(4)

Sub-50 nm X-ray Lithography with Application to a Coupled Quantum Dot Device

by

David John Donat Carter

A.B., Dartmouth College, (1988)
M.S., Dartmouth College, (1990)

Submitted to the Department of Electrical Engineering and Computer Science in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1998

© Massachusetts Institute of Technology 1998. All rights reserved.

Author ....
Department of Electrical Engineering and Computer Science
May 15, 1998

Certified by ....
Henry I. Smith
Keithley Professor of Electrical Engineering
Thesis Supervisor

Certified by ....
Terry P. Orlando
Professor of Electrical Engineering
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Students

LIBRARIES

Sub-50 nm X-ray Lithography with Application to a Coupled Quantum Dot Device

by

David John Donat Carter

Submitted to the Department of Electrical Engineering and Computer Science
on May 15, 1998, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Electrical Engineering

Abstract

This thesis describes the development of an x-ray lithography process capable of reliably and repeatably exposing arbitrary patterns aligned onto a substrate with feature sizes well below 50 nm. The end result is a robust process which allows device fabrication in a previously inaccessible regime.

The device that motivated this work was a coupled-quantum-dot structure requiring a fine-line Schottky gate with a width less than 50 nm. An overview is given of the device physics and testing, as well as the process required to go from a conceptual device design to an actual pattern on a substrate using x-ray lithography. The work done to improve the technology of soft-contact x-ray replication, both for mask-to-mask replication and for aligned chip or wafer exposures, is described. The first direct experimental observations of the effect of substrate photoelectrons during x-ray exposure, and their detrimental effect on pattern replication, are presented and modeled. Modifications to the x-ray mask replication process which enable reliable mask reproduction in the presence of these substrate-generated electrons are described. Exposure results for sub-50 nm features are shown and statistical analysis of these exposures presented which indicate excellent process latitude in this deep-sub 100 nm regime. Electron transport measurements on devices in GaAs/AlGaAs with fine-line Schottky gate structures are presented, showing that the lithography was successful at producing devices with the desired geometry and electrical characteristics.

Thesis Supervisor: Henry I. Smith
Title: Keithley Professor of Electrical Engineering

Thesis Supervisor: Terry P. Orlando
Title: Professor of Electrical Engineering

Mountains should be climbed with as little effort as possible and without desire. The reality of your own nature should determine the speed. If you become restless, speed up. If you become winded, slow down. You climb the mountain in an equilibrium between restlessness and exhaustion. Then, when you're no longer thinking ahead, each footstep isn't just a means to and end but a unique event in itself. *This* leaf has jagged edges. *This* rock looks loose. From *this* place the snow is less visible, even though closer. These are some things you should notice anyway. To live only for some future goal is shallow. It's the sides of the mountain which sustain life, not the top. Here's where things grow ...

– Robert Pirsig, *Zen and the Art of Motorcycle Maintenance*

Do everything beautifully; don't worry about the outcome – it will take care of itself.

– David Stolper, my high school physics (among other things) teacher

Acknowledgments

The five years spent pursuing this thesis have truly been a wonderful experience, although there were times in those five years when I might not have fully appreciated that fact. Over the course of those years there have been many people to whom I could turn for advice and support during the difficult times, and with whom I could share the good times.

First and foremost I must thank my two co-advisors, Hank Smith and Terry Orlando. Throughout this continually-evolving project, they have both been extremely encouraging and supportive of me and my ideas.

I entered MIT with what I thought was a solid knowledge of microfabrication. From Hank, however, I have learned how to take fabrication to the next level: to really understand, create, debug, and improve upon a process. Hank's infectious enthusiasm and attention to detail permeates the NanoStructures Laboratory, and I feel truly grateful and lucky to have him as a mentor.

I applied to MIT in large part because Terry was here, and have been thrilled to be able to work with him since I arrived. From our many discussions of Coulomb blockade and (early on) Andreev reflection and through his teaching, I have been able to more completely synthesize the engineering mathematics and analysis on which I was weaned scientifically with more traditional physics. This has allowed me to find both an engineer's path around the physics and a physicist's path around the engineering of the electronic properties of materials, microelectronic devices, x-ray interactions with matter, and many other areas of inquiry.

Mark Schattenburg has been very helpful and supportive in the later parts of this work and as a reader of this manuscript. I am grateful that his office is on the indoor route between my office and the NSL, for many occasions when I have popped in to ask a quick question have resulted in the mention of a reference which I might not otherwise have found or the clarification of a point which would have taken me much longer to understand. Dmitri Antoniadis was also very supportive during earlier parts of this work.

I have benefited tremendously from interactions with members of the former QEE research group. Working with Anne Pepin when she was here as a post-doc made tracking down the substrate photoelectron problem a much more enjoyable experience than it would have been had I been tracking alone. I will always remember office discussions with Chagarn Whan about the Coulomb blockade, capacitance matrices, and Mongolian barbecue. Mark Schweizer was a tremendous asset as a colleague and friend – my home computer would not exist without his free almost 24-hour tech support! I have also enjoyed interactions with David Berman and Ilia Sokolinski in discussing various aspects of device fabrication and measurement.

Upon my arrival at MIT, Arvind Kumar was the person who showed me the ropes, from understanding how to traverse the three-dimensional maze of buildings to how to process GaAs and measure quantum dots. I am grateful to him for his patience and for devising the coupled quantum dot experiment. Similarly I thank Martin Burkhardt for sharing his extensive knowledge of x-ray lithography. Most of the process development that was done for this thesis built upon his work.

I have relished the opportunity to collaborate with scientists from other institutions. Kee Rhee at the Naval Research Laboratory has been extremely responsive and attentive to the x-ray mother mask writing requests which I would periodically e-mail to her. I have always been amazed at the quick turnaround time and the high quality of the e-beam writing which she provided for me. When Leo Ocola at the University of Wisconsin (now at Bell Laboratories, Lucent Technologies) gave a talk to our research group in which he showed a plot of a simulation of dose versus height above the substrate for an x-ray exposure, I hoped that he would be interested in a collaboration, and I am very happy to have had the opportunity to work with him. Professor Mike Melloch at Purdue University has been very responsive to requests for substrate materials on which to fabricate devices and in providing ideas for new types of structures.

This work would also not have been possible without the the help of the employees of the NSL. Jimmy Carter has always been quick to lend a hand or sage advice when needed. Jim Daley has been very helpful and understanding, particularly in these

last months when I have been perhaps less amiable than I might have been otherwise. Mark Mondol, Scott Silverman, Cindy Lewis, and Ed Murphy have all at times gone out of their way to aid me in my work.

This project would also not have been possible without financial support from the Air Force Office of Scientific Research and the Defense Advanced Research Projects Agency through Naval Air Systems Command as part of the Defense Advanced Lithography Program. I am particularly grateful to the Department of Electrical Engineering and Computer Science for support during my last semester through the Vinton Hayes Endowed Fund.

Probably the best thing about coming to MIT has been the opportunity to work with and befriend some truly intelligent, stimulating, and enjoyable fellow students. Working in the NanoStructures Laboratory has allowed me to cross paths with many people. I have in particular enjoyed close interactions and camaraderie with Mike Lim, Juan Ferrera, Tom Murphy, and Tim Savas. Two of those four (to remain unnamed) have provided much enjoyment in their creative flaming e-mails that chastise NSL denizens who stray from the path of righteousness. The former cleanroom manager in me is supportive of their efforts. I am also glad to have had the opportunity to get to know and work with James Goodberlet, Maya Farhoud, Keith Jackson, and Euclid Moon.

When not under yellow lights and wearing a bunny suit, I have had the pleasure of occupying office and lab space with Amy Duwel and Enrique Trias. I have enjoyed sharing liquid nitrogen, birthday cakes, and many non-science-related discussions with both of them. Summertime lunches by the river with Enrique were a particularly high point of my tenure as a student.

Perhaps what truly allows one to get through graduate school is being able to commiserate with other students who are not quite as disgruntled as you may be at a particular moment. Thankfully, my mood swings and those of Mark Somerville and Steve Patterson have been slightly out of phase for the most part, which has allowed them to become two of my closest friends. I thank both of them and Enrique for sharing group camping trips and Friday afternoon beers. I have been amazed to

watch Mark at work in a kitchen (or on a beach on a Boston Harbor island), and both my palate and my annual solstice parties have benefited greatly from his skill as a chef. I have (usually) enjoyed being challenged in my many discussion with Steve ranging from the political to the personal and everything in-between.

