the dislocation network period and the natural lateral dimension of Ge self-organized quantum dots, only one quantum dot is expected per period. The point is that, if the strain field is efficient enough for laterally organizing the final 3D islands, it has first to order such intermediate probes. By moving the probe, we can investigate if the strain field is efficient enough to allow optimal locations to be determined. These locations are the minima of the reduced normalized elastic energy E_{red} :

$$E_{\text{red}} = \frac{E_{\text{tot, is}}/N_{\text{is}}}{E_{\text{tot, wl}}/N_{\text{wl}}},$$

Table 2. The samples studied.

Twist Boundary	Angle	d (nm)	Bonded Layer Thickness (atomic planes/nm)	TLB/ d
Σ_{265}	4.98°	4.78	16/2.17	0.45
Σ_{613}	3.27°	6.72	16/2.17	0.32
Σ_{925}	2.66°	8.26	16/2.17	0.26
Σ_{925}	2.66°	8.26	22/2.99	0.36
Σ_{925}	2.66°	8.26	26/3.53	0.42

where $E_{\text{tot, is}}$ is the elastic energy of the system (substrate plus wetting layer) *with* the probe minus the elastic energy of the system (substrate plus wetting layer) *without* the probe, N_{is} the number of atoms in the probe, $E_{\text{tot, wl}}$ is the elastic energy of the purely 2D system equivalent to the wetting layer plus the probe, and N_{wl} the number of atoms in this 2D equivalent layer.

The variations of E_{red} as the probe moves over the surface (Fig. 2) show that the island always preferentially nucleates at the vertical of the centre of the squares defined by the dislocation array. A closer look teaches that the shear distorted areas at the vertical of TWIDs are re-

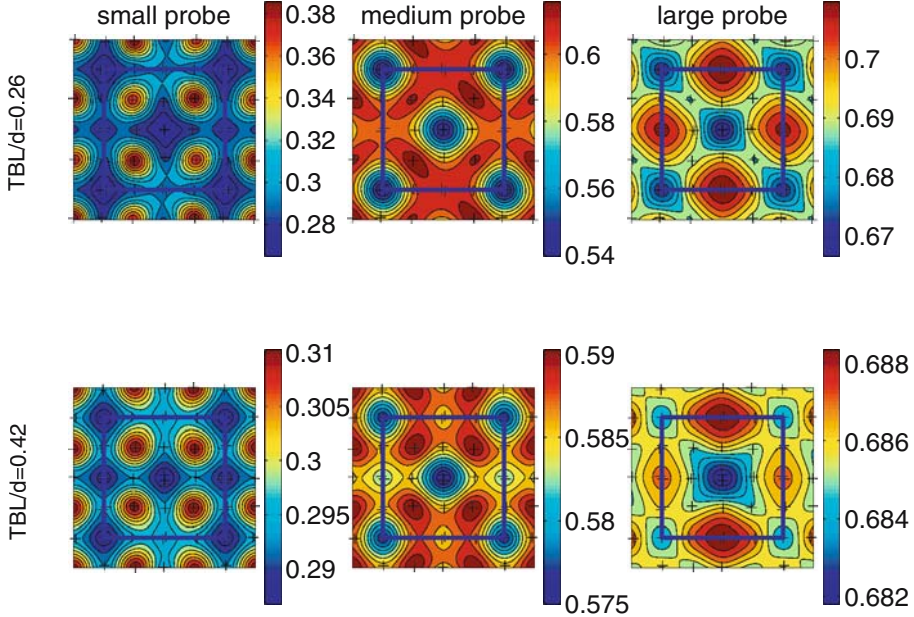


Figure 2. E_{red} variation maps for two samples (twist angle = 2.7° ; $\text{TBL}/d = 0.26$ (top) and $\text{TBL}/d = 0.42$ (bottom) and three probe sizes (left hand: small; center: medium, right hand: large)

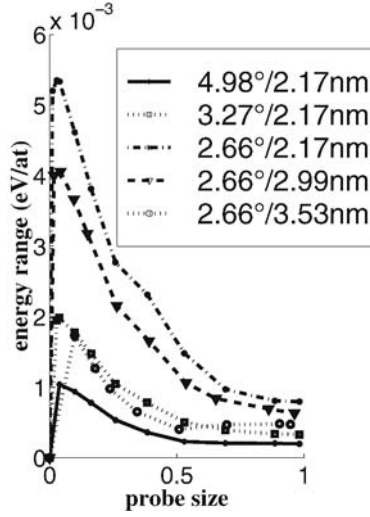


Figure 3. Organization strength *versus* probe size (in units reduced to the distance between two TWIDs) for the 5 samples in Table 2 (twist angle and bonded layer thickness are labelled in the inset). The vertical axis is the reduced energy difference between the optimal and less favorable probe locations.

pulsive for the island, because the island relaxes much more easily when located where it is simply biaxially strained. In fact, the organization effectiveness can be quantified by the energy difference between the optimal and the less favorable location. This quantity variations *versus* probe size have been reported in Fig. 3, for the five samples described in Table 2. From Fig. 3 one can deduce first that the smallest twist angle is the better for dot organization and second, that the efficiency is enhanced when the bonded layer becomes thinner. Thus, the optimal design parameters correspond to a twist angle of one degree (in this case, the distance between buried TWIDs is about 30-40 nm, so that one gets only one single dot per pattern) and a bonded layer about 10 nm thick (a thickness reachable by current technology [3]) is thin enough for providing strong enough organization efficiency. This result extends the pioneer work by Bourret[2] who has calculated the TWID induced strained field but has not considered the active role of the nucleating strained dots. To date, experimentalists have managed to get 40 nm periods with 10-nanometer thick bonded layers [12], a system which proves to be sufficient for reasonably organizing quantum dots.

It is important to note that in this section we have assumed that the surface of the bonded layer substrate is flat. This is not the case with real samples, simply due to the fact that fabricating such bonded

layer samples requires several technological steps, the last one being the thinning of the bonded layer by chemical etching. This chemical etching reveals the strain field, providing a periodically rough surface. The surface is thus not only strain-patterned, but also morphologically patterned. This is the purpose of next section

3. Strained Growth on a Nanomesa

When the twist-bonded layer becomes thinner, the surface stress is enhanced and the etching becomes more selective. As detailed in [13, 14, 15], a regular square network of rather sharp nanomesas can be obtained from such an etching process. Moreover the etching depth can be chosen by varying etching time and mixtures. For a deep enough etching process, the TWIDs can even be removed. From a theoretical point of view, the second step is thus to study how strained heteroepitaxy evolves when changing from a flat substrate to one with a nanomesa network. Let us emphasize that describing the growth mode requires a model that includes not only elastic, but also *surface* energies. Even if Keating's potential correctly describes elastic strain energy, it was not originally designed for straightforwardly taking into account surface energy. That is the reason why (001) surfaces are 2×1 reconstructed, and mesa facets are chosen to minimize the number of dangling bonds. Attributing to each dangling bond an energetic cost (roughly estimated as 350 meV from published surface energy values[16]) leads to a reasonable estimation of the elastic plus surface energy of the systems we look at.

In Fig. 4, the systems we consider are shown schematically. Figure 4(a) shows the Si nanomesa *before* Ge deposition. The question of substrate wetting is a tricky point which would warrant extensive investigations. As a starting point one can say that the bottom of the very sharp grooves between the mesas is not an interesting sticking site for Ge atoms, as they would be too highly strained in this area [Fig. 4(b)]. The edges of the mesas appear much more attractive from an elastic point of view and the early Ge atom deposition has a chance to adopt a profile not too far away from the one shown in Fig. 4(c). However, for the sake of simplicity, in the present study we assume that only the tops of mesas are Ge covered [Fig. 4(d)]. This is probably *not* what exactly occurs, but one can reasonably guess that this will be sufficient for displaying the main effect concerning Ge stress relaxation from the mesa edges. Systems shown in Fig. 4(d) will be called below “pseudo 2D growth mode on nanomesa” whereas those shown in Fig. 4(e) will be called “3D growth mode on nanomesa”. In this latter case, a 3D island nucleates on the top of the wetting layer which covers the (001) tops of

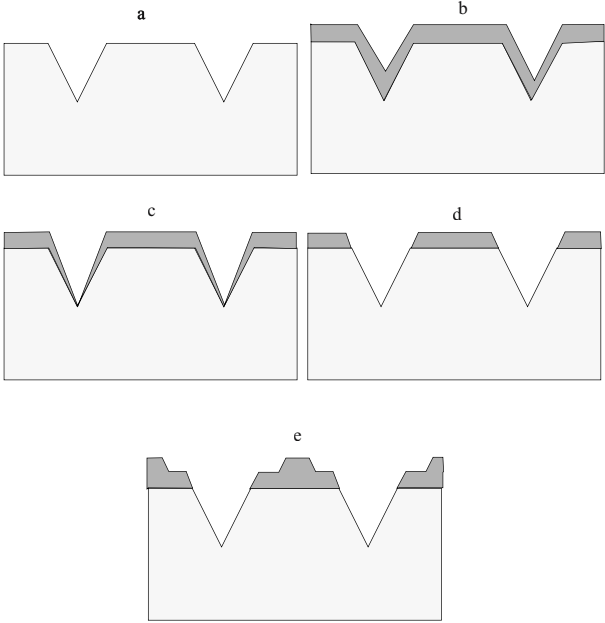


Figure 4. Schematic cross-sectional views of several nanomesa systems (see text). Light gray corresponds to the Si nanomesa and dark grey to germanium.

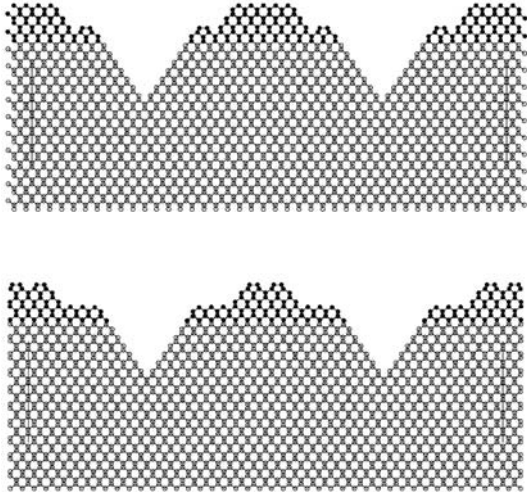


Figure 5. Atomic scale cross-sectional views of a (7.7 nm wide, 1.6 nm deep) Si (open circles) mesa array with Ge (closed points) deposited. Flat (001) areas are 2×1 reconstructed.

mesas. Figure 5 displays the cross section views of one of the smallest calculated nanomesa (with such Ge “3D growth”).

For estimating the 2D-3D critical thickness one must compare the sum of elastic and surface energies for the Ge/Si nanomesa system with and without a 3D island nucleus. Up to now we have restricted our investigations to square-based-truncated-pyramid shaped nuclei (varying their width, height and facets orientations) located at the centre of the nanomesa. Further investigation of different shapes (e.g. rings) will be performed soon. Figure 6 summarizes the pseudo-2D growth mode on nanomesas. The solid line gives the reduced energy variations on a flat substrate, whereas dotted and dashed lines correspond to 7.7- and 15.7-wide nanomesas (etched deeper for crosses than for triangles). This figure provides evidence of the elastic relaxation of the Ge layers: pseudo 2D layers relax so well that, the more Ge is deposited, the more the reduced energy decreases. Figure 6 also indicates that it is not the mesa size that is the most important point, but its aspect ratio (obviously, the higher mesa the better the elastic relaxation).

For investigating the 2D-3D transition, we show in Fig. 7 the energy density variations for pseudo-2D and 3D growth modes for several amounts of deposited Ge. The dashed curve corresponds to wetting plus the calculated most stable 4 ML-high island nuclei. In the case of a flat substrate (Fig. 7, upper panels), starting from 3 ML, the 3D curve lies above the 2D interpolation, whereas, starting from 4 ML, the 3D curves lie below the 2D interpolation. This indicates that the growth mode be-

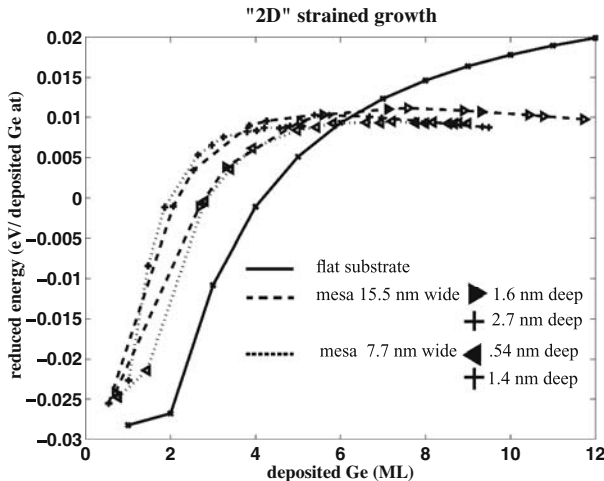


Figure 6. Reduced energy variations for a “pseudo 2D” Ge/Si growth mode on a flat substrate and a nanomesa.

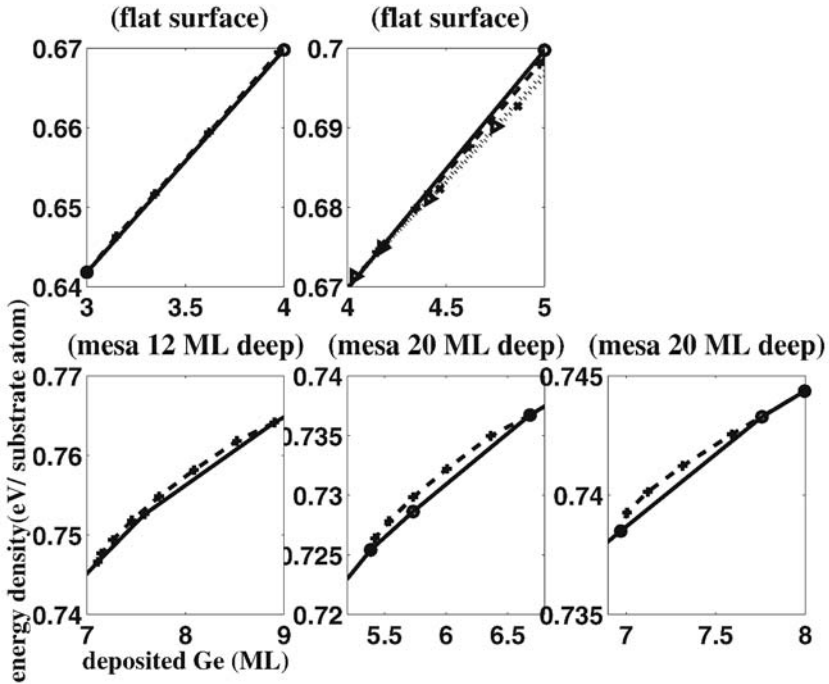


Figure 7. Energy density variations for a “pseudo 2D” (solid curves) and 3D (dashed curves) Ge/Si growth mode on a flat substrate and a nanomesa.

comes 3D just after 4 Ge ML have been deposited. The island will not remain 4 ML high but will grow upward because of the better stability of 5 and 6 ML high islands, as it can be seen from Fig. 7 (dotted lines, triangles and cross markers). The behaviour is quite different for growth on nanomesas because, as can be seen by the bottom line in Fig. 7, the 3D mode is always less stable than the pseudo-2D mode. From this we deduce that, on such nanomesas, there is no driving force for 3D island nucleation. Hence, the 2D-3D transition is inhibited, simply due to the elastic relaxation allowed by the edges of the mesa. This mechanism is active for mesa sizes around 10–20 nm, which corresponds to the latest experimental experiments[13]. Obviously, if the mesa enlarges towards the micron range, the relaxation by the mesa edges is no longer efficient enough, and behaviour similar to growth on flat surfaces will be brought back. However, keeping within the nanomesa range, it has to be noted that these “pseudo-2D wetting layers” at the mesa tops are to a certain extent 3D shaped due to their small (nanometer range) lateral dimensions, and thus can be viewed as perfectly well calibrated 3D-shaped dots, more relaxed than those obtained on a flat substrate.

This study is still at its beginning and further investigation will be performed in the future: even if the micromesa cannot be calculated because the model system would contain about 10^{10} atoms or more, it will be interesting to estimate, by making use of extrapolations, for which mesa sizes the usual 2D-3D transition would be recovered. Another highly interesting topic is the case of even smaller nanomesas (which would be even more attractive from a “3D dots” fabrication point of view, but one has first to check the wetting mechanism on such narrow mesas to make sure that in this case Ge does not first fill in the grooves).

4. Summary

In this work we have focused, from a theoretical point of view, on the role of nanopatterning in strained heteroepitaxy. We have shown that twist bonded interfaces present attractive possibilities from two points of view: either as creating a very regular strain field at the (flat) surface or as favoring, by the use of stress selective chemical etching, the fabrication of a very regular array of well calibrated square nanomesa. In the former case the classical 3D islands which nucleate interact with the strain field and thus reach high quality lateral organization which is a first step for dot size calibration. In the latter case the early stage of a more detailed study indicates that, due to the elastic relaxation of the wetting layer by the edges of the mesa, the usual 2D-3D transition is inhibited, but “3D-shaped” layers are provided by the mesa network geometry itself.

Acknowledgments

The author would like to thank D. Vvedensky and his co-organizers of the NATO ARW “Quantum dots: fundamentals, application, frontiers” for the invitation to this workshop and thus incitement to write this paper. She also wants to thank Joel Eymery for having provided her with unpublished experimental results, and Geneviève Grenet for her friendly critical reading of this manuscript.

References

- [1] J. Tersoff, C. Teichert, and M. G. Lagally. Self-organization in growth of quantum dot superlattices. *Phys. Rev. Lett.* **76**: 1675–1678, 1996.
- [2] A. Bourret. How to control the self-organization of nanoparticles by bonded thin layers? *Surf. Sci.* **432**: 37–53, 1999.
- [3] F. Fournel, H. Moriceau, N. Magnea, J. Eymery, J. L. Rouvière, K. Rousseau and B. Aspar. Ultra thin silicon films directly bonded onto silicon wafers. *Mat. Sci. Eng. B* **73**: 42–46, 2000.

- [4] F. Fournel, H. Moriceau, B. Aspar, K. Rousseau, J. Eymery, J. L. Rouvière, and, N. Magnea. Accurate control of the interface tilt and twist disorientations in direct wafer bonding. *Appl. Phys. Lett.* **80**: 793–795, 2002.
- [5] S. Rohart, C. Priester and G. Grenet. On compliant effect in twist-bonded systems. *Appl Surf. Sci.* **188**: 193–201, 2002.
- [6] C. Priester and G. Grenet. Plastic relaxation mechanisms in systems with a twist-bonded layer. *Nanomaterials for Structural Applications*, edited by C. C. Berndt, T. E. Fischer, I. Ovid’ko, G. Skandan, and T. Tsakalakos, Materials Research Society Symposium Proceedings **740**: I13.4.1–I13.4.6, 2003.
- [7] E. Tarnow, P. Dallot, P. D. Bristowe, and J. D. Joannopoulos, G. P. Francis, and M. C. Payne. Structural complexity in grain boundaries with covalent bonding. *Phys. Rev. B* **42**: 3644–3657, 1990.
- [8] R. D. Kamien and T. C. Lubensky. Minimal surfaces, screw dislocations, and twist grain boundaries. *Phys. Rev. Lett.* **82**: 2892–2895, 1990.
- [9] J. Thibault, J.L. Rouvière, A. Bourret. Grain boundaries in semiconductors. *Handbook of Semiconductor Technology: Electronic Structure and Properties of Semiconductors*, edited by K. A. Jackson and W. Schröter (Wiley-VCH, Weinheim, Germany, 2000) pp. 379–451.
- [10] J. L. Rouvière, K. Rousseau, F. Fournel, and H. Moriceau. Huge differences between low- and high-angle twist grain boundaries: The case of ultrathin (001) Si films bonded to (001) Si wafers. *Appl. Phys. Lett.* **77**: 1135–1137, 2000.
- [11] C. Priester. Modified 2D-3D growth transition process in multistacked self-organized quantum dots. *Phys. Rev. B* **63**: 153303–153303, 2001.
- [12] J. Eymery, private communication.
- [13] F. Leroy, J. Eymery, P. Gentile and F. Fournel. Controlled surface nanopatterning with buried dislocation arrays. *Surf. Sci.* **545**: 211–219, 2003.
- [14] F. Leroy, J. Eymery, P. Gentile and F. Fournel. Ordering of Ge quantum dots with buried Si dislocation networks. *Appl. Phys. Lett.* **80**: 3078–3080, 2002.
- [15] D. Buttard, J. Eymery, F. Fournel, P. Gentile, F. Leroy, N. Magnea, H. Moriceau, G. Renaud, F. Rieutord, K. Rousseau, and J. L. Rouvière. Toward two-dimensional self-organization of nanostructures using wafer bonding and nanopatterned silicon surfaces. *IEEE J. Quantum Elect.* **38**: 995–1005, 2002.
- [16] All the results reported in this section keep qualitatively unchanged when the dangling bond cost varies from 0.3 eV to 0.4 eV.

SOME THERMODYNAMIC ASPECTS OF SELF-ASSEMBLY OF QUANTUM DOT ARRAYS

José Emilio Prieto

*Institut für Experimentalphysik, Freie Universität
Germany*

n Markov

Institute of Physical Chemistry, Bulgarian Academy of Sciences, 1113 Sofia, Bulgaria

Abstract We have studied the relative adhesion (the wetting) of dislocation-free three-dimensional (3D) islands belonging to an array of islands to the wetting layer in the Stranski–Krastanov growth mode. The array has been simulated as a chain of islands in $1 + 1$ dimensions placed on top of a wetting layer. In addition to the critical size of the two-dimensional (2D) islands for the 2D-3D transformation to occur, we find that the wetting depends strongly on the density of the array, the size distribution and the shape of the islands.

1. Introduction

The instability of planar films against coherently strained three-dimensional (3D) islands in highly mismatched epitaxy is a subject of intense research in recent time owing to their possible optoelectronic applications as quantum dots [1]. The term “coherent Stranski–Krastanov (SK) growth” has been coined for this case of formation of 3D islands that are strained to fit the underlying wetting layer at the interface but are largely strain-free near their top and side walls [2, 3]. This term was introduced in order to distinguish this case from the “classical” SK growth in which the lattice misfit is accommodated by misfit dislocations at the interface [4].

Experimental studies of arrays of coherent 3D islands in SK growth of highly mismatched semiconductor materials one on top of the other have shown surprisingly narrow size distributions of the islands,[5, 6, 7, 8]

(see also Ref. [3] and the references therein). (It is worth noting that a narrow size distribution has been established also in the Volmer–Weber growth of metals on insulators in the absence of a wetting layer [9].) This phenomenon, known in the literature as self-assembly (for a review see Ref. [10]), is highly desirable as it guarantees a specific optical wavelength of the array of quantum dots. The physics of this self-assembly is still not understood in spite of the numerous thermodynamic and kinetic studies [11, 12, 13, 14, 15]. For example, Priester and Lannoo found that two-dimensional (2D) islands with a monolayer height act as precursors of the 3D pyramidal islands [16], (see also Ref. [17]). The energy per atom of the 2D islands possesses a minimum for a certain volume, but the 3D islands become energetically favored at a smaller size. Thus, at some critical surface coverage, the 2D islands spontaneously transform into 3D islands preserving a nearly constant volume during the 2D–3D transformation. The resulting size distribution reflects that of the 2D islands which is very narrow. This picture has been recently corroborated by Ebiko *et al.* [18], who found that the volume distribution of InAs/GaAs self-assembled quantum dots agrees well with the scaling function that is characteristic for the two-dimensional submonolayer homoepitaxy [19]. Korutcheva *et al.* [20] and Markov and Prieto [21] reached the same conclusion except that the 2D–3D transformation was found to take place through a series of intermediate states with discretely increasing thickness (one, two, three, *etc.* monolayers-thick islands) that are stable in separate consecutive intervals of volume. Khor and Das Sarma arrived at the same conclusion by using Monte Carlo simulations [22].

In a recent paper, Prieto and Markov discussed the formation of coherent 3D islands within the framework of the traditional concept of wetting [23]. As is well known the wetting parameter which accounts for the energetic influence of a crystal B in the heteroepitaxial growth of a crystal A on top of it is defined as (for a review, see Ref. [24])

$$\Phi = 1 - \frac{E_{AB}}{E_{AA}} \quad (1)$$

where E_{AA} and E_{AB} are the energies per atom required to disjoin a half-crystal A from a like half-crystal A and from an unlike half-crystal B, respectively.

The mode of growth of a thin film is determined by the difference $\Delta\mu = \mu(n) - \mu_{3D}^0$, where $\mu(n)$ and μ_{3D}^0 are the chemical potentials of the film (as a function of its thickness n) and of the bulk material A, respectively [24]. The chemical potential of the bulk crystal A is given at zero temperature by the work ϕ_{AA} (taken with a negative sign) to detach an atom from the well known kink or half-crystal position. The

latter name is due to the fact that an atom at this position is bound to a half-atomic row, a half-crystal plane and a half-crystal block [25, 26]. In the case of a monolayer-thick film of A on the surface of B the chemical potential of A is given by the analogous work ϕ_{AB} with the exception that the underlying half-crystal block of A is replaced by a half-crystal block of B . Thus $\Delta\mu = \phi_{AA} - \phi_{AB}$. In the simplest case of additivity of bond energies the difference $\phi_{AA} - \phi_{AB}$ reduces to $E_{AA} - E_{AB}$ as the lateral bondings cancel each other. Then $\Delta\mu$ is proportional to Φ , i.e. $\Delta\mu = E_{AA}\Phi$ [24]. It follows that it is the wetting parameter Φ which determines the mechanism of growth of A on B [27]. In the two limiting cases of growth of isolated 3D islands of A directly on the surface of B [Volmer–Weber growth, characterized by incomplete wetting ($0 < \Phi < 1$) and any misfit $\varepsilon_0 = \Delta a/a$], or by consecutive formation of monolayers of A on B [Frank–van der Merwe growth, with complete wetting ($\Phi \leq 0$) and $\varepsilon_0 \approx 0$], $\Delta\mu$ goes asymptotically to zero from above and from below, respectively, but it changes sign in the case of growth of 3D islands of A on a thin wetting layer of A on the substrate B [Stranski–Krastanov growth, complete wetting at the beginning ($\Phi \leq 0$) and $\varepsilon_0 \neq 0$] [23, 24]. The equation $\Delta\mu = E_{AA}\Phi$ is thus equivalent to the familiar 3- σ criterion of Bauer [24, 28].

The Stranski–Krastanov morphology appears as a results of the interplay of the film-substrate bonding, misfit strain and the surface energies. A wetting layer with a thickness of the order of the range of the interatomic forces is first formed (owing to the interplay of the A - B interaction and the strain energy accumulation) on top of which partially or completely relaxed 3D islands nucleate and grow. The 3D islands and the thermodynamically stable wetting layer represent necessarily different phases. If this were not the case, the growth would continue by 2D layers. Thus, we can consider as a useful approximation to regard the 3D islanding on top of the uniformly strained wetting layer as a Volmer–Weber growth. That requires the mean adhesion of the atoms that belong to the base plane of the 3D islands to the stable wetting layer to be smaller than the cohesion between them. In other words, the wetting of the underlying wetting layer by the 3D islands must be incomplete. Otherwise, 3D islanding will not occur [27]. In the Volmer–Weber growth the incomplete wetting is due mainly to the difference in bonding ($E_{AB} < E_{AA}$), the supplementary effect of the lattice misfit being usually smaller. In the coherent SK growth ($E_{AB} \approx E_{AA}$), the incomplete wetting is due to the lattice misfit, which leads to the atoms at the edges of the islands to displace from the bottoms of the potential troughs provided by the atoms in the layer underneath [29, 30, 23].

As has been shown elsewhere [20, 23], it is the incomplete wetting which determines the formation of dislocation-free 3D islands on top of the wetting layer in the case of coherent SK growth. In this case, however, the relation between $\Delta\mu$ and Φ is not as simple as given above. The bond energies are generally not additive, the misfit strain is relaxed mostly near the side and top walls and increasing the island's thickness leads to larger displacements of the edge atoms from the bottoms of the potential troughs provided by the wetting layer and in turn to a decrease of the wetting. For this reason, in this work we define the wetting parameter Φ as the difference of the interaction energies with the wetting layer of misfitting and non-misfitting 3D islands.

In the present paper we study the behavior of Φ for islands which belong to an array of islands. We study the effect of the density of the array (the distance to the nearest neighbor islands), the size distribution (the difference in size of the neighboring islands), and the shape distribution (the slope of the side walls of the neighboring islands) on the wetting parameter Φ of the considered island.

2. Model

We consider an atomistic model in $1 + 1$ dimensions (lateral size + height), which we treat as a cross section of the real $2 + 1$ dimensional case. An implicit assumption is that in the real $2 + 1$ dimensional model the monolayer islands have a compact rather than a fractal shape and that the lattice misfit is the same in both orthogonal directions. The 3D islands are represented by linear chains of atoms stacked one upon the other [31, 32]. Each upper chain is shorter than the lower one. The shape of the islands in our model is given by the slope of the side walls. For, example, an island which consists of consecutive chains with $N, N - 1, N - 2, \dots$ atoms has a 60° slope of the side walls, whereas an island with chains consisting of, say, $N, N - 5, N - 10, \dots$ atoms has a slope of 19.1° , *etc.*, where N is the number of atoms in the base chain. The array in the $1 + 1$ dimensional space is represented by a row of 3 or 5 islands on a wetting layer consisting of several monolayers (Fig. 1). The distance between two neighboring islands is given by the number n of vacant atomic positions between the ends of their base chains and can be varied from one to infinity.

In order to simplify the computational procedure, the “wetting layer” in our model is composed of several monolayers of the true wetting layer consisting of atoms of the overlayer material A plus several monolayers of the unlike substrate material B . This composite wetting layer has the atom spacing of the substrate material B as in the real case, but for the

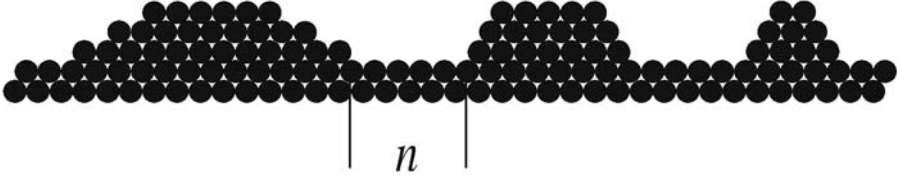


Figure 1. Schematic view of an array of islands on a wetting layer. The central island is surrounded by two islands with different shapes and sizes. The spacing between neighboring islands, denoted by n , is a measure of the density of the array.

sake of simplicity the atom bonding is that of the overlayer material A . We believe that the latter does not introduce a perceptible error as the energetic influence of the substrate B is screened by the true wetting layer A . We expect that this approximation underestimates to some extent the wetting parameter Φ by making the composite substrate a bit softer than the real one (the A - A bonding is weaker than the B - B bonding [33] (for a later review, see Ref. [24]). We found that beyond 10 monolayers the studied parameter Φ saturates its value. This is why in all cases given below, unless otherwise stated, we allowed 10 monolayers to relax.

As in our previous work [21], we make use of a simple minimization procedure. The atoms interact through a potential that can be easily generalized to vary its anharmonicity by adjusting two constants μ and ν ($\mu > \nu$) that govern separately the repulsive and attractive branches, respectively,[34]

$$V(x) = V_o \left[\frac{\nu}{\mu - \nu} e^{-\mu(x-b)} - \frac{\mu}{\mu - \nu} e^{-\nu(x-b)} \right], \quad (2)$$

where b is the equilibrium atom separation. In this work, we have used $\mu = 2\nu$ with $\nu = 6$, which turns the potential (2) into the familiar Morse form.

Our program calculates the interaction energy of all the atoms as well as its gradient with respect to the atomic coordinates, i.e. the forces. Atoms in the islands and in the wetting layer are then allowed to relax. Relaxation of the system is performed by allowing the atoms to displace in the direction of the gradient in an iterative procedure until the forces fall below some negligible cutoff value. Periodic boundary conditions are applied in the lateral direction. We consider only interactions in the first coordination sphere in order to mimic the directional bonds that are characteristic for most semiconductor materials [35].

3. Results

Figure 2(a) shows the horizontal displacements of the atoms of the base chain from the bottoms of the potential troughs provided by the homogeneously strained wetting layer for a misfit of 7%. The considered island has two identical ones at a distance of $n = 5$. This is the same behavior as predicted by the one-dimensional model of Frank and van der Merwe [29, 30]. The horizontal displacements increase with increasing island thickness (measured in number of monolayers) precisely as in the case of a rigid substrate and non-interacting islands.[21] In contrast to the rigid substrate case [23], the vertical displacements of the edge atoms of the base chain of the islands and the underlying atoms of the uppermost monolayer of the wetting layer are directed downwards (Fig. 2(b)). It is worth noting that the same result has been found by Lysenko *et al.* in the case of homoepitaxial metal growth by using a computational method within the framework of the tight-binding model [36].

In spite of their downwards vertical displacements, the edge atoms are again more weakly bound to the underlying wetting layer, as in the rigid substrate models of Refs. [20] and [21] (Fig. 3). Shown in the same figure for comparison is an island without neighbors. As seen, the edge atoms of the single island adhere more strongly to the substrate. This is in fact the essential physics behind the effect that the neighboring islands exert on the middle island. The central island looses to some degree contact with the substrate (in this case the wetting layer) and the wetting parameter is increased. We can interpret this as the wetting layer becoming stiffer under the influence of the neighboring islands.

The influence of the density of the array is demonstrated in Fig. 4. The values for 3 and 5 islands were calculated assuming equally spaced islands. These can be thus treated as a self-organized array. As expected the wetting parameter increases with decreasing distance between the islands or, in other words, with increasing array density.

Figure 5 shows the wetting parameter of the central island as a function of the size of the side islands. For this calculation, we considered three islands with the same thickness of 3 ML. Furthermore, both side islands have one and the same volume. Increasing the volume of the side islands leads to an increase of the elastic fields around them and to a further reduction of the bonding between the edge atoms of the central island and the wetting layer.

Figure 6 demonstrates one of the most important results, the effect of the size distribution on the wetting of the islands. It shows the behavior of the wetting parameter Φ of the central island as a function of the number of atoms in the base chain (which is a measure of the volume) of

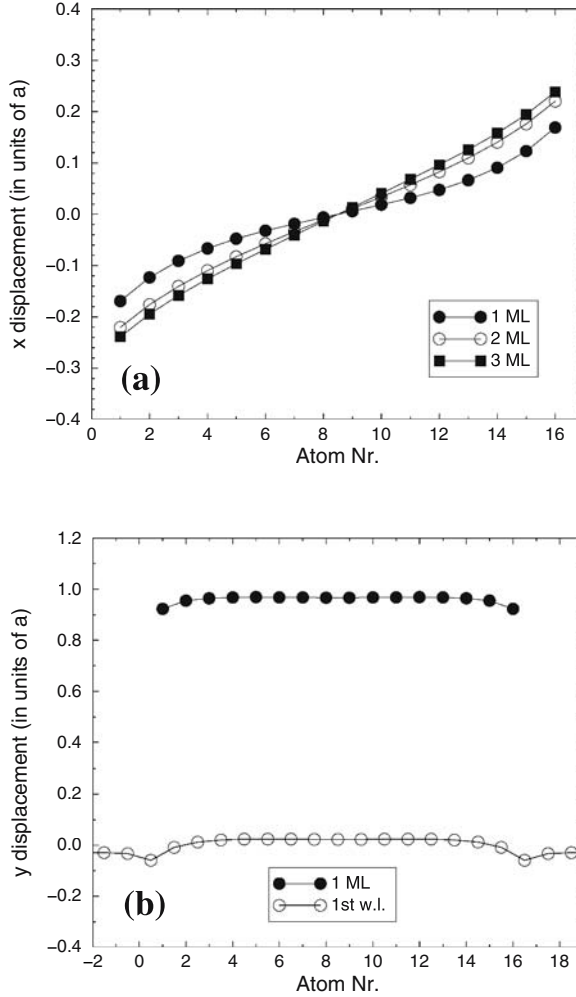


Figure 2. Horizontal (a) and vertical (b) displacements of the atoms of the base chain from the bottoms of the potential troughs provided by the homogeneously strained wetting layer, for a lattice misfit of 7%. The considered island has 16 atoms in the base chain and is located between two identical ones at a distance of $n = 5$. The displacements are given in units of the lattice parameter a of the composite wetting layer. The horizontal displacements increase with increasing island thickness taken in number of monolayers. Contrary to the rigid substrate case, the vertical displacements of the edge atoms of the base chain of the islands and those of the underlying atoms of the uppermost monolayer of the wetting layer, included in (b), are directed downwards.

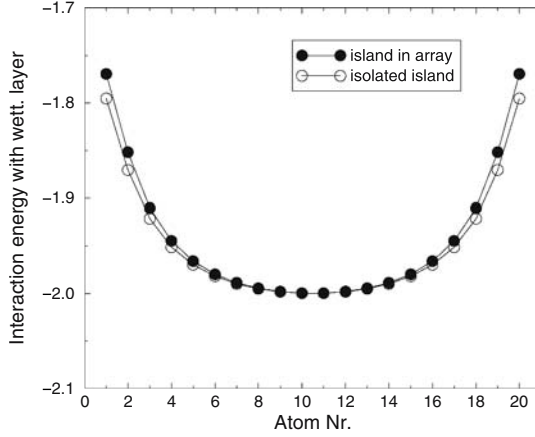


Figure 3. Distribution of the interaction energy (in units of V_0) between the atoms of the base chain A of a 3 ML-high, coherent island with 20 atoms in the base chain, and the underlying wetting layer B , for a positive misfit of 7%. Full circles correspond to an island separated by a distance $n = 5$ from two identical ones, the empty ones correspond to a reference isolated island.

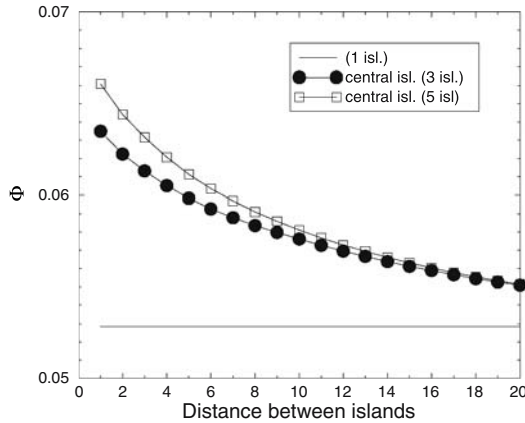


Figure 4. Dependence of the wetting parameter of the central island on the distance n between the islands. Results for arrays of 3 and 5 islands are given, as well as for a reference isolated island. All islands are 3 ML-high, have 20 atoms in their base chains and the lattice misfit amounts to 7%. As seen, the next-nearest neighbors play a smaller but not negligible role.

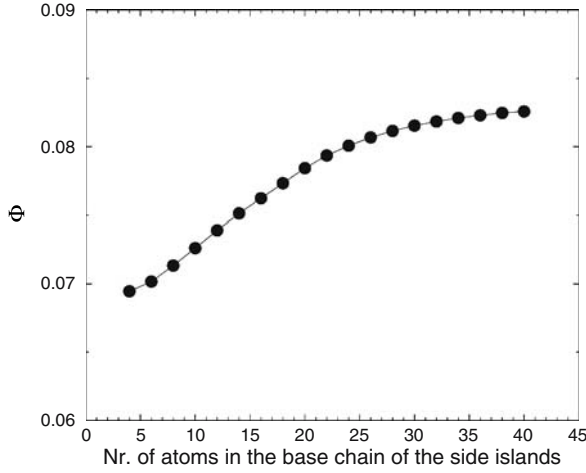


Figure 5. Dependence of the wetting parameter of the central island on the size of the base chains of the two side islands. These two have the same volume and are separated from the central one by a distance $n = 5$. All the islands are 3 ML high, the central one having 20 atoms in the base chain, the misfit amounts to 7% and the wetting layer consists of 3 ML that are allowed to relax.

the left island. In this case, the sum of the volumes (the total number of atoms) of the left and right islands is kept constant and precisely equal to the doubled volume of the central island. All three islands have the same thickness. The side walls of all the three islands make an angle of 60° with their bases. Thus, the first point (and, by symmetry, also the last one) gives the maximum asymmetry in the size distribution of the array, the left island consists of 9 atoms whereas the right island is built of 105 atoms. The point at the maximum of wetting describes the monodisperse distribution – the three islands have one and the same volume of 57 atoms. As seen in the case of perfect self-assembly of the array the wetting parameter, or in other words, the tendency to clustering displays a maximum value.

The effect of the shape of the side islands, i.e. their facet angles, on the wetting parameter of the central island is demonstrated in Fig. 7. The slope of the facets of the central island is 60° . The effect is greatest when the side islands have the steepest walls. The same result (not shown) is obtained when the central island has a different facet angle, e.g. 11° . The explanation follows the same line as the one given above. The side islands with larger-angle side walls exert a greater elastic effect

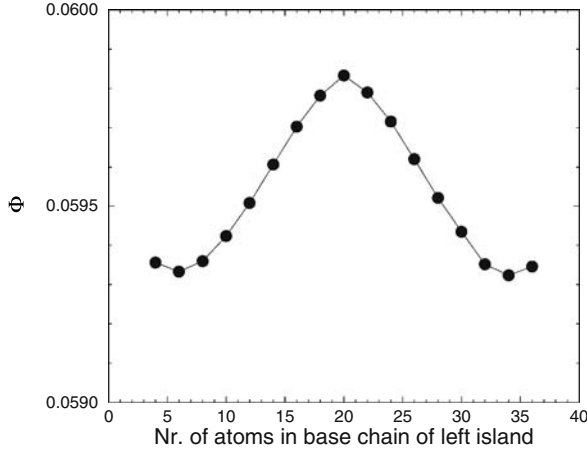


Figure 6. Dependence of the wetting parameter of the central island on the size distribution of the side islands. The central island is three monolayers thick and has 20 atoms in the base chain thus containing a total of 57 atoms. The lattice misfit is 7%. The x -axis represents the number of atoms in the base chain of the left island. The sum of the volumes of the left and right islands is kept constant and equal to the doubled volume of the central island. We increase the volume of the left island and decrease the volume of the right island thus passing through the middle point at the maximum where the three islands have equal volumes.

on the wetting layer and in turn on the displacements and the bonding of the edge atoms of the central island.

We studied also the stability of islands with a thickness increasing by one monolayer in the presence of two islands on both sides on a deformable substrate. The result in Fig. 8 shows the same behavior observed in Refs. [20] and [21] where rigid substrates were assumed. This means that the overall transformation from the precursor 2D islands to a macroscopic 3D islands takes place in consecutive stages in each of which the islands thicken by one monolayer. As shown in Refs. [20] and [21], the latter leads in turn to a critical misfit beyond which coherent 3D islanding takes place, and below which the lattice mismatch is accommodated by misfit dislocations. The existence of a critical misfit has been experimentally observed in a series of different systems [5, 37, 38, 39]. The energies computed in the case of the reference single islands always lie below the curves of the islands in an array. The difference obviously gives the energy of repulsion between the neighboring islands. It follows from the above that the presence of neighboring islands leads to a slight decrease of the critical misfit.

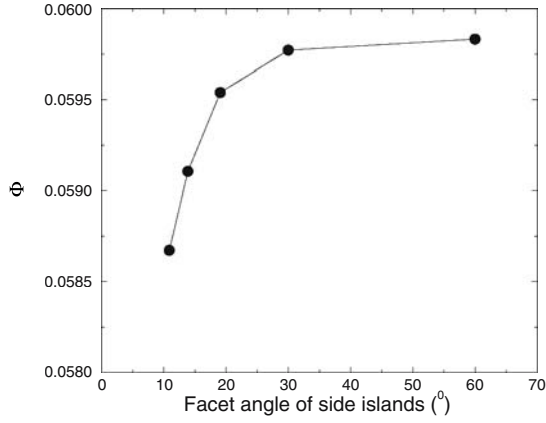


Figure 7. Dependence of the wetting parameter of the central island on the shape of the neighboring islands, measured in degrees of their facet angles. The central island has a slope of 60° of its side walls. All islands are 3 ML high, have 20 atoms in their base chains and are separated by a distance $n = 5$. The lattice misfit amounts to 7%.

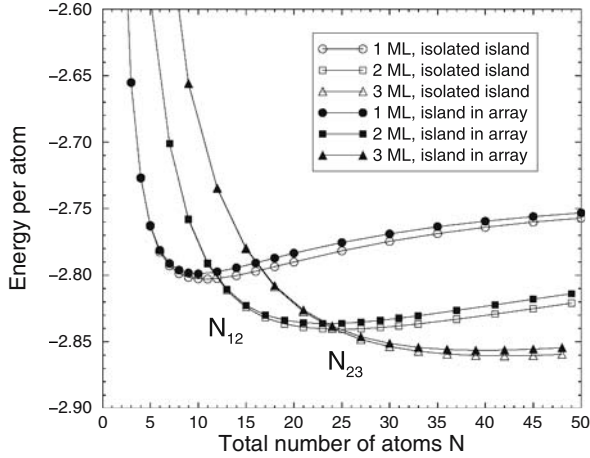


Figure 8. Dependence of the energy per atom on the total number of atoms in compressed coherently strained islands with different thicknesses in monolayers denoted by the figures at each curve for a misfit of 7%. The considered island has two identical neighbours at a distance $n = 5$. The analogous curves for single isolated islands (empty symbols) are also given for comparison. The wetting layer consists in all cases of 3 ML which are allowed to relax.

4. Discussion

For the discussion of the above results we have to bear in mind that a positive wetting parameter shows in fact a tendency of the deposit to form 3D clusters rather than a planar film. In the case of coherent SK growth, the non-zero wetting parameter is due to the weaker adhesion of the atoms that are closer to the islands edges. The presence of other islands, particularly with large angle facets, in the near vicinity of the considered island makes this effect stronger as seen in Fig. 3. The transformation of two-dimensional islands with a monolayer height into bilayer three-dimensional islands takes place by detachment of atoms from the edges and their subsequent jumping and collision on the top island's surface [40]. Thus this edge effect clearly demonstrates the influence of the lattice misfit on the rate of second layer nucleation and in turn on the kinetics of the 2D-3D transformation. citeFil, Lin. The presence of neighboring islands facilitates and thus accelerates the formation of 3D clusters and their further growth. In a self-assembled population of islands the tendency to clustering is thus enhanced.

We can think of the flatter islands in our model (11° facet angle) as the famous “hut” clusters discovered by Mo *et al.*, [20] and of the clusters with 60° facet angles as the “dome” clusters. It is well known that clusters with steeper side walls relieve the strain much more efficiently than the flatter clusters (see the discussion in Ref. [44]; the planar film, which is the limiting case of the flatter islands with a facet angle equal to zero, does not relieve the strain at all.) We see that large-angle facet islands affect more strongly the growth of the neighboring islands, leading thus to a more narrow size distribution.

We further conclude that a self-assembled population of quantum dots is expected at comparatively low temperatures such that the critical wetting-layer thickness for 3D islanding to take place approaches an integer number of monolayers. In InAs/GaAs quantum dots, the reported values of the critical thickness were found to vary from 1.2 to 2 monolayers [45], (see also the discussion in Ref. [46] and references therein). The critical wetting-layer thickness should be given by an integer number of monolayers plus the product of the 2D island density and the critical volume (or area) N_{12} . The 2D island density increases steeply with decreasing temperature [14]. In such a case, a dense population of 2D islands will overcome simultaneously the critical size N_{12} to produce 3D bilayer islands. The latter will interact maximally with each other from the very beginning of the 2D-3D transformation giving rise to a maximum wetting parameter and, in turn, to large-angle facets and a narrow size distribution. This is in agreement with the observations of

Le Tanh *et al.* in the case of nucleation and growth of self-assembled Ge quantum dots on Si(001) [47]. At 700°C, a population of islands with a concentration of the order of 1×10^7 – 1×10^8 cm⁻² is obtained; the islands have the shape of a truncated square pyramid with four side wall facets formed by (105) planes with an inclination angle of about 11° and the size distribution of the islands is quite broad. On the other hand, at 550°C, a population of islands with an areal density of the order of 1×10^9 – 1×10^{10} cm⁻² is observed, the islands have larger angle (113) facets and their size distribution is much more narrow.

In summary, we have shown that the presence of neighboring islands decreases the wetting of the substrate (in this case the wetting layer) by the 3D islands. The larger the density of the array, the weaker the wetting. Neighboring islands with steeper side walls reduce more strongly the wetting of the considered island. The wetting parameter displays a maximum (implying a minimal wetting) when the array shows a monodisperse size distribution. We should expect optimum self-assembled islanding at lower temperatures such that the 2D-3D transformation takes place at the maximum possible island density.

Acknowledgments

J.E.P. gratefully acknowledges financial support from the Alexander-von-Humboldt Stiftung and the Spanish Ministerio de Educación y Cultura (grant No. EX2001 11808094).

References

- [1] P. Politi, G. Grenet, A. Marty, A. Ponchet, and J. Villain. Instabilities in crystal growth by atomic or molecular beams. *Phys. Rep.* **324**: 271–404, 2000.
- [2] D. J. Eaglesham and M. Cerullo. Dislocation-free Stranski–Krastanov growth of Ge on Si(001). *Phys. Rev. Lett.* **64**: 1943–1946, 1990.
- [3] V. A. Shchukin and D. Bimberg. Spontaneous ordering of nanostructures on crystal surfaces. *Rev. Mod. Phys.* **71**: 1125–1171, 1999.
- [4] J. W. Matthews, D. C. Jackson, and A. Chambers. Effect of coherency strain and misfit dislocations on mode of growth of thin films. *Thin Solid Films* **29**: 129–134, 1975.
- [5] D. Leonard, M. Krishnamurthy, C. M. 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**: 3203–3205, 1993.
- [6] J. M. Moison, F. Houzay, F. Barthe, L. Leprince, E. André, and O. Vatel. Self-organized growth of regular nanometer-scale InAs dots on GaAs. *Appl. Phys. Lett.* **64**: 196–198, 1994.

- [7] M. Grundmann, J. Christen, N. N. Ledentsov, J. Böhrer, D. Bimberg, S. S. Ruvimov, P. Werner, U. Richter, U. Gösele, J. Heidenreich, V. M. Ustinov, A. Egorov, A. E. Zhukov, P. S. Kop'ev, and Zh. I. Alferov. Ultra-narrow luminescence lines from single quantum dots. *Phys. Rev. Lett.* **74**: 4043–4046, 1995.
- [8] Z. Jiang, H. Zhu, F. Lu, J. Kin, D. Huang, X. Wang, C. Hu, Y. Chen, Z. Zhu, and T. Yao. Self-organized germanium quantum dots grown by molecular beam epitaxy on Si(100). *Thin Solid Films* **321**: 60–64, 1998.
- [9] Z. Gai, B. Wu, J. P. Pierce, G. A. Farnan, D. Shu, M. Wang, Z. Zhang, and J. Shen. Self-assembly of nanometer-scale magnetic dots with narrow size distribution on an insulating substrate. *Phys. Rev. Lett.* **89**: art. no. 235502, 2002.
- [10] C. Teichert. Self-organization of nanostructures in semiconductor heteroepitaxy. *Phys. Rep.* **365**: 335–432, 2002.
- [11] J. Tersoff and R. M. Tromp. Shape transition in growth of strained islands: Spontaneous transition of quantum wires. *Phys. Rev. Lett.* **70**: 2782–2785, 1993.
- [12] J. Tersoff and F. K. LeGoues. Competing relaxation mechanisms in strained layers. *Phys. Rev. Lett.* **72**, 3570–3573, 1994.
- [13] V. A. Shchukin, N. N. Ledentsov, P. S. Kop'ev, and D. Bimberg. Spontaneous ordering of arrays of coherent strained islands. *Phys. Rev. Lett.* **75**: 2968–2971, 1995.
- [14] B. A. Joyce, J. L. Sudijono, J. G. Belk, H. Yamaguchi, X. M. Zhang, H. T. Dobbs, A. Zangwill, D. D. Vvedensky, and T. S. Jones. A scanning tunneling microscopy–Reflection high energy electron diffraction–Rate equation study of the molecular beam epitaxial growth of InAs on GaAs(001), (110) and (111)A – quantum dots and two-dimensional modes. *Jpn. J. Appl. Phys.* **36**: 4111–4117, 1997.
- [15] H. M. Koduvely and A. Zangwill. Epitaxial growth kinetics with interacting coherent islands. *Phys. Rev. B* **60**: R2204–R2207, 1999.
- [16] C. Priester and M. Lannoo. Origin of self-assembled quantum dots in highly mismatched heteroepitaxy. *Phys. Rev. Lett.* **75**: 93–96, 1995.
- [17] Y. Chen and J. Washburn. Structural transition in large-lattice-mismatch heteroepitaxy. *Phys. Rev. Lett.* **77**: 4046–4049, 1996.
- [18] Y. Ebiko, S. Muto, D. Suzuki, S. Itoh, H. Yamakoshi, K. Shiramine, T. Haga, K. Unno, and M. Ikeda. Scaling properties of InAs/GaAs self-assembled quantum dots. *Phys. Rev. B* **60**: 8234–8237, 1999.
- [19] J. G. Amar and F. Family. Critical cluster size: Island morphology and size distribution in submonolayer epitaxial growth. *Phys. Rev. Lett.* **74**: 2066–2069, 1995.
- [20] E. Korutcheva, A. M. Turiel, and I. Markov. Coherent Stranski–Krastanov in 1 + 1 dimensions with anharmonic interactions: An equilibrium study. *Phys. Rev. B* **61**: 16890–16901, 2000.
- [21] I. Markov and J. E. Prieto. Dislocation-free 3D islands in highly mismatched epitaxy: An equilibrium study with anharmonic interactions. *Atomistic aspects of Epitaxial Growth*, vol. 65, edited by M. Kotrla, N. Papanicolaou, D. D. Vvedensky and L. T. Wille, (Kluwer, Dordrecht, 2002), pp. 411–428.
- [22] K. E. Khor and S. Das Sarma. Quantum dots self-assembly in growth of strained-layer thin films: A kinetic Monte Carlo study. *Phys. Rev. B* **62**: 16657–16664, 2000.

- [23] J. E. Prieto and I. Markov. Thermodynamic driving force of formation of coherent three-dimensional islands in Stranski–Krastanov growth. *Phys. Rev. B* **66**: art. no. 073408, 2002.
- [24] I. Markov. *Crystal Growth for Beginners, Fundamentals of Nucleation, Crystal Growth and Epitaxy*, 2nd edition (World Scientific, Singapore, 2003).
- [25] W. Kossel. *Nachrichten der Gesellschaft der Wissenschaften Göttingen, Mathematisch-Physikalische Klasse*, Band 135, 1927.
- [26] I. N. Stranski. Zur Theorie des Kristallwachstums. *Z. phys. Chem.* **136**: 259–278, 1928.
- [27] R. Peierls. Clustering in adsorbed films. *Phys. Rev. B* **18**: 2013–2015, 1978.
- [28] E. Bauer. Phänomenologische Theorie der Kristallabscheidung an Oberflächen I. *Z. Kristallogr.* **110**: 372–394, 1958.
- [29] F. C. Frank and J. H. van der Merwe. One-dimensional dislocations I. Static theory. *Proc. Roy. Soc. London* **A198**: 205–216, 1949.
- [30] F. C. Frank and J. H. van der Merwe. II. Misfitting monolayers and oriented overgrowth. *Proc. Roy. Soc. London* **A198**: 216–225, 1949.
- [31] L. C. A. Stoop and J. H. van der Merwe. A simple model for layered growth in small epitaxial islands. *Thin Solid Films* **17**: 291–309, 1973.
- [32] C. Ratsch and A. Zangwill. A. Equilibrium theory of the Stranski–Krastanov epitaxial morphology. *Surf. Sci.* **293**: 123–131, 1993.
- [33] J. H. Van der Merwe. (1963) Crystal interfaces. Part I. Semi-infinite crystals. *J. Appl. Phys.* **34**: 117–122, 1963.
- [34] I. Markov. Static multikink solutions in a Frenkel - Kontorova model with anharmonic interactions. *Phys. Rev. B* **48**: 14016–14019, 1993.
- [35] J. Tersoff. New empirical model for the structural properties of silicon. *Phys. Rev. Lett.* **56**: 632–635, 1986.
- [36] O. V. Lysenko, V. S. Stepanyuk, W. Hergert, and J. Kirschner. Mesoscopic relaxation in homoepitaxial metal growth. *Phys. Rev. Lett.* **89**: art. no. 126102, 2002.
- [37] Y. J. Xie, G. H. Gilmer, C. Roland, P. J. Silverman, S. K. Buratto, J. Y. Cheng, E. A. Fitzgerald, A. R. Kortan, S. Schuppler, M. A. Marcus, and P. H. Citrin. Semiconductor surface roughness: Dependence on sign and magnitude of bulk strain. *Phys. Rev. Lett.* **73**: 3006–3009, 1994.
- [38] T. Walther, A. G. Cullis, D. J. Norris, and M. Hopkinson. Nature of the Stranski–Krastanov transition during epitaxy of InGaAs on GaAs. *Phys. Rev. Lett.* **86**: 2381–2384, 2001.
- [39] M. Pinczolis, G. Springholz, and G. Bauer. Direct formation of self-assembled quantum dots under tensile strain by heteroepitaxy of PbSe on PbTe(111). *Appl. Phys. Lett.* **73**: 250–252, 1998.
- [40] S. Stoyanov and I. Markov. On the 2D–3D transition in epitaxial thin film growth. *Surf. Sci.* **116**: 313–337, 1982.
- [41] S. N. Filimonov and Yu. Yu. Hervieu. Statistics of second layer nucleation in heteroepitaxial growth. *Surf. Sci.* **507–510**: 270–275, 2002.

- [42] C. H. Lin and Y.-C. Tsai. A platelet-to-pyramid shape transition under the influence of the adsorbate-substrate interfacial energy. *Surf. Sci.* **512**: 287–299, 2002.
- [43] Y.-W. Mo, D. E. Savage, B. S. Swartzentruber, and M.G. Lagally. Kinetic pathway in Stranski–Krastanov growth of Ge on Si(001). *Phys. Rev. Lett.* **65**: 1020–1023, 1990.
- [44] M. Horn-von Hoegen. Ge on Si(001): A heteroepitaxial playground for surface science. *Surf. Sci.* **537**: 1–3, 2003.
- [45] A. Polimeni, A. Patane, M. Capizzi, F. Martelli, L. Nasi, and G. Salviati. Self-aggregation of quantum dots for very thin InAs layers grown on GaAs. *Phys. Rev. B* **53**: R4213–R4216, 1996.
- [46] L. G. Wang, P. Kratzer, M. Scheffler, and N. Moll. Formation and stability of self-assembled coherent islands in highly mismatched heteroepitaxy. *Phys. Rev. Lett.* **82**: 4042–4045, 1999.
- [47] V. Le Tanh, P. Boucaud, D. Débarre, Y. Zheng, D. Bouchier, and J.-M. Lourtioz. Nucleation and growth of self-assembled Ge/Si(001) quantum dots. *Phys. Rev. B* **58**: 13115–13120, 1998.

THE SEARCH FOR MATERIALS WITH SELF-ASSEMBLING PROPERTIES: THE CASE OF Si-BASED NANOSTRUCTURES

Ilan Goldfarb

Department of Solid Mechanics, Materials and Systems The Fleischman Faculty of Engineering and University Research Institute for Nanoscience and Nanotechnology, Tel Aviv University, Ramat Aviv 69978, Israel

ilang@eng.tau.ac.il

Abstract This work describes an effort to seek for new materials capable of self-assembly of nanostructures, such as dots and wires, for electron-confined devices. A good candidate is a metal-semiconductor compounds group, most notably metal silicides. As will be shown below, the disilicides of cobalt and titanium form distinct nanodot arrays on silicon, of a significantly smaller mean size, and substantially improved size and shape uniformity, as compared to Ge/Si arrays. In order to achieve that, not only the deposition parameters were carefully controlled, but the Si-substrate orientation seemed to play an important role, as well. In the case of CoSi_2 , small and uniform dots resulted from reactive deposition epitaxy at 800 K on Si(001), whereas TiSi_2 dots required Si(111) substrate orientation to form even more uniform, small and isotropic nanodot array. In both cases reacting the metal with silicon in a solid state, and/or on differently oriented substrates did not produce the desired result.

Keywords: Epitaxial silicide nanocrystals, scanning tunneling microscopy and spectroscopy, elastic and surface and interface energies, strain relaxation

1. Introduction

The continuous shrinking of electronic devices inevitably leads into the nanoscale, where quantum effects, e.g. intense photoluminescence [1], single-electron tunneling (SET) and the Coulomb blockade [2, 3], dominate over classical behaviour. However there is at least an order of magnitude gap between the level of miniaturisation that can be reached by optical lithography and the 5–10 nm structures that have their quasi-atomic orbitals sufficiently spread relative to thermal excitation at room-temperature (RT). Even the still-developing lithographic methods, such

as electron-beam, X-ray, and soft-imprint are too crude, while scanning probe lithography is too time-consuming to be suitable for large-area mass-production. In addition, the nanostructures must be different in composition from the surrounding matrix to achieve sufficient band offsets, and they must be uniform in size not to “smear-out” the orbitals of the confined charge carriers.

Stranski-Krastanow (SK) or Volmer-Weber (VW) growth, where nanocrystals self-assemble at the initially 2D wetting layer or directly at the substrate, respectively, provides a more sophisticated approach to fabrication of large-area nanostructure arrays. In this approach the size, shape and uniformity of the nanostructures are controlled by a correct selection of the materials system and the resulting mismatch strain (e.g. by varying x in the $\text{Ge}_x\text{Si}_{1-x}$ alloy), and direct manipulation of growth parameters [4a, 4b, 5], such as growth temperature, deposition rate, etc. So far the main effort has been directed largely towards group IV (i.e. Ge/Si), and III-V compound (i.e. InAs/GaAs) semiconductors, and mainly for optoelectronics [1]. The layer/substrate lattice mismatch, which causes the strain and the respective elastic energy to be stored in heteroepitaxial layers, is one the main factors determining the islanding behaviour. The additional terms that counterbalance the energy are the surface and interface [4a, 4b], as well as edge and corner terms in faceted islands [6, 7]. Unfortunately these energy terms are rarely known. As the nanocrystal volume is inversely proportional to the sixth power of strain, i.e. mean lateral dimension $L \sim \varepsilon^{-2}$ [4a, 4b], one can reduce (increase) the dot size by increasing (decreasing) the mismatch strain. However in the case of the $\text{Ge}_x\text{Si}_{1-x}$ /Si alloy, the minimum dot size cannot be further reduced by varying x , as the maximum 4.2% mismatch can only be achieved with $x = 1$, i.e. by growing pure Ge/Si, which corresponds to about $L = 15$ nm [8]. The same applies to the In(Ga)As/GaAs system, i.e. at 0% Ga, the mismatch is about 7% and diminishes upon Ga increase.

Some metals and/or their silicides can also grow epitaxially on semiconductors, e.g. Ni, Co, Fe, etc. However, smooth monocrystalline growth can sometimes be achieved only with certain substrate orientations, while 3D islanding will result on the others. Metal silicides, Me_xSi_y , have been used in microelectronics as contacts, interconnects, gate and Schottky diode materials [9]. Perhaps because of this reason until recently they have not been considered for nanodevices, although they offer continuous variation of lattice mismatches with Si, ranging from -0.64% for NiSi_2 up to 10% and more for group-IV transition metals. For example, CoSi_2 [10, 11] and TiSi_2 [12–15] are known to grow in a VW fashion on Si, thus providing self-assembled arrays of nanocrystals [10–

15]. Additional degree of freedom in choosing the lattice mismatch can be achieved by reacting a metal not directly with a Si substrate, but with thin Ge/Si layer [16]. Clearly, quantum-size effects need to be demonstrated, and not just favourable structural characteristics. Indeed, SET and the Coulomb staircase have been measured in TiSi_2 nanoislands by Oh *et al.* [3]. Thus, silicides and other metal-semiconductor compounds seem to be strong candidates for nanostructure materials.

2. Experimental Procedures

The experiments were performed in an ultra-high vacuum (UHV) variable-temperature (VT) scanning probe [scanning tunneling (STM) and atomic force (AFM)] microscope, equipped with reflection high-energy electron-diffraction (RHEED) and low-energy electron-diffraction (LEED)/Auger spectrometer, and capable of operation up to 1500 K by direct-current heating. Silicon (001) and vicinal (111) wafers were chemically treated *ex vacuo*, by repeated etch-and-regrowth procedure, to produce clean and homogeneous oxide at the top. In UHV (base pressure 1×10^{-8} Pa), after thorough degassing, this oxide was evaporated by repeated flashes at 1400 K, and the clean Si surface was left to order during a slow cool to the desired temperature, as measured by infrared pyrometer with ± 30 K accuracy. Such treatment has been effective in producing well-ordered (001)- (1×2) and (111)- (7×7) surfaces, as was indeed verified at this time by LEED, RHEED and STM.

Ge hut-cluster nanoislands were grown on Si(001) by gas-source molecular beam epitaxy (GSMBE) from GeH_4 at 630 K. Ge growth was monitored in real time by continuous STM acquisition in a constant-current mode using electrochemically etched W tips, as described elsewhere (cf. Ref. [8]), and terminated when Ge dots attained a dense array configuration, as shown in Fig. 1(a).

CoSi_2 and TiSi_2 were grown by reactive deposition (RDE) and solid-phase (SPE) epitaxy. In the former, the metal (Co or Ti) was deposited from a precise e-beam evaporator onto the Si substrate [(001) for Co, and (111) for Ti] held at 800 K in the VT-STM stage, while scanning, as described in previous publications (cf. Refs. [11, 16]). Hence, the silicide growth was also monitored in real-time. Thereafter the evaporated metal atoms land and react with Si atoms from the substrate to create the respective silicide, while the process is being continuously observed with an STM. This allows for the most intimate insight into the processes of nucleation and subsequent evolution of the 3D nanocrystals. In this case too, the growth was terminated immediately after the appearance of a dense nanodot array, as shown in the upper-right part of Fig. 2.

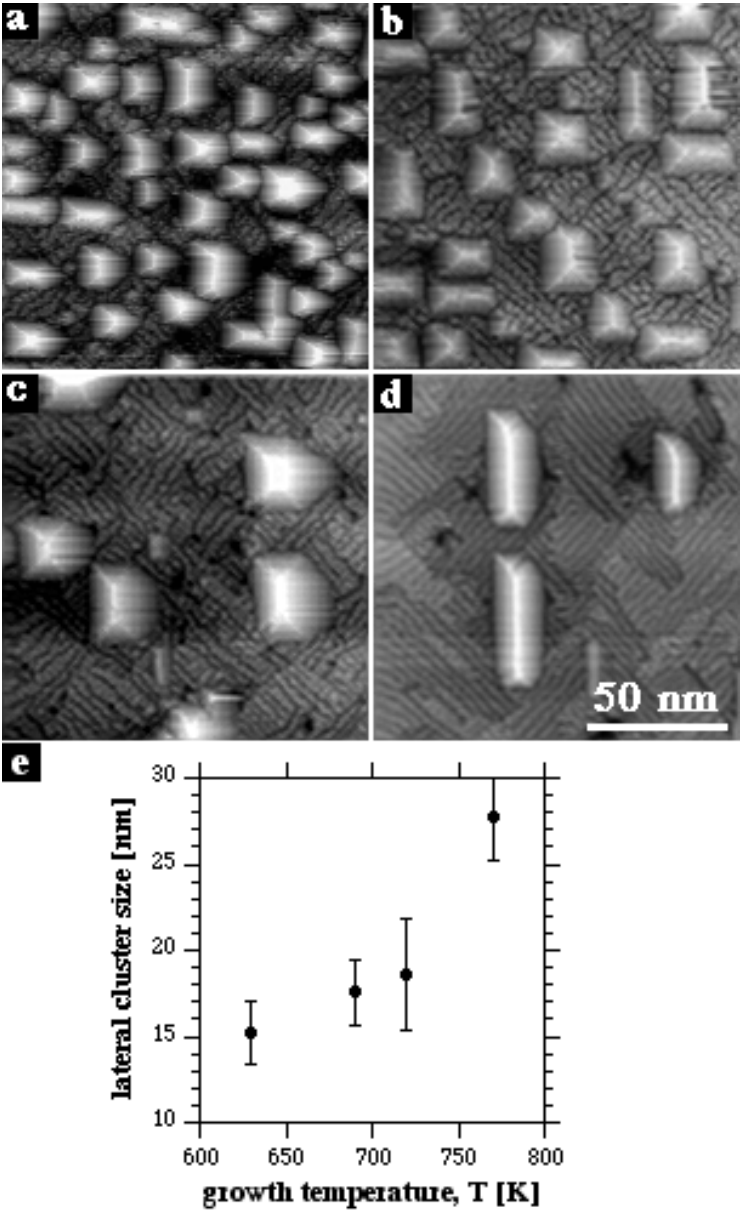


Figure 1. Constant-current STM micrographs of Ge/Si(001) nanodots grown at different temperatures: (a) 630 K, (b) 690 K, (c) 720 K, and (d) 770 K. (e) Dependence of the lateral nanodot dimension (taken as an average of the long and short base sides) on the growth temperature.

In the SPE experiment, the metal was evaporated at RT, and then the resulting metal/silicon system underwent a series of increasingly elevated annealing treatments to promote phase formation (see left-hand side of Fig. 2).

3. Results and Discussion

3.1 Ge Dots

As was already mentioned in the Introduction, one of the important requirements for dots to become useful, quantum ones, is their small size, which is directly linked to mismatch strain as $L \sim \varepsilon^{-2}$. Therefore growing pure Ge on Si represents the highest possible mismatch of about 4.2%, and hence the smallest possible dot size that can be achieved using this tool. Another tool that can be used to tailor the dot size and shape is the growth temperature. Since the dot nucleation rate exponentially-inversely depends on temperature, it might be expected that at higher temperatures the dominant surface process would be growth of already formed dots, rather than continuous nucleation of the new ones. Therefore, as the temperature is increased, the mean dot size increases as well, at the expense of dot number density, as indeed can be seen when going from 630 K [Fig. 1(a)] to 770 K [Fig. 1(d)]. The variation of the Ge dot mean size is displayed graphically in Fig. 1(e). Furthermore, not only the size of the dot increases with temperature, but also its shape anisotropy, due to hut-shape instabilities [5, 8], that may ultimately lead to strikingly elongated shapes shown in Fig. 1(d). Thus, small dot size and uniformity require Ge growth at the lowest possible temperature, which is about 630 K in GSMBE (below that, growth is impossible due to the hydrogen blocking of surface diffusion [17]).

3.2 CoSi₂ Dots

The growth of CoSi₂ dots is only possible on a Si(001) surface and only by RDE. On the Si(111) surface CoSi₂ grows in a smooth continuous fashion, as indeed could have been expected in view of a rather low lattice mismatch with Si (about -1.2%). It has been known for some time that small, even though somewhat elongated, nanodots can be produced when growing the silicide on the Si(001) surface [10]. However, it turns out that merely choosing the substrate orientation is not enough: it is also important how the silicide is actually grown. For example, it was found that three-dimensional CoSi₂ nanodots form only during reactive deposition (at about 800 K [11]), while two-dimensional layers, even though rather rough on the atomic scale, result when the Co-Si reaction

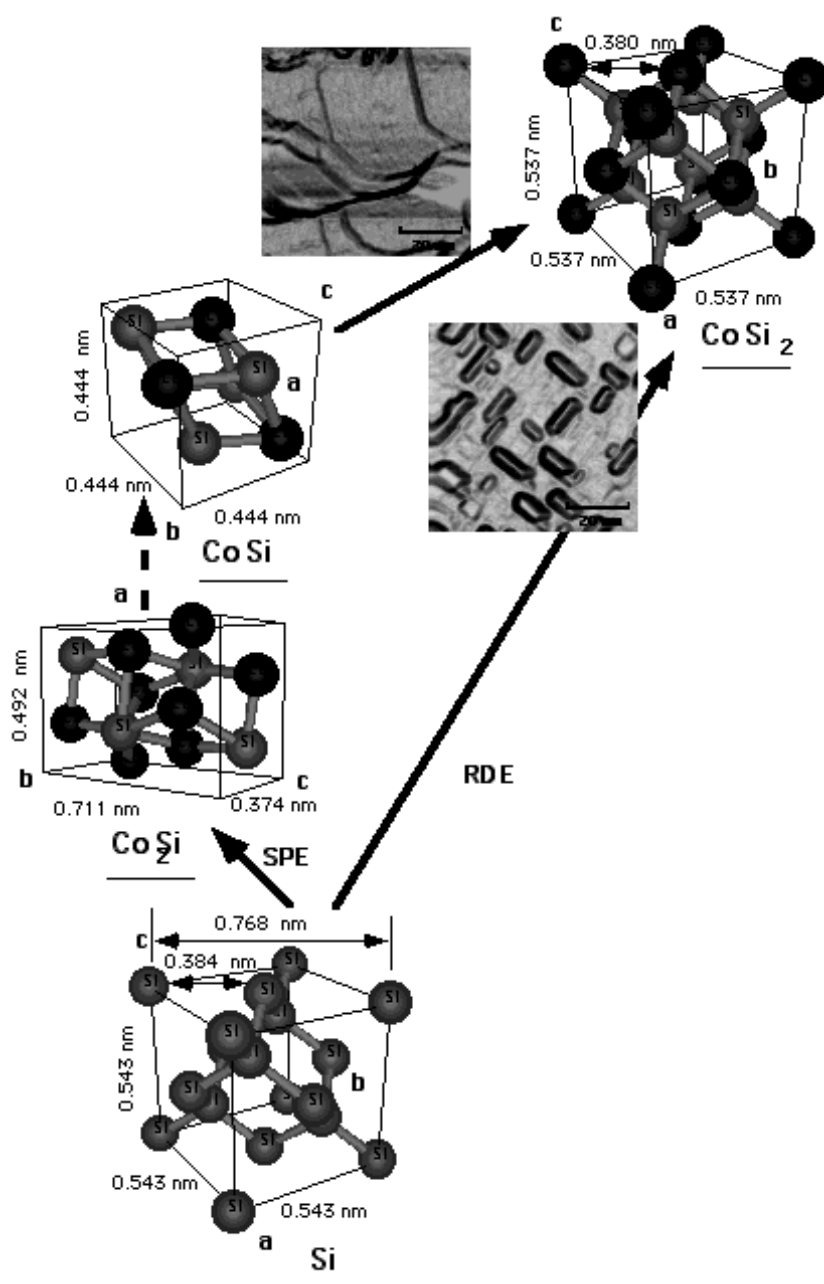


Figure 2. Schematic of possible pathways of phase formation in RDE (right-hand part and upper inset) and SPE (left-hand part and lower inset) silicide growth in a Co-Si system. The STM image insets were filtered for enhancement (scale bars in both images are 20 nm long).

takes place in the solid state [18]. This morphological variation can be explained by the different kinetic pathways of phase-formation dominant in each process, as shown in Fig. 2. Typical RDE-grown CoSi_2 nanodots are compared with the Ge ones in Fig. 3(a,b), and their respective size distributions in Fig. 1(d,e). In RDE, the metal atoms land on a hot substrate and react immediately to form the silicide, metal accumulation at the growing metal/silicon interface is avoided, and the first silicide phase formed is relatively Si-rich, i.e. CoSi_2 . In SPE, on the other hand, the unreacted metal (Co in this case) is first accumulated at the Co/Si

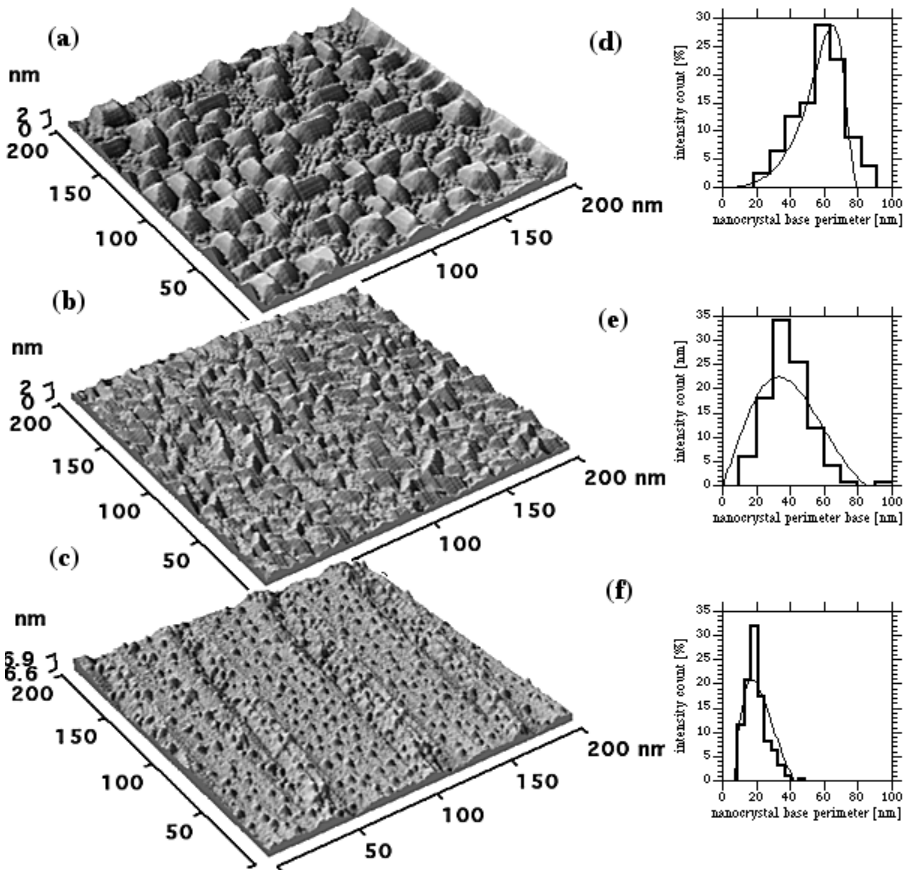


Figure 3. Comparison between the appearance and statistical characteristics of the Ge/Si(001), $\text{CoSi}_2/\text{Si}(001)$, and $\text{TiSi}_2/\text{Si}(111)$ dot arrays. (a)-(c) 3D representation of constant-current STM micrographs, and (d)-(f) their respective size distributions. Note the significant shift towards smaller size and narrower distribution when going from Ge/Si(001) (d), through $\text{CoSi}_2/\text{Si}(001)$ (e), towards $\text{TiSi}_2/\text{Si}(111)$ (f).

interface, and when annealed the first phase formed is Co-rich, e.g. Co_2Si , then perhaps also CoSi , before terminating at the disilicide, CoSi_2 . In the process the coherency with the substrate is lost, consequently the CoSi_2 layer is no longer strained, and hence it does not require strain relaxation by three-dimensional islands, where a flat morphology appears instead.

3.3 TiSi_2 dots

The phenomena of TiSi_2 dot growth is on the one hand similar to that of the CoSi_2 dots since, in order to obtain them, RDE rather than SPE must be used, but different on the other, in the sense that the more successful TiSi_2 dots seem to grow on $\text{Si}(111)$ rather than on $\text{Si}(001)$ surfaces. RDE and chemical vapour deposition (CVD) of Ti onto $\text{Si}(001)$ preheated to about 900 K, seem to produce rather large and irregularly shaped islands of C49- TiSi_2 (apparently transformed into C54- TiSi_2 after $T > 1100$ K anneals), incommensurate with the underlying substrate [14, 15]. The lack of coherency with the substrate is scarcely surprising, in view of very large nominal mismatch values with silicon. SPE on the both (001)- and (111)-oriented Si substrates does not seem to yield better results [12, 13]. Hence very small TiSi_2 dots of uniform equiaxed shape, resulting from RDE growth at 800 K on $\text{Si}(111)$ in this study, as shown in Fig. 3(c,f), are certainly nicely surprising.

4. Summary and Conclusions

The experiments in this work were aimed at growing various types of nanodot in a controllable manner, using the nominal layer/substrate lattice mismatch and substrate orientation as tools for producing the conditions for self-assembled nanodot formation. Two substrate orientations, (001) and (111), and three materials systems, $\text{Ge}/\text{Si}(001)$, $\text{CoSi}_2/\text{Si}(001)$, and $\text{TiSi}_2/\text{Si}(111)$, were chosen to represent the entire range of mismatches suitable for self-assembly. Ge/Si is the most studied and known system that represents a moderate mismatch, a classical case of self-assembled SK dot growth. CoSi_2/Si represents the lower end of the allowed mismatch range that, obviously, is still sufficient to induce VW self-assembled dot growth, and TiSi_2/Si stands for a very high mismatch, perhaps too high for coherent growth. However the nominal mismatch is not a precise enough instrument, because even with a large nominal mismatch a crystal can still be matched to a certain atomic plane of a substrate, provided there is a close similarity in symmetry and interatomic spacings in the particular crystallographic directions of that plane. Hence, the conditions for dot self-assembly are apparently

met in the case of $\text{CoSi}_2/\text{Si}(001)$ and $\text{TiSi}_2/\text{Si}(111)$, provided the growth method is suitable, as well. While $\text{Ge}/\text{Si}(001)$ self-assembles in the SK mode independently of the growth method (even though the kinetics may slightly vary between MBE and GSMBE, or in surfactant-mediated growth), the way to grow silicide dots seem to be solely RDE.

The comparison between the three above systems, shown in Fig. 3, is striking. Not only the mean CoSi_2 -dot size is merely two-thirds of the Ge-dot size, and TiSi_2 -dot size is even one-third of it (taking the dot base perimeter as a measure of a size), but the uniformity of the silicide dots (in particular TiSi_2) is superior by a far, as can be judged from the size distributions presented in Fig. 3(d,e,f)! These findings are very important for the fundamental understanding of oriented growth and self-assembled behaviour. However, for the more practical applications in devices, knowledge of the individual and collective electronic properties of the dot arrays is a must. The work on measuring some of these properties, e.g. probing the electronic states in scanning tunneling I-V spectra, is still in progress, however SET manifested through the Coulomb blockade in TiSi_2 dots of a lesser than here quality has already been demonstrated [3]. The results presented here are encouraging and facilitate further research into both growth behaviour and the corresponding physical properties of silicide nanodots.

Acknowledgements

The author wishes to thank the Israel Science Foundation (ISF GR 9043/00) for its generous contribution towards the equipment costs. The help of S. Grossman, G. Cohen-Taguri, and M. Levinshtein in this work is also gratefully acknowledged.

References

- [1] K. Eberl, M. O. Lipinski, Y. M. Manz, W. Winter, and N. Y. Jin-Phillipp, O. G. Schmidt. Self-assembling quantum dots for optoelectronic devices on Si and GaAs. *Physica E* **9**: 164–174, 2001.
- [2] T. J. Thornton. Mesoscopic devices. *Rep. Prog. Phys.* **58**: 311–364, 1995.
- [3] J. Oh, V. Meunier, H. Ham, and R. J. Nemanich. Single electron tunneling of nanoscale TiSi_2 islands on Si. *J. Appl. Phys.* **92**: 3332–3337, 2002.
- [4] J. Tersoff and F. K. LeGoues. Competing relaxation mechanisms in strained layers. *Phys. Rev. Lett.* **72**: 3570–3573, 1994.
- [5] J. Tersoff and R. M. Tromp. Shape transition in growth of strained islands: Spontaneous formation of quantum wires. *ibid.* **70**: 2782–2785, 1993.
- [6] D. E. Jesson, K. M. Chen and S. J. Pennycook. Kinetic pathways to strain relaxation in the Si-Ge system. *MRS Bull.* **21**(4): 31–37, 1996.

- [7] V. A. Shchukin, N. N. Ledentsov, P. S. Kopev, and D. Bimberg. Spontaneous ordering of arrays of coherent strained islands. *Phys. Rev. Lett.* **75**: 2968–2971, 1995.
- [8] G. Medeiros-Ribeiro, A. M. Bratkovski, T. I. Kamins, D. A. A. Ohlberg, and R. S. Williams. Shape transition of germanium nanocrystals on a silicon (001) surface from pyramids to domes. *Science* **279**: 353–355, 1998.
- [9] I. Goldfarb, P. T. Hayden, J. H. G. Owen, and G. A. D. Briggs. Competing growth mechanisms of Ge/Si(001) coherent clusters. *Phys. Rev. B* **56**: 10459–10468, 1997.
- [10] S. P. Murarka. *Silicides for VLSI Applications* (Academic Press, New York, 1983).
- [11] V. Scheuch, B. Voigtländer, and H. P. Bonzel. Nucleation and growth of CoSi₂ on Si(100) studied by scanning tunneling microscopy. *Surf. Sci.* **372**: 71–82, 1997.
- [12] I. Goldfarb and G. A. D. Briggs. Reactive deposition epitaxy of CoSi₂ nanostructures on Si(001): Nucleation and growth and evolution of dots during anneal. *Phys. Rev. B* **60**: 4800–4809, 1999.
- [13] A. W. Stephenson and M. E. Welland. Scanning tunneling microscope crystallography of titanium silicide on Si(100) substrates. *J. Appl. Phys.* **77**: 563–571, 1995.
- [14] A. W. Stephenson and M. E. Welland. Scanning tunneling microscope investigation of the growth morphology of titanium silicide on Si(111) substrates. *J. Appl. Phys.* **78**: 5143–5154, 1995.
- [15] G. Medeiros-Ribeiro, D. A. A. Ohlberg, D. R. Bowler, R. E. Tanner, G. A. D. Briggs, and R. S. Williams. Titanium disilicide nanostructures: two phases and their surfaces. *Surf. Sci.* **431**: 116–127, 1999.
- [16] G. A. D. Briggs, D. P. Basile, G. Medeiros-Ribeiro, T. I. Kamins, D. A. A. Ohlberg, and R. S. Williams. The incommensurate nature of epitaxial titanium disilicide islands on Si(001). *Surf. Sci.* **457**: 147–156, 2000.
- [17] I. Goldfarb and G. A. D. Briggs. Surface studies of phase formation in Co-Ge system: Reactive deposition epitaxy versus solid-phase epitaxy. *J. Mater. Res.* **16**: 744–752, 2001.
- [18] I. Goldfarb, J. H. G. Owen, P. T. Hayden, D. R. Bowler, K. Miki, and G. A. D. Briggs. Gas-source growth of group-IV semiconductors: III. Nucleation and growth of Ge/Si(001). *Surf. Sci.* **394**: 105–118, 1997.
- [19] I. Goldfarb and G. A. D. Briggs. Morphological evolution of epitaxial cobalt-semiconductor compound layers during growth in a scanning tunneling microscope. *J. Vac. Sci. Technol. B* **20**: 1419–1426, 2002.

X-RAY SCATTERING METHODS FOR THE STUDY OF EPITAXIAL SELF-ASSEMBLED QUANTUM DOTS

J. Stangl, T. Schüllli*, A. Hesse, and G. Bauer

Institute of Semiconductor Physics, Johannes Kepler University Linz, Austria
guenther.bauer@jku.at

V. Holý

Institute of Condensed Matter Physics, Masaryk University Brno, Czech Republic
holy@physics.muni.cz

Abstract Several x-ray diffraction methods are presented to determine the local chemical composition of self-assembled islands and to discriminate them from strain gradients. Two different routes are followed: in the first approach, the scattered intensities are simulated using numerical fitting to a suitable structure model for the islands (indirect methods). In the second one, the structural data are directly derived from the experimental ones. This direct approach is based on the so-called iso-strain scattering method and/or on the anomalous diffraction technique, which uses the strong enhancement and suppression of the scattered intensity close to the absorption edge of one of the chemical elements in the island. We show that for Ge dome-shaped islands the different techniques give similar results.

1. Introduction

The electronic and optical properties of self-assembled semiconductor quantum dots are substantially affected by their structural parameters, especially by their shape, chemical composition and by the inhomogeneous elastic strain [1, 2, 3]. The shape of uncapped quantum dots can be investigated non-destructively by atomic force microscopy (AFM),

*
tale, CEA, Grenoble, France

épartement de Recherche Fondamen-

for quantum dots buried below the surface plan-view or cross-sectional transmission electron microscopy (TEM) is often used. Chemical composition and elastic strain can also be determined by TEM using a Z -sensitive mode or a numerical analysis of the TEM micrographs[4].

Transmission electron microscopy requires a complicated and very laborious sample preparation. X-ray scattering has been employed to determine the shape and chemical composition of uncapped and buried islands, as well [5, 6, 7, 8, 9]. This method is non-destructive and rather simple experimentally, however, in it is *nonlocal*, as it probes the distribution of scattered intensity in *reciprocal* space instead of imaging real space like TEM. For the quantitative analysis of the scattering data a structure model of the dot must be assumed *a priori*, the model parameters can be obtained indirectly by a numerical fitting of the measured intensity distribution in reciprocal space to the intensity distribution simulated on the basis of the model [7, 8].

A particular advantage of x-ray techniques is that due to the rather large illuminated area, typically many thousands of islands contribute to the measured signal, hence a rather high statistical relevance is obtained. The techniques are based on high resolution in reciprocal space, obtained by small divergence and small wavelength spread of the primary beam, and a high angular resolution of the set-up, typically in the range of few arcsec. Due to the small scattering volume of islands, high brilliance, as provided by synchrotron radiation sources, is required. So far, despite recent progress with microbeam installations, the spatial resolution of the x-ray scattering set-ups is not sufficient to probe individual islands.

Shape, size and positional correlation of nanostructures are obtained by x-ray reflectivity (XRR) [10] and grazing incidence small angle scattering (GISAXS) measurements [9, 11, 12, 13]. Both techniques probe the reciprocal space close to its center (000) and thus the scattered signal depends on the electron density averaged over a unit cell. Consequently, the scattered intensities are nearly insensitive to strain fields. Composition and strain are obtained from x-ray diffraction techniques both in conventional x-ray diffraction (XRD) [6, 7] as well as in grazing incidence geometry (GID) [5, 14]. In XRD, information on in-plane strains as well as on strains parallel to the surface normal is obtained. An advantage of the grazing incidence geometry is the enhanced discrimination with respect to signals originating from the substrate: Here, both the incidence and the exit angles with respect to the sample surface are small, allowing for a tuning of the penetration depth of the x-rays into the sample. If these angles are kept below the critical angle of total external reflection, the penetration depth is of the order of several nm only, and increases to several micrometers for larger angles of incidence/exit.

The diffraction process in GID takes place at lattice planes perpendicular to the sample surface. Therefore only in-plane strain components can be investigated. In order to obtain information on the strain component along growth direction, but keeping the surface sensitivity, a scattering geometry with small incidence but large exit angle can be applied.

In principle, a direct determination of the island structure from the measured intensities is not possible, since the phase of the scattered radiation is not recorded. However, several approaches have been published recently promising a *direct* determination of the shape and chemical composition of quantum dots from their scattering data. In the coherent scattering method suggested by Williams et al. [15] a highly coherent primary x-ray beam irradiates one single nano-object and the phase of the measured intensity can be recovered by a numerical Gerchberg-Saxton algorithm [16]. Up to now, this method has been used only in a small-angle scattering geometry for the determination of the shape of Au-particles and its application for x-ray diffraction has not been published yet.

The direct methods described below use the measured intensity data without phase recovery. Instead, in order to distinguish the influence of strain from the shape-dependent scattering, one has to assume that the scattering objects are sufficiently large and/or the strain gradient is large enough. Under these assumptions, the shape, strain, and chemical composition of uncapped quantum dots are determined using the "iso-strain scattering" method [17, 18]. The inhomogeneous chemical composition of quantum dots can also be obtained directly by anomalous x-ray scattering using a pair of wavelengths close to and far from the absorption edge of a chosen element present in the dot lattice [19, 20, 21].

In this paper we compare indirect and direct methods and we demonstrate their applicability on SiGe islands as an example. We restrict ourselves to uncapped islands, for which the direct methods are applicable.

2. Kinematical Scattering from Islands

A typical island is much smaller than the extinction length of x-rays in matter, therefore the diffraction in the island can be described using kinematical approximation (the first Born approximation) [10]. Using this assumption, the amplitude of the plane-wave component of the scattered beam having the wave vector $\mathbf{K}_f$ is proportional to the Fourier transformation of the electron density with the argument $\mathbf{Q} = \mathbf{K}_f - \mathbf{K}_i$, where $\mathbf{K}_i$ is the wave vector of the primary x-ray beam (assumed per-

fectly monochromatic and plane) and $\mathbf{Q}$ is the scattering vector. If we assume that $\mathbf{Q}$ is close to a given vector $\mathbf{h}$ of the reciprocal lattice (diffraction vector), we use the two-beam approximation, in which the amplitude of the diffracted beam is proportional to the structure factor of an island [10]

$$F_{\mathbf{h}}(\mathbf{q}) = \int d^3\mathbf{r} \chi_{\mathbf{h}}(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} e^{-i\mathbf{h}\cdot\mathbf{u}(\mathbf{r})}, \quad (1)$$

where $\mathbf{q} = \mathbf{Q} - \mathbf{h}$ is the reduced scattering vector, $\mathbf{u}(\mathbf{r})$ is the displacement field in the dot and its neighborhood and $\chi_{\mathbf{h}}(\mathbf{r})$ is the $\mathbf{h}$ -th Fourier component of the crystal polarizability proportional to the complex atomic scattering factors $f(\mathbf{h})$ of the atoms constituting the unit cell. The scattering factor is a sum of three components $f(\mathbf{h}) = f_0(\mathbf{h}) + f' + if''$, where $f_0(\mathbf{h})$ is the classical scattering factor proportional to the electron density, this factor decreases with increasing $|\mathbf{h}|$. The other two components represent the real and imaginary dispersion corrections; these corrections sensitively depend on the x-ray energy and they do not depend on $\mathbf{h}$. The position dependence of $\chi_{\mathbf{h}}$ expresses the inhomogeneity of the chemical composition of the island and the island shape, in vacuum $\chi_{\mathbf{h}}(\mathbf{r}) = 0$ holds. Equation (1) was derived assuming that $\varepsilon_{jk} = \partial u_j / \partial x_k \ll 1$ for all the components of the strain tensor.

Equation (1) can be simplified if one assumes that the island shape and the chemical composition exhibit a cylindrical symmetry with respect to the vertical axis. If, in addition, one neglects the elastic anisotropy of the crystal, the displacement field $\mathbf{u}(\mathbf{r})$ is cylindrically symmetric as well. In the *coplanar arrangement*, the diffraction vector $\mathbf{h}$ and the reduced scattering vector $\mathbf{q}$ lie in the same plane perpendicular to the surface (the scattering plane). Then, using the symmetry simplification above, we obtain

$$F_{\mathbf{h}}(\mathbf{q}) = 2\pi \int_0^\infty d\rho \rho \int_{-\infty}^\infty dz e^{-iq_z z - ih_z u_z(\rho, z)} J_0(q_{\parallel} \rho + h_{\parallel} u_{\parallel}(\rho, z)) \left[H(z) \chi_{\mathbf{h}}^{(\text{island})}(\rho, z) + H(-z) \chi_{\mathbf{h}}^{(\text{sub})} \right]. \quad (2)$$

Here, ρ and z are the cylindrical coordinates, $h_{\parallel, z}$ and $u_{\parallel, z}$ are the in-plane and z coordinates of the diffraction vector and the displacement, respectively. J_0 is the Bessel function of the zero-th order, $H(z)$ is the Heaviside step function (unity for $z \geq 0$ and zero for $z < 0$). The first term in the square brackets represents the scattering from the island itself, the second term describes the scattering from the distorted part of the substrate under the island. A similar expression for the scattering factor can be derived, if the in-plane components $\mathbf{q}_{\parallel}$ and $\mathbf{h}_{\parallel}$ are mutually

perpendicular:

$$F_{\mathbf{h}}(\mathbf{q}) = 2\pi \int_0^\infty d\varrho \varrho \int_{-\infty}^\infty dz e^{-iq_z z - ih_z u_z(\varrho, z)} J_0 \left(\sqrt{(q_{\parallel} \varrho)^2 + (h_{\parallel} u_{\parallel}(\varrho, z))^2} \right) \\ \times \left[H(z) \chi_{\mathbf{h}}^{(\text{island})}(\varrho, z) + H(-z) \chi_{\mathbf{h}}^{(\text{sub})} \right]. \quad (3)$$

In the kinematical approximation, the intensity of the plane wave component of the diffracted beam with the wave vector $\mathbf{K}_f = \mathbf{K}_i + \mathbf{h} + \mathbf{q}$ is given by the following approximation

$$I_{\mathbf{h}}(\mathbf{q}) \approx A I_i |F_{\mathbf{h}}(\mathbf{q})|^2 G(\mathbf{q}), \quad (4)$$

where A is a constant and

$$G(\mathbf{q}) = \left\langle \sum_{m,n} e^{-i\mathbf{q} \cdot (\mathbf{R}_m - \mathbf{R}_n)} \right\rangle$$

is the correlation function of the island positions, $\mathbf{R}_m$ is the random position vector of the island m , the averaging $\langle \cdot \rangle$ is performed over all configurations of the island positions. The approximation in Eq. (4) holds if the strain fields of different islands do not overlap and/or the positions of the islands are not strongly correlated. If the island positions are completely uncorrelated and their density at the sample surface is rather small, the correlation function G is constant and simply equals the total number N of irradiated islands. This simplified situation will be assumed in the following. Correlated positions of islands have been investigated for instance in Ref. [22].

If the islands are placed on a free surface, they are irradiated not only by the primary x-ray beam itself, but also by the beam specularly reflected from the flat sample surface. Similarly, a part of the measured diffracted beam is specularly reflected from the flat surface after being scattered by the islands. If the angles of incidence and/or exit are small, these processes cannot be neglected and the measured intensity of the diffracted beam is a coherent superposition of four scattering processes [9], as sketched in Fig. 1:

$$I_{\mathbf{h}}(\mathbf{q}) = AN I_i \left| F_{\mathbf{h}}(\mathbf{q}) + r_i F_{\mathbf{h}}(\mathbf{q}') + r_f F_{\mathbf{h}}(\mathbf{q}'') + r_i r_f F_{\mathbf{h}}(\mathbf{q}''') \right|^2. \quad (5)$$

The scattering vectors of these processes are:

$$\mathbf{q} = \mathbf{K}_f - \mathbf{K}_i - \mathbf{h}, \quad \mathbf{q}' = \mathbf{K}_f - \mathbf{K}_{iR} - \mathbf{h}, \\ \mathbf{q}'' = \mathbf{K}_{fR} - \mathbf{K}_i - \mathbf{h}, \quad \mathbf{q}''' = \mathbf{K}_{fR} - \mathbf{K}_{iR} - \mathbf{h}, \quad (6)$$

$r_{i,f}$ are the complex Fresnel reflection amplitudes for the primary waves $\mathbf{K}_{i,f}$, $\mathbf{K}_{iR}$ and $\mathbf{K}_{fR}$ are the wave vectors of the reflected waves belonging

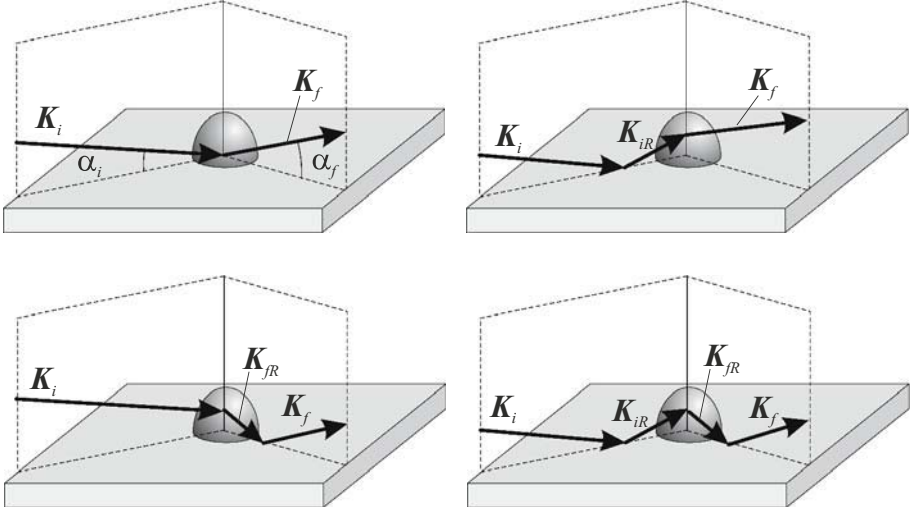


Figure 1. Illustration of the four scattering processes for an uncapped quantum island.

to the primary waves $\mathbf{K}_i$ and $\mathbf{K}_f$, respectively. The z -components of the vectors $\mathbf{K}_{i,fR}$ perpendicular to the sample surface have opposite signs, i.e. $K_{i,fRz} = -K_{i,fz}$. Equation (5) can be derived using the so called Distorted-Wave Born Approximation (DWBA) [9, 23].

Even in the case of uncapped islands, the scattering originates not only from the island volume, but also from the distorted part of the substrate under the island (the second term in the square brackets in Eqs. (1,2)). The buried part of the sample is irradiated by the *transmitted* waves and not by the reflected ones, and the corresponding intensity is calculated using the kinematical expression in Eq. (4), but taking refraction into account and multiplying with the factor $|t_i t_f|^2$. $t_{i,f}$ are the Fresnel transmittivities of the free surface belonging to the primary and scattered beams, respectively.

Using the indirect method, one has to assume both, a certain chemical composition profile in the island, and the island shape; these assumptions determine the function $\chi_{\mathbf{h}}(\mathbf{r})$. From the composition profile the elastic displacement field $\mathbf{u}(\mathbf{r})$ can be calculated using either a continuum elasticity approach or a suitable atomistic method, see Ref. [24] and citations therein. Both methods are based on a numerical minimization of the elastic energy expressed either by continuum displacements and macroscopic elastic constants in the former case, or by atomic displacement and suitable inter-atomic potentials in the latter case.

The x-ray data are analyzed using an iterative "fitting" procedure: starting from assumptions on the shape and Ge distribution of the islands, we calculate the displacement fields $\mathbf{u}(\mathbf{r})$ using a commercial finite element method (FEM) package. For this purpose, the island and the surrounding matrix are modelled using a three-dimensional grid with several 10 000 nodes. The grid is very fine within the islands and in its close proximity, and coarser with increasing distance from the island. For cubic semiconductors with the surface oriented along [001] direction, it is sufficient to simulate a 45° wedge limited laterally by a $\langle 100 \rangle$ and a $\langle 110 \rangle$ plane, and use the following boundary conditions: (i) nodes in the side planes (including the outer circumference) are allowed to move only within the plane, but not perpendicular to it. (ii) nodes at the bottom of the simulation cell are fixed. The size of the simulation cell is increased until no further influence on the resulting strain fields from the size is detected. The elastic properties of the materials, including the elastic anisotropy, are included in the simulations. For SiGe alloys, the values of pure Si and Ge are linearly interpolated.

Having calculated the displacement field, the diffracted intensity is obtained by a numerical Fourier transformation according to Eqs. (1,4) or (1,5). The island shape and the chemical composition distribution are changed iteratively, until a reasonable correspondence between experiment and simulation is achieved.

3. The Direct Methods

The main problem in a direct determination of the shape and strain profile in an island is to distinguish the contribution of the chemical contrast caused by $\chi_{\mathbf{h}}(\mathbf{r})$ from the strain contrast due to $\exp(-i\mathbf{q} \cdot \mathbf{u}(\mathbf{r}))$. Both factors enter the Fourier transformation in Eq. (1) for the structure factor in a product.

Let us define the *iso-strain volume* as the volume of the part of the island, where the values ε_{jk} of the strain tensor can be assumed nearly constant. The whole island volume can be expressed as a set of these iso-strain volumes defined for various strains (Fig. 2). Each iso-strain volume contributes to the total intensity distribution $I_{\mathbf{h}}(\mathbf{q})$ by a partial intensity distribution $I_{\mathbf{h}}^{(\varepsilon)}(\mathbf{q})$, with a maximum in the point $\mathbf{q}_{\max}^{(\varepsilon)} = -\hat{\varepsilon}\mathbf{h}$. For many semiconductor islands, the shape of the iso-strain volumes can be approximated by rather flat disks with diameters $L^{(\varepsilon)}$. Then the lateral width of the central peak in reciprocal space is inversely proportional to $L^{(\varepsilon)}$: $\Delta q^{(\varepsilon)} = 2\pi/L^{(\varepsilon)}$. The contributions of two iso-strain volumes with strains ε and $\varepsilon + \Delta\varepsilon$ can be distinguished, if their

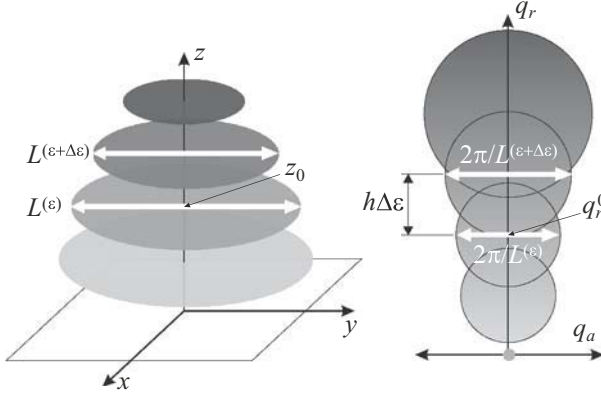


Figure 2. Sketch of the iso-strain areas in an uncapped island and the corresponding intensity distribution in reciprocal space.

contributions $I_{\mathbf{h}}^{(\varepsilon)}(\mathbf{q})$ and $I_{\mathbf{h}}^{(\varepsilon+\Delta\varepsilon)}(\mathbf{q})$ do not overlap, i.e. if

$$\frac{2\pi}{L(\varepsilon)} \ll h\Delta\varepsilon.$$

If, for instance, the strain gradient in the island is oriented mainly parallel to the vertical axis z , the shape contrast and the strain can be distinguished if

$$HLh \frac{d\varepsilon}{dz} \gg 1, \quad (7)$$

where H is the island height and L is the island width.

If condition (7) is fulfilled, the integral in Eq. (1) for the structure factor of an island can be calculated using the stationary-phase method [25]. For the sake of simplicity we assume a symmetric GID geometry, where the diffraction vector $\mathbf{h}$ is parallel to the x -axis lying in the sample surface, the strain changes mainly along the growth direction (parallel to the z -axis), and we neglect the shear terms of the strain tensor ε_{jk} , $j \neq k$. Then, the structure factor can be approximated by

$$F_{\mathbf{h}}(\mathbf{q}) \approx \frac{2\pi}{h} \left[\frac{d\varepsilon_{xx}}{dz} \Big|_{x=x_0, z=z_0} \right]^{-1} e^{-iq_z z_0} \int_{-L(x_0, z_0)/2}^{L(x_0, z_0)/2} dy \chi_{\mathbf{h}}(x_0, y, z_0) e^{-iq_y y}. \quad (8)$$

Here we have denoted (x_0, z_0) the coordinates of the stationary point in the xz -plane obeying the following expressions:

$$q_x + h\varepsilon_{xx}(x_0, z_0) = 0, \quad q_z + hx_0 \frac{d\varepsilon_{xx}}{dz} \Big|_{x=x_0, z=z_0} = 0. \quad (9)$$

In Eq. (8), $L(x_0, z_0)$ is the island width along y through the stationary point (x_0, z_0) . Equation (8) can be further simplified, if we neglect a lateral chemical composition gradient, and consequently the polarizability $\chi_{\mathbf{h}}$ does not depend on y :

$$F_{\mathbf{h}}(\mathbf{q}) \approx \frac{2\pi}{h} L(x_0, z_0) \chi_{\mathbf{h}}(x_0, z_0) \text{sinc}\left[\frac{1}{2}L(x_0, z_0)q_y\right] \times \left[\frac{d\varepsilon_{xx}}{dz} \Big|_{x=x_0, z=z_0} \right]^{-1} e^{-iq_z z_0}, \quad (10)$$

where $\text{sinc}(x) = \sin(x)/x$. The stationary phase method can be used only if only *one* stationary point exists for a given $\mathbf{q}$. This is certainly the case if the in-plane strain varies monotonic with height.

Recently, several publications demonstrated the existence of lateral composition gradients [4, 26, 27]. From Eq. (8) it is obvious that the structure factor of a laterally inhomogeneous island is proportional to the Fourier transformation

$$\int_{-\infty}^{\infty} dy \chi_{\mathbf{h}}(x, y, z) e^{-iq_y y} \Big|_{x=x_0, z=z_0},$$

i.e. the chemical composition can be determined from the intensity distribution along q_y in reciprocal space, referred to as "angular scans" in GID geometry.

Equation (10) makes it possible to reconstruct the mean shape of the quantum dots from the measured distribution $I_{\mathbf{h}}(\mathbf{q})$. The contribution to the scattered intensity in a point $\mathbf{q} = (q_x, q_y, q_z)$ depends on the size of the iso-strain volume, having the strain component $\varepsilon_{xx} = -q_x/h$. The intensity distribution along q_y is determined by the function $\text{sinc}[L(x_0, z_0)q_y]$, from its width the size $L^{(\varepsilon)} \equiv L(x_0, z_0)$ can be determined.

The vertical position z_0 of the iso-strain volume affects the phase factor $\exp(-iq_z z_0)$. Within the simple kinematical approximation as in Eq. (4), the phase factor vanishes and z_0 cannot be determined. However, this factor affects the measured intensity for uncapped islands within DWBA: Since all scattering processes included in Eq. (5) are added coherently, the resulting scattered intensity depends on the phase $-iq_z z_0$. Thus, from an intensity distribution $I_{\mathbf{h}}(\mathbf{q})$ measured in a three-dimensional region in reciprocal space it is possible to determine the widths $L^{(\varepsilon)}$ of a certain iso-strain volume in y -direction, and the corresponding vertical position $z_0^{(\varepsilon)}$ of this volume. We can therefore reconstruct both the shape and the strain distribution in the island.

In order to determine the local chemical composition within an island, a pair of measurements of $I_{\mathbf{h}}(\mathbf{q})$ has to be performed, in which the

values of the polarizability coefficients $\chi_{\mathbf{h}}(\mathbf{r})$ are different for different elements. In Ge or SiGe islands grown on Si, the spatial variation of the polarizability coefficient $\chi_{\mathbf{h}}(\mathbf{r})$ is a weighted average of the coefficients of Si and Ge:

$$\chi_{\mathbf{h}}(\mathbf{r}) = \left[\chi_{\mathbf{h}}^{(\text{Ge})} c_{\text{Ge}}(\mathbf{r}) + \chi_{\mathbf{h}}^{(\text{Si})} (1 - c_{\text{Ge}}(\mathbf{r})) \right] \Omega(\mathbf{r}), \quad (11)$$

where $c_{\text{Ge}}(\mathbf{r})$ is the local concentration of Ge and $\Omega(\mathbf{r})$ is the shape function of the island (unity in the island and substrate volumes and zero outside it). The variation of $\chi_{\mathbf{h}}$ can be achieved in two different ways:

- 1 One can use two different wavelengths λ_1 and λ_2 and the same diffraction vector $\mathbf{h}$ in so called *anomalous x-ray scattering*. Usually the wavelengths are chosen so that one of them is close to the absorption edge of one element in the island. For a SiGe island on Si, the Ge K-edge at 11.1036 keV can be used as one wavelength, the other one can be chosen arbitrarily, sufficiently away from the edge. Then, the intensity map measured at a wavelength close to the edge is determined mainly by the Si-rich regions in the sample, since the real correction f' to the atomic scattering factor f_{Ge} exhibits a minimum. For a wavelength far away from the edge, scattering from Ge, which has a higher Z , gives a stronger contribution, since $f_{\text{Ge}} > f_{\text{Si}}$, due to the difference in the Z -values [19, 20, 21].
- 2 Another possibility is to choose only one wavelength and to compare measurements with different diffraction vectors $\mathbf{h}_1 \parallel \mathbf{h}_2$. In particular, this method is suitable for a zinc-blende structure, where the strong 400 and the quasi-forbidden 200 diffractions can be chosen. The polarizability coefficient for 400 contains the sum of the atomic scattering factors of both elements constituting the unit cell, whereas the *difference* of these factors appears in the expression for $\chi_{\mathbf{h}}$ for 200. For the InAs/GaAs system, as an example, the 200 polarizability of GaAs is very small compared with that of InAs, therefore the intensity distribution measured in 200 diffraction stems mainly from the InAs-rich parts of the structure. This method suffers, however, from the difference in the phase terms $\exp(-i\mathbf{h}_{1,2} \cdot \mathbf{u})$ and therefore it is not as precise as the previous one [17, 18].

Of course both methods may be combined in order to enhance the sensitivity.

Using Eq. (10) one obtains for the ratio of the scattered intensities

$$\frac{I_{\mathbf{h}1}(\mathbf{q})}{I_{\mathbf{h}2}(\mathbf{q})} \approx \left| \frac{\chi_{\mathbf{h}1}}{\chi_{\mathbf{h}2}} \right|^2 \bigg|_{x=x_0, z=z_0}, \quad (12)$$

from which the local composition $c(x_0, z_0)$ in the stationary point can be determined.

4. Experiments

Our investigations focus on a sample with uncapped SiGe islands grown by molecular beam epitaxy on a Si(001) substrate. 6 ML of Ge were deposited at a growth temperature of 600°C, resulting in dome-shaped islands with a density of about $4 \times 10^9 \text{ cm}^{-2}$, and a height and base diameter of about 13 nm and 110 nm, respectively.

In the following we will present diffraction data taken with three different methods: strongly asymmetric coplanar diffraction (CXRD), general non-coplanar diffraction (NCXRD), and grazing incidence diffraction (GID). For the latter, the direct analysis is performed using only one wavelength but different reflections, as well as using anomalous scattering at a high index reflection.

4.1 Indirect Methods

Coplanar Diffraction. In order to determine the structural parameters of SiGe islands, we used x-ray diffraction around asymmetrical Bragg reflections. The measurements have been performed in the coplanar scattering geometry, with small incident angles in order to reduce scattering from the substrate, which otherwise blurs the intensity scattered from the Ge islands. Measuring the lattice parameter in growth direction *and* parallel to the sample surface, and thus reconstructing the unit cell, the strain state and the composition of the islands is obtained.

Figure 3(a) shows the reciprocal space map (RSM) recorded around the (113) Si reflection, at a wavelength of 1.55 Å. We chose this reflection because it has the lowest incidence angle (i.e. when measured in grazing incidence and steep exit geometry), and hence shows the best ratio of scattering from the surface island layer and the substrate compared to all other available *coplanar* reflections.

From a first qualitative evaluation of peak positions using kinematical scattering theory, the Ge content and the in-plane strain $\varepsilon_{\parallel} = \frac{a_{\parallel} - a_{\text{bulk}}}{a_{\text{bulk}}}$ of the islands can be calculated, and we obtain the averaged values of $\bar{\varepsilon}_{\parallel} = -0.012$ and $\bar{x}_{\text{Ge}} = 0.76$. However, for a more detailed evaluation, simulations of the RSMs taking the complex strain distribution into

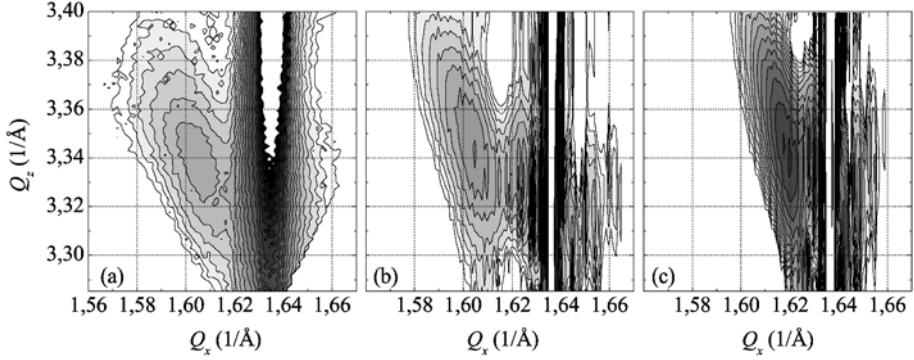


Figure 3. (a) RSM around the (113) reflection of Si of the SiGe island sample. The substrate has been excluded from the measurement as it would saturate the detector. (b-c) simulations using FEM calculations as input, (b) for a island with a varying content from $x_{\text{Ge}} = 0.45$ at the island base and $x_{\text{Ge}} = 0.80$, and (c) for a constant Ge content of $x_{\text{Ge}} = 0.6$.

account are required [7, 8]. The strain distribution was obtained from FEM calculations, assuming different models on the interdiffusion of the islands. For the shape we used the simple model of a parabolic island with a circular base, the height and base width were taken from AFM: $H = 13$ nm, $L = 80$ nm.

In the simulation we assume that the displacement field, the polarizability χ_{h} and the shape function have cylindrical symmetry with respect to the z axis. This is a simplification, as even for a cylindrically symmetric shape of the islands, the elastic relaxation will not be cylindrically symmetric due to the elastic anisotropies of Si and Ge [30]. These anisotropies are, however, small, and may be neglected here. Then, the intensity scattered by an individual island was calculated using Eqs. (1)

Some of our simulation results are shown Fig. 3(b-c). For a island with a constant composition $x_{\text{Ge}} = 0.6$ the calculated intensity distribution [Fig. 3(c)] has a maximum at the correct value of $q_{\perp}$, the value of $q_{\parallel}$ is too small. The composition of the islands cannot be uniform throughout, since by changing x_{Ge} , both the lateral and vertical positions of the intensity maximum change simultaneously. Assuming a gradient of the Ge content in the islands, we found a better agreement with the experiment, the optimum was found with a Ge concentration profile starting at $x_{\text{Ge},1} = 0.45$ at the base and increasing up to $x_{\text{Ge},2} = 0.80$ at the island top in a square-root manner: $x_{\text{Ge}}(z) = x_{\text{Ge},1} + (x_{\text{Ge},2} - x_{\text{Ge},1})(z/H)^{0.5}$. The resulting strain distribution gives rise to a peak in the calculated intensity distribution at the correct *position* [see Fig. 3(b)]. Assuming a linear Ge profile $x_{\text{Ge}}(z) = x_{\text{Ge},1} + (x_{\text{Ge},2} - x_{\text{Ge},1})(z/H)^1$ or a quadratic profile

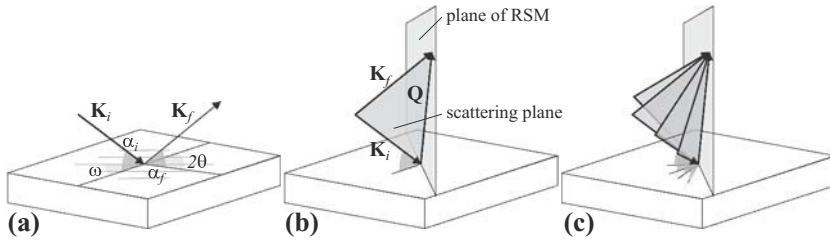


Figure 4. Sketch of the NCXRD geometry in real space (a) and in reciprocal space (b). Note that the scattering plane and the plane of RSM do not coincide. In (c) several possible combinations of $\mathbf{K}_i$ and $\mathbf{K}_f$ are shown resulting in the same scattering vector $\mathbf{Q}$, but with different angles of incidence α_i .

$x_{\text{Ge}}(z) = x_{\text{Ge},1} + (x_{\text{Ge},2} - x_{\text{Ge},1})(z/H)^2$ yields values for $x_{\text{Ge},1,2}$ different by about 5 to 10%, but with a slightly worse correspondence between simulation and experiment. Hence we conclude that the Ge content increase with height is steeper at the island base than at the island apex.

General Non-Coplanar Diffraction. The scattering geometry for general non-coplanar x-ray diffraction (NCXRD) is shown in Fig. 4(a). The scattering experiment is described by four angles, the in-plane angles ω and 2θ , and the incidence and exit angles α_i and α_f . Figure 4(b) shows the scattering geometry in reciprocal space. The momentum transfer $\mathbf{Q} = \mathbf{K}_f - \mathbf{K}_i$ can be achieved with different combinations of the wave vectors of the incident and scattered beams, $\mathbf{K}_{i,f}$, as $\mathbf{Q}$ is a vector in three-dimensional reciprocal space, but the experiment has four degrees of freedom. This is illustrated in Fig. 4(c): The scattering plane, i.e. the plane defined by $\mathbf{K}_i$ and $\mathbf{K}_f$, can be rotated around $\mathbf{Q}$, which changes (among others) the incidence angle α_i , or *vice versa* allows a selection of a certain value of α_i . The minimum value of α_i is reached in coplanar geometry, i.e. with the sample's surface normal lying in the scattering plane. The advantage of this geometry is that the incidence angle can be kept constant below the critical angle of external reflection, which results in a constant and small penetration depth, reducing the scattering signal from the substrate as compared to CXRD. Hence this geometry combines advantages of CXRD (measuring in-plane *and* vertical lattice parameters) and GID (tunable penetration depth).

Using synchrotron radiation, Stangl *et al.* have recorded RSMs around the (202) reciprocal lattice point of Si [28], which is inaccessible at a wavelength of 1.55 Å in coplanar geometry because α_i would be negative. Similar experiments on InAs islands on GaAs were performed by Zhang *et al.*[29].

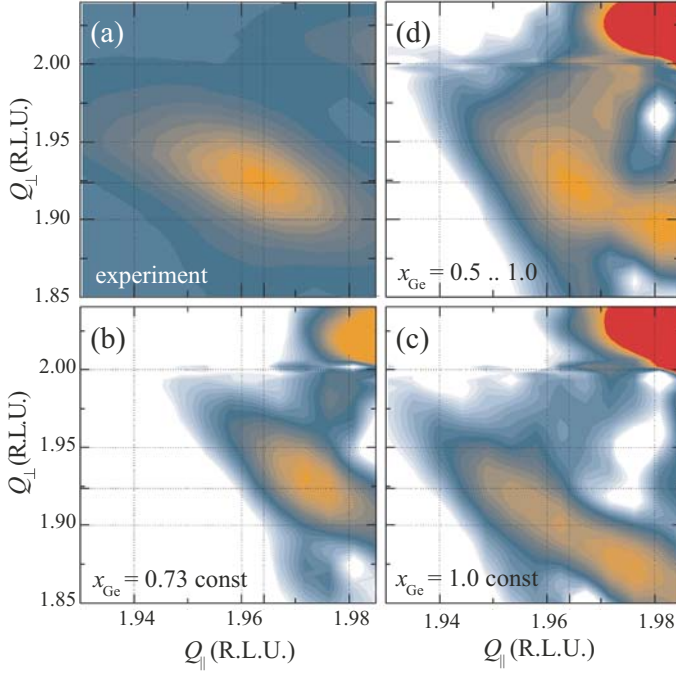


Figure 5. (a) RSM around the (202) reflection of Si of the SiGe island sample. The substrate has been excluded from the measurement as it would saturate the detector. (b-d) Simulations with different model parameters. The coordinates are expressed in reciprocal lattice units of Si.

Figure 5(a) shows the measured intensity distribution. RSMs have been measured at $\lambda = 1.55 \text{ \AA}$ in two different azimuths $[202]$ and $[\bar{2}02]$ in order to account for possible anisotropies, which have, however, not been observed. Taking simply the peak position as a measure of the lattice parameters and the composition, we obtain average parameters $\bar{\varepsilon}_{\parallel} = -0.011$ and $\bar{x}_{\text{Ge}} = 0.73$. A more elaborate sample analysis has been carried out in a similar way as for the (113) RSM: the strain distribution has been calculated by FEM for different assumptions on the island's composition profile, and the reciprocal space maps calculated in a similar manner than before:

Similarly to the previous section, we assume a cylindrical symmetry of the island shape, its chemical composition and elastic strain field. We calculate the intensity distribution for $\mathbf{q}$ within the plane defined by the surface normal and $\mathbf{h} = (2, 0, 2)$. In this case, one of the three integrals in Eq. (1) can be calculated analytically, and using cylindrical coordinates we obtain the expression for the structure factor in Eq. (1).

If we subtract the contribution of the coherent scattering, this expression will be slightly modified:

$$F_{\mathbf{h}}(\mathbf{q}) = 2\pi \int_0^\infty d\rho \rho \int_{-\infty}^\infty dz \chi_{\mathbf{h}}(\rho, z) e^{-iq_z z} \left[e^{-ih_z u_z(\rho, z)} J_0(q_{\parallel} \rho + h_{\parallel} u_{\parallel}(\rho, z)) - H(-z) J_0(q_{\parallel} \rho) \right]. \quad (13)$$

In the calculation, we set $\chi_{\mathbf{h}}(\rho, z)$ constant within the island and the substrate, and zero in vacuum. This is possible because we do not calculate absolute intensities, but only the scattering from a single island on an arbitrary intensity scale. Neglecting the difference in $\chi_{\mathbf{h}}$ between Si and Ge has no significant influence on the result of the simulations. The resolution of the simulations along q_x and q_z is determined by the area used for the FEM calculations, in our case $\delta q_x \simeq 0.1 \text{ \AA}^{-1}$ and $\delta q_z \simeq 0.2 \text{ \AA}^{-1}$, while the step width of the FEM calculations determine the maximum size of the simulated map, here $\Delta q_x = \Delta q_z \simeq 45 \text{ \AA}^{-1}$.

The results of our simulations are shown in Fig. 5(b-d). As in the case of coplanar diffraction, assuming a constant composition of the SiGe islands in the FEM calculations finally yields no correspondence with the experiment: for $x_{\text{Ge}} = 0.73$, the simulation yields a peak at the correct value of Q_z , but at a too large value of Q_x [Fig. 5(b)]. At $x_{\text{Ge}} = 1.0$ [Fig. 5(c)], the peak position is almost correct along Q_x , but Q_z is too small. Thus we again varied the Ge composition distribution within the islands from a value $x_{\text{Ge},1}$ at the base to a value $x_{\text{Ge},2}$ at the top. The best correspondence between simulation and experiment was found for a steeper increase of the Ge content at the base of the islands, as for the analysis of the CXRD data. Figure 5(d) shows the calculated RSM for a Ge distribution which reproduces the peak position best, with a maximum content of $x_{\text{Ge},2} = 1$ at the top of the island, and $x_{\text{Ge},1} = 0.5$ at the base.

The corresponding strain distribution is shown in Fig. 6. In contrast to, e.g. InAs islands on GaAs, which have a higher lattice mismatch and a higher aspect ratio, in the SiGe islands the elastic relaxation is only about 50% even at the top: $\varepsilon_{xx} \simeq 0.02$ is about half of the lattice mismatch between Si and Ge.

From this analysis it is evident that the variations of strain and composition take place on a length scale which is too small to allow a treatment as an island with *average* properties. Calculating the average Ge composition of the island from the obtained Ge concentration profile yields $x_{\text{Ge,av}} = 0.78$, which is slightly but significantly different from the value of $\bar{x}_{\text{Ge}} = 0.73$ obtained from a simple evaluation of the peak position alone.

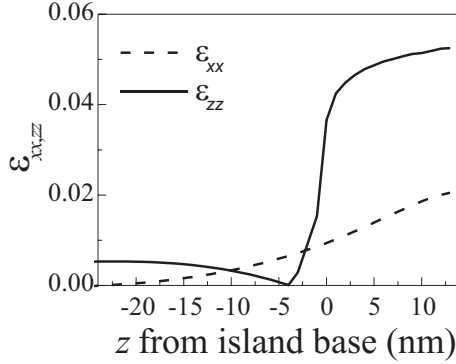


Figure 6. The profiles of the $\varepsilon_{xx,zz}$ components of the strain tensor along the vertical axis of the SiGe island following from the CXRD and NCXRD measurements.

Both for CXRD and NCXRD measurements, the shape of the intensity distribution around the island peak is not perfectly reproduced by our simulations, which we attribute to several simplifying assumptions in our approach. We have neglected the lateral Ge concentration gradient in the islands and we have assumed that all irradiated islands have the same shape, whereas AFM shows a certain size distribution. We have also assumed the same Ge composition of the wetting layer as that at the base of the islands, $x_{\text{Ge}} = 0.5$, neglecting possible lateral variations, for instance due to trenches around the domes [31]. However, the in-plane lattice parameter of the wetting layer matches that of Si and hence the scattered intensity appears only close to the truncation rod. We also investigated the influence of the island shape, assuming a truncated pyramid, a truncated cone, and a paraboloid; the island shape has some influence on the peak shape, but virtually none on its position, and the parabolic shape gives the best fit to the measurement. Changing the island size by $\pm 15\%$ influences the peak shape as well. We observed a better agreement with the experiment assuming island sizes slightly smaller than those observed by AFM. This indicates that the diffracting crystalline core of the islands is smaller than their outer dimensions observed with AFM, probably due to Ge oxidation.

4.2 Direct Methods

In order to compare different methods, we have measured the intensity distribution in GID around the (220) and (400) in-plane reflections for the same sample, and have performed an analysis using the iso-strain scattering approach.

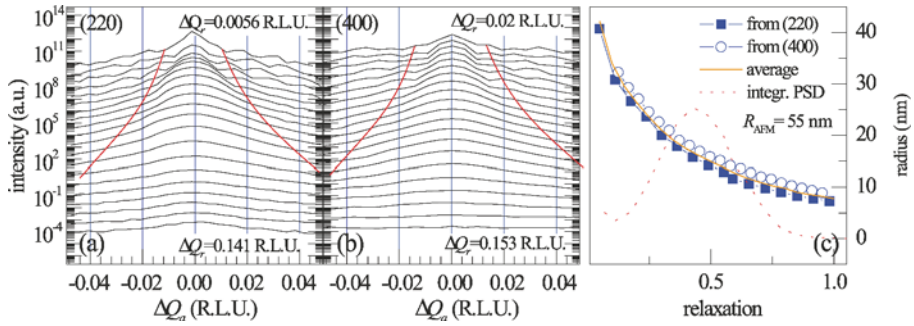


Figure 7. q_a -scans recorded at various values of q_r and $q_z = 0$, for the (220) RLP (a), and the (400) RLP (b). From the distance of the side maxima (indicated by the red lines) the lateral extents of the iso-strain areas have been determined, and are plotted as a function of the relaxation in panel (c). The dotted red line shows the PSD spectrum at $q_a = 0$ integrated over the full α_f range.

Iso-Strain Method. The GID experiments have been carried out using a position sensitive detector (PSD) oriented vertical to the sample surface, the wavelength was $\lambda = 1.55$ Å. 3D-RSMs around the (220) and (400) RLPs have been recorded by performing scans along $q_a \equiv q_y$ for various settings of $q_r \equiv q_x$, shown in Fig. 7(a,b). Here the intensities along α_f for $\alpha_f < \alpha_c$ ($q_z \simeq 0$) have been integrated to achieve better counting statistics. The lateral sample dimensions have been obtained from the position of the first side maxima in the scans, and are plotted as a function of the relaxation in Fig. 7(c) for the two reflections.

The heights of the iso-strain areas above the surface were determined from α_f -scans at $q_a = 0$, which are shown in Fig. 8(a,b). To achieve reasonable counting statistics, the regions of the central maxima along q_a in the RSMs have been integrated. The obtained values are plotted in Fig. 8(c), again as a function of the relaxation.

As is obvious from Fig. 8, the obtained values differ significantly for the two reflections, (220) and (400). Although the strain gradients are smaller in SiGe islands than in InAs islands, and the islands have a lower aspect ratio, the condition Eq. (7) is well fulfilled. But there is a limitation inherent to the iso-strain analysis: q_a - and q_z -scans are performed for various values of q_r , corresponding to various relaxation states. The intensity distribution is spread out over a certain distance along q_r , as is indicated by the red dotted line in Figs. 7(c) and 8(c). The maximum in this curve can, however, not directly be interpreted as the predominant presence of SiGe with the corresponding strain, as was demonstrated in Ref. [18], and it is not intrinsically clear up to which q_r

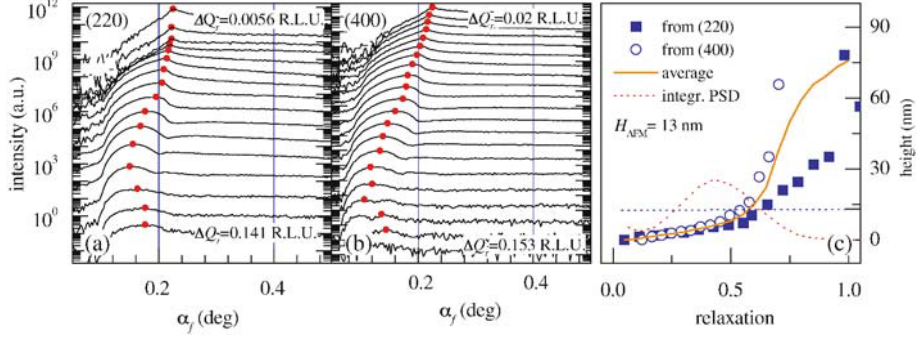


Figure 8. α_f -scans recorded at various values of q_r and $q_a = 0$, for the (220) RLP (a), and the (400) RLP (b). From the position of the pronounced maxima (indicated by the red dots) the height of the iso-strain areas above the surface have been evaluated, they are plotted as a function of the relaxation in panel (c). The dotted red line shows the PSD spectrum at $q_a = 0$ integrated over the full α_f range.

the scans may be identified with a certain iso-strain area: also for a single strain value the finite shape of the island gives rise to a certain extent of the diffracted intensity in reciprocal space. A "natural" limit would be to evaluate the data up to the value of q_r corresponding to pure, relaxed Ge, which is the plotted range (note that the measurements do not start exactly at the reciprocal lattice point (RLP) of Si, which is the reason why the obtained values for the island radius in Fig. 7 at the base are lower than the 55 nm found by AFM). But if the spectra in Fig. 8(a,b) are inspected, it is reasonable to assume that not the full range of relaxation states between Si and Ge lattice parameters are actually present in the islands, as the shape of the curves for higher values of Δq_r differs from the theoretical prediction (shift of the maximum back to *larger* values in α_f , two maxima in the region below $\alpha_c \simeq 0.22^\circ$). Furthermore the wide spread of data between the (220) and (400) reflections in Fig. 8(c) occurs for relaxation values above about 0.7. At this relaxation, the derived height value equals that obtained from AFM. It seems therefore reasonable that the divergence of height data from the two reflections and the reversal of peak shift in the q_z -spectra indicates the limit in relaxation that is actually present in the investigated sample.

Combining the data of height and lateral extent of the islands as a function of the relaxation, we obtain the in-plane strain distribution within the islands. Figure 9 compares the z -dependence of the in-plane strain values obtained with the iso-strain scattering technique with the laterally averaged values of the FEM calculations. The strain distributions obtained by the two methods show a good agreement.

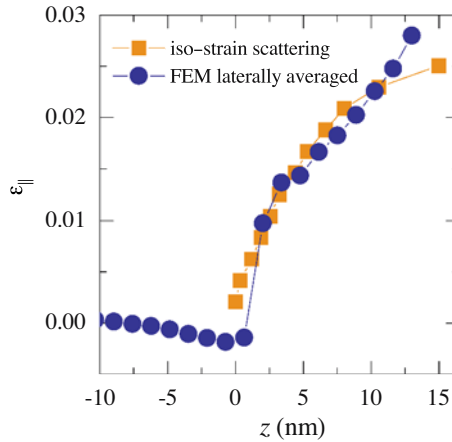


Figure 9. In-plane strain distribution ($\varepsilon_{xx} = (a_{\parallel} - a_{\text{Si}})/a_{\text{Si}}$) within the uncapped SiGe islands obtained by the iso-strain scattering technique and from FEM simulations.

From the iso-strain analysis presented so far, the Ge composition cannot be obtained unambiguously. As the maximum relaxation is about 0.7, the islands are either not fully relaxed even at the top, or the island composition does not reach pure Ge. If we assumed that the islands are fully relaxed at the top, we would obtain a maximum Ge concentration of $x_{\text{Ge,max}} \simeq 0.7$. Accordingly, if we assumed pure Ge at the top of the islands, the relaxation would not exceed a value of 0.7. Where between these two extreme situations the actual maximum content and relaxation lie, cannot be judged from the analysis above. As in SiGe no quasi-forbidden reflections such as (200) exist, anomalous diffraction has been used to clarify this point.

Anomalous Diffraction. In anomalous diffraction, the intensity distribution is recorded at different energies in the vicinity of an absorption edge of one of the sample's constituent elements. The scattering power of this element changes for the two energies, while all other details of the experimental setup remain unchanged: for a small energy variation, the scattering power of all other elements, which do not exhibit absorption edges near the employed energies, remains practically constant, and small changes in scattering angles have virtually no influence on, e.g. the illuminated sample area. Hence, from the ratio of intensities measured at different energies the composition of the sample can be obtained.

For InAs islands on GaAs, anomalous scattering was employed for the determination of composition by Schüllli *et al.*[19]. Anomalous diffraction here is an alternative to the measurement at weak and strong reflections in a zincblende structure like GaAs in the iso-strain method. In 200 diffraction for GaAs, where the polarizability coefficient $\chi_{\mathbf{h}}$ is proportional to the difference $f_{\text{Ga}} - f_{\text{As}}$, an energy can be found (12.352 keV) where $\text{Re}(\chi_{\mathbf{h}}) = 0$. At this energy, the contribution of the GaAs lattice is negligible in comparison to InAs. The intensity measured at this energy is compared with that at 11.8 keV below the As absorption edge. Using this procedure, the InAs local concentration was directly obtained [19].

For Ge islands on Si, no weak reflections exist, since both materials crystallize in diamond lattice, and anomalous scattering provides the only possibility to use chemical sensitive iso-strain scattering. Such studies have been performed by Magalhaes *et al.*[20] and Schüllli *et al.* [21]. In Ref. [20] the 220 diffraction was used with the x-ray energies at 11.103 keV (slightly below the Ge K absorption edge, to avoid x-ray fluorescence) and at 11.000 keV. Magalhaes *et al.* performed radial scans along the [110] direction using both energies and angular scans along $[1\bar{1}0]$ for various q_r . Such measurements were performed for two different samples, one with pyramids with a height of about 3 nm and diameter of 50 nm, the other with domes with a height of 14 nm and diameter of 65 nm. From the data they determined the vertical dependence of the laterally averaged chemical composition and found that the vertical gradient in Ge content is much steeper for pyramids than for the domes.

Schüllli *et al.* [21] used energies of 11.103 keV and 11.043 keV (below the edge). In order to improve the chemical sensitivity, a series of measurements for different diffraction vectors was performed. For higher order reflections (i.e. for longer $\mathbf{h}$) the local minimum in f_{Ge} is deeper, since the monotonic part $f_0(\mathbf{h})$ of f_{Ge} decreases with increasing $|\mathbf{h}|$. In addition, the condition in Eq. (7) is better fulfilled for larger diffraction vector.

This technique was employed for the study of dome-shaped SiGe islands grown by depositing 7 ML of Ge on Si(001) at a temperature of 600°C, at nearly identical growth conditions to those used for the sample investigated in Sec. 4.1. Within a height interval of 3 nm the Ge content rises steeply from 0 to about 80%. Above this transition the composition remains at about 80% up to the island top at a height of about 12 nm (see Fig. 10). A comparison between the vertical Ge profiles obtained by the indirect and direct methods on Ge domes on Si is presented in Fig. 11. Apparently, both methods yield rather comparable results.

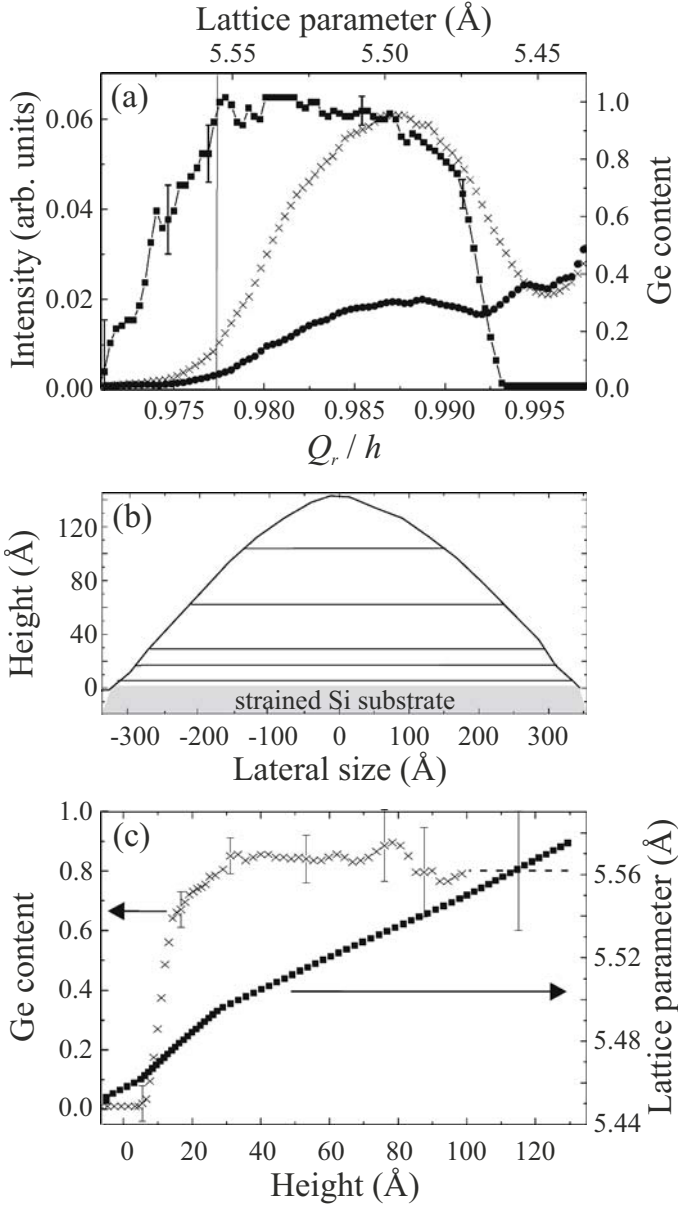


Figure 10. (a) Radial scans in GID geometry around the (800) Bragg reflection, recorded at 11043 eV (crosses) and 11103 eV (dots). The intensity ratio is plotted as squares; (b) lateral dimension and (c) Ge content of SiGe islands as a function of island height obtained from anomalous x-ray scattering (from [21]).

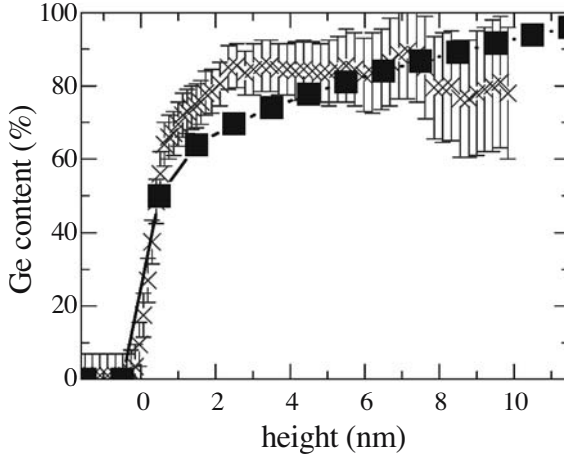


Figure 11. Vertical profiles of the Ge concentration x_{Ge} in a SiGe dome determined by the indirect method (squares) and by anomalous scattering (crosses with error bars).

So far, the direct methods have been used only for uncapped islands. Their application for buried islands is complicated by the fact that deformed material contributing to the diffracted intensity surrounds the islands entirely, and the strain variation with height will be no longer monotonic. Secondly, the vertical resolution of the iso-strain method following from the interference of the primary and reflected waves in Eq. (5) is absent in the case of buried islands, due to the lack of the surface-reflected reference beam.

5. Conclusions and Outlook

From the different experiments, it is clear that albeit pure Ge is deposited during growth, already the uncapped islands are alloyed, and the Ge content varies from about $x_{\text{Ge}} = 0.5$ at the base of the islands to $x_{\text{Ge}} = 1.0$ at the top. The reduced elastic energy of an interdiffused island, as compared to a island of pure Ge, is believed to be the main driving force for alloying [32, 33]. With the different scattering techniques presented above, we could establish the strain distribution *and* the composition distribution in uncapped SiGe islands. The strain distributions obtained with indirect methods based on model calculations, and with both the iso-strain and anomalous scattering methods show a good agreement.

For any device structure, the SiGe islands have to be overgrown, which is expected to lead to a further intermixing of Ge and Si. To determine

the composition and the strain distributions in *buried* nanostructures, the indirect methods can be used without modification.

Acknowledgments

This work was supported by the contracts HPRI-CT-1999-00040/2001-00140 and HPRN-CT-1999-00123 of the European Commission, the FWF Vienna (contract # 14668), the BMWV, and the GMe, Vienna and by the Grant Agency of Czech Republic (project 202/03/0148). We are grateful to T. H. Metzger for helpful discussions, the staff of beamlines W1 (HASYLAB, Hamburg) and ID10B (ESRF, Grenoble) for their support with the experiments.

References

- [1] For a series of articles on self-organized dots, see *MRS Bulletin* **23(2)**, (1998).
- [2] D. Bimberg, M. Grundmann, and N. Ledentsov, *Quantum Dot Heterostructures* (Wiley, Chichester, 1999).
- [3] O. G. Schmidt, U. Denker, S. Christiansen, and F. Ernst. Composition of self-assembled Ge/Si islands in single and multiple layers. *Appl. Phys. Lett.* **81**: 2614–2616, 2002.
- [4] T. Walther, A. G. Cullis, D. J. Norris, and M. Hopkinson. Nature of the Stranski-Krastanow transition during epitaxy of InGaAs on GaAs *Phys. Rev. Lett.* **86**: 2381–2384, 2001.
- [5] A. J. Steinfert, P. M. L. O. Scholte, A. Ettema, F. Tuinstra, M. Nielsen, E. Landemark, D.-M. Smilgies, R. Feidenhansl, G. Falkenberg, L. Seehofer, and R. L. Johnson. Strain in nanoscale germanium hut clusters on Si(001) studied by x-ray diffraction. *Phys. Rev. Lett.* **77**: 2009–2012, 1996.
- [6] A. A. Darhuber, P. Schittenhelm, V. Holy, J. Stangl, G. Bauer, and A. Abstreiter. High-resolution x-ray diffraction from multilayered self-assembled Ge dots. *Phys. Rev. B* **55**: 15652–15663, 1997.
- [7] Th. Wiebach, M. Schmidbauer, M. Hanke, H. Raidt, R. Köhler, and H. Wawra. Strain and composition in SiGe nanoscale islands studied by x-ray scattering. *Phys. Rev. B* **61**: 5571–5578, 2000.
- [8] A. Hesse, J. Stangl, V. Holy, T. Roch, G. Bauer, O. G. Schmidt, U. Denker, and B. Struth. Effect of overgrowth on shape, composition, and strain of SiGe islands on Si(001). *Phys. Rev. B* **66**: art. no. 085321, 2002.
- [9] M. Rauscher, R. Paniago, H. Metzger, Z. Kovats, J. Domke, J. Peisl, H.-D. Pfannes, J. Schulze, and I. Eisele. Grazing incidence small angle x-ray scattering from free-standing nanostructures. *J. Appl. Phys.* **86**: 6763–6769, 1999.
- [10] V. Holý, U. Pietsch, and T. Baumbach, *High-Resolution X-Ray Scattering From Thin Films and Multilayers* (Springer, Berlin, 1999).
- [11] M. Schmidbauer, Th. Wiebach, H. Raidt, M. Hanke, R. Köhler, and H. Wawra. Ordering of self-assembled $\text{Si}_{1-x}\text{Ge}_x$ islands studied by grazing incidence small-angle x-ray scattering and atomic force microscopy *Phys. Rev. B* **58**: 10523–10531, 1998.

- [12] M. Schmidbauer, M. Hanke, and R. Köhler. X-ray diffuse scattering on self-organized mesoscopic structures. *Cryst. Res. Technol.* **36**: 3–34, 2002.
- [13] J. Stangl, V. Holy, T. Roch, A. Daniel, G. Bauer, J. Zhu, K. Brunner, and G. Abstreiter. Grazing incidence small-angle x-ray scattering study of buried and free-standing SiGe islands in a SiGe/Si superlattice. *Phys. Rev. B* **62**: 7229–7236, 2000.
- [14] V. Holy, A. A. Darhuber, J. Stangl, S. Zerlauth, F. Schaeffler, G. Bauer, N. Darowski, D. Luebbert, U. Pietsch, and I. Vavra. Coplanar and grazing incidence x-ray-diffraction investigation of self-organized SiGe quantum dot multilayers. *Phys. Rev. B* **58**: 7934–7943, 1998.
- [15] G. J. Williams, M. A. Pfeifer, I. A. Vartanyants, and I. K. Robinson. Three-dimensional imaging of microstructure in Au nanocrystals. *Phys. Rev. Lett.* **90**: art. no. 175501, 2003.
- [16] R. W. Gerchberg and W. O. Saxton. A practical algorithm for the determination of phase from image and diffraction plane pictures. *Optik (Stuttgart)* **35**: 237–246, 1972.
- [17] I. Kegel, T. H. Metzger, A. Lorke, J. Peisl, J. Stangl, G. Bauer, J. M. Garcia, and P. M. Petroff. Nanometer-scale resolution of strain and interdiffusion in self-assembled InAs/GaAs quantum dots. *Phys. Rev. Lett.* **85**: 1694–1697, 2000.
- [18] I. Kegel, T. H. Metzger, A. Lorke, J. Peisl, J. Stangl, G. Bauer, K. Nordlund, W.V. Schoenfeld, and P.M. Petroff. Determination of strain fields and composition of self-organized quantum dots using x-ray diffraction. *Phys. Rev. B* **63**: art. no. 035318, 2001.
- [19] T. U. Schüllli, M. Sztucki, V. Chamard, T. H. Metzger, and D. Schuh. Anomalous x-ray diffraction on InAs/GaAs quantum dot systems. *Appl. Phys. Lett.* **81**: 448–450, 2002.
- [20] R. Magalhaes-Paniago, G. Medeiros-Ribeiro, A. Malachias, S. Kycia, T. I. Kamins, and R. Stanley Williams. Direct evaluation of composition profile, strain relaxation, and elastic energy of Ge:Si(001) self-assembled islands by anomalous x-ray scattering. *Phys. Rev. B* **66**: art. no. 245312, 2002.
- [21] T. U. Schüllli, J. Stangl, Z. Zhong, R. T. Lechner, M. Sztucki, T. H. Metzger, and G. Bauer. Direct determination of strain and composition profiles in SiGe islands by anomalous X-ray diffraction at high momentum transfer. *Phys. Rev. Lett.* **90**: art. no. 066105, 2003.
- [22] V. Holy, J. Stangl, S. Zerlauth, G. Bauer, N. Darowski, D. Luebbert, and U. Pietsch. Lateral arrangement of self-assembled quantum dots in an SiGe Si superlattice. *J. Phys. D* **32**: A234–A238, 1999.
- [23] S. K. Sinha, E. B. Sirota, S. Garoff, and H. B. Stanley. X-ray and neutron scattering from rough surfaces. *Phys. Rev. B* **38**: 2297–2311, 1988.
- [24] C. Pryor, J. Kim, L. W. Wang, A. J. Williamson, and A. Zunger. Comparison of two methods for describing the strain profiles in quantum dots. *J. Appl. Phys.* **83**: 2548–2554, 1998.
- [25] V. A. Borovikov, *Uniform Stationary Phase Method* (IEE Publishing, 1994).
- [26] Ch.-P. Liu, J. Murray Gibson, D. G. Cahill, T. I. Kamins, D. P. Basile, and R. Stanley Williams. Strain evolution in coherent Ge/Si islands. *Phys. Rev. Lett.* **84**: 1958–1961, 2000.