I would also not have been able to complete this work without the influence of people in my past. My years at Harvard working in the group of Professor Michael Tinkham were where I learned how to think like a physicist. My career as a scientist began at Harvard, and I thank Professor Tinkham and his research group (including in particular Mark Tuominen, Tom Tighe, Rich Fitzgerald, Lydia Sohn, and Jack Hergenrother) for training me in my "pre-doc." If I learned to think like a physicist at Harvard, I learned to think like an engineer at Dartmouth. Professor Al Henning was responsible for most of what I took away from the Thayer School of Engineering at Dartmouth, and I continue to enjoy his friendship. Before learning to think as a physicist or as an engineer, I needed to learn to think analytically. I did that under the tutelage of David Stolper, who was probably the most exciting physics teacher to ever come through Lin-Wood High School. His influence on me, both scientifically and personally, has been profound. I must also thank my step-father, Michel Bujeaud, whose general knowledge of seemingly all things mechanical and electrical was what helped make me curious about the world.

I have had the good fortune to enjoy tremendous support from my mother and father. My mother has always allowed me to make my own choices, and the fact that many of those choices have been good ones is due in no small part to her example as a role model in my life. Especially during these past five years, I have appreciated the growing relationship I have had with my father, whose friendship is something I value dearly.

In addition, numerous friends have supported me outside my work. Kevin Comeau has over the years been a roommate, friend, and soul-mate. I could not place a value on his friendship, support, love, and ability to point things out to me about myself that nobody else could. Carl Deblois is my home-boy, even though we never knew each other when we both lived in the 'hood of the White Mountains of New Hampshire.

I have very much valued the time spent visiting him and his wife Sue Carter and his support and advice over the years. I thank Stewart Lussky for being a friend, roommate, and landlord who provided palatial living quarters for a pauper's rent during my first three years at MIT, making the transition back to the student life infinitely easier than it might have been otherwise. I will always remember her famous Putnam Ave. dinner parties and I cherish her friendship and the friends I met through her. I thank Paul Wickboldt for his friendship, time spent at his family's place in New Hampshire, and for being a calm source of perspective when I've needed it. Tom Daniels and Mickey Zibello have also been great friends. If anything in my life relieves stress and provides a non-scientific creative outlet, it is music, and the many years we have spent playing and more recently performing together as Sacramento Street Acoustic Trio have provided me with much joy.

Lastly, I wish to acknowledge April Robb, with whom I have had the honor and pleasure of sharing my life for the past two years. In meeting her I found what I had been looking for, and her love and support have been instrumental both in this work and in helping me to grow in ways I had never imagined.

Contents

Abstract 2

Acknowledgements 4

1 Preface 19

2 Background and Motivation 23

2.1 A New Paradigm for Computation 23

2.2 Quantum Dots and the Coulomb Blockade 26

2.2.1 Classical Coulomb Blockade 26

2.2.2 The Quantum Dot 31

2.3 A Quantum Dot Cellular Automata Scheme 36

2.4 Quantum Dot Interactions 38

2.4.1 Interdot Interactions 38

2.4.2 Dot-to-Lead Interactions 39

2.5 Coulomb Blockade at High Temperatures 40

2.6 A Novel Coupled Quantum Dot Structure 43

2.6.1 A Non-Point-Contact Tunnel Barrier 43

2.6.2 Fine-Line Barrier Width Considerations 43

2.7 Motivation for Improved Sub-50 nm X-ray Lithographic Process . . . 47

3 X-ray Nanolithography 49

3.1 Alternative Lithographies 49

3.2 Overview of the Lithographic Process 51

3.3 Cu_L X-Ray Exposure Technology 53
3.3.1 Lithographic Input in X-ray Lithography 53
3.3.2 Resist Dose 61
3.4 Device Fabrication with X-ray Nanolithography 68
3.4.1 Overview of Full Fabrication Process 69
3.4.2 Mother Mask Fabrication 70
3.4.3 Daughter Mask Replication 74
3.4.4 Device Patterning 78
3.5 Previous Pattern Replication Results 78
3.5.1 Pattern Replication Below 50 nm 78
3.5.2 Previous Device Patterning Results 79

4 Soft Contact Technology **81**
4.1 Microgap X-ray Lithography and Limits to Replicable Feature Sizes . 81
4.2 Previous Methods for Reducing Mask-to-Sample Gap 82
4.2.1 Gap Reduction for Mask Daughtering 82
4.2.2 Gap Reduction for Small Samples 84
4.3 Full Vacuum Contact 89
4.3.1 Vacuum Contact for Daughter Exposures 89
4.3.2 Aligned Vacuum Contact for Small Samples 91
4.4 Analysis of Issues in Soft-Contact 95
4.4.1 Squeeze-Film Analysis of Membrane Dynamics 95
4.4.2 Membrane Strength 101

5 Substrate Photo- and Auger Electrons **105**
5.1 Unsuccessful Cu_L X-ray Daughter Exposures 106
5.2 Position Dependence of Dose in X-ray Lithography 109
5.3 Substrate Effects Reported in the Literature 111
5.4 Experiments and Simulations 112
5.4.1 Experimental Setup 112
5.4.2 Simulation Parameters 113

5.4.3 Results 115
5.5 Successful Daughter Mask Replication 119

6 Sub-50 nm Exposure Results **125**
6.1 Individual Sub-50 nm Replication Results 125
6.2 Statistical Analysis of Sub-50 nm Replication 128
6.2.1 Description of Experiment 128
6.2.2 Linewidth Measurement Results 133

7 Transport Measurements **144**
7.1 Conductance vs. Gate Voltage Curves 145
7.2 QPC Pinchoff Characteristic 148
7.3 Tunnel Barrier Pinchoff Characteristic 148

8 Conclusions **152**

A Mask Fabrication **156**
A.1 Mother Mask Preparation for NRL 157
A.2 Developing and Plating a Mother Mask 159
A.3 Suck-and-Puff Daughtering 166
A.4 Developing and Plating a Daughter Mask 168
A.5 Polyimide Coating of Masks 174
A.6 Clearing Polyimide from a Mask or Part of a Mask 178

B Device Fabrication **179**
B.1 Mesa Isolation 180
B.1.1 Solvent Cleaning of Sample and Surface Clean 180
B.1.2 Spin and Bake Photoresist 181
B.1.3 Photolithography and Development 181
B.1.4 Etch AlGaAs 182
B.2 Ohmic Contacts Processing 184
B.2.1 Solvent Cleaning of Sample and Surface Clean 184

B.2.2 Spin and Bake Photoresist 184
B.2.3 Photolithography 185
B.2.4 Evaporation of Ohmic Contact Metal 185
B.2.5 Annealing Contacts 187
B.2.6 Check Contact Resistances 187
B.3 Gate Processing 188
B.3.1 Solvent Cleaning of Sample and Surface Clean 188
B.3.2 Spin and Bake PMMA 188
B.3.3 Small Sample Vacuum Contact Exposure 189
B.3.4 Removal of Small Chip from 2" Wafer 191
B.3.5 Development and Schottky Gate Evaporation 191
B.4 Patching Holes and Fixing Shorts 194

C Electronics and Cryogenic Design for Device Measurements 195
C.1 Cryogenic and Electronic System Design 195
C.1.1 Cryogenic System 196
C.1.2 Design of Measurement Electronics 200
C.2 ^{3}He Probe Operation 204
C.3 Device Measurement Difficulties 209

Bibliography 214

List of Figures

1-1 Minimum semiconductor device feature size versus year 20

2-1 Schematic of cellular automata interactions 24

2-2 Schematic of small metal island 27

2-3 Energy of small metal island versus gate voltage, with schematic Coulomb blockade peaks 28

2-4 GaAs/AlGaAs heterostructure and simulated conduction band level at 4K 32

2-5 Schematic of laterally-patterned constriction (a quantum point contact) in a 2DEG 34

2-6 Quantum dot cellular automata computation scheme, showing the unit cell or automata and a programmable AND/OR gate 36

2-7 Schematic of coupled quantum-dot device with fine-line tunnel barrier separating the two dots 44

2-8 Simulations of the conduction band energy for 40- and 80-nm-wide fine-line gate structures 45

3-1 Schematic of optical and e-beam lithography 50

3-2 Schematic overview of lithographic process 52

3-3 Schematic of x-ray exposure 54

3-4 Schematic of Auger electron and x-ray generation 55

3-5 Copper target x-ray photon spectrum 57

3-6 Attenuation of several materials at soft x-ray wavelengths 59

3-7 Schematic of x-ray mask 60

3-8 X-ray photon spectra in exposed and dark regions 62
3-9 Photon point spread function (PSF) for a 1.3 keV photon absorbed in PMMA 65
3-10 Penumbral blur 66
3-11 Mask-to-sample gap vs. minimum printable linewidth 67
3-12 Full x-ray fabrication sequence for a device 69
3-13 SEM image of test pattern on mother mask showing remarkable lack of proximity effect 72
3-14 SEM image of device pattern on mother mask 73
3-15 AFM image of device pattern on mother mask 74
3-16 Schematic of gap monitoring in microgap exposures 75
3-17 Simulation of reflectance of mother/daughter mask pair vs. angle with and without $1\,\mu$m polyimide coating on the mother mask 76

4-1 Schematic of electromechanical model for electrostatic contact 86
4-2 Optical micrograph of GaAs chip and membrane shard destroyed by arcing when attempting electrostatic contact 88
4-3 Schematic of "suck and puff" technique for performing daughter exposures in intimate contact 90
4-4 Schematic of vacuum contact technique for performing aligned small-chip exposures in intimate contact 92
4-5 Photograph of small GaAs chip in contact with x-ray mask 94
4-6 Schematic of membrane deflection for single mask and mother-daughter mask pair when placed under vacuum 95
4-7 Time-lapse photograph sequence of two masks coming into vacuum contact 96
4-8 Model system and initial pressure profile for squeeze-film analysis . . 98
4-9 Plot of pressure profile as a function of time for simplified squeeze-film model 100
4-10 Photograph of membrane vs. wooden-sticked cotton swab 103

5-1 Scanning electron micrograph of unsuccessful daughter replication on a mask with "thick-gold" plating base with PMMA still present . . . 106
5-2 Scanning electron micrograph of unsuccessful daughter replication on a mask with "thick-gold" plating base with PMMA removed 107
5-3 Scanning electron micrograph of a daughter pattern in PMMA on silicon coated with "thick-gold" plating base 108
5-4 Schematic of x-ray absorption and photo- and Auger electron generation during x-ray exposure showing the effect of the substrate 110
5-5 Screen shot of AFM image and height histogram 114
5-6 Measured PMMA dissolution rate versus height above the substrate for substrates of thick-gold plating base on silicon, thin-gold plating base on silicon, and bare silicon 116
5-7 Simulations of exposures for exposed and "dark" regions 117
5-8 Two-dimensional simulation of a half-plane x-ray exposure on thick-gold plating base 118
5-9 Scanning electron micrograph of a plated daughter exposure on thick-gold plating base coated with a thin layer of polyimide and PMMA . 120
5-10 Scanning electron micrograph of a successfully-plated daughter with ~25 nm feature sizes done on thin-gold plating base 122
5-11 Scanning electron micrograph of a successfully plated daughter with a ~20 nm fine line 123
5-12 Scanning electron micrograph of a sub-50 nm starburst pattern on a daughter mask 124

6-1 Scanning electron micrographs of device pattern replication by liftoff onto a substrate with sub-40 nm linewidths 126
6-2 Scanning electron micrograph of a sub-50 nm test pattern replicated by liftoff onto a substrate 127
6-3 One cell of mask design for sub-50 nm exposure experiments 129

6-4 Screen shot of the program `extraxt` used to extract linewidths from a TIFF image of a grating 131
6-5 Plot of standard deviation of linewidth measurement on mother mask versus measured linewidth 134
6-6 Plot of average standard deviation of linewidth measurement on daughter mask versus measured linewidth 135
6-7 Plot of difference between two sets of linewidth measurements of the same die on the daughter mask 136
6-8 Plots of difference between mask and liftoff pattern linewidth measurements versus mask linewidth for differing development times 138
6-9 Plot of difference between lifted-off pattern linewidth measurements versus linewidth for differing development times 140
6-10 Plot of linewidth change from mask versus binned mask linewidths for differing development times 141
6-11 Plot of linewidth change from nominal versus development time for mask feature sizes of 30, 40, 50, and 60 $\pm$ 2.5 nm 142

7-1 Scanning electron micrograph of coupled quantum dot Schottky gates on GaAs/AlGaAs heterostructure 146
7-2 Conductance vs. Gate Voltage for the fine-line tunnel barrier and a QPC on the coupled quantum dot device 147
7-3 Conductance vs. Gate Voltage for QPCs 149
7-4 Conductance vs. Gate Voltage for a fine-line tunnel barrier 150

8-1 Scanning electron micrographs of the same structure patterned on a mother mask, daughter mask, and device substrate 153

C-1 Schematic of Oxford Instruments Heliox insertion probe for 300 mK device measurements 197
C-2 Schematic of modifications made to the sample mount for improved ease of use and performance 199

C-3 Schematic of measurement circuit designed for measuring QPCs, tunnel barriers, and quantum dot devices 202
C-4 Pinchoff characteristics of QPCs and tunnel barrier showing change in pinchoff voltages and nature of tunnel barrier conductance with time 210
C-5 Measurements of Schottky barrier characteristics at 4.2 K 212

List of Tables

2.1 Table of parameters to fit simulated fine-line tunnel barriers and tunneling probabilities for the barriers calculated using the WKB approximation 47

3.1 Calculated absorbed dose in J/cm^3 in 250 nm PMMA for a 12 hour x-ray exposure at a 10.5 cm source-to-substrate distance. 63

3.2 Table of single-pass electron-beam linewidths as a function of dose. . 70

4.1 Minimum feature size vs. gap presented in tabular form. $\alpha = 1.5$ gives the largest gap which has been shown to yield good processing latitude. 81

Chapter 1

Preface

The technology of a product and the technology used in its manufacture often advance hand in hand. Developments in manufacturing techniques create improved products, whose use will inspire ideas for further changes to the product. To make these changes, modifications to the manufacturing processes are often required, which begins the cycle again. This has been true from the earliest discoveries of basic metallurgy to early assembly line manufacturing of automobiles and guns, and it continues to be true today.

Microelectronic device technology and microlithography are an excellent example of this partnership. As devices are designed that require feature sizes which bump up against the limits of available fabrication technology, lithography is called upon to create smaller structures. Once smaller features are manufacturable, the device technology makes good use of them, which again pushes the lithographic technology.

The fruitfulness of this interaction is immediately evident when one examines the semiconductor industry. From the beginnings of integrated circuits (ICs) in the late 1950's [1], the performance, density, and cost of semiconductor devices have improved exponentially with time. This trend was first noted in 1965 by Gordon Moore (then Director of Research and Development for Fairchild Semiconductor, now Chairman Emeritus of Intel Corporation), who noted that the number of components on an IC had doubled every two years since the invention of the planar transistor in 1959 [2]. Since then the so-called Moore's "Law" has been used to describe "almost

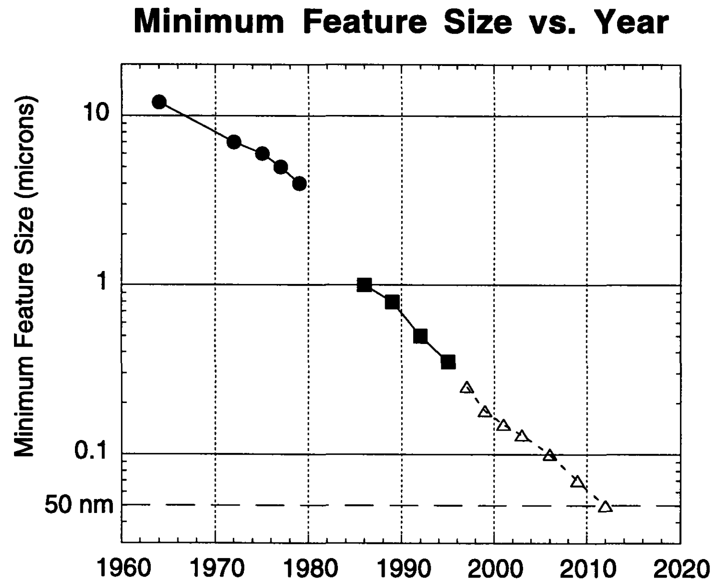


Figure 1-1: Semiconductor device minimum feature size versus year. Note the straight-line trend on the semi-log plot, with a prediction that 50 nm feature sizes will be in commercial production in the year 2012. Data from 1980 and before, plotted with filled circles, are from reference [5]. Numbers after 1980 are for dynamic random-access memory (DRAM) gate lengths. Data from 1986-1995, plotted with filled squares, are from reference [6]. Open triangles represent predictions from reference [4].

anything related to the semiconductor industry which when plotted on semi-log paper approximates a straight line" [3].