- [27] U. Denker, M. Stoffel, and O. G. Schmidt. Probing the lateral composition profile of self-assembled islands. *Phys. Rev. Lett.* **90**: art. no. 196102, 2003.
- [28] J. Stangl, A. Daniel, V. Holy, T. Roch, G. Bauer, I. Kegel, T. H. Metzger, Th. Wiebach, O. G. Schmidt, and K. Eberl. Strain and composition distribution in uncapped SiGe islands from x-ray diffraction. *Appl. Phys. Lett.* **79**: 1474–1476, 2001.
- [29] K. Zhang, C. Heyn, W. Hansen, T. Schmidt, and J. Falta. Strain status of self-assembled InAs quantum dots. *Appl. Phys. Lett.* **77**: 1295–1297, 2000.
- [30] Landoldt-Börnstein, *Technology of Si, Ge and SiC*, in *Numerical Data and Functional Relationships in Science and Technology*, ed. by M. Schulz and H. Weiss (Springer, Berlin, 1987).
- [31] S. A. Chaparro, Y. Zhang, J. Drucker, D. Chandrasekhar, and D. J. Smith. Evolution of Ge/Si(100) islands: Island size and temperature dependence. *J. Appl. Phys.* **87**: 2245–2254, 2000.
- [32] S. A. Chaparro, J. Drucker, Y. Zhang, D. Chandrasekhar, M. R. McCartney, and D.J. Smith. Strain-driven alloying in Ge/Si(100) coherent islands. *Phys. Rev. Lett.* **83**: 1199–1202, 1999.
- [33] P. Sonnet and P. C. Kelires. Monte Carlo studies of stress fields and intermixing in Ge/Si(100) quantum dots. *Phys. Rev. B* **66**: art. no. 205307, 2002.

CARBON-INDUCED Ge DOTS ON Si(100): INTERPLAY OF STRAIN AND CHEMICAL EFFECTS

G. Hadjisavvas

Physics Department, University of Crete, P.O. Box 2208, 710 03 Heraklion, Crete, Greece

Ph. Sonnet

Laboratoire de Physique et de Spectroscopie Electronique, CNRS-UMR 7014, 68093 Mulhouse Cedex, France

P. C. Kelires

Physics Department, University of Crete, P.O. Box 2208, 710 03 Heraklion, Crete, Greece, and Foundation for Research and Technology-Hellas (FORTH), P.O. Box 1527, 711 10 Heraklion, Crete, Greece

Abstract Carbon system plays a twofold role in the SiGe, inducing both high stress fields and strong chemical effects. Our Monte Carlo simulations, based on a novel algorithm enabling C-insertion and equilibration, shed light on the stress field and composition of C-induced Ge islands on Si(100), a prototypical case where these two effects operate. It is shown that the dots do not contain C under any conditions of temperature and coverage, but have a gradual composition profile from SiGe at the bottom to Ge at the apex. The average compressive stress in the islands is considerably reduced, compared to the pure Ge/Si case. At low Ge coverage, the terrace around the dots is enriched with Si-C dimers. At high Ge contents, Ge wets the surface and covers the pre-deposited C geometries. We predict enhancement of Ge content in the islands upon C incorporation.

1. Introduction

When C enters substitutionally into the Si lattice major strains are generated around the sites of insertion, leading to long-ranged stress

fields [1]. In the case of SiGe alloys, C produces an additional effect [2]: It preferentially bounds to Si atoms, and repels the Ge atoms. This is due to the weakness of the Ge–C bond, having a high positive enthalpy of formation (~ 2 eV/atom), and to the well-known great strength of the Si–C bond. Thus, in the general case of the ternary SiGeC system, strain and chemical effects compete with each other in shaping up the structure and the composition profiles.

A striking manifestation of such C-induced interplay is seen in the case of Ge islands formed on a Si(100) surface, which is precovered with a small amount of C. These islands are a special and important class of Si-based nanostructures. They are fabricated by molecular-beam epitaxy [3, 4, 5, 6] and attracted interest because they are remarkably small (typical sizes are 10-15 nm in diameter and 1-2 nm in height) and exhibit intense photoluminescence (PL) [4]. The growth of these islands proceeds without the formation of a wetting layer (Volmer-Weber mode), contrary to the islands grown on bare Si(100) which follow the Stranski-Krastanov growth mode. This is attributed to the predeposited C atoms. Their small size reduces the lattice constant of the alloyed Si surface and exaggerates the mismatch with the Ge overlayer. In addition, C atoms modify the surface and interface energetics.

Since C atoms play such a vital role, it is essential to know how they are distributed in the surface region, and especially whether they occupy sites in and below the dots. The interpretation of PL data strongly depends on this information, but the issue is controversial. There have been two different models drawn from experimental work. Schmidt and Eberl [4], interpreted PL data as suggesting that the dots have a gradual composition profile from homogeneous SiGeC below and at the bottom to pure Ge towards their apex. On the other hand, Grützmacher and co-workers [6] interpreted their STM images as suggesting that the dots are free of C, have a gradual composition profile from SiGe at the bottom to Ge at the apex, and are located between C-rich patches.

We present here a review of our recent studies [7] which gave a theoretical answer to the problem, and resolved the experimental controversy. We first describe in detail the Monte Carlo algorithm on which the simulations are based. We then extract the stress fields and the associated composition profiles in small pyramidal islands and the surrounding surface region, at typical growth temperatures and for various carbon contents. Our results indicate that the Grützmacher model is more plausible. We predict enhancement of Ge content in the islands, compared to the carbon-free case. The stress pattern in the dots is different from the bare Ge/Si profile.

2. Methodology

2.1 Monte Carlo Algorithms

The simulations are based on a *biased* MC algorithm, within the semi-grand canonical (SGC) ensemble, devised by Kelires to deal with C incorporation in the Si, Ge, and SiGe lattices, or with any case where large size-mismatch exists between the constituent atoms [2, 8, 9].

In this ensemble, denoted as $(\Delta\mu, N, P, T)$, one requires that the total number of atoms N , the pressure P , the temperature T , and the chemical potential differences $\Delta\mu$ remain fixed. These conditions allow fluctuations Δn in the number of atoms of each species as a result of exchanges of particles within the system, which are driven by the appropriate chemical potential differences $\Delta\mu$ ($=\mu_i - \mu_j$, $i, j \equiv \text{Si, Ge, C}$ in the present general case). At the same time, we have also exchanges of volume with the heat bath, as well as the traditional MC moves involving random atomic displacements. Thus, the SGC ensemble can be considered as a combination of the grand canonical (GC) and the more familiar isobaric-isothermal (N, P, T) ensemble. It can be shown that the SGC ensemble is obtained from the GC ensemble upon imposing the constraint that $N = \sum_i N_i$ is fixed and changing to constant pressure [10]. The resulting partition function (only configurational part) for an n -component mixture, which couples volume changes and atom-identity flips, is given by

$$Q_{\text{semi}} = \beta P \int dV e^{-\beta P V} \frac{V^N}{N!} \sum_{\text{identities}} \prod_{i=1}^n \left(\frac{\lambda_i}{\lambda_1} \right)^{N_i} \int ds^N e^{-\beta U(s^N)}, \quad (1)$$

where $\lambda_i = e^{\mu_i/k_B T}$ are the fugacities in the system, $U(s^N)$ is the potential energy associated with both atom identities and displacements, and is a function of the $3N$ scaled atomic coordinates s , and $i=1$ is the arbitrarily fixed identity to which all chemical potential differences (there are $n-1$ independent fugacities in the system) are referred. For a detailed derivation of Eq. (1) see Ref. [10].

The implementation of this ensemble for MC simulations is done through the Metropolis algorithm [11], in the following way: The traditional random atomic moves ($s^N \rightarrow s'^N$), leading to a change in the potential energy $U \rightarrow U'$, and the volume changes $V \rightarrow V'$ are accepted with a probability

$$P_{\text{acc}} \sim e^{-\Delta W/k_B T} = \text{Min}[1, \exp(-\beta \Delta W)] \quad (2)$$

$$\Delta W = (U' - U) + P(V' - V) - N k_B T \ln(V'/V), \quad (3)$$

as in the (N, P, T) ensemble. For the trial moves which select one of the N particles at random, and with equal probability change its identity into one of the other possible identities of the system, the acceptance probability is given by

$$P_{\text{acc}}^{\text{idn}}(i \rightarrow i') = \text{Min} \left[1, \frac{\lambda_{i'}}{\lambda_i} \exp(-\beta \Delta U(s^N)) \right] \sim e^{\beta \Delta \mu} e^{-\beta \Delta U(s^N)}. \quad (4)$$

where

$$\Delta \mu = \mu_i \Delta N_i + \mu_j \Delta N_j + \mu_k \Delta N_k + \dots \quad (5)$$

and $\Delta U(s^N)$ denotes the change in potential energy due to the identity ($i \rightarrow i'$) flip. For the SiGeC system $\Delta \mu$ would be equal to

$$\mu_{\text{Si}} \Delta N_{\text{Si}} + \mu_{\text{Ge}} \Delta N_{\text{Ge}} + \mu_{\text{C}} \Delta N_{\text{C}}. \quad (6)$$

The identity switches can be viewed as Ising-type flips which incorporate the elastic degrees of freedom (effects of strain) explicitly. Eq. (4) shows that the system compositions are dictated by the imposed chemical potential differences and by the potential energy changes. In the latter, chemical-bond preference plays a significant role. To maximize $P_{\text{acc}}^{\text{idn}}$ it is necessary that the attempted flips are considerably less than the attempted random displacements, and always followed by volume changes, so that all degrees of freedom in the system are sufficiently relaxed.

Equilibration within this ensemble is straightforward in systems where the atomic size mismatch among the constituents is small. In this respect, Foiles used the method to study metal alloy surfaces [12], while Kelires and Tersoff studied both bulk and surface properties of SiGe alloys [13]. Later work also studied this system within the SGCMC ensemble [14, 15]. However, semiconductor alloys containing carbon represent a challenging case since the atomic size mismatch is now huge. A carbon atom incorporated substitutionally into the Si, or SiGe lattice, produces severely strained bonds in the neighborhood of the insertion, so that very few attempted MC flips involving formation or elimination of a C site would be accepted, and equilibration is impossible to achieve.

To overcome the large formation energies and diffusion barriers associated with C incorporation, Kelires devised a modification of the SGCMC algorithm that introduces appropriate relaxations of first-nearest-neighbor (nn) atoms to accompany each attempted move [2, 8, 9]. This makes the flips less costly since the “exchange” barriers are effectively reduced. With this modification the change in the potential energy of the alloy now becomes a sum of three terms

$$\Delta U(s^N) = \Delta U_{\text{displ}}(s^N \rightarrow s'^N) + \Delta U_{\text{flip}}(s^N) + \Delta U_{\text{relax}}(s^N \rightarrow s'^N). \quad (7)$$

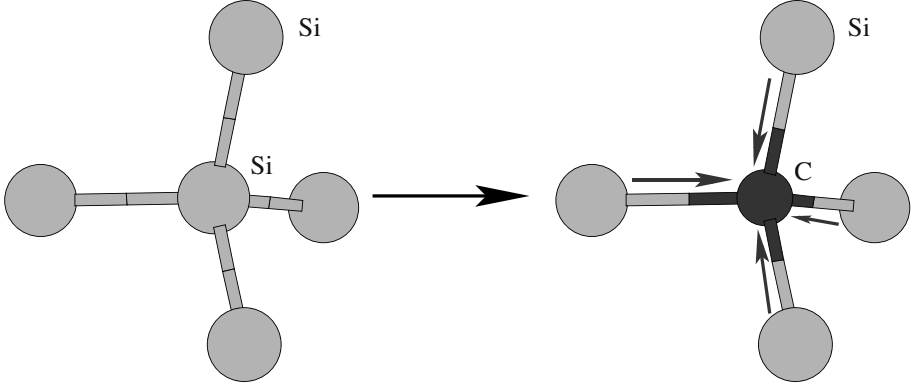


Figure 1. An identity switch of a site from being Si to C, or *vice versa*, produces strained bonds. In order to make the flips less costly, appropriate relaxations of first-nearest-neighbor atoms are introduced.

The first term is the change due to random displacements, the second to flips (alterations in chemical bonding), and the third to the accompanying relaxations. The last two terms substitute now for $\Delta U(s^N)$ in Eq. (4). The combined effect of these two terms is expressed as

$$\Delta U(s^N) = E_{\text{cluster}} \left[i \rightarrow i', \sum_{k=1}^{nn} \sum_{j=1}^3 \Delta s_k^j(r_{0k}^j) \right] - E_{\text{cluster}}^0. \quad (8)$$

The energy is actually estimated over the cluster of atoms affected by the move and the relaxations, instead over the whole system, before and after the move. The capability for doing this stems from the fact that in the Tersoff potential, as well as in other empirical potentials, e.g. the Stillinger-Weber potential [16], we can decompose the total energy of the system into atomic contributions, and so only the energies of atoms affected by the moves have to be recalculated. This saves an enormous amount of computational time. Each nearest neighbor is relaxed away or towards the central atom (which changes identity from i to i' and is labeled 0) in the bond direction $\mathbf{r}_{0k}$. This means that every scaled coordinate s^j is altered according to the scheme

$$\Delta s_k^j(r_{0k}^j) = A_{\text{bond}} r_{0k}^j, \quad (9)$$

$$A_{\text{bond}} = (b_{0k} [i'(0), i(k)] - |\mathbf{r}_{0k}|) \chi_{\text{rel}} / |\mathbf{r}_{0k}|, \quad (10)$$

where b_{0k} is the bulk equilibrium bond length among atoms 0 (after the flip) and k . In principle one could choose b_{0k} to be the bond length associated with the specific environment at hand. For example, the value

of b_{0k} could differ for bulk and surface bonds, or for graphite-like and diamond-like carbon bonds. Provided that flips are followed by a large number of random moves and volume changes, to relax completely the structures, the initial choice has no effect on the acceptance rate. The relaxation parameter χ_{rel} , ranging from 0.0 to 1.0, decides how large the relaxation (expressed by A_{bond}) should be. We find that intermediate values of χ_{rel} make the best effect. For values approaching unity, i.e. when the bond relaxes to its ideal bulk value, the success rate drops (but still remains higher than for $\chi_{\text{rel}} = 0.0$) due to straining of the back-bonds in the neighboring atoms.

2.2 Energetics and Structural Model

In order to model the interactions and make the simulations tractable, we use the well established interatomic potentials of Tersoff for multi-component systems [17] extended by Kelires to the ternary SiGeC system [2]. These potentials have been used with success in similar contexts [2]. In particular, bonding and strain fields induced by the surface reconstruction are accurately described, and various predictions about C interactions and stress compensation are verified experimentally.

To model the experimental conditions as much as possible, we first “predeposit” C atoms in the Si(100) surface by the incorporation process described above at various sub-monolayer coverages and at 900 K. The Si substrate in our simulational cell contains 9 monolayers (ML) of size $200 \text{ \AA} \times 200 \text{ \AA}$. The bottom layer is kept fixed throughout the simulation. The cell is constrained to have laterally the Si lattice dimensions, with relaxation occurring vertically. Periodic boundary conditions are imposed in the lateral directions.

It is now well established that C induces the $c(4 \times 4)$ reconstruction of the surface, although its structure is still a matter of debate [5, 18, 19]. At low coverages, the surface is partly covered with the $c(4 \times 4)$ configurations involving C atoms, while the rest retains the (2×1) symmetry. It has been shown by Leifeld *et al.* [6] that Ge dots nucleate on such C-free areas. The theoretical argument to support this originates from the work of Kelires who unraveled a repulsive interaction between Ge and C atoms in the Si lattice [2]. To conform with this experimental observation, we do not allow incorporation of C in the central portion of the surface in our simulational cell, producing so a C-free area. In the C-containing region, the dominant structures are Si-C dimers on the terrace and C atoms in the third and fourth sub-surface layers, at sites below the surface dimers. These are believed to be the main configurations in the $c(4 \times 4)$ reconstruction at low C contents [18, 19]. To investigate the

effect of the amount of C predeposited, we generate cells with 0.16 and 0.36 ML C coverage by varying the C chemical potential.

In the second stage, we form a coherent pure Ge island on top of the C-free area. The dot has a pyramidal shape with a square base and $\{105\}$ facets, and is oriented at an angle of 45° with respect to the dimer rows of the surface [20]. The dot contains 1750 atoms arranged in 7 ML. The base width is ~ 92 Å. Typically, C-induced dots seen in experiment have similar dimensions. The amount of Ge in the cell is equivalent to ~ 0.5 ML. With such Ge coverage, the model simulates islands produced by experiment at low substrate temperatures (~ 625 K), at which no Ge wetting layer is formed on the terrace [6].

3. Results and Discussion

3.1 Stress Field

We first analyze the stress state of the initial configuration by invoking the concept of atomic level stresses σ_i , as in our recent work on bare Ge/Si islands [21]. We can map the stress field site by site, and we can infer the average stress $\bar{\sigma}$ in each ML in the dot and the substrate, or over the whole dot, by summing the σ_i . This analysis reveals that, in the presence of C, the stress pattern deviates significantly from the bare Ge/Si case (Fig. 2). It is clear that the overall compressive stress in the dot is significantly reduced under the influence of C, and that this reduction is enhanced as the amount of predeposited C increases. The average stress in the dot $\bar{\sigma}_{QD}$ for the bare case is 1.4 GPa/atom [21]. For 0.16 ML C, $\bar{\sigma}_{QD} = 0.45$ GPa/atom, while for 0.36 ML C, $\bar{\sigma}_{QD} = 0.274$ GPa/atom. In the substrate, tensile conditions prevail, especially on the terrace, because the C atoms experience large local tensile stresses. This is partly compensated in the first subsurface layer, where the surface reconstruction induces compressive conditions [13].

The most noticeable change due to C is observed in the bottom layer of the dot. It is found to be under slight tension, contrary to the bare case in which this layer is very compressed, and originates from certain Ge atoms that are under excessive tensile stress. To get more insight into this effect, we resort to the site-by-site analysis of the stress pattern. This is illustrated in Fig. 3 for selected layers of the cell with 0.16 ML C content. It is clearly shown that a large number of Ge atoms in the base layer of the dot are under tension, not only at the periphery but also in regions inside, while the central region is compressed. In the bare case [21] only the peripheral Ge atoms are under tension. We can think of this as “dragging out” the exterior regions of the base layer to conform with the contracted surrounding lattice due to C incorporation. This

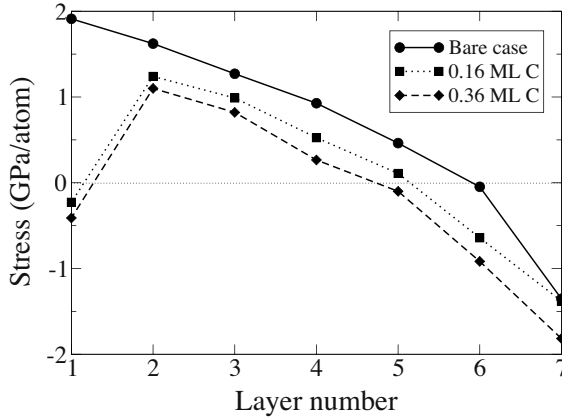


Figure 2. Variation of average stress layer by layer in the dot at 625 K. Layer numbered 1 denotes the base in the dot. Positive (negative) sign indicates compressive (tensile) stress, respectively.

effect considerably weakens in the second layer (not shown), which is mostly under compression (but still lower than in the bare case.) On the terrace, tensile stresses dominate around the dot, but the region just below the dot is mostly compressed, as in the base dot layer. The first sub-surface layer exhibits regions of tensile stress, especially under the dot. This compensates for the compression above on the terrace and the base layer.

3.2 Composition Profiles

At higher temperatures (~ 800 K), intermixing of Ge with C in the Si-C areas, which was initially inhibited due to Ge-C repulsion, is believed to take place [6]. To address this possibility and obtain the appropriate equilibrium distribution of species at high temperatures, the initial configuration is further relaxed at 800 K by interdiffusion. To model this, a pair of dissimilar atoms is chosen randomly and an attempt is made to switch their identities. In the ergodic limit of many thousand attempted flips per site, the *average* site occupancies are calculated for each species. These are compared to the respective *random* occupancies that would be the result of a random distribution of atoms on the lattice sites to infer the average identity of each site, i.e. the overwhelming occupancy.

The application of this analysis to the cell with 0.36 ML C content is shown in Fig. 4, for selected layers. The atoms are shaded according to their average identity. The outstanding feature revealed in these graphs is the complete absence of carbon from the dot layers (only the base layer is shown.) We repeated the simulation for an even higher T (1000 K),

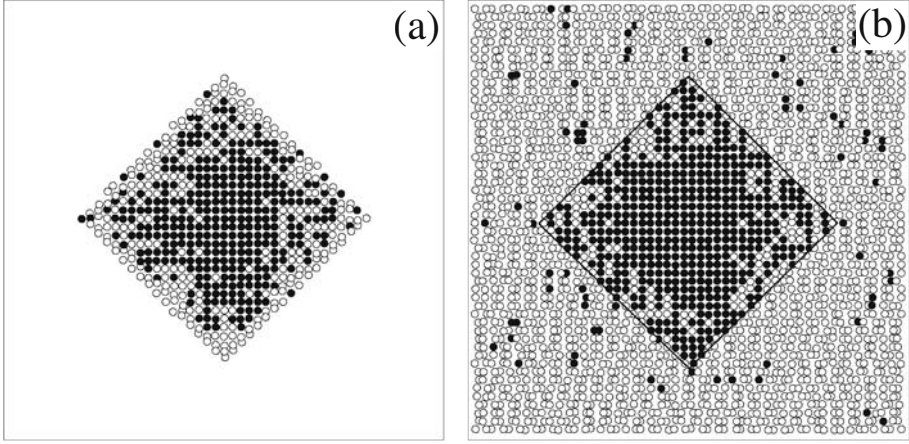


Figure 3. Atomic sites shaded according to their local stress. Filled spheres: compressive; open spheres: tensile. (a) Base dot layer. (b) Top substrate layer. Solid lines (guide to the eye) enclose the area below the island.

to increase further the acceptance of flipping moves, and found the same result. We can point out two factors responsible for this: (a) C has to break Ge-Ge bonds in the dot and form instead Ge-C bonds, which are unstable as we said. (b) The compressive stress in the dot is better compensated by the segregation of Si, rather than C. The latter would induce tensile conditions. So, Si atoms diffuse into the island, while Ge atoms out-diffuse from the interior to wet the terrace due to their low surface energy, mostly forming Ge-Ge dimers and few Ge-Si dimers. C atoms on the terrace are solely involved in Si-C dimers.

Interestingly, the areas on the terrace and in deeper layers, which are underneath the island base, also remain free of carbon after the redistribution of species. Note that the Ge and Si composition profiles in the island are quite similar to the respective profiles in the bare Ge/Si case [21]. Si atoms enrich the bottom layers, mainly in the central regions. This is consistent with the compressive conditions shown in Fig. 3. The Ge content is slowly varying in the bottom and rapidly varying when approaching the top of the island (not shown here.) Taking all these observations into account, we conclude that there is no evidence for SiGeC alloying in the dot or below, and that our theoretical model is in general lines consistent with the experimental model of Grützmacher.

It is reported that at higher Ge coverages the C-induced dots show improved PL [4, 22]. We therefore investigated the distribution of species in this important case. We start with a configuration where Ge fully wets the terrace, covering the predeposited C-rich areas surrounding

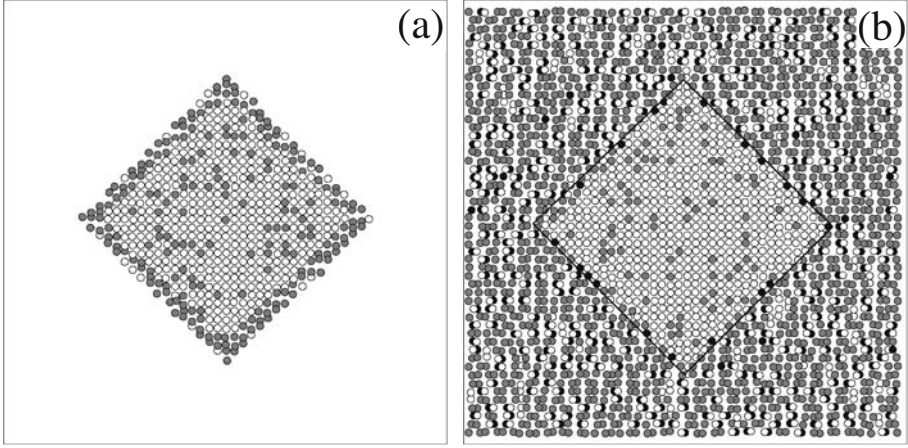


Figure 4. Atomic sites shaded according to their average occupancy. Open spheres show Si atoms, grey spheres denote Ge atoms, and black spheres show C atoms. (a) Base dot layer. (b) Top substrate layer. Solid lines (guide to the eye) enclose the area below the island.

the island, as is done in experiment. This is equivalent to 1.6 ML Ge coverage. The structure is then equilibrated and relaxed by intermixing at 800 K. Again, we find that the dot remains free of C. But, most importantly, the terrace is also free of C, and the layer below contains a very small amount of it. Instead, the C content is maximized in the third layer. This is clearly demonstrated in Fig. 5, which compares this case with the C profile for 0.5 ML Ge coverage. Obviously, the Ge-C repulsion forces C into deeper layers. On the other hand, Ge atoms are found to stay on top. We conclude that the relaxed Ge overlayer covers the Si-C geometries.

This has an important consequence. The average of the site occupancies over the whole island yields its Ge content. This comes out to be $\sim 60\%$, for 0.36 ML C, compared to $\sim 50\%$ for the bare case [21]. We interpret this to mean that C atoms in deeper layers act as a trap of Si atoms. This decelerates the diffusion of Si in the dot, and the out-diffusion of Ge, so enhancing the Ge content and providing better confinement conditions. We expect that for 2.5 ML Ge, where the PL signal is at the maximum [22], the enhancement of Ge content in the dot will be even stronger.

4. Conclusions

We have described an efficient MC algorithm which has been able to provide firm answers about the composition and stress field of C-induced

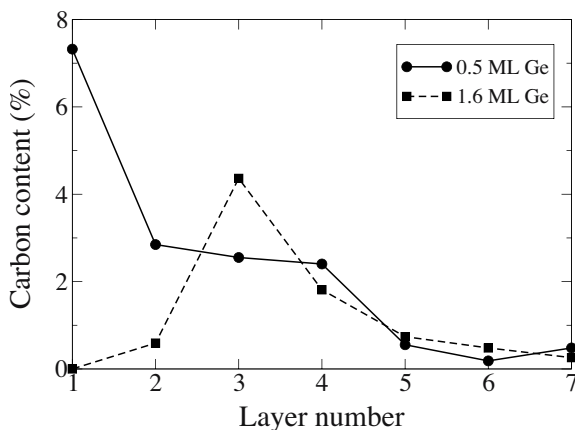


Figure 5. Carbon content in substrate layers. Layer numbered 1 denotes the top layer. The C coverage is 0.36 ML.

Ge dots, and resolve the controversy between experimental studies. The dots are found to be free of C, and are located between C-rich regions. We predict enhancement of Ge content in the islands compared to the C-free case.

Acknowledgments

This work was supported by the EU RTN program under Grant No. HPRN-CT-1999-00123.

References

- [1] P. C. Kelires and E. Kaxiras. Energetics and equilibrium properties of thin pseudomorphic $\text{Si}_{1-x}\text{C}_x$ layers in Si. *Phys. Rev. Lett.* **78**: 3479–3482, 1997.
- [2] P. C. Kelires. Monte Carlo studies of ternary semiconductor alloys: Application to the $\text{Si}_{1-x-y}\text{Ge}_x\text{C}_y$ system. *Phys. Rev. Lett.* **75**: 1114–1117, 1995.
- [3] O. G. Schmidt, C. Lange, K. Eberl, O. Kienzle, and F. Ernst. Formation of carbon-induced germanium dots. *Appl. Phys. Lett.* **71**: 2340–2342, 1997.
- [4] O. G. Schmidt and K. Eberl. Photoluminescence and band edge alignment of C-induced Ge islands and related SiGeC structures. *Appl. Phys. Lett.* **73**: 2790–2792, 1998.
- [5] O. Leifeld, D. Grützmacher, B. Müller, K. Kern, E. Kaxiras, and P. C. Kelires. Dimer Pairing on the C-Alloyed Si(001) Surface. *Phys. Rev. Lett.* **82**: 972–975, 1999.
- [6] O. Leifeld, A. Beyer, D. Grützmacher, and K. Kern. Nucleation of Ge dots on the C-alloyed Si(001) surface. *Phys. Rev. B* **66**: art. no. 125312, 2002.
- [7] G. Hadjisavvas, Ph. Sonnet, and P. C. Kelires. Stress and composition of C-induced Ge dots on Si(100). *Phys. Rev. B* **67**: art. no. 241302(R), 2003.

- [8] P. C. Kelires. Simulations of carbon containing semiconductor alloys: Bonding, strain compensation, and surface structure. *Int. J. Mod. Phys. C* **9**: 357–389, 1998.
- [9] P. C. Kelires. Theoretical investigation of the equilibrium surface structure of $\text{Si}_{1-x-y}\text{Ge}_x\text{C}_y$ alloys. *Surf. Sci.* **418**: L62–L67, 1998.
- [10] D. Frenkel. Lectures notes on free-energy calculations. in *Computer Simulations in Materials Science*, edited by M. Mayer and V. Pontikis, NATO ASI Ser. E, Vol. 205 (Kluwer Academic, Dordrecht, 1991), p. 85–117.
- [11] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller. Equation of State Calculations for Fast Computing Machines. *J. Chem. Phys.* **21**: 1087–1092, 1953.
- [12] S. M. Foiles. Calculation of the surface segregation of Ni-Cu alloys with the use of the embedded-atom method. *Phys. Rev. B* **32**: 7685–7693, 1985.
- [13] P. C. Kelires and J. Tersoff. Equilibrium alloy properties by direct simulation. *Phys. Rev. Lett.* **63**: 1164–1167, 1989.
- [14] B. Dünweg and D. P. Landau. Phase diagram and critical behavior of the Si-Ge unmixing transition: A Monte Carlo study of a model with elastic degrees of freedom. *Phys. Rev. B* **48**: 14182–14197, 1993.
- [15] M. Laradji, D. P. Landau, and B. Dünweg. Structural properties of $\text{Si}_{1-x}\text{Ge}_x$ alloys: A Monte Carlo simulation with the Stillinger-Weber potential. *Phys. Rev. B* **51**: 4894–4902, 1995.
- [16] F. Stillinger and T. Weber. Computer simulation of local order in condensed phases of silicon. *Phys. Rev. B* **31**: 5262–5271, 1985.
- [17] J. Tersoff. Modeling solid-state chemistry: Interatomic potentials for multicomponent systems. *Phys. Rev. B* **39**: 5566–5568, 1989.
- [18] I. N. Remediakis, E. Kaxiras, and P. C. Kelires. Thermodynamics of C incorporation on Si(100) from *ab initio* calculations. *Phys. Rev. Lett.* **86**: 4556–4559, 2001.
- [19] Ph. Sonnet, L. Stauffer, A. Selloni, A. De Vita, R. Car, L. Simon, M. Stoffel, and L. Kubler. Energetics of surface and subsurface carbon incorporation in Si(100). *Phys. Rev. B* **62**: 6881–6884, 2000.
- [20] Y.-W. Mo, D. E. Savage, B. S. Swartzentruber, and M. G. Lagally. Kinetic pathway in Stranski-Krastanov growth of Ge on Si(001). *Phys. Rev. Lett.* **65**: 1020–1023, 1990.
- [21] Ph. Sonnet and P. C. Kelires. Monte Carlo studies of stress fields and intermixing in Ge/Si(100) quantum dots. *Phys. Rev. B* **66**: art. no. 205307, 2002.
- [22] A. Beyer, O. Leifeld, S. Stutz, E. Müller, and D. Grützmacher. *In-situ* STM analysis and photoluminescence of C-induced Ge dots. *Nanotechnology* **11**: 298–304, 2000.

GROWTH INFORMATION CARRIED BY REFLECTION HIGH-ENERGY ELECTRON DIFFRACTION

Ákos Nemcsics

Hungarian Academy of Sciences, Research Institute for Technical Physics, and Materials Science, MTA-MFA, P.O.Box 49, H-1525 Budapest, Hungary

nemcsics@mfa.kfki.hu

Abstract Scientific and technological developments have made it possible to grow materials with different properties onto each other, and this way we can build quantum wells, quantum islands, quantum dots (QDs), etc., which leads to the possibility of creating novel devices and applications. Molecular-beam-epitaxy (MBE) is the nearly exclusive technique of growth of the above mentioned low-dimensional structures. The technology of growth under UHV made the *in-situ* observation of the growth process possible, which is widely realized by reflection high-energy electron-diffraction (RHEED). The growth of perfect crystal layers and low-dimensional structures is basically conditioned by the control of epitaxy. We need the knowledge and understanding of the growth mechanism for this, and the RHEED pattern and its intensity oscillations carry information to help us attain this goal. We will briefly deal with the basic information that is carried by RHEED. After that we investigate the dependence of mechanical strain appearing in the layer, material dependence, and other particular behaviour on RHEED. Finally, we discuss the relation between observed RHEED and the QD formation.

1. Some Aspects of RHEED Phenomena

Reflection high-energy electron-diffraction (RHEED) is a widely used monitoring technique during molecular-beam-epitaxial (MBE) growth. The orientation, quality and reconstruction of the grown surface can be determined by the RHEED pattern. Compared to other *in situ* investigations, the glancing-incidence-angle geometry of RHEED is ideally for the observation of growth process and furthermore has high surface sensitivity. The penetration of the electron beam into the surface can be varied by changing the incidence angle. The intensity of

the RHEED pattern oscillates under appropriate conditions during the growth process, a lucid geometrical explanation of which has been given by Joyce [1]. One period of these oscillations corresponds exactly to the growth of one complete monolayer (ML) in a layer-by-layer growth mode. The growth rate and the composition, in the case of alloy materials, can be determined from these oscillations.

RHEED and its intensity oscillations are very complex phenomena. This technique is a versatile tool for *in situ* monitoring, in spite of the fact that we do not know many details of its nature. RHEED oscillations are characterized by their period, amplitude, phase and damping, the behaviour at the initiation of growth, the recovery after growth and the frequency distribution in the Fourier spectrum of the oscillations. The origin and description of RHEED oscillations have been investigated by several authors [2–9]. Many properties and behaviours of the oscillations are not yet understood. For example, some outstanding issues are the different phases of the specular and non-specular beams [10], and the different behaviour of the oscillations in the case of III–V and II–VI materials [11]. There are still more interesting and open problems on the topic of the decay of intensity oscillations and of the initial phase change, etc. Several authors have attempted to describe these phenomena. References [12, 13] give a review of their models. Several effects can be interpreted by the above mentioned geometrical description: the oscillations in the case of two-dimensional growth, the disappearance of oscillations by step propagation, and the exponential decay of the oscillations [12, 14].

2. Direction of the Incident Electron Beam and the Resulting Information

2.1 Experimental Observations and Models

RHEED is an appropriate technique to investigate the binding properties on the surface. Furthermore it is appropriate to distinguish special crystallographical directions, for which X-ray diffraction is unsuitable. These directions are, e.g. on the (001) surface, where the [110] and $[\bar{1}\bar{1}0]$ directions have different properties crystallographically and from the technological view point of view as well. However, dynamical RHEED is appropriate to determine and calibrate the growth parameters, such as, e.g. growth rate and molecule flux [14–16]. One of the interesting problems is the initial phase change of the oscillations, which depends on the direction of the incident electron beam and the growth parameters.

The amplitude and period of the initial RHEED oscillations are different from those later during growth. Except for the first period, the measured decay of the oscillations fits well to an exponential function [14]. The incident electron beam impinges on the surface with low angle. If we change the incident or the azimuthal angle, the initial phase of the oscillations changes as well. For our investigations of phase changes, we will use GaAs(001) as our model material. The measured “rocking” data for the specular spot is presented in Refs. [10, 12, 13, 17]. There are several models that describe this phenomenon over a broad range of incidence angles [9, 13, 18–21]. These, generally scattering-based models, take into account several MLs and describe qualitatively well the phase switch in the range of incident angles up to about 4° , but the fitting, e.g. at low angle under 2° , has some deviations from the measured curve.

In the following section we will present a model based on a geometrical picture, where we take into account the change of coherence length of electron beam [22, 23]. Data points were obtained by measuring the time $t_{3/2}$ to the second minimum and normalizing by the period T at steady state. These data are obtained *versus* the incident angle at two different azimuthal directions for the GaAs(001) surface (Fig. 1). The substrate temperature during growth was 580°C [12] and 600°C [10], respectively. The energy of the electron beam was given only in the first case [12], which was 12.5 keV.

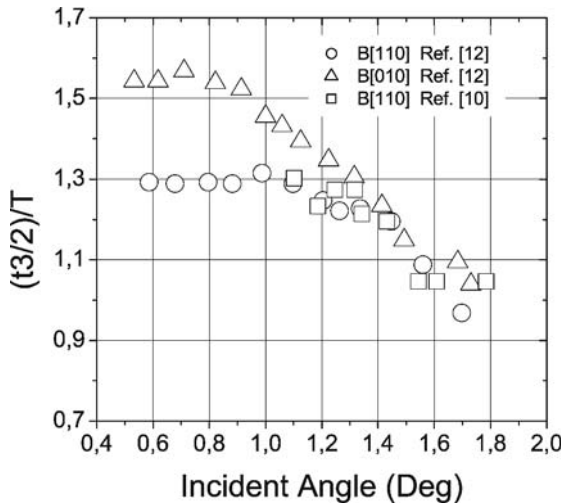


Figure 1. The experimental “rocking” data of RHEED oscillations along the [110] and [010] azimuthal directions of the GaAs(001) surface.

2.2 Description Based on the Coherence Length

The simple kinematic theory does not predict the phase shift of the oscillations, which depends on the condition of the electron beam. The contribution of inelastic processes such as Kikuchi scattering to the phase shift phenomenon is not completely taken into account [10]. The RHEED phenomenon is partly reflection-like and partly diffraction-like. The effect of the phase shift is described by the position of the minima of the oscillations. The behaviour of the minima and maxima of the oscillations can also be explained with a geometrical picture, which will be employed in this case. Because the specular spot is not a reflected beam, the interaction of the electron beam and the target surface must be described quantum mechanically. The glancing-incidence-angle electron beam touches the surface over a large area. The reflected-diffracted information obtained does not come from the whole area. The interaction between the surface and the electron beam exists only under special conditions, so we need to consider the surface coherence length w . Beeby introduced the quantum mechanically reasoned surface coherence length [24]. The wave function after the interaction $|X_a\rangle$ is written as follows:

$$|X_a\rangle = W_B(\mathbf{r}_1)W_T(\mathbf{r}_2)|X_{\bar{a}}\rangle, \quad (1)$$

where $W_B(\mathbf{r})$ and $W_T(\mathbf{r})$ are the wave packets of the electron beam and the interacting surface of the target, respectively. From the solution of this equation it can be shown that the surface coherence length depends on the interaction potential between the incident electron and the target and depends less on the wave packet [24].

We can suppose that the surface coherence length is of the same order as the coherence length Λ of the beam. In the above-mentioned experiments, the energy E of the electron beam is of the order of 10 keV, with a de Broglie wavelength $\lambda = 12.2 \times 10^{-12}$ m. The coherence length of the electron beam is between 12.2 nm and 3.7 nm [22]. The spot size of the illuminating electron beam on the surface in the incident direction depends strongly on the incident angle. The size of the touching area between the beam and the surface in the case of unit beam width can be seen in Fig. 2. We can suppose that the surface coherence length depends on the incident angle as well.

The relation between the size s of characteristic growth terrace and the surface coherence length w in the case of a polycrystalline surface was investigated in Ref. [24]. This concept can be applied in our case. An estimate of the characteristic dimension of a growth terrace can be given from experiments. The terrace average width s and the migration length ℓ of Ga depend on the substrate temperature. RHEED oscillations are

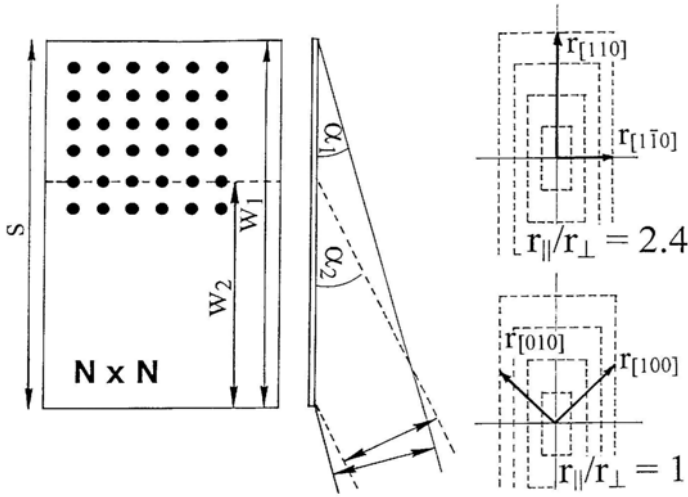


Figure 2. The change of touching length of the electron beam *versus* incident angle. The left side of the figure shows one island with lattice nodes. The right side illustrates the ratio of $r_{\parallel}/r_{\perp}$ if the observation direction changes towards $[110]$ and $[010]$.

present if $\ell \leq s$ and absent if $\ell \geq s$. In our case, the migration length is 7 nm because the substrate temperature is 580°C [12].

The binding energy on the (001) surface in the $[110]$ and $[1\bar{1}0]$ directions is not the same, which explains the dangling bonds. This anisotropy is manifested by the different growth rates. The growth rate in the $[110]$ direction is larger than that in the perpendicular direction [25]. This anisotropy is apparent not only in the growth of the crystal (in other words, the composition of the crystal) but also in the etching (that is, decomposition) of the crystal. The growth rates in the $[110]$ and $[1\bar{1}0]$ directions are different by more than a factor of two. This factor can be estimated with the help of the ratio of the etching rates along the two directions [26].

We can suppose that the surface coherence length and the average terrace width have commensurate dimensions at glancing-incidence angles ($w \approx s$). This supposition seems right, because the touching length of the electron beam (also the surface coherence length after our supposition) changes very abruptly at angles less than 1° , and in this region the function $t_{3/2}/T$ is constant accordingly, as $w > s$. The relation between the surface coherence length and average terrace width is changed with changes of the incident angle. If the incident angle increases, the surface coherence length becomes smaller than the average terrace width

($w < s$), so thus reflected-diffracted information comes from only a part of the average terrace.

For the calculation we used the polynuclear growth model in the two-dimensional case [6]. The simplified picture takes into consideration diffraction contributions only from the top most layer and the RHEED intensity is taken as proportional to the smooth part of the surface top layer [14]. The computing model is based on lattice node arrangement [23]. The relation between the terrace size and the area of surface coherence is shown in left part of Fig. 2. The surface area supplying the information decreases with the increasing incident angle of the beam. The different crystallographic directions mean different growth rates. Here, the ratio $r_{[110]}/r_{[\bar{1}\bar{1}0]}$ is estimated to be 2.4 at a growth temperature of 600°C. The oscillations were computed for two different ratios of $r_{\parallel}/r_{\perp}$, where $r_{\parallel}$ and $r_{\perp}$ are the components of the growth rate in the observation direction (parallel with the electron beam) and the perpendicular direction, respectively (right part of Fig. 2.). The calculated function of $t_{3/2}/T$ versus azimuthal angle in the two different directions is shown in Fig. 3. The growth time for one complete monolayer in the two different directions is the same (T), but the phase is different ($t_{3/2}$) because of the anisotropic growth rate. These curves correspond with the measured data. If the surface coherence length is larger than the average terrace width then the $t_{3/2}/T$ ratio remains constant (which

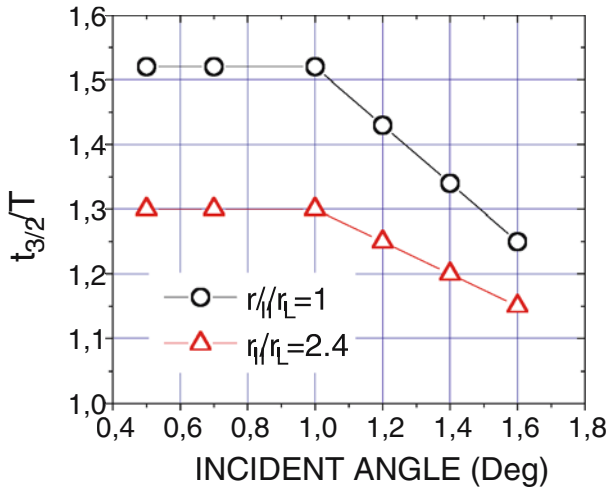


Figure 3. The computed $t_{3/2}/T$ ratio versus incident angle in different crystallographic directions (in the case of $r_{\parallel}/r_{\perp} = 2.4$ and $r_{\parallel}/r_{\perp} = 1$).

constant value is determined by the $r_{\parallel}/r_{\perp}$ ratio). If the surface coherence length is smaller than the average terrace width, then the $t_{3/2}/T$ ratio decreases as well.

The behaviour of $t_{3/2}/T$ versus incident angle was investigated under 1.8° glancing-incidence-angles. In real situations, the diffracted-reflected electron beam gets information not only from the topmost monolayer. A larger incident angle causes a larger penetration depth. The description of this phenomenon probably can be improved in either range by considering more monolayers below the surface during the growth process. The calculated curves for the decay of the oscillations correspond very well with the measured data in the investigated range.

3. The 2D-3D Transition and the Damping of Oscillations

3.1 Intensity Measurements

It is very important to investigate the 2D-3D transition during the growth process to understand the formation of quantum dots (QDs). The study of the damping of the RHEED intensity oscillation is a possible and useful *in situ* technique to observe 2D-3D transition of the grown layer. In the next section, the information which is carried by the intensity damping will be demonstrated.

InGaAs/GaAs and InAs/GaAs heterostructures are of a great importance not only in semiconductor technology, but because they serve as model system for the investigation of lattice mismatch growth and of QD formation. The growth of InGaAs layers on GaAs (001) substrate was carried out by MBE at 490°C deposition temperature. In this case the accurate intensity of RHEED oscillations was measured directly by means of a high resolution Faraday cup. The growth rate and the composition were determined by the help of RHEED oscillations. The composition change was caused by the increase of In flux while the other growth parameters remained unchanged. More details about this growth technology are described in Ref. [14].

The RHEED pattern and oscillation intensities are recorded by video camera focused on the fluorescent screen of MBE chamber. The fluorescent film on the window and the camera have non-linear behaviour in the intensity, which can deform the observed intensity distribution. The direct measurement of electron current is more advantageous than the indirect mode but its application and adjustment are more cumbersome. If we want to investigate real intensity decay of the oscillation we must measure the electron current directly.

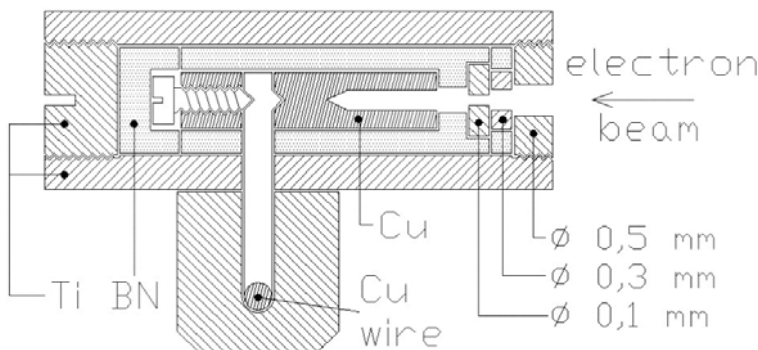


Figure 4. Cross section of the Faraday cup used for the direct measurement of electron current during the MBE growth.

In this experiment the RHEED-intensity oscillations were measured directly by means of a Faraday-cup (Fig. 4.). This cup had three pin-holes along a line of decreasing diameters (0.5, 0.3 and 0.1 mm) in order to obtain good angular resolution and suppression of the electron background. The cup was attached to a precision manipulator. The distance (about 30 cm) between the cup and the sample resulted in a good angle resolution (0.33 mrad). The current which was typically in the 100 pA range was measured with an electrometer. We have received exact exponential function for the decay with the help of direct measurement of electron current.

3.2 Interpretation of the Damping

Let us investigate the decay of oscillation in the case of InGaAs growth on GaAs(001). With the increase of In composition the decay of the oscillations became faster. The least-squares-method was used to fit an exponential function to the decay of intensity. Except for the first period the measured decays fits well to an exponential function. As mentioned above, the first period is the initial phase of growth: it is characterized not only by a period length different from that of the following oscillations, but also by an unusual intensity.

The exponential decay can be roughly modeled within a simplified picture, which only takes into consideration diffraction contributions from the topmost layer. Let us suppose that at the start of the growth the surface of area A is perfectly smooth, i.e. $\theta_0 = 1$, where θ_0 is the coverage of the uppermost layer. In order to describe the increase of surface roughness during growth, we introduce a coverage reduction factor b

that gives the reduction in coverage from layer to layer. This factor may be regarded as a measure for the smoothness, therefore b can also be considered as an epilayer quality ideality factor [14]. The first layer grown on the perfect surface is not complete, but has a coverage $\theta_1 = b$ with $b < 1$, but close to unity. The next layer has $\theta_2 = b^2$, and for the n th layer,

$$\theta_n = b^n. \quad (2)$$

If, in a first approximation, the RHEED intensity I is taken as proportional to the smooth part of the surface top layer, the intensity of the perfectly smooth surface is given by $I_0 = cA$, where c is a constant that characterizes the diffraction power. After the growth of the first layer the intensity is $I_1 = bcA$ and, after the n th layer,

$$I_n = b^n cA. \quad (3)$$

A continuous description is obtained by replacing of n by rt , which yields $I(t) = b^{rt} cA$, where r is the growth rate, given by $r = 1/\tau_{\text{InGaAs}}(x)$ and t is the growth time. This can be written in the following form:

$$I = cAe^{-t/\tau_d}, \quad (4)$$

where τ_d is the decay time:

$$\tau_d = -\frac{1}{r \ln b}, \quad (5)$$

with $\tau_d > 0$, since $b < 1$. In the fit of the experimental intensity decay the fit parameter was the coverage reduction factor, the growth rate being calculated from the period of the RHEED-oscillations. The growth rate increases with increasing composition x and therefore the decay time τ_d should decrease with x , if b is constant. The decay time constant is a function of composition, in agreement with our experimental results as shown in Fig. 5. Calculating the coverage reduction factor from the experimentally observed decay time shown in Fig. 5 by Eq. (5), one can see that b is not a constant with respect to composition x , but decreases with increasing x . A qualitative explanation can be given for this behavior. If the composition increases, the growth rate increases due to the preparation conditions. Therefore, the Ga and In atoms at the surface have less time for diffusion along the surface, resulting in less perfect layers, i.e. in a smaller coverage reduction factor b . If a linear relation between b and $1/r$ is assumed, this results in:

$$\frac{b(x_i)}{\tau_{\text{InGaAs}}(x_i)} = \frac{b(x_j)}{\tau_{\text{InGaAs}}(x_j)}, \quad (6)$$

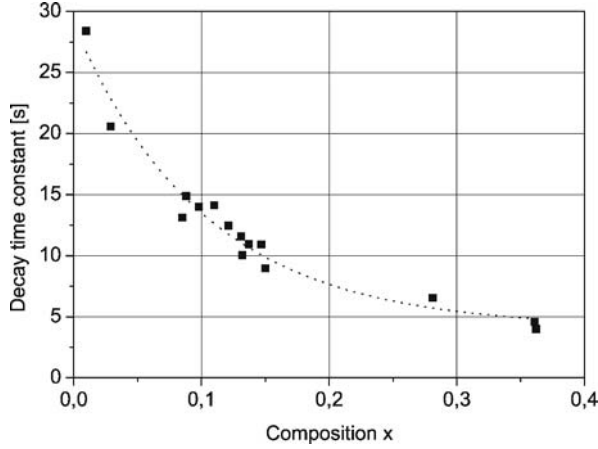


Figure 5. Relationship between decay time constant with the intensity of RHEED oscillation and composition, where symbols and dashed line represent the measured data and the fitted function, respectively.

where $b(x_i)$ and $b(x_j)$ are the coverage reduction factors and $\tau_{\text{InGaAs}}(x_i)$ and $\tau_{\text{InGaAs}}(x_j)$ the periods of RHEED oscillations for compositions x_i and x_j , respectively. Using Eq. (5) for the decay time as a function of composition, i.e. $\tau_d(x) = -1/r(x) \ln[b(x)]$, from the experimental results shown in Fig. 5, one can calculate $b(x)$. With $\tau_{\text{InGaAs}}(x) = 1/r(x)$ known from the RHEED oscillations, we obtain $b(x)\tau_{\text{InGaAs}} = 0.30 \text{ s}^{-1}$ within an accuracy of 10%. From this we may conclude that the model used here, in spite of its simplicity, gives a satisfying first-order explanation of the RHEED-intensities observed in the experiment.

3.3 Estimation of the Critical Layer Thickness

The above treated model material system will be used henceforward in the following section as well. The critical layer thickness (CLT) during heteroepitaxy was investigated experimentally as well as theoretically [27]. Under real growth conditions the measured CLT depends not only on the misfit but on the growth parameters as well [28]. RHEED intensity oscillations are used for accurate determination of the threshold layer thickness in the 2D growth mode which is also growth-condition dependent. The lattice mismatch also affects the behaviour of RHEED oscillations. The oscillation amplitude decreases rapidly during the InGaAs growth due to its strong dependence on the InAs mole fraction. This decrease also indicates the formation of the 3D growth mode.

The magnitude of intensity oscillations continuously decreases during the growth process. With the increase of In composition the growth rate and the decay of oscillations both increase. We have investigated the decay of amplitude oscillations (peak to peak). The least-squares method was used to fit an exponential function to the experimental intensity decay data [see the first part of Eq. (7)], where τ_d is a decay time constant and A_0 is a fitting factor. The measured decay fits very well to an exponential function except for the first period. This description for $I(t)$ is valid for all x values in the investigated range ($0 \leq x \leq 0.4$) [14, 29]. With increasing composition x the decay time τ_d decreases. The measured value of the decay time constant is a function of composition $\tau_d(x)$ (Fig. 5.), where the exponential function was fitted by the least-squares method. The fitted function is $\tau_d = a \exp(-x/b) + c$, where $a = 24.85$, $b = 0.103$ and $c = 4.1$ [29].

The measured values of the decay time constants depend on the composition. Most probably there is a relationship between the behaviour of the RHEED oscillation decay and CLT. The oscillation amplitude decreases rapidly during InGaAs growth due to the 3D growth, which has a strong dependence on the InAs mole fraction. Not only the mismatch is responsible for this decay but the other growth conditions as well. The decay of the oscillation amplitude exists even without mismatch e.g. at pure GaAs layer growth. The changes of the In content modify the mismatch and the growth conditions (e.g. growth rate). Therefore, the misfit and the growth parameters have a joint influence on the behaviour of amplitude decay.

The exponential function is a good approximation to the $I(t)$ function at every x value. Provided that both processes are independent from each other, the phenomenon of decay at arbitrary x values can be described by two time constants as follows:

$$I(t) = A_0 \exp\left(-\frac{t}{\tau_d}\right) = A_0 \exp\left(-\frac{t}{\tau_M} - \frac{t}{\tau_G}\right), \quad (7)$$

where τ_M and τ_G are the assumed time constants of the separate effects that are responsible for the influence of misfit and the growth, respectively. According to this model, the x -dependent decay from misfit can be written as follows:

$$\frac{1}{\tau_M(x)} = \frac{1}{\tau_d(x)} - e(x), \quad (8)$$

where the $\tau_d(x)$ and $\tau_M(x)$ time constants are functions of the composition x , and $e(x)$ contains the dependence from decay constants. In the case of pure GaAs growth ($x=0$), there is no decay caused by misfit.

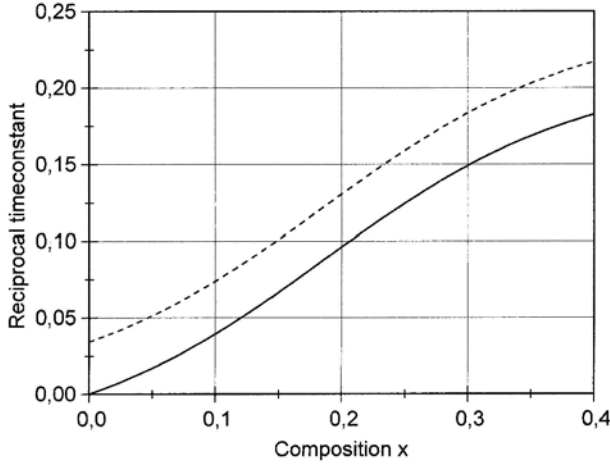


Figure 6. Reciprocal function of decay time constant where the solid and dashed lines represent the $1/\tau_M$ and $1/\tau_d$, respectively.

This means that, at $x=0$, the reciprocal value of the decay time constant originates entirely from crystal growth phenomenon. The function $\tau_d(x)$ is known. The value of $1/\tau_M$ at $x=0$ is zero. As a first approximation, we can neglect the x dependence of negative part in Eq. (8), so the $1/\tau_M$ function describing the mismatch can be obtained by moving the $1/\tau_d$ function in the negative ordinate direction (Fig. 6). We can easily determine $\tau_M(x)$ from its reciprocal form. In order to be able to compare the calculated decay with the theoretical CLT, we should change to layer thickness instead of period time, because now we only have function of time *versus* x [29]. The grown thickness under one period of oscillation is one monolayer, which corresponds to half of the lattice parameter. After evaluating the function of thickness *versus* x instead of the function of time, we can compare the calculated threshold thickness with the theoretical CLT. As can be seen in Fig. 7. there is a relatively good agreement between our calculation and the isotropic and anisotropic Matthews model [30].

As a first approximation, it was supposed that the dependence of $1/\tau_G$ on the composition x is not significant, i.e. the second term in Eq. (2) is constant. Although we have received similar course between the theoretical CLT and our calculated threshold layer thickness, one can find difference between curves as well. The reason of the deviance can be in the fact that the τ_G also depends on the composition.

We can investigate the influence of growth conditions in the case of pure GaAs growth. We can increase the growth rate by increasing of

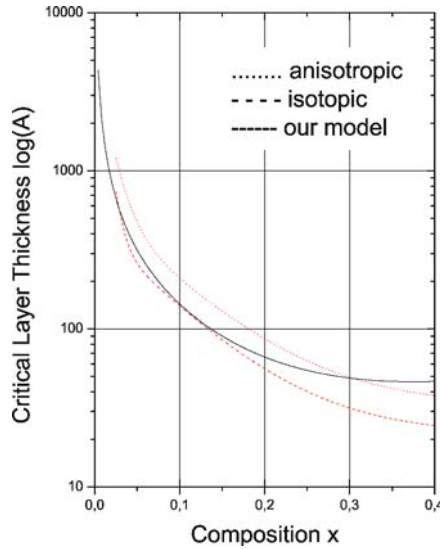


Figure 7. Threshold layer thickness calculated by us is compared by isotropic and anisotropic Matthews model.

the Ga flux. Let us see an example! At the Ga source temperature of 935°C the growth rate is $r = 3.85 \text{ s}^{-1}$ ($\tau_{\text{GaAs}} = 2.61 \text{ s}$) [3]. In this case the decay of oscillations can be described with decay constant $\tau_d = 15.3 \text{ s}$, where, because of the homoepitaxy, τ_M is missing, therefore τ_d and τ_G are equal. The above mentioned growth rate corresponds to the case of InGaAs growth ($\tau_{\text{InGaAs}} = 2.61 \text{ s}$) $x = 0.13$ where ($T_{\text{Ga}} = 925^{\circ}\text{C}$ and $T_{\text{In}} = 540^{\circ}\text{C}$). With the increase of the growth rate the decay constant decreases. In the case of InGaAs growth the increase of growth rate means a simultaneous increase of InAs mole fraction. The more pronounced decay of oscillations can be traced back not only to the change of growth rate but to change in the In content as well. The course of RHEED intensity oscillations strongly depends on the observed material. The observed oscillation is weaker and has larger decay in the case of InAs growth on InAs (001) than in the case of GaAs homoepitaxy under similar growth conditions [31]. The change of τ_G can be caused by the change of growth rate and/or change of In/Ga ratio. The supposed tendency of τ_d is supported by both the above mentioned phenomena. The threshold thickness is defined as the thickness where the oscillations become weak enough. The supposed dependence of τ_G on the growth rate/composition seems valid because the deviation between CLT and the calculated data diminishes. The growth rate dependence was proved

in the case of InGaAs growth by *in situ* XRD measurement where the deposition temperature was 490°C (the same as in our case) [32].

This description shows only the correct trend of the τ_G dependence. Reality is more complicated. We have not taken into account, e.g. the material contrast and the different sticking coefficient of Ga and In. The CLT derived from this assumption shows relatively good agreement with CLT data from the literature. We can improve this agreement by including the composition/growth rate dependence of τ_G .

3.4 RHEED Studies of QD Formation

RHEED is a many sided *in situ* investigation technique during growth, which is appropriate to observe the QD formation and to determine some parameters of QD. Besides the above treated 2D-3D transition [33, 34], we can investigate the wetting layer [33, 35], characteristic times of the formation process [34, 36], critical coverage of the layer [33, 35, 37], segregation of the components during QD formation [38], shape of the dot, change of the lattice mismatch and surface lattice parameter [39]. For example, the characteristic times of the QD formation can be determined with the help of the pattern change and the intensity of the RHEED pattern [37]. The shape of the forming dot can be investigated also by the pattern and by the so-called chevron-shape spot of RHEED [40]. RHEED gives more possibilities to investigate the parameters. For example, the 2D-3D transition can be investigated not only with the help of oscillation decay [33] but also with the help of change of diffraction pattern [37, 40, 41] and chevron-shape spot [41].

4. Summary

RHEED is a many-sided *in situ* monitoring technique, which is necessary for growth technology and understanding growth mechanisms. The information for us is supplied by the electron beam, which interacts with the growing crystal surface. The information concerning the growth must be selected from the diffracted-reflected electron image. The investigation and understanding of this interaction is not only basic research, but is also indispensable for crystal growth. This investigation technique is also useful for the study of QD formation.

Acknowledgments

The author is indebted to the organizers for the invitation to this workshop. Some work mentioned here was supported by Hungarian National Scientific Foundation (OTKA) through Grant Nos. T030426 and T037509, which are very gratefully acknowledged.

References

- [1] B. A. Joyce. in *Two-dimensional systems: Physics and New Devices*, G. Bauer, F. Kucher, H. Heinrich, editors, (Springer Series in Solid State Sciences **67**, Springer-Verlag, Heidelberg, 1986).
- [2] C. S. Lent and P. I. Cohen. Diffraction from stepped surfaces I: Reversible surfaces. *Surf. Sci.* **139**: 121–154, 1984.
- [3] P. J. Dobson, B. A. Joyce, J. H. Neave, and J. Zhang. Current understanding and applications of the RHEED intensity oscillation technique. *J. Cryst. Growth* **81**: 1–8, 1987.
- [4] S. Clarke and D. D. Vvedensky. Origin of reflection high-energy electron-diffraction intensity oscillations during molecular-beam-epitaxy: A computational modelling approach. *Phys. Rev. Lett.* **58** 2235–2238, 1987.
- [5] L.-M. Peng and M. J. Whelan. Dynamic RHEED from MBE growing surfaces. *Surf. Sci.* **238**: L446–L452, 1990.
- [6] Y. Ishibashi. Intensity oscillation of reflection high-energy electron-diffraction (RHEED) by a polynuclear growth model. *J. Phys. Soc. Jap.* **60**: 3215–3217, 1991.
- [7] J. Resh, J. Strozier, A. Bensaoula, and A. Ignatiev. Rate equation models of epitaxial growth on stepped surfaces. *J. Vac. Sci. Technol. A* **9**: 1551–1557, 1991.
- [8] L.-M. Peng, S. L. Dudarev, and M. J. Whelan. Bethe potentials in dynamical RHEED calculations. *Surf. Sci.* **351**: L245–L252, 1996.
- [9] W. Braun, L. Däweritz, and K. H. Ploog. Origin of electron diffraction oscillations during crystal growth. *Phys. Rev. Lett.* **80** 4935–4938, 1998.
- [10] J. Resh, K. D. Jamison, J. Strozier, A. Bensaoula, and A. Ignatiev. Phase of reflection high-energy electron diffraction intensity oscillations during molecular-beam epitaxy growth of GaAs(100). *Phys. Rev. B* **40**: 11799–11803, 1989.
- [11] J. Griesche, N. Hoffmann, M. Rabe, and K. Jacobs. Developments in the use of RHEED for interpreting growth processes in the MBE of wide gap II-VI semiconductors. *Prog. Cryst. Growth* **37**: 103–121, 1998.
- [12] B. A. Joyce, J. H. Neave, J. Zhang, P. J. Dobson, in: P. K. Larsen, P. J. Dobson (Eds) *Reflection High-Energy Electron Diffraction and Reflection Imaging of Surfaces*, NATO ASI Series B: Physics: Vol. 188 Plenum Press, New York 1987
- [13] W. Braun. *Applied RHEED, Reflection High-Energy Electron Diffraction during Crystal Growth* (Springer-Verlag, Berlin, 1999).
- [14] Á. Nemcsics, J. Olde, M. Geyer, R. Schnurpfeil, R. Manzke, and M. Skibowski. MBE growth of strained $\text{In}_x\text{Ga}_{1-x}\text{As}$ on GaAs(001). *Phys. Stat. Sol. (a)* **155**: 427–437, 1996.
- [15] A. J. Springthorpe and A. Majeed. Epitaxial growth rate measurements during molecular-beam epitaxy. *J. Vac. Sci Technol. B* **8** 266–270, 1990.
- [16] Ch. Heyn and M. Harsdorff. Flux control and calibration of an As effusion cell in a molecular-beam epitaxy system for GaAs and AlGaAs with a quadrupole mass spectrometer. *J. Cryst. Growth* **133** 241–245, 1993.
- [17] J. Zhang, J. H. Neave, P. J. Dobson, and B. A. Joyce. Effects of diffraction conditions and processes on RHEED intensity oscillations during the MBE growth of GaAs. *Appl. Phys. A* **42** 317–326, 1987.

- [18] Z. Mitura, S. L. Dudarev, and M. J. Whelan. Phase of RHEED oscillations. *Phys. Rev. B* **57**: 6309–6312, 1998.
- [19] Z. Mitura. Iterative method of calculating reflection high-energy-electron-diffraction intensities. *Phys. Rev. B* **59**: 4642–4645, 1999.
- [20] Z. Mitura, S. L. Dudarev, and M. J. Whelan. Interpretation of reflection high-energy electron diffraction oscillation phase. *J. Cryst. Growth* **198/199** 905–910, 1999.
- [21] J. M. McCoy, U. Korte, P. A. Maksym, and G. Meyer-Ehmsen. Multiple scattering evaluation of RHEED intensities from the GaAs(001)-(2 × 4) surface: Evidence for subsurface relaxation. *Phys. Rev B* **48**: 4721–4728, 1993.
- [22] Á. Nemcsics. The initial phase shift phenomenon of RHEED oscillations. *J. Cryst. Growth* **217**: 223–227, 2000.
- [23] Á. Nemcsics. Valuing of the critical layer thickness from the deading time constant of RHEED oscillation in the case of $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ heterojunction. *Appl. Surf. Sci.* **190** 294–297, 2002.
- [24] J. L. Beeby. Theory of RHEED for general surfaces. *Surf. Sci.* **80**: 56–61, 1979.
- [25] Y. Horikoshi, H. Yamaguchi, F. Briones, and M. Kawashima. Growth process of III-V compound semiconductors by migration-enhanced epitaxy. *J. Cryst. Growth* **105**: 326–338, 1990.
- [26] E. R. Messmer. *Selective epitaxy and in-situ etching on III-V semiconductor surfaces*. (Ph.D. thesis, Stockholm, 2000).
- [27] E. A. Fitzgerald. Dislocations in strained-layer epitaxy: Theory, experiment, and applications. *Mater. Sci. Rep.* **7**: 91–142, 1991.
- [28] N. Grandjean and J. Massies. Epitaxial growth of highly strained $\text{In}_x\text{Ga}_{1-x}\text{As}$ on GaAs(001): The role of surface diffusion length. *J. Cryst. Growth* **134**: 51–62, 1993.
- [29] Á. Nemcsics. Growth control of the strained InGaAs/GaAs heterostructures for device purposes by decay of RHEED oscillation. *Proceedings of the society of photo-optical instrumentation engineers (SPIE)* **3975**: 264–267, 2000.
- [30] J. W. Matthews. Coherent interfaces and misfit dislocations. *Epitaxial Growth*, Part B, (Academic Press, New York, 1975) pp. 559–609.
- [31] A. Schreckenbach, J. Olde, Á. Nemcsics, R. Mancke, and M. Skibowski. *Verhandl. DPG (VI)* **26**: 1161, 1991.
- [32] M. U. Gonzalez, Y. Gonzalez, and L. Gonzalez. *In situ* detection of an initial elastic relaxation stage during growth of $\text{In}_{0.2}\text{Ga}_{0.8}\text{As}$ on GaAs(001). *Appl. Surf. Sci.* **188**: 128–133, 2002.
- [33] T. J. Krzyzewski, P. B. Joyce, G. R. Bell, and T. S. Jones. Wetting layer evolution in InAs/GaAs(001) heteroepitaxy: effects of surface reconstruction and strain. *Surf. Sci.* **517**: 8–16, 2002.
- [34] Ch. Heyn and W. Hansen. Desorption of InAs quantum dots. *J. Cryst. Growth* **251**: 218–222, 2003.
- [35] T. J. Krzyzewski, P. B. Joyce, G. R. Bell, and T. S. Jones. Understanding the growth mode transition in InAs/GaAs(001) quantum dot formation. *Surf. Sci.* **532/535**: 822–827, 2003.

- [36] Ch. Heyn, D. Endler, K. Zhang, and W. Hansen. Formation and dissolution of InAs quantum dots on GaAs. *J. Cryst. Growth* **210**: 421–428, 2000.
- [37] Ch. Heyn. Critical coverage for strain-induced formation of InAs quantum dots. *Phys. Rev. B* **64**: art. no. 165306, 2001.
- [38] Ch. Heyn and W. Hansen. Ga/In-intermixing and segregation during InAs quantum dot formation. *J. Cryst. Growth* **251**: 140–144, 2003.
- [39] B. H. Koo, T. Hanada, H. Makino, J. H. Chang, and T. Yao. RHEED investigation of the formation process of InAs quantum dots on (100) InAlAs/InP for application to photonic devices in the 1.55 μm range. *J. Cryst Growth* **229**: 142–146, 2001.
- [40] T. Hanada, B. H. Koo, H. Totsuka, and T. Yao. Anisotropic shape of self-assembled InAs quantum dots: Refraction effect on spot shape of reflection high-energy electron diffraction. *Phys. Rev. B* **64**: art. no. 165307, 2001.
- [41] J. G. Belk, C. F. McConville, J. L. Sudijono, T. S. Jones, and B. A. Joyce. Surface alloying at InAs-GaAs interfaces grown on (001) surfaces by molecular beam epitaxy. *Surf. Sci.* **387**: 213–226, 1997.

EFFICIENT CALCULATION OF ELECTRON STATES IN SELF-ASSEMBLED QUANTUM DOTS: APPLICATION TO AUGER RELAXATION

D. Chaney, M. Roy and P. A. Maksym

University of Leicester, Leicester LE1 7RH, United Kingdom

Abstract An efficient method for calculation of self-assembled dot states within the effective mass approximation is described and its application to the calculation of Auger relaxation rates is detailed. The method is based on expansion of the dot states in a harmonic oscillator basis whose parameters are optimised to improve the convergence rate. This results in at least an order of magnitude reduction in the number of basis states required to represent electron states accurately compared to the conventional plane wave approach. Auger relaxation rates are calculated for harmonic oscillator model states and exact states for various pyramidal models. The dipole approximation, previously used to calculate Auger rates, is found to be inaccurate by a factor of around 2-3. The harmonic oscillator states do not reproduce the rates for the more realistic pyramidal models very well and even within the set of pyramidal models variations in the dot shape and size can change the rates by up to an order of magnitude. Typical Auger relaxation rates are on a picosecond timescale but the actual value is strongly dependent on the density of electron inside the dot.

Keywords: Self-assembled dot, electron states, Auger relaxation

1. Introduction

Self-assembled quantum dots are being studied intensively because of their potential applications to optoelectronics, quantum cryptography and quantum computing [1]. A number of methods for calculating the eigenstates of electrons and holes in self-assembled dots have been developed [2, 3, 4, 5] but most of the techniques used so far involve very large scale computations. As applications of dots become more important it will be necessary to go beyond the computation of energies and eigenstates and develop techniques to compute other physical properties of

dots, such as relaxation times. The calculation of the eigenstates is just one step in the problem of calculating the physical properties of a dot so it is advantageous to have a method which generates them efficiently. The purpose of this work is to review one particular approach to efficient calculation of electron states [6] together with its application to Auger relaxation [7].

One of the reasons why the calculation of self-assembled dot states involves large scale computations is that there is no natural basis in which to expand the dot eigenstates. The problem can of course be solved by expanding in any basis set, such as plane waves, but this generally leads to Hamiltonian matrices of very large dimension and makes the computation expensive. We have developed an alternative approach in which a harmonic oscillator basis is used and its parameters are optimised variationally before the Hamiltonian matrix is found. This generally leads to a large reduction in the number of basis states needed to represent the dot states accurately because the low lying basis states (i.e. the ones with a few nodes) are localised on a length scale similar to the real dot states. In Sec. 2 we describe our approach and detail the dot models to which it has been applied.

Following the discussion of dot states we consider Auger relaxation rates in Sec. 3. Until now, most work on Auger relaxation has involved model dot states with rates calculated in the dipole approximation [8]. At first sight the dipole approximation should be very accurate because of the small spatial extent of the dot states but we have found a surprising discrepancy between exact Auger rates and rates calculated in the dipole approximation. The nature of this discrepancy and the reason why it occurs are detailed in Sec. 3.2. Following this we give Auger rates for a number of pyramidal dot models, compare them with rates for a harmonic oscillator model and comment on the suitability of this model. Our conclusions are summarised in Sec. 4.

2. Harmonic Oscillator Basis Calculation of Electron States

2.1 Dot Model and Schrödinger Equation

The structure and composition of self-assembled dots have been under investigation for some time and knowledge of these systems is still evolving, see [2, 9, 10, 11] for example. In our work we have used scanning tunnelling microscopy results of Bruls *et al.* [12] as a starting point for constructing dot models. In this section we discuss the calculation of electron states for a particular model that we take to be typical, although not definitive, and in Sec. 3 we consider Auger relaxation rates for

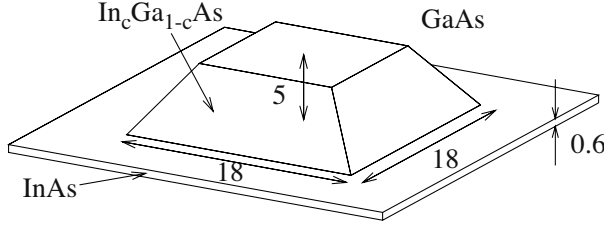


Figure 1. Dot model based on results of Ref. [12].

a number of variations of this model. However our calculation method is completely general and can be used to find the quantum states of dots of arbitrary size, shape and composition.

The particular dot studied by Bruls *et al.* (Fig. 1) is a truncated pyramid composed of $\text{In}_c\text{Ga}_{1-c}\text{As}$ with the Indium fraction c varying linearly from 0.6 at the base to 1.0 at the top. We have no information about the composition variation in the lateral direction so we have assumed that the only composition variation is in the vertical direction. The dot base dimensions are $18 \text{ nm} \times 18 \text{ nm}$, the dot height is 5 nm and the top face dimensions are $10.6 \text{ nm} \times 10.6 \text{ nm}$. The dot stands on a 0.6 nm InAs wetting layer on a GaAs substrate and is surrounded by GaAs.