In fact Moore's "Law" has become self-enforced, as it is used as a guideline for planning future technologies in the Semiconductor Industry Association (SIA) Technology Roadmap [4]. Figure 1-1 shows a plot of minimum feature sizes of commercial semiconductor products versus year. The data points decrease more or less exponentially, with a prediction that 50 nm feature sizes will be in production in the year 2012.

The payoff from this exponential trend has been enormous. The miniaturization of silicon metal-oxide-semiconductor field-effect transistors (MOSFETs) has given us today's powerful computers which perform high-speed digital logic operations while consuming very little power. The public takes for granted that next year a faster and more powerful model will make current computers obsolete. As the turn of the millennium approaches, however, the continuation of this trend is not certain.

As conventional silicon MOSFETs are shrunk further into the deep-submicron regime, their performance degrades because of short-channel and hot-electron effects. Perhaps a more pressing problem is the wiring crisis: as the density of transistors increases with decreasing feature size, the interconnects between transistors become more congested, and crosstalk between signals can increase, while increasing clock speeds can make delays due to wiring the limiting factor in chip speed.

The potential limitations posed by these issues are specific to MOSFET technology which has to this point proven to be so highly successful. As the problems with this technology become more difficult to overcome, alternatives must be explored which may allow a way around the roadblocks mentioned above.

Even in the absence of MOSFET-specific problems, however, the exponential trend in device size reduction becomes increasingly difficult to maintain as device dimensions are pushed into the deep-submicron regime. Optical lithography is the technology which is used at present to define features on ICs. Although it has been tremendously successful to date, it is unlikely to be a viable manufacturing vehicle for feature sizes below 130 nm, and there is some question as to whether it will work below 150 nm [4].

In addition to microelectronics, there are other technologies, such as integrated optics and microelectromechanical systems (MEMS), which are pushing lithographic requirements. Many of these technologies were born from microfabrication capabilities developed in large part by and for the semiconductor industry. These technologies are now themselves requiring smaller feature sizes and dimensional control whose requirements can exceed those of microelectronics.

In order to manufacture micro- and nanometer-scale devices in the "post-optical" era, alternative technologies will have to be developed to a state where they are suitable for high-volume manufacturing. Several potential candidate technologies exist, including x-ray lithography, electron-beam lithography, and extreme-ultraviolet (EUV) lithography. None has been successfully used even for pilot production in the sub-100 nm regime. Of these technologies, I believe that x-ray nanolithography offers the best combination of present development and potential for scaling to production levels.

This thesis describes an example of an interaction between lithographic technology and microelectronic device technology. An interest in exploring non-MOSFET microelectronics led to a proposed device which required sub-50 nm features in order to operate properly.

Although sub-50 nm features had been created with x-ray nanolithography in the past, it had been done only as a demonstration to show its capability. The extension of this earlier work to create a robust x-ray lithography process capable of fabricating device patterns on a wafer with feature sizes below 50 nm is the subject of this thesis.

Chapter 2

Background and Motivation

2.1 A New Paradigm for Computation

One set of candidate technologies which could potentially overcome problems associated with MOSFET devices at very small size scales is quantum-effect electronics [7]. These technologies seek to exploit quantum-mechanical phenomena in semiconductor nanostructures to achieve enhanced functionality of devices. Several devices which use quantum-mechanical principles have been proposed and tested [8–13].

If a suitable quantum-mechanical transistor were available, using it to replace conventional transistors would seem to be an attractive possibility. However, the problems associated with increased interconnect density would still exist. A more radical departure from conventional electronics would be a design in which computational elements interacted directly with one another.

One possible scheme for computation in this manner is a cellular automata computer. A conventional cellular automata system is a mathematically-defined structure comprised of an array of individual units (automata) which interact only with their (usually nearest) neighbors and only at controlled times. Each element or "computational primitive" chooses its next internal state based upon its present internal state and the state of its neighbors. Translating from this abstract concept to an actual computer comprised of physical elements, however, is not an easy task.

The main problem is that while the idealized cellular automata of the mathe-

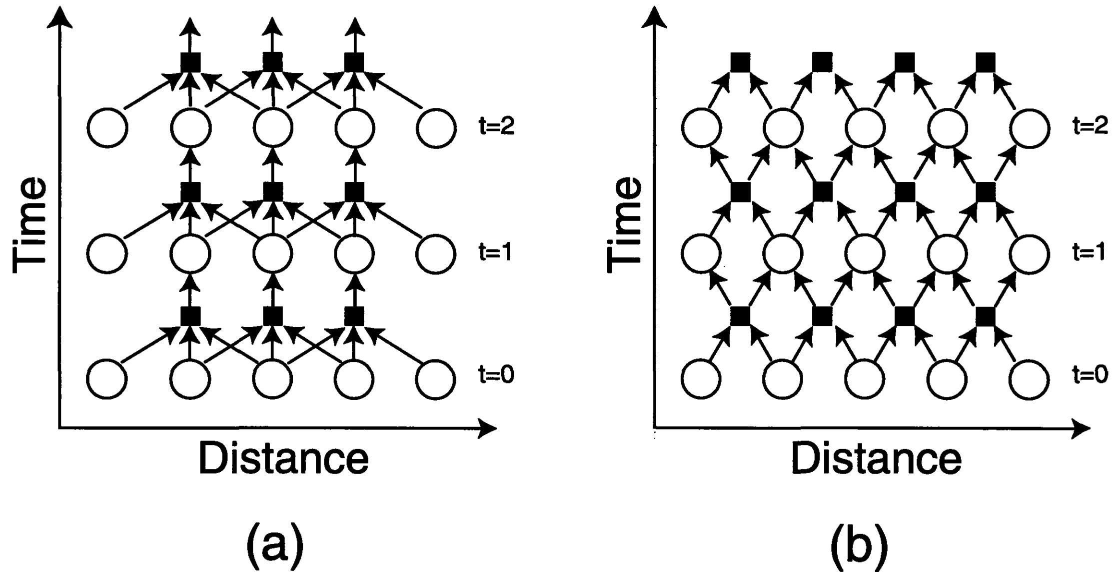


Figure 2-1: Schematic of cellular automata (CA) interactions. Each cell or computational primitive is represented by an open circle. Interactions are represented by a filled square. (a) Conventional CA. Elements interact with their nearest neighbors at regular time intervals according to defined rules. Each element chooses its next state based upon its present state and the state of its nearest neighbors. (b) Few-body CA. Elements only interact when their information is brought into proximity with their nearest neighbors. In the two-body scheme pictured, alternating pairs interact at each clock cycle. This type of interaction resembles a sequence of scattering events.

matician interacts only with its neighbors and only at controlled times, real, physical systems are rarely so well-behaved. They tend to interact in a more long-range manner than only by nearest-neighbor interactions, and they tend to interact continuously. There is one cellular automata structure which has been proposed which makes a serious attempt to get around these problems. It is known as a few-body cellular automata [14–16]. It allows the use of real, physical entities for cellular automata computation.

The idea behind the few-body cellular automata is similar to that of scattering processes between particles. In the scattering case, particles approach each other and interact when they are in close proximity. After the interaction, they separate and effectively no longer interact. The way in which this is implemented in the few-body cellular automata scheme is to place the computational primitives at a far enough distance that they would not normally interact. When interaction is desired, the

information in each primitive is transported in some manner to a location where the interactions are allowed. Once the interaction occurs, the resulting primitives and their information are again removed from each other.

Given a primitive which has an appropriate interaction range and which interacts with other primitives in a predictable manner, computation with cellular automata can be achieved. A schematic picture of the differences between a conventional cellular automata interaction and a few-body cellular automata interaction is shown in Figure 2-1.

Even though the concept of a few-body cellular automata brings us much closer to being able to physically realize a cellular automata computer, it is still as of yet an abstract concept with no physical implementation. However, by analyzing its requirements we can begin to understand how we might use available nanostructures to make a suitable computational primitive to realize the model.

Any available nanostructure must meet the following criteria [16]:

- There must be localized physical states that can represent binary logical states.
- It must be possible to transport the localized physical states without changing the part that represents the logical state.
- The localized physical states must interact via a potential which is not too long-ranged.
- The effect of the interaction on the part of the physical state used to represent the logical state must correspond to a deterministic few-body rule.

From this list we see that the nature and the range of interactions between nanostructures is critically important.

We will focus on a quantum dot device. In order to better understand how such a device might be suitable for use in cellular automata computation, it is necessary to understand how quantum dots interact with each other and with their environment. First, however, we discuss the physics behind the Coulomb blockade and describe what a quantum dot is.

2.2 Quantum Dots and the Coulomb Blockade

Quantum dots in semiconductor nanostructures exhibit the effects of the Coulomb blockade. Some general discussions of the Coulomb blockade in quantum dots and metallic structures are given by van Houten *et al.* [17], Kastner [18], Kouwenhoven *et al.* [19], and Kouwenhoven and McEuen [20]. We begin with a review of the basic theory of the Coulomb blockade in metal grains. We then discuss the formation of a quantum dot in a GaAs/AlGaAs heterostructure, and investigate how it differs from a metallic structure.

2.2.1 Classical Coulomb Blockade

The Coulomb blockade arises because of the electrostatic energy required to place an electron on an isolated conducting island. In the normal, macroscopic world, Coulomb blockade effects are not seen because this energy is very small compared to the thermal energy $k_B T$ available at room temperature. When the size of a conducting island becomes very small or if the temperature is reduced to near absolute zero, however, the effect of this Coulomb energy can be seen. Theoretical analyses of the Coulomb blockade which are more complete than and complementary to what is presented below are given in Averin and Likharev [21] and van Houten *et al.* [17].

Energy of an Isolated Conducting Island

Let us imagine a small metal grain or island which is isolated from its environment. This grain is capacitively coupled to three conductors which are labeled l (for left), r (for right), and g (for gate). This is pictured schematically in Figure 2-2. We assume that any charge on the island is imaged solely on the three conductors so that the total capacitance of the island to the outside world is $C_d = C_l + C_r + C_g$.

The electrostatic energy U of this island as a function of its charge Q_d can be written as:

$$U(Q_d) = \int_0^{Q_d} V_d(Q_d')dQ_d',$$

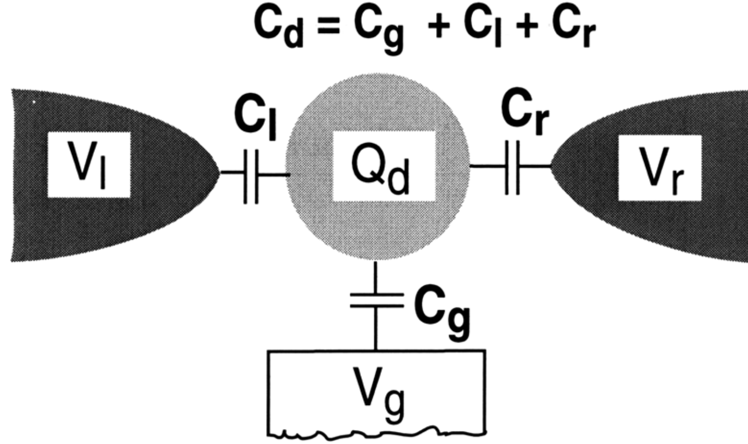


Figure 2-2: Schematic of metal island which experiences Coulomb blockade. Quantities related to the island (or dot) are identified with the subscript d. Also shown are leads l (left) and r (right), and a gate g.

where V_d represents the voltage on the island (relative to the voltage on l and r, which are at ground). Let us assume that there is some voltage applied to the gate g. In this case we have

$$V_d = \left(\frac{C_g}{C_d}\right) V_g + \frac{Q_d}{C_d}.$$

We also know that charge is quantized, so

$$Q_d = -Ne,$$

where e is the absolute value of the electron charge.

Performing the integration with the above substitutions and constraints gives the energy of the island as a function of normalized gate voltage V_g/e parameterized by the number of electrons on the island N:

$$\frac{U_N(\frac{V_g}{e})}{e^2} = -N\frac{C_g}{C_d}\left(\frac{V_g}{e}\right) + \frac{N^2}{2C_d}.$$

For a given N, this is a straight line as a function of V_g. Lines of different N will have slopes which become more negative and y-intercepts which increase with increasing N. A plot of the family of lines for the values $C_g = 1$ and $C_l = C_r = 0.5$

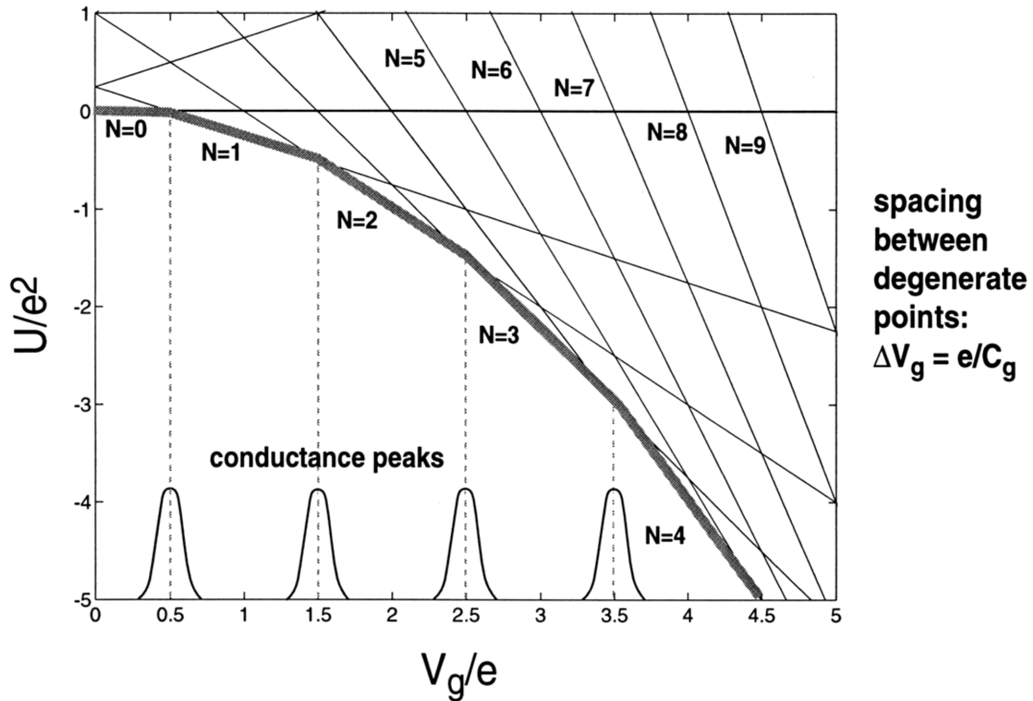


Figure 2-3: Energy of small metal island versus gate voltage. For a given number of electrons N on the island, the energy is linear with gate voltage. Increasing N yields lines with a more negative slope and a larger y-intercept. At low temperature and small bias voltage, current can only flow through the island when the energy of the state with N electrons is equal to the energy of the state with $N+1$ electrons. "Zero-bias" conductance peaks which have been broadened by temperature and/or bias voltage are schematically shown at the bottom of the plot.

is shown in Figure 2-3.[1]

Electron Population of Island and Current Flow

If we start out with one electron on the island and sweep the gate voltage, the system will want to remain in its lowest-energy state. However, if no electrons are allowed

[1]The usual picture one sees in the literature is of parabolas instead of lines corresponding to different N's. This picture is equivalent to the model presented here at a given gate voltage, *i.e.* the *relative* energies for a given gate voltage are the same in the two models. The difference lies in the fact that here we have neglected the electrostatic energy of the gate charge $Q_g^2/2C_d = C_g^2V_g^2/2C_d$. Not including that energy makes it look like the system, if left to itself, would quickly run to infinite gate voltage, which is clearly non-physical. The advantage of ignoring the gate charge in our model is that, particularly for double-dot systems, the results of changing parameters are easier to see graphically with lines than with parabolas.

on or off the island, N cannot change and as the gate is swept the dot energy will follow the $N = 1$ line. We can see, for example, that at zero gate voltage that $N = 0$ would be energetically favorable.