The dot is strained, so the strain-dependent effective mass and confinement potential are needed to find the quantum states. We use a commercial finite element package, Abaqus, to find the strain within the continuum elasticity approximation. The strain is calculated for a dot in the centre of a large block of material of size $140 \text{ nm} \times 140 \text{ nm} \times 65.6 \text{ nm}$, with periodic boundary conditions. The large size of the periodic cell ensures that the strain in the dot is unaffected by boundary effects. The lattice mismatch strain is taken to be 6.7%, the Young's moduli and Poisson ratios for GaAs are 85.3 GPa and 0.32 while for InAs they are 51.4 GPa and 0.35 [13]. Values for $\text{In}_c\text{Ga}_{1-c}\text{As}$ are obtained by linear interpolation.

The confinement potential and effective mass are found in the standard way from the strain [14]. First, the electron and hole confinement potentials are found from the composition dependent conduction band offset, band gap and deformation potentials:

$$V_e(\mathbf{r}) = V_o(\mathbf{r}) + V_c(\mathbf{r}),$$

$$V_o(\mathbf{r}) = -1.178c(\mathbf{r}) + 0.381c(\mathbf{r})^2,$$

$$V_c(\mathbf{r}) = a_c \epsilon_h(\mathbf{r}),$$

$$\begin{aligned}
V_{hh}(\mathbf{r}) &= a_v \epsilon_h(\mathbf{r}) + \frac{1}{2} b \epsilon_b(\mathbf{r}), \\
V_{lh}(\mathbf{r}) &= a_v \epsilon_h(\mathbf{r}) - \frac{1}{2} b \epsilon_b(\mathbf{r}).
\end{aligned} \tag{1}$$

Here, the electron potential has been written as the sum of the conduction band offset, V_o , and a strain-dependent part, V_c , and an empirical relation [4] is used to obtain the composition dependence of the conduction band offset. The energy zero is at the unstrained GaAs conduction band edge. V_{hh} and V_{lh} respectively are the strain dependent contributions to the heavy and light hole potentials, ϵ_h and ϵ_b respectively are hydrostatic and biaxial strains and a_c , a_v and b are deformation potentials and their values in eV are [14]

$$\begin{aligned}
a_c(\mathbf{r}) &= -5.08c(\mathbf{r}) - 7.17[1 - c(\mathbf{r})], \\
a_v(\mathbf{r}) &= 1.00c(\mathbf{r}) + 1.16[1 - c(\mathbf{r})], \\
b(\mathbf{r}) &= -1.80c(\mathbf{r}) - 1.70[1 - c(\mathbf{r})].
\end{aligned} \tag{2}$$

The piezoelectric contribution to the confinement potential is not included in the present work but this does not affect our conclusions about the efficiency of our method. The effect of the piezoelectric potential on the energies of the dot states is thought to be small [2].

Once the potentials are known the position dependent electron effective mass is found from first order perturbation theory [14]

$$\begin{aligned}
m_z^*(\mathbf{r}) &= m^* \frac{V_c + E_g^{\text{GaAs}} - V_{lh}}{E_g}, \\
m_{xy}^*(\mathbf{r}) &= m^* \frac{(V_c + E_g^{\text{GaAs}} - V_{hh})(V_c + E_g^{\text{GaAs}} - V_{lh})}{E_g(V_c + E_g^{\text{GaAs}} - \frac{3}{4}V_{lh} - \frac{1}{4}V_{hh})},
\end{aligned} \tag{3}$$

where m_z^* and m_{xy}^* are respectively perpendicular and lateral effective masses and E_g^{GaAs} is the bulk GaAs band gap. E_g and m^* are the position dependent band gap and effective mass given by [14]

$$\begin{aligned}
E_g(\mathbf{r}) &= 0.41c(\mathbf{r}) + 1.52(1 - c(\mathbf{r})), \\
m^*(\mathbf{r}) &= 0.023c(\mathbf{r}) + 0.067(1 - c(\mathbf{r})).
\end{aligned} \tag{4}$$

Within the single band effective mass approximation the Schrödinger equation for the electron states is then

$$H\Psi = \left[\frac{1}{2}(-i\hbar\nabla + e\mathbf{A})M^{-1}(-i\hbar\nabla + e\mathbf{A}) + V(\mathbf{r}) \right] \Psi = E\Psi, \tag{5}$$

where M is the effective mass tensor and we have allowed for the possibility that the dot is in a magnetic field.

2.2 Harmonic Oscillator Basis for Solution of Schrödinger's Equation

The Schrödinger equation can always be solved by expanding the wave function in a suitable orthonormal basis, that is by putting

$$\Psi(\mathbf{r}) = \sum_i a_i \psi_i(\mathbf{r}), \quad (6)$$

where the ψ_i are the basis states. Usually it is advantageous to choose the basis to be the eigenstates of a Hamiltonian that is close to the one of interest so that the equation $H\Psi = E\Psi$ becomes $(H_0 + \Delta H)\Psi = E\Psi$ and the basis states are chosen to be the eigenstates of H_0 . If ΔH is in some sense small the eigenstates of H can be found with only a few terms in the basis set expansion and the process is computationally efficient. In many cases there is a natural separation of H into $H_0 + \Delta H$, where the eigenstates of H_0 are easy to find and ΔH is small, but this does not happen for self-assembled dots because of the complicated nature of the confinement potential. Instead most authors have used a plane wave basis [3, 4, 5] or a real space mesh [2, 15] to find the dot states, however both approaches are computationally expensive.

In our alternative approach the basis is optimised before use so that the lowest few basis states (the ones with only a few nodes) are localised on about the same length scale as the dot states. We take the basis states to be products of lateral and vertical harmonic oscillator states so that

$$\begin{aligned} \psi_i(\mathbf{r}) &= \Phi_{l_i}(\phi) Z_{m_i}(z) R_{n_i l_i}(r), \\ \Phi_{l_i}(\phi) &= \frac{1}{\sqrt{2\pi}} e^{il_i\phi}, \\ Z_{m_i}(z) &= \left(\frac{1}{2^{m_i} m_i! \lambda_z \sqrt{\pi}} \right)^{\frac{1}{2}} H_{m_i} \left[\frac{(z - z_o)}{\lambda_z} \right] e^{-(z - z_o)^2 / 2\lambda_z^2}, \\ R_{n_i l_i}(r) &= \left(\frac{n_i!}{\lambda_r^2 (n_i + |l_i|)!} \right)^{\frac{1}{2}} e^{-r^2 / 4\lambda_r^2} \left(\frac{r}{\sqrt{2}\lambda_r} \right)^{|l_i|} L_{n_i}^{|l_i|} \left[\frac{r^2}{2\lambda_r^2} \right]. \quad (7) \end{aligned}$$

Here, H_{m_i} is a Hermite polynomial, $L_{n_i}^{|l_i|}$ is a Laguerre polynomial, λ_z and λ_r are vertical and lateral length parameters and z_o is an offset parameter. These basis states form a complete orthonormal set for all values of λ_z , λ_r and z_o so the sum in Eq. (6) always converges to an exact dot state provided enough terms are included. We reduce the number of terms by choosing the length and offset parameters to minimise $\langle \psi_0 | H | \psi_0 \rangle$, where ψ_0 has quantum numbers $l = 0$, $m = 0$ and $n = 0$.

Thus ψ_0 is a variational approximation to the exact dot ground state and the length scale of the oscillator states is approximately the same as that of the exact dot states. This enables us to obtain accurate dot states by retaining relatively few oscillator states in the basis state expansion. The dot states and energies are found in the usual way by numerically diagonalising the resulting Hamiltonian matrix. A second advantage of the harmonic oscillator basis is that the effect of a perpendicular magnetic field can be included with relatively little additional computational effort. This is explained in Ref. [6] which also gives full details of our method.

2.3 Comparison of Harmonic Oscillator and Plane Wave Methods

Energies of electron states for the dot shown in Fig. 1, calculated with the harmonic oscillator and plane wave expansions, are given in Table 1. The harmonic oscillator results were obtained with $l_{\max} = 12$, $m_{\max} = 20$ and $n_{\max} = 20$ which corresponds to a total of 11,025 basis states. The plane wave results are for the same dot in a $60 \text{ nm} \times 60 \text{ nm} \times 60 \text{ nm}$ periodic cell. 41 plane waves are used in the vertical direction and 21 in each of the two lateral directions which corresponds to a total of 18,081 basis states. The ground state energy calculated with the harmonic oscillator basis is accurate to about 0.1 meV but the plane wave ground state energy is about 0.6% higher. The agreement between the two sets of results very good, with only a small discrepancy that results from a combination of the poorer convergence of the plane wave method and the difficulty of calculating the plane wave Hamiltonian matrix elements accurately [6].

The superior convergence of the harmonic oscillator basis calculation is illustrated in Fig. 2. The figure shows the convergence of the energies

Table 1. Comparison of energies calculated with the harmonic oscillator and plane wave bases.

State	Energy (eV)	Energy (eV)	% difference
1	-0.2247	-0.2234	0.6
2	-0.1563	-0.1564	0.1
3	-0.1563	-0.1564	0.1
4	-0.0947	-0.0960	1.4
5	-0.0720	-0.0726	0.8
6	-0.0594	-0.0606	2.0

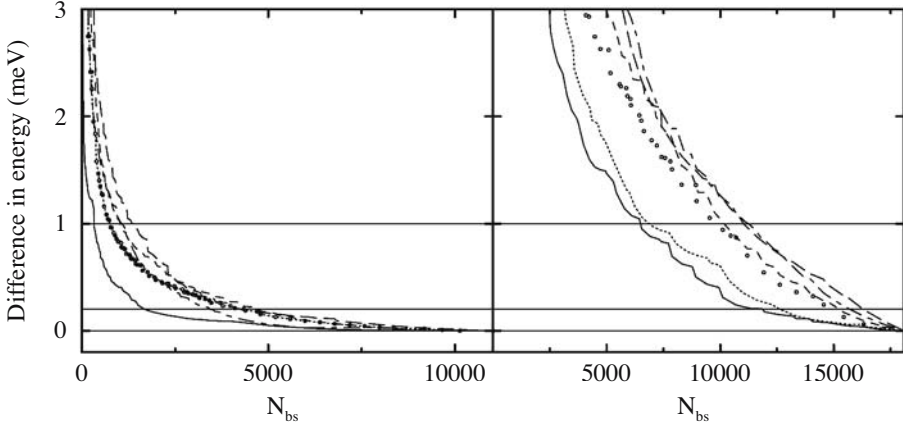


Figure 2. Rate of convergence of bound states as a function of number of basis functions included in the calculation. Solid line (ground state), dotted line (state 2), points (state 3), dashed line (state 4), long dashed line (state 5), dot dash line (state 6). Horizontal solid lines show energy differences of 0, 0.2 meV and 1 meV between the energy calculated with the largest number of basis functions and the energies calculated with N_{bs} basis functions. Left frame: harmonic oscillator basis; right frame: plane wave basis.

of the states that are bound in the dot. It is clear that the convergence of the harmonic oscillator basis calculation is the fastest. The accuracy needed for meaningful comparison with experiment is about 1 meV or better. This can be achieved with the harmonic oscillator basis with just 324 basis states ($l_{\max}=4$, $m_{\max}=8$ and $n_{\max}=3$) while the plane wave calculation requires 7,425 basis states (33 in the vertical direction and 15 in each of the lateral directions) for the same accuracy. Since the CPU time needed to diagonalise a $N \times N$ matrix scales like N^3 this leads to an enormous reduction in CPU time.

Another difficulty with the plane wave method is that the results are sensitive to the size of the periodic cell used for the calculation. Care has to be taken to ensure that the tail of the wave function is sufficiently small at the cell boundary and ground state energies can be around 10% too low if this condition is not satisfied. See Ref. [6] for details.

3. Auger Relaxation

3.1 Auger Processes and the Dipole Approximation

The relaxation rate of a dot electron from an excited state to the ground state is a quantity that is very relevant to the operation of opto-

electronic devices [1, 16]. When the electron or hole density outside the dot, either in the capping material or in the wetting layer, is sufficiently high the dominant relaxation process is expected to be an Auger process where energy is transferred from the electron in the dot to the external particle. We consider Auger processes that involve electrons external to the dot, both in the wetting layer and the capping layer.

The rate of Auger transitions is found from the Fermi Golden Rule in the standard way by summing over all the transitions that conserve energy:

$$\frac{2\pi}{\hbar} \sum_{\mathbf{k}_i, \mathbf{k}_f} |\langle \Psi_f | W | \Psi_i \rangle|^2 f(k_i) \delta(E(\mathbf{k}_f) - E(\mathbf{k}_i) + \Delta E_{QD}), \quad (8)$$

where Ψ_i and Ψ_f are the initial and final states. Exchange effects are neglected which allows us to take these states to be products of dot states and external electron states. The external electron states are taken to be either 2D or 3D plane waves. In either case $\mathbf{k}_i$ and $\mathbf{k}_f$ are the initial and final wave-vectors of the external electrons and ΔE_{QD} is the change in energy of the dot electron. $f(k_i)$ is the energy distribution function of the external electrons in their initial state and we take it to be a Maxwell-Boltzmann distribution. We suppose that the dot electron interacts with the external electron via the screened Coulomb interaction,

$$W(\mathbf{r} - \mathbf{r}_e) = \frac{e^2}{4\pi\epsilon\epsilon_0|\mathbf{r} - \mathbf{r}_e|} \exp(-\kappa|\mathbf{r} - \mathbf{r}_e|), \quad (9)$$

where $\mathbf{r}$ and $\mathbf{r}_e$ are the dot and external electron coordinates, ϵ is the dielectric constant, $\kappa = \sqrt{4\pi e^2 n_B / (\epsilon\epsilon_0 k_B T_B)}$ is the Debye screening constant, and n_B and T_B are the bulk electron density and temperature. The screening effect of the wetting layer electrons is neglected [8].

To calculate the relaxation rate from Eq. (8) we need to know the electron densities in the wetting layer and the bulk. We suppose that electrons are injected into the bulk conduction band and relax down to the conduction band edge. After this they can be captured into the wetting layer and then recombine with holes or they can recombine directly from the conduction band. This process is described by the following rate equations

$$\begin{aligned} \frac{dn_p(t)}{dt} &= G(t) - \frac{n_p(t)}{\tau_{\text{rel}}}, \\ \frac{dn_1(t)}{dt} &= \frac{n_p(t)}{\tau_{\text{rel}}} - \frac{n_1(t)}{\tau_{r1}} - \frac{n_1(t)}{\tau_{\text{cap}}}, \\ \frac{dn_{\text{WL}}(t)}{dt} &= \frac{n_1(t)}{\tau_{\text{cap}}} - \frac{n_{\text{WL}}(t)}{\tau_{r2}}, \end{aligned} \quad (10)$$

where the time constants are taken from experimental data of Liu *et al.* [17] ($\tau_{\text{rel}} = 2$ ps, $\tau_{\text{cap}} = 7$ ps, $\tau_{r1} = 1$ ns, $\tau_{r2} = 0.25$ ns) and the injection rate $G(t)$ is chosen to give the desired number of injected electrons.

Auger relaxation rate calculations for electrons in a self-assembled dot were first performed by Uskov *et al.* [8]. They used an axially symmetric dot model so the dot states contain an angular momentum factor, identical to the factor Φ_{l_i} that appears in the basis states defined in Eq. (7). In addition they approximated the interaction potential by Taylor expanding it about $\mathbf{r} = \mathbf{0}$ so

$$W(\mathbf{r} - \mathbf{r}_e) \sim W(-\mathbf{r}_e) + \mathbf{r} \cdot \nabla W(-\mathbf{r}_e) + \dots \quad (11)$$

At first sight this is physically reasonable because the spatial extent of the external electron state is much greater than that of the localised dot electron state, so that typically $r \ll r_e$. The first term in the Taylor expansion does not contribute to the relaxation rate because of the orthogonality of the initial and final states. So if only the second term is retained the matrix element for the transition becomes a dipole matrix element and the approximation is equivalent to the dipole approximation used in atomic physics to find radiative transition rates. The advantage of making this approximation is that the Auger relaxation rate can be found analytically in some cases but we have found a discrepancy between the exact and dipole relaxation rates.

3.2 Breakdown of the Dipole Approximation

We have examined the validity of making the dipole approximation for the lateral coordinates, in other words making the Taylor expansion with respect to x and y only. We use an axially symmetric dot model and take the dot states to be products of lateral and vertical harmonic oscillator states. That is, the states we use for testing the dipole approximation are identical to the basis states defined in Eq. (7). Instead of Taylor expanding the interaction potential in real space we work with its Fourier representation,

$$W(\mathbf{r} - \mathbf{r}_e) = \frac{2\pi}{A} \frac{e^2}{4\pi\epsilon\epsilon_0} \sum_{\mathbf{p}} (\kappa^2 + p^2)^{-1/2} \times \exp\left(-|z - z_e| \sqrt{\kappa^2 + p^2} + i\mathbf{p} \cdot (\mathbf{r} - \mathbf{r}_e)\right), \quad (12)$$

where $\mathbf{p}$ is a two dimensional vector in the x, y plane, the system is taken to be periodic with area A , and z_e is the vertical position of the wetting layer. The exact matrix element for an Auger process in which the momentum change of a wetting layer electron is $\hbar\mathbf{q} = \hbar(\mathbf{k}_f - \mathbf{k}_i)$ is

then

$$\frac{2\pi}{A} \frac{e^2}{4\pi\epsilon\epsilon_0} \frac{1}{(\kappa^2 + q^2)^{1/2}} \times \left\langle \psi_f \left| \exp\left(-|z - z_e|\sqrt{\kappa^2 + q^2}\right) \exp(-i\mathbf{q} \cdot \mathbf{r}) \right| \psi_i \right\rangle, \quad (13)$$

while the dipole approximation is made by expanding the second exponential factor as $\exp(-i\mathbf{q} \cdot \mathbf{r}) \sim 1 - i\mathbf{q} \cdot \mathbf{r} \dots$. Here, ψ_i and ψ_f are initial and final dot states of the form defined by Eq. (7).

In the case when the initial state is the first excited state of lateral motion and the final state is the ground state the exact matrix element is

$$\frac{2\pi}{A} \frac{e^2}{4\pi\epsilon\epsilon_0} \frac{1}{(\kappa^2 + q^2)^{1/2}} \times F(q) \frac{q\lambda_r}{i\sqrt{2}} \exp(-q^2\lambda_r^2/2), \quad (14)$$

where $F(q)$ is the matrix element between vertical states:

$$F(q) = \left\langle Z_f \left| \exp\left(-|z - z_e|\sqrt{\kappa^2 + q^2}\right) \right| Z_i \right\rangle, \quad (15)$$

which do not change in the case we are considering. The corresponding matrix element in the dipole approximation is

$$\frac{2\pi}{A} \frac{e^2}{4\pi\epsilon\epsilon_0} \frac{1}{(\kappa^2 + q^2)^{1/2}} \times F(q) \frac{q\lambda_r}{i\sqrt{2}} (1 - q^2\lambda_r^2/2 + \dots). \quad (16)$$

Strictly speaking, the dipole approximation corresponds to the first term in the square brackets but we have retained terms up to cubic order in $\mathbf{q} \cdot \mathbf{r}$ to allow comparison of the exact and approximate matrix elements. It is clear that the term in square brackets in the approximate matrix element is just the power series expansion of the exponential factor in the exact matrix element. Therefore the approximation is valid when $\lambda_r^2 q^2/2 \ll 1$. When the temperature is low and the energy transfer is high, $k_f \gg k_i$ so $|\mathbf{q}| = |\mathbf{k}_f - \mathbf{k}_i| \sim k_f$ and $\Delta E_{QD} = \hbar\omega \sim \hbar^2 q^2/2m^*$. In addition, λ_r and ω for the harmonic oscillator states are related via $\lambda_r^2 = \hbar/2m^*\omega$ therefore $\lambda_r q \sim 1$. Thus the dipole approximation is never valid no matter how small the dot is. A similar result holds in the case of relaxation of electrons in harmonic oscillator states interacting with bulk electrons. It is also likely that the dipole approximation is not valid for real dot states. The lateral length scale of a typical dot state is of order $l = \sqrt{\langle r^2 \rangle - \langle r \rangle^2}$. We have calculated the product lq for a number of pyramidal dot models and found that it is typically of order 1. This suggests that the dipole approximation is typically not valid but as there is no advantage to using

the approximation when the matrix elements are calculated numerically we have not pursued this issue.

Quantitatively, the discrepancy between the exact and dipole relaxation rates is a constant factor. This is illustrated in Figs. 3 and 4 which show exact and dipole rates for relaxation of an electron in the first lateral excited state of a harmonic oscillator as a function of lateral confinement energy $\hbar\omega$. The rates were computed by numerically evaluating the sum in Eq. (8) as a multidimensional integral to an accuracy of about 0.1%. The vertical length parameter is 2.41 nm, the distance to the wetting layer is 2.5 nm, the material is InAs ($m^* = 0.023$, $\epsilon = 15.15$), the bulk electron temperature is 1 K and the injected electron density is 10^{22} m^{-3} . Rates for screened and unscreened interactions are given. Qualitatively, the dipole rates show the same trends as the exact ones in all cases but quantitatively they differ from the exact rates by a factor of order 2 - 3 that is nearly independent of the confinement energy. This is consistent with the arguments in the previous paragraph. The lack of dependence on confinement energy occurs because the product $\lambda_r q$ is approximately constant. The numerical value of the factor can be understood by considering the matrix element. The exact and dipole matrix elements differ by the factor $\exp(-\lambda_r^2 q^2/2) \sim \exp(-0.5)$. There-

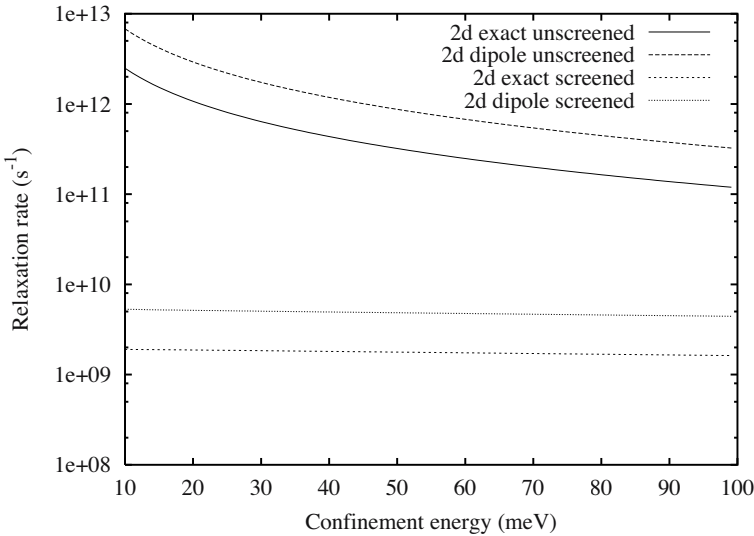


Figure 3. Relaxation rates for a dot electron in the first excited lateral state interacting with wetting layer electrons. Rates are shown for the harmonic oscillator model as a function of lateral confinement energy.

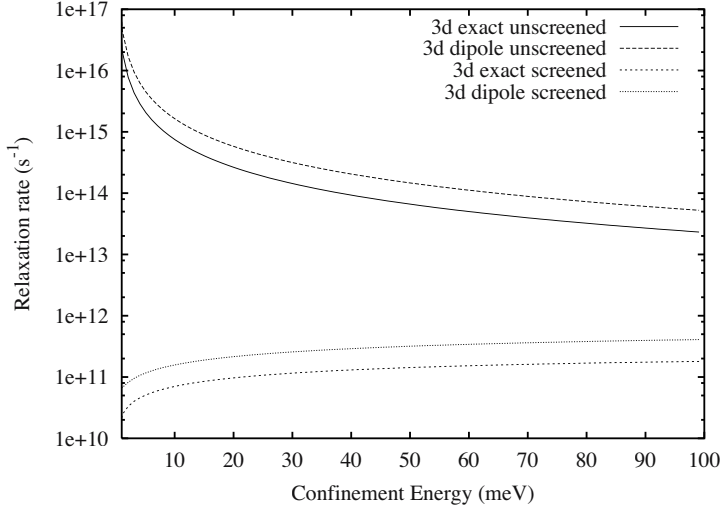


Figure 4. Relaxation rates for a dot electron in the first excited lateral state interacting with bulk electrons. Rates are shown for the harmonic oscillator model as a function of lateral confinement energy.

fore the rates differ by a factor of order e as found numerically. The dependence of the harmonic oscillator model relaxation rates on confinement energy can be understood by considering the q dependence of the matrix elements, as detailed in Ref. [7].

3.3 Effect of Dot Geometry on Relaxation Rate

To examine the effect of dot geometry on the Auger relaxation from the first excited state to the ground state we have computed the relaxation rate as a function of electron density for a number of pyramidal dot models [7]. The models include square and rectangular based truncated pyramids and a square based full pyramid, dimensions are given in Table 2. In each case the dot is taken to be composed of InAs. The small effect of different dielectric constants of the dot and the capping material is ignored. As harmonic oscillator states have been proposed as a model for self-assembled dot states (see Ref. [1] for a review) we have also computed rates for the harmonic oscillator model. The results are shown in Figs. 5 (interaction with wetting layer electrons) and 6 (interaction with bulk electrons). The temperature is 1 K. The injected electron density range is the same for the calculations leading to Figs. 5 and 6 so the figures are directly comparable. All of the curves have a

Table 2. Dimensions of quantum dot models used in this work.

Model geometry	$l \times w \times h$ (nm)
Square based full pyramid	$6.0 \times 6.0 \times 6.0$
Rectangular based truncated pyramid	$14.0 \times 21.0 \times 5.0$
Square based truncated pyramid	$17.7 \times 17.7 \times 5.0$

similar structure: the rate first increases with electron density because the increasing density increases the probability of an interaction and then decreases because of screening.

The length parameters for the harmonic oscillator model have been chosen to be similar to the radial and vertical RMS displacements, $\sqrt{\langle r^2 \rangle - \langle r \rangle^2}$ and $\sqrt{\langle z^2 \rangle - \langle z \rangle^2}$ of the pyramidal models. The vertical length parameter is 2.24 nm, the lateral length parameter is 4.5 nm, and all other parameters are as in the previous section. In the case of relaxation via interactions with wetting layer electrons (Fig. 5) the harmonic oscillator model rate is one to two orders of magnitude *larger* than any of the pyramidal model rates. This is a consequence of the fact that a harmonic oscillator potential does not confine electrons as effectively as the potential in the pyramidal model, which rises abruptly at the dot boundary. Consequently the pyramidal dot states have a smaller overlap than the harmonic oscillator states with the wetting layer. On the other hand, for relaxation via interactions with bulk electrons (Fig. 6) the harmonic oscillator model rate is one to two orders of magnitude *smaller* than any of the pyramidal model rates. This is a consequence of the form of the matrix element for interactions with bulk electrons,

$$\frac{4\pi}{V} \frac{e^2}{4\pi\epsilon\epsilon_0} \frac{1}{(\kappa^2 + q^2)} \langle \psi_f | \exp(-i\mathbf{q} \cdot \mathbf{r}) | \psi_i \rangle, \quad (17)$$

where V is the volume of the system and $\hbar\mathbf{q}$ is a 3D momentum change. As explained in the previous section $q^2 \sim 2m^*\Delta E/\hbar^2$ at low temperature, so the matrix element is of order $1/\Delta E$ in the low density, low screening limit ($\kappa \sim 0$). ΔE for the harmonic oscillator model is about 2 to 3 times larger than for the pyramidal models and this results in the rate being reduced by about an order of magnitude. In addition, there is some further reduction through the value of the matrix element of $\exp(-i\mathbf{q} \cdot \mathbf{r})$. This result illustrates the difficulty of reproducing self-assembled dot physics accurately with a harmonic oscillator model. The length parameter and excitation gap of the harmonic oscillator are not independent so it is not possible to match the length scale and gap of the pyramidal dot states accurately with an oscillator model, however

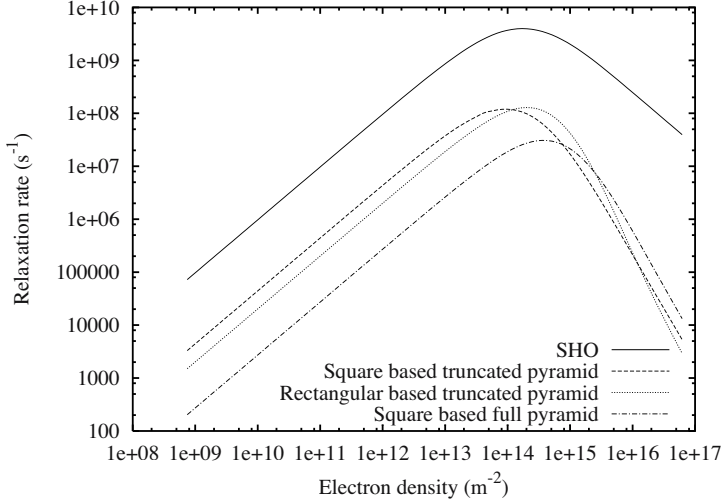


Figure 5. Relaxation rates as a function of wetting layer electron density for a dot electron in the first excited lateral state interacting with wetting layer electrons.

the oscillator model does reproduce the Auger relaxation physics qualitatively, as illustrated in Figs. 5 and 6.

Figures 5 and 6 also show that there is considerable variability in the rates for the various pyramid models, particularly at low density. The rate for relaxation via bulk interactions is always the highest but the ratio of the bulk and wetting layer relaxation rates is very sensitive to the time constants in Eq. (10), which are probably system dependent. Composition variations are likely to decrease the relaxation rate for wetting layer interactions by a small factor because a composition gradient with the largest In fraction at the top of the dot causes a small shift of the electron wave function away from the wetting layer. Some of the relaxation rates predicted here are comparable to those measured experimentally by Morris *et al.* [16], though it is difficult to make a quantitative comparison with these results because the electron densities corresponding to the experimental conditions of Ref. [16] are not known. It would be very interesting to have experimental data for a system in which both the relaxation rate and the electron densities have been measured.

4. Conclusion

We have reviewed an efficient method for the computation of quantum states in self-assembled dots and its application to Auger relaxation.

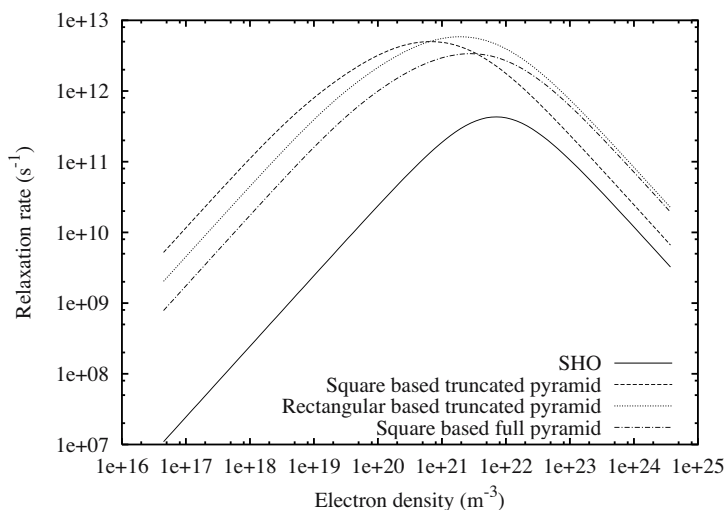


Figure 6. Relaxation rates as a function of bulk electron density for a dot electron in the first excited lateral state interacting with bulk electrons.

Our method enables dot states to be computed economically within the effective mass approximation and seems to be useful for computing physical properties of dots. Our results on Auger relaxation indicate that the dipole approximation does not give an accurate description of the relaxation rate. We have also indicated the difficulties of using harmonic oscillator model states for calculating Auger rates. The difficulties of using the harmonic oscillator model to describe self-assembled dots containing more than a few electrons have also been discussed by Petroff *et al.* [1] and in both cases the difficulties are caused by the very simple excitation spectrum of the harmonic oscillator model. Our theory predicts relaxation rates can be around 10 ps, as seen experimentally, but the rates depend critically on the bulk and wetting layer electron densities and it would be very useful to have measurements of Auger relaxation rates in a system for which these electron densities are known.

Acknowledgments

The numerical calculations described here were performed on a supercomputer at the Leicester Mathematical Modelling Centre which was purchased through the EPSRC strategic equipment initiative. DC is grateful for the award of an EPSRC research studentship.

References

- [1] P. M. Petroff, A. Lorke, and A. Imamoglu. Epitaxially self-assembled quantum dots. *Physics Today* **54**(5): 46–52, 2001.
- [2] M. Grundmann, O. Stier, and D. Bimberg. InAs/GaAs pyramidal quantum dots: Strain distribution, optical phonons, and electronic-structure,” *Phys. Rev. B* **52**: 11969–11981, 1995.
- [3] M. A. Cusack, P. R. Briddon, and M. Jaros. Electronic structure of InAs/GaAs self-assembled quantum dots. *Phys. Rev. B*. **54**: R2300–R2303, 1996.
- [4] J. A. Barker and E. P. O’Reilly. The influence of inter-diffusion on electron states in quantum dots. *Physica E* **4**: 231–237, 1999.
- [5] J. A. Barker and E. P. O’Reilly. Theoretical analysis of electron-hole alignment in InAs-GaAs quantum dots. *Phys. Rev. B* **61**: 13840–13851, 2000.
- [6] M. Roy and P. A. Maksym (2003). Efficient method for calculating electronic states in self-assembled quantum dots. *Phys. Rev. B* **68**: art. no. 235308 (2003)
- [7] D. Chaney, M. Roy, P. A. Maksym, and F. Long. The effect of self-assembled quantum dot geometry on Auger relaxation rate. *Proc. 26th Int. Conf. on the Physics of Semiconductors*, edited by A. R. Long and J. H. Davies (Bristol, IOP Publishing, 2003).
- [8] A. V. Uskov, F. Adler, H. Schweizer, and M. H. Pikuhn. Auger carrier relaxation in self-assembled quantum dots by collisions with two-dimensional carriers. *J. Appl. Phys.* **81**: 7895–7899, 1997.
- [9] P. B. Joyce, T. J. Krzyzewski, G. R. Bell, B. A. Joyce and T. S. Jones. Composition of InAs quantum dots on GaAs(001): Direct evidence for (In,Ga)As alloying,” *Phys. Rev. B* **58**: R15981–R15984, 1998.
- [10] N. Liu, J. Tersoff, O. Baklenov, A. L. Holmes, and C. K. Shih. Nonuniform composition profile in $\text{In}_{0.5}\text{Ga}_{0.5}\text{As}$ alloy quantum dots. *Phys. Rev. Lett.* **84**: 334–337, 2000.
- [11] P. W. Fry, I. E. Itskevich , D. J. Mowbray, M. S. Skolnick, J. J. Finley, J. A. Barker, E. P. O’Reilly, L. R. Wilson, I. A. Larkin, P. A. Maksym, M. Hopkinson, M. Al-Khafaji, J. P. R. David, A. G. Cullis, G. Hill, and J. C. Clark. Inverted electron-hole alignment in InAs-GaAs self-assembled quantum dots. *Phys. Rev. Lett.* **84**: 733–736, 2000.
- [12] D. M. Bruls, J. W. A. M. Vugs, P. M. Koenraad, M. S. Skolnick, M. Hopkinson, F. Long, S. P. A. Gill, and J. H. Wolter. Determination of the shape and indium distribution of low-growth-rate InAs quantum dots by cross-sectional scanning tunnelling microscopy. *Appl. Phys. Lett.* **81**: 1708–1710, 2002.
- [13] M. R. Bruni, A. Lapicciarella, G. Scavia, M. G. Simeone, S. Viticoli, and N. Tomassini. Thermodynamic study of molecular-beam epitaxial-growth of In-GaAs/GaAs strained layer superlattices. *Thermochemica Acta*. **210**, 49–65, 1992.
- [14] L. R. C. Fonseca, J. L. Jimenez, J. P. Leburton, and R. M. Martin. Self-consistent calculation of the electronic structure and electron-electron interaction in self-assembled InAs-GaAs quantum dot structures,” *Phys. Rev. B* **57**: 4017–4026, 1998.

- [15] O. Stier, M. Grundmann, and D. Bimberg. Electronic and optical properties of strained quantum dots modelled by 8-band $\mathbf{k} \cdot \mathbf{p}$ theory. *Phys. Rev. B* **59**: 5688–5701, 1999.
- [16] D. Morris, N. Perret and S. Fafard. Carrier energy relaxation by means of Auger processes in InAs/InGaAs self-assembled quantum dots. *Appl. Phys. Lett.* **75**: 3593–3595, 1999.
- [17] B. Liu, Q. Li, Z. Xu and W. E. Ge. Detection of efficient carrier capture in ultrathin InAs/GaAs layers using a degenerate pump-probe technique. *J. Phys: Condens. Matter* **13**: 3923–3930, 2001.

QUANTUM DOT MOLECULES AND CHAINS

W. Jaskólski, M. Zieliński, A. Stróżecka

Instytut Fizyki UMK, Grudziądzka 5, 87-100 Toruń, Poland

Garnett W. Bryant and J. Aizpurua

National Institute of Standards and Technology, Gaithersburg MD, USA

Abstract A review of results from theoretical investigations of several systems composed of two or more coupled quantum dots (known as *artificial molecules* or *quantum dot solids*) is presented. All the calculations are performed within an empirical tight-binding theory. It is shown that coupling between nanocrystals can split and reorder energy levels and change state symmetries. The results help to understand and explain differences observed in optical spectra of arrays of quantum dots in comparison to the spectra obtained for non-interacting nanocrystals. We show also how an external electric field influences the properties of coupled quantum dots.

1. Introduction

By analogy to diatomic molecules and crystalline solids, *artificial* and *quantum dot solids* can be built from coupled semiconductor nanocrystals. Coupled dots created by etching techniques or electrical confinement from quantum-well systems were studied first due to their possible applications as single-electron resonant tunneling devices [1, 2, 3, 4, 5, 6, 7, 8]. Systems of vertically stacked self-assembled quantum dots are intensively investigated now to determine how the coupling between nanocrystals in a dense array of such dots influences properties and performance of quantum dot lasers [9, 10, 11, 12]. Linear chains of electrically confined quantum dots have been proposed recently for a realization of multi-qubit gates [13].

The densest ensembles of quantum dots are obtained for chemically synthesized nanocrystals [14, 15, 16, 17, 18]. Kagan et al [14] first found luminescence from close-packed CdSe nanocrystals totally different than

the luminescence of non-interacting dots in a dilute solution. Several other experiments [16, 17] have shown that optical properties of dense arrays of chemically synthesized nanocrystals are significantly different from properties of individual dots.

We present here a review of the results of our investigations on the formation of delocalized states in coupled quantum dots and linear chains of nanocrystals. We work within the empirical tight-binding theory which, as a microscopic approach, is well suited for precise investigation of the coupling on the atomic scale. The presented results help to explain why the optical properties of coupled nanocrystals are so different from the properties of individual quantum dots. We present also some preliminary results showing how an external electric field influences delocalized states in coupled systems.

2. Theory

In the empirical tight-binding approach (TB) the one-particle wavefunction is represented in an orthogonal basis set of atomic orbitals $\phi_\alpha(\mathbf{r} - \mathbf{R}_J)$ [19],

$$\Psi(\mathbf{r}) = \sum_J \sum_\alpha c_{\alpha,J} \phi_\alpha(\mathbf{r} - \mathbf{R}_J), \quad (1)$$

where α denotes an orbital of a given symmetry and $\mathbf{R}_J$ is an atomic site. In our model, each atom is described by 5 orbitals (s, p_x, p_y, p_z , and s^*). We assume that atoms occupy the sites of a zinc-blend structure. Only atoms embedded in a volume of a given system are taken into account. The interaction in the Hamiltonian is restricted to on-site and nearest neighbors only. The Hamiltonian matrix elements

$$t_{\alpha'\alpha R'_J R_J} = \langle \phi_{\alpha'}(\mathbf{r} - \mathbf{R}'_J) | H | \phi_\alpha(\mathbf{r} - \mathbf{R}_J) \rangle \quad (2)$$

are treated as empirical parameters obtained by fitting bulk-band structure to experimentally known band gaps and effective masses. There are 13 different TB parameters for each material. The surface dangling bonds are passivated by shifting their energies high above the conduction band edges. The one particle quantum dot states are found by diagonalizing the TB Hamiltonian matrix with the use of an iterative eigenvalue solver.

To study linear infinite chains of nanocrystals we first define the supercell of a given chain. The TB wavefunction components corresponding to atomic sites at opposite supercell boundaries differ in phase by e^{iqD} , where D is the chain period. Therefore, the TB elements for $\mathbf{R}_J$ and $\mathbf{R}'_J$ at adjacent supercells are $t_{\alpha'\alpha R'_J R_J} = e^{iqD} t_{\alpha'\alpha R'_J R_{J+D}}$. A static external electric field is described by adding $e\vec{F}\vec{r}$, where F is the electric

field, to the one particle Hamiltonian. This is accomplished by including dipole matrix elements between atomic orbitals on the same and neighboring atoms, in the TB Hamiltonian matrix. The off-site dipole matrix elements are naturally much smaller than the on-site ones and are neglected. The on-site dipole moments are

$$\langle \alpha J | \mathbf{r} | \beta J \rangle = \delta_{\alpha\beta} \mathbf{R}_J + \langle \alpha J | \mathbf{r} - \mathbf{R}_J | \beta J \rangle. \quad (3)$$

The dipole matrix elements between different orbitals on the same site (the second term in Eq. 3) are taken from [21].

3. Artificial Molecule of Two CdS Nanocrystals

We first consider a quantum dot molecule built of two spherical CdS nanocrystals of radii 2.9 nm and 1.9 nm, respectively. The sizes of the nanocrystals are similar to CdSe nanocrystals studied in [14] (3.1 nm and 1.9 nm), for which very different luminescence was observed depending on whether the sample was liquid mixture of such dots or a close-packed quantum dot solid. The lowest conduction band energy level (Ls) of the larger nanocrystal has energy 2.6548 eV above the valence band-edge of the bulk CdS [20]. The TB wavefunction of this state reveals a global s -type symmetry. The first excited state energy level (Lp), with energy 2.8010 eV, is triply degenerate: the corresponding wavefunctions have global p -type symmetry. The lowest conduction band energy level (Ss) of the smaller dot is also of s -type and has energy 2.7981 eV, i.e. close to the first excited level of the larger nanocrystal.

When the two nanocrystals are close but not directly connected, the energy spectrum of such a system is a simple sum of the spectra of individual nanocrystals. This is because the surrounding medium is not represented in the TB wavefunction, and no interaction is mediated between the quantum dots. The case of nanocrystals that just touch each other, is modeled by putting the dot centers at distance D that is only 0.1 nm smaller than the sum of the dots radii. In such a case the nanocrystals have only one common atom and 16 new chemical bonds created between the nearest atoms¹. Although the spherical symmetry is broken, the lowest electron state is still localized entirely in the larger nanocrystal and has s -type symmetry. However, the degeneracy of the p -type excited state (Lp) is lifted and four new states of quantum dot molecule are formed from three (Lp) states of the larger dot and the (Ss) state of the smaller nanocrystal. Their energies are: 2.7913, 2.8009, 2.8010, 2.8059 eV (see Fig. 2). Figure 1 shows charge density isosurfaces of these states. The lowest state of the quartet has clearly bonding-like character, while the highest one is antibonding-like. The two almost degenerate levels (the second and the third) have p -type symmetry in the

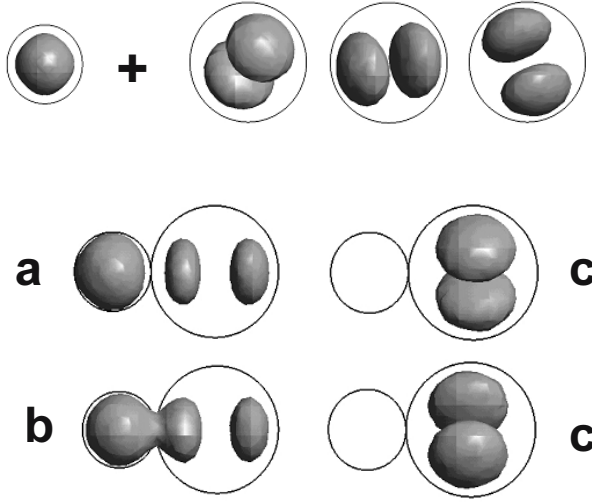


Figure 1. Quantum dot molecule built of two CdS spherical nanocrystals of radii 1.9 and 2.9 nm; dot centers separation distance $D = 4.7$ nm. Density isosurfaces (50%) of the four states (bottom) of double dot created by s -type state of smaller dot and three p -type states of larger nanocrystal (top). (b) bonding-like state, (a) antibonding-like state, (c) p -like states.

plane perpendicular to the artificial molecule axis and are entirely localized in the larger dot. The appearance of delocalized states in systems of nanocrystals that merely touch each other explains how the excitation transfer can occur in close-packed solids of chemically synthesized nanocrystals [14].

When the two nanocrystals are closer connected the energy of the ground state decreases only slightly, the energies of the quasi-degenerate doublet do not change, while the bonding- and antibonding-like states dramatically change their energies. This is shown in Fig. 2. For $D = 3.6$ nm, the level splitting of the *artificial molecule* is about 50 meV.

How an external electric field influences the energy spectra and charge densities of coupled dots is important because electric field can be used to control electron transfer between coupled dots and thus to control elementary qubit operations in such systems [22, 23, 24]. On the other hand, experimental mapping of the electron wavefunctions in quantum dots has been shown possible by using scanning tunneling microscopy (STM) [25]. In such experiments, quantum dot systems are also subjected to strong electric fields. Here, we study Stark effects in the above-mentioned double CdS quantum dot. We assume that a homogeneous external electric field, F is applied in the direction of the axis of the *artificial molecule*. The problem of finding the field inside a nanocrystal

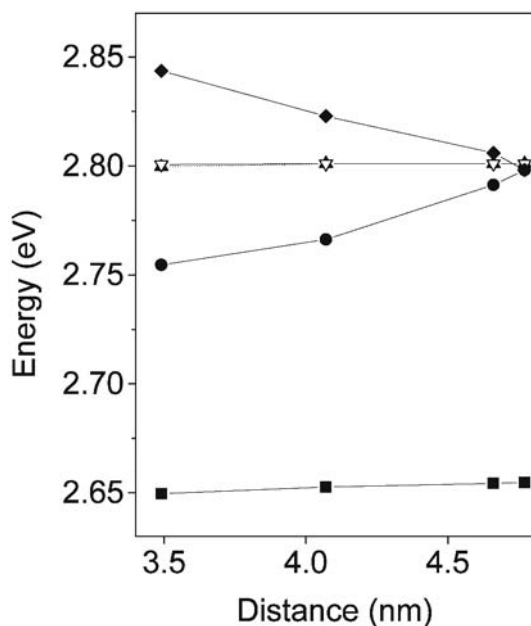


Figure 2. Four lowest electron energy levels of a double CdS quantum dot (radii 1.9 and 2.9 nm) versus distance between dot centers. The rightmost values correspond to the case when dots have only a single common atom.

is not trivial. The field can be found analytically for a limited number of cases, such as the spherical isotropic dielectric dot[26]. As for the spherical nanocrystal, we assume that the electric field inside a double dot is homogeneous. In addition, we assume that the same relation between internal and external fields, $F_{\text{ins}} = 3F/(2 + \epsilon)$ where ϵ is the relative dielectric constant, holds for double and single dots. In the STM experiment [27] a bias up to 3 V was applied between the tip and the substrate. Assuming the tip-substrate distance ~ 10 nm, we study fields up to 3×10^6 V/cm.

Several lowest electron energy levels of the double CdS quantum dot versus electric field are shown in Fig. 3. Since the double dot system is not symmetric, the energy levels split and behave differently for different signs of the field. A strong splitting of the bonding- and antibonding-like states is observed for even moderate fields. For a field of 2.5×10^5 V/cm (0.25 V applied bias) these states completely lose their delocalized character (point marked as z in Fig. 3): the lowest one transforms into the s -type state, localized almost entirely in the larger dot, while the upper one becomes an s -type state located in the smaller dot. The localization occurs because the energy splitting of this bonding-like state is only

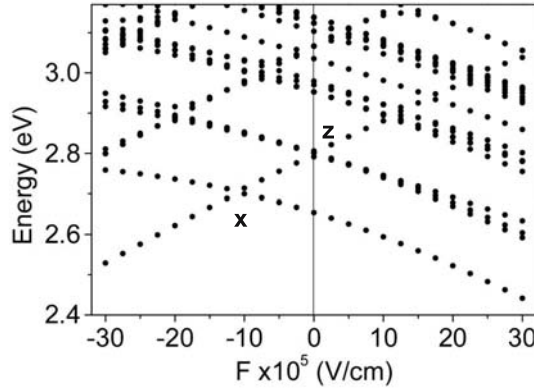


Figure 3. Electron energy levels of a double CdS nanocrystal *versus* electric field applied along the axis of the quantum dot molecule.

about 7 meV (see Fig. 2), while the energy shift by such a field is about 70 meV. Figures 3 and 4 show also that stronger field can lead to formation of another field tunable state (point marked as *x* in Fig. 3): for $F = -10 \times 10^5$ V/cm a delocalized state built of *s*-type states of both nanocrystals is created.

4. Linear Chain of ZnS/CdS Nanocrystals

In this section we study the formation of bands for infinite linear chains of two-layer nanocrystals. We study nanocrystals built of an internal ZnS core of radius 5 nm and an external CdS clad of thickness 1 nm. Such nanocrystals, called also *quantum dot-quantum wells* are synthesized by the wet-chemistry methods [28, 29]. Since the energy

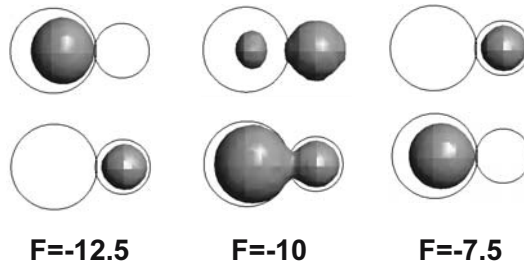


Figure 4. Quantum dot molecule built of two CdS spherical nanocrystals of radii 1.9 and 2.9 nm; dot centers separation distance $D = 4.7$ nm. Density isosurfaces of two lowest electron states *versus* electric field F (corresponds to point marked as *x* in Fig. 3).

gap of the bulk ZnS is 3.7 eV and the gap of CdS is 2.5 eV, the ZnS core act as a barrier and the CdS clad plays a role of a thin spherical external well. The lowest electron and hole states have densities localized mainly in the CdS shell, i.e. pushed away from the dot center. The coupling in such nanocrystals is then stronger than in the case of uniform dots.

In Fig. 5 the minibands, originating from several lowest electron energy levels of individual ZnS/CdS nanocrystals, are shown for three different values of the superlattice period $D = 12$ nm, 11 nm and 10 nm (note that $D = 12$ nm corresponds to a chain formed by the nanocrystals that just touch each other). For $D = 12$ nm the minibands are very narrow; their energies are almost the same as the energy levels of individual dots. When D decreases, excited state minibands become wider. However, the most striking effect is that the ground miniband detaches down from the spectrum and becomes extremely narrow. This can be understood by looking at the charge density of the ground state for a system built of two such nanocrystals. This is shown in Fig. 6. When the distance between dot centers equals $2R$, where R is the dot radius, the ground state is an extended state, delocalized in all nanocrystals. When the distance between dot centers decreases, the ground state (and thus all states in the ground miniband) becomes more localized in the region where the dots overlap because the CdS well in this region be-

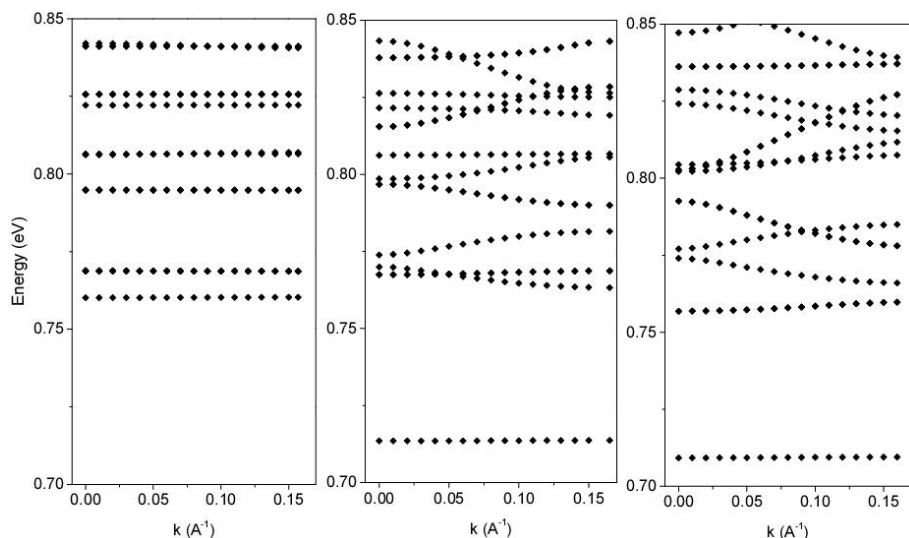


Figure 5. Minibands $E(k)$ for a linear chain of ZnS/CdS nanocrystals; ZnS core radius 5 nm, total dot radius 6 nm. Superlattice period (from left to right): 12 nm, 11 nm and 10 nm.

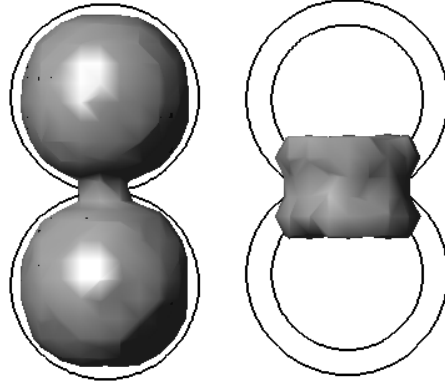


Figure 6. Density isosurfaces of the lowest electron state for a ZnS/CdS quantum dot molecule (as in Fig. 5). Top: nanocrystals just touch each other, bottom: nanocrystals overlap on in the range of CdS clad thickness.

comes wider in the (x, y) plane and the energy of this state decreases. When the overlap d reaches the thickness of the CdS clad (i.e. 1 nm), the ground state is no longer an extended state, it becomes highly localized in the overlap region; the miniband converts into strongly degenerate energy level [30].

The hole states behave in an analogous way. As a result the effective gap decreases significantly and the absorption edge is strongly redshifted [20], in agreement with experimental observations [16, 17].

5. Two Vertically Stacked Self-Organized InAs/GaAs Quantum Dots

For the chemically synthesized nanocrystals studied in the previous sections, the coupling and interaction between quantum dots could be described in the TB method only when the nanocrystals at least touched each other. However, this is not a significant limitation to modeling the coupling between nanocrystals made by wet synthesis. These nanocrystals are usually surrounded by a large-gap dielectric medium and the interaction of separated dots is negligibly weak in such a medium. In contrast, the coupling between self-organized quantum dots is through the surrounding barrier material. In this case the coupling between spatially separated dots can be described within the TB approach because the coupling is mediated by the barrier atoms.

We have studied two vertically stacked lens-shape InAs quantum dots of base size 6 nm and height 1.2 nm on a 2 monolayer (ML) thick wetting layers (WL), embedded in a large box of GaAs [31]. In such quantum dots the strain effects can be significant and can influence coupling effects between the dots [32]. Here we neglect the strain effects ², since we want

to show how the coupling between two vertically stacked quantum dots influences their energy spectra.

For very large separation between the dots, the energy spectrum of the double system is the same as the spectrum of an individual dot, with all the energy levels doubly degenerate. When the separation distance decreases, the degeneracy is lifted and splitting caused by the coupling between dots appears. This coupling is strong enough that even for distance $d_c = 18$ ML separating the dot centers, the splitting of the electron energy levels is about 10 meV. When $d_c = 8$ ML (i.e. there are still 6 ML of GaAs separating the two InAs wetting layers) the splitting of the electron energy levels reaches ~ 200 meV. As a consequence, the split ground energy level starts to cross with the excited level, which also splits, and the symmetry of the consecutive states changes. This is shown in Fig. 7, where isosurfaces of densities of the two lowest electron states are shown for $d_c = 10$ and 6 ML. The hole levels start to cross for larger d_c because the hole energy spectrum is denser. Isosurfaces of densities of the four lowest hole states for $d_c = 12$ and 6 ML are shown in Fig. 8. Note that for $d_c = 6$ ML the top of the lower dot touches the bottom of the upper wetting layer, while for $d_c = 12$ ML the distance between these two points is only 1.8 nm. However, the consecutive hole states have completely different symmetries and polarizations for these two separation distances. The figure shows also that when the self-organized quantum dots are close enough in a vertical stack, well delocalized states are formed. Since changes of symmetries of the electron

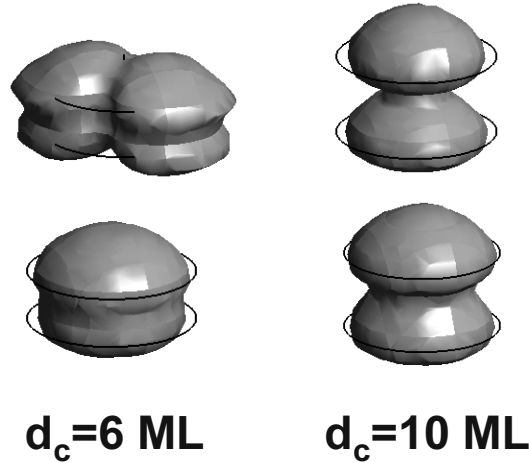


Figure 7. Density isosurfaces (50%) of the ground (bottom) and first excited (top) electron states for two vertically stacked self-organized quantum dots *versus* distance d_c between dot centers.

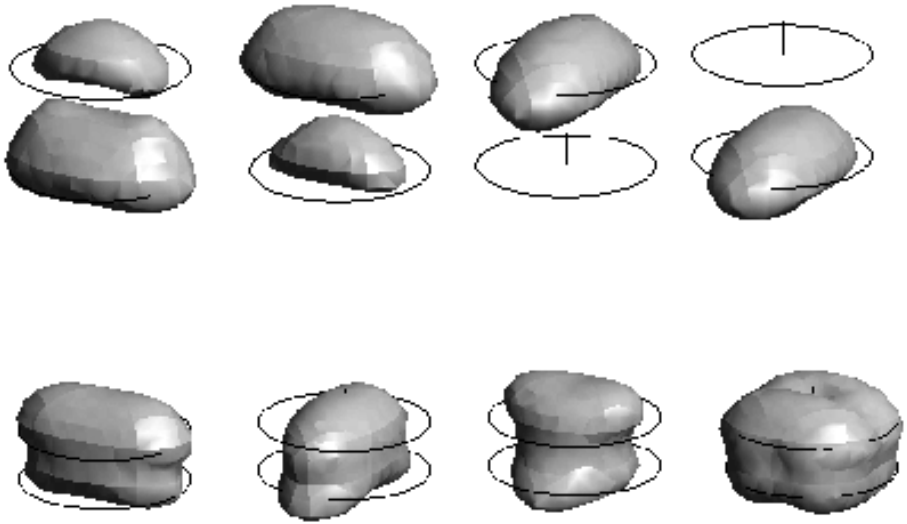


Figure 8. Density isosurfaces (50%) of several lowest hole states (from left to right) for two vertically stacked self-organized lens-type quantum dots for two distances between dot centers: 12 ML (top) and 6 ML (bottom). Circles and vertical bars mark lens base position and height, respectively.

and hole states occur for different d_c , the optical properties of quantum dot lasers can strongly depend on the separation distance between the layers of self-organized quantum dots.

6. Conclusions

The results of theoretical investigations of several systems composed of two or more coupled quantum dots have been reviewed. We have investigated both chemically synthesized nanocrystals and self-organized quantum dots. By studying charge densities of the states of coupled nanocrystals we have shown the formation of delocalized states in such systems. Our results provide an understanding of differences observed between optical spectra of close-packed nanocrystals and spectra of non-interacting quantum dots. For vertically stacked self-organized quantum dots, the order of energy levels and symmetry of the states depends strongly on the distance between the neighboring dots. A weak electric field can destroy delocalization of states in quantum dot molecules and solids.

Notes

1. The total number of atoms in a double quantum dot is 5333.
2. they can be investigated using the *valence force field method* [33].

References

- [1] G. W. Bryant. Electrons in coupled vertical quantum dots: Interdot tunneling and Coulomb correlation. *Phys. Rev. B* **48**: 8024–8034, 1993.
- [2] R. H. Blick, R. J. Haug, J. Weis, D. Pfannkuche, K. v. Klitzing, and K. Eberl. Single-electron tunneling through a double quantum dot: The artificial molecule. *Phys. Rev. B* **53**: 7899–7902, 1996.
- [3] G. W. Bryant. Electronic band structure of semiconductor nanostructure arrays. *Phys. Rev. B* **40**: 1620–1629, 1989.
- [4] G. W. Bryant. Resonant tunneling through coupled, double-quantum-box nanostructures. *Phys. Rev. B* **44**: 3064–3069, 1991.
- [5] R. Aguado and L. P. Kouwenhoven. Double Quantum dots as detectors of high-frequency quantum noise in mesoscopic conductors. *Phys. Rev. Lett.* **84**: 1986–1989, 2000.
- [6] Y. Tokura, D. G. Austing, and S. Tarucha. Single-electron tunnelling in two vertically coupled quantum dots. *J. Phys.: Condens. Matter* **11**: 6023–6034, 1999.
- [7] R. Zigler, C. Bruder, and H. Schoeller. Transport through double quantum dots. *Phys. Rev. B* **62**: 1961–1970, 2000.
- [8] H. Imamura, P. A. Maksym, and H. Aoki. Vertically coupled double quantum dots in magnetic fields. *Phys. Rev. B* **59**: 5817–5825, 1999.
- [9] G. S. Solomon, J. A. Trezza, A. F. Marshall, and J. S. Harris, Jr. Vertically aligned and electronically coupled growth induced InAs islands in GaAs. *Phys. Rev. Lett.* **76**: 952–955, (1996).
- [10] P. Yu, W. Langbein, K. Leosson, J. M. Hvam, N. N. Ledentsov, D. Bimberg, V. M. Ustinov, A. Yu. Egorov, A. E. Zhukov, A. F. Tsatsul'nikov, and Yu. G. Musikhin. Optical anisotropy in vertically coupled quantum dots. *Phys. Rev. B* **60**: 16680–16685 (1999).
- [11] B. Grandidier, Y. M. Niquet, B. Legrand, J. P. Nys, C. Priester, D. Stiévenard, J. M. Gérard, and V. Thierry-Mieg. Imaging the wave-function amplitudes in cleaved semiconductor quantum boxes. *Phys. Rev. Lett.* **85**: 1068–1071, (2000).
- [12] M. Bayer, P. Hawrylak, K. Hinzer, S. Fafard, M. Korkusinski, Z. R. Wasilewski, O. Stern, and A. Forchel. Coupling and entangling of quantum states in quantum dot molecules. *Science* **291**: 451–453, 2001.
- [13] G. Burkard and D. Loss. Spin qubits in solid-state structures. *Europhysics News* **33**(5): 166, 2002.
- [14] C. R. Kagan, C. B. Murray, M. Nirmal, and M. G. Bawendi. Electronic energy transfer in CdSe quantum dot solids. *Phys. Rev. Lett.* **76**: 1517–1520, 1996.
- [15] A. P. Alivisatos. Semiconductor clusters, nanocrystals, and quantum dots. *Science* **271**: 933–937, 1996.

- [16] M. V. Artemyev, A. I. Bibik, L. I. Gurinovich, S. V. Gaponenko, and U. Woggon. Evolution from individual to collective electron states in a dense quantum dot ensemble. *Phys. Rev. B* **60**: 1504–1506, 1999.
- [17] H. Döllefeld, H. Weller, and A. Eychmüller. Particle-particle interactions in semiconductor nanocrystal assemblies. *Nano Lett.* **1**: 267–269, 2001.
- [18] D. Yu, C. J. Wang, and P. Guyot-Sionnest. *n*-type conducting CdSe nanocrystal solids. *Science* **300**: 1277–1280, 2003.
- [19] G. W. Bryant and W. Jaskólski. Tight-binding theory of quantum-dot quantum wells: Single-particle effects and near-band-edge structure. *Phys. Rev. B* **67**, 205320, 2003.
- [20] W. Jaskolski, G. W. Bryant, J. Planelles, and M. Zieliński. Artificial molecules. *Int. J. Quantum Chem.* **90**: 1075–1082, 2002.
- [21] S. Fraga and J. Muszyńska. *Atoms in external fields* (Elsevier, New York 1981).
- [22] I. Shtrichman, C. Metzner, B. D. Gerardot, W. V. Schoenfeld, and P. M. Petroff. Photoluminescence of a single InAs quantum dot molecule under applied electric field. *Phys. Rev. B* **65**, 081303, 2002.
- [23] G. Burkard, G. Seelig, and D. Loss. Spin interactions and switching in vertically tunnel-coupled quantum dots. *Phys. Rev. B* **62**: 2581–2592, 2000.
- [24] P. Zhang and X.-G. Zhao. Localization and entanglement of two interacting electrons in a double quantum dot. *J. Phys.: Condens. Matter* **13**: 8389–8403, 2001.
- [25] U. Banin, Y. W. Cao, D. Katz, and O. Millo. Identification of atomic-like electronic states in indium arsenide nanocrystal quantum dots. *Nature* **400**: 542–544, 1999.
- [26] E. M. Purcell. *Electricity and Magnetism* (McGraw Hill, New York, 1965).
- [27] O. Millo, D. Katz, Y. Cao, and U. Banin. Scanning tunneling spectroscopy of InAs nanocrystal quantum dots. *Phys. Rev. B* **61**, 16773–16777, 2000.
- [28] A. Mews, A. Eychmüller, M. Giersig, D. Schooss, and H. Weller. Preparation, characterization, and photophysics of the quantum-dot quantum-well system Cds/HgS/CdS. *J. Phys. Chem.* **98**: 934–941, 1994.
- [29] R. B. Little, M. A. El-Sayed, G. W. Bryant, and S. Burke. Formation of quantum-dot quantum-well heteronanostructures with large lattice mismatch: ZnS/CdS/ZnS. *J. Chem. Phys.* **114**: 1813–1822, 2001.
- [30] J. G. Diaz, W. Jaskólski, J. Planelles, J. Aizpurua, and G. W. Bryant. Nanocrystal molecules and chains. *J. Chem. Phys.* **119**: 7484–7490, 2003.
- [31] W. Jaskólski, M. Zieliński, and G. W. Bryant, submitted
- [32] W. Sheng and J.-P. Leburton. Anomalous quantum-confined Stark effects in stacked InAs/GaAs self-assembled quantum dots. *Phys. Rev. Lett.* **88**: 167401, 2002.
- [33] T. Saito and Y. Arakawa. Electronic structure of piezoelectric In_{0.2}Ga_{0.8}N quantum dots in GaN calculated using a tight-binding method. *Physica E* **15**: 169–181, 2002.

COLLECTIVE PROPERTIES OF ELECTRONS AND HOLES IN COUPLED QUANTUM DOTS

Guido Goldoni, Filippo Troiani, Massimo Rontani, Devis Bellucci,
and Elisa Molinari

*INFN National Research Center on nanoStructures and bioSystems at Surfaces (S3)
Dipartimento di Fisica, Università degli Studi di Modena e Reggio Emilia,
Via Campi 213/A, 41100 Modena, Italy*

Ulrich Hohenester

*Institute für Theoretische Physik, Karl-Franzens-Universität Graz,
Universitätsplatz 5, 8010 Graz, Austria*

Abstract We discuss the properties of few electrons and electron-hole pairs confined in coupled semiconductor quantum dots, with emphasis on correlation effects and the role of tunneling. We discuss, in particular, exact diagonalization results for biexciton binding energy, electron-hole localization, magnetic-field induced Wigner molecules, and spin ordering.

1. Introduction

The similarity between quantum dots (QDs) and natural atoms, originating in the discrete density of states, is often pointed out [1, 2, 3, 4, 5]. Shell structure [3, 4], correlation effects [6], and Kondo physics [7, 8] are among the most striking demonstrations. Aside from these similarities, and in addition to the huge technological interest in quantum systems that can be grown with precise control, there are two main differences between natural and artificial atoms that make QDs especially interesting from a fundamental point of view. First, while in natural atoms Coulomb interactions are typically of the order of the lowest energy single-particle gaps, in QDs these interactions can be made larger than the latter, due to the different scaling of the kinetic and Coulomb energies with confinement length. Secondly, magnetic fields available in laboratories are

associated with energy scales that are comparable with single-particle gaps. Therefore, they can be used to manipulate the quantum states and the ratio between kinetic and Coulomb contributions, so that in QDs one can reach regimes that are unattainable in natural atoms.

Coupled quantum dots, also called artificial molecules (AMs), extend to the molecular realm the similarity between quantum dots and artificial atoms [9, 10, 11, 12, 13, 14]. Inter-dot tunneling introduces a kinetic contribution that must be added to the QD confinement energy, which can be tuned by structural engineering and transverse magnetic fields. In addition, inter-dot and intra-dot Coulomb interactions couple different electronic configurations (Slater determinants), the ratio between different contributions being a function, e.g. of the external magnetic field strength and direction. Therefore, AMs constitute an interesting laboratory to study the physics of correlation. A specific aspect of AMs is that carrier interactions are tied to localization. Indeed, while kinetic energy favors delocalization over the diatomic AM, Coulomb interaction favors localization in opposite dots for charges of the same sign; the ratio between the two contributions and, therefore, the localized or delocalized character of the correlated state, is controlled by the tunneling energy.

From the theoretical point of view, highly correlated quantum systems are obviously a challenge. In many instances one is not allowed to use mean-field methods [15, 16]. Reliable results can be obtained by a configuration interaction approach, where the calculated single-particle states are used to form a large basis of Slater determinants to represent the Coulomb interaction. This method has limitations in the number of free-carriers which can be treated. On the other hand, it provides accurate results which can be quantitatively compared with experiments and which represent a benchmark for more approximate methods.

In this work we review some theoretical results obtained for few electrons and electron-hole pairs in AMs. We specifically focus on the effects of carrier interaction, and the interplay of tunneling with other energy scales. All calculations are obtained by exact diagonalization methods of the Coulomb interaction. We consider N_e electrons and N_h holes with the effective-mass Hamiltonian $H = H_e + H_h + H_{eh}$, where

$$H_\alpha = \sum_{i=1}^{N_\alpha} \left\{ -\frac{\hbar^2}{2m_\alpha^*} \left[\nabla_i + \frac{e}{c} \mathbf{A}(\mathbf{r}_i^\alpha) \right]^2 + V_\alpha(\mathbf{r}_i^\alpha) \right\} \\ + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N_\alpha} \frac{e^2}{\epsilon^* |\mathbf{r}_i^\alpha - \mathbf{r}_j^\alpha|^2} + g^* \mu_B \boldsymbol{\sigma} \cdot \mathbf{B},$$

$$H_{eh} = - \sum_{i=1, j=1}^{N_e, N_h} \frac{e^2}{\epsilon^* |\mathbf{r}_i^e - \mathbf{r}_j^h|^2}, \quad (1)$$

with $\alpha = e, h$. Here, m_α^* and ϵ^* are effective masses and the dielectric constant, respectively, g^* is the effective giromagnetic factor, μ_B the Bohr magneton, $\mathbf{A}$ is the vector potential generating the magnetic field $\mathbf{B}$, and $e = |e|$; all parameters are taken for the AlGaAs class of materials. Equations (1) neglect non-parabolicity and spin-orbit effects, but otherwise describe samples with realistic properties, such as layer width and finite band offsets, by means of the effective potentials $V_\alpha(\mathbf{r})$. Typical samples that we shall consider consist of two identical QDs obtained from symmetric coupled quantum wells, grown, say, along the z direction, of width L_W , separated by a barrier d , and with band-offset V_0 between wells and barriers. The lateral confinement can be often taken as a two-dimensional (2D) parabolic potential characterized by an energy $\hbar\omega_0$ which may assume rather different values, depending on whether QDs are obtained by gating or by etching of a 2D electron/hole gas, by self-assembly, etc. The lateral confinement energy $\hbar\omega_0$ is typically in the range $1 \div 10$ meV and often much weaker than the confinement along the growth direction.

Our numerical approach consists of mapping the single-particle terms H_α in a real-space grid, leading to a large sparse matrix that is diagonalized by Lanczos methods. Single-particle spin-orbitals thus obtained are then used to build a basis of Slater determinants for the N -particle problem, which is then used to include the two-body term, in the familiar configuration interaction approach. Coulomb matrix elements are calculated numerically. The resulting matrix, which can be very large, is again sparse and can be diagonalized via the Lanczos method [17].

The paper is organized as follows. In Sec. 2 we discuss electron-hole complexes, and particularly biexciton binding and localization. In Sec. 3 we discuss the phase diagram of few-electron systems in a magnetic field. In particular, in Sec. 3.1 we consider the effects of tunneling on these systems in a large vertical field, where carriers are localized in the so-called Wigner molecule. Finally, in Sec. 3.2, we show how spin-ordering of two electrons can be manipulated by a magnetic field of arbitrary strength and direction.

2. Electron-Hole Complexes in Artificial Molecules

In this section we are concerned with the biexciton states in two identical vertically-coupled QDs. The main focus is on the effects of the

Coulomb-induced interdot correlations, which are investigated as a function of the barrier width d [18]. The features of the biexcitonic states are shown to critically depend on the detailed balance between three phenomena: (i) the interdot tunneling, which is a single-particle feature, tending to delocalize the carriers and to spread them over the artificial molecule; (ii) the homopolar Coulomb interaction, which is a few-particle effect, resulting in spatial correlations between the two electrons (holes) that tend to minimize the repulsion energy by localizing the identical carriers in opposite dots; (iii) the heteropolar Coulomb interaction, that induces spatial correlations between electrons and holes in order to maximize their overlap. Interestingly enough, the interplay between these trends can be widely tuned, either by modifying the structural parameters of the dots or by applying external fields, to induce non-trivial behaviours already at the few-particle level.

In the following we fix all the physical parameters of the AM except the interdot distance d in order to explore the different coupling (tunneling) regimes. At the smallest interdot distance, the symmetric-antisymmetric (S-AS) splitting $2t^{(e,h)} = \epsilon_{AS}^{(e,h)} - \epsilon_S^{(e,h)}$ is maximized, where ϵ_S (ϵ_{AS}) is the energy of the single-particle S (AS) level. For the sample of Fig. 1, for example, where d is varied in the range $1 \div 3$ nm, the tunneling splitting amounts to about 12 and 6 meV for the electrons and the holes, respectively, at $d = 1$ nm. As a result, both electrons are frozen in the S state and the spatial distribution of each electron is uncorrelated both with respect to the other electron position and to that of the two holes. Due to their smaller effective mass, the holes tunnel less efficiently than the electrons [$t^{(h)} < t^{(e)}$]. Therefore, the interdot spatial correlation, which requires the occupation of the AS states, is favored and acts in such a way that the carriers are always localized in different dots. This is shown in the left-hand insets of Fig. 1(b-d), where we plot the dependence of the pair-correlation on the relative position between the carriers.

By progressively increasing the barrier width, the effects of the homopolar Coulomb interactions between the carriers continuously increase, whereas electrons and holes are not correlated with each other. In this kind of configuration, resulting in a factorized wavefunction

$$\psi(\mathbf{r}_{e_1}, \mathbf{r}_{e_2}, \mathbf{r}_{h_1}, \mathbf{r}_{h_2}) \simeq \phi^{(e)}(\mathbf{r}_{e_1}, \mathbf{r}_{e_2}) \phi^{(h)}(\mathbf{r}_{h_1}, \mathbf{r}_{h_2}), \quad (2)$$

the probability of having double occupancies (two electrons or two holes) in each dot is strongly suppressed, i.e. the two excitons are localized in different dots. In approaching the weak-coupling regime, the biexciton ground state undergoes a rapid transition towards a maximally correlated configuration, where all the carriers are localized in the same dot

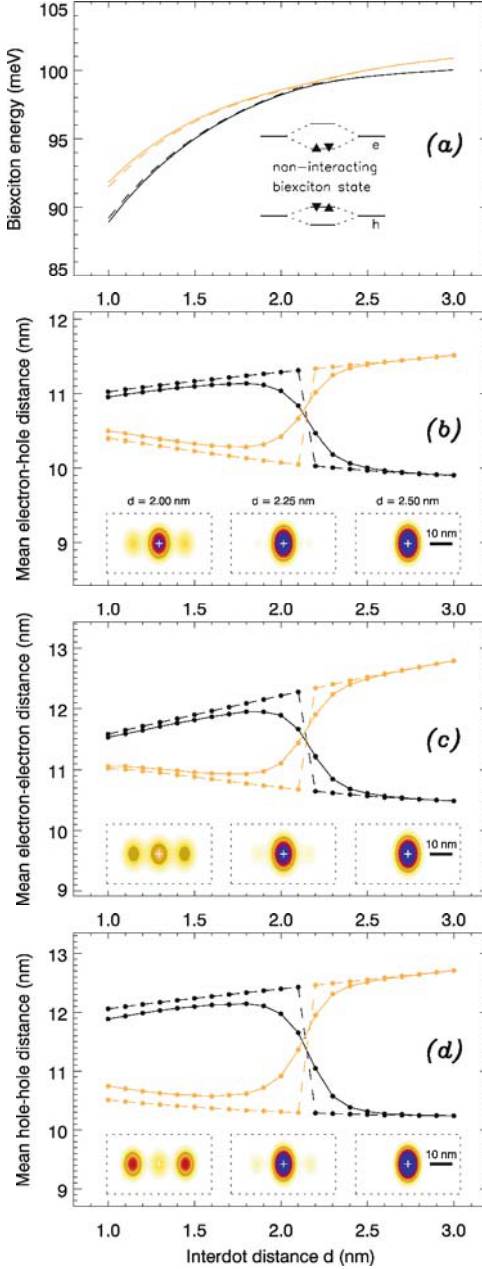


Figure 1. Biexciton energies (a) and mean distances between carriers (b-d) as a function of the interdot distance d . The plots refer to the four states lowest in energy, with even (continuous lines) or odd (dotted lines) parity. The insets (b-d) show the dependence of the pair-correlation functions on the relative position $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$ of the two carriers. In particular, we take $y=0$ and show the dependence on the z (horizontal axis) and x (vertical axis) coordinates. The two vertically-coupled GaAs/AlGaAs QDs are modeled by means of a confinement potential which is double-well like in the growth direction (the well width and band offsets are $L_W = 10$ nm and $V_0^{(e,h)} = 400, 215$ meV respectively) and parabolic in the plane ($\hbar\omega_0^{(e,h)} = 20, 3.5$ meV, $m_{(e,h)}^* = 0.067, 0.38 m_0$).

[see the right-hand insets of Fig. 1(b-d)]. The reason why this arrangement is energetically favored as compared to the previously discussed one is entirely related to the in-plane correlations between the carriers. In fact, due to the substantial symmetry of the electron and hole wavefunctions, the Coulomb energy of the two configurations is the same for both types of particles in the mean-field limit. The increase in the Coulomb repulsion arising from the stronger degree of localization is cancelled by that of the Coulomb attraction. The occupation of the higher Fock-Darwin states, however, gives rise to additional (in-plane) spatial correlations, such as those that cause the biexciton binding energies in single QDs. This kind of interaction, whose closest classical analogue is an induced dipole-induced dipole force, is short-ranged and is therefore ineffective as far as the two excitons are localized in different dots.

In Fig. 1 we plot the energy levels [panel (a)] and the average distance between each pair of carriers [panels (b-d)] for the four lowest biexciton states. The continuous (dotted) lines correspond to the two states of even (odd) symmetry. Two features emerge from the plots: (i) the transition between the weakly- and the highly-correlated configurations occurs quite abruptly for $d \simeq 2$ nm, and (ii) in a limited interval around this value, each of the four states swaps its features with a state of equal symmetry, and correspondingly an anticrossing occurs between their energy levels.

Interdot correlation is therefore seen to play a crucial role in determining the carrier localization in artificial molecules, which exhibit a fully three-dimensional nature and novel behavior occurs as compared to the single dots. Indeed, such features have to be taken into account within the design of exciton-based quantum computation schemes and devices [19], where the double occupancies are known to spoil the required (tensorial) Hilbert-space structure.

3. Interacting Electrons in Artificial Molecules in Magnetic Fields

In this section we shall consider the effects of a magnetic field of arbitrary strength and direction on the few-electron states in AMs. Let us first summarize a few well-known effects of a field that is parallel to the axis of the AM.

Figure 3 shows the single-particle levels at zero field, separated by $\hbar\omega_0$, resulting from the 2D parabolic potential of the in-plane confinement. The two sets of shells are separated by the tunneling energy. As a vertical field is switched on, the Hamiltonian can still be analytically solved [20],

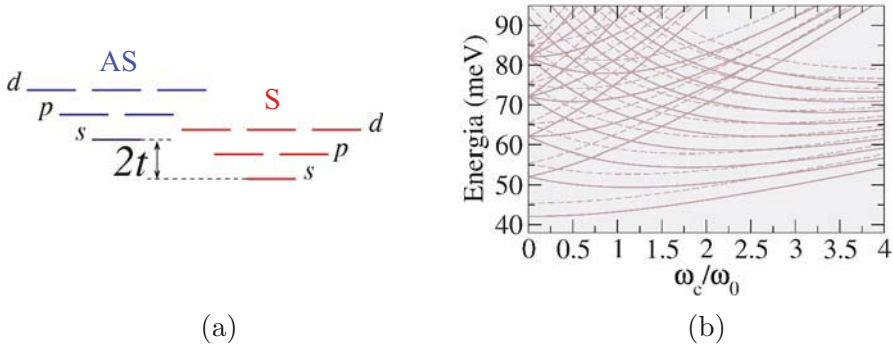


Figure 2. (a) Sketch of the energy levels in a 2D parabolic potential for an AM at zero magnetic field. The labels s, p, d indicate the value of the single-particle magnetic moment consistent with the corresponding atomic notation. S and AS denote symmetric and antisymmetric levels arising from two coupled QDs with tunneling energy $2t$. (b) Fock-Darwin levels (Eq. 3) for an AM with $L_w = 10$ nm, $d = 3$ nm, $V_0 = 300$ meV, and $\hbar\omega_0 = 10$ meV.

and gives rise to the Fock–Darwin (FD) levels

$$\varepsilon_{nm}^i = \varepsilon_i + \hbar\Omega(2n + |m| + 1) - \frac{1}{2}\hbar\omega_c m. \quad (3)$$

Here, n is the principal quantum number describing the radial distribution of the wavefunction, and m the azimuthal quantum number labelling the angular momentum, which is conserved in this cylindrically symmetric configuration. The oscillator frequency $\Omega = (\omega_0^2 + \frac{1}{4}\omega_c^2)^{1/2}$, with the cyclotron frequency $\omega_c = eB/m^*c$, shows the competition between the confinement energy gaps, $\hbar\omega_0$, and the field-induced gaps, $\hbar\omega_c$. The energy ε_i is the confinement energy along the growth direction. For AMs with sufficiently narrow symmetric quantum wells, we can limit our considerations to two levels, which are the ground S and AS states ($i = S, AS$). The Fock-Darwin states for an AM are shown with lines in Fig. 3. At zero field one recovers the $s, p, \dots$ shells of Fig. 3. At finite field, the $\pm m$ orbital degeneracy splits, and at large fields the levels with highest m tend to form highly degenerate Landau levels.

Figure 3 shows a qualitative sketch of the ground state phase diagram of few-electrons in a AM which will be discussed in the next sections. Let us focus, for the time being, on the vertical field effects. When a field is applied parallel to the growth axis, the effect is to squeeze the states in the QD plane. By doing so, the Coulomb energy increases and, at sufficiently large field, the system becomes spin-polarized to gain in exchange energy. At even larger fields, the kinetic energy is quenched, since all single-particle levels tend to become degenerate [Fig. 3], and correlation dominates the system. Accordingly, the electron system first

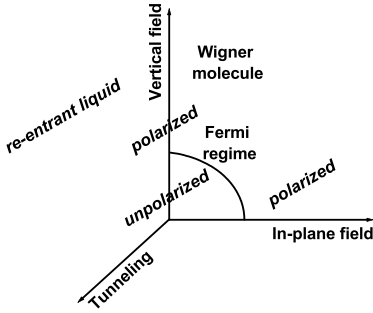


Figure 3. Sketch of the ground state phase diagram for an AM with respect to field strength and direction, and tunneling.

becomes a spin-polarized confined Fermi sea, laterally delocalized over the QDs. Then, at very high fields, the systems transforms to a confined Wigner crystal, in which the carriers are localized in the plane of the QD. The latter regime is often called Wigner molecule [21].

There are two obvious directions along which the phase diagram of a AM can be extended: the tunneling energy and an in-plane component of the magnetic field, as shown in Fig. 3. These two regions will be discussed next.

3.1 Wigner Molecules and Tunneling

As an example of the effects of tunneling in the high field regime, we show in Fig. 4 the calculated ground-state energy *versus* the inter-dot distance d for six electrons [22]. According to the previous discussion, we assume that at this field carriers are spin polarized. Furthermore, at the high field considered here, electrons in a single QD are expected to be localized, and quantum fluctuations to play a minor effect. In the inset we also show the geometrical configuration of the localized carriers that will be discussed in more detail in Fig. 5. At small distances the energy increases with d (phase I), because the kinetic energy exponentially grows due to the progressive localization of the wavefunction into the dots.: electrons occupy only S orbitals, whose energies increase. At these small values of d the AM behaves as a single QD, and electrons sit at the vertices of a centered pentagon, which minimizes the Coulomb energy, as predicted by a classical static calculation [23]. There is another obvious regime: when the barrier is large, the six electrons sit, three per QD, at the vertices of two triangles staggered by 60° (Phase III). Close to $d = 5$ nm inter-dot tunneling stabilizes Phase II, in which electrons sit at the vertices of an hexagon. The figure also shows the calculated average radius $\langle \varrho \rangle$ [$\varrho \equiv (x, y)$]. Figure 4 shows that these states are incompressible, in the same sense of Laughlin's states of the FQHE [24]. Indeed, varying d acts like an external pressure applied in

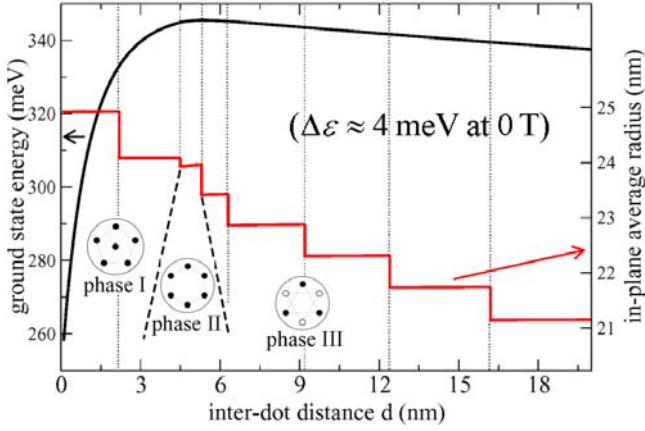


Figure 4. Ground-state energy (left axis) and in-plane average radius $\langle \varrho \rangle$ (right axis) versus inter-dot distance d for six electrons at $B = 25$ T. Sample parameters are: $\hbar\omega_0 = 3.7$ meV, $V_0 = 250$ meV, and $L_W = 12$ nm. Insets show the electron arrangements in the different phases.

the z direction, forcing the wavefunction to change. However, due to a cusp-like structure of the energy spectrum [1], this happens only in a discontinuous way, except for Phase II. We show below that Phase II has very special properties and is stabilized by tunneling fluctuations.

To analyze the ground state of the artificial molecule in the different regimes, we show in Fig. 5 the pair correlation function

$$P(\varrho, z; \varrho_0, z_0) = \frac{1}{N_e(N_e - 1)} \sum_{i \neq j} \langle \delta(\varrho - \varrho_i) \delta(z - z_i) \delta(\varrho_0 - \varrho_j) \delta(z_0 - z_j) \rangle, \quad (4)$$

where the average is over the ground state. Figure 5 shows $P(\varrho, z; \varrho_0, z_0)$ along a circle in the same dot (solid line) or in the opposite dot (dashed line) with respect to the position of a reference electron, taken at the maximum of its charge density, (ϱ_0, z_0) . The right column shows the electron arrangement in the QDs as inferred from the maxima of $P(\varrho, z; \varrho_0, z_0)$.

At small d (Phase I) the whole system is coherent, i.e. it behaves as a unique QD. Indeed, the correlation function peaks, forming the outer shell of the centered pentagon, have the same height in both QDs. At intermediate values of the tunneling energy (Phase II) the peaks, corresponding to the vertices of a regular hexagon, have different heights. Finally, when d is sufficiently large (Phase III), the structure evolves into

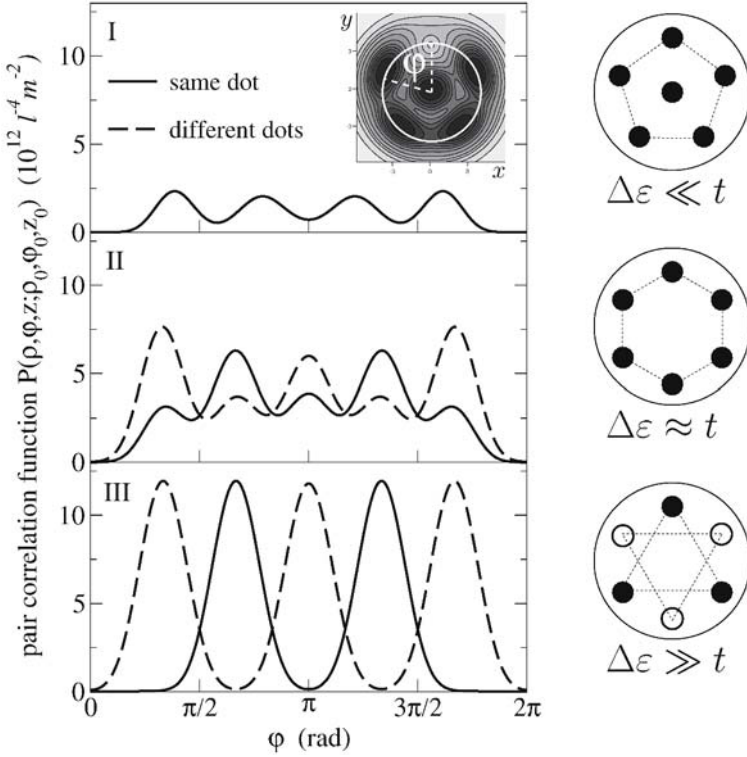


Figure 5. Angular correlation function of the three phases of Fig. 4. The three phases correspond to the interdot distances $d = 2, 4.6, 8$ nm, respectively.

two isolated dots coupled only via Coulomb interaction. Accordingly, the peaks are again of the same height, but shifted by 60° .

It is important to note from Fig. 5 that Phase I and III are strongly localized phases, as demonstrated by the high peak-to-valley ratio of the correlation function, and quantum fluctuations play a minor role. Therefore, electron configurations are basically determined by Coulomb interactions, and have completely classical counterparts [23]. On the contrary, in Phase II, tunneling fluctuations prevent electrons from localizing and the configuration has a “liquid” character. Such a phase cannot be explained in terms of Coulomb interactions alone and, in fact, the hexagonal arrangement shown in Fig. 5 is classically unstable. Thus, tunneling fluctuations may induce melting of the otherwise well-localized Wigner molecule in the high-field regime and induce reentrant liquid phases, as schematically indicated in Fig. 3. A discussion of the possible experimental signatures of the different phases in inelastic light scattering experiments can be found in Ref. [22].

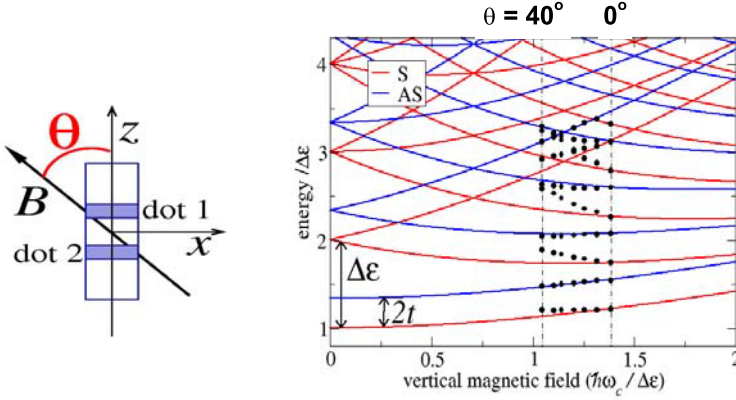


Figure 6. Evolution of single particle states in an AM when the field is rotated from the vertical direction by an angle θ . Sample parameters are: $L_w = 10$ nm, $d = 3$ nm, $V_0 = 300$ meV, $\hbar\omega_0 = 10$ meV. Solid lines represent the Fock-Darwin states ε_{nm}^i induced by a strictly vertical field.

3.2 Effects of an In-Plane Magnetic Field

Next, we discuss the effect of a finite in-plane component of the field, $B_{\parallel}$. To show the effect on the single-particle states, let us suppose that the total field is tilted by an angle θ with respect to the growth direction, so that $B_{\parallel} \neq 0$. As shown in Fig. 6, with increasing θ , energy levels tend to follow the FD states backward, which are shown for comparison, since of course $B_{\perp}$ decreases with increasing angle. In addition, however, the splitting between S and AS levels decreases. This shows that an in-plane component of the field suppresses tunneling. Note that this effect is larger for higher levels. Note also that here the S/AS labelling is used for brevity; obviously, for a general field direction with respect to the tunneling direction wavefunctions do not have a well defined S/AS symmetry. It is important to stress that deviations from the FD states are expected because the energy scale associated with the in-plane field is comparable to the tunneling gap. For single QDs, for example, a reasonable in-plane field will not affect the FD states, since the in-plane field energy scale is typically much smaller than the single-particle gaps induced by the quantum well confinement.

In AMs carriers sitting on either dot are not only electrostatically coupled, but also have their spin interlaced when tunneling is allowed [25]. This is sketched in Fig. 7(a). For two electrons in a singlet state it is possible to tunnel into the same dot. By doing so, they gain the tunneling energy t , which may compensate for the loss in the Coulomb energy U . This process is forbidden for two electrons in the triplet state

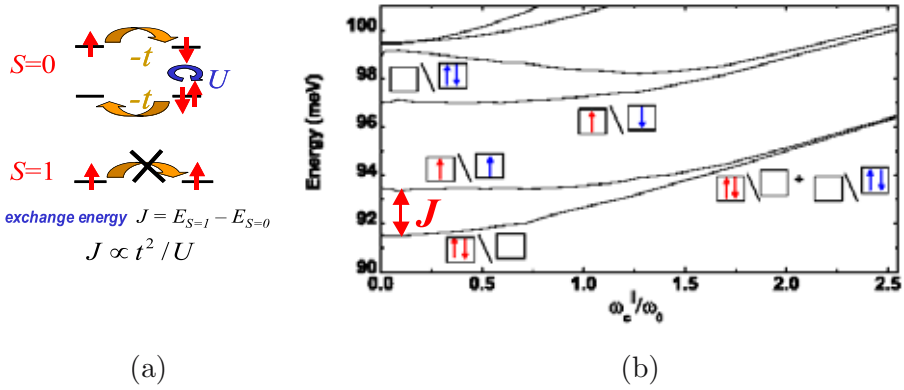


Figure 7. (a) Sketch of the energy contributions to the tunneling-induced spin-spin interaction. (b) Two-electron levels, with indication of the main component of the wavefunctions in terms of S (left boxes) and AS (right boxes) states. Parameters are: $L_w = 10$ nm, $d = 3$ nm, $V_0 = 300$ meV, and $\hbar\omega_0 = 10$ meV.

by Pauli blocking. Different spin orderings, therefore, are associated with an exchange energy J which, to lowest order in perturbation theory, is $J \propto t^2 / U$. While in real molecules J is fixed by the bond length, in AMs it is possible to tune almost all energy scales by sample engineering and external fields. Since an in-plane field affects tunneling, we expect the exchange energy to be affected by an in-plane field as well.

Guided by the above considerations, we next consider the two-electron system [26]. At low vertical fields the ground state of single and coupled QDs is known to be a singlet state [27, 28]. In the moderate field regime, therefore, the lowest energy levels are nearly unaffected by the rotation, except for the shift due to the reduction of the tunneling energy, with the singlet state being the lowest.

At sufficiently high vertical field singlet-triplet transitions take place at a given threshold field. Since the exchange energy is proportional to tunneling, we expect that the threshold fields will be lowered as $B_{\parallel}$ increases [26]. This is shown in Fig. 8. The singlet state is stable in the low-field regime. The triplet state becomes favored in the large-field regime, whether the field is perpendicular or parallel to the plane of the QD. It should be noted, however, that this happens by different mechanisms whether $B_{\perp}$ or $B_{\parallel}$ is large. In the former case, the squeezing of the wavefunction has a Coulomb energy cost which can only be avoided by triplet spin order. This is analogous to single QDs. However, while a finite $B_{\parallel}$ would not affect very much electronic states in single QDs, where single-particle gaps are large, in AMs the in-plane field affects the

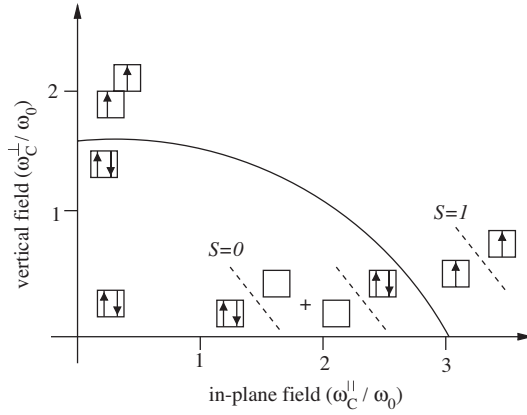


Figure 8. Singlet-triplet phase diagram calculated for a GaAs AM with $L_W = 10$ nm, $d = 3$ nm, $\hbar\omega_0 = 4$ meV. The insets show the single-particle occupation in terms of S (left to the dashed lines) and AS orbitals (right to the dashed lines). At low $B_{||}$ only S orbitals are occupied.

S/AS gap. When this vanishes, no tunneling energy is lost by localizing in each dot. The spin ordering then becomes irrelevant, and the triplet state is favored only due to Zeeman energy [Fig. 7(b)].

In Figs. 7(b) and 8 we show the character of the two-electron wavefunction of the ground state. In the low $B_{||}$ regime, the two electrons occupy only the S state, either with the s symmetry with opposite spin (low field) or the s and p symmetry levels with the same spin orientation (high field), since a large vertical field reduces the s - p gaps. In the large $B_{||}$ regime, on the contrary, S and AS states become degenerate, and are equally occupied by the two electrons, due to Coulomb correlations.

In Fig. 9 we show the exchange energy J defined as the difference between the energy of the lowest triplet and the singlet levels, as a function of the in-plane field at zero vertical field. This is positive (i.e. the

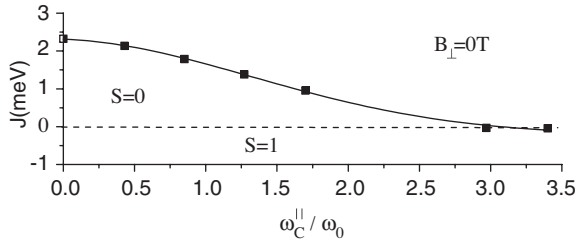


Figure 9. Exchange energy $J = E_{S=1} - E_{S=0}$ for a GaAs AM. Same parameters as in Fig. 8.

singlet is the ground state) at low fields, but rapidly decreases as the field increases. At large fields, the exchange energy changes sign, being eventually dominated by Zeeman energy.

Acknowledgements

This work was supported in part by MIUR-FIRB *Quantum phases of ultra-low electron density semiconductor heterostructures*, and by INFN I. T. *Calcolo Parallelo* (2003). Con il contributo del Ministero degli Affari Esteri, Direzione Generale per la Promozione e la Cooperazione Culturale.

References

- [1] P. A. Maksym and T. Chakraborty. Quantum dots in a magnetic field: Role of electron-electron interactions. *Phys. Rev. Lett.* **65**: 108–111, 1990.
- [2] M. A. Kastner. Artificial atoms. *Phys. Today* **46**(1): 24–31, 1993.
- [3] R. Ashoori. (1996). Electrons in artificial atoms *Nature (London)* **379**: 413–419, 1996.
- [4] S. Tarucha, D. G. Austing, T. Honda, R. J. van der Hage, and L. P. Kouwenhoven. Shell filling and spin effects in a few electron quantum dot *Phys. Rev. Lett.* **77**: 3613–3616, 1996.
- [5] M. Rontani, G. Goldoni, and E. Molinari. *New Directions in Mesoscopic Physics (towards Nanoscience)*. (Kluwer, Dordrecht, 2003). See also cond-mat/0212626.
- [6] G. W. Bryant. Electronic structure of ultrasmall quantum-well boxes. *Phys. Rev. Lett.* **59**: 1140–1143, 1987.
- [7] D. G. Goldhaber-Gordon, H. Shtrikman, D. Mahalu, D. Abusch-Magder, U. Meirav, and M. A. Kastner. Kondo effect in a single-electron transistor. *Nature (London)* **391**: 156–159, 1998.
- [8] S. M. Cronenwett, T. H. Oosterkamp, and L. P. Kouwenhoven. A tunable Kondo effect in quantum dots. *Science* **281**: 540–544, 1998.
- [9] L. P. Kouwenhoven, F. W. J. Hekking, B. J. van Wees, C. J. P. M. Harmans, C. E. Timmering, and C. T. Foxon. Transport through a finite one-dimensional crystal. *Phys. Rev. Lett.* **65**: 361–364, 1990.
- [10] C. Livermore, C. H. Crouch, R. M. Westervelt, K. L. Campman, and A. C. Gossard. The Coulomb blockade in coupled quantum dots. *Science* **274**: 1332–1335, 1996.
- [11] R. H. Blick, D. Pfannkuche, R. J. Haug, K. von Klitzing, and K. Eberl. Formation of a coherent mode in a double quantum dot. *Phys. Rev. Lett.* **80**: 4032–4035, 1998.
- [12] M. Brodsky, N. B. Zhitenev, R. C. Ashoori, L. N. Pfeiffer, and K. W. West. Localization in artificial disorder: Two coupled quantum dots. *Phys. Rev. Lett.* **85**: 2356–2359, 2000.
- [13] M. Rontani, F. Troiani, U. Hohenester, and E. Molinari. Quantum phases in artificial molecules. *Solid State Commun.* **119**: 309–321, 2001.

- [14] M. Pi, A. Emperador, M. Barranco, F. Garcias, K. Muraki, S. Tarucha, and D. G. Austing. Dissociation of vertical semiconductor diatomic artificial molecules. *Phys. Rev. Lett.* **87**: art. no. 066801, 2001.
- [15] D. Pfannkuche, V. Gudmundsson, and P. A. Maksym. Comparison of a Hartree, a Hartree-Fock, and an exact treatment of quantum-dot helium. *Phys. Rev. B* **47**: 2244–2250, 1993.
- [16] B. Szafran, S. Bednarek, and J Adamowski. Correlation effects in vertical gated quantum dots. *Phys. Rev. B* **67**: art. no. 115323, 2003.
- [17] R. B. Lehoucq, K. Maschhoff, D. C. Sorensen, and C. Yang. *ARPACK computer code*. Available at <http://www.caam.rice.edu/software/ARPACK/>.
- [18] F. Troiani, U. Hohenester, and E. Molinari. Electron-hole localization in coupled quantum dots. *Phys. Rev. B* **65**: art. no. 161301, 2002.
- [19] F. Troiani, U. Hohenester, and E. Molinari. Exploiting exciton-exciton interactions in semiconductor quantum dots for quantum-information processing. *Phys. Rev. B* **62**: R2263–R2266, 2000.
- [20] L. Jacak, P. Hawrylak, and A. Wójs. *Quantum dots*. (Springer-Verlag, Berlin, 1998).
- [21] P. A. Maksym, H. Imamura, G. P. Mallon, and H. Aoki, H. Molecular aspects of electron correlation in quantum dots. *J. Phys.: Condens. Matter* **12**: R299–R334, 2000.
- [22] M. Rontani, G. Goldoni, F. Manghi, and E. Molinari. Raman signatures of classical and quantum phases in coupled dots: A theoretical prediction. *Europhys. Lett.* **58**: 555–561, 2002.
- [23] B. Partoens, V. A. Schweigert, and F. M. Peeters. Classical double-layer atoms: Artificial molecules. *Phys. Rev. Lett.* **79**: 3990–3993, 1997.
- [24] R. B. Laughlin. (1983). Quantized motion of three two-dimensional electrons in a strong magnetic field. *Phys. Rev. B* **27**: 3383–3389, 1983.
- [25] G. Burkard, G. Seelig, and D. Loss. Spin interactions and switching in vertically tunnel-coupled quantum dots. *Phys. Rev. B* **62**: 2581–2592, 2000.
- [26] D. Bellucci, M. Rontani, G. Goldoni, F. Troiani, and E. Molinari. Spin-spin interaction in artificial molecules with in-plane magnetic field. *Physica E* (in press).
- [27] U. Merkt, J. Huser, and M. Wagner. Energy spectra of two electrons in a harmonic quantum dot. *Phys. Rev. B* **43**: 7320–7323, 1991.
- [28] J. H. Oh, K. J. Chang, G. Ihm, and S. J. Lee. Electronic structure and optical properties of coupled quantum dots. *Phys. Rev. B* **53**: 13264–13267, 1996.

PHASE TRANSITIONS IN WIGNER MOLECULES

J. Adamowski, B. Szafran and S. Bednarek

Faculty of Physics and Nuclear Techniques,

AGH University of Science and Technology, Kraków, Poland

adamowski@ftj.agh.edu.pl

Abstract Electrons confined in quantum dots can form island-like space structures called *Wigner molecules*. We discuss the ground-state properties of two-dimensional Wigner molecules. In particular, we consider the formation of different phases (isomers) of the Wigner molecules at high magnetic fields. The N -electron system, confined in the quantum dot and subject to a sufficiently strong magnetic field, forms a fully spin-polarized maximum density droplet (MDD). At high enough magnetic field the MDD decays and the Wigner molecule is formed with an island-like electron distribution. For $N \geq 6$ several different phases of the N -electron Wigner molecule have been predicted. At extremely high magnetic fields, the spatial distribution of the electrons is the same as that in a classical system of equal point charges. Possible mechanisms of MDD breakdown, i.e. hole formation in the occupation number distribution and an edge reconstruction, are addressed. We also consider the creation of Wigner molecules without applying an external magnetic field in an electron system confined within a single quantum dot of large enough size and compare these Wigner molecules with artificial molecules formed in coupled quantum dots. Possible experimental evidence is examined for the formation of different phases of Wigner molecules.

Keywords: Quantum dot, Wigner crystal, artificial atoms and molecules.

1. Introduction

Electrons in strongly correlated systems can form so-called Wigner phases. In the many-electron system interacting with the positive background a Wigner crystal is created if the density of electrons is sufficiently low [1]. In a few-electron system confined in the quantum dot the existence of Wigner molecules [2, 3], also called [5], has been studied [2–12]. In two-dimensional (2D) space, the Wigner

crystal possesses only one stable phase with a triangular (hexagonal) crystalline lattice [13, 14]. Nevertheless, the energy of a 2D Wigner crystal with a quadratic lattice is only slightly higher [13, 14]. The ground-state space symmetry of 2D Wigner molecules has been predicted [11, 12] to be different in different magnetic-field regimes.

In an external magnetic field, the electrons confined in the quantum dot undergo several ground-state transformations [15] that are associated with the change of spin-orbital configuration [16]. If the magnetic field increases, the electron system becomes spin-polarized and forms a phase, called a maximum-density droplet (MDD) [17, 18, 19]. In this phase, the electrons occupy Fock-Darwin orbitals with consecutive orbital angular momenta. In vertical pillar-shape quantum dots [15], the electron density distribution possesses cylindrical symmetry. If the magnetic field is sufficiently strong, the MDD phase decays and a molecular phase is formed with the Wigner-type of electron localization. MDD breakdown was observed in vertical gated quantum dots [20]. Recently, we have predicted [11] the existence of different phases (isomers) of a Wigner molecule with the different spatial symmetry. According to our results [11], with increasing magnetic field, the 2D Wigner molecule, which consists of at least six electrons, undergoes several phase transitions between isomers with different electron density distribution. The most stable isomer of the Wigner molecule created at extremely high magnetic field possesses the same spatial configuration as its classical counterpart. For the six-electron Wigner molecule, Hartree-Fock results [11, 12] agree with those obtained by exact configuration interaction calculations [21]. Similarly to the Wigner crystal, a Wigner molecule can also be created in the absence of the external magnetic field by lowering the electron density, i.e. increasing the quantum-dot size.

In this article, we present a review of the properties of the Wigner molecules. In Sec. 1, we discuss the classical Wigner molecules, and in Sec. 2, we present a brief account of the quantum theory of the Wigner molecules, Section 3 contains the most important results, obtained in Refs. [11, 12, 22], Sec. 4 contains a discussion, and our results are summarized in Sec. 5.

2. Classical Wigner Molecules

The classical Wigner molecule is the system of N equally charged point particles confined in a certain external potential. The results of computer simulations for this system are presented in Refs. [23, 24, 25]. The total potential energy of the N -particle classical Wigner molecule

is given by

$$U_{\text{tot}}^c = \sum_{i=1}^N \left[V_{\text{conf}}(\mathbf{R}_i) + \sum_{j<i}^N \frac{\kappa}{|\mathbf{R}_i - \mathbf{R}_j|} \right], \quad (1)$$

where $\mathbf{R}_i$ are the position vectors of the particles, $V_{\text{conf}}(\mathbf{R}_i)$ is the confinement potential energy, $\kappa = e^2/4\pi\epsilon_0\epsilon$, ϵ is the dielectric constant of the medium, e is the elementary charge, and all the particles have charge equal to either $+e$ or $-e$. Let us assume the confinement potential to be

$$V_{\text{conf}}(R) = \frac{1}{2}m_e\omega_0^2 R^2, \quad (2)$$

where m_e is the mass of the particle (in the following it will be the effective mass of the electron) and ω_0 is the confining frequency. The stable equilibrium configuration of the classical Wigner molecule corresponds to the global minimum of potential energy (1) and the metastable configurations correspond to the local minima of (1).

Figure 1 displays the stable and metastable configurations of the 2D classical Wigner molecules found by the simulated annealing method. The different configurations of the N -particle Wigner molecule are labelled as follows: N_1 - N_2 - N_3 denotes the isomer, in which N_1 , N_2 , and N_3 point charges are located on the inner, middle, and outer ring, respectively, whereby $N = N_1 + N_2 + N_3$. If only two rings are occupied, we denote the corresponding configuration by N_1 - N_2 . In Fig. 1, the configuration depicted in the left panel is the stable configuration, while the right panel displays the metastable configuration. The stable equilibrium configuration is usually realized if a larger number of particles is gathered inside the molecule. If the number of particles increases, the classical point-charge system steadily goes over into the triangular Wigner crystal lattice [23, 24, 25].

The results of the computer simulations [23, 24, 25] have been confirmed experimentally [26]. The authors [26] studied the stable and metastable configurations of the system of N equally charged millimeter-size steel balls in the electrostatic field. The agreement between the measured and calculated configurations is good. The few discrepancies can be explained by the fact that the electrostatic interaction between the finite-size metallic balls is different than that between the point charges.

3. Quantum Theory of Wigner Molecules

First we consider a single electron moving on the $x - y$ plane in a magnetic field $\mathbf{B} = (0, 0, B)$. In the non-symmetric (Landau) gauge the

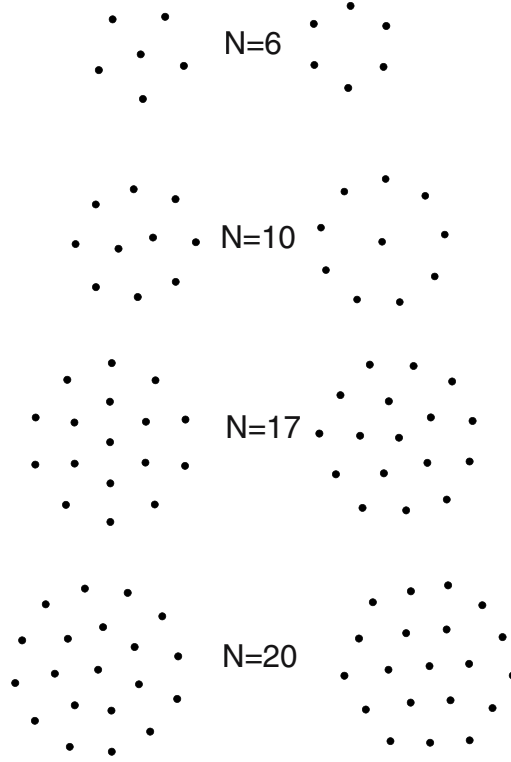


Figure 1. Stable (left panel) and metastable (right panel) configurations of N -electron classical Wigner molecules.

one-electron Hamiltonian has the form

$$H = -\frac{\hbar^2}{2m_e} \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) + i\hbar\omega_c y \frac{\partial}{\partial x} + \frac{1}{2}m_e\omega_c^2 y^2, \quad (3)$$

where $\omega_c = eB/m_e$ is the cyclotron frequency.

As each Landau level, the ground-state energy level, $E_0 = \frac{1}{2}\hbar\omega_c$, is infinitely degenerate. Therefore, the ground-state wavefunction can be chosen in many ways. We have chosen [11, 12] the one-electron wavefunction in the form of a displaced Gaussian with a phase factor. Explicitly,

$$\psi_{\mathbf{R}}(\mathbf{r}) = \left(\frac{\alpha}{2\pi} \right)^{1/2} \exp \left[-\frac{1}{4}\alpha(\mathbf{r} - \mathbf{R})^2 + \frac{1}{2}i\alpha(x - X)(y + Y) \right], \quad (4)$$

where $\mathbf{r} = (x, y)$, $\mathbf{R} = (X, Y)$ is an arbitrary vector, and $\alpha = eB/\hbar$. This wavefunction leads to an electron density distribution having the shape of the Gaussian centered at site $\mathbf{R}$, which can be treated as the center

of the Landau orbital. Such wavefunctions centered at different $\mathbf{R} = \mathbf{R}_i$, where $i = 1, \dots, N$, have been applied [11] to the construction of a suitable multicenter basis for the N -electron problem.

Let us consider a system of N spin polarized electrons confined within the parabolic quantum dot in magnetic field $\mathbf{B}$. The Hamiltonian of the system has the form

$$H = \sum_{i=1}^N \left[h_i + V_{\text{conf}}(r_i) + \sum_{j<i}^N \frac{\kappa e^2}{|\mathbf{r}_i - \mathbf{r}_j|} \right] + \frac{1}{2} N g^* \mu_B B, \quad (5)$$

where V_{conf} is given by Eq. (2), g^* is the effective Landé factor of the semiconductor, μ_B is the Bohr magneton, and the last term is the spin Zeeman energy.

The N -electron problem has been solved [11, 12] by the unrestricted Hartree-Fock method. The Slater determinant is built from the one-electron wavefunctions Ψ^μ of the form

$$\Psi^\mu(\mathbf{r}) = \sum_{i=1}^N c_i^\mu \tilde{\psi}_{\mathbf{R}_i}(\mathbf{r}), \quad (6)$$

where $\mu = 1, \dots, N$, c_i^μ are the linear variational parameters, and basis wavefunctions $\tilde{\psi}_{\mathbf{R}_i}(\mathbf{r})$ have form (4) with parameter α replaced by the nonlinear variational parameter α^* .

The positions $\mathbf{R}_i$ of N centers in wavefunction (6) are determined from classical configurations $\mathbf{R}_i^c$, which correspond to the local minima of the potential energy (1). In the quantum-mechanical calculations, we apply [11, 12] the following uniform scaling: $\mathbf{R}_i = \sigma \mathbf{R}_i^c$, where σ is the second nonlinear variational parameter.

4. Properties of the Wigner Molecules

In this Section, we present the results of the Hartree-Fock calculations [11, 12] performed for $N = 2, \dots, 20$ electrons confined in the 2D parabolic quantum dot and subject to an external magnetic field. The material parameters employed are those of GaAs. We use the same notation for the different isomers of the quantum Wigner molecule as that introduced in Sec. 2 for their classical counterparts, i.e. N_1 - N_2 - N_3 . However, the localization of the electrons within the subsequent rings (shells) should be interpreted in terms of the quantum localization.

Figures 2 and 3 display the magnetic-field induced phase transitions for $N = 6$ and 10 electrons, respectively. The ground-state electron density distribution is also shown on these figures. At magnetic fields lower

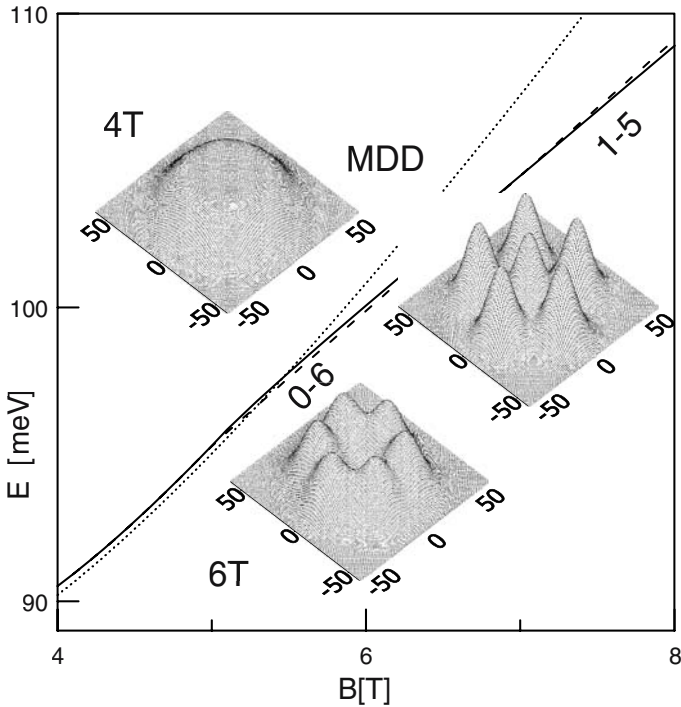


Figure 2. Ground-state energy E of the Wigner molecule isomers 1–5 (solid curve) and 0–6 (dashed curve) and the six-electron MDD (dotted curve) as a function of magnetic field B . Inset: the electron spatial distribution on the $x - y$ plane. Length is expressed in nanometers. (From Ref. [12]).

than ~ 5 T, the MDD is the lowest-energy phase of the electron system. For higher magnetic fields stable molecular phases are formed. The Wigner-molecule isomer formed just after the breakdown of the MDD possesses configuration 0– N with all the electrons localized within the outer shell. If the magnetic field increases further, the electrons are forced to occupy the inner shells. The high-field isomers 1–5 and 2–8 possess the same shape and symmetry as their stable classical analogues (cf. Fig. 1). In Fig. 3, ΔE is the difference between the ground-state energy of the considered isomer and the high-field 2–8 isomer, the classical counterpart of which is the most stable. It is interesting to note that, below point (a), the MDD phase is reproduced in the calculations with multicenter basis (6) using either of the configurations of Gaussian centers, i.e. 0–10, 1–9, and 2–8. At point (b) the 1–9 isomer becomes the most stable and above point (c) we obtain the most stable 2–8 isomer.

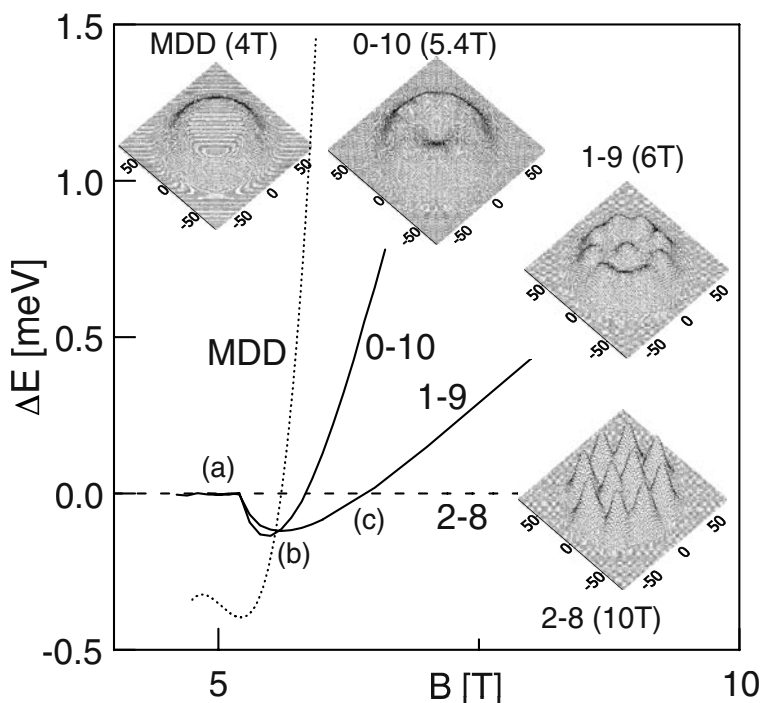


Figure 3. Energy separation ΔE (solid curves) between the ground-state energy of the 0-10 and 1-9 isomers of the Wigner molecule and that of the 2-8 isomer (dashed line). The dotted curve shows this energy difference for the MDD. Inset: the electron spatial distribution. Length is expressed in nanometers. (From Ref. [12]).

The results for $N = 2, \dots, 20$ are summarized in a phase diagram (Fig. 4). According to Fig. 4, the Wigner molecule, which consists of $N \geq 6$ electrons, can appear in at least two phases. In the different magnetic-field regimes, the different isomers are the most stable. For each N the symmetry of the high-magnetic-field phase is the same as that of the stable classical Wigner molecule.

The formation of Wigner molecules and their quasi-classical properties can be demonstrated if we consider the expectation value of the electron-electron interaction energy in the N -electron system (Fig. 5) [12]. In the MDD regime, i.e. for $B \leq 5.1$ T, the increasing magnetic field causes the shrinking of the electron spatial distribution, which in turn leads to a rapid increase of the electron-electron repulsion. For $B = 5.1$ T the MDD phase breaks down and a molecular phase with island-like charge density is created, which leads to the rapid decrease of the repulsive interactions in the electron system. Therefore, we conclude that

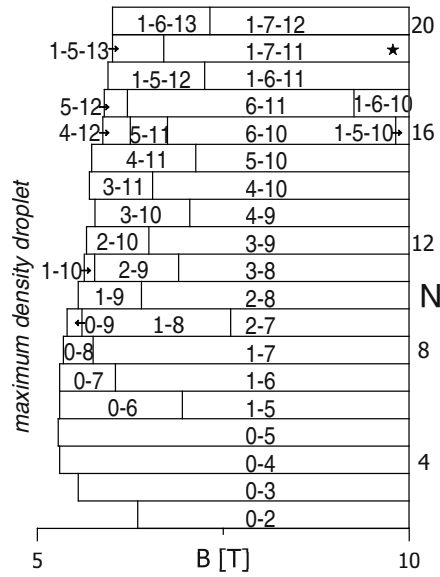


Figure 4. Phase diagram for N -electron Wigner molecules. The vertical lines correspond to magnetic fields B , for which the phase transitions are predicted. $\star$ For $N = 19$ the most stable isomer 1-6-12 is created at $B = 15.8$ T. (From Ref. [11]).

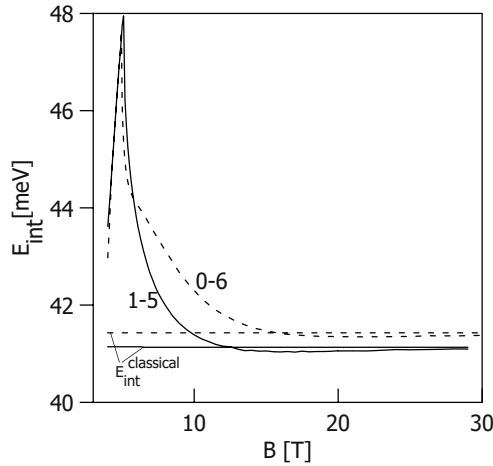


Figure 5. The expectation value E_{int} of the electron-electron interaction energy in the six-electron system as a function of magnetic field B . Solid (dashed) curve shows the results for the 1-5 (0-6) isomer. The horizontal lines display the interaction energy $E_{\text{int}}^{\text{classical}}$ for the classical counterparts of both the isomers. (From Ref. [12]).

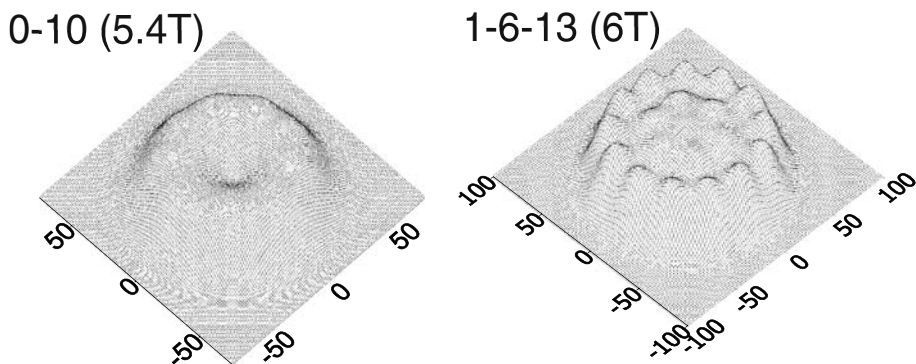


Figure 6. Electron spatial distribution on the $x - y$ plane obtained just after the MDD decay for isomers 0-10 and 1-6-13 of the 10- and 20-electron Wigner molecules. The left (right) panel shows the central hole formation (edge reconstruction). Length is expressed in nanometers. (From Ref. [12]).

the MDD-Wigner molecule phase transition is associated with a cusp of the interaction energy. Moreover, in the high magnetic-field limit, the quantum-mechanical expectation value of the electron-electron interaction energy asymptotically approaches the classical interaction energy for the considered configuration. These asymptotics are a signature of quasi-classical behavior of Wigner molecules at high magnetic fields.

In Refs. [18, 19, 27], two different mechanisms of MDD decay were proposed. Reimann *et al.* [18, 19] suggested that MDD breaks down via an edge reconstruction, while Yang and MacDonald [27] interpreted their results in terms of hole formation in the occupation number distribution. An illustration of both the mechanisms is shown in Fig. 6. The left panel shows hole formation in the 10-electron density distribution, which is followed by the formation of the molecular isomer 0-10, and the right panel shows the edge reconstruction in the 20-electron system. The results of Ref. [12] suggest that, depending on N , we can deal with either of these mechanisms. For $N \leq 12$ a central hole is formed in the electron density distribution during the MDD decay, while for $N > 12$, the electrons at the edge of the density distribution first become more strongly localized and then a molecular-type spatial distribution is formed for all the electrons.

The Wigner molecules considered so far were formed as the result of the joint effect of the external magnetic field and electron-electron repulsion – quantum confinement played a secondary role. Nevertheless, Wigner molecules can be created in quantum dots in the absence of a

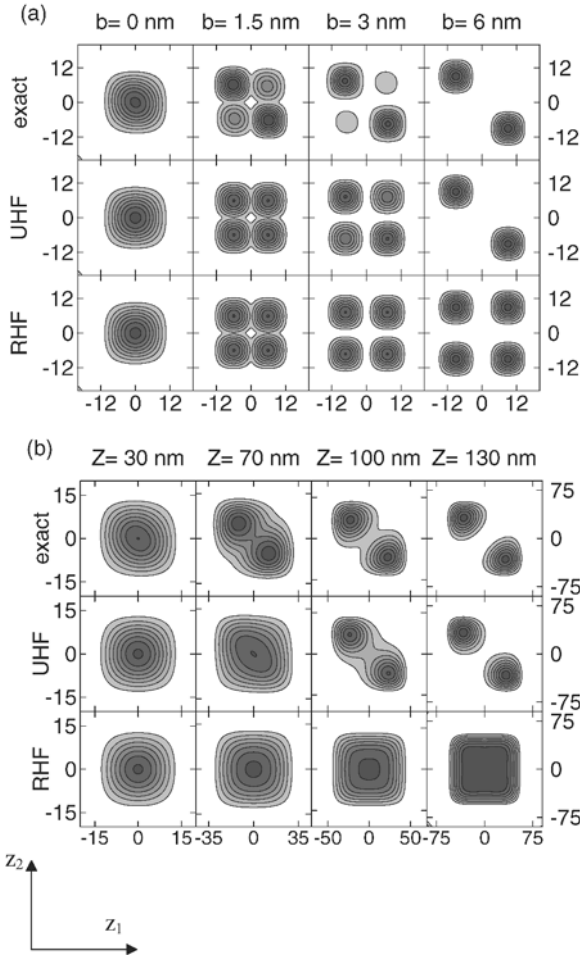


Figure 7. Contours of the two-electron density distribution as functions of vertical coordinates z_1 and z_2 of the electrons in (a) the two quantum dots separated by the barrier of thickness b and (b) the single quantum dot of vertical size Z . The darker the shade of gray, the larger the electron density. Length is expressed in nanometers. (From Ref. [22]).

magnetic field. Then, the quantum confinement plays an important role [22]. In Ref. [22] a one-dimensional (1D) model has been proposed to describe the electron states in quantum wires and vertically coupled quantum dots. Two-electron systems were studied [22] by an exact numerical method and by unrestricted (UHF) and restricted (RHF) Hartree-Fock methods. The electrons confined in the coupled quantum dots form states called *artificial molecules*. Figure 7 shows the two-electron elec-

tron density distribution in a double coupled quantum dot with barrier thickness b [Fig. 7(a)] and in a single quantum dot with linear size Z [Fig. 7(b)]. Figure 7(a) shows the transition from the artificial atom (for $b=0$) to the artificial molecule ($b \neq 0$). For $b=6$ nm the electrons are localized within the different quantum dots. Figure 7(b) demonstrates the size-induced formation of the Wigner molecule. In the small quantum dot ($Z=30$ nm), the electron density distribution corresponds to that of the MDD phase. The electrons become spatially separated, i.e. they form the Wigner molecule, if the quantum dot size increases, i.e. the confinement becomes weaker. In a large enough quantum dot ($Z=130$ nm), we observe the clear separation of the electron spatial distribution. It is interesting that the electron density distributions in the artificial molecule and Wigner molecule exhibit a remarkable similarity [cf. the top right panels in Figs. 7(a,b)]. Figure 7 also shows that the UHF method very well reproduces the exact results, which is not the case of the RHF approach, which fails to reproduce the molecular-type of electron localization. This confirms the reliability of the results [11, 12], obtained by the UHF method, in the Wigner-molecule regime.

5. Discussion

The breakdown of MDD was reported by Oosterkamp *et al.* [20] in vertical gated quantum dots. The authors measured single-electron transport through the quantum dot and observed kinks on the plots of the differential conductance versus the magnetic field. Since the conditions of the single-electron tunnelling [28] are determined by the chemical potential of the N -electron confined system, each kink of the ground-state energy, which is related with the phase transition (cf. Figs. 2 and 3), causes the corresponding kink of the chemical potential, i.e. the kink on the single-electron transport plot. The kinks [15, 20] at low magnetic fields correspond to transitions between the different spin orbitals [16], while the kinks at higher magnetic fields were interpreted [20] as corresponding to the decay of the MDD. Additional kinks observed [20] at still higher magnetic fields can be interpreted in terms of the phase transitions between the different Wigner-molecule isomers. A direct observation of the electron spatial distribution requires another experimental technique, e.g. wavefunction imaging [29]. We note that, due to the use of the 2D model with parabolic confinement, the results of calculations [11, 12] possess mainly a qualitative character. Therefore, the relation with experimental data [20] also has only qualitative meaning.

Based on the results in Refs. [11, 12] we can speculate about possible properties of Wigner molecules in three-dimensional (3D) nanostruc-

tures. In this case, the effective confinement of the electrons is weaker and the electron-electron Coulomb interaction is also weaker than in the 2D quantum dots. Therefore, we can predict that possible phase transitions between the different isomers of the 3D Wigner molecule will also occur, but at higher magnetic fields.

In the Wigner-molecule magnetic-field regime, the UHF method applied to the electron system confined in a potential with circular symmetry yields broken symmetry solutions [11, 12], which however are degenerate with respect to rotation by an arbitrary angle. The exact ground state of the Wigner molecule can be constructed as the superposition of the Hartree-Fock states, obtained in Refs. [11, 12], which restores the circular symmetry of the system [30]. The island-like electron density distribution in the Wigner molecular states should then be understood in terms of the relative electron-electron positions.

In vertical quantum dots [20], a slight anisotropy is inevitable, e.g. due to the roughness of the surface. The rotational degeneracy is then lifted and the Wigner molecule can be pinned under fixed orientation. We note that the pinning of the Wigner molecule under a fixed angle is a favorable condition for the direct observation of the electron density distribution by the wavefunction imaging [29].

The prediction of the different phases of the Wigner molecules is based on the results of the UHF calculations [11, 12]. Contrary to the RHF method, the UHF method takes into account a large amount of the electron-electron correlation (cf. Fig. 7) [22]. Moreover, the UHF results [11, 12] exactly reproduce the classical high-field behavior of the Wigner molecules. For $N=6$ two different isomers of the Wigner molecule have been obtained by the exact configuration interaction method [21]. However, some small corrections of the UHF results [11, 12] are possible for lower magnetic fields. Therefore, the UHF results [11, 12] can be treated as sufficiently reliable for the purposes of the discussion of the qualitative properties of the Wigner molecules. According to these results [11, 12], several different phases can be formed in the spin-polarized N -electron quantum-dot confined system in different magnetic-field regimes. These phases differ between themselves in the electron density distribution. The ground-state energy exhibits kinks as a function of the magnetic field, which leads to jumps of its first derivative. Therefore, we deal with the phase transitions in the N -electron system. At critical magnetic fields, which correspond to the phase transitions, the degeneracy of different molecular ground states can occur. Just after the MDD breakdown, the energy separations between the different phases of the Wigner molecule are rather small and grow with increasing B , reaching the classical values at extremely high magnetic fields [12].

6. Summary

Theoretical studies [4, 5, 10, 11, 12] show that the magnetic field imposed on the N -electron quantum-dot confined system leads to the breakdown of MDD and the formation of a Wigner molecule. For $N \geq 6$ the Wigner-molecule isomer, created just after the decay of the MDD, differs from the high-field quasi-classical isomer. The various isomers of the N -electron Wigner molecule possess the different ground-state energies and the different electron spatial distribution. We note that, for the cylindrical quantum dots, the exact N -electron wavefunction is rotationally symmetric for each Wigner-molecule phase. The broken symmetry solutions, obtained by the Hartree-Fock method [11, 12], should be interpreted in terms of the inner (relative) electron coordinates, i.e. in terms of the pair correlation function. These solutions reproduce the symmetry of the pair correlation function, obtained by the exact calculations [21]. The kinks, observed in transport spectroscopy with vertical quantum dots [20], can be interpreted as resulting from the decay of the MDD with the creation of the Wigner molecule and the subsequent transitions between the different molecular phases. The Wigner molecules can also be created in the absence of the magnetic field if the size of the quantum dot is sufficiently large [22]. Finally, we note that the Wigner molecule is the electron system for which we can study the continuous transition from the quantum to classical behavior.

Acknowledgments

This work has been supported in part by the Polish Government Scientific Research Committee (KBN).

References

- [1] E. P. Wigner. On the interaction of electrons in metals (1934). *Phys. Rev. B* **46**: 1002–1011, 1934.
- [2] W. Häusler, B. Kramer, and J. Masek. The influence of Coulomb interaction on transport through mesoscopic two-barrier structures. *Z. Phys.* **85**: 435–442, 1991.
- [3] K. Jauregui, W. Häusler, and B. Kramer. Wigner molecules in nanostructures. *Europhys. Lett.* **24**: 581–587, 1993.
- [4] H.-M. Müller and S. E. Koonin. Phase transitions in quantum dots. *Phys. Rev. B* **54**: 14532–14539, 1996.
- [5] P. A. Maksym, H. Imamura, G. P. Mallon, and H. Aoki. Molecular aspects of electron correlation in quantum dots. *J. Phys.: Condens. Matter* **12**: R299–R334, 2000.

- [6] H. Imamura, P. A. Maksym, and H. Aoki. Symmetry of 'molecular configurations' of interacting electrons in a quantum dot in strong magnetic fields. *Physica B* **249-251**: 214–219, 1998.
- [7] R. Egger, W. Häusler, C. H. Mak, and H. Grabert. Crossover from Fermi liquid to Wigner molecule behavior in quantum dots. *Phys. Rev. Lett.* **82**: 3320–3323, 1999.
- [8] B. Reusch, W. Häusler, and H. Grabert. Wigner molecules in quantum dots. *Phys. Rev. B* **63**: art. no. 113313, 2001.
- [9] A. V. Filinov, M. Bonitz, and Yu. E. Lozovik. Wigner crystallization in mesoscopic 2D electron systems. *Phys. Rev. Lett.* **86**: 3851–3854, 2001.
- [10] S. M. Reimann and M. Manninen. Electronic structure of quantum dots. *Rev. Mod. Phys.* **74**: 1283–1342, 2002.
- [11] B. Szafran, S. Bednarek, and J. Adamowski. Magnetic-field-induced transformations of Wigner molecule symmetry in quantum dots. *Phys. Rev. B* **67**: art. no. 045311 (erratum 159902(E)), 2003.
- [12] B. Szafran, S. Bednarek, and J. Adamowski. Magnetic-field-induced phase transitions in Wigner molecules. *J. Phys.: Condens. Matter* **15**: 4189–4205, 2003.
- [13] G. Meissner, H. Namaizawa, and M. Voss. Stability and image-potential-induced screening of electron vibrational excitations in a three-layer structure. *Phys. Rev. B* **13**: 1370–1376, 1976.
- [14] L. Bonsall and A. A. Maradudin. Some static and dynamical properties of a two-dimensional Wigner crystal. *Phys. Rev. B* **15**: 1959–1973, 1977.
- [15] L. P. Kouwenhoven, T. H. Oosterkamp, M. W. S. Danoesastro, M. Eto, D. G. Austing, T. Honda, and S. Tarucha. Excitation spectra of circular, few-electron quantum dots. *Science* **278**: 1788–1792, 1997.
- [16] B. Szafran, S. Bednarek, and J. Adamowski. Electric- and magnetic-field-induced evolution of transport windows in a vertical quantum dot. *Phys. Rev. B* **65**: art.no. 035316, 2002.
- [17] A. H. MacDonald, S.-R. Eric Yang, and M. D. Johnson. Quantum dots in strong magnetic fields: stability criteria for the maximum density droplet. *Aust. J. Phys.* **46**: 345–358, 1993.
- [18] S. M. Reimann, M. Koskinen, M. Manninen, and B. R. Mottelson. Quantum dots in magnetic fields: phase diagram and broken symmetry at the maximum-density-droplet edge. *Phys. Rev. Lett.* **83**: 3270–3273, 1999.
- [19] S. M. Reimann, M. Koskinen, and M. Manninen. Formation of Wigner molecules in small quantum dots. *Phys. Rev. B* **62**: 8108–8113, 2000.
- [20] T. H. Oosterkamp, J. W. Janssen, L. P. Kouwenhoven, D. G. Austing, T. Honda, and S. Tarucha. Maximum-density droplet and charge redistribution in quantum dots at high magnetic fields. *Phys. Rev. Lett.* **82**: 2931–2934, 1999.
- [21] M. Manninen, M. Koskinen, S. M. Reimann, and B. Mottelson. Magnetic properties of quantum dots and rings. *Eur. Phys. J. D* **16**: 381–386, 2001.
- [22] S. Bednarek, T. Chwiej, J. Adamowski, and B. Szafran. Artificial molecules in coupled and single quantum dots. *Phys. Rev. B* **67**: art. no. 205316, 2003.
- [23] V. Bedanov and F. M. Peeters. Ordering and phase transitions of charged particles in a classical two-dimensional system. *Phys. Rev. B* **49**: 2667–2676, 1994.

- [24] F. Bolton and U. Rössler. Classical model of a Wigner crystal in a quantum dot. *Superlatt. Microstr.* **13**: 139–145, 1993.
- [25] V. A. Schweigert and F. M. Peeters. Spectral properties of classical two-dimensional clusters. *Phys. Rev. B* **51**: 7700–7713, 1995.
- [26] M. Saint Jean, C. Even, and C. Guthmann. Macroscopic 2D Wigner islands. *Europhys. Lett.* **55**: 45–51, 2001.
- [27] S.-R. Eric Yang and A. H. MacDonald. Coupling between edge and bulk in strong-field quantum dots. *Phys. Rev. B* **66**: art. no. 041304, 2002.
- [28] S. Bednarek, B. Szafran, and J. Adamowski. Theoretical description of electronic properties of vertical gated quantum dots. *Phys. Rev. B* **64**: art. no. 195303, 2001.
- [29] E. E. Vdovin, A. Levin, A. Patanè, L. Eaves, P. C. Main, Yu. N. Khanin, Yu. Y. Dubrovkii, M. Henini, and G. Hill. Imaging the electron wavefunction in self-assembled quantum dots. *Science* **290**: 122–124, 2000.
- [30] C. Yannouleas and U. Landman. Trial wavefunctions with long-range Coulomb correlations for two-dimensional N-electron systems in high magnetic fields. *Phys. Rev. B* **66**: art. no. 115315, 2002.

FAST CONTROL OF QUANTUM STATES IN QUANTUM DOTS: LIMITS DUE TO DECOHERENCE

Lucjan Jacak, Paweł Machnikowski

Institute of Physics, Wrocław University of Technology, 50-370 Wrocław, Poland

Jurij Krasnyj

*Institute of Mathematics, University of Opole, 45-051 Opole, Poland**

Abstract We study the kinetics of confined carrier-phonon system in a quantum dot under fast optical driving and discuss the resulting limitations to fast coherent control over the quantum state in such systems.

1. Introduction

Unlike natural atoms, semiconductor quantum dots (QDs) always form part of a macroscopic crystal. The interaction with the quasi-continuum of lattice degrees of freedom (phonons) constitutes an inherent feature of these nanometer-size systems and cannot be neglected in any realistic modeling of QD properties, especially when the coherence of confined carriers is of importance. The understanding of the decisive role played by the QD coupling to lattice modes has increased recently due to both experimental and theoretical study (spectrum reconstruction [1, 2, 3, 4], relaxation [5, 6, 7, 8, 9, 10], phonon replicas and phonon-assisted transitions [11, 12, 13, 14, 15, 16], phonon-induced pure dephasing upon ultrafast excitation [17, 18, 19, 20, 21, 22]). The phonon-induced decoherence seems to be crucial for any quantum information processing application and for any nanotechnological device relying on quantum coherence of confined carriers [23, 24].

There are three major mechanisms of carrier-phonon interaction [25]: (1) Coulomb interaction with the lattice polarization induced by the relative shift of the positive and negative sub-lattices of the polar compound,

*On leave from Odessa University, Ukraine

described upon quantization by longitudinal optical (LO) phonons; (2) deformation potential coupling describing the band shifts due to lattice compression, i.e. longitudinal acoustical (LA) phonons; (3) Coulomb interaction with piezoelectric field generated by shear crystal deformation (transversal acoustical, TA, phonons). The lattermost effect is weak in InAs/GaAs systems but may be of more importance for the properties, e.g. of GaN dots [26, 27, 28].

Due to the carrier-phonon interaction, any change in the carrier subsystem must be accompanied by the appropriate modification of the lattice state. An example of this effect may be the creation of the exciton-polaron states: excitons accompanied by lattice polarization (coherent LO phonon field) with energy shifted down by a few meV [3, 22]. The interaction with acoustical branches leads to the creation of a similar deformation field which, although much less pronounced in terms of its energy shift, strongly influences the dynamical properties of the interacting system due to the gapless spectrum and low characteristic frequencies of the acoustical phonons.

The lattice response to any manipulation of the charge distribution confined in a QD leads to a discrepancy between the desired and actual states of the quantum confined system. This effect may be stronger than any decoherence process in an undriven system. In terms of the quantum information processing schemes, such a discrepancy is manifested by the loss of fidelity of the quantum operation. In quantum nano-technology applications, it will reduce the efficiency of nano-devices.

In the present paper we discuss the carrier-lattice kinetics induced by optical excitation of a confined exciton. The paper is organized as follows: In the next section we present the formal model of the exciton confined in a QD. Section 3 describes the system kinetics in response to an abrupt change of the charge distribution. In Sec. 4 we discuss the trade-off between the dynamically induced error and other processes limiting the coherence time of a quantum state. The final section contains concluding discussion.

2. The Model

We consider the Hamiltonian describing a single exciton interacting with phonons,

$$H = \sum_n \epsilon_n a_n^\dagger a_n + \sum_{\mathbf{k}, s} \omega_{s, \mathbf{k}} b_{s, \mathbf{k}}^\dagger b_{s, \mathbf{k}} + \frac{1}{\sqrt{N}} \sum_{\mathbf{k}, n, n', s} F_{n, n'}^{(s)}(\mathbf{k}) a_n^\dagger a_{n'} \left(b_{s, \mathbf{k}} + b_{s, -\mathbf{k}}^\dagger \right),$$

Electron mass	m_e	0.067 m_0
Hole mass	m_h	0.38 m_0
Static dielectric constant	ε_s	13.2
Effective dielectric constant	$\tilde{\varepsilon}$	62.6
Optical phonon energy	Ω_0	36 meV
Longitudinal sound speed	c	5150 m/s
Deformation potential for electrons	σ_e	6.7 eV
Deformation potential for holes	σ_h	-2.7 eV
Unit crystal cell volume	v	0.044 nm ³
LO phonon dispersion parameter	β	0.03 meV nm ²

Table 1. The material parameters used in the calculations (partly after Refs. [29, 30]).

where $b_{s,\mathbf{k}}, b_{s,\mathbf{k}}^\dagger$ are bosonic annihilation and creation operators for LO ($s=o$) and LA ($s=a$) phonons with quasi-momentum $\mathbf{k}$. The corresponding frequencies within the effectively coupled wavevector range may be modeled by $\omega_{o,\mathbf{k}} = \Omega_{\mathbf{k}} \simeq \Omega - \beta k^2$ and $\omega_{a,\mathbf{k}} = ck$. The coupling constants for the LO and LA phonon branches are given by

$$F_{n,n'}^{(o)}(\mathbf{k}) = -\frac{e}{k} \sqrt{\frac{2\pi\Omega}{v\tilde{\varepsilon}}} \int \Phi_n^*(\mathbf{r}_e, \mathbf{r}_h) \left(e^{i\mathbf{k}\cdot\mathbf{r}_e} - e^{i\mathbf{k}\cdot\mathbf{r}_h} \right) \Phi_{n'}(\mathbf{r}_e, \mathbf{r}_h) d^3\mathbf{r}_e d^3\mathbf{r}_h$$

and

$$F_{n,n'}^{(a)}(\mathbf{k}) = -\sqrt{\frac{k}{2\rho vc}} \int \Phi_n^*(\mathbf{r}_e, \mathbf{r}_h) \left(\sigma_e e^{i\mathbf{k}\cdot\mathbf{r}_e} - \sigma_h e^{i\mathbf{k}\cdot\mathbf{r}_h} \right) \Phi_{n'}(\mathbf{r}_e, \mathbf{r}_h) d^3\mathbf{r}_e d^3\mathbf{r}_h, \quad (1)$$

where $\mathbf{r}_e, \mathbf{r}_h$ denote the coordinates of the electron and hole, respectively, and $\Phi_n(\mathbf{r}_e, \mathbf{r}_h)$ is the exciton wavefunction. The other elements of the notation are described in Table 1, along with values (corresponding to InAs/GaAs system) used in the calculations.

Numerical diagonalization of the interacting electron-hole system in parabolic confinement, under the assumption that non-interacting electrons and holes would have the same wavefunctions, leads to the spectrum shown in Fig. 1(a) (M is the conserved total angular momentum). The dominant contribution to the lowest excited states of the exciton comes from the excited hole states, while the electron wavefunction is only slightly modified. The corresponding electron and hole distributions are shown in Fig. 1(b-d).

The exciton-phonon coupling functions may be calculated using the wavefunctions found numerically. Figure 2 (left panels) shows the results

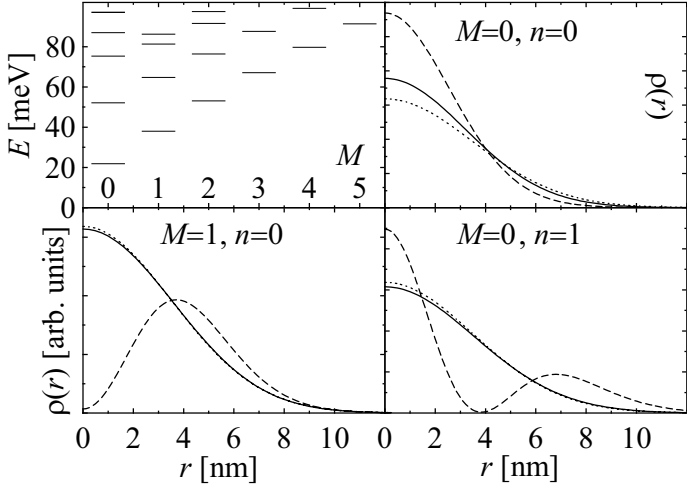


Figure 1. The spectrum of the exciton in a QD (upper left) and the electron (solid lines) and hole (dashed lines) probability densities for the three lowest exciton states, compared to the ground state probability density for a noninteracting particle (dotted lines). Here the noninteracting particle wavefunction width is $l_{\perp} = 4.9$ nm in-plane and $l_z = 2$ nm in the growth direction.

for coupling between the ground state and a few lowest states averaged over angles. The coupling to LO phonons is much stronger than to LA phonons. In the case of LO phonons, the coupling strength increases for excited states due to less charge cancellation between the electron and hole in these states [cf. Fig. 1(b-d)].