If we allow tunneling of electrons onto and off of the island (say through the left and right capacitors), then as the gate voltage is swept, the number of electrons on the island should increase as the gate voltage is raised and the island seeks its minimum energy.

Let us now assume that we have applied a very small bias voltage on the left lead (this voltage is small enough that the imaged charge on the island due to it is negligible compared to the charge from the gate electrode). We continue to assume that electrons can tunnel through the left and right capacitors if a tunneling event is energetically favorable. Under what conditions will current flow through the island?

Current will be prevented from flowing unless the charge on the island can fluctuate between N and $N+1$. For very low temperatures and a very small "zero-bias" voltage, this can only happen when the energy of having N and $N+1$ electrons on the island is degenerate. We would therefore expect current to flow through the island only when $V_g/e = n + 1/2$, where n is an integer, as this is where the energies are equal, as can be seen from the diagram.

The current through the island as a function of gate voltage is a series of peaks, separated in voltage by e/C_g. Finite temperature and applied bias voltage will smear out the conductance peaks [22], which are sketched at the bottom of the plot in Figure 2-3. By controlling a gate voltage, one can change the number of electrons on the island one at a time. For this reason Coulomb blockade effects are often referred to as single electronics, and a gated island which exhibits the effect is referred to as a single-electron transistor (SET).

In addition to the zero-bias conductance, the I-V characteristic of a small metal island can be affected by the Coulomb blockade. At a gate voltage which is not at a degeneracy point in energy for the states N and $N+1$, a finite bias voltage will be required in order for current to flow. For a gate bias which corresponds to N electrons on the dot (placing us exactly in the middle between two zero-bias

conductance peaks), a bias voltage of $e/2C_d$ is necessary to overcome the Coulomb repulsion. This non-linearity in the I-V characteristics is referred to as the Coulomb blockade of current. More details on this effect can be found in Averin *et al.* [21].

Coupling to the Environment and Destruction of Coulomb Blockade Effects

One of the assumptions which we made in developing the theory of the Coulomb blockade was that the conducting island was isolated from its environment. Making that assumption meant that the number of electrons on the island was well-defined. If the island were connected to the leads by a strongly conducting link, it would no longer make sense to discuss the number of electrons on the island – this would be analogous to discussing the number of electrons on a mustard-seed-sized piece of copper arbitrarily chosen from the interior of a large copper block. If the resistance of the link from our island were increased, however, at some point we would expect that it would become isolated enough that its electron population would be a well-defined number.

We can estimate what this resistance might be using the uncertainty principle for energy and time,

$$\Delta E \Delta t \geq \hbar,$$

where $\hbar$ is Planck's constant divided by 2π. This says that a particle can exist at an energy ΔE outside of its classically-allowed energy for a time Δt subject to the above constraint. The relevant energy scale in our problem is the energy required to add an electron to the island, $e^2/2C_d$. If we have current flowing, $I = ef$, where f is the frequency of an electron entering or exiting the island. By taking $\Delta t = 1/f$, we can substitute for our uncertainty relation and get

$$\frac{e^2}{2C_d}\frac{e}{I} \geq \hbar.$$

We also know that $I = V/R$, so we can substitute that into the above expression to

yield

$$\frac{e^3 R}{2C_d V} \geq \hbar.$$

Capacitance times voltage is equal to charge, and the relevant charge on the island for our purposes is one electron, e. Substituting this and solving for R, we arrive at

$$R \geq \frac{2\hbar}{e^2} \quad \text{or} \quad \sigma = \frac{1}{R} = \frac{e^2}{2\hbar}.$$

This is quite close to a more rigorously-derived value of $R \geq h/2e^2$ which is derived by Büttiker [23] as the resistance of a one-dimensional conductor with only one transmission channel which has a transmission $T = 1$ (the point at which the number of electrons on the dot is no longer well-defined). This value of R is referred to as the quantum of resistance or the resistance quantum, R_Q.

2.2.2 The Quantum Dot

So far we have discussed the Coulomb blockade in a metallic island. *Quantum dots* are conducting islands which are defined in a semiconductor. Although the basic physics of the Coulomb blockade remains the same in both cases, there are additional issues which arise because of the differences between the two systems.

Constricting the Electrons to a Plane: The Two-Dimensional Electron Gas

The starting point for our quantum dot is a GaAs/AlGaAs heterostructure grown in such a way that the conduction electrons are confined in a two-dimensional layer a few hundred angstroms below the surface. A typical heterostructure and a simulation of the conduction band energy at 4K is shown in Figure 2-4. A triangular well in the conduction band dips below the Fermi level at the GaAs/AlGaAs interface. This creates a confining potential in one dimension for electrons, while they are free to move in the other two dimensions.

It is often convenient to think of the 2DEG as a two-dimensional metal, as it remains conducting at low temperatures. However, some important differences exist

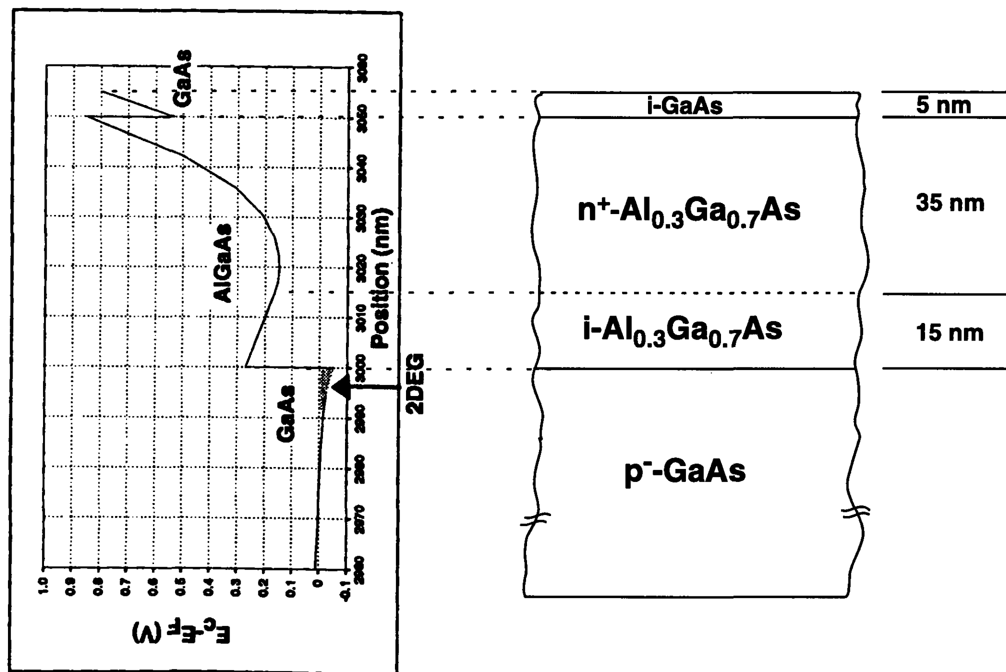


Figure 2-4: GaAs/AlGaAs heterostructure and simulated conduction band level at 4 K. At the lower GaAs/AlGaAs interface, the conduction band dips below the Fermi level, creating a population of electrons which is confined in one direction, but free to travel in the other two dimensions, termed a two-dimensional electron gas (2DEG). Donors are separated from the 2DEG by an undoped layer of AlGaAs to reduce scattering of the electrons. From reference [24].

between a 2DEG and an elemental metal.

Typical sheet densities n_s in the 2DEG are $\sim$3-5$\times 10^{11}\,\mathrm{cm}^{-2}$. This means that the Fermi wavelength, $\lambda_f = \sqrt{2\pi/n_s} \approx$ 35-40 nm. The Fermi wavelength can be considered in some sense to be the "size" of the electron. In a 2DEG, electrons are much "bigger" than in metals, where λ_f is typically on the order of a few angstroms, or a factor of about 100 smaller.

In addition, the heterostructure is grown with the dopants removed from the plane of the 2DEG – there is a layer of undoped AlGaAs between the doped AlGaAs and the 2DEG. This means that there is very little Coulomb scattering from ionized dopants. At low temperatures when phonon scattering is negligible, the mobility of the electrons in the 2DEG can be quite high. Typical mean free paths between scattering events in GaAs/AlGaAs heterostructures at cryogenic temperatures are several microns. This means that over the size scale of a quantum dot, transport can be ballistic. This contrasts rather strongly with a metal where mean free paths at room temperature are typically one to a few tens of nanometers.

Patterning Laterally: Schottky Gates, Quantum Point Contacts, and Quantum Dots

In order to create a quantum dot, the 2DEG must be patterned laterally. This is done by placing metal Schottky gates on the surface of the heterostructure which, when biased negatively, electrostatically deplete the 2DEG below. An example of a structure patterned in this manner is shown in Figure 2-5.

In this structure, the surface Schottky gates create a constriction in the 2DEG. If the gap between the lithographically-defined gates is comparable to a few Fermi wavelengths of the electrons in the 2DEG and its length is shorter than the mean free path between scattering events, the constriction will act as a one-dimensional waveguide for the electrons. As the gates are biased more negatively, fringing fields push out laterally from the gates and deplete more of the 2DEG, narrowing the constriction through which the electrons must pass.

As the constriction narrows, fewer electron modes in the waveguide are below the

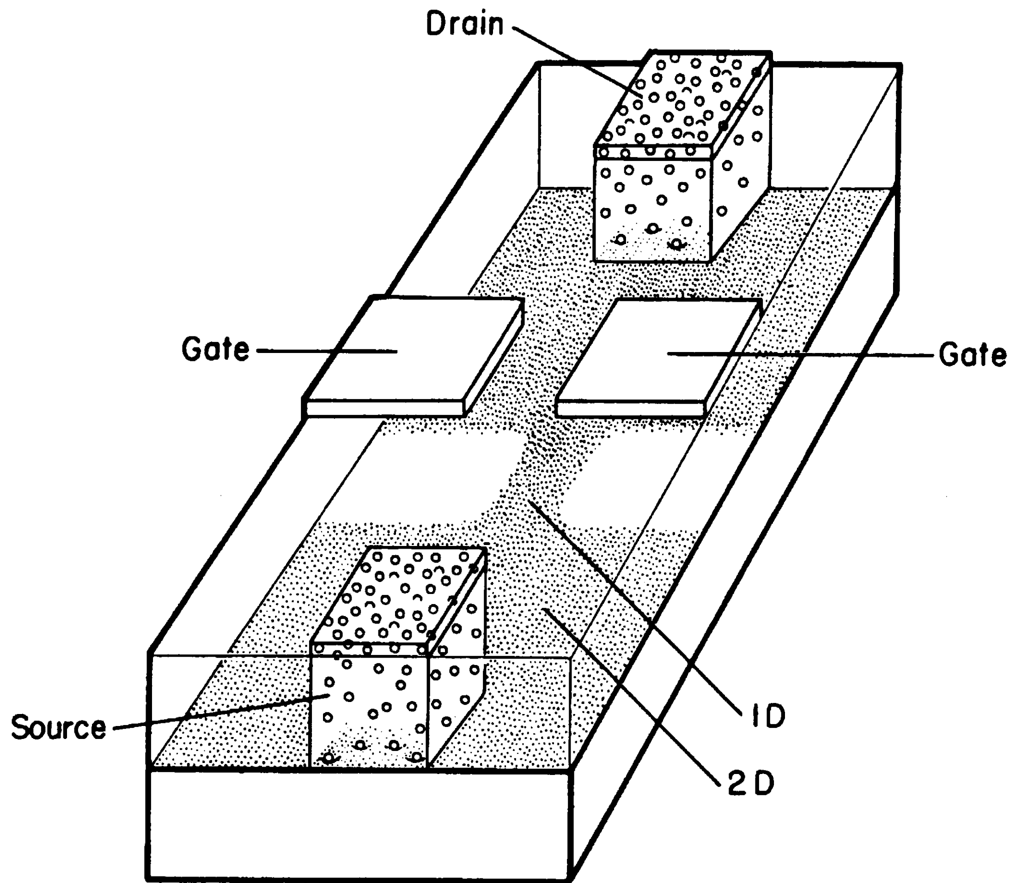


Figure 2-5: Schematic of laterally-patterned constriction (a quantum point contact) in a 2DEG. Schottky gates on the surface of the heterostructure electrostatically deplete the 2DEG (indicated with shading) below. If the width of the constriction is on the order of the Fermi wavelength of the electrons in the 2DEG and its length is shorter than the mean free path between scattering events, the constriction can act as a one-dimensional waveguide for the electrons, or a quantum point contact (QPC). Figure from reference [25].

Fermi level of the 2DEG and therefore fewer transmission channels are populated and available for conduction. A step-like structure is seen in the conductance versus gate voltage of the constriction, with each step having a conductance value of $1/R_Q = 2e^2/h \approx 77\,\mu\text{S} \approx 1/13\,\text{k}\Omega$. This one-dimensional waveguide structure is called a quantum point contact (QPC) [25,26].

A quantum dot is typically formed by isolating a small puddle of electrons from a 2DEG in a GaAs/AlGaAs heterostructure with a set of Schottky gates. Typically two QPCs allow tunneling current to flow into and out of the dot. By adjusting the QPCs to a resistance value greater than the resistance quantum, the quantum dot can be isolated from its environment well enough that the number of electrons on it is well-defined. In this way, Coulomb blockade effects can be observed in semiconductor quantum dots.

Since the electrons in a quantum dot are confined, they will occupy quantized energy levels. These single-particle energy states will have a spacing of approximately $\Delta E(E_f) = 1/(L^2 g(E_f))$, where L is the length of a side of the electron box, and $g(E_f)$ is the two-dimensional density of states at the Fermi level. For a two-dimensional box, the characteristic energy spacing is $\hbar^2\pi/m^* L^2$, where m^* is the two-dimensional effective mass of the electron, which is 0.067 times the free electron mass in (100) GaAs [27].

For a 100 nm quantum dot in GaAs, this corresponds to a single-particle level spacing of 0.36 meV, or a temperature equivalent of about 4 K. Single-particle energy states can be important in very small quantum dots at low temperature. For a metal, which has a much smaller Fermi wavelength and is three-dimensional, the single-particle energy level spacings are much smaller – a 100 nm metallic island would have an energy level spacing only of about $5\,\mu$V [20].

There are several possible regimes of operation of a quantum dot, depending on the relative energy scales of temperature, $k_B T$, the charging energy, $E_c = e^2/2C_d$, and the difference between the single-particle energy levels, ΔE. Quantum dots designed for this work were in the metallic regime, where $\Delta E < k_B T < E_c$.

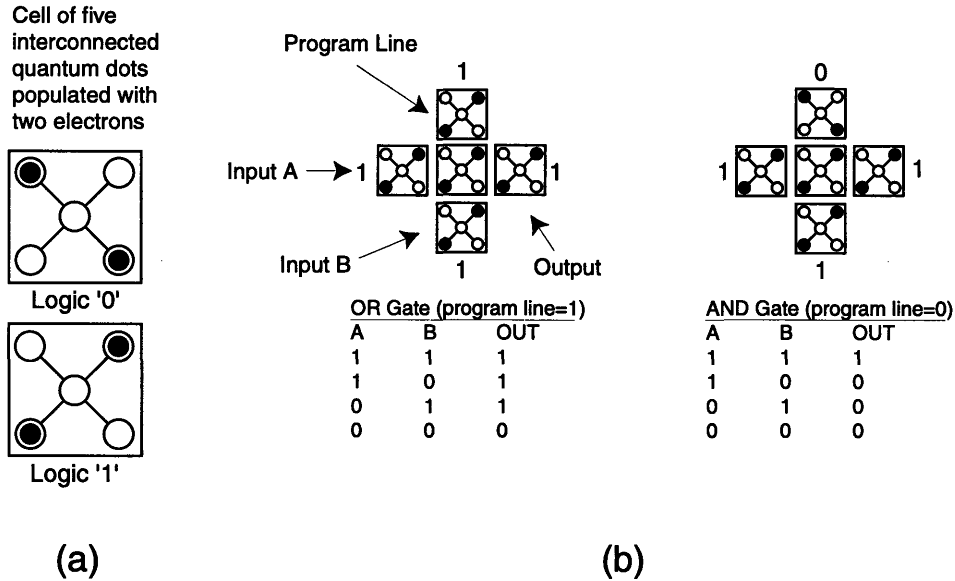


Figure 2-6: A scheme for cellular automata computation with quantum dots. (a) The automaton consists of five interconnected quantum dots populated with two electrons. The two electrons will repel each other and will want to sit in one of two polarizations, defined as logical one and zero. (b) A programmable AND/OR gate. Depending on the value of the program line, the gate acts as either an AND gate or an OR gate. From references [28–30].