3. System Kinetics After An Ultrafast Pulse

Let us study the kinetics of the system after an excitation performed by an extremely short (formally infinitely short) laser pulse. The information on the dynamics of the interacting carrier-lattice system is contained in the exciton single-particle causal Green function,

$$G_{n,n'}(t) = -i\langle T\{a_n(t)a_{n'}^{\dagger}(0)\} \rangle.$$

The averaging $\langle \dots \rangle$ is temperature-dependent with respect to the phonon degrees of freedom and the vacuum of the exciton (cf. Ref. [31]; it corresponds to the case when the grand canonical averaging sector without the exciton, vacuum, is energetically distant (~ 1 eV) from the next sectors). For $t \geq 0$ this Green function coincides (up to a constant factor) with the correlation function $\langle a_n(t)a_{n'}^{\dagger}(0) \rangle$ which, for $n_1 = n_2$ may

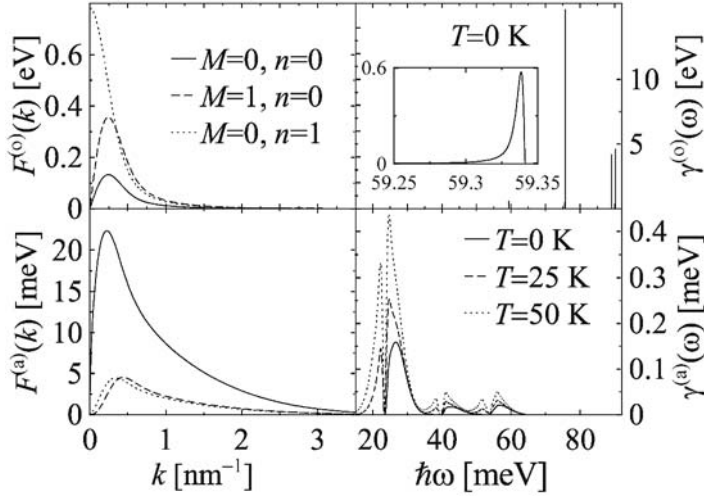


Figure 2. Coupling constants for LA (a) and LO (b) phonon modes between the ground state and the three lowest states and the resulting structure of the imaginary part of the mass operator (c,d).

be interpreted as a measure of integrity of the excitonic state: it corresponds to the overlap of the state at time t with this state at the initial time $t=0$. The Fourier transform of the correlation function, $I_{n,n'}(\omega) = \int_{-\infty}^{\infty} \langle a_n(t) a_{n'}^\dagger(0) \rangle e^{i\omega t} dt$, is usually called the spectral density [32, 33], and it can be expressed by the imaginary part of the causal Green function:

$$\text{Im} G_{n,n'}(\omega) = -\frac{1}{2} \left(1 + e^{-\omega/k_B T} \right)^{-1} I_{n,n'}(\omega), \quad (2)$$

where $G_{n,n'}(\omega) = \int_{-\infty}^{\infty} G_{n,n'}(t) e^{i\omega t} dt$.

The equation of motion for the causal Green function may be rewritten as a Dyson-type equation,

$$(\omega - \epsilon_n) G_{n,n'}(\omega) - \sum_{n_3} M_{n,n''}(\omega) G_{n'',n'}(\omega) = \delta_{n,n'}, \quad (3)$$

with the mass operator

$$M_{n,n'}(\omega) = \frac{i}{2\pi N} \sum_{\mathbf{k}, s, n'', n'''} F_{n,n''}^{(s)}(\mathbf{k}) F_{n''',n'}^{(s)}(-\mathbf{k}) \\ \times \int d\omega' G_{n'',n'''}(\omega + \omega') D^0(\mathbf{k}, s, \omega'),$$

where we have restricted ourselves to single-phonon processes by replacing the vertex function by its zeroth order approximation and using the free phonon Green function (LO phonon damping may be easily included later: see below). Since $F_{n,n'}^{(s)}(\mathbf{k}) \sim g_s$, then $|M_{n,n'}(\omega)|^2 \sim g_s^2$. Hence, up to g_s^2 one has from Eq. (3),

$$G_{n,n}(\omega) = \frac{1}{\omega - \epsilon_n - M_{n,n}(\omega)} \quad (4)$$

and

$$\begin{aligned} M_{nn}(\omega) &= \Delta_n(\omega) - i\gamma(\omega) \\ &= \frac{i}{2\pi N} \sum_{n', \mathbf{k}} |F_{nn'}^{(s)}(\mathbf{k})|^2 \int d\omega' G_{n'n'}(\omega + \omega') D^{(0)}(s, \mathbf{k}, \omega'). \end{aligned} \quad (5)$$

For the real and the imaginary parts of the mass operator we obtain from (5)

$$\begin{aligned} \Delta_n(\omega) &= \frac{1}{N} \sum_{\mathbf{k}, s, n} |F_{nn'}^{(s)}(\mathbf{k})|^2 \\ &\times \left[\frac{(1 + n_{s,\mathbf{k}})(\omega - \epsilon_{n'} - \Delta_{n'}(\omega - \omega_{s,\mathbf{k}}) - \omega_{s,\mathbf{k}})}{[\omega - \epsilon_{n'} - \Delta_{n'}(\omega - \omega_{s,\mathbf{k}}) - \omega_{s,\mathbf{k}}]^2 + \gamma_{n'}^2(\omega - \omega_{s,\mathbf{k}})} \right. \\ &\left. + \frac{n_{s,\mathbf{k}}(\omega - \epsilon_{n'} - \Delta_{n'}(\omega + \omega_{s,\mathbf{k}}) + \omega_{s,\mathbf{k}})}{[\omega - \epsilon_{n'} - \Delta_{n'}(\omega + \omega_{s,\mathbf{k}}) + \omega_{s,\mathbf{k}}]^2 + \gamma_{n'}^2(\omega + \omega_{s,\mathbf{k}})} \right] \end{aligned} \quad (6)$$

and

$$\begin{aligned} \gamma_n(\omega) &= \frac{1}{N} \sum_{\mathbf{k}, s, n'} |F_{nn'}^{(s)}(\mathbf{k})|^2 \\ &\times \left[\frac{(1 + n_{s,\mathbf{k}})\gamma_{n'}(\omega - \omega_{s,\mathbf{k}})}{[\omega - \epsilon_{n'} - \Delta_{n'}(\omega - \omega_{s,\mathbf{k}}) - \omega_{s,\mathbf{k}}]^2 + \gamma_{n'}^2(\omega - \omega_{s,\mathbf{k}})} \right. \\ &\left. + \frac{n_{s,\mathbf{k}}\gamma_{n'}(\omega + \omega_{s,\mathbf{k}})}{[\omega - \epsilon_{n'} - \Delta_{n'}(\omega + \omega_{s,\mathbf{k}}) + \omega_{s,\mathbf{k}}]^2 + \gamma_{n'}^2(\omega + \omega_{s,\mathbf{k}})} \right]. \end{aligned} \quad (7)$$

In this equation the first term determines the energy shift due to dressing with LO phonons, while the second one corresponds to LA phonons. The former dominates energetically: the energy shift is mainly due to the interaction of exciton with optical phonons (dressing of exciton with LO phonons, creating together an exciton-polaron). The latter is small and its energetic effect may be safely neglected but it contributes considerably to the system kinetics.

To the lowest order, one may neglect $\Delta_{n'}$ and $\gamma_{n'}$ in the right-hand-side of (6) and, neglecting the acoustic phonon term, write the equation for the exciton-polaron energy levels E_n as the poles of the Green function (4), $E_n - \epsilon_n - \Delta_n(E_n) = 0$, i.e.

$$E_n - \epsilon_n - \frac{1}{N} \sum_{\mathbf{k}, n} |F_{nn}^{(o)}(\mathbf{k})|^2 \left[\frac{1 + n_{o,\mathbf{k}}}{E_n - \epsilon_n - \Omega} + \frac{n_{o,\mathbf{k}}}{E_n - \epsilon_n + \Omega} \right] = 0. \quad (8)$$

At $k_B T \ll \Omega$, the above equation is equivalent to that found by Davydov diagonalization of the Fröhlich Hamiltonian for an exciton [34]. The Eq. (8) for $n=0$ gives the ground state energy shift $\Delta_0 \sim -5$ meV for the QD with parameters listed in Table 1 (see Ref. [22]).

The imaginary part of the mass operator is given to leading order by the equation (putting $\gamma_{n'} = 0$ in the right-hand side of Eq. (7)):

$$\begin{aligned} \gamma_n(\omega) = & \quad (9) \\ & \frac{\pi}{N} \sum_{\mathbf{k}, n} \left\{ |F_{nn'}^{(o)}(\mathbf{k})|^2 [(1 + n_{o,\mathbf{k}})\delta(\omega - E_{n'} - \Omega_{\mathbf{k}}) + n_{o,\mathbf{k}}\delta(\omega - E_{n'} + \Omega_{\mathbf{k}})] \right. \\ & \left. + |F_{nn'}^{(a)}(\mathbf{k})|^2 [(1 + n_{a,\mathbf{k}})\delta(\omega - E_{n'} - ck) + n_{a,\mathbf{k}}\delta(\omega - E_{n'} + ck)] \right\}, \end{aligned}$$

where we use the fact that the equation $\omega - \epsilon_{n'} - \Delta_{n'}(\omega \pm \omega_{s,\mathbf{k}}) \pm \omega_{s,\mathbf{k}} = 0$ is solved by $\omega = E_{n'} \mp \omega_{s,\mathbf{k}}$ and neglected higher-order corrections resulting from resolving the Dirac δ . The first term in Eq. (9) describes the energy transfer to the LO phonon sea, while the second one corresponds to the energy transfer from gradually dressing exciton to the LA phonon sea. The form of $\gamma_0(\omega)$ for our model, including a few lowest states, obtained using the numerical wavefunctions is shown for various temperatures in Fig. 2(c,d).

Using (4), one finds

$$\text{Im}G_{n,n}(\omega) = \frac{\gamma_n(\omega)}{[\omega - \epsilon_n - \Delta_n(\omega)]^2 + \gamma_n^2(\omega)}. \quad (10)$$

In the following, we will focus on the ground exciton state, $n=0$. Usually, this state is long-living at low temperatures: the broadening of the corresponding spectral line, related to radiative lifetime and thermally activated processes, does not exceed 0.1 meV for $T < 100$ K [35, 17].

Let us note that the term $\gamma_0^2(\omega)$ in the denominator of (10) is important only near E_0 , where the other term vanishes (otherwise, it is a higher-order correction). Therefore, it may be approximated by a constant $\gamma_0 = \gamma_0(E_0)$. In a similar manner, the correction $\Delta_0(\omega)$ may be replaced by $\Delta_0(\omega) \approx \Delta_0(E_0) + (\omega - E_0)(d\Delta_0(\omega)/d\omega)_{\omega=E_0}$ and combined with ϵ_0 to give E_0 , in accordance with (8).

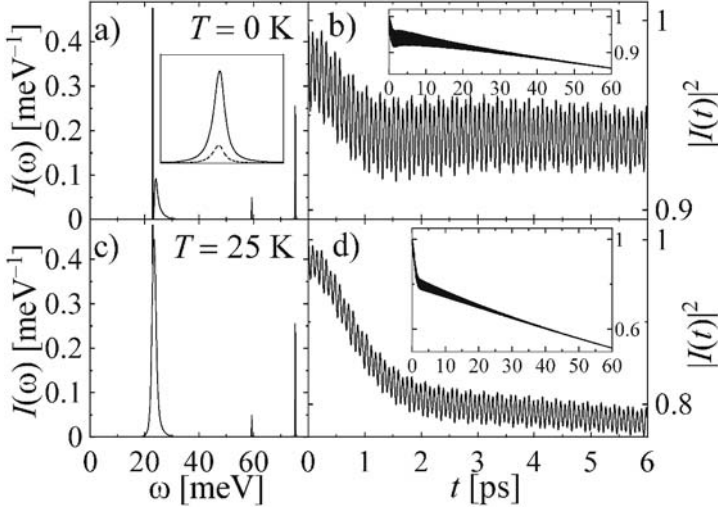


Figure 3. Left: Spectral density *versus* energy (inset shows the shape of the LO phonon replica) Right: The corresponding evolution of the correlation function for two temperatures, as shown (Insets show the evolution on longer timescales). The calculation is limited to the two lowest exciton states.

Thus, one may write for $I_0(\omega) = -(1/\pi)\text{Im}G_{00}(\omega)$

$$I_0(\omega) = \quad (11)$$

$$Z^{-1} \frac{1}{\pi} \frac{\gamma_0/2}{(\omega - E_0)^2 + \gamma_0^2/4}$$

$$+ \sum_{n,s,\mathbf{k}} |F_{0n}^{(s)}(\mathbf{k})|^2 \frac{(n_{s,\mathbf{k}} + 1)\delta(\omega - E_n - \omega_{s,\mathbf{k}}) + n_{s,\mathbf{k}}\delta(\omega - E_n + \omega_{s,\mathbf{k}})}{(\omega - E_0)^2 + \gamma_0^2},$$

where

$$Z = 1 - \frac{d\Delta(\omega)}{d\omega}|_{\omega=E_0} \approx 1 + \frac{1}{N} \sum_{\mathbf{k},s,n'} \left| \frac{F_{nn'}^{(s)}(\mathbf{k})}{E_{n'} - E_n + \omega_{s,\mathbf{k}}} \right|^2 [1 + 2n_{s,\mathbf{k}}].$$

Figure 3(a,c) shows the result.

At $T = 0$, $\gamma_0(\omega) = 0$ at $\omega = E_0$ (neglecting recombination process) and this point is a well-defined pole of the causal Green function. It corresponds to a quasiparticle: the dressed exciton (i.e. the exciton-polaron, if neglecting LA phonons). For $T > 0$, this spectral density has a Lorentzian shape around $\omega \sim E_0$, with width γ_0 . Our model is unable to quantitatively account for the broadening found in experiment. In

fact, to our knowledge, its origin has not been explained so far. We will use the experimental values of γ_0 corresponding to 630 ps and 170 ps exciton lifetimes (including radiative processes and thermally induced transitions) at $T = 0$ K and $T = 25$ K, respectively [17].

Apart from the central peak, the spectral density shows acoustic and optical phonon sidebands. To the former, only the ground state itself contributes: the magnitude of the features visible in $\gamma(\omega)$ around excited states [Fig. 2(c,d)] is much smaller than the energy distance between these states. Therefore, as far as the interaction with acoustical phonons is concerned, a single level (independent boson) model supplemented by central line broadening is very accurate.

The features resulting from the LO phonon branch behave in a different manner: the contribution from nonadiabatic coupling to excited states is considerable. This results from the fact that the coupling to LO phonons is much stronger in general, the energies of LO phonons are comparable to the exciton energy spacing and the coupling to excited levels is stronger than to the ground state itself due to partial charge cancellation in the latter. In this case, the contribution from the higher states dominates over the ground state contribution and the independent boson approach is not valid.

The time evolution of the dressing process is described by the inverse Fourier transform of the correlation function given by (2). Taking into account that for ω in the energy sector of exciton (of order of eV) $\omega \gg k_B T$, one has

$$I_n(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} d\omega \text{Im} G_{n,n}(\omega) e^{-i\omega t}.$$

The time-dependent correlation function may be obtained from (11),

$$\begin{aligned} I_0(t) = & Z^{-1} e^{-i(E_0 - i\gamma_0/2)t} \\ & + \frac{1}{N} \sum_{n,s,\mathbf{k}} |F_{0n}^{(s)}(\mathbf{k})|^2 \left[\frac{(n_{\mathbf{k}} + 1) e^{-i(E_n + \omega_{s,\mathbf{k}})t}}{(\Delta E_{n0} + \omega_{s,\mathbf{k}})^2 + \gamma_0^2} + \frac{n_{\mathbf{k}} e^{-i(E_n - \omega_{s,\mathbf{k}})t}}{(\Delta E_{n0} - \omega_{s,\mathbf{k}})^2 + \gamma_0^2} \right], \end{aligned}$$

where $\Delta E_{n0} = E_n - E_0$. The result, calculated numerically for various temperatures, is plotted in Fig. 3(b,d).

The smooth sidebands of $I(\omega)$ correspond to the initial correlation decay on 1 ps timescale. The LO phonon peaks lead to slowly decaying fast oscillations manifesting phonon beats with frequencies corresponding to the energy differences $E_n - E_0 - \Omega$. Without including anharmonic effects, these oscillations would decay very slowly due to the weak LO phonon dispersion. When the anharmonic phonon damping (corresponding to $\tau_{\text{LO}} = 9$ ps decay of an LO phonon [36, 37]) is included by

substituting $\omega_{o,\mathbf{k}} \rightarrow \omega_{o,\mathbf{k}} \pm i\gamma_{\text{LO}}$, the LO phonon beats decay with characteristic time slightly below τ_{LO} , which results from the joint effect of damping and dispersion. The long-time dynamics is governed by the Lorentzian feature around $\omega = E_0$ and shows an exponential decay (as assumed beforehand).

4. System Response for Finite-Duration Pulses

In the previous section we have shown that any change in the carrier distribution in a QD is followed by lattice relaxation (dressing) processes that lead to some loss of the quantum coherence. To preserve the coherence while operating on the carrier states, the action must be adiabatic with respect to lattice timescales, i.e. phonon periods. Slowing down the dynamics leads, however, to increasing effect of other decoherence mechanisms, like radiative recombination or thermally activated transitions to higher states, as the decoherence caused by such effects increases linearly with time.

In order to study the interplay of these decoherence effects we have studied the evolution of an exciton coupled to acoustic phonons under optical excitation by a finite-duration laser pulse [23]. The effect of coupling to phonons was quantified in terms of the error of a coherent operation, $\delta = 1 - F$, where the fidelity is defined as an overlap between the actual state and the desired one (the latter was taken to be the dressed final state, obtained by adiabatic operation), $F = \langle \phi | \rho | \phi \rangle$, where $|\phi\rangle$ is the ideal final state and ρ is the final density matrix.

The density matrix for the final state is calculated using second order expansion for the evolution operator of the total system and tracing over the lattice degrees of freedom (see Refs. [23, 38] for details). For timescales relevant here, it is sufficient to consider acoustic phonons and only one (ground) exciton state. The error averaged over initial states can be represented as the overlap of two functions

$$\delta = \int \frac{d\omega}{\omega^2} R(\omega) S(\omega),$$

where $R(\omega)$ is the spectral density of the reservoir

$$R(\omega) = \sum_{\mathbf{k}} [\delta(\omega_{\mathbf{k}} - \omega)(n_{\mathbf{k}} + 1) + \delta(\omega_{\mathbf{k}} + \omega)n_{\mathbf{k}}] |F(\mathbf{k})|^2.$$

The function $S(\omega)$ represents spectral characteristics of the system and is given by

$$S(\omega) = \frac{1}{3}(|F_-(\omega)|^2 + |F_+(\omega)|^2), \quad |F_{\pm}(\omega)|^2 \approx \alpha^2 e^{-\frac{1}{2}\tau_g^2 \left(\omega \pm \frac{\alpha}{\sqrt{2\pi}\tau_g}\right)^2}, \quad (12)$$

for a Gaussian driving pulse $f(t) = \alpha/(\sqrt{2\pi}\tau_g)e^{-\frac{1}{2}(t/\tau_g)^2}$, rotating the state by the angle α on the Bloch sphere.

For the coupling (1) to LA phonons, one finds at $T=0$ for $\omega \ll cl_\perp^{-1}$

$$R(\omega) \simeq R_0\omega^3, \quad R_0 = \frac{(\sigma_e - \sigma_h)^2}{16\pi^2\rho c^5}, \quad (13)$$

and we obtain the averaged error $\delta = \alpha^2 R_0 \tau_g^{-2}/3$ (taking into account the upper cut-off, the error will be finite even for an infinitely fast gate (Fig. 4)).

Hence, this non-Markovian error increases as the gate speed increases. This could result in obtaining an arbitrarily low error by choosing suitably low gate speeds. However, if our system is also subjected to other type of noise this becomes impossible. Indeed, if we assume that our system undergoes additional amplitude damping with rate γ_M , the total error per gate is

$$\delta = \frac{\gamma_{nM}}{\tau_g^2} + \gamma_M \tau_g, \quad \gamma_{nM} = \frac{1}{3}\alpha^2 R_0, \quad \gamma_M = \frac{1}{\tau_r}, \quad (14)$$

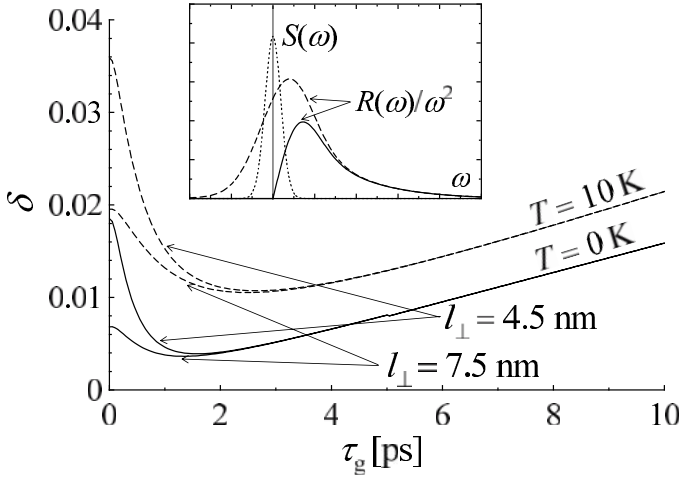


Figure 4. Combined Markovian and non-Markovian error for a $\alpha = \pi/2$ rotation on a qubit implemented as a confined exciton in a InAs/GaAs quantum dot, for $T=0$ (solid lines) and $T=10$ K (dashed lines), for two dot sizes (dot height is 20% of its diameter). The Markovian decoherence times are inferred from the experimental data [17]. Inset: Spectral density of the phonon reservoir $R(\omega)$ at these two temperatures and the gate profile $S(\omega)$ for $\alpha = \frac{1}{2}\pi$.

where τ_r is the characteristic time of exponential decay (Markovian error). As a result, the overall error is unavoidable and optimization is needed. The above formulas lead to

$$\delta_{\min} = \frac{3}{2}(2\gamma_M^2\gamma_{nM})^{1/3} = \frac{3}{2}\left(\frac{2\alpha^2 R_0}{3\tau_r^2}\right)^{1/3} = \alpha^{2/3}0.0035,$$

for

$$\tau_g = \left(2\frac{\gamma_{nM}}{\gamma_M}\right)^{1/3} = \left(\frac{2}{3}\alpha^2 R_0\tau_r\right)^{1/3} = \alpha^{2/3}1.5\text{ps},$$

where we have used GaAs material parameters (Table 1).

The exact solution within the proposed model, taking into account the cut-off and anisotropy of the dot shape and allowing finite temperatures, is shown in Fig. 4. The size-dependent cut-off is reflected by a shift of the optimal parameters for the two dot sizes: larger dots admit faster gates and lead to lower error. Interestingly, the trade-off becomes more apparent at nonzero temperature. It should be noted that these optimal times are longer than the limits imposed by level separation [39, 40, 41, 42, 43]. Thus, the non-Markovian reservoir effects (dressing) seem to be the essential limitation to the gate speed.

5. Conclusions

We have shown that an action performed on the carrier system (e.g. exciton) confined in a quantum dot must be accompanied by an appropriate reconstruction of the lattice state. Such a dressing process takes place spontaneously if the charge distribution is changed on times short compared to timescales of the lattice dynamics (phonon oscillation periods). This spontaneous process destroys the quantum coherence of the system and may preclude quantum information processing or coherent nanotechnology applications. On the other hand, sufficiently adiabatic operation would take long time compared to the exciton lifetime and other exponential decay processes, again leading to large coherence loss. Therefore, optimal operating conditions for maximum preservation of coherence must be searched for.

We have evaluated the minimal error for typical parameters, showing that the optimal gating time (~ 1 ps) lies within the current experimental possibility. This optimal time sets up the limit beyond which further gate speed-up is unfavorable. It is remarkable that minimal error δ we found (0.004 – 0.01), although not extremely high, is still of 1-2 orders of magnitude higher than the one admitted by fault-tolerant schemes known so far ($\sim 10^{-5}$). However, possible improvements of the latter schemes cannot be excluded.

Acknowledgments

This work was supported by the Polish KBN under Grants No. 2 P03B 024 24 and No. 2 P03B 085 25.

References

- [1] S. Hameau, Y. Guldner, O. Verzelen, R. Ferreira, and G. Bastard. Strong electron-phonon coupling regime in quantum dots: evidence for everlasting resonant polarons. *Phys. Rev. Lett.* **83**: 4152–4155, 1999.
- [2] S. Hameau, J. M. Isaia, Y. Guldner, E. Deleporte, O. Verzelen, R. Ferreira, G. Bastard, J. Zeman, and J. M. Gérard. Far-infrared magnetospectroscopy of polaron states in self-assembled InAs/GaAs quantum dots. *Phys. Rev. B* **65**: art. no. 085316, 2002.
- [3] O. Verzelen, R. Ferreira, and G. Bastard. Excitonic polarons in semiconductor quantum dots. *Phys. Rev. Lett.* **88**: art. no. 146803, 2002.
- [4] L. Jacak, J. Krasnyj, D. Jacak, and P. Machnikowski. Magneto-polaron in a weakly elliptical InAs/GaAs quantum dot. *Phys. Rev. B* **67**: art. no. 035303, 2003.
- [5] R. Heitz, M. Veit, N. N. Ledentsov, A. Hoffmann, D. Bimberg, P. S. Kop'ev, and Zh. I. Alferov. Energy relaxation by multiphonon processes in InAs/GaAs quantum dots. *Phys. Rev. B* **56**: 10435–10445, 1997.
- [6] R. Heitz, H. Born, F. Guffarth, O. Stier, A. Schliwa, A. Hoffmann, and D. Bimberg. Existence of a phonon bottleneck for excitons in quantum dots. *Phys. Rev. B* **64**: art. no. 241305, 2001.
- [7] I. V. Ignatiev, I. E. Kozin, V. G. Davydov, S. V. Nair, J.-S. Lee, H.-W. Ren, S. Sugou, and Y. Masumoto. Phonon resonances in photoluminescence spectra of self-assembled quantum dots in an electric field. *Phys. Rev. B* **63**: art. no. 075316, 2001.
- [8] O. Verzelen, R. Ferreira, and G. Bastard. Polaron lifetime and energy relaxation in semiconductor quantum dots. *Phys. Rev. B* **62**: R4809–R4812, 2000.
- [9] L. Jacak, J. Krasnyj, D. Jacak, and P. Machnikowski. Anharmonicity induced polaron relaxation in InAs/GaAs quantum dots. *Phys. Rev. B* **65**: art. no. 113305, 2002.
- [10] O. Verzelen, G. Bastard, and R. Ferreira. Energy relaxation in quantum dots. *Phys. Rev. B* **65**: art. no. 081308, 2002.
- [11] R. Heitz, I. Mukhametzhanov, O. Stier, A. Madhukar, and D. Bimberg. Enhanced polar exciton-phonon interaction in quantum dots. *Phys. Rev. Lett.* **83**: 4654–4657, 1999.
- [12] R. Heitz, I. Mukhametzhanov, O. Stier, A. Madhukar, and D. Bimberg. Phonon-assisted polar exciton transitions in self-organized InAs/GaAs quantum dots. *Physica E* **7**: 398–402, 2000.
- [13] A. Lemaître, A. D. Ashmore, J. J. Finley, D. J. Mowbray, M. S. Skolnick, M. Hopkinson, and T. F. Krauss. Enhanced phonon-assisted absorption in single InAs/GaAs quantum dots. *Phys. Rev. B* **63**: art. no. 161309, 2001.
- [14] F. Findeis, A. Zrenner, G. Böhm, and G. Abstreiter. Phonon-assisted biexciton generation in a single quantum dot. *Phys. Rev. B* **61**: R10579–R10582, 2000.

- [15] V. M. Fomin, V. N. Gladilin, S. N. Klimin, and J. T. Devreese. Enhanced phonon-assisted photoluminescence in InAs/GaAs parallelepiped quantum dots. *Phys. Rev. B* **61**: R2436–R2439, 2000.
- [16] L. Jacak, J. Krasnyj, and W. Jacak. Renormalization of the Fröhlich constant for electrons in a quantum dot. *Phys. Lett. A* **304**: 168–171, 2002.
- [17] P. Borri, W. Langbein, S. Schneider, U. Woggon, R. L. Sellin, D. Ouyang, and D. Bimberg. Ultralong dephasing time in InGaAs quantum dots. *Phys. Rev. Lett.* **87**: art. no. 157401, 2001.
- [18] P. Borri, W. Langbein, S. Schneider, U. Woggon, R. L. Sellin, D. Ouyang, and D. Bimberg. Rabi oscillations in the excitonic ground-state transition of InGaAs quantum dots. *Phys. Rev. B* **66**: art. no. 081306, 2002.
- [19] B. Krummheuer, V. M. Axt, and T. Kuhn. Theory of pure dephasing and the resulting absorption line shape in semiconductor quantum dots. *Phys. Rev. B* **65**: art. no. 195313, 2002.
- [20] A. Vagov, V. M. Axt, and T. Kuhn. Electron-phonon dynamics in optically excited quantum dots: Exact solution for multiple ultrashort laser pulses. *Phys. Rev. B* **66**: art. no. 165312, 2002.
- [21] A. Vagov, V. M. Axt, and T. Kuhn. Impact of pure dephasing on the nonlinear optical response of single quantum dots and dot ensembles. *Phys. Rev. B* **67**: art. no. 115338, 2003.
- [22] L. Jacak, P. Machnikowski, J. Krasnyj, and P. Zoller. Coherent and incoherent phonon processes in artificial atoms. *Eur. Phys. J. D* **22**: 319–331, 2003.
- [23] R. Alicki, M. Horodecki, P. Horodecki, R. Horodecki, L. Jacak, and P. Machnikowski. Optimal strategy of quantum computing and trade-off between opposite types of decoherence. quant-ph/0302058, 2003.
- [24] P. Machnikowski and L. Jacak. Limitation to coherent control in quantum dots due to lattice dynamics. cond-mat/0305165, 2002.
- [25] Gerald D. Mahan. *Many-Particle Physics*. (Kluwer, New York, 2000).
- [26] U. Hohenester, R. Di Felice, E. Molinari, and F. Rossi. Optical spectra of nitride quantum dots: Quantum confinement and electron-hole coupling. *Appl. Phys. Lett.* **75**: 3449–3451, 1999.
- [27] U. Hohenester, R. Di Felice, E. Molinari, and F. Rossi. Theoretical analysis of the optical spectra of $\text{In}_x\text{Ga}_{1-x}\text{N}$ quantum dots in $\text{In}_y\text{Ga}_{1-y}\text{N}$ layers. *Physica E* **7**: 934–938, 2000.
- [28] S. De Rinaldis, I. D’Amico, E. Biolatti, R. Rinaldi, R. Cingolani, and F. Rossi. Intrinsic exciton-exciton coupling in GaN-based quantum dots: Application to solid-state quantum computing. *Phys. Rev. B* **65**: art. no. 081309, 2002.
- [29] S. Adachi. GaAs, AlAs and $\text{Al}_x\text{Ga}_{1-x}\text{As}$: Material parameters for use in research and device applications. *J. Appl. Phys.* **5**: R1–R29, 1985.
- [30] D. Strauch and B. Dorner. Phonon dispersion in GaAs. *J. Phys: Condens. Matter* **2**: 1457–1483, 1990.
- [31] A. Suna. Green’s function approach to exciton–phonon interaction. *Phys. Rev.* **135**: A111–A123, 1964.
- [32] A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski. *Methods of Quantum Field Theory in Statistical Physics*. (Dover, New York, 1975).

- [33] V. L. Bonch-Bruевич and S. V. Tyablikov. *The Green Function Method in Statistical Mechanics*. (North-Holland, Amsterdam, 1962).
- [34] A. C. Davydov and G. M. Pestryakov. On the canonical transformation in the theory of exciton-phonon interaction. *Phys. Stat. Sol. (b)* **49**: 505–, 1972.
- [35] M. Bayer and A. Forchel. Temperature dependence of the homogeneous linewidth in $\text{In}_{0.60}\text{Ga}_{0.40}\text{As}/\text{GaAs}$ self-assembled quantum dots. *Phys. Rev. B* **65**: art. no. 041308, 2002.
- [36] F. Vallée and F. Bogani. Coherent time-resolved investigation of LO-phonon dynamics in GaAs. *Phys. Rev. B* **43**: 12049–12052, 1991.
- [37] F. Vallée. Time-resolved investigation of coherent LO-phonon relaxation in III-V semiconductors. *Phys. Rev. B* **49**: 2460–2468, 1994.
- [38] R. Alicki, M. Horodecki, P. Horodecki, and R. Horodecki. Dynamical description of quantum computing: generic nonlocality of quantum noise. *Phys. Rev. A* **65**: art. no. 062101, 2002.
- [39] E. Biolatti, R. C. Lotti, P. Zanardi, and F. Rossi. Quantum information processing with semiconductor macroatoms. *Phys. Rev. Lett.* **85**: 5647–5650, 2000.
- [40] P. Chen, C. Piermarocchi, and L. J. Sham. Control of exciton dynamics in nanodots for quantum operations. *Phys. Rev. Lett.* **87**: art. no. 067401, 2001.
- [41] C. Piermarocchi, Pochung Chen, Y. S. Dale, and L. J. Sham. Theory of fast quantum control of exciton dynamics in semiconductor quantum dots. *Phys. Rev. B* **65**: art. no. 075307, 2002.
- [42] L. Tian and S. Lloyd. Resonant cancellation of off-resonant effects in a multilevel qubit. *Phys. Rev. A* **62**: art. no. 050301, 2000.
- [43] L.-A. Wu, M. S. Byrd, and D. A. Lidar. Efficient universal leakage elimination for physical and encoded qubits. *Phys. Rev. Lett.* **89**: art. no. 127901, 2002.

REAL SPACE *AB INITIO* CALCULATIONS OF EXCITATION ENERGIES IN SMALL SILICON QUANTUM DOTS

Aristides D. Zdetsis and C. S. Garoufalidis

Department of Physics, University of Patras, 26500 Patras, Greece

Stefan Grimme

Organisch-Chemisches Institut, Westfälische Wilhelms Universität, Corrensstraße 40, D-48149 Münster, Germany

Abstract

We report accurate *ab initio* calculations, performed over the last three years, of the optical gap of small silicon nanocrystals (Si dots). In these calculations the Si dangling bonds on the surface of the nanocrystals are passivated by hydrogen and (more recently) by hydrogen and oxygen. The actual diameters of the nanocrystals range from about 2 Å (SiH₄) to 25 ° 281H₁₇₂). The results for the optical gap can be safely (judging from the quality of the fit) extrapolated up to much larger diameters.

We have used a large variety of real-space theoretical techniques, including configuration-interaction singlets (CIS) and multi-reference second-order perturbation theory (MR-MP2). The bulk of the calculations has been performed in the framework of density functional theory (DFT) using advanced methods ranging from screened Coulomb interactions and ΔSCF calculations to DFT/CIS and time-dependent DFT (TDDFT).

In all DFT methods we have used the hybrid nonlocal exchange-correlation functional of Becke and Lee, Yang and Parr, which includes partially exact Hartree-Fock exchange (B3LYP). Our results are in excellent agreement with accurate recent and earlier experimental data. We have found that the diameter of the smallest oxygen-free nanocrystal that could emit photoluminescence in the visible region of the spectrum is around 22 Å, whereas the largest diameter falls in the range of 84–85 Å. The high level and the resulting high accuracy (estimated within 0.3 eV for the optical gap) of our calculations have led to the resolution of existing experimental and/or theoretical discrepancies. Our (most recent) results also clarify unambiguously and confirm earlier predictions about the role of oxygen on the gap size.

1. Introduction

The study of the optical properties of porous silicon and silicon nanocrystals has been a very promising field of research for the last decade. A large number of experimental and theoretical approaches have been used to explore the properties and resolve the puzzle about the origins of the observed visible photoluminescence (PL). However, until recently [1], several issues were not unambiguously settled. For instance, the majority of the experimental work so far gives diverse results for the size of the Si dots capable of emitting in the visible and the role of surface passivants. Recent experimental data suggest that the diameters of the dots able that emit in the visible are much less than 20 Å. The results of Wolkin *et al.* [3] revealed optical gaps as small as 2.2 eV for nanoclusters with a diameter of 18 Å. For nanoclusters of about the same size Wilcoxon *et al.* [2] obtained a similar result (2.5 eV), together with a much larger gap of about 3.2 eV for highly purified samples of the same dot diameter. Furthermore, Schupler *et al.* [4, 5] have estimated the critical diameter for visible PL to be less than 15 Å.

In addition to this body of work, recent studies about the role of surface oxygen on the optical properties of silicon nanocrystals have reported conflicting levels of importance, ranging from minimal to crucial. In many experiments, the presence of oxygen is considered as a means of effectively passivating the surface dangling bonds or as a means of reducing (through oxidation) the size of the silicon nanoparticles. In most of the cases oxygen is just a contaminant, which is very difficult to remove. Gole *et al.* [6], have performed an extensive series of time-dependent excitation spectroscopy experiments on porous silicon, coupled with quantum chemical calculations. This work gave strong evidence that the observed PL is due to oxyhydride-like fluorophors bound to PS surface [6]. The same results, complemented by quantum chemical calculations on candidate molecular emitters led Gole *et al.* [6] to a model interpretation of PL in which the key role is played by the Si=O bonds.

Apart from the role of oxygen, there are two controversial issues in this field: (1) the mechanism responsible for the visible PL and (2) the variation of the optical gap with the dot diameter d . These issues are not independent of each other. A settlement of the second issue will settle at the same time the issue of the mechanism of the visible PL. The main source of error in the experimental results, besides the purity of the sample, is the determination of the diameter of the dot, which is very difficult for very small dots. As a result, the optical gap is a multivalued function of d , which is different in different works (see, for

instance, Figs. 10, 4, and 5 in Refs. [2, 3, 7], respectively). In fact, in some cases, the diameter d of the dot is determined by comparison to theoretical calculations of the gap versus d [3].

From the theoretical point of view, the *ab initio* investigation of the optical properties of the dot demands an accurate account of electron correlation and involves a high computational cost, which scales at least with the fifth or sixth power of the cluster size d . As a result, the early attempts to implement existing theoretical and computational techniques were focused mainly on empirical and semiempirical approximations. More recently however, several *ab initio* calculations have appeared [1, 8, 9], providing a more realistic description of the emission of light from Si nanocrystals. Although quantum confinement retains its key role in the heart of any explanation of the phenomenon, there is still place for additional or complementary channels of visible radiation (especially in the small cluster regime). Despite the evidence given by the experiments [2, 3, 4, 5, 6, 10, 11] about the role of oxygen, there has been a lack of realistic model calculations towards this direction until recently [9, 12, 13]. For this reason we also present the largest, so far, accurate comparative calculations of the optical gap based on time dependent density functional theory (TDDFT) [20] employing the functionals of Becke and Perdew BP86 [14, 15] and the hybrid nonlocal exchange-correlation functional B3LYP [16]. It has been demonstrated [17] that this functional for small Si nanoparticles gives superior results for electronic and structural properties.

2. Results for Oxygen-Free Nanocrystals

2.1 Early Attempts

As was explained earlier, the need for low-computational-cost correlation methods led us to the exploration of several theoretical techniques for the description of the optical properties of Si dots. As was proven later by the application and comparison with more sophisticated methods, some of these techniques could indeed be reliable and therefore very promising for extended calculations in larger dots (compare for example Tables 1 and 2). Some of these techniques (with increasing level of complexity), are: the semiempirical MNDO and AM1 methods, pure LDA calculations, B3LYP/ Δ SCF, Hartree-Fock configuration interaction singlets (HF/CIS) and DFT/B3LYP configuration interaction Singlets (DFT/CIS) [18, 19]. A relatively simple method which we have also employed is the indirect estimation of the optical gap ($\varepsilon_g^{\text{opt}} = \varepsilon_g^{\text{HL}} - E_{\text{Coul}}$) through the calculation of the exciton binding energy using a size-dependent dielectric constant. With this method, the

screened electron-hole interaction E_{coul} is calculated as:

$$E_{\text{Coul}} = \int \frac{|\Psi_H(\mathbf{r}_1)|^2 |\Psi_L(\mathbf{r}_2)|^2}{\varepsilon_{\infty}^{\text{dot}} |\mathbf{r}_1 - \mathbf{r}_2|} d\mathbf{r}_1 d\mathbf{r}_2, \quad (1)$$

where $\varepsilon_{\infty}^{\text{dot}}$ is a size-dependent dielectric function of the form [19]

$$\varepsilon_{\infty}^{\text{dot}} = \frac{\varepsilon_0 - 1}{1 + (R_0/R)^\eta}. \quad (2)$$

In Table 1 we have collected some of the simple, but promising, “low-cost” methods. As was mentioned before, without the accurate high level *ab initio* calculations, the results of Table 1 would be very close to be considered as naive solutions of a complex problem.

Table 1. Comparison of the optical gap for small Si nanoclusters with relatively “low cost” and relatively reliable correlation methods.

Cluster	B3LYP/ ΔSCF (eV)	HF/CIS (eV)	DFT/CIS ^a (eV)	$\varepsilon_g^{\text{HL}} - E_{\text{Coul}}^{\text{a}}$ (eV)	Expt (eV)
Si ₁ H ₄	9.1	9.22	9.05	8.55	8.8
Si ₂ H ₆	7.0	7.3	7.5	—	7.6
Si ₅ H ₁₇	6.16	6.56	6.0	6.4	6.5
Si ₁₇ H ₃₆	4.68	5.46	5.25	5.0	—
Si ₂₉ H ₃₆	4.45	5.44	5.01	4.7	—

^a The B3LYP functional was used.

2.2 High-Level *Ab Initio* Calculations

High-level calculations of the optical gap include time-dependent density functional theory (TDDFT) [20] and the multi-reference second-order perturbation theory (MR-MP2) [21, 22]. As mentioned before, the size of the quantum dots considered here ranges from 1 to 281 Si atoms, with 4 to 172 H atoms (a total of about 453 atoms). Some of these systems are shown in Fig. 1. The diameter of the larger clusters falls in the range of 12–25 Å for which visible photoluminescence has been reported [3, 4, 5, 10, 11, 23, 24]. All dots have T_d symmetry and their geometries have been fully optimised within this symmetry constraint using the B3LYP exchange-correlation functional [16]. We have shown, by performing high-level MR-MP2 [21, 22] and additional TDDFT calculations using the well-known functional of Becke and Perdew (BP86) [14, 15], that the partially exact Hartree-Fock (HF) exchange, which is included in the B3LYP method, is very important for the correct description of the optical properties. The BP86 functional does not include exact (or

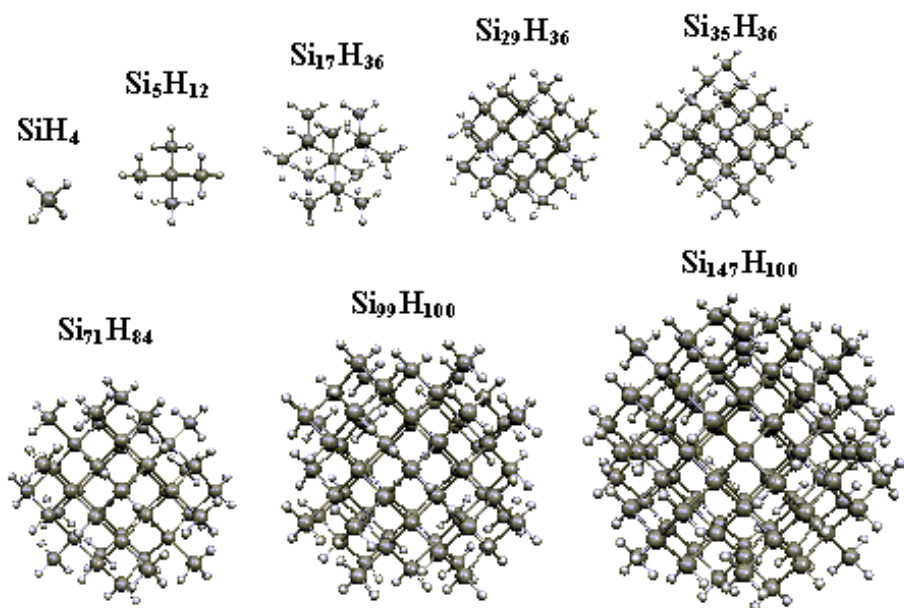


Figure 1. Some of the Si_xH_y nanocrystals studied here represented in terms of ball-and-stick models. Large solid spheres denote Si atoms and the smaller spheres denote surface H atoms.

partially exact) exchange which, as a result, produces a systematic underestimate of the optical gap calculated with the TDDFT method. The inclusion of exact HF exchange also remedies the well-known deficiency of the local-density approximation (LDA) to underestimate the band gap (in this case the HOMO-LUMO gap). The importance of exchange interaction for a non-metallic material is rather surprising and is related to quantum confinement.

The DFT and the TDDFT calculations were performed with the TURBOMOLE [25] suite of programs using Gaussian atomic orbital basis sets of split valence [SV(P)]: $[4s3p1d]/2s$ [26] quality which involves 5400 basis functions for the largest system studied. The TDDFT calculations have been performed as described in detail in Ref. [27] using the B3LYP functional consistently for both, the self-consistent solution of the Kohn-Sham equation for the ground state, and the solution of the linear response problem. The optical gap in TDDFT (and MR-MP2) is identified as the energy of the lowest allowed electronic transition (i.e. with nonzero oscillator strength). The MR-MP2 calculations were performed as described in Ref. [28] and can be outlined as follows:

- The zeroth-order configuration-interaction (CI) reference wavefunction is expressed in spin and space symmetry adapted configuration state functions (CSF):

$$\Psi^{(0)} = \sum_i \Psi_i^0.$$

- The first-order correction is constructed by CSF generated by simple and double excitations of reference configurations:

$$\Psi^{(1)} = \sum_{\alpha} C_{\alpha} \Psi_{\alpha}^1.$$

- The expansion coefficients C_{α} are obtained by solving the inhomogeneous set of linear equations:

$$\sum_{\alpha} C_{\alpha} \langle \Psi_{\beta}^1 | H - E_0 | \Psi_{\alpha}^1 \rangle = -\langle \Psi_{\beta}^1 | H | \Psi^{(0)} \rangle.$$

- The second-order energy correction $E(\text{MR-MP2})$ is evaluated as

$$E(\text{MR-MP2}) = -\sum_{\alpha} C_{\alpha} \langle \Psi_{\alpha}^1 | H | \Psi^{(0)} \rangle.$$

The general procedure adopted for the selection of the reference configuration space can be outlined as follows:

- Substitute HF virtual orbitals with improved virtual orbitals (IVO)*.
- R(O)HF calculation of the ground state.
- Determination of IVOs.
- Choose an initial restricted active space (RAS) and the excitation level.
- Quick MR-MP2 calculation with truncated external orbital space.
- Determine approximate natural orbital occupation numbers. Include orbitals which differ by more than 0.02–0.03 from integers 0, 1, 2 in the RAS.
- Repeat the MR-MP2 calculation with full external orbital space.

In these calculations, ground state HF molecular orbitals were used and the RI (resolution of the identity) approximation for the two electron integrals (see Ref. [28]) was employed. Again, the SV(P) basis set was

utilized. All valence electrons (32 for Si_5H_{12} , 104 for $\text{Si}_{17}\text{H}_{36}$, and 152 for $\text{Si}_{29}\text{H}_{36}$) were correlated. The zeroth-order wave functions in the MR-MP2 treatment were constructed from single excitation CI wave functions including about 30 of the frontier orbitals. With this choice, the largest amplitude in the first-order corrected MP2 wave functions is always smaller than 0.03, a point where MR-MP2 relative energies are usually converged to about 0.1 eV.

As can easily be seen in Table 2, the MR-MP2 results practically coincide with the TDDFT/B3LYP results. Therefore, for computer time and storage economy, the MR-MP2 calculations have been restricted here to nanoclusters containing up to 29 Si atoms (with 36 additional H atoms). The silicon nanoparticles considered here have a closed-shell electronic structure in the ground state with vanishing non-dynamical electron correlation contributions.

Table 2. Comparison of the optical gap for small Si nanoclusters.

No. of Si atoms	Total No. of atoms	H-L ^a gap (eV)	TDDFT ^b (eV)	MR-MP2 (eV)	f^c
5	17	7.6	6.66	6.76	0.0065
17	353	5.72	5.03	5.02	0.1507
29	65	5.15	4.53	4.45	0.0189
35	71	5.04	4.42	—	0.0052
47	107	4.64	4.04	—	0.1151
71	155	4.20	3.64	—	0.0124
99	199	3.89	3.39	—	0.0009
147	247	3.66	3.19	—	0.0007
281	453	3.11	—	—	—

^a HOMO–LUMO gap from Kohn–Sham eigenvalues of the DFT/B3LYP ground state.

^b TDDFT/B3LYP.

^c Oscillator strength.

The lowest excited states of T_2 symmetry are dominated by a small number of single excitations between orbitals of pure valence character; the double excitation contributions are less than 8% in the MR-MP2 calculations. For states of this type, according to prior experience [27, 28], MR-MP2 and TDDFT/B3LYP methods have an estimated accuracy of about 0.3 eV for the excitation energies. Test calculations at the MR-MP2 level using a larger triple-zeta atomic orbitals basis ($[5s4p1d]/[3s1p]$) for Si_5H_{12} give an excitation energy of 6.51 eV, smaller by only 0.25 eV, whereas for the largest clusters, the basis set difference is expected to be negligible. Therefore, 0.3 eV is a conservative estimate of the error margin.

The sizes of the nanoclusters compiled in Table 2 range from about 7 Å to 25 Å. This includes the sizes of 15 Å, 18 Å, and 20 Å, for which PL has been reported [3, 4, 5, 10] with peaks between 2.1 eV and 2.5 eV. Our calculated optical gaps in this region range from about 3.4 eV to 6.7 eV, with a possible error of about 0.3 eV. The results from the two very different theoretical approaches agree to within 0.1 eV for the three smallest systems. These values are clearly well above the observed limits of 2.1–2.5 eV. Therefore, since the observed PL could not have its origin in quantum dots of this type and/or size, we suggest that either the samples are contaminated with oxygen (mainly) or their size is not correctly determined. As will be illustrated below, this conclusion is not in conflict but, on the contrary, is supported by recent experimental data from two different groups [2, 3], who have demonstrated that these measured values refer to oxygen-containing samples. Furthermore, as we have mentioned earlier, the experimental determination of the nanocluster sizes is difficult and rather ambiguous for very small nanoclusters (dots). It is very difficult to be certain, without large statistical samples, that the diameters obtained (usually) from TEM images are truly representative. In several cases where a statistical analysis was performed, it was found that nanoclusters samples exhibit a broad distribution of sizes [2].

To avoid this ambiguity, Wolkin *et al.* [3] have determined the sizes of their Si dots by equating their measured peak PL energies with calculated excitonic band gaps, using a simple tight binding model. This model, which is similar to and produces almost identical results with those of Reboredo *et al.* [7], cannot match the accuracy of the present calculations. On the other hand, it is obvious that, if our calculations were used to determine the sizes of the Si dots, the agreement with the experimental results of Wolkin *et al.* [3] for oxygen-free samples, as is shown in Fig. 2, would be perfect by definition. This illustrates an alternative role our results could play, namely, a well-defined reference “yard stick” for the actual size of the dots. Also, it is well known that surface oxidation of the samples in air could dramatically change their PL spectra. Wolkin *et al.* [3] have demonstrated that even a 3 min exposure of the samples in air produces a redshift as large as 1 eV due to the formation of oxygen bonds on the surface. Thus, the great majority of experimental work deals with nanoclusters covered with a few layers of silicon oxides and the dimensions of the inner crystalline core are not strictly defined. The last two observations can easily explain the discrepancy between our optical gap and experimental measurements such as those of Schuppler *et al.* [4, 5], Wilcoxon *et al.* [2], and Brus *et al.* [10].

Figure 10 of Wilcoxon *et al.* [2] summarizes most of the existing experimental results for PL peak energies as a function of the diameter d

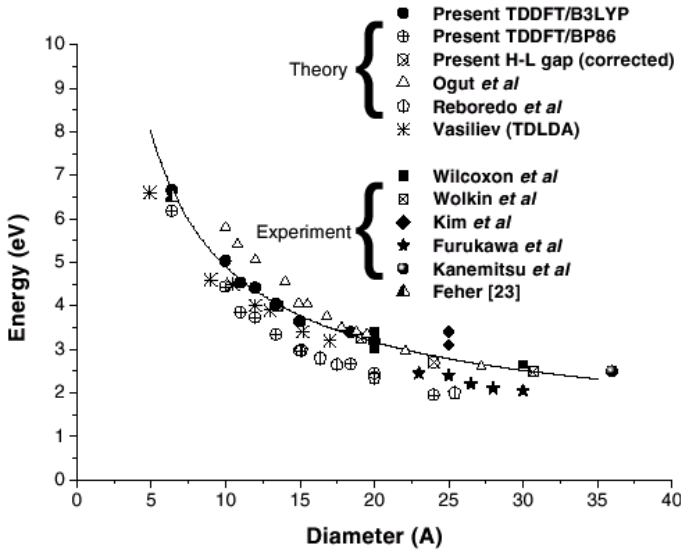


Figure 2. Comparison of the calculated optical gap in this work with different experimental results and theoretical calculations. The solid line is a least square fit to the TDDFT/B3LYP results of Table 2.

of the nanoclusters. We can clearly see that all these measurements fall into the shaded region of the diagram. As explained by these authors, all PL peaks of SiO_2 capped nanoclusters or nanoclusters embedded in glass matrices fall into this region. The central issue for this region [2], which contains a wide band of energies for each value of d , is the role of SiO_2 or glass and suboxide layer that almost certainly exists at the interface. As Wilcoxon *et al.* [2] point out, there are undoubtedly defects and surface states at this interface that could play a significant role in the luminescence from these samples. The experimental data of Ref. [2] for the oxygen-free samples (for the same value of the diameter) fall above this shaded region. Obviously, this is the region on which we shall focus here to make contact with our calculations. This region corresponds to direct electron-hole recombination, as Wilcoxon *et al.* have suggested.

For dots with diameters around 18–20 Å, which correspond to our Si dots with 99 and 147 Si atoms, respectively, the experimental measurements of Wilcoxon *et al.* on size selected and purified samples give the energy of the major PL peak at 3.40 eV (365 nm). This is in excellent agreement with our results (3.2 eV, 3.4 eV). Wilcoxon *et al.* have attributed this peak to the direct $\Gamma_{25}-\Gamma_{15}$ transition, which in real space corresponds to the t_1-t_2 transition. This is also in agreement with our results, which verifies the observation of this direct transition for the

first time. The measured exciton binding energy E_B for the same sample is 0.4 eV, whereas the calculated value from the difference of the optical gap from the (HOMO-LUMO or H-L) gap is about 0.5 eV. As we can clearly see from Table 2, this value of E_B varies from about 1 eV for Si_5H_{12} with $d=7$ Å, to about half this value for nanoclusters in the range of 18–25 Å. The value $E_B=0.45$ eV remains practically constant over this large range of d . This can be used to calculate the optical gap of somewhat larger nanoclusters by performing simple ground state DFT/B3LYP calculations only. For the $\text{Si}_{281}\text{H}_{172}$ cluster, we have used this correction of the H-L gap to estimate the optical gap.

In Fig. 3 we show a plot of the calculated optical gap as a function of the dot diameter. The smallest possible dot diameter for visible PL (PL energy 3.0 eV) is about 22 Å. On the other hand, the diameter of the largest nanocrystal capable of emitting in the visible can be estimated by an extrapolation of these results. For the TDDFT/B3LYP optical gap, this leads to a maximum nanocrystal diameter of 84–85 Å. In Fig. 2, we have also included TDDFT results, using the BP86 [14, 15] exchange-correlation functional, which, however, unlike B3LYP, does not include exact or partially exact exchange. The BP86 results practically coincide with the results of Reboredo *et al.* [7], although on average they are about 0.6–0.7 eV lower compared to the more accurate TDDFT/B3LYP. The

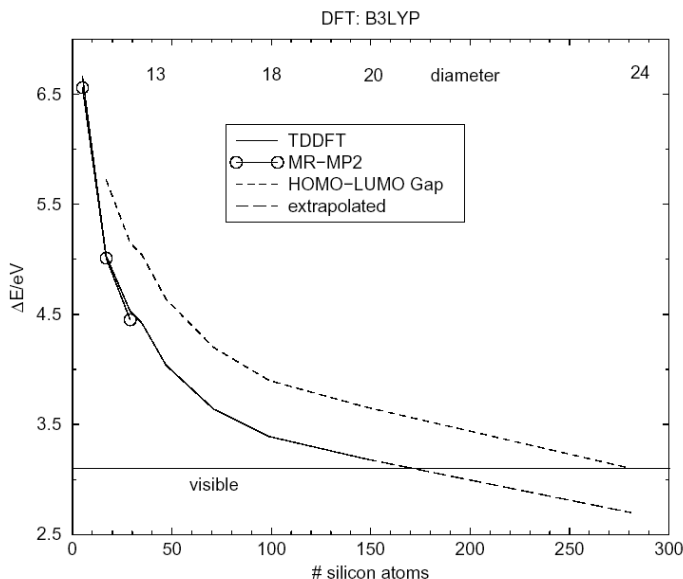


Figure 3. Extrapolation of the optical gap and the HOMO-LUMO gap.

similarity of our TDDFT/BP86 results with those of Reboredo *et al.* [7], which are based on tight-binding supercell calculations using empirical pseudopotentials, is rather surprising. This is highly suggestive that, apart from the technical differences between the two calculations, the real reason for the difference of their supercell results from our present real-space TDDFT/B3LYP results could be their improper treatment of exchange. On the other hand, the theoretical results (especially for very small dots) of Ögüt *et al.* [29] for the excitonic energy gap overestimate the optical gap (with the meaning of the optical gap given here, namely, the energy of the lowest allowed excitation). This, according to Reboredo *et al.*, is due to the underestimation of the screening function and overestimation of the quasiparticle energy gap [29]. However, for dots larger than 17 Å in diameter, the discrepancies in the calculations of Ref. [29] seem to cancel out and have no real influence on the calculated optical gap, which is very close to our TDDFT/B3LYP results.

3. Oxygen-Containing Nanocrystals

To examine the role of oxygen passivation (and especially the existence of Si=O bonds) on the optical gap of silicon nanocrystals, we had to take some simplifying steps to make the calculation feasible. The most important of these stems from the necessity of maintaining the T_d symmetry of the nanocrystals, since any deviation from this point group would greatly affect the computational cost, making the use of sophisticated *ab initio* techniques impractical. The starting point was the determination of the geometry of a carefully selected set of nanocrystals obtained by oxygen substitution of hydrogen atoms from the corresponding oxygen-free [1] structures. We have chosen surface silicon atoms having a bonding scheme of the form $R_N\text{Si}_2\langle\text{Si}\rangle\text{H}_2$, where R_N signifies the main body of the nanocluster, Si_2 are the two silicon atoms bonded to the Si atom of interest, $\langle\text{Si}\rangle$, and H_2 represents the two hydrogen passivants. For such a site, we identify the symmetrically-equivalent Si atoms and replace the two hydrogen passivants by one doubly bonded oxygen atom. For these sites the bonding scheme becomes $R_N\text{Si}_2\langle\text{Si}\rangle\text{O}$. It is clear that, for the silicon atoms of interest, $\langle\text{Si}\rangle$, the bonding changes from sp^3 to sp^2 and that, for the nanocrystals considered, the number of surface oxygen atoms is dictated by the need to maintain T_d symmetry. The quantum dots considered here are $\text{Si}_{17}\text{H}_{12}\text{O}_{12}$, $\text{Si}_{71}\text{H}_{72}\text{O}_6$, $\text{Si}_{99}\text{H}_{76}\text{O}_{12}$, and $\text{Si}_{147}\text{H}_{52}\text{O}_{24}$ (Fig. 4). Their geometries have been fully optimised within the T_d symmetry constraint using the B3LYP functional [16].

Our TDDFT/B3LYP results on oxygen-containing dots are in very good agreement with the recent quantum Monte Carlo ΔSCF -like cal-

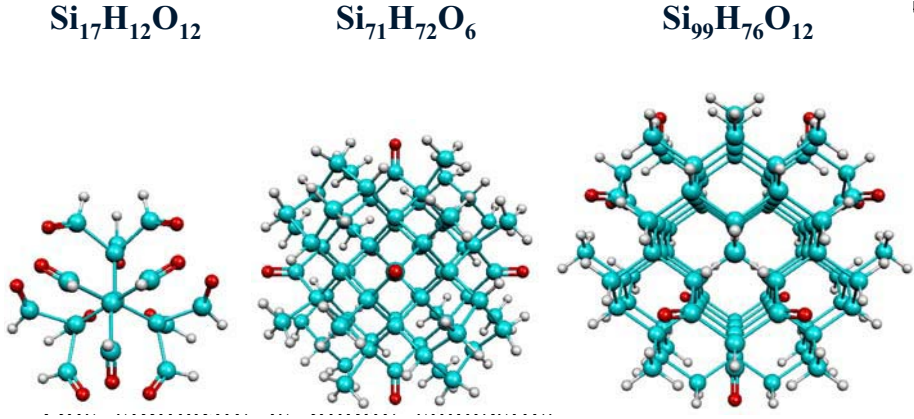


Figure 4. Some of the oxygen containing nanocrystals with T_d symmetry

culations of Puzder *et al.* [12] and the experimental data of Wilcoxon *et al.* [2]. Additionally, they establish the same trends with the corresponding results of Ref. [9]. Compared to the corresponding optical gap of the oxygen-free nanocrystals [1], the nanoclusters $\text{Si}_{17}\text{H}_{12}\text{O}_{12}$, $\text{Si}_{71}\text{H}_{72}\text{O}_6$, $\text{Si}_{99}\text{H}_{76}\text{O}_{12}$, and $\text{Si}_{147}\text{H}_{52}\text{O}_{24}$ show a gap decrease of the order of 1 eV. For the larger dots, this is consistent with the result of Wolkin *et al.* [3], who have demonstrated that even a 3 min exposure of the samples in air produces a red shift as large as 1 eV due to the formation of oxygen bonds on the surface. Our results reproduce the experimental results of Wilcoxon *et al.* [2], as depicted in their Fig. 10, which shows oxygen-containing nanocrystals, lying in a shaded area, exhibiting a significantly lower optical gap. Our results clearly fall in this shaded region. It becomes clear, therefore, that the diverse experimental data on the size of the optical gap *versus* the size of the nanocrystals (capable of emitting in the visible region of the spectrum) can be easily understood from the existence of doubly-bonded oxygen atoms on the surface of the nanoparticles. For example, as we have demonstrated, in the case where $\text{Si}=\text{O}$ bonds are present, the minimum dimensions of dots capable of emitting in the visible can be even smaller than 15 Å. This is consistent with the claims of Schuppler *et al.* [4, 5] (accepting that their samples contain oxygen, as the authors themselves admit). On the other hand, in cases where the nanocrystals are free of oxygen, the critical dimension for light emission can be significantly larger and well above 22 Å. These findings bridge the differences of existing experimental results, making them consistent with each other.

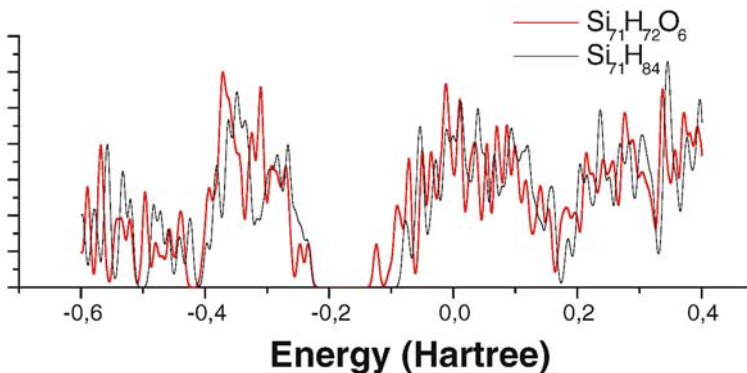


Figure 5. Densities of states of $\text{Si}_{71}\text{H}_{84}$ and $\text{Si}_{71}\text{H}_{72}\text{O}_6$ nanoclusters.

In Fig. 5 we have plotted the densities of states of $\text{Si}_{71}\text{H}_{84}$ and $\text{Si}_{71}\text{H}_{72}\text{O}_6$ nanoclusters. These curves were generated from the eigenstates of the ground state with a suitable Gaussian broadening. It can be seen that the presence of oxygen affects mainly the tail of the conduction band by inserting a new peak dominated by oxygen. Thus, the decrease of the optical gap is consistent with an analogous decrease of the HOMO–LUMO gap. From these plots can be easily seen that, for $\text{Si}_{71}\text{H}_{72}\text{O}_6$, the oxygen-related peak is located well within the gap and in the energy region of -3.4 eV. The presence of these peaks inside the gap makes the formation of stable $\text{Si}=\text{O}$ bonds possible. For the small dots regime, where the gap is bigger due to stronger confinement, the oxygen related states do not overlap with the conduction band. Thus, oxygen plays the key role in the determination of the optical properties of Si nanocrystals. As the size of the dots increases the band gap decreases and the oxygen peaks gradually overlap with the conduction band. This destabilises the $\text{Si}=\text{O}$ bonds and, for bigger nanocrystals, it even prevents their formation. In this case oxygen serves as a simple passivant and it no longer determines the optical properties of the quantum dot.

4. Conclusions

We have shown that

- 1 Quantum confinement is responsible for the visible PL from oxygen-free samples of silicon nanoclusters (Si quantum dots).
- 2 The diameter of the smallest dots, which could emit PL in the visible (in the limits between violet and ultraviolet ~ 3.0 eV), is around 22 Å.

- 3 The upper limit in the size of dots emitting in the visible is estimated to be around 85 Å. The existing discrepancies in the experimental results are due either to oxygen contamination and/or size uncertainties. The discrepancies in the theoretical calculations are mainly due to the poor treatment of the exchange interaction and the non-transferability of bulk parameters to small size dots.
- 4 In the small dot regime, oxygen can form stable double bonds with the silicon atoms. The existence of such bonds introduces energy levels, which lie well within the gap and they are responsible for visible light emission.

Our present calculations for oxygen containing Si nanocrystals are the largest and most accurate, with an estimated error margin of 0.3 eV. These calculations are in excellent agreement with existing experimental measurements and clarify the role of oxygen.

Acknowledgements

This work was partially supported by a “KARATHEODORIS” grant at the University of Patras, Greece.

References

- [1] C. S. Garoufalis, A. D. Zdetsis, and S. Grimme. High level *ab initio* calculations of the optical gap of small silicon quantum dots. *Phys. Rev. Lett.* **87**: art. no. 276402, 2001.
- [2] J.P. Wilcoxon, G. A. Samara, and P. N. Provencio. Optical and electronic properties of Si nanoclusters synthesized in inverse micelles. *Phys. Rev. B* **60**: 2704–2714, 1999.
- [3] M. V. Wolkin, J. Jorne and P. M. Fauchet, G. Allan, and C. Delerue. Electronic states and luminescence in porous silicon quantum dots: The role of oxygen. *Phys. Rev. Lett.* **82**: 197–200, 1999.
- [4] S. Schuppler, S. L. Friedman, M. A. Marcus, D. L. Adler, Y.-H. Xie, F. M. Ross, Y. L. Chabal, T. D. Harris, L. E. Brus, W. L. Brown, E. E. Chaban, P. F. Szaajowski S. B. Christman, and P. H. Citrin. Dimensions of luminescent oxidized and porous silicon structures. *Phys. Rev. Lett.* **72**: 2648–2651, 1994.
- [5] S. Schuppler, S. L. Friedman, M. A. Marcus, D. L. Adler, and Y.-H. Xie, F. M. Ross, Y. J. Chabal, T. D. Harris, L. E. Brus, W. L. Brown, E. E. Chaban, P. F. Szaajowski, S. B. Christman, and P. H. Citrin. Size, shape, and composition of luminescent species in oxidized Si nanocrystals and H-passivated porous Si. *Phys. Rev. B* **52**: 4910–4925, 1995.
- [6] J. L. Gole and D. A. Dixon. Potential role of silanones in the photoluminescence-excitation, visible-photoluminescence-emission, and infrared spectra of porous silicon. *Phys. Rev. B* **57**: 12002–12016, 1998.

- [7] F. A. Reboredo A. Franceschetti, and A. Zunger. Dark excitons due to direct Coulomb interactions in silicon quantum dots. *Phys. Rev. B* **61**: 13073–13087, 2000.
- [8] I. Vasiliev, S. Ögüt, and J. Chelikowsky. *Ab initio* absorption spectra and optical gaps in nanocrystalline silicon. *Phys. Rev. Lett.* **86**: 1813–1816, 2001.
- [9] I. Vasiliev, J. R. Chelikowsky and R. M. Martin. Surface oxidation effects on the optical properties of silicon nanocrystals. *Phys. Rev. B* **65**: art. no. 121302, 2002.
- [10] L. E. Brus, P. F. Szajowski, W. L. Wilson, T. D. Harris, S. Schuppler, and P. H. Citrin. Electronic spectroscopy and photophysics of Si nanocrystals: Relationship to bulk C-Si and porous Si. *J. Am. Chem. Soc.* **117**: 2915–2922, 1995.
- [11] Y. Kanemitsu. Luminescence properties of nanometer-sized Si crystallites: Core and surface states. *Phys. Rev. B* **49**: 16845–16848, 1994.
- [12] A. Puzder, A. J. Williamson, J. C. Grossman and G. Galli. Surface chemistry of silicon nanoclusters *Phys. Rev. Lett.* **88**: art. no. 097401, 2002.
- [13] C. S. Garoufalidis, A. D. Zdetsis and S. Grimme, to be published.
- [14] A. D. Becke. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**: 3098–3100, 1988.
- [15] J. P. Perdew. Density-functional approximation for the correlation energy of the inhomogeneous electron gas. *Phys. Rev. B* **33**: 8822–8824, 1986.
- [16] P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch. *Ab initio* calculation of vibrational absorption and circular dichroism spectra using density-functional force fields. *J. Phys. Chem.* **98**: 11623–11627, 1994.
- [17] A. D. Zdetsis. The real structure of the Si₆ cluster *Phys. Rev. A* **64**: art. no. 023202, 2001.
- [18] C. S. Garoufalidis, A. D. Zdetsis, and J. P. Xanthakis. in *Microelectronics Microsystems and Nanotechnology*, edited by A. G. Nassiopoulou and X. Zianni (World Scientific, Singapore, 2000), pp 85–88.
- [19] C. S. Garoufalidis and A. D. Zdetsis. in *Microelectronics Microsystems and Nanotechnology*, edited by A. G. Nassiopoulou and X. Zianni (World Scientific, Singapore, 2000), pp 81–84.
- [20] M. E. Casida. Time-dependent density-functional response theory for molecules. in *Recent Advances in Density Functional Methods*, Vol. 1, edited by D. P. Chong (World Scientific, Singapore, 1995), pp 155–188.
- [21] R. B. Murphy and R. P. Messmer. Generalized Moller–Plesset Perturbation theory applied to general MCSCF reference wave functions. *Chem. Phys. Lett.* **183**: 443–448, 1991.
- [22] R. B. Murphy and R. P. Messmer. Generalized Moller–Plesset and Epstein–Nesbet perturbation theory applied to multiply bonded molecules. *J. Chem. Phys.* **97**: 4170–4184, 1992.
- [23] K. Kim. Visible light emissions and single-electron tunneling from silicon quantum dots embedded in Si-rich SiO₂ deposited in plasma phase. *Phys. Rev B* **57**: 13072–13076, 1998.
- [24] S. Furukawa and T. Miyasato. Quantum size effects on the optical band gap of microcrystalline Si:H. *Phys. Rev. B* **38**: 5726–5729, 1988.

- [25] TURBOMOLE (Vers. 5.3), Universität Karlsruhe, 2000.
- [26] A. Schäfer, H. Horn and R. Ahlrichs. Fully optimized contracted Gaussian basis sets for atoms Li to Kr. *J. Chem. Phys.* **97**: 2571–2577, 1992.
- [27] R. Bauernschmitt and R. Ahlrichs. Treatment of electronic excitations within the adiabatic approximation of time dependent density functional theory. *Chem. Phys. Lett.* **256**: 454–464, 1996.
- [28] S. Grimme and M. Waletzke. Multi-reference Moller-Plesset theory: Computational strategies for large molecules. *Phys. Chem. Chem. Phys.* **2**: 2075–2081, 2000.
- [29] S. Ögüt, J. Chelikowsky, and S. G. Louie. Quantum confinement and optical gaps in Si nanocrystals. *Phys. Rev. Lett.* **79**: 1770–1773, 1997. *Phys. Rev. Lett.* **80**: 3162, 1998, and references therein.

GeSi/Si(001) STRUCTURES WITH SELF-ASSEMBLED ISLANDS: GROWTH AND OPTICAL PROPERTIES

N. V. Vostokov, Yu. N. Drozdov, D. N. Lobanov, A. V. Novikov, M. V. Shaleev, A. N. Yablonskii and Z. F. Krasilnik

*Institute for Physics of microstructures RAS, 603950, Nizhny Novgorod,
GSP-105, Russia*

zfk@ipm.sci-nnov.ru

A. N. Ankudinov, M. S. Dunaevskii and A. N. Titkov

Ioffe Physico-Technicheskii Institute, Polytekhnicheskaya 26, St.Petersburg, Russia

P. Lytvyn, V. U. Yukhymchuk and M. Ya. Valakh

Institute of Semiconductor Physics, NASU, 03028 Kiev, Ukraine

Abstract The fabrication of Ge(Si) self-assembled islands on Si substrates has attracted much attention in the last several years because of their great potential for practical application. In this paper we present the results of investigations into the growth and optical properties of single and multilayer structures with Ge(Si)/Si(001) self-assembled islands, grown at different Ge deposition temperatures (T_{Ge}). It is shown that the sizes and surface density of the islands can be varied in a wide range by altering T_{Ge} : from nanometer scale islands for high T_{Ge} to small quantum dots with a height of ~ 1 nm and a surface density of $> 10^{11} \text{ cm}^{-2}$. The GeSi alloy formation in islands grown at $T_{\text{Ge}} > 580^\circ \text{C}$ was observed by X-ray diffraction, Raman scattering and selective etching. The alloy formation is caused by a strain-driven Si diffusion in islands. The dependence of island composition on T_{Ge} was experimentally measured and found to drastically affect the sizes and shapes of the islands.

1. Introduction

The first publications devoted to the effect of spontaneous self-assembly of three-dimensional germanium nanoislands on silicon appeared more than 10 years ago [1–3]. Since that time a number of papers and re-

views has been published about general questions of the growth of quantum dots and nanoislands on Ge/Si heterostructures by the Stranski–Krastanov growth mode, as well as to original study of the composition and properties of these nano-objects grown under different conditions.

The considerable interest in GeSi nanostructures grown on silicon is due to several reasons, one of which is the extraordinary activity during the last decade on the development of the physical basics of Si optoelectronics based on nanostructures. Microelectronics has already met significant limitations in further increasing the operating speed of integrated circuits and optical fiber communication line performance. The gravity of plans related to future Si optoelectronics is represented in particular in the review by Kimerling [4], where Si optoelectronics in the nearest decades is predicted to be a “killer” of modern microelectronics.

Due to its indirect band structure, bulk silicon is not easily applicable for light emission. On the other hand a transfer to nanostructures makes possible band engineering and the quantum confinement of wavefunctions and, as a consequence, a partial abolition of the selection rules, prohibiting direct light-emitting recombination of electron-hole pairs. Today, luminescence from Si-based nanostructures is observed in a wide wavelength range.

In this work we present an analysis of growth features, characterization and optical properties of nanostructures based on GeSi/Si heterostructures obtained by molecular-beam epitaxy (MBE) with electron-beam evaporation of silicon and germanium. Our first observations of the low-temperature photoluminescence at the wavelength $1.55\text{ }\mu\text{m}$ were reported in 1998 [5]. Now, photoluminescence from these nanostructures is observed in the range $1.3\text{--}2\text{ }\mu\text{m}$ up to room temperature [6].

2. Experiment

The GeSi/Si structures under investigation were grown on *p*-Si (001) substrates by solid-source MBE. Silicon and germanium were evaporated by electron-beam evaporators. Si and Ge deposition rates varied as $0.1\text{--}0.15\text{ nm/s}$ and $0.01\text{--}0.015\text{ nm/s}$, respectively. Single layer structures with self-assembled nanoislands consisted of a 200 nm Si buffer layer and Ge layer with an equivalent thickness of 5 to 10 monolayers (ML) ($1\text{ monolayer} \approx 0.14\text{ nm}$). The Ge deposition temperature (T_{Ge}) was varied over the range $460^\circ\text{C}\text{--}750^\circ\text{C}$. Single layer structures with islands grown for photoluminescence measurements had a Si cap layer of $80\text{--}100\text{ nm}$ thickness grown at the same temperatures as used for Ge deposition. The multilayer structures were grown at 600°C , 650°C and 700°C to

as many as five Ge layers with an equivalent thickness of ~ 7.5 ML, separated by 30 nm, 40 nm and 60 nm Si spacer layers, respectively.