2.3 A Quantum Dot Cellular Automata Scheme

In Section 2.1 we described cellular automata and a scheme for perhaps performing computation with few-body cellular automata. It was, however, an abstract concept and still a far cry from an actual computer.

A much more intuitively appealing scheme for cellular automata computation with quantum dots was developed at Notre Dame University [28]. The automaton, pictured in Figure 2-6(a), consists of five interconnected quantum dots with a population of two excess electrons.

Simulations confirm what is intuitively expected: it is energetically favorable for the two electrons to polarize in opposite corners of the structure. There are two possible polarizations, each of which can represent a binary state. What makes the structure computationally useful is the fact that in response to even small perturbations of external electronic charge (such as might be found in a polarized neighboring cell), the cell responds in a highly nonlinear and stable manner by polarizing strongly

in one of the two states [29,30].

Schemes for making binary signal lines [31], line crossings, and more importantly logic gates [32] have been proposed. Figure 2-6(b) shows a configuration which can, through the use of a program line, be used to selectively perform AND and OR operations. Both static and dynamic simulations of these devices have been run, and the results are quite promising [32,33].

Understanding how one could perform computation with this scheme is immediately obvious. However, in order to function as simulated, the quantum dots making up this structure must be extremely small – much smaller than intentionally-positioned quantum dots can be fabricated today. All of the simulations to date have assumed dots of 10 nm diameter with a 20 nm center-to-center pitch. The authors even mention that shrinking the size more would improve performance [33]! Given current nanofabrication techniques, this is unlikely to happen in the immediate future. However, alternative schemes do exist for creating small islands of charge (as will be discussed in Section 2.5) which may someday allow for implementation of this structure.

Another difficulty which would be encountered in actually fabricating a quantum cellular automata system is that the system requires that each block of dots have exactly two electrons in it, which is quite difficult to do even for a single block, let alone a chip full of them. Although some experiments have shown that dot electron population can be quite accurately controlled an a large array of dots [34], this by no means assures that evenly populating dots in cellular automata will be a trivial task.

Despite the difficulties encountered in fabrication, the progress made by the Notre Dame group toward actually making these devices has been quite impressive. Both a readout scheme [35] and a working automata cell [36] have recently been implemented in metallic SET structures.

2.4 Quantum Dot Interactions

In any cellular automata computation scheme, the manner in which the automata interact is critical. For the Notre Dame scheme, both the interactions of the five quantum dots making up the automaton and the interaction of the dots between automata will determine the ability to successfully perform computations. For the many-body cellular automata scheme, the range of the interaction between the cells is critical, and a configuration of quantum dots in each automata that could yield a deterministic few-body rule is still undetermined.

If we are to pursue cellular automata computation with quantum dots, it is necessary to understand how quantum dots interact with each other and with their environment (for example, leads). Possible schemes for clocking cellular automata interactions involve opening up the barriers between dots in an automaton between inter-automata interactions. This could work for both the Notre Dame scheme [37] or for a few-body cellular automata, where one might want to transport charge into and out of various dots. In such a clocking scheme it would be necessary to understand how the Coulomb blockade is modified as a quantum dot goes from being weakly coupled to its environment (therefore having a well-defined charge) to being strongly coupled to its environment (therefore not having a well-defined charge).

Several experiments have been done which investigate how Coulomb blockade effects are modified as the quantum dot's coupling to the outside world is increased, and there has been good theoretical modeling to understand the results of the experiments. The work can be broken up into two general classes: dot-to-dot interactions, and dot-to-lead interactions.

2.4.1 Interdot Interactions

A series of experiments done at Harvard University showed the effect of opening up a barrier between two matched quantum dots which were connected in series. As the barriers were opened, the Coulomb blockade peaks split into pairs until one dot was formed with a higher peak frequency in gate voltage, as would be expected for

a larger gate-to-dot capacitance [38–43]. Some splitting is to be expected purely due to increased interdot capacitance [6], but the observed peak splitting is greater than would be expected due to capacitance alone.

In order to fit the data, increased interdot tunneling must be taken into account. The tunneling allows for quantum charge fluctuations, which can reduce the energy of a polarized state (where one dot in the pair has an electron and the other does not) and allow transport in a regime where it would be forbidden if only the capacitance effects were accounted for [44, 45].

If the series quantum dots are not matched or the electrostatic potentials are not swept uniformly in each dot, then as the temperature decreases, the number of Coulomb blockade peaks also decreases. This is called the stochastic Coulomb blockade, and it can be understood from a purely capacitive model [46, 47]. It has been demonstrated in several experiments [48, 49].

Other experiments with coupled dots have investigated electrostatic effects in which the addition of a single electron to one dot can affect the charge state of (and therefore transport through) another dot which is situated in close proximity to the first. In these experiments, the dots are typically not connected in series, but one dot acts as an electrometer to monitor the charge state of the other dot [50–52]. A similar experiment used a metal single-electron transistor to measure the charge on a semiconductor quantum dot [53].

2.4.2 Dot-to-Lead Interactions

Other experiments have been done in which a quantum dot is opened up to a lead. Several experiments have been done in which one QPC in a quantum dot is opened beyond the point where one transmission channel is fully opened, where theory would predict the Coulomb blockade would vanish. Results, however, are conflicting, with some authors reporting the vanishing of the Coulomb blockade [54] while others report the persistence of effects at higher QPC conductances [55, 56].

In a novel quantum dot structure with three leads, with current input through one lead and with the output current split between the other two leads, an oscillation

was seen between the currents in the two output QPCs as their barrier strength was lowered [57]. This was attributed to coupling to the leads resulting in fluctuations of their local chemical potential [24, 58]. The model, in effect, replaced ohmic leads with sections which were capacitively isolated from the rest of the 2DEG. The model explained the data, but the physical explanation for the isolated leads was not well-motivated.

These experiments were all done on quantum dots formed in GaAs/AlGaAs heterostructures. The way in which barriers between dots or between a dot and a lead was opened up was to decrease the negative bias on the Schottky gates which formed a QPC. As discussed in Section 2.2.1, when a QPC is opened to the point where its resistance is equal to R_Q, the QPC forms a one-dimensional electron waveguide which has one transmission channel populated with electrons, and that channel has a transmission probability of one. This contrasts with other systems which exhibit Coulomb blockade.

2.5 Coulomb Blockade at High Temperatures

Quantum dot structures formed in GaAs/AlGaAs heterostructures have a limitation on the minimum capacitance of a quantum dot to the outside world. This is due to the minimum size dot which can be created by biasing Schottky gates which are removed from the 2DEG by several tens of nanometers, and by the capacitance of the dot to metallic gates which necessarily surround the dot. This minimum capacitance sets a maximum energy $e^2/2C_d$ which sets a maximum temperature at which Coulomb blockade effects are visible. Room temperature operation of a quantum dot structure in GaAs/AlGaAs is probably impossible. In fact, creating a quantum dot capable of 4K operation in GaAs/AlGaAs is difficult [59, 60].

This has led to much research into other systems in which Coulomb blockade effects can be observed. A short overview of various technologies for fabrication of single electron devices is given by Ahmed [61]. Here we will mention some of the technologies in use which yield structures which show Coulomb blockade effects at

higher temperatures.

There are several technologies for creating metallic structures which exhibit the Coulomb blockade. Fabricated metallic single-electron transistors have been around for many years [62], and have recently been shown to be operable at 30 K [63]. The Coulomb blockade has recently been observed at room temperature in metallic structures with oxide tunnel barriers formed by STM oxidation [64]. Few-atom gold clusters have been shown to exhibit the Coulomb blockade by depositing (usually randomly) an array of clusters between electrodes on a substrate [65–68], where effects have been clearly seen at 77 K and persist in some form to room temperature. Room temperature coulomb blockade is also seen in STM measurements on single clusters [69].

There has been much recent work in making devices in silicon which exhibit Coulomb blockade effects. Silicon-on-insulator (SOI) wafers allow for structures with low capacitances, which helps to boost the operating temperatures of these devices. Typically oxidation of patterned silicon is used to decrease the size of an island below its lithographically-defined dimensions [70–75]. These