Investigations of the sample surface morphology were performed with the atomic force microscope (AFM) “Solver” P4. X-ray diffraction measurements were performed with the double-crystal diffractometer DRON-4. Raman spectra from structures with islands were recorded at room temperature by the Jobin Ivon 64000 Raman microprobe system with 488 nm wavelength of exciting radiation. The photoluminescence (PL) spectra were recorded by the Fourier-spectrometer BOMEM DA3.36 utilizing cooled Ge and InSb detectors. An Ar^+ laser (514.5 nm line) was used for PL excitation.

3. Results and Discussion

3.1 Dependence of Parameters of Uncapped Islands on Ge Deposition Temperature

Analysis of AFM images of the grown structures (Fig. 1) has shown that the island surface density increases monotonically with decreasing T_{Ge} [Fig. 2(a)], but the dependence of the average island height on T_{Ge} has a singularity in the range 550°C – 600° [Fig. 2(b)]. One can distinguish two temperature intervals in the AFM images and the dependence of the island height on T_{Ge} [Fig. 2(b)]: $T_{\text{Ge}} \geq 600^\circ\text{C}$ and $T_{\text{Ge}} < 600^\circ\text{C}$.

3.2 Sizes, Shape and Surface Density of Islands

The results of AFM investigations of single layer structures with islands grown over the range $T_{\text{Ge}} = 460^\circ\text{C}$ – 750°C show (Figs. 1 and 2) that T_{Ge} determines the island size, shape and surface density. By varying the only growth parameter (T_{Ge}), it is possible to change the surface density of the islands by more than 3 orders of magnitude, their height nearly 30-fold, and to form nanoislands with low surface density at high T_{Ge} and “quantum dots” with a surface density $> 10^{11} \text{ cm}^{-2}$ at low T_{Ge} .

It was shown earlier [7,8] that island growth at $T_{\text{Ge}} \geq 600^\circ\text{C}$ begins at an equivalent Ge thickness $d_{\text{Ge}} = 4$ – 5 ML, with the formation of pyramid islands with a small tilt angle ($\sim 11^\circ$) between the base and the lateral faces. With increasing Ge deposition the pyramid islands grow in volume, retaining their shape. After reaching some critical volume they transform into dome islands with a larger tilt angle between the base and the lateral faces than in pyramid islands.

The transformation of island shape is associated with more effective elastic strain relaxation in dome islands than in pyramid islands due to the larger aspect ratio [9]. The maximum (critical) size of pyramid

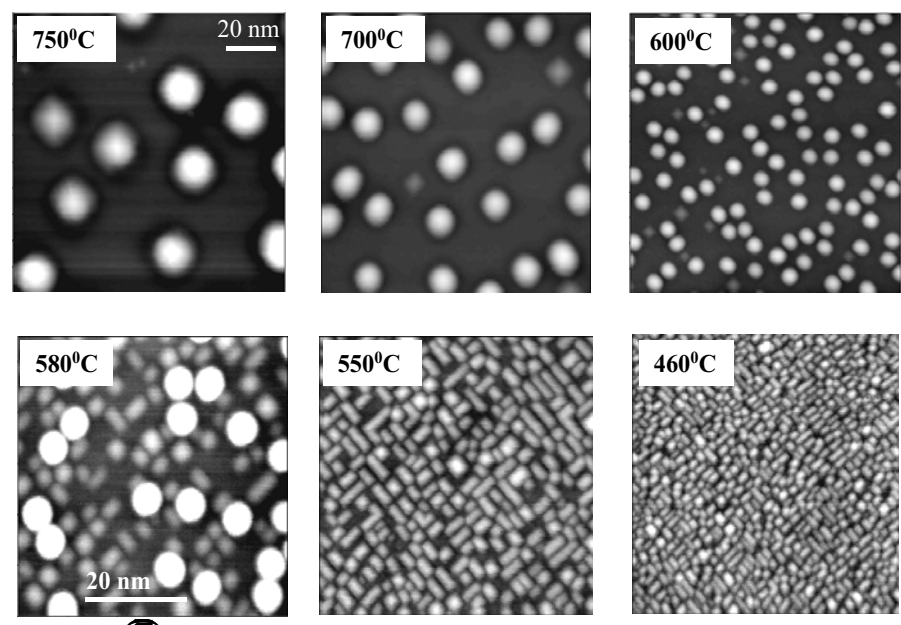


Figure 1. AFM images of GeSi self-assembled islands grown at different temperatures and equivalent thicknesses $d_{\text{Ge}} = 7\text{--}8$ ML of Ge deposition. The image scale is different for structures grown at $T_{\text{Ge}} \geq 600^\circ\text{C}$ and at $T_{\text{Ge}} < 600^\circ\text{C}$.

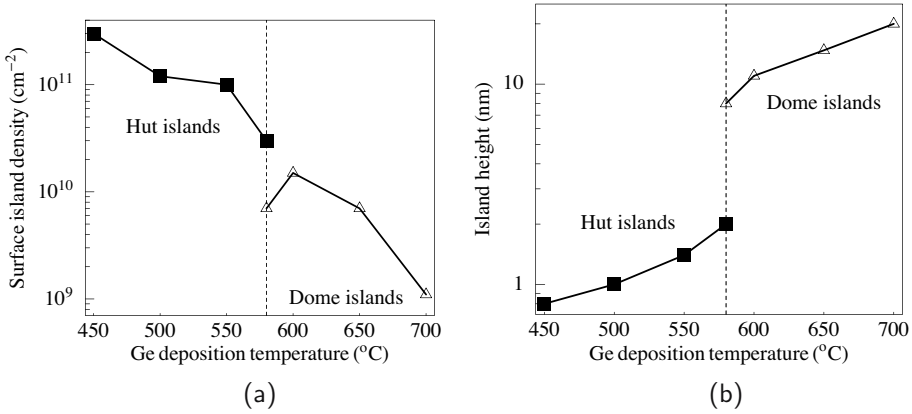


Figure 2. Dependence of (a) the surface density of islands and (b) their average height on the Ge deposition temperature. The dotted line at $T_{\text{Ge}} = 580^\circ\text{C}$ separates the domains for hut and dome islands.

islands decreases with decreasing T_{Ge} , for reasons that are discussed below. AFM image analysis of the structures grown at $T_{\text{Ge}} \geq 600^\circ\text{C}$ has shown that the growth of dome islands proceeds mainly through an increase of their height. Since the maximum lateral size of pyramid islands decreases as T_{Ge} is lowered, the dome island size also decreases (Figs. 1 and 2). The growth of the dome island volume slows down considerably once their aspect ratio has reached 0.2–0.25. At $d_{\text{Ge}} = 7\text{--}8$ ML we have observed pyramid islands and islands of some intermediate shape on the surface of structures grown at $T_{\text{Ge}} \geq 600^\circ\text{C}$, together with dome islands which dominate at given conditions (Fig. 1). The transformation of pyramid islands to dome islands ceases at $d_{\text{Ge}} = 9\text{--}10$ ML, with most of the dome islands having reached their ultimate size. As a result, there is a much smaller spread in both lateral sizes and heights of the islands. Dome islands with a small ($\leq 10\%$) size dispersion were observed for the structures with $d_{\text{Ge}} = 9\text{--}10$ ML grown in the temperature range $T_{\text{Ge}} = 600^\circ\text{C}\text{--}700^\circ\text{C}$ [10].

As the Ge deposition temperature is reduced from 600°C to 550°C , the growth of self-assembled islands undergoes a qualitative change (Figs. 1 and 2). A decrease of T_{Ge} by just 50°C causes a 5-fold decrease of island height [Fig. 2(b)]. This abrupt decrease of island sizes is due to a change of the island type that dominates on the surface at the given ($d_{\text{Ge}} = 7\text{--}8$ ML) equivalent thickness. As mentioned above, the prevailing type of islands at $d_{\text{Ge}} = 7\text{--}8$ ML is the dome island with nanometer size. According to the AFM images in Fig. 1, there are no dome islands on the surface of the structures grown at $T_{\text{Ge}} \leq 550^\circ\text{C}$. The prevailing type here is the hut island [2] having a small height (< 2 nm) and a rectangular base extended in the $\langle 100 \rangle$ or $\langle 010 \rangle$ directions. Hut islands appear on the surface at $T_{\text{Ge}} = 580^\circ\text{C}$ [11]. All three types of islands, dome, pyramid and hut, were observed at this Ge deposition temperature. The height of hut islands decreases monotonically with T_{Ge} from 550°C to 460°C , but their surface density increases (Fig. 2). The hut island height at $T_{\text{Ge}} = 460^\circ\text{C}$ rises to 0.7–1 nm, the lateral size varies over the range 15–30 nm, and the surface density reaches the value $2.5 \times 10^{11} \text{ cm}^{-2}$.

3.3 Composition and Elastic Strain in Islands

AFM investigations of GeSi/Si(001) self-assembled islands have shown [Fig. 2(b)] that the height of a dome island is more than an order of magnitude higher than the critical thickness of coherent growth of a two-dimensional (2D) layer of pure Ge on Si(001), which is ~ 1 nm [12]. This means that either dome islands have had almost complete relaxation of their elastic strain, or they consist of a SiGe alloy rather than pure

Ge. Studies of island structures with a variety of methods have shown [10,13,14] that the latter supposition is correct. The position of a weak peak from the islands in the (004) $\omega/2\theta$ X-ray diffraction spectra of the structures [Fig. 3(a)], and lines appearing in the Raman spectra in relation to GeSi and SiSi atom vibrations in the islands [Fig. 3(b)], all indicate the formation of a GeSi alloy in the islands during deposition of pure Ge on Si(001) at $T_{\text{Ge}} \geq 600^\circ\text{C}$. The formation of the GeSi alloy in the islands is associated with a bulk and surface strain driven diffusion of Si into the islands [13,14]. It should be noted that for structures grown at $T_{\text{Ge}} \geq 600^\circ\text{C}$, the intensity of the X-ray diffraction peak from the islands and of the island-related lines in the Raman spectra is enhanced with an increase of the total volume and surface density of dome islands [10]. This fact allows us to relate the observed signal in the X-ray diffraction and Raman spectra to the dome islands.

With a decrease of T_{Ge} , the X-ray peak from the islands shifts from the diffraction area of Si substrate toward the diffraction area of Ge layer [Fig. 3(a)] and the Raman spectra exhibit a noticeable increase of an intensity of the GeGe line while the SiSi line is getting less intensive. These changes in the X-ray diffraction and Raman spectra correspond to an increase of Ge content in the islands with a decrease of growth temperature. Specific values of composition and residual elastic strain of the islands were obtained by X-ray analysis of the reciprocal space maps around the (004) and (440) reciprocal lattice points from the Si substrate, and from the frequency dependence of different Raman modes

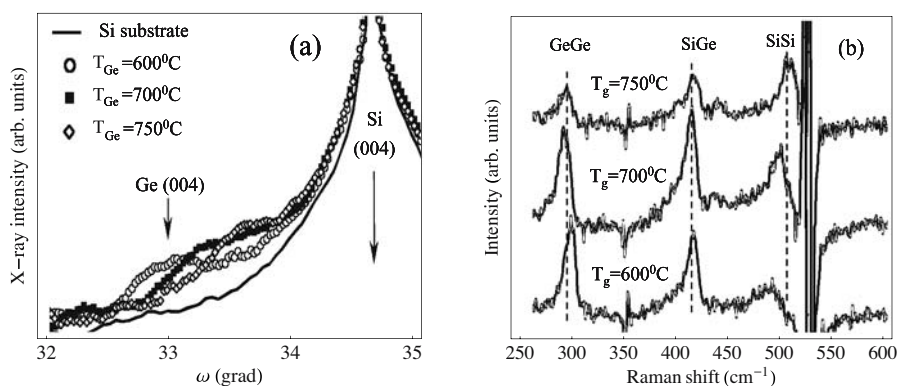


Figure 3. (a) (004) $\omega/2\theta$ X-ray diffraction spectra around the reflection from a Si substrate and (b) Raman spectra of the structures grown at different T_{Ge} . The spectrum from the Si substrate is subtracted from the Raman spectra of the structures with islands. The arrows in figure (a) show the position of the peak from the Si substrate and the calculated peak position for unstrained Ge layer.

on composition and elastic strain of the islands [8,10]. The values of Ge content and elastic strain in the islands, obtained by two methods, agreed within an experimental error. The value obtained for the residual elastic strain was $50\% \pm 10\%$ and was found to be independent of T_{Ge} . This is due to the fact that the dome island shape (aspect ratio), which largely determines the degree of the elastic strain relaxation, does not depend on Ge deposition temperature.

The experimental dependence of the average composition of dome islands on Ge deposition temperature is shown in Fig. 4(a). It is readily seen that the Si content in uncapped dome islands grown at $T_{\text{Ge}} = 600^\circ\text{C}$ amounts to 25%–30% and increases up to 55%–60% for the structures grown at $T_{\text{Ge}} = 750^\circ\text{C}$. As mentioned above, such a high content of silicon is due to Si diffusion enhanced by nonuniform elastic strain fields from the islands [13,14]. The AFM analysis of islands growth at $T_{\text{Ge}} \geq 600^\circ\text{C}$ (these results are discussed in Sec. 3.2) have shown that a general reduction of an island size with a lowering of the growth temperature is caused by a decrease of the critical volume of pyramid islands at which they transform into dome islands. A possible reason for this may be the observed dependence of island composition on Ge deposition temperature. According to the model of the island shape transformation suggested in Ref. [9], the increase of Ge content observed in the islands with a lowering of T_{Ge} [Fig. 4(a)] should cause a decrease in the critical volume of pyramid islands.

The results of the study into the dependence of sizes, shape and composition of GeSi/Si(001) self-assembled islands on growth temperature in the range $T_{\text{Ge}} = 600^\circ\text{C} - 750^\circ\text{C}$ can be presented as a phase diagram of the dependence of island shape on its volume and composition [Fig. 4(b)].

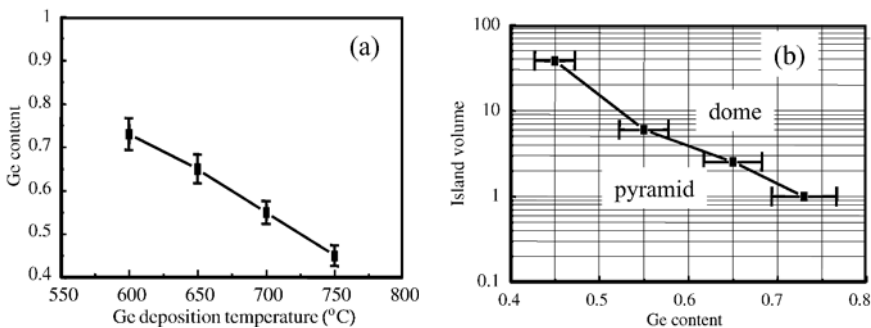


Figure 4. The dependence of Ge content in dome islands on Ge deposition temperature (a) and the relation between the dome island volume and Ge content in them (b). Vertical and horizontal intervals show the error in determining of Ge content in islands. The lines connecting experimental dots are guides for the eye.

The experimental dependence of the average volume of a dome island, as calculated from AFM images, on its composition is given in Fig. 4(b). The dome island volume is normalized to the average volume of dome islands formed at $T_{\text{Ge}} = 600^\circ\text{C}$, which is $\sim 8700 \text{ nm}^3$. It is believed that islands with volume and composition values that are close to the experimental curve have a dome shape. Those with volume and composition values that are far below the experimental curve have a pyramid shape. The exact position of the line corresponding to the onset of the pyramid-to-dome shape transformation is unknown because there are no data on the Ge content in pyramid islands that have reached the critical volume.

The weak intensity and large width of the X-ray diffraction peak from hut islands, being due to their small height, make determination of the composition of islands grown at $T_{\text{Ge}} < 600^\circ\text{C}$ by the methods of X-ray diffraction impossible. The small height of hut islands also prevents an estimation of their composition from Raman spectra. This happens because the position of the line from hut islands in the Raman spectra is affected not only by composition and elastic strain of the islands but also by their size [15]. Although in this work we have not obtained exact values for the island composition at $T_{\text{Ge}} < 600^\circ\text{C}$, one may suggest that the Ge content in such islands is higher than 70%–75%, typical for Ge content in islands grown at $T_{\text{Ge}} = 600^\circ\text{C}$, with the tendency to increase by further lowering the Ge deposition temperature.

4. Structures with a Si Cap Layer

4.1 Composition and Elastic Strain in Overgrown Islands

In the previous section we presented the results of an analysis of the structures with uncapped GeSi/Si(001) self-assembled islands. However, most practical applications need structures with a Si cap layer. It is well known that during overgrowth of islands with Si at $T > 390^\circ\text{C}$, the island height decreases with the average content of Si growing higher [16,17]. X-ray diffraction measurements were carried out on single and multilayer structures grown at $T_{\text{Ge}} = 600^\circ\text{C}$, 650°C and 700°C to obtain the values of composition and elastic strain in overgrown islands.

It was found out that overgrowth of islands with Si in the single-layer structures causes reduction of the intensity and broadening of the island-related X-ray diffraction peak. These effects are due to a decrease of the island height through overgrowth and to greater inhomogeneity of composition and elastic strain in capped islands. The low intensity and large width of the X-ray diffraction peak do not allow the average

Ge content or residual elastic strain to be determined in single layer structures with overgrown islands.

However, in multilayer structures with islands, the intensity of the X-ray diffraction peak related to islands is increased by adding the signal from different layers so that it becomes sufficient for determination of island parameters. The $\omega/2\theta$ X-ray diffraction spectrum near the (004) Si reflection for a multilayer structure (5 layers) with islands grown at $T_{\text{Ge}} = 600^\circ\text{C}$ is shown in Fig. 5(a). For comparison, the X-ray diffraction spectrum for a single layer structure with uncapped islands is also shown. Besides the satellite peaks related to the lattice of the wetting layers, the X-ray diffraction spectrum from the multilayer structure exhibits a wide peak associated with the islands. As compared to the signal from uncapped islands, the peak from the islands in the multilayer structure is shifted towards that from the Si substrate, which indicates a higher average Si content in the islands of multilayer structures.

Similar to the case of X-ray diffraction measurements of uncapped islands, the real values of the Ge content and residual elastic strain in the islands have been obtained by an X-ray analysis from the reciprocal space maps around the (004) and (440) reciprocal lattice points of the Si substrate. The values of the average Ge content in the islands of multilayer structures were found to be $\sim 20\%$ smaller than the Ge content in uncapped islands [Fig. 5(b)]. There could be two reasons for the drastic decrease of Ge content in the islands of multilayer structures. First, there is Si diffusion into the islands during overgrowth [17]. The second reason is the accumulation of the elastic strain through the growth of

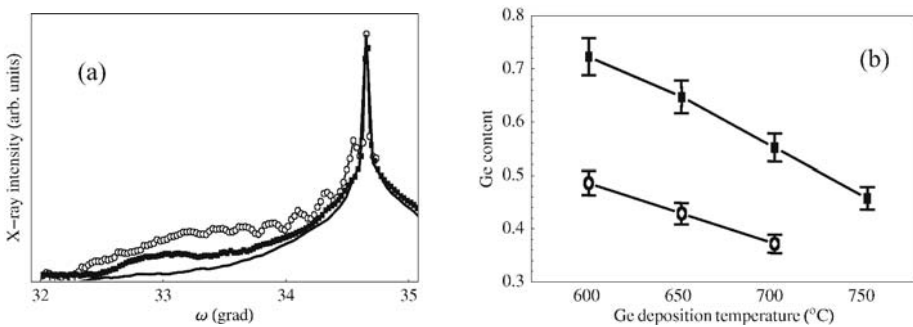


Figure 5. (a) $\omega/2\theta$ X-ray diffraction spectra near Si (004) reflection for Si substrate (solid line), single layer structure with uncapped islands (open circles), and multilayer structure (filled squares), grown at $T_{\text{Ge}} = 600^\circ\text{C}$. (b) Dependence of Ge content in islands on Ge deposition temperature for structures with uncapped islands (filled squares) and for multilayer structures (open circles). Error bars in (b) show the error in Ge content in the islands.

a multilayer structure with islands. The latter effect enhances Si diffusion into the islands and, as a consequence, causes the Si content to increase even in uncapped islands in the second and subsequent layers of the structure [18]. The residual elastic strain in islands obtained by the X-ray analysis was independent of the growth temperature and equal to $80\% \pm 20\%$. The higher value of the residual elastic strain in capped as compared to the uncapped islands ($50\% \pm 10\%$) is accounted for by the disappearance, through the Si cap layer growth, of the free surface area of the island, which is responsible for relaxation of most of the elastic strain in uncapped structures.

4.2 AFM Microscopy of Cleavage Structures with GeSi Islands

The structures were cleaved along the [111] and [110] crystallographic directions. Prior to the chipping, the structures were thinned by substrate polishing to a thickness of less than $100\text{ }\mu\text{m}$ in order to obtain an atomically smooth surface free from step-like formations. AFM images of the cleavages have exposed the features of their topography such as local elevations and depressions of the surface (Fig. 6). The distance between the local elevations on the cleavage surface of the single layer structures and the cleavage edge coincides with the thickness of the Si cap layer [Figs. 6(a) and 6(b)]. Cleavages of the multilayer structures display a periodic alternation of local elevations and depressions of the surface. The number of local elevations on the cleavages coincides with the number of periods in the multilayer structure, and the distance between the neighboring local elevations is equal to the thickness of a Si space layer between the neighboring layers with islands.

The above facts allow one to attribute the observed features of cleavage topographic to the elastic strain relaxation in the layer of GeSi self-assembled islands and their vicinity. This interpretation of topography features on the surface of cleavages of the single layer structures with GeSi islands and of their relation to the elastic strain relaxation of GeSi islands and the surrounding Si matrix are shown schematically in Fig. 6(e). When the structure is cleaved right across a GeSi island, the relaxation of the elastic compressive strain in the island causes a local elevation above the originally smooth free surface of the cleavage. At the same time, the relaxation of the elastic tensile strain in the Si matrix surrounding the island leads to a lowering of the island-adjacent areas on the cleavage surface. As a result, when the cleavage passes across a GeSi island, the elastic strain relaxation gives rise to a local elevation on the cleavage surface, which is surrounded from both sides by local de-

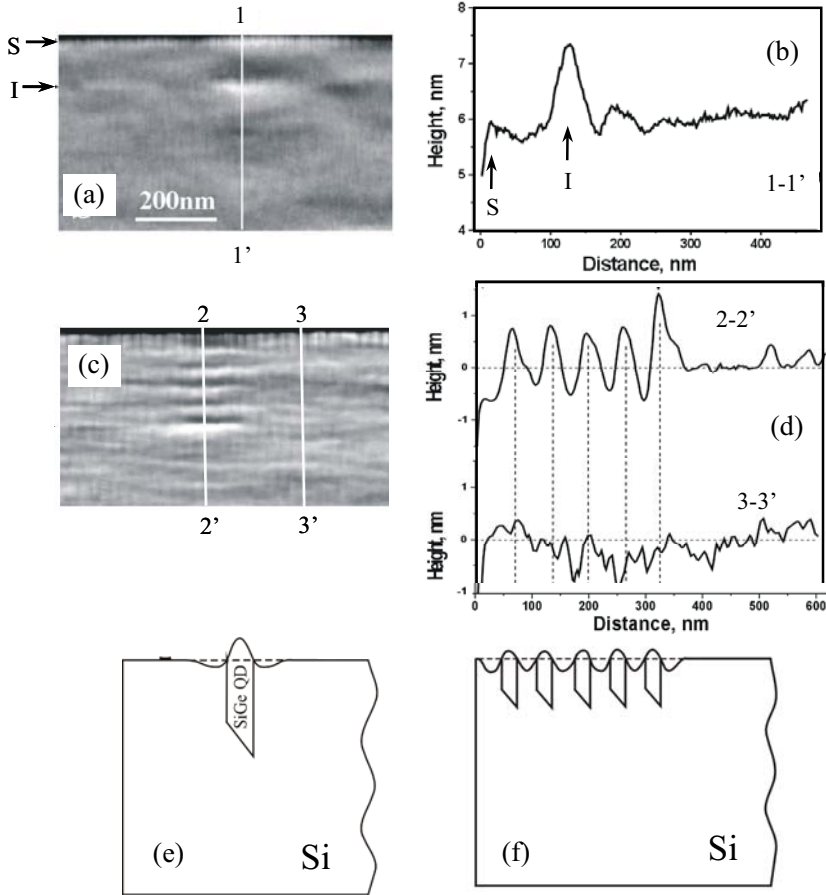


Figure 6. (a), (c) AFM images, and (b), (d) topography of cleavages along the lines shown on the AFM images for single and multilayer structures with islands, respectively. (e), (f) Schematic representation of the topographic features on a cleavage surface as originating from the elastic strain relaxation in islands and in Si matrix for single (e) and multilayer (f) structures.

pressions of the surface [Figs. 6(b,e)]. The AFM cleavage images of the multilayer structures with islands exhibit periodicity in the growth direction for the topography features related to the islands [Figs. 6(c,d)]. This periodicity is associated with a vertical correlation of islands in neighboring layers of the structure [Fig. 6(f)], arising through the propagation of the elastic strain fields from the islands in the lower layers of a multilayer structure into the upper layers [19]. A local lowering of the cleavage surface between the island-related topography features indicates strain in the Si layers separating the layers with islands (Fig. 6).

This lowering of the cleavage surface occurs through the relaxation of the elastic tensile strain in Si layers.

5. Photoluminescence of Structures with GeSi Islands

5.1 Single Layer Structures

In this section we present results of investigations of the photoluminescence of single layer structures with islands ($d_{\text{Ge}} = 7\text{--}8$ ML) grown over the range $T_{\text{Ge}} = 460^\circ\text{C}\text{--}700^\circ\text{C}$. The structures used for the photoluminescence measurements had a Si cap layer of 80–100 nm thickness grown at T_{Ge} in this range.

Measured PL spectra, except for expected substrate-related spectral lines and signals from Si buffer and cap layers, had a wide peak at low energies for all studied structures (Fig. 7) associated with a real-space indirect optical recombination of holes localized in islands and electrons localized in Si on a type-II heterojunction with the island [5]. This indirect optical transition is shown schematically in Fig. 8. Besides the wide PL band related to islands, PL spectra of the structures grown at $T_{\text{Ge}} \geq 600^\circ\text{C}$ display two spectral lines from a wetting layer [Fig. 7(a)], related to the non-phonon and transverse optical phonon assisted recombination. PL lines originated from the wetting layer are red-shifted and their intensity decreases with a lowering of T_{Ge} . The red shift of PL signal from the wetting layer is caused by suppression of the diffusion process at lower growth temperatures and by increase of Ge content in the 2D layer. A lower PL signal intensity from the 2D layer is associated

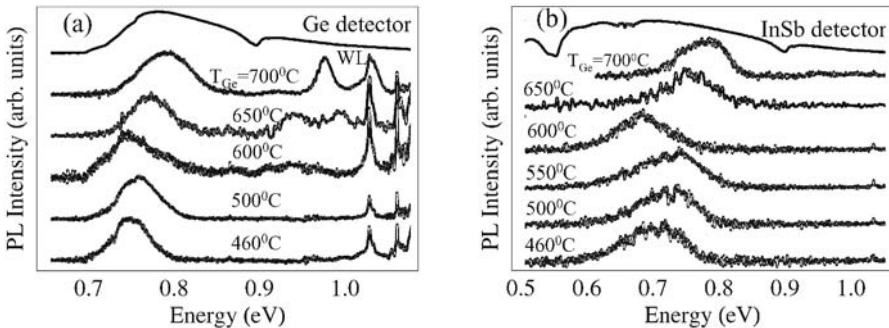


Figure 7. PL spectra recorded by (a) Ge and (b) InSb detectors for structures with islands grown at different Ge deposition temperatures. The upper curves show detector spectral sensitivities. WL indicates lines from the wetting layer. The narrow lines with the energies 1.03 and 1.06 eV are related to the optical recombination in Si layers. PL spectra have been recorded at 4 K. The excitation power was 50 W/cm².

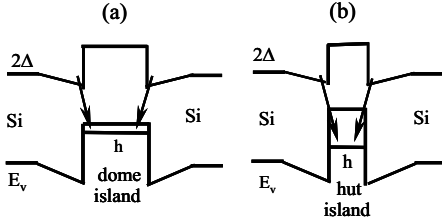


Figure 8. Schematic diagram of the spatially indirect optical transition for (a) dome and (b) hut islands. The valence band edge and the position of 2Δ electron valleys in the islands and their vicinity are depicted on the figure. The arrows show the spatially indirect optical transition.

with an increase of island surface density by an order of magnitude with a lowering of T_{Ge} from 700°C to 600°C [Fig. 2(a)]. When the island surface density grows higher, increasingly more charge carriers are in time to reach the islands and recombine there. The high surface density of the islands in structures grown at $T_{\text{Ge}} < 600^\circ\text{C}$ is responsible for disappearance of PL signal from the wetting layer [11].

PL spectra measured by the Ge detector show [Fig. 7(a)] a non-monotonic dependence of the island-related PL peak position on Ge deposition temperature. However, a comparison of PL spectra from the structures and the spectral sensitivity of the Ge detector show that low energy edge of the PL signal from islands grown at $T_{\text{Ge}} \leq 650^\circ\text{C}$ is determined by low energy limit of Ge detector. To overcome this limitation we used cooled InSb detector with a longer-wave length limit, although less sensitive. PL spectra, recorded with a InSb detector [Fig. 7(b)] show that a considerable part of PL signal from islands in structures grown at $T_{\text{Ge}} \leq 650^\circ\text{C}$ falls within the range of energies lower than the band gap of bulk Ge (0.72 eV at 4 K). This observation supports the model [5] where PL signal from islands is associated with spatially indirect optical transition between holes localized in the islands and electrons localized in Si on type-II heterojunction with the island (Fig. 8).

As seen from Fig. 9, the dependence of the maximum PL peak position from the islands is non-monotonic and has a singularity in the range $T_{\text{Ge}} = 550^\circ\text{C}$ – 600°C . Similar to the AFM investigations of island growth (Sec. 3.2), the experimental dependence can be divided into two parts.

In the first part ($T_{\text{Ge}} \geq 600^\circ\text{C}$) the maximum $E_m(T_{\text{Ge}})$ of the PL peak from islands is monotonically red-shifted with a decrease of the growth temperature. According to results of AFM investigations (Sec. 3.2), the dominating type of islands at $T_{\text{Ge}} \geq 600^\circ\text{C}$ and $d_{\text{Ge}} = 7$ – 8 ML are dome islands (Fig. 1). Therefore, the observed PL signal from the structures with $T_{\text{Ge}} \geq 600^\circ\text{C}$ can be due to optical recombination of charge carriers in the dome islands. There are two main factors affecting the energy of optical transitions in dome islands with a lowering of the growth temperature: the reduction of sizes [Fig. 2(b)] and the change of composition

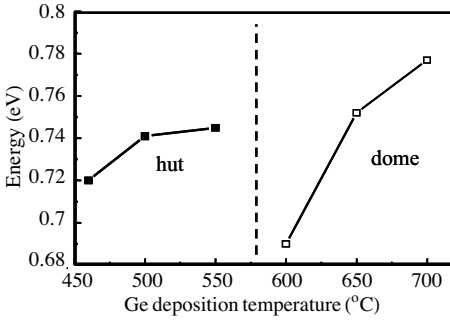


Figure 9. Dependence of the maximum position of PL peak from the islands on the Ge deposition temperature. The dotted line on the plot separates the domains of existence for hut and dome islands and corresponds to $T_{\text{Ge}} = 580^\circ\text{C}$.

[Fig. 4(a)] of the islands. At $T_{\text{Ge}} \geq 600^\circ\text{C}$ the first factor has a weak influence on the energy of optical transitions in islands because the height of uncapped dome islands, even at $T_{\text{Ge}} = 600^\circ\text{C}$, is larger than 10 nm [Fig. 2(b)], and the quantum confinement effects on the position of the first level of quantum confinement of holes in the islands is weak. As a result, the hole energy in dome islands lies near the valence band edge [Fig. 8(a)]. The major factor affecting position of PL peak from the dome islands is a dependence of the island composition on T_{Ge} . The red shift of PL peak from islands is associated with the suppression of Si diffusion into the islands with decreasing growth temperature and, hence, with an increase of Ge content in the islands [6]. The buildup of Ge content in islands causes an increase in the valence band offset on the silicon-island heterojunction and, consequently, a decrease of the energy of indirect optical transition [5] [Fig. 8(a)]. A signal from islands of the structures grown at $T_{\text{Ge}} = 600^\circ\text{C}$ was observed down to an energy of 0.6 eV ($\sim 2 \mu\text{m}$) [6].

As mentioned above, $E_m(T_{\text{Ge}})$ has a singularity in the range $T_{\text{Ge}} = 550^\circ\text{C}$ – 600°C . The PL peak from the islands is blue-shifted by ~ 50 meV with T_{Ge} lowering from 600°C to 550°C (Fig. 9). Investigations into the growth of uncapped islands have shown [11] that essential changes in the island morphology occur precisely in this T_{Ge} range: dome islands with a height > 10 nm dominate on the surface of structures for $T_{\text{Ge}} > 580^\circ\text{C}$ and hut islands with a height ≤ 2 nm dominate on the surface of structures for $T_{\text{Ge}} < 580^\circ\text{C}$ [Figs. 1 and 2(b)]. A sharp reduction of an island height enhances quantum confinement effects on the energy levels of holes in the islands. As a result, in spite of greater offset of the valence band on the Si-island heterojunction due to an increase of Ge content in islands with a lowering of T_{Ge} , the first quantum confinement level of holes in hut islands shifts (is “pushed out”) towards the edge of the Si valence band [Fig. 8(b)]. Therefore, the energy of spatially indirect

optical transition in the islands increases, which leads to the observed blue shift of the PL from the islands.

The monotonic decrease of $E_m(T_{\text{Ge}})$ as T_{Ge} is lowered from 550°C to 460°C is associated with the suppression of an island parameter variation during the overgrowth. It is well known [16] that the Si content in the islands increases and their height diminishes during the growth of a Si cap layer. It has been shown in [17] that these processes are largely suppressed as the temperature of the island overgrowth decreases from 500°C to 390°C. Thus, when the growth temperature goes down from 550°C to 460°C, there is an increase in the height and average Ge content of the overgrown islands. Both these processes induce a decrease of the energy of spatially indirect optical transition in islands (Fig. 8) and, consequently, a red shift of the PL peak from the islands.

5.2 Multilayer Structures with Islands

In this section, results of investigations of PL spectra from multilayer structures with GeSi islands are presented. The multilayer structures were grown at $T_{\text{Ge}} = 600^\circ\text{C}$ and 700°C and consisted of 5 periods of Ge layers with an equivalent thickness of 7.5 ML, separated by Si space layers with a thickness of 30 nm and 60 nm, respectively.

A comparative analysis of PL spectra from single and multilayer structures with islands has shown (Fig. 10) that, besides a considerable enhancement of PL signal from the islands of multilayer structures, there is a blue shift of the PL signal from the islands by 15–25 meV. This blue shift is associated with lower Ge content in the islands of multilayer structures than the same of single layer structures. As mentioned above

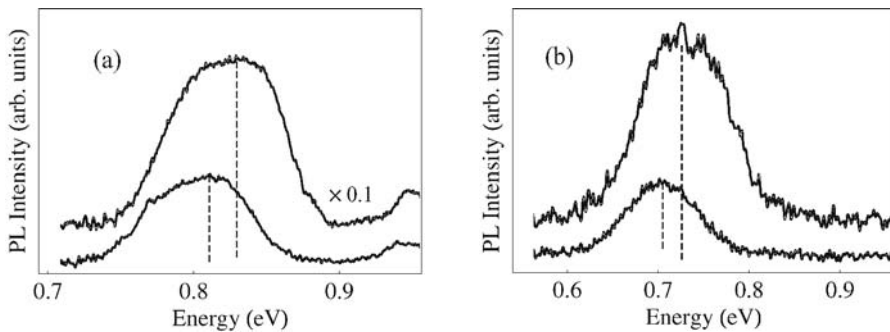


Figure 10. PL spectra for single (lower curves) and multilayer (upper curves) structures with islands grown at (a) $T_{\text{Ge}} = 700^\circ\text{C}$ and (b) $T_{\text{Ge}} = 600^\circ\text{C}$ (b). The PL spectra were recorded at 4 K by (a) Ge and (b) InSb detectors. The excitation power was 50 W/cm².

(Sec. 4.1), the decrease of the average Ge content in the islands of upper layers of multilayer structures could be caused by Si diffusion enhanced by elastic strain. Accumulation of the elastic strain from the islands during the growth of a multilayer structure results in an enhancement of this diffusion component [18]. The low Ge content in the islands of lower layers of multilayer structures is accounted for by the bulk Si diffusion from space layers into the islands during the longer growth time as compared with single layer structures [20]. The widening of the PL peak from the islands in multilayer structures (Fig. 10) could also be related to the difference in island composition in different layers. The width of the PL line from the islands of multilayer structures increases, in spite of the smaller size distribution and space ordering of the islands observed for these structures [19]. Apparently, the width of the PL line in this case is determined by the difference in island composition of different layers rather than the size distribution of islands.

Using methods of X-ray diffraction we were able to measure the average Ge content and residual elastic strain in the islands of multilayer structures (Sec. 4.1 and Fig. 5). The energy band position in the islands and the adjacent areas has been calculated based on the values obtained of the composition and residual elastic strain. The layer with islands in the calculations was approximated by a uniformly strained layer of the same composition and residual elastic strain as in the islands. This approximation is possible because, according to the AFM data, the lateral size of the dome-islands exceeds 50 nm and is 5 times their height (Fig. 1). In our calculations we took into account the change of Si band structure near the island due to the propagation of elastic strain from the islands into Si layers (Fig. 8). According to our calculations, the energy of optical transition between electrons and holes localized in the islands (direct in real space) is $0.83 \text{ eV} \pm 0.05 \text{ eV}$ at $T_{\text{Ge}} = 600^\circ\text{C}$ and $0.89 \text{ eV} \pm 0.04 \text{ eV}$ at $T_{\text{Ge}} = 700^\circ\text{C}$. The calculated values are higher than the energies of the PL signal from the islands. At the same time, the calculated values of the energy of spatially indirect optical transition between holes confined in the islands and electrons localized in Si matrix on a heterojunction with islands (Fig. 8) are $0.76 \text{ eV} \pm 0.05 \text{ eV}$ for $T_{\text{Ge}} = 600^\circ$ and $0.86 \text{ eV} \pm 0.04 \text{ eV}$ for $T_{\text{Ge}} = 700^\circ\text{C}$. These energy data are in a good agreement with the experimentally observed energy position of the PL peaks from the islands in multilayer structures (Fig. 10), which confirms the model [5] associating the island-related photoluminescence signal with spatially indirect optical recombination.

Calculations of the band structure for samples with islands, taking into account the experimentally obtained values of the composition and residual elastic strain in the islands, showed that the valence band off-

set at the silicon-island heterojunction depends on T_{Ge} , increasing from 250 meV at 700°C to 350 meV at 600°C. These energies are an order of magnitude higher than the thermal energy of charged carriers at room temperature (~ 25 meV). A considerable number of holes, therefore, remains localized in the islands even at room temperature, which allows observation of the PL signal from the islands up to room temperature (Fig. 11). As at low temperatures, a considerable part of the PL signal from islands in multilayer structures grown at $T_{\text{Ge}} = 600^\circ\text{C}$ falls within a range of energies lower than the band gap of bulk Ge (curve 1 in Fig. 11) [6]. The intensity of the PL signal from islands grown at 600° is about 10 times higher than that for the islands grown at 700° (Fig. 11). This is due to better localization of holes in the islands grown at the lower temperature and, hence, having smaller sizes, higher Ge content and a larger valence band offset on the silicon-island heterojunction. A decrease of Ge content in islands with increasing growth temperature also causes a reduction of the potential well for electrons located in Si on the heterojunction with islands (Fig. 8). The potential well for electrons is induced by the penetration of the elastic strain from the islands into Si layers, its depth increasing with larger lattice mismatch between the Si matrix and the island. Since, according to the X-ray data, the residual elastic strain does not depend on the growth temperature, the mismatch is greater for islands with higher Ge content. Hence, the potential well for electrons is larger near the islands with a higher Ge content.

Comparison of PL spectra from structures with islands recorded at 4 K and 300 K (Figs. 10, 11) shows that the energy position of the island-related PL peak has a weak dependence on the measurement temperature. This is because the band gap narrowing with increasing measurement temperature is compensated by an increase of the hole population in the excited levels in the islands. The optical transition of holes from these levels, having a higher energy, accounts for the weak temperature dependence of the energy position of the island-related PL peak.

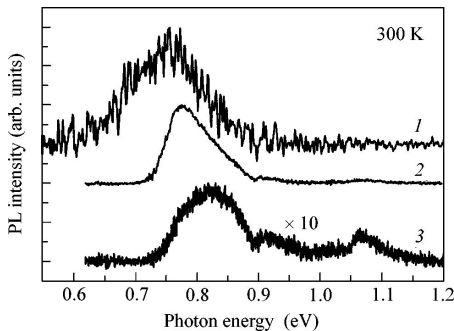


Figure 11. PL spectra for multilayer structures with islands grown at 600°C (curves 1 & 2) and 700°C (curve 3) recorded at room temperature by InSb (curve 1) and Ge (curves 2 & 3) detectors. All spectra were recorded at excitation power of 50 W/cm².

6. Conclusion

The investigations of the temperature-dependent growth of GeSi/Si(001) self-assembled islands has been performed. The investigated parameters of the GeSi/Si(001) islands include size, shape, surface density, composition, elastic strain, and photoluminescence (PL) properties. It has been shown that by varying the Ge deposition temperature one can change the sizes and surface density of the islands over a wide range, thus forming nanometer scale islands at high temperatures and quantum dots of ~ 1 nm height and over 10^{11} cm $^{-2}$ surface density at low temperatures.

Single and multilayer structures with GeSi/Si(001) self-assembled islands having a PL signal at 1.4–2 μ m up to room temperature have been grown. A wide PL peak observed in Ge(Si)/Si(001) structures with islands in the energy range 0.6–0.85 eV (1.4–2 μ m) is associated with spatially indirect optical transition between holes confined in the islands and electrons localized in Si matrix on a type-II heterojunction with islands. It has been shown that the red shift of the PL peak from the islands is due to an increase of Ge content in islands for structures grown at temperatures $\geq 600^\circ\text{C}$. The island-related PL peak position was found to have a non-monotonic dependence on the Ge deposition temperature, which is associated with a change of island morphology over a Ge deposition temperature range of 550°C – 600°C .

Acknowledgements

This work was supported by RFBR Grant No. 02-02-16792, Programs of the Ministry for Industry and Science, Grant INTAS NANO No. 01-444, BRHE Program and Scientific and Educational Center for Physics of Solid-State Nanostructures of Nizhny Novgorod State University.

References

- [1] D. J. Eaglesham and M. Cerullo, Dislocation-free Stranski-Krastanow growth of Ge on Si(001). *Phys. Rev. Lett.* **64**: 1943–1946, 1990.
- [2] Y.-M. Mo, D. E. Savage, B. S. Swartzentruber and M. G. Lagally. Kinetic pathway in Stranski–Krastanov growth of Ge on Si(001). *Phys. Rev. Lett.* **65**: 1020–1023, 1990.
- [3] A. I. Yakimov, V. A. Markov, A. V. Dvurechenskii and O. P. Phelyakov. Coulomb staircase in a Si/Ge structures. *Phil. Mag. B* **65**: 701–705, 1992.
- [4] L. C. Kimerling. Silicon microphotonics: The next killer technology. in *Towards the First Silicon Laser*, L. Pavesi, S. Gaponenko and L. DalNegro, editors (Kluwer, Dordrecht, 2003), pp. 465–476.
- [5] V. Ya. Aleshkin, N. A. Bekin, N.G. Kalugin, Z. F. Krasilnik, A. V. Novikov, V. V. Postnikov and H. Seyringer. Self-organization of germanium nanoislands in silicon obtained by molecular beam epitaxy. *JETP Lett.* **67**: 48–53, 1998.

- [6] N. V. Vostokov, Yu. N. Drozdov, Z. F. Krasil'nik, D. N. Lobanov, A. V. Novikov and A. N. Yablonskii. Low-energy photoluminescence of structures with GeSi/Si(001) self-assembled nanoislands. *JETP Lett.* **76**: 365–369, 2002.
- [7] T. I. Kamins, E. C. Carr, R. S. Williams and S. J. Rosner. Deposition of three-dimensional Ge islands on Si(001) by chemical vapor deposition at atmospheric and reduced pressures. *J. Appl. Phys.* **81**: 211–219, 1997.
- [8] A. V. Novikov, B. A. Andreev, N. V. Vostokov, Yu. N. Drozdov, Z. F. Krasilnik, D. N. Lobanov, L. D. Moldavskaya, A. N. Yablonskii, M. Miura, N. Usami, Y. Shiraki, M. Ya. Valakh, N. Mesters and J. Pascual. Strain-driven alloying: Effect on sizes, shape and photoluminescence of GeSi/Si(001) self-assembled islands. *Mater. Sci. Eng. B* **89**: 62–65, 2002.
- [9] F. M. Ross, J. Tersoff and R. M. Tromp. Coarsening of self-assembled Ge quantum dots on Si(001). *Phys. Rev. Lett.* **80**: 984–987, 1998.
- [10] Z. F. Krasilnik, P. Lytvyn, D. N. Lobanov, N. Mesters, A. V. Novikov, J. Pascual, M. Ya. Valakh and V. U. Yukhymchuk. Microscopic and optical investigation of Ge nanoislands on silicon substrates. *Nanotechnology* **13**: 81–85, 2002.
- [11] O. G. Schmidt, C. Lange and K. Eberl. Photoluminescence study of the 2D-3D growth mode change over for different Ge/Si island phases. *Phys. Stat. Solid. B* **215**: 319–324, 1999.
- [12] J. H. van der Merwe. Crystal interfaces. Part II. Finite overgrowths. *J. Appl. Phys.* **34**: 123–127, 1963.
- [13] S. A. Chaparro, J. Drucker, Y. Zhang, D. Chandrasekhar, M. R. McCartney and D. J. Smith. Strain-driven alloying in Ge/Si(001) coherent islands. *Phys. Rev. Lett.* **83**: 1199–1202, 1999.
- [14] Z. F. Krasil'nik, S. A. Gusev, I. V. Dolgov, Yu. N. Drozdov, D. N. Lobanov, L. D. Moldavskaya, A. V. Novikov, V. V. Postnikov and D. O. Filatov. The elastic strain and composition of self-assembled GeSi islands on Si (001). *Thin Solid Films* **367**: 171–175, 2000.
- [15] C. Sheng, T. Zhou, Q. Cai, Yu. M. Dawei-Gong, X. Zhang and X. Wang. Suppression of Ge-Si interfacial mode in the Raman spectrum of a Si₆Ge₄ superlattice. *Phys. Rev. B* **53**: 10771–10774, 1996.
- [16] P. Sutter, E. Mateeva, J. S. Sullivan and M. G. Lagally. Low-energy electron microscopy of nanoscale three-dimensional SiGe islands on Si(001). *Thin Solid Films* **336**: 262–270, 1998.
- [17] A. Rastelli, E. Muller and H. von Kanel. Shape preservation of Ge/Si(001) islands during Si capping. *Appl. Phys. Lett.* **80**: 1438–1440, 2002.
- [18] O. G. Schmidt, U. Denker, S. Christiansen and F. Ernst. Composition of self-assembled Ge/Si islands in single and multiple layers. *Appl. Phys. Lett.* **81**: 2614–2616, 2002.
- [19] J. Tersoff, C. Teichert and M. G. Lagally. Self-organization in growth of quantum dot superlattices. *Phys. Rev. Lett.* **76**: 1675–1678, 1996.
- [20] X. Z. Liao, J. Zou and D. J. H. Cockayne. Annealing effect on the microstructures of Ge/Si(001) quantum dots. *Appl. Phys. Lett.* **79**: 1258–1260, 2001.

QUANTUM DOTS IN HIGH ELECTRIC FIELDS: FIELD AND PHOTOFIELD EMISSION FROM Ge NANOCCLUSERS ON Si(100)

A. A. Dadykin and A. G. Naumovets

National Academy of Sciences of Ukraine Institute of Physics, 46 Prospect Nauki, UA-03028, Kiev, Ukraine

Yu. N. Kozyrev and M. Yu. Rubezhanska

Institute of Surface Chemistry, 17 Gen. Naumova Str., UA-03680, Kiev, Ukraine

Yu. M. Litvin

Institute of Semiconductor Physics, 45 Prospect Nauki, UA-03028, Kiev, Ukraine

Abstract A stable field electron emission was obtained from Ge nanocluster structures grown on Si(100) by molecular-beam epitaxy. The size of the clusters was ≤ 10 nm and cluster density was $\sim 10^{10}$ cm $^{-2}$. The emission current-voltage characteristics showed current peaks, presumably due to resonance electron tunneling *via* the energy levels of the nanocluster potential well. For Ge cluster multilayers embedded in Si, the field emission current showed a considerable temperature sensitivity, as well as photosensitivity in the wavelength range from 0.4 to 10 μ m.

1. Intr n

Ge-Si nanostructures attract much attention of both physicists and technologists as systems that allow introduction of band gap engineering into silicon technology. This opens promising ways for the development of new Si-based high-frequency and optoelectronic devices [1–6]. We present here experimental results that seem to reveal quantum size effects in field electron emission from Ge quantum dots (QDs) grown on the Si(100) substrate. The application of the external electric field to

the QDs creates new conditions for electron confinement, which can be exploited to some practical use.

Starting from 1960s, the field emission from semiconductor cathodes was widely investigated [7–12]. The aim was to study the physical processes that occur in semiconductors under high (~ 1 V/nm) electric fields and to develop new electronic devices, especially photoemission detectors for infrared (IR) range. In those studies, semiconductor single-tip field emitters and tip arrays were prepared by chemical etching and other techniques producing the cathodes with typical apex curvature radius of ~ 100 nm.

With the progress in nanotechnologies new possibilities were opened for the easy creation of nanosized objects which, due to their geometry, ensure the realisation of electric fields ~ 1 – 10 V/nm at moderate voltages $\sim 10^3$ – 10^4 V. Such fields are necessary to make the surface potential barrier sufficiently transparent for electrons. It was thus an appealing idea to examine Ge nanoclusters grown on Si as potential field emitters. To our knowledge, the first field emission measurements with this system were reported in 2000 by Tondare *et al.* [13]. They obtained a rather stable current of ~ 40 μ A at 12.5 kV from a Ge nanocluster array nearly 1 cm² in area. We also investigated such emitters [14] and found interesting peculiarities in the emission current-voltage (CV) characteristics that had not been considered in Ref. [13], i.e. the existence of current maxima in the CV curves that are observed in some range of the cluster sizes.

2. Experiment

2.1 Preparation of Samples

The Ge-Si(100) nanostructures investigated in this work were grown by molecular-beam epitaxy (MBE). The substrates were *n*-Si plates with the resistivity of 7.5 and 20 Ohm-cm. The accuracy of the (100) surface orientation was better than 1° . The Si and Ge evaporators were heated by electron bombardment. To check the chemical composition of the molecular beams, we used a quadrupole mass-spectrometer. The elemental composition of the deposited layers was monitored by Auger spectrometry. The background pressure in the MBE chamber was $\leq 10^{-10}$ Torr. The pressure increased by about 3×10^{-10} Torr when Si and Ge evaporators were operated.

The growth rate was determined by RHEED (reflection high-energy electron-diffraction) from the intensity oscillations of the central reflected beam. One oscillation period marks the formation of one full Ge or Si monolayer about 0.135 nm thick. The growth rates used in our experi-

ments were 0.02 to 0.15 nm/s for Si and 0.01 to 0.05 nm/s for Ge. The substrate temperatures T_s were varied from 300 to 850° C.

The preparation of the nanostructure samples was started with desorption of the passivating Si oxide film from the Si plate, which was carried out at $T_s = 800^\circ$ C in the presence of a Si beam of weak intensity directed to the surface. Then the surface was covered with a buffer Si film about 0.1–0.5 μm thick. This process resulted in formation of a high-contrast Si(100)–(2×1) RHEED pattern characteristic of clean Si. Since the lattice misfit between Ge and Si amounts to 4.2%, the growth of Ge films on Si occurs by the Stranski–Krastanov mechanism, in which a layer-by-layer growth changes for three-dimensional (3D) growth at some critical thickness corresponding to a few monolayers [6, 15, 16]. This allows the minimisation of the elastic strain energy in the Ge-Si heterojunction. To achieve an appropriate ordering of Ge film, the substrate should be kept at temperature that ensures a sufficient mobility of Ge atoms. If this temperature exceeds 500–600° C, the growth of Ge film on Si is accompanied by rather intense intermixing of Ge and Si, which gives rise to formation of a $\text{Ge}_x\text{Si}_{1-x}$ alloy. It is actually this alloy that serves as a substrate to the growing Ge film.

The lattice periodicity and the band structure of the alloy can be varied by varying x , which gives a favourable possibility to tailor the Ge-Si interface in the sense of the distribution of strain fields and band structure engineering. Using also such parameters as substrate temperature and deposition rate, one can rather effectively control the size structure, composition and density of the nanoclusters grown on the Si substrate. In the case when the lattice misfit between the clusters and the substrate surface is not too high, the clusters may relieve the stress without forming dislocations [6, 15, 16]. In order to achieve this, we first covered the Si surface with a series of intermediate $\text{Ge}_x\text{Si}_{1-x}$ layers (SIL), in which the mole content of Ge was increased step by step from 0 to about 0.6. The number of the intermediate layers in our experiments was varied from 3 to 10, and their total thickness lay in the range 50–180 nm. The depth profile of the composition was recorded by Auger spectroscopy. The substrate temperature during deposition of successive intermediate layers was decreased step by step from 800° C to 350–400° C. More details on the preparation procedure are given in Refs. [17, 18].

As the Ge content was increased, the diffraction pattern (2×1) changed gradually to the pattern (1×1), and the diffraction rods from the flat surface transformed into point-like reflections from 3D clusters. The geometry of the clusters and their distribution over the substrate was investigated by atomic force microscopy (AFM). The AFM observations showed that the deposition of Ge onto $\text{Ge}_x\text{Si}_{1-x}$ surfaces with

higher x values gives more uniformly sized and densely arranged nanoclusters. In the structures studied here the height of the clusters was ≤ 10 nm and their surface density comprised 10^{10} – 10^{11} cm $^{-2}$. Such samples were used in the measurements of field emission.

To investigate the photofield emission, we prepared multilayer Ge-Si nanocluster structures, which provide a higher photosensitivity. With this aim the procedure elaborated for growing a monolayer of Ge nanoclusters was repeated 5 to 10 times. It was found important in this case to keep the SIL thickness between 5 and 15 nm and to catch the stage when the clusters assume the hut shape. We developed earlier a technique of operating Ge and Si molecular beam sources with high stability and switching precision [19]. In this work, the relative accuracy in the periodicity of Ge-Si multilayer nanostructures was maintained at the level of 1–2%. It is known that Ge quantum dots embedded in Si exhibit a strong tendency toward self-ordering and that such 3D quantum dot superlattices manifest interesting electronic and optical properties [3].

2.2 Field and Photofield Emission Measurements

The Ge-Si nanostructures were tested for field and photofield emission in experimental devices of two types. The first of them represented diode cells, in which Ge-Si sample was the cathode and a transparent conductive SnO $_2$ layer, deposited onto a glass plate and covered with ZnS cathodoluminophore, served as an anode. The cathode-anode distance was 50–300 μ m. Such devices allowed visualization of the emission current distribution over the cathode surface. In the second type of the test devices, the anode represented a molybdenum rod (1 mm in diameter) thanks to which one could avoid the unwanted deposition of ZnS, sputtered from the cathodoluminescent screen, onto the Ge-Si cathode. The cathode-anode gap in the devices of this geometry could be varied from 10 to 200 μ m.

The Ge-Si nanostructures, prepared in the MBE chamber, had been exposed to atmosphere before mounting them in the vacuum devices where field emission measurements were made. No additional sample cleaning procedure was applied prior to the measurements, which were carried out at residual gas pressures $\leq 10^{-8}$ Torr.

The photosensitivity of the field emission from the multilayer Ge-Si nanocluster structures was tested in the wavelength from 0.4 to 10 μ m. In the range from 0.4 to 2 μ m, we used a 10 W electric bulb with a system of filters, which illuminated the Ge-Si sample through the transparent cathodoluminescent screen. In the range of 2 to 10 μ m, the sample was

irradiated from 100 mW light emitting diodes placed immediately in the vacuum cell (with the Mo rod anode) close to the cathode-anode gap.

3. Results and Discussion

3.1 Field Emission

The AFM images of the Ge clusters of different sizes grown in our experiments are shown in Figs. 1 to 3. Note that the scales in the surface plane (xy -plane) and normal to the surface (z -axis) are strongly different. Thus, the Ge clusters are shaped as pyramids (“huts”) whose base sizes are much larger than their heights.

These cluster arrays were found to give a stable field electron emission. The high electric field near the cluster tip (~ 1 V/nm), necessary for the field emission to occur, can be achieved even at moderate voltages ($\sim 10^3$ V), because of the small curvature radius of the tip (~ 1 nm). The emission current-voltage (I - V) characteristics are also given in Figs. 1 to 3. It was noted that the I - V curves for some arrays contain reproducible current peaks whose number depends on the cluster height. For example, two current peaks are seen in the I - V curve for the clusters with the average height of ≈ 3 nm (Fig. 1) and four peaks for the clusters ≈ 5 nm high (Fig. 2). It is pertinent to mention that a careful analysis of the results of Tondare *et al.* [13] on field emission from self-assembled Ge nanoclusters on Si(111) also reveals some hints of the current peaks in

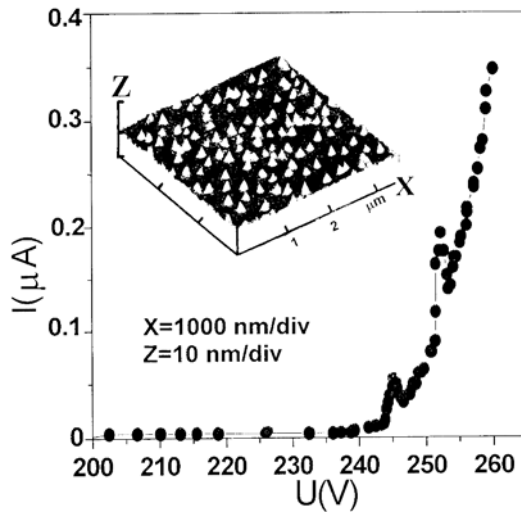


Figure 1. Typical AFM image and field emission current-voltage curve for Ge-Si nanostructure with Ge clusters whose average height is ≈ 3 nm.

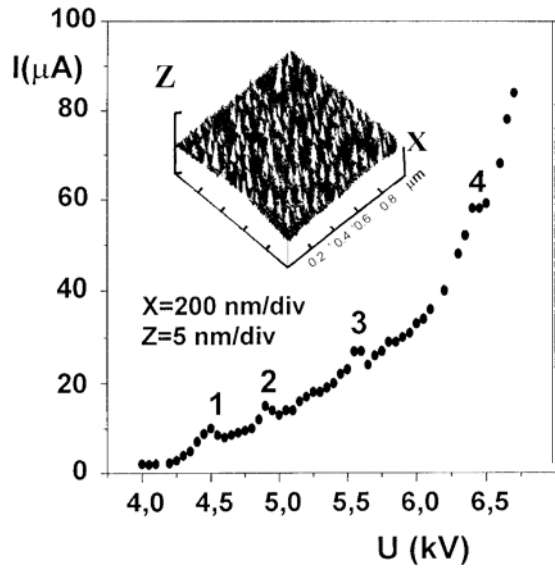


Figure 2. Typical AFM image and field emission current-voltage curve for Ge-Si nanostructure with Ge clusters whose average height is ≈ 5 nm.

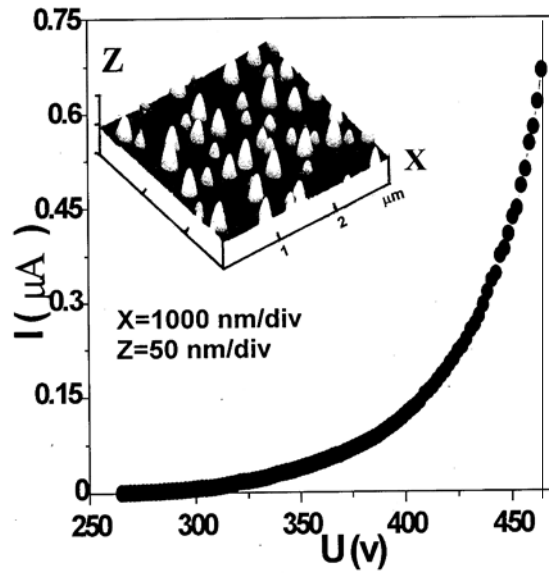


Figure 3. Typical AFM image and field emission current-voltage curve for Ge-Si nanostructure with Ge clusters whose average height is ≥ 10 nm.

their I - V curves. The curve for the clusters sized above 10 nm (Fig. 3) displays no distinct peaks at all. We speculate that this curve may actually contain a large number of overlapping small peaks, which are not discerned in the scatter of points. The peaks likewise disappear in the case when the height of the clusters increases owing to deposition of the sputtered phosphor material of the screen. On the other hand, the I - V curves for the smallest Ge clusters studied here (≈ 1 nm) also show no peaks.

All the I - V curves presented above, when plotted in Fowler–Nordheim (FN) coordinates $\log(I/V^2)$ versus $1/V$, appear strongly non-linear, and the current peaks seen in Figs. 1 and 2 are even more pronounced [14]. Generally, the non-linearity of FN plots is typical of semiconductor field emitters [7–12]. We suggest that the appearance of the current maxima in the field emission I - V curves may be related to the quantization of electron energy in the potential well corresponding to Ge cluster. The well is confined by the Ge-Si interface from one side and by the surface barrier (Ge electron affinity) from the other side, and is therefore highly asymmetric even in the absence of the electric field (Fig. 4). The use of the one-dimensional potential model is justified by the circumstance that, in the consideration of the process of field emission, we deal mainly the motion of electrons normal to the surface. The potential diagram is constructed under assumption that both Si and Ge, after their high-

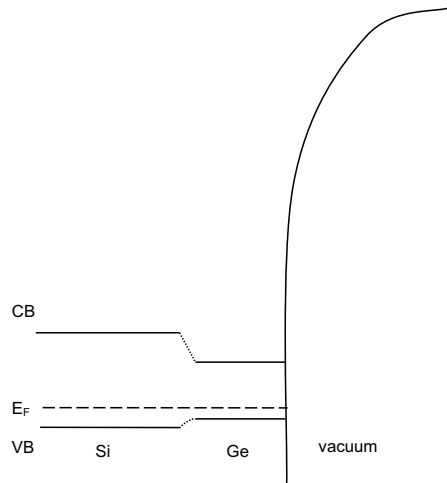


Figure 4. Schematic potential diagram of Ge layer on Si in absence of electric field. This diagram is drawn with assumptions that Si and Ge have p-conductivity and their bandgaps are 1.12 and 0.67 eV, respectively. the Ge work function is ≈ 4.8 eV. Possible surface states and band bending are not considered.

temperature treatment during the sample preparations, possessed a p -conductivity. The p -conductivity in this case is attributed to appearance of the defect centres containing vacancies [20]. Thus, the well depth can be approximately taken equal to the difference of the bandgaps of Si (1.12 eV) and Ge (0.67 eV).

In a rectangular and infinitely deep potential well, the electron energy levels lie at

$$E_n = \frac{n^2 h^2}{8m^* \ell^2}, \quad (1)$$

where n is the quantum number, h is Planck's constant, m^* the effective electron mass and ℓ the width of the potential well (see e.g. [21, 22]). As stated above, the depth of the well corresponding to Ge clusters in the absence of electric field is ≈ 0.45 eV (with respect to Si) and $m^* \approx 0.2 m_e$. The ground-state energy in such a well ≤ 10 nm wide is estimated at $\sim 10^2$ meV and, depending on the cluster size ℓ , a few energy levels can exist in it with energy gaps between them of the order of kT at $T = 300$ K. The criteria for the dimensions of quantum dots are discussed in detail in Ref. [23].

It is understood, however, that the presence of the high electric field near the surface changes radically the shape of the well due to band bending and formation of a semitransparent barrier on the vacuum side (Fig. 5). An inversion layer, with a specific quantization of the electron gas in the direction normal to the surface, arises at the surface (see, e.g. Refs. [10, 11, 24]). As the electric field is increased, the energy levels in the inversion layer change in a complicated way. This problem was considered in detail in connection with the investigation of field emission from p -type semiconductor tips [10]. However, to our knowledge, in none of the works that were carried out with the tips prepared of uniform semiconductor materials the maxima in the emission I - V curves were observed. The observation of such maxima in our experiments is likely to be connected with the distinction of the emitters studied here: they represented quantum dots grown on a foreign substrate. In the process of field emission, the dots are “transilluminated” by electrons supplied from the substrate volume.

The tunneling current density J for a certain energy E can be represented as the product

$$J(E) = S(E)T(E), \quad (2)$$

where $S(E)$ is the supply function and $T(E)$ the transmission function [25, 26]. The existence of the discrete energy levels in quantum dots should be most clearly reflected in $J(E)$, i.e. in the energy distribution of the emitted electrons [25, 27, 28]. However, if the peculiarities in $S(E)$

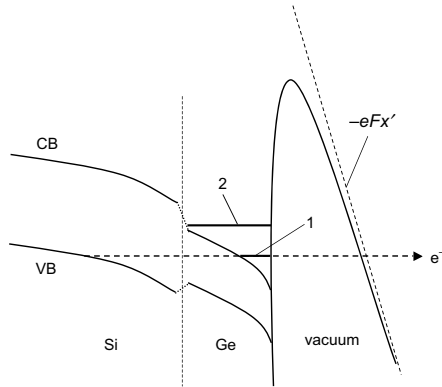


Figure 5. Possible potential diagram explaining field emission from Ge nanoclusters (quantum dots) on Si. Potential of electric field in vacuum is $-eFx'$. The hypothetical case is illustrated, where two energy levels (denoted 1 and 2) exist in the quantum dot and resonance tunneling occurs from level 1. VB denotes the valence band and CB the conduction band. Surface and interface states are not shown.

and $T(E)$ are strongly pronounced, they can be manifested in the I - V curves as well. We suppose that the current peaks in the I - V curves may be attributed to the resonance tunneling of electrons via energy levels existing in the Ge clusters. The electrons, originating from the valence band in the emitter volume, pass in their way to vacuum through the substrate/cluster interface, the cluster itself and the surface barrier. Generally, in each of these regions, the density of states distributions have some features relating to interface states, quantized levels in the cluster, and surface states. As the electric field is increased, the energy levels change their position in respect to each other (moreover, the field and confinement may alter the characteristics of the states in a fundamental way [26, 29]). Under favourable superposition, a maximum can appear in the supply function $S(E)$, giving a current peak in the I - V curve due to resonance tunneling. The problem is somewhat similar to the resonance tunneling through a combination of the potential barriers, where current maxima are observed [30].

The external field, in order to induce significant field emission, must be ~ 1 V/nm. Considering that the cluster size is ≤ 10 nm and the levels in the cluster well are separated by the gaps $\sim 10^2$ meV, the mean field inside the clusters that can be estimated from the above model of resonant tunneling should be $\sim 10^{-2}$ V/nm. As the size is increased, the number of the levels in the cluster should grow, and the gaps between them decrease. Our findings for the clusters of different size (Figs. 1 to 3) seem to agree with this prediction. We recall also that the I - V

curves for the smallest (~ 1 nm) Ge clusters, as well as for the large ones, showed no peaks, which is consistent with theoretically predicted absence of bound states in too narrow asymmetric potential wells [22].

One may ask why the clear-cut current peaks are recorded for ensembles of Ge clusters having some size spread. The most probable reason is the exponential dependence of the emission current on the field strength, due to which the group of the clusters with the smallest tip radius and the largest height gives the dominant contribution to the current. The I - V curves of field emission from Ge nanocluster multilayers have an intricate shape with a number of peaks, which are yet unexplained. Obviously, the potential diagram is highly complex for such structures.

It is important to note that rather good vacuum conditions ($P \leq 10^{-8}$ Torr) are necessary to obtain a stable field emission current from the Ge-Si nanocluster structures. By contrast, dielectric thin-film field emitters can generate a stable electron emission even at $P \sim 10^{-6}$ – 10^{-3} Torr [31, 32]. This is attributed to effective regime of negative electron affinity that can be attained in the latter case due to strong band bending, with the result that surface conditions are not decisive in the emission. Contrary to this, surface processes seem to be important in field emission from Ge nanoclusters on Si, as illustrated in Fig. 5.

3.2 Photofield Emission

Consider now the photofield emission from multilayer Ge-Si nanocluster structures shown schematically in Fig. 6. The field emission current, measured at room temperature from a 10-layer structure of Ge dots sized below 10 nm, increased several-fold under the irradiation by light with $\lambda = 2$ and $10 \mu\text{m}$ (Fig. 7). As mentioned in Sec. 1, interest in photofield

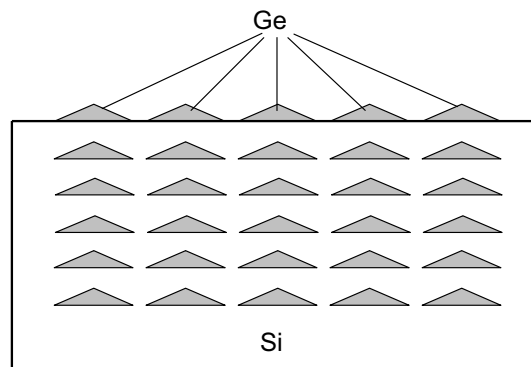


Figure 6. Schematic sketch of multilayer Ge nanocluster structure in Si.

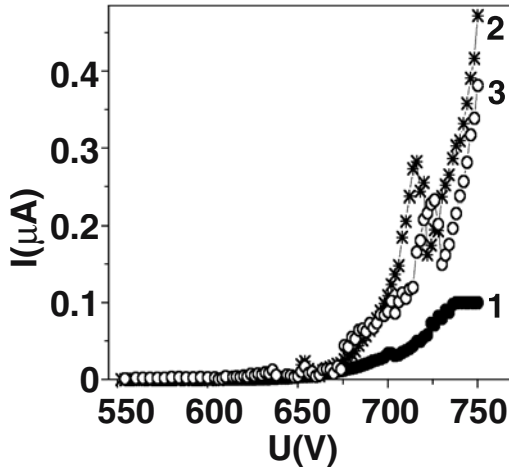


Figure 7. Effect of IR irradiation on field emission current-voltage characteristic for a multilayer Ge-Si nanocluster structure. (1) Dark current; (2) $\lambda \approx 2 \mu\text{m}$; (3) $\lambda \approx 10 \mu\text{m}$. $T = 300 \text{ K}$.

emission from semiconductors was stimulated by the intention to create efficient IR devices with high quantum yield [2, 4]. It was found [7, 9] that close to the intrinsic absorption edge ($\lambda \sim 1 \mu\text{m}$), the quantum yield from high-resistance Ge and Si samples amounts to ten per cent at liquid nitrogen cooling. However, at room temperature, the photosensitivity is practically absent for all the resistivity values studied in the interval $\rho \sim 10^{-1} - 10^5 \text{ Ohm-cm}$. The situation can be understood using a zero-current model of photo-field emission, illustrated in Fig. 8 (see, e.g. Ref. [10]). The shadowed region in this diagram corresponds to the region where the conductivity is close to the intrinsic one (depletion layer). Its extension depends on the doping degree. The quantum transitions induced by the light polarised in the detector surface plane are forbidden [2], and there exists only a low-efficient absorption due to indirect (phonon-assisted) transitions [33]. For Ge and Si, the absorption coefficient near the red edge, i.e. in the range interesting for IR detectors, is only about 10 cm^{-1} . It is therefore necessary to utilise high-resistance samples in which the extension of the depletion layer is 1 mm. The low photosensitivity of such emitters at 300 K is caused by the generation of a rather large number of electron-hole pairs at this temperature.

The above drawbacks of the traditional tip emitters can be obviated in the systems containing quantum dots (“artificial atoms”) as prepared in this work. In the clusters sized about 10 nm, the lateral electron con-

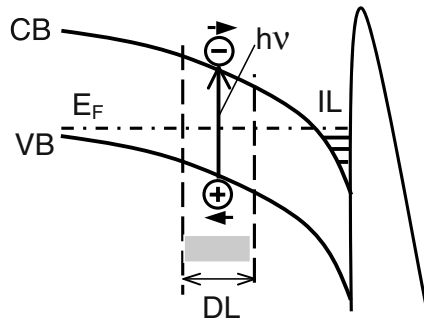


Figure 8. Potential diagram of band bending in process of field emission from uniform p-type semiconductors. E_F is Fermi level, VB valence band, CB conduction band, DL depletion layer, and IL inversion layer. Excitation of an electron-hole pair in the DL region by the quantum $h\nu$ is shown by the arrow.

finement gives rise to the discrete energy levels with gaps between them of the order of kT corresponding to 300 K. This lifts the restriction on optical transitions induced by normally incident light [2, 5, 33] and, as a result, can ensure the photosensitivity at 300 K which is 3 to 4 orders of magnitude higher than that of uniform tip samples. The spectral distribution of the photosensitivity can be controlled by variation of the cluster size. The strain, induced in Ge clusters due to the large lattice misfit (4.2%) between Ge and Si, can also significantly affect the cluster electronic structure. According to our data, which will be published elsewhere, the high strains can even, under some conditions, induce polymorphous transformations in Ge clusters in Si. Such an observation for Ge clusters embedded in SiO_2 was reported by A. Nylandsted Larsen in his talk at Nanomeeting-2003 in Minsk [34] (similar results were obtained for metal films [35]). As a consequence, the strain factor can also contribute to the broadening of the photosensitivity spectral range of field emission from the Ge-Si nanostructures.

The multilayer structures of Ge nanoclusters in Si prepared in this work showed considerable photosensitivity in the spectral range of 0.4 to 10 μm at room temperature. It was also found that liquid nitrogen cooling of the emitter allows the dark current to be decreased by about an order of magnitude [18].

4. Conclusions

The results of this work show that the size and the surface density of Ge nanoclusters on Si(100) can be effectively controlled by deposition of

a series of the intermediate Ge-Si layers. These layers allow a gradual relaxation of the elastic strains at the surface that serves as the substrate for the growth of Ge quantum dots. In the presence of high electric field, the dots can generate a stable field emission, which seems to exhibit some effects of energy quantization in the dots.

Electron emission from Ge clusters on and in Si is an interesting, but intricate process sensitive to cluster size, electric field, electromagnetic irradiation, temperature and, probably, to misfit interface strains. Further relevant factors can be the Fermi level inclination due to intense current take-off, the nonequilibrium heating of electrons by the internal field, and surface conditions, in particular the characteristics of surface and interface states [26, 29, 36]. It is pertinent to note that manifestations of surface states and resonances were revealed in field and photofield emission from metals [37–39].

The properties of the nanostructures reported here may find some application in novel optoelectronic devices operating at room temperatures.

Acknowledgements

We acknowledge the support of this investigation in the framework of the Program of Basic Research of National Academy of Sciences of Ukraine. We thank P. M. Litvin for obtaining the AFM images of Ge-Si nanostructures.

References

- [1] D. Gruetzmacher. Si/SiGe nanostructures: challenges and future perspectives. in V.E. Borisenko, S.V. Gaponenko and V.S. Gurin eds., *Physics, Chemistry and Application of Nanostructures* (World Scientific, Singapore, 2003), pp. 3–10.
- [2] G. Masini, L. Colace and G. Assanto. (2003) Germanium-on-silicon infrared detectors, in H.S. Nalwa ed., *Encyclopedia of Nanoscience and Nanotechnology* Vol. X (American Scientific Publishers, 2003), pp. 1–14.
- [3] V. Le Thanh. Mechanisms of island vertical alignment in Ge/Si(001) quantum dot multilayers. in V. E. Borisenko, S. V. Gaponenko and V. S. Gurin (eds.), *Physics, Chemistry and Application of Nanostructures* (World Scientific, Singapore, 2003), pp. 447459.
- [4] P. Boucaud, V. Le Thanh, S. Sauvage, Débarre and D. Bouchier. (1999) Intra-band absorption in Ge/Si self-assembled quantum dots. *Appl. Phys. Lett.* **74**: 401–404, 1999.
- [5] A. I. Yakimov, A. V. Dvurechenskii, A. I. Nikiforov, O. P. Pchelyakov and A. A. Nenashev. Evidence for a negative interband photoconductivity in arrays of Ge/Si type-II quantum dots. *Phys. Rev. B* **62**: R16283–R16286, 2000.
- [6] C. Teichert. Self-organization of nanostructures in semiconductor heteroepitaxy. *Phys. Rep.* **365**: 335–432, 2002.

- [7] J. R. Arthur. Photosensitive field emission from p-type germanium. *J. Appl. Phys.* **36**: 3221–3227, 1965.
- [8] R. Fischer and H. Neumann. Feld Emission aus Halbleitern, *Fortschritte d. Phys.* **14**: SS. 603–692, 1966.
- [9] P. G. Borzyak, A. F. Jatsenko and L. S. Miroshnichenko. Photo-field-emission from high-resistance silicon and germanium. *Phys. Stat. Sol.* **14**: 403–411, 1966.
- [10] A. F. Yatsenko. On a model of photo-field-emission from p-type semiconductors. *Phys. Stat. Sol. (a)* **1**: 333–348, 1970.
- [11] M. H. Herman and T. T. Tsong. Photon excited field emission from a semiconductor surface. *Phys. Lett.* **71A**: 461–463, 1979.
- [12] G. N. Fursei. Early field emission studies of semiconductors. *Appl. Surf. Sci.* **94/95**: 44–59, 1996.
- [13] V. N. Tondare, B. I. Birajdar, N. Pradeep, D. S. Joag, A. Lobo and S. K. Kulkarni. Self-assembled Ge nanostructures as field emitters. *Appl. Phys. Lett.* **77**: 2394–2396, 2000.
- [14] A. A. Dadykin, Yu. N. Kozyrev and A. G. Naumovets. Field electron emission from Ge-Si nanostructures with quantum dots. *JETP Lett.* **76**: 472–474, 2002.
- [15] V. A. Shchukin and D. Bimberg. Spontaneous ordering of nanostructures on crystal surfaces. *Rev. Mod. Phys.* **71**: 1125–1171, 1999.
- [16] E. Korutcheva, A. M. Turiel and I. Markov. Coherent Stranski-Krastanov growth in 1 + 1 dimensions with anharmonic interactions: An equilibrium study, *Phys. Rev. B* **61**: 16890–16901, 2000.
- [17] Yu. Kozyrev, M. Yu. Rubezhanska, O. O. Chuiko and V. M. Ogenko. Epitaxial formation of Ge dot system in Si(100) substrate. *Dopovidi NAN Ukrainy (Reports of Natl. Acad. Sci. of Ukraine)*, No. 1, pp. 76–78, 2002 (in Ukrainian).
- [18] A. A. Dadykin, A. G. Naumovets, Yu. N. Kozyrev, M. Yu. Rubezhanska, P. M. Lytvyn and Yu. M. Litvin. Field and photofield electron emission from self-assembled Ge-Si nanostructures with quantum dots. *Prog. Surf. Sci.* **74**: 305–318, 2003.
- [19] F. F. Sizov, V. P. Kladko, S. V. Plyatsko, S. A. Shevlyakov, Yu. N. Kozyrev and V. M. Ogenko. Si/Si_{1-x}Ge_x epitaxial layers and superlattices. Growth and structural characteristics. *Semiconductors* **31**: 786–788, 1997.
- [20] V. F. Bibik, A. A. Dadykin and V. A. Titov. Influence of high-temperature heating in vacuum on the near-surface layer of silicon crystals. *Ukrainskii Fiz. Zhurnal (Ukrainian Phys. Journ.)* **20**: 1684–1688 (in Russian), 1975.
- [21] L. I. Schiff. *Quantum Mechanics* (McGraw-Hill, New York, 1955), § 9.
- [22] L. D. Landau and E. M. Lifshits. (1981) *Quantum Mechanics* (Pergamon, Oxford, 1981), § 22.
- [23] N. N. Ledentsov, V. M. Ustinov, V. A. Shchukin, P. S. Kopev, Zh. I. Alferov and D. Bimberg. Quantum dot heterostructures: fabrication, properties, lasers. *Semiconductors* **32**: 343–365, 1998.
- [24] T. Ando, A. B. Fowler and F. Stern. Electronic properties of two-dimensional systems. *Rev. Mod. Phys.* **54**: 437–672, 1982.
- [25] A. Modinos. *Field, Thermionic, and Secondary Electron Emission Spectroscopy* (Plenum Press, New York, 1984), Ch. 7.

- [26] M. Kolar and S. G. Davison. Field emission from band gap states. *J. Vac. Sci. Technol. A* **3**: 392–397, 1985.
- [27] C. B. Duke and M. E. Alferieff. Field emission through atoms adsorbed on a metal surface. *J. Chem. Phys.* **46**: 923–937, 1967.
- [28] J. W. Gadzuk and E. W. Plummer. Field emission energy distribution. *Rev. Mod. Phys.* **45**: 487–548, 1973.
- [29] R. J. Jerrard, H. Ueba and S. G. Davison. Virtual surface states of thin films. *Phys. Stat. Sol. (b)* **103**: 353–359, 1981.
- [30] B. Ricco and M. Ya. Azbel. Physics of resonance tunneling. The one-dimensional double-barrier case. *Phys. Rev. B* **29**: 1970–1981, 1984.
- [31] A. A. Dadykin, A. G. Naumovets, V. D. Andreev, T. A. Nachalnaya and V. A. Semenovich. A study of stable low-field electron emission from diamond-like films. *Diamond and Related Materials* **5**: 771–774, 1996.
- [32] A. A. Dadykin and A. G. Naumovets. Low-macroscopic-field electron emission from piezoelectric thin films and crystals. *Mater. Sci. Eng. A* **353**: 12–21, 2003.
- [33] K. Seeger. *Semiconductor Physics* (Springer, Wien, 1973), Ch. 11.
- [34] A. Nylandsted Larsen, A. Kanjilal, J. Lundsgaard Hansen, P. Gaiduk, N. Cherkashin, A. Claverie, P. Normand, E. Kapelanakakis, D. Tsoukalas and K.-H. Heinig. Germanium quantum dots in SiO₂: Fabrication and characterization. in V. E. Borisenko, S. V. Gaponenko and V. S. Gurin (eds.), *Physics, Chemistry and Application of Nanostructures* (World Scientific, Singapore, 2003), pp. 439–446.
- [35] P. Turban, L. Hennet and S. Andrieu. In-plane lattice spacing oscillatory behaviour during the two-dimensional hetero- and homoepitaxy of metals. *Surf. Sci.* **446**: 241–253, 2000.
- [36] S. G. Davison and M. Steslicka. *Basic Theory of Surface States* (Clarendon Press, Oxford, 1996).
- [37] Z. A. Ibrahim and M. J. G. Lee. Surface states and resonances in field emission from low-index facets of clean tungsten. *Prog. Surf. Sci.* **67**: 309–322, 2001.
- [38] T. Radon. Photofield emission spectroscopy and surface density of states. *Prog. Surf. Sci.* **67**: 339–346, 2001.
- [39] T. Radon, L. Jureczyszyn, P. Hadzel. Size effect in photofield emission from (001) tungsten. *Surf. Sci.* **513**: 549–554, 2002.

OPTICAL EMISSION BEHAVIOR OF Si QUANTUM DOTS

X. Zianni and A. G. Nassiopoulou

IMEL/NCSR 'Demokritos', 153 10 Aghia Paraskevi, Athens, Greece

Abstract Silicon plays a dominant role in microelectronics. Si nanostructures are intensively investigated because of their compatibility with Si technology. In the present work, we report on the optical emission properties of Si quantum dots. The energy states are calculated within the effective mass approximation. Both first and second order transitions are included in the calculation. The sizes of the quantum dots that we have considered are in the range 2–5 nm. The magnitude of the lifetime is found to be very sensitive to the structural parameters of the dots such as the shape, the cross section and the crystallographic direction. Lifetimes varying from the order of μsec to the order of msec have been obtained. The results of our calculations provide further insight in the photoluminescence behavior of Si nanostructures and porous Si.

1. Introduction

The recent progress in the fabrication of low-dimensional semiconductors in the nanometric scale for applications in transistor and memory devices has boosted the research in the properties of Si nanostructures. The electrical properties of Si nanowires, nanopillars and arrays of dots is the subject of current research. Their optical properties are also investigated for applications in the field of photonics and sensing including important perspectives in biosensing.

Different systems containing silicon quantum dots or nanowires have been investigated so as the intrinsic properties due to dots or wires have been demonstrated [1–6]. They include both porous silicon, which consists of a network of crystalline silicon wires and/or dots in the nanometer range with dispersion of sizes and shapes, and other more ordered systems, as for example two-dimensional arrays of silicon quantum dots in SiO_2 , fabricated by ion implantation of silicon into the SiO_2 [7–13]. Empirical, semi-empirical and first principles techniques have been used to study theoretically the effects of the reduction in dimensions, of the

geometry and of the structure on the optical properties of silicon nanostructures [14–29]. All the calculations found a widening of the band gap from the near-infrared wavelength region to and beyond the visible range. It has been shown that the confinement results to an enhancement of the dipole matrix element that is responsible for the radiative transitions. Both direct and indirect band gap have been reported.

Most calculations are restricted to particular cases of geometries and sizes of the dots and of the wires. Continuous values can be used for the structural parameters within the effective mass approximation (EMA) and so further evidence of new physical mechanisms that can be directly connected with the experiment can be obtained. The predictions of EMA for the energy band structure of electrons in Si nanowires, have been discussed in detail in [30, 31] where the findings of EMA are also compared with results of more sophisticated calculations. The implications of the features of the electron band structure on the spontaneous emission have been discussed in [32] and the behavior of the PL lifetime in Si quantum wires with dimensions in the range of a few nanometers has been discussed in [33]. In this paper, we present our results on Si quantum dots. The PL lifetime is calculated as a function of the crystallographic direction and the size of the dots. We have considered sizes in the range 1–4 nm and crystallographic directions varying in the (100) plane from the [100] to the [110] direction. The theoretical model is presented in short in Sec. 2 together with the results on the energy levels and the lifetimes. Conclusions are drawn in Sec. 3.

2. Model and Results

We consider free standing dots of rectangular cross section. Electron and hole eigenstates are obtained by solving Schrödingers equation for each ellipsoid valley of bulk Si conduction band minimum. An infinitely deep confining potential is assumed. We define a system of coordinates (x, y, z) as follows: the z -axis is along the [001] direction and the x - and y -axes are rotated anticlockwise by an angle ϕ relative to the [100] and [010] directions respectively.

In Si quantum dots the anisotropy of bulk Si valleys causes a deviation of the electron states from the simple sinusoidal functions in the case of cubic dots. Schrödingers equation is solved numerically [34]. The six valleys of bulk Si give distinct levels. The order and the separation of the energy levels depend on the size and on the crystallographic direction of the dot. Representative results are plotted in Fig. 1. In Fig. 1(a) are shown the first three calculated energy levels for the three sets of valleys versus the crystallographic direction for a cubic dot of width $L = 2$ nm.

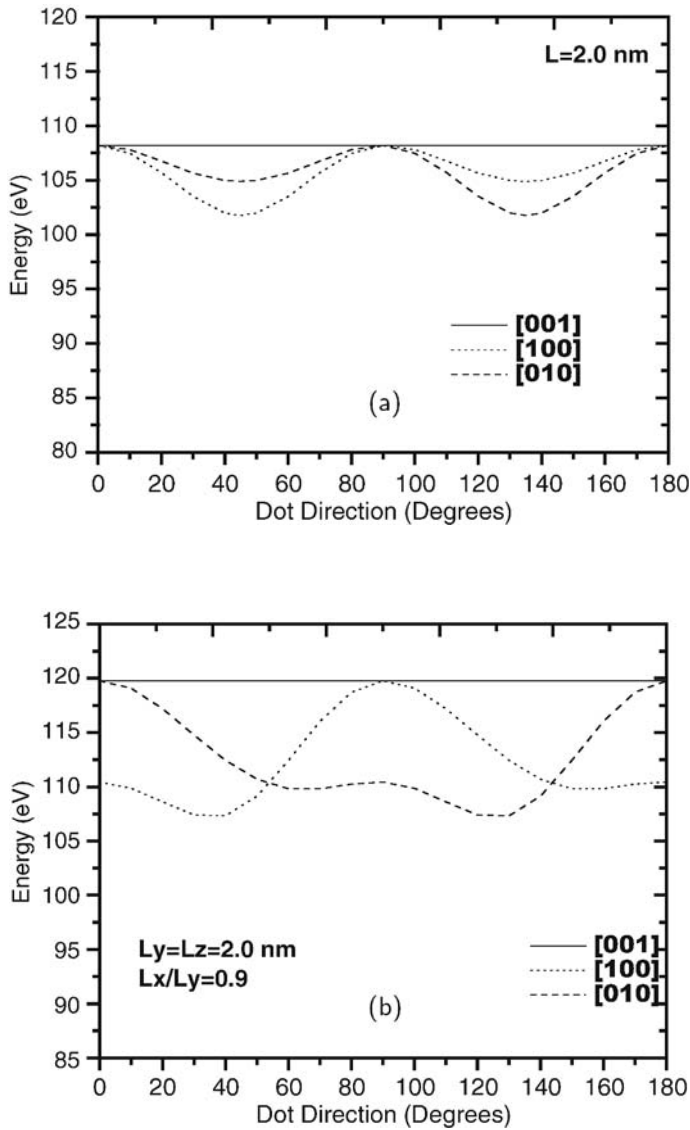


Figure 1. The electron energy levels versus the dot direction from the [001] valleys (solid line), the [100] valleys (dotted lines) and the [010] valleys (dashed line) for (a) a cubic dot of width 2 nm, and (b) for an orthogonal dot with $L_y = L_z = 2$ nm and $L_x = 0.9L_y$.

It can be seen that the energy levels do not depend on the direction of the dot for the energy levels from the valleys [001] (the [001]-levels) whereas they exhibit dependence for the other two sets of valleys. The exact dependence and the order of the levels depends on the shape of the dots, as it can be seen in Fig. 1(b), where are shown the results for a dot of orthogonal shape with $L_y = L_z$ and $L_x = 0.9 L_y$.

In Si quantum dots of dimensions of a few nanometers due to considerable confinement, the number of excited electron-hole ($e-h$) pair is small and our system consists of independent $e-h$ pairs. Therefore, our discussion refers to a single $e-h$ pair recombination. The lifetime τ is defined as the inverse of the recombination rate between a hole in the highest occupied molecular orbital (HOMO) and an electron in the lowest unoccupied molecular orbital (LUMO) or in an excited state near LUMO as described in the following discussion. Both first and second order (phonon-assisted) transitions are included in the calculations. The calculated lifetimes for a cubic dot with size 2 nm are shown in Fig. 2. The lifetime for the [001]-levels do not depend on the direction and it is of the order of μsec . Its magnitude depends on the overlap integral. The overlap integral is determined by the size of the dot and by the phase of the wave function. The phase of the electron wave function is explicitly determined by the conduction band minimum of bulk Si and results to an overlap integral smaller than unity. The lifetime for the

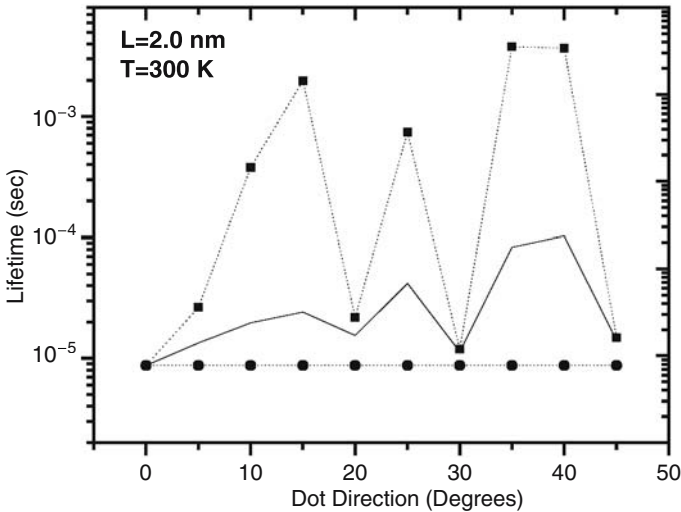


Figure 2. The lifetime versus the dot direction for a cubic dot of width 2 nm for the [001]-level (dots), the [100]-level (squares) and the total lifetime (solid line). The dashed curves are guide for the eye.

[100]-levels shows a strong directional dependence and as it can be seen in Fig. 2, it varies in the range μsec – msec . Now, the phase of the wave function additionally depends on the crystallographic direction and this dependence explains the oscillations seen in Fig. 2.

In the previous paragraph, it has been shown that the character of the electronic levels strongly affects the PL lifetime. Thermal activation from one level to the other is also important. At low temperatures, only the ground level is occupied and it determines the PL lifetime. At room temperature, the electron can be excited to an excited level, due to its small separation from the ground level as it can be seen in Fig. 1, and recombine from there. Then, the PL lifetime is determined by both levels as is expected by Boltzmann statistics:

$$\tau^{-1} = \frac{\tau_F^{-1} e^{-\Delta E/k_B T} + \tau_S^{-1}}{1 + e^{-\Delta E/k_B T}}, \quad (1)$$

where τ stands for the total PL lifetime and $\Delta E = E_F - E_S$ for the energy separation of the two levels. The subscripts F and S are for “fast” and “slow”, respectively, and refer to the ground and the excited level depending on which is faster or slower than the other. The total lifetime calculated at room temperature is shown in Fig. 2 by a solid line. It can be seen that the lifetimes features of the [100]-level are smoother in the total lifetime due to the effect of the [001]-level that has a lifetime independent of the crystallographic direction. However, in so small dots a considerable dispersion in the magnitudes of the lifetime is found as a function of the crystallographic direction. This effect is less remarkable in bigger dots, as can be seen in Fig. 3, where are presented the calculated lifetimes for a dot with size 4 nm. In this case, fewer oscillations are found in the [100]-level lifetimes. Moreover, the separation between the ground level and the excited level is smaller and the effect of the [001]-level in the total lifetime is more important as it is expected by Eq. (1).

3. Conclusion

In the above discussion, it has been shown that in Si quantum dots the separation between the energy levels is not determined solely by the confinement dimension but also by the crystallographic direction due to the character of the anisotropic valleys of bulk Si conduction band minimum. Recombination from electron states originating from different valleys can have quite different lifetimes. The energy separation between states originating from different valleys differs depending on the size and the crystallographic direction of the dots. Electrons can be thermally excited from the ground to an excited level and recombine from there with a smaller PL lifetime. It can therefore be concluded

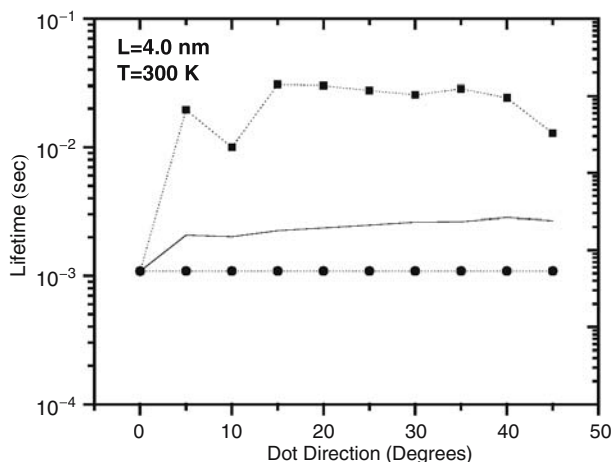


Figure 3. The lifetime versus the dot direction for a cubic dot of width 4 nm for the [001]-level (dots), the [100]-level (squares) and the total lifetime (solid line). The dashed curves are guide for the eye.

that non-uniformity in the size and/or in the crystallographic direction of a network of quantum dots cause dispersion in the magnitude of the PL lifetime.

References

- [1] A. G. Nassiopoulou, S. Grigoropoulou, E. Gogolides and P. Papadimitriou. Visible luminescence from one- and two-dimensional silicon structures produced by conventional lithographic and reactive ion etching techniques. *Appl. Phys. Lett.* **66**: 1114–1116, 1995.
- [2] A. G. Nassiopoulou, S. Grigoropoulos and P. Papadimitriou. Electroluminescent device based on silicon nanopillars. *Appl. Phys. Lett.* **69**: 2267–2269, 1996.
- [3] D. J. Lockwood, Z. H. Lu and J. M. Baribeau. Quantum confined luminescence in Si/SiO₂ superlattices *Phys. Rev. Lett.* **76**: 539–541, 1996).
- [4] V. Ioannou-Sougleridis, A. G. Nassiopoulou, T. Ouisse, F. Bassani and F. Arnaud d'Avitaya. Electroluminescence from silicon nanocrystals in Si/CaF₂ superlattices. *Appl. Phys. Lett.* **79**: 2076–2078, 2001.
- [5] B. V. Kamenev and A. G. Nassiopoulou. Self-trapped excitons in silicon nanocrystals with sizes below 1.5 nm in Si/SiO₂ multilayers. *J. Appl. Phys.* **90**: 5735–5740, 2001.
- [6] T. Ouisse and A. G. Nassiopoulou. Dependence of the radiative recombination lifetime upon electric field in silicon quantum dots embedded into SiO₂ *Europhys. Lett.* **51**: 168–173, 2000.
- [7] L. T. Canham (ed.). *Properties of Porous Silicon*, Vol. 18 (INSPEC, London, 1997).

- [8] L. T. Canham. Silicon quantum wire array fabrication by electrochemical and chemical dissolution of wafers. *Appl. Phys. Lett.* **57**: 1046–1048, 1990.
- [9] C.-Y. Yeh, S. B. Zhang and A. Zunger. Identity of the light-emitting states in porous silicon wires. *Appl. Phys. Lett.* **63**: 3455–3457, 1993.
- [10] S. Ossicini. Porous silicon modeled as idealized quantum wires. in *Properties of Porous Silicon*, Vol. 18, edited by L. T. Canham (INSPEC, London, 1997).
- [11] P. Photopoulos, A. G. Nassiopoulou, D. N. Kouvatsos and A. Travlos. Photoluminescence from nanocrystalline silicon in Si/SiO₂ superlattices. *Appl. Phys. Lett.* **76**: 3588–3590, 2000.
- [12] K. J. Nash. Porous silicon modeled as undulating quantum wires. in *Properties of Porous Silicon*, Vol. 18, edited by L. T. Canham (INSPEC, London, 1997).
- [13] Z. Wu, T. Nakayama, S. Qiao and M. Aono. Strong linear polarization in scanning tunneling microscopy-induced luminescence from porous silicon. *Appl. Phys. Lett.* **74**: 3842–3844, 1999).
- [14] A. J. Read, R. J. Needs, K. J. Nash, L. T. Canham, P. D. J. Calcott and A. Oteish. First-principles calculations of the electronic properties of silicon quantum wires. *Phys. Rev. Lett.* **69**: 1232–1235, 1992).
- [15] T. Ohno, K. Shiraishi and T. Ogawa. Intrinsic origin of visible light emission from silicon quantum wires: Electronic structure and geometrically restricted exciton. *Phys. Rev. Lett.* **69**: 2400–2403, 1992.
- [16] F. Buda, J. Kohanoff and M. Parinello. Optical properties of porous silicon: A first-principles study. *Phys. Rev. Lett.* **69**: 1272–1275, 1992.
- [17] M. S. Hybertsen and M. Needels. First-principles analysis of electronic states in silicon nanoscale quantum wires. *Phys. Rev. B* **48**: 4608–4611, 1993.
- [18] C. Delerue, G. Allan and M. Lannoo. Theoretical aspects of the luminescence of porous silicon. *Phys. Rev. B* **48**: 11024–11036, 1993.
- [19] L. W. Wang and A. Zunger. Electronic structure pseudopotential calculations of large (approx. 1000 atoms) Si quantum dots. *J. Phys. Chem.* **98**: 2158–2165, 1993.
- [20] B. Delley and E. F. Steigmeier. Size dependence of band gaps in silicon nanostructures. *Appl. Phys. Lett.* **67**: 2370–2372, 1995.
- [21] A. M. Saitta, F. Buda, G. Fiumara and P. V. Giaquinta. *Ab initio* molecular-dynamics study of electronic and optical properties of silicon quantum wires: Orientational effects. *Phys. Rev. B* **53**: 1446–1451, 1996.
- [22] L. Dorigoni, O. Bisi, F. Bernardini and S. Ossicini. Electron states and luminescence transition in porous silicon. *Phys. Rev. B* **53**: 4557–4564, 1996.
- [23] G. Allan, C. Delerue and M. Lannoo. Electronic structure of amorphous silicon nanoclusters. *Phys. Rev. Lett.* **78**: 3161–3164, 1997.
- [24] J.-B. Xia and K. W. Cheah. Quantum confinement effect in thin quantum wires. *Phys. Rev. B* **55**: 15688–15693, 1997.
- [25] S. Ossicini. Optical properties of confined Si structures. *Phys. Stat. Sol. (a)* **170**: 377–390, 1998.
- [26] J. Taguena-Martinez, Y. G. Rubo, M. Cruz M, M. R. Beltran and C. Wang. Tight-binding description of disordered nanostructures: An application to porous silicon. *Appl. Surf. Sci.* **142**: 564–568, 1999.

- [27] Y. M. Niquet, C. Delerue, G. Allan and M. Lannoo. Method for tight-binding parametrization: Application to silicon nanostructures. *Phys. Rev. B* **62**: 5109–5116, 2000.
- [28] E. Degoli and S. Ossicini. The electronic and optical properties of Si/SiO₂ superlattices: role of confined and defect states. *Surf. Sci.* **470**: 32–42, 2000.
- [29] E. Degoli, M. Luppi and S. Ossicini. From undulating Si quantum wires to Si quantum dots: A model for porous silicon. *Phys. Stat. Sol. (a)* **182**: 301–306, 2000).
- [30] S. Horiguchi, Y. Nakajima, Y. Takahashi and M. Tabe. Energy eigenvalues and quantized conductance values of electrons in Si quantum wires on (100) plane. *Jpn. J. Appl. Phys.* **34**: 5489–5498, 1995.
- [31] S. Horiguchi S. Conditions for a direct band gap in Si quantum wires. *Superlattices Microst.* **23**: 355–364, 1998.
- [32] X. Zianni and A. G. Nassiopoulou. Directional dependence of the spontaneous emission of Si quantum wires. *Phys. Rev. B* **65**: art. no. 035236, 2002.
- [33] X. Zianni and A. G. Nassiopoulou. Photoluminescence lifetimes of Si quantum wires. *Phys. Rev. B* **66**: art. no. 205323, 2002.
- [34] X. Zianni and A. G. Nassiopoulou. unpublished.

STRAIN-DRIVEN PHENOMENA UPON OVERGROWTH OF QUANTUM DOTS: ACTIVATED SPINODAL DECOMPOSITION AND DEFECT REDUCTION

M. V. Maximov

Ioffe Physical-Technical Institute, Politekhnikeskaya 26, 194021, St. Petersburg, Russia

N. N. Ledentsov

Ioffe Physical-Technical Institute, Politekhnikeskaya 26, 194021, St. Petersburg, Russia, and Institut für Festkörperphysik, Technische Universität Berlin, Hardenbergstr. 36, D-10623 Berlin, Germany

Abstract Strain-driven decomposition of an alloy layer is investigated as a means to control the structural and electronic properties of self-organized quantum dots. Coherent InAs/GaAs islands overgrown with an InGaAs alloy layer serve as a model system. Cross-sectional and plan-view transmission electron microscopy, as well as photoluminescence studies, consistently indicate an increase in height and width of the quantum dots with increasing indium content and/or thickness of the alloy layer. The increase in dot size is attributed to the phase separation of the alloy layer driven by the surface strain introduced by the initial InAs islands. The density of large dislocated clusters in the quantum dot samples can be dramatically reduced by special *in situ* annealing (defect reduction) techniques. A laser based on multiple (up to 10) layers of InAs/GaAs quantum dots grown under optimized conditions with the use of defect reduction techniques show considerably enhanced optical gain and improved performance. Differential efficiency as high as 88% is achieved in the lasers. An emission wavelength of 1280 nm, a threshold current density of 147 A/cm², a differential efficiency of 80%, and a characteristic temperature of 150 K are realized simultaneously in one device. Wavelength extension up to 1500 nm for InAs quantum dot lasers on GaAs substrates is possible by using metamorphic (plastically relaxed) buffer InGaAs layers. In cases when the threading dislocations are avoided in the active region, high-performance operation with quantum efficiency exceeding 60–70% and pulsed output powers up to 7 W are realized.

1. Introduction

The formation of InAs/GaAs quantum dots (QDs) in highly lattice-mismatched epitaxy in the Stranski–Krastanow growth mode has attracted increasing interest in recent years. This growth mode provides an efficient means to fabricate dense arrays of virtually defect-free uniform QDs. Such self-organized QDs have enabled, for the first time, full three-dimensional confinement to be used in injection lasers [1]. Self-organized growth provides, however, only limited control of the inter-related size and density of the islands, making it difficult to access the technologically important $1.3\ \mu\text{m}$ region (for a review, see, e.g. Ref. [2] and references therein). Extensive investigations have been conducted to overcome these limitations and to optimize the properties of QD arrays. Several approaches have been exploited to generate high area densities of QDs, providing sufficient localization for electrons and holes. In molecular-beam epitaxy (MBE) the alternate deposition of group-III atoms and group-V molecules, e.g. by alternating deposition beams of In, As₂, Ga, and As₂ [3,4], as well as the use of a seed layer [5], has been successful.

The growth of large QDs is mostly based on exploiting the strain fields generated by nanoislands, either free-standing or covered with a thin cap layer. For typical growth conditions, the bulk diffusion coefficients of atoms are several orders of magnitude smaller than surface ones and, thus, bulk diffusion plays only a limited role during growth. Structural inhomogeneities, e.g. inclusions of a foreign material in a host matrix or composition modulations persist after overgrowth. Such inhomogeneities create long-range strain fields, which direct the surface migration of adatoms. These strain fields are known to enhance phase separation in growing alloys, both lattice-matched [6] and lattice-mismatched [7], to the substrate. For example, the growth of multi-sheet arrays of QDs proceeds in such a way that formation of new islands is driven by the strain fields of the buried islands, resulting either in vertical correlation [8,9] or vertical anticorrelation [10] between the QDs of successive sheets depending on the spacer thickness [11].

Here we report on the surface strain driven phase separation of an InGa(Al)As alloy cap layer activated by predeposited InAs islands formed by Stranski–Krastanow growth. We describe technological approaches to reduce the density of dislocations and large dislocated clusters in QD samples (defect reduction techniques) and show that defect reduction results in a dramatic improvement of QD lasers performance. Finally we describe first results on $1.5\ \mu\text{m}$ InAs QD lasers on GaAs substrates.

2. Sample Growth

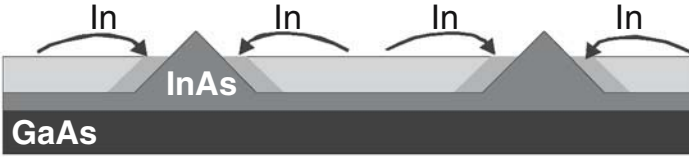
The investigated samples were grown by solid-source molecular-beam epitaxy (MBE) on GaAs (001) substrates using a Riber-32 MBE machine. After oxide desorption, an $0.5\ \mu\text{m}$ thick GaAs buffer was grown at 600°C followed by a $2\ \text{nm}/2\ \text{nm}$ GaAs/AlAs short-period superlattice with six periods and a $100\ \text{nm}$ GaAs spacer. Subsequently, the substrate temperature was lowered to 480°C for the deposition of an InAs QDs. The first $100\ \text{nm}$ of the subsequent GaAs cap layer were grown at the same temperature before the substrate temperature was increased back to 600°C for the growth of $20\ \text{nm}$ of GaAs, a $2\ \text{nm}/2\ \text{nm}$ GaAs/AlAs short-period superlattice with six periods, and a final $5\ \text{nm}$ thick GaAs layer for surface protection. The investigated laser structures were grown on Si-doped GaAs(001) substrate as described in detail in Ref. [12]. $\text{Al}_{0.13}\text{Ga}_{0.87}\text{As}$ waveguide of $0.5\ \mu\text{m}$ thickness and $1.5\ \mu\text{m}$ -thick Si- and Be-doped $\text{Al}_{0.7}\text{Ga}_{0.3}\text{As}$ cladding layers were used. Transmission electron microscopy (TEM) measurements were performed in plan-view and cross-sectional geometry with a Philips EM 420 microscope at an acceleration voltage of $100\ \text{kV}$. For each sample the average QD width was determined from the histogram of size distribution plotted using TEM images with $100\text{--}200$ dots.

The QDs were formed by the method that we refer to as activated alloy phase separation (AAPS) [13]. A sheet of strained InAs islands is formed on the GaAs surface [Fig. 1(a)] by the deposition of $D_{\text{IS}} = 1.8\text{--}4\ \text{ML}$ InAs in the conventional Stranski–Krastanow mode, and subsequently overgrown with an $\text{In}_x\text{Ga}_{1-x}\text{As}$ alloy layer [Fig. 1(b,c)]. The laterally varying surface strain caused by the InAs islands affects the overgrowth with an InGaAs alloy layer by strain-driven surface migration. One may expect the following qualitative picture of the alloy phase separation activated by stressors. For the conventional overgrowth of InAs QDs by pure GaAs, it has been found [14] that Ga atoms prefer to migrate away from the QDs towards pseudomorphically strained regions having an in-plane local lattice constant equal to that of unstrained GaAs. A similar effect should occur during the overgrowth of InAs QDs by an InGaAs alloy: Indium atoms will accumulate at the InAs QDs, increasing their lateral size. When the QDs are completely covered, the tensile strain on top of the QDs will favor In accumulation from the growing alloy, probably increasing the effective height of the islands. Thus, AAPS increases the In concentration in the vicinity of the islands at the expense of the In concentration in the InGaAs alloy. In other words, AAPS increases the width and height of the InAs islands, providing for enhanced localization of electrons and holes.

(a) $D_{IS}=1.8 - 4$ ML InAs



(b) $\text{In}_x\text{Ga}_{1-x}\text{As}$



(c)



Figure 1. Schematic diagram illustrating activated alloy phase separation. (a) Initial InAs stressors, (b) partial decomposition of the alloy layer due to strain-driven surface migration of Indium, and (c) the final structure after GaAs capping.

Figure 2 supports the expected trend, showing plan-view TEM pictures of the original InAs islands covered with GaAs [panel (a)], as well as of such islands covered with a 5 nm $\text{In}_{0.15}\text{Ga}_{0.75}\text{As}$ layer [panels (b,c)]. Images (a) and (b), taken at $\mathbf{g} = 220$, show a strong contribution of the local strain, image (c), taken under diffraction conditions far away from exact Bragg orientation, allows for a more precise determination of the island size and shape. The TEM images show a clear increase of the width for islands capped with an InGaAs alloy. The estimated increase of the width from ~ 12 nm to ~ 18 nm supports the above described qualitative model of AAPS. Below we will always quote the QD widths and heights derived from TEM images taken under diffraction conditions far away from exact Bragg orientation, similar to Fig. 2(c). Figure 3

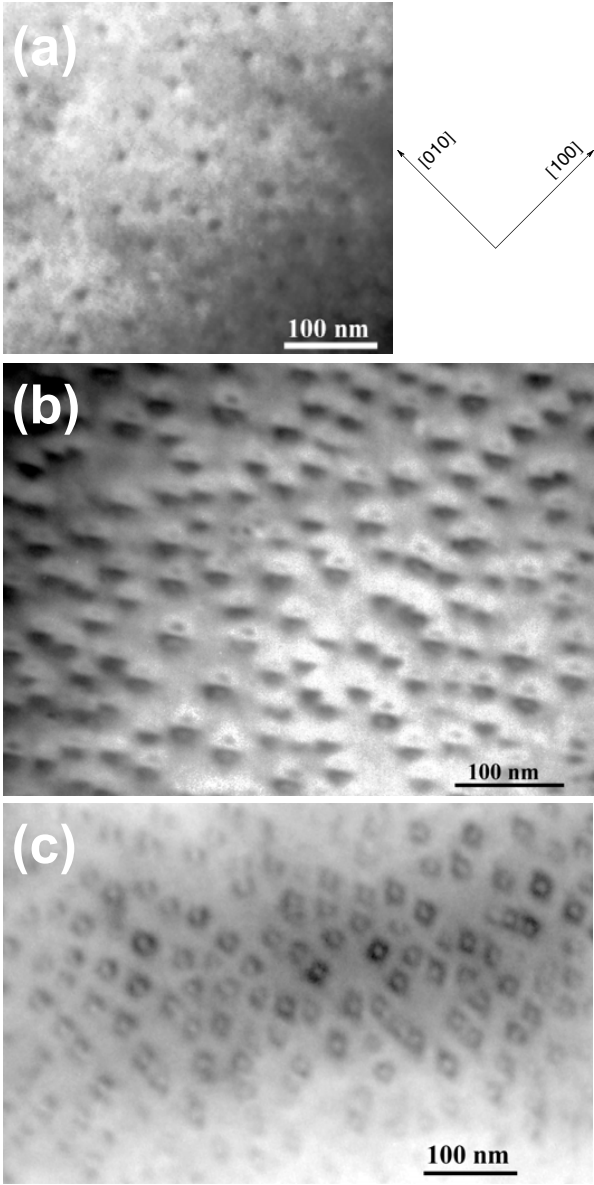


Figure 2. Plan-view electron micrographs of the initial Stranski-Krastanow islands (a) and QDs formed during their overgrowth with 5 nm of $\text{In}_{0.15}\text{Ga}_{0.85}\text{As}$ (b,c). Images (a) and (b) are taken at $\mathbf{g} = 220$, whereas image (c) was taken at diffraction conditions far away from the exact Bragg orientation.

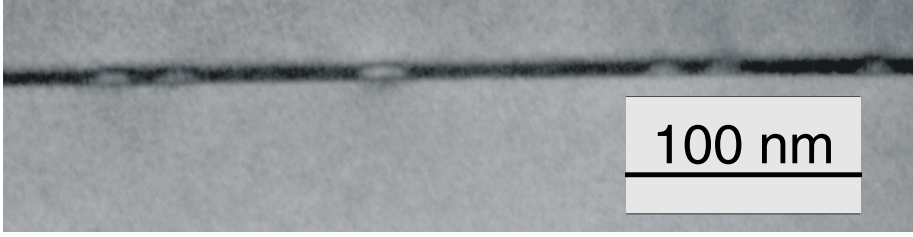


Figure 3. Cross-sectional transmission electron microscopy images of samples with an initial InAs deposition $D_{\text{IS}} = 2.1$ ML overgrown with 5 nm $\text{In}_{0.15}\text{Ga}_{0.85}\text{As}$.

shows a cross-sectional TEM image of the QDs formed by AAPS. Note that the cross-sectional TEM images provide only a rough estimate for the height of the islands due to the uncertain In concentration profile generated during the AASP process.

3. Optimization of Growth parameters during AAPS

From the above discussion it is clear that the AASP process will sensitively depend on the size and density of the InAs nanostressors, i.e. the InAs deposition amount, as well as on the thickness (H), and composition (x) of the $\text{In}_x\text{Ga}_{1-x}\text{As}$ alloy layer. Below we will describe in detail the dependence of the structural and photoluminescence properties of QDs formed by AAPS on the above mentioned growth parameters. To first order, the QD ground state transition energy can be taken as a measure for the island volume to relate structural and optical data. However, the exact shape as well as the In concentration profile in and around the QDs has to be taken into account for detailed predictions of the transition energies.

Figure 4 shows the effect of the amount D_{IS} of InAs deposited in the submonolayer (SML) growth mode on the PL spectra of samples overgrown with 2.5 nm $\text{In}_{0.18}\text{Ga}_{0.82}\text{As}$. Note that the SML growth mode generates larger islands than the conventional continuous one. The QD PL shifts to lower energies with increasing D_{IS} . The concomitant variation of QD size and height obtained from TEM studies (the inset to Fig. 4) is in qualitative agreement with the PL data. The width (height) of the QDs grows from 18 nm (4.9 nm) to 19.6 nm (6.0 nm) when D_{IS} increases from 2.2 ML to 2.7 ML. When D_{IS} exceeds 2.7 ML, the red shift saturates and the PL intensity starts to decrease. TEM images suggest a sudden increase of the dislocation density above a critical deposition of ~ 2.7 ML, providing an efficient means for local strain relaxation. The dislocation density was estimated to $1 \times 10^8 \text{ cm}^{-2}$ for 2.7 ML InAs depo-

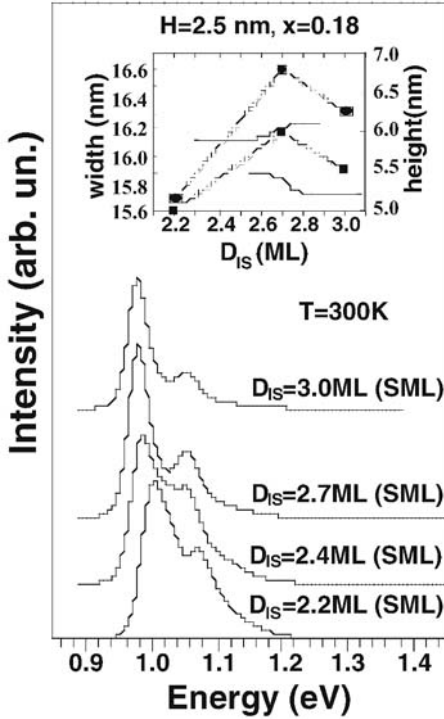


Figure 4. PL spectra of AAPS QDs for various InAs deposition amounts D_{IS} (SML mode) for constant thickness ($H = 2.5$ nm) and composition ($x = 0.18$) of the $\text{In}_x\text{Ga}_{1-x}\text{As}$ alloy. The inset shows the width and height of the QDs determined from TEM images.

sition. Such dislocations have a significant effect on the island evolution on the surface since they attract InAs from both the alloy and already formed QDs. Ultimately, the QD size and height saturates or may even decrease again. Simultaneously, the red shift of PL line saturates at 2.7 ML and the PL intensity drops. For $D_{IS} = 3.0$ ML InAs, the dislocation density increases to $3 \times 10^8 \text{ cm}^{-2}$. Note that the initially elastically strained InAs islands become dislocated only during overgrowth. The total strain energy increases due to the additional InAs accumulation at the islands favoring dislocation formation. Figure 4 suggests an optimum InAs deposition amount for achieving efficient long wavelength emission wavelength maintaining the low dislocation density. The optimum will, however, depend on the composition (x) and thickness (H) of the alloy layer.

Figure 5 compares PL spectra of samples for which the initially deposited 2.1 ML InAs (continuous deposition) have been overgrown with

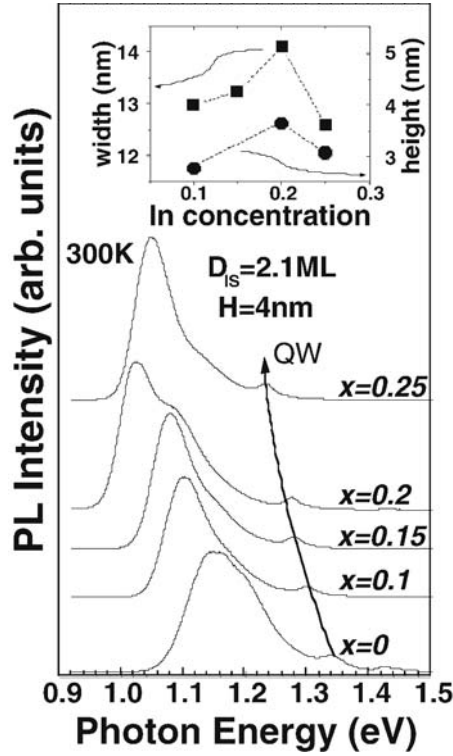


Figure 5. PL spectra of AAPS QDs for various In concentrations x of the alloy for constant InAs deposition amount ($D_{\text{IS}} = 2.1$ ML) and thickness ($H = 4$ nm) of the $\text{In}_x\text{Ga}_{1-x}\text{As}$ alloy. The inset shows the width and height of the QDs determined from TEM images.

a 4 nm thick $\text{In}_x\text{Ga}_{1-x}\text{As}$ alloy layer with In compositions x between 0 and 0.25. The initial islands ($x=0$) show PL at 1.14 eV with a full width at half maximum (FWHM) of 100 meV. With increasing Indium content x the PL peak shifts to lower energies and becomes narrower. The weak emission line appearing ~ 200 meV above the QD ground state transition is attributed to the quantum well (QW) formed by the InAs wetting layer covered by the residual InGaAs alloy layer. The identification of the QW peak is confirmed by PLE studies. The thickness of the QW layer is constant (4 nm) and, thus, the red shift of the QW emission peak with increasing Indium composition is attributed to the decreasing band gap of the alloy layer. Interestingly, the QD and QW PL peaks behave differently with increasing In content suggesting also structural changes. For Indium concentrations x in excess of 0.2, the QW emission is still red shifted, reflecting the increased In content of the residual alloy

layer, whereas the QD emission is again blue shifted. The insert shows the QD width and height obtained from TEM images via the indium content. The QD width (height) increases from 13.2 (2.8) nm to 14.1 (3.7) nm with increase in the In content from 0.1 to 0.2 and subsequently decreases to 12.6 (3.1) nm for $x = 0.25$, which is in qualitative agreement with the evolution of the QD transition energy. The changes in QD volume correlate to a changing localization of the exciton wave function and transition energy. The decrease of QD width and height (as well as the corresponding blue shift) observed for $x > 0.2$ is attributed to the increased formation of dislocations that absorb Indium from the alloy and maybe also from already formed AAPS QDs.

The effect of the thickness of the $\text{In}_x\text{Ga}_{1-x}\text{As}$ alloy layer on the AAPS-formed QDs is shown in Fig. 6 for an In composition $x = 0.18$ and a SML InAs deposition to form initial islands $D_{\text{IS}} = 2.2$ ML. The increase of the thickness of the $\text{In}_x\text{Ga}_{1-x}\text{As}$ alloy layer from 1 nm to 2.5 nm leads to an ~ 100 meV red shift of the ground state transition. Further increasing the thickness results only in a decrease in the PL intensity for practically constant transition energy. The initial low energy shift for small alloy thickness is connected to an increase of the island width (insert). The increase in QD width and corresponding red shift of PL line saturates at a certain critical thickness. Further increasing the thickness leads only to a decrease of the PL yield suggesting enhanced dislocation formation. Again, this effect is attributed to the total strain energy exceeding a critical value for dislocation formation.

The PL and TEM data presented in Figs. 4–6 suggest a narrow range of interrelated values for the InAs deposition (D_{IS}) as well as the $\text{In}_x\text{Ga}_{1-x}\text{As}$ alloy thickness (H) and composition (x) allowing to achieve emission in $1.3 \mu\text{m}$ spectral region maintaining a high PL efficiency and low dislocation density. For the used growth conditions the optimal parameters are $D_{\text{IS}} = 2.5\text{--}2.7$ ML, $x = 12\text{--}15\%$, and $H = 4\text{--}6$ nm.

4. Defect Reduction Techniques

Growth of QDs, especially large-sized with high confinement energy for electrons and holes, may be accompanied by the formation of defects and large dislocated clusters. This problem can be eliminated by applying specially developed *in situ* defect-reduction techniques, which permits selective elimination of dislocated objects without affecting of coherent QDs [15]. A Schematic diagram illustrating reduction of dislocation density is shown in Fig. 7. The QDs are overgrown with a thin (1 nm) GaAs layer [Fig. 7(b)] at the temperature corresponding to that of QD formation (480°C). Then the substrate temperature is raised to

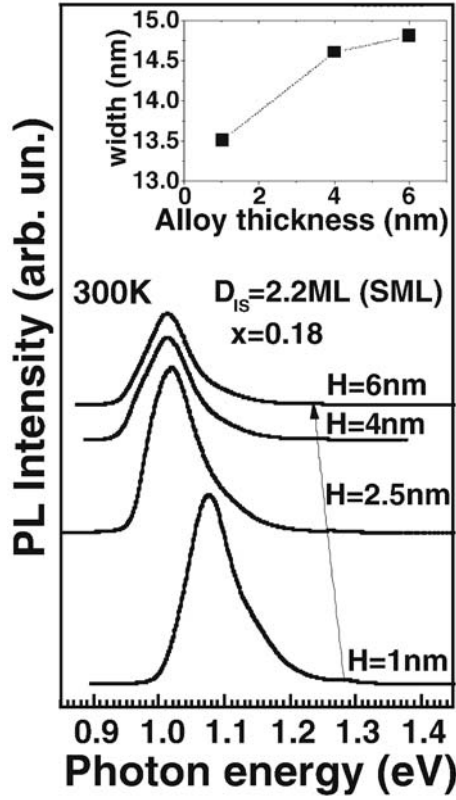


Figure 6. PL spectra of AAPS QDs for various In concentrations x of the alloy for constant InAs deposition amount ($D_{\text{IS}} = 2.1\text{ ML}$) and thickness ($H = 4\text{ nm}$) of the $\text{In}_x\text{Ga}_{1-x}\text{As}$ alloy. The inset shows the width and height of the QDs determined from TEM images.

600°C. The thin cap layer only partly covers the large dislocated InAs clusters allowing the InAs to evaporate in the annealing step [Fig. 7(c)]. The smaller coherent QDs, being completely covered, are only weakly affected. Fig. 8 shows transmission electron microscopy (TEM) images for the samples grown with and without annealing step. The density of dislocated islands in the sample grown without the annealing step is $\sim 10^9\text{ cm}^{-2}$. No dislocated islands are revealed in the structure grown with the annealing step.

5. Lasers Based on QDs Formed by AAPS

The need for $1.3\text{ }\mu\text{m}$ cost-efficient high-performance lasers is currently a powerful driving force for the development of novel GaAs-based QD

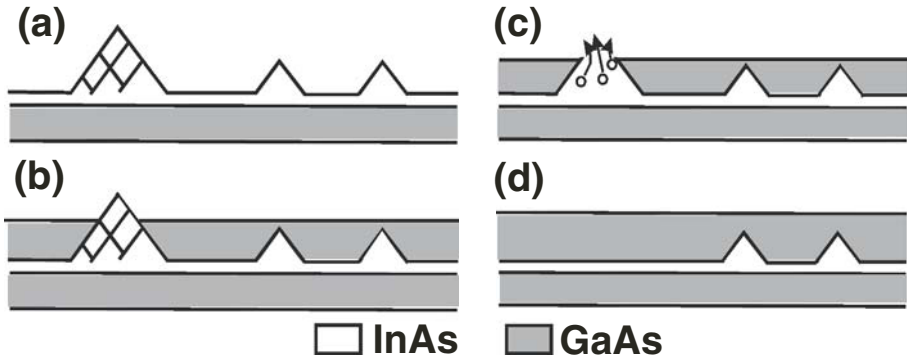


Figure 7. A schematic illustration of the reduction of dislocation density.

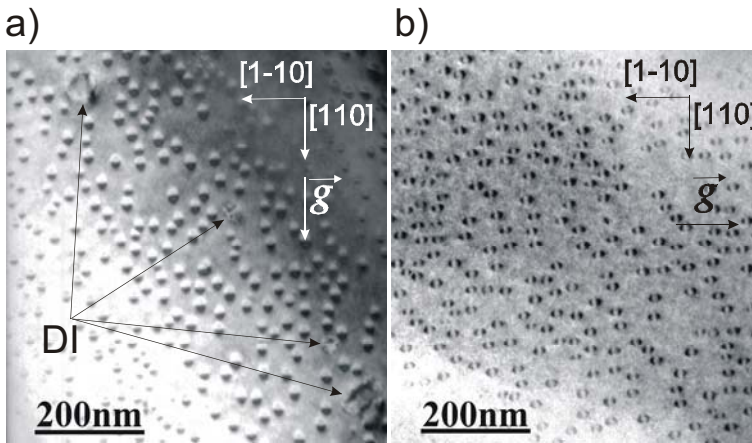


Figure 8. Plan view bright field (220) TEM images of the structures grown (a) without and (b) with (b) the annealing step. DI denotes dislocated islands.

heterostructures. Long-wavelength QD lasers have demonstrated very low threshold current density of 16 A/cm^2 [16]. On the other hand, the low surface density of such QDs results in a small maximum optical gain of the QD ground state. This motivated the use of low-loss cavity design (very long cavities and/or HR/HR facet coatings) at the expense of low external differential efficiency (η_D) and output power. Increasing the number of QD stacks can enhance the saturated optical gain in QD lasers and thus, to improve the efficiency of laser diodes and open new possibilities for their VCSEL application. However in case of large number of QD stacks even small density of dislocated clusters in the first QD sheet results in a drastic deterioration of the properties of upper QD sheets. The use of the defect reduction techniques results in a dramatic

improvement in the characteristics of lasers based on multiple QD stacks (Fig. 9). For instance more than 7-fold decrease in the threshold current is observed for a structure based on 10 QD sheets.

Figure 10(a) shows the dependence of output power on drive current for a 1.5 mm-long laser based on 10 QDs layers along with electroluminescence spectra recorded at different drive currents. The lasing wavelength of $1.28\ \mu\text{m}$ corresponds to the maximum of ground-state emission. High output power (5W) operation considerably widens the lasing spectrum. The threshold current is 220 mA and slope efficiency is $0.78\ \text{W/A}$, corresponding to the threshold current density (J_{th}) of $147\ \text{A/cm}^2$ and η_D of 80%. Temperature dependences of the threshold current presented in Fig. 11(b) is described by the characteristic temperature (T_0) of 150 K in 20–50 °C range. Starting at $1.28\ \mu\text{m}$ at room temperature lasing wavelength gradually shifts toward longer wavelength and reaches $1.3\ \mu\text{m}$ at 64°C [Fig.10(b)].

The dependences of the threshold current density, lasing wavelength and differential efficiency on the cavity length (L) are summarized in Fig. 11. For the 5-QD- and 10-QD-lasers, the threshold current density is in the 100–200 A/cm^2 range for L longer than 1 mm. Stronger dependence of J_{th} on L and shorter emission wavelength of 2-QD-laser are caused by the lower value of maximum optical gain in this structure. Lasers based on 5- and 10-QD layers demonstrate ground-state lasing for

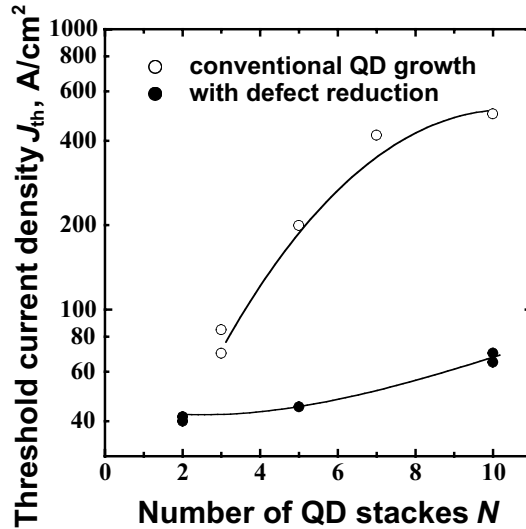


Figure 9. Threshold current densities for the structures with various numbers of QD stacks grown without and with defect reduction. A design with negligible external losses is used.

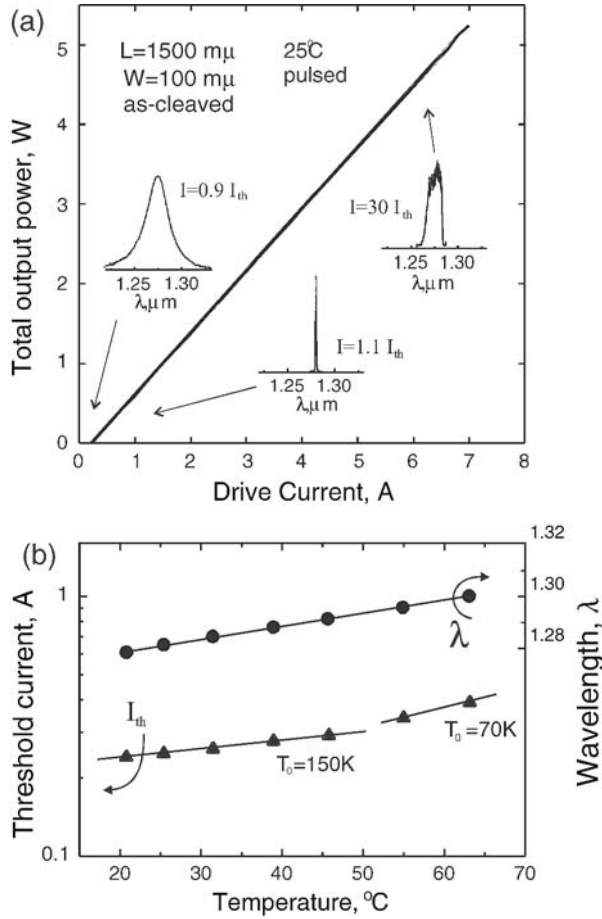


Figure 10. Characteristics of 100- μm -wide and 1.5 mm-long diode based on 10 QD layers. (a) Dependence of the output power, P_{out} , from both facets on drive current, I , at room temperature and emission spectra recorded at different currents. (b) Temperature dependences of lasing wavelength, λ , (solid circles) and threshold current, I_{th} , (solid triangles).

cavities as short as $550\text{ }\mu\text{m}$. For shorter cavity diodes excited state lasing was found. J_{th} increases significantly for diodes shorter than 1 mm, which is accompanied by the blue shift of the lasing wavelength and by steep decrease of differential efficiency. The higher number of QD layers results in a higher value of the differential efficiency for diodes of all measured cavity lengths. We believe that increasing number of QD planes prevents the carrier pile-up in the waveguide. It lowers the internal loss associated with the free carrier absorption and thereby increases η_D for the diodes of the same cavity length. In combination with the extended

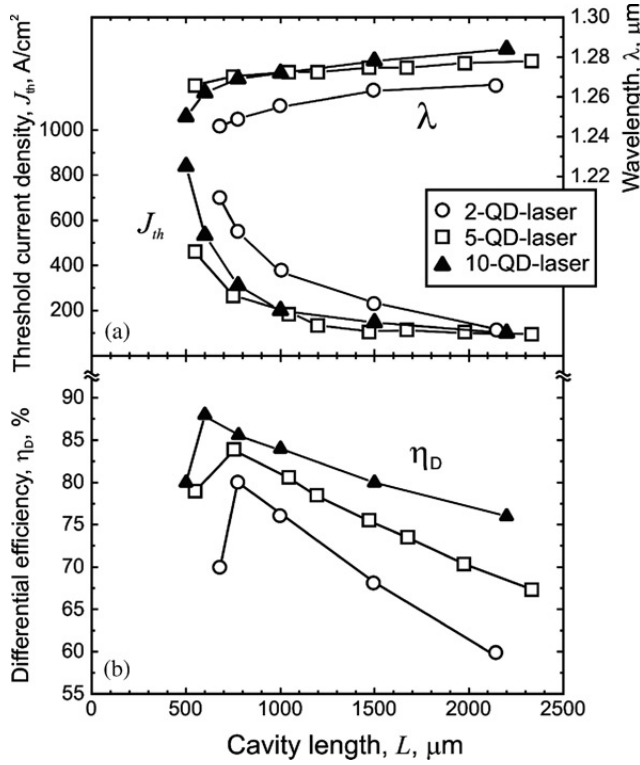


Figure 11. Dependences of the threshold current density, J_{th} , lasing wavelength, λ , and differential efficiency, η_D , on the cavity length, L , for lasers based on 2 (open circles), 5 (open squares) and 10 (solid triangles) layers of QDs.

range of the ground-state lasing (shorter diodes) this leads to the improved maximum external differential efficiency, which is equal to 80%, 84% and 88% for 2-, 5- and 10-QD lasers, respectively. Such high values of η_D indicate that internal quantum efficiencies are close to 100%.

From the dependence of J_{th} on $1/L$ (mirror loss), we estimated that the transparency current density (J_0) is about $60 \text{ A}/\text{cm}^2$ and the maximum gain is well above 30 cm^{-1} for the 10-QD-laser. (The internal loss was found for each L under assumption of $\eta_i = 90\%$ and was taken into account). The obtained value of J_0 corresponds to $6 \text{ A}/\text{cm}^2$ per QD layer. This is in agreement with a simple estimate of the current corresponding to the radiative recombination via the ground state of QDs with a surface density of $3\text{--}4 \times 10^{10} \text{ cm}^{-2}$ and a radiative lifetime of 1–2 ns. This agreement indicates that non-radiative recombination is negligible although as many as 10 layers of QDs are used.

6. 1.5 μm QD Lasers on GaAs Substrates

InAs QDs on GaAs substrates emitting at 1.55 μm are very promising for datacom and telecom applications. However, to achieve this spectral range, InAs QDs of a very large size must be formed [17]. Until recently, no QD lasing in this spectral range was reported, presumably due to the high defect density in such InAs-GaAs structures. Very recently we have shown that good performance 1.5 μm range broad area lasers QD lasers can be realized by using metamorphic $\text{In}_x\text{Ga}_{1-x}\text{As}$ buffer layer [18].

Laser structures were grown at 500°C using elemental-source MBE on Si-doped GaAs(100) substrates. The active region comprised multiple QD sheets deposited on top of a metamorphic $\text{In}_{0.2}\text{Ga}_{0.8}\text{As}$ layer and formed using 2.7 ML InAs deposition and a 4 nm $\text{In}_{0.4}\text{Ga}_{0.6}\text{As}$ layer overgrowth. 45 nm $\text{In}_{0.2}\text{Ga}_{0.8}\text{As}$ spacers were used to separate 10 planes with QDs. The 0.8-1 μm active region with QDs was confined by $\text{In}_x\text{Al}_y\text{Ga}_{1-x-y}\text{As}$ layers doped by Si and Be to $\sim 5 \times 10^{17} \text{ cm}^{-3}$ from the substrate and the surface side, respectively. The 1.6 μm *p*-doped $\text{In}_{0.2}\text{Al}_{0.3}\text{Ga}_{0.5}\text{As}$ cladding layer was followed by an 0.4 μm Be doped (10^{19} cm^{-3}) contact layer. The two structures grown differ in the *n*-doped layers. In one, a 1.2 μm $\text{In}_{0.2}\text{Ga}_{0.8}\text{As}$ Si-doped ($\sim 2 \times 10^{18} \text{ cm}^{-3}$) buffer layer was initially deposited on the GaAs substrate, followed by a 1.6 μm $\text{In}_{0.2}\text{Al}_{0.3}\text{Ga}_{0.5}\text{As}$ Si-doped cladding layer. A secondary electron microscopy image of this structure with the superimposed band gap alignment is shown in the inset to Fig. 12(a). In the other, the two bottom layers were replaced by a 1 μm InGaAlAs layer with a lower molar AlAs content (15%). Both structures have demonstrated similar parameters and we do not distinguish between them in the following. Four-side cleaved samples and broad-area lasers with a 100 μm stripe width were fabricated. No facet coatings were deposited. The lasers were characterised in the 20–85°C range in the pulsed mode (0.2 μs).

Figure 12(a) shows the dependence of the inverse differential efficiency on the cavity length. A high internal efficiency ($> 60\%$) with low losses ($\sim 3 \text{ cm}^{-1}$) are derived. Very high uniformity of the parameters was observed with differential efficiency $\sim 50\%$ for 1 mm-long devices. In Fig. 12(b) we show the dependence of the threshold current density on the inverse cavity length. 2 mm-long cavities demonstrate threshold current densities $\sim 1.5 \text{ kA/cm}^2$. The transparency current, estimated by extrapolation of the threshold current density to the infinite cavity length is $\sim 700 \text{ A/cm}^2$ ($\sim 70 \text{ A/cm}^2$ per single QD sheet). Both the electroluminescence and the lasing spectra in long stripe lasers are shifted towards the longer wavelength with respect to the photoluminescence maximum at 1.45 μm measured in a sample with etched-off contact layer.

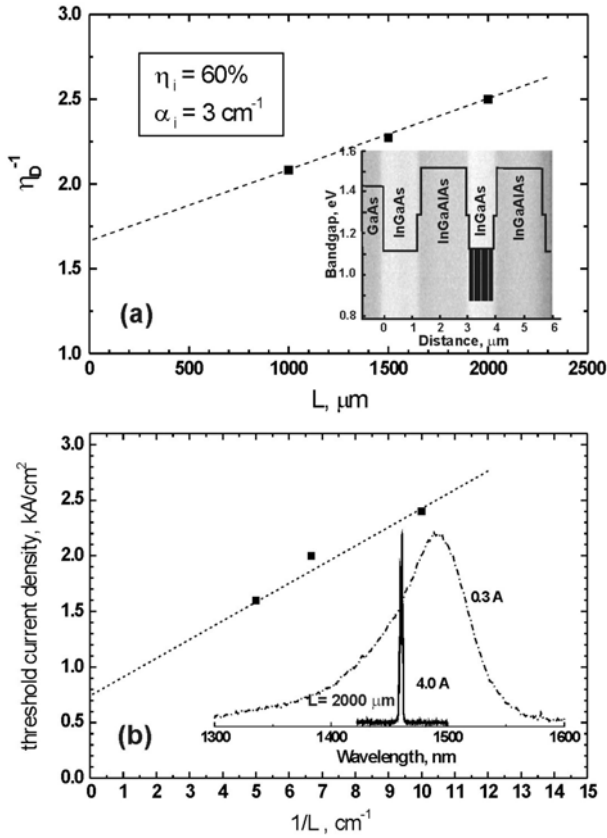


Figure 12. Dependences of (a) the inverse differential efficiency on the cavity length and (b) the threshold current density on the inverse cavity length measured at room temperature. The stripe width is 100 μm . Inset (a): a secondary electron microscopy image of the cleaved metamorphic QD laser structure with the superimposed band gap alignment. Inset (b): electroluminescence and lasing spectra for the sample with a 2 mm cavity length.

The lasing wavelength [inset to Fig. 12(b)] ranges between 1.46 μm and 1.49 μm , depending on the cavity length.

The estimated transparency current density is in a good agreement with the threshold current density measured in four-side-cleaved devices (inset in Fig. 13). The threshold current density increases monotonically with temperature in the temperature range studied (20°C–85°C), limited by the maximum heat sink temperature. The characteristic temperature for the threshold current is 60 K. The emission wavelength exceeds 1.51 μm above 60°C. Output power as high as 7 W is achieved in pulsed mode at room temperature (Fig. 14).

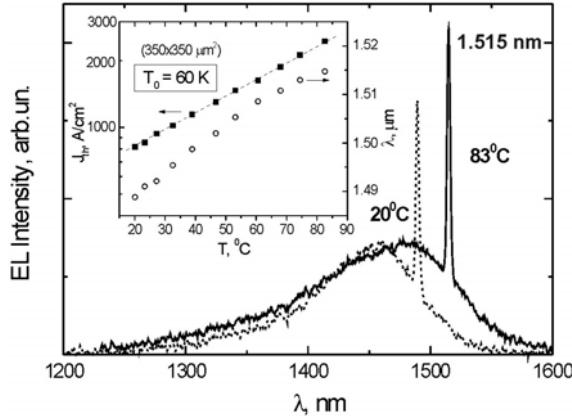


Figure 13. Lasing spectra at different temperatures. Inset: temperature dependences of threshold current density, J_{th} , and lasing wavelength, λ .

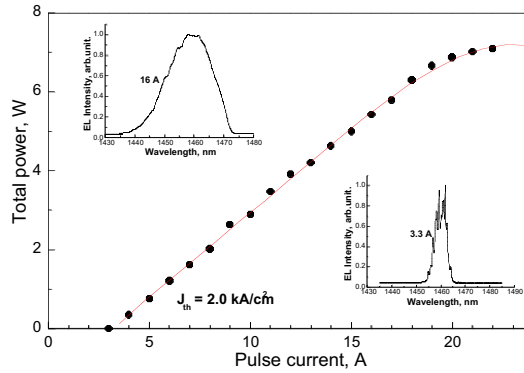


Figure 14. Light-power characteristics of a metamorphic QD laser. Stripe width is $100 \mu\text{m}$; stripe length is 1.5 mm . Inserts show lasing spectra at different drive currents.

7. Conclusions

We have studied QDs formed by activated phase separation of an InGaAs alloy layer grown on InAs nanostressors. Strain-driven surface kinetics was shown to lead to a partial decomposition of the alloy layer, increasing the size of the initial InAs nanostressors. This approach, in conjunction with the use of defect reduction techniques, enables high performance QD lasers to be realised at $1.3 \mu\text{m}$. Quantum efficiency above 60%, low losses and high-temperature operation at $\sim 1.5 \mu\text{m}$ is demonstrated for InAs/InGaAs QD lasers formed on InGaAs metamorphic buffer on GaAs substrates.

Acknowledgements

This work is supported in different parts by the Volkswagen Foundation, INTAS and Nanosemiconductor GmbH, Germany.

References

- [1] N. Kirstaedter, N. N. Ledentsov, M. Grundmann, D. Bimberg, U. Richter, S. S. Ruvimov, P. Werner, J. Heydenreich, V. M. Ustinov, M. V. Maximov, P. S. Kop'ev, and Zh. I. Alferov. Low threshold, large T0 injection laser emission from (InGa)As quantum dots. *Electron. Lett.* **30**: 1416–1417, 1994.
- [2] D. Bimberg, M. Grundmann and N. N. Ledentsov. *Quantum Dot Heterostructures* (Wiley, Chichester, 1999).
- [3] G. E. Tsyrlin, A. O. Golubok, S. Ya. Tipisev and N. N. Ledentsov. InAs/GaAs quantum dots obtained by submonolayer migration-enhanced epitaxy. *Semiconductors* **29**: 884–886, 1995. (*Fizika i Tekhnika Poluprovodn.* **29**: 1697–1701, 1995).
- [4] R. P. Mirin, J. P. Ibbetson, K. Nishi, A. C. Gossard and J. E. Bowers. 1.3 μm photoluminescence from InGaAs quantum dots on GaAs. *Appl. Phys. Lett.* **67**: 3795–3797, 1995.
- [5] I. Mukhametzhanov, R. Heitz, J. Zeng, P. Chen and A. Madhukar. Independent manipulation of density and size of stress-driven self-assembled quantum dots. *Appl. Phys. Lett.* **73**: 1841–1843, 1998.
- [6] V. G. Malyshkin and V. A. Shchukin. Development of composition inhomogeneities in layer-by-layer growth of an epitaxial film of a solid solution of III-V semiconductors. *Semiconductors* **27**: 1062–1068, 1993.
- [7] J. E. Guyer and P. W. Voorhees. Morphological stability of alloy thin films. *Phys. Rev. Lett.* **74**: 4031–4034, 1995.
- [8] Q. Xie, A. Madhukar, P. Chen, and N. P. Kobayashi. Vertically self-organized InAs quantum box islands on GaAs(100). *Phys. Rev. Lett.* **75**: 2542–2545, 1995.
- [9] N. N. Ledentsov, V. A. Shchukin, M. Grundmann, N. Kirstaedter, J. Böhrer, O. Schmidt, D. Bimberg, S. V. Zaitsev, V. M. Ustinov, A. E. Zhukov, P. S. Kop'ev, Zh. I. Alferov, A. O. Kosogov, S. S. Ruvimov, P. Werner, U. Gösele and J. Heydenreich. Direct formation of vertically coupled quantum dots in Stranski–Krastanow growth. *Phys. Rev. B* **54**: 8743–8750, 1996.
- [10] I. L. Krestnikov, M. Strassburg, M. Caesar, A. Hoffmann, U. W. Pohl, D. Bimberg, N. N. Ledentsov, P. S. Kop'ev, Zh. I. Alferov, D. Litvinov, A. Rosenauer, and D. Gerthsen. Control of the electronic properties of CdSe submonolayer superlattices via vertical correlation of quantum dots. *Phys. Rev. B* **60**: 8695–8703, 1999.
- [11] V. A. Shchukin, D. Bimberg, V. G. Malyshkin and N. N. Ledentsov. Vertical correlations and anticorrelations in multisheet arrays of two-dimensional islands. *Phys. Rev. B* **57**: 12262–12274, 1998.
- [12] V. M. Ustinov, N. A. Maleev, A. E. Zhukov, A. R. Kovsh, A. Yu. Egorov, A. V. Lunev, B. V. Volovik, I. L. Krestnikov, Yu. G. Musikhin, N. A. Bert, P. S. Kop'ev, and Zh. I. Alferov, N. N. Ledentsov and D. Bimberg. InAs/InGaAs

- quantum dot structures on GaAs substrates emitting at 1.3 μm . *Appl. Phys. Lett.* **74**: 2815–2817, 1999.
- [13] M. V. Maximov, A. F. Tsatsul'nikov, B. V. Volovik, D. S. Sizov, Yu. M. Shernyakov, I. N. Kaiander, A. E. Zhukov, A. R. Kovsh, S. S. Mikhlin, V. M. Ustinov, Zh. I. Alferov, R. Heitz, V. A. Shchukin, N. N. Ledentsov, D. Bimberg, Yu. G. Musikhin and W. Neumann. Tuning quantum dot properties by activated phase separation of an InGa(Al)As alloy grown on InAs stressors. *Phys. Rev. B* **62**: 16671–16680, 2000.
- [14] Q. Xie, P. Chen, and A. Madhukar. InAs island-induced-strain driven adatom migration during GaAs overlayer growth. *Appl. Phys. Lett.* **65**: 2051–2053, 1994.
- [15] N. N. Ledentsov, M. V. Maximov, D. Bimberg, T. Maka, C. M. Sotomayor Torres, I. V. Kochnev, I. L. Krestnikov, V. M. Lantratov, N. A. Cherkashin, Yu. M. Musikhin and Zh. I. Alferov. 1.3 μm luminescence and gain from defect-free InGaAs-GaAs quantum dots grown by metal-organic chemical vapour deposition. *Semicond. Sci. and Technol.* **15**: 604–607, 2000.
- [16] G. T. Liu, A. Stinz, H. Li, T. C. Newell, A. L. Gray, P. M. Varangis, K. J. Malloy, and L. F. Lester. The influence of quantum-well composition on the performance of quantum dot lasers using InAs/InGaAs dots-in-a-well (DWELL) structure. *IEEE J. Quant. Electron.* **36**: 1272–1279, 2000.
- [17] M. V. Maximov, A. F. Tsatsul'nikov, B. V. Volovik, D. A. Bedarev, A. Yu. Egorov, A. E. Zhukov, A. R. Kovsh, N. A. Bert, V. M. Ustinov, P. S. Kop'ev, and Zh. I. Alferov, N. N. Ledentsov, D. Bimberg, I. P. Soshnikov and P. Werner. Optical and structural properties of InAs quantum dots in a GaAs matrix for a spectral range up to 1.7 μm . *Appl. Phys. Lett.* **75**: 2347–2349, 1999.
- [18] N. N. Ledentsov, A. E. Zhukov, A. R. Kovsh, N. A. Maleev, S. S. Mikhlin, A. P. Vasil'ev, E. S. Semenova, M. V. Maximov, Yu. M. Shernyakov, N. V. Kryzhanovskaya, V. M. Ustinov, and D. Bimberg. High performance quantum dot lasers on GaAs substrates operating in the 1.5 μm range. *Electron. Lett.* **39**: 1126–1128, 2003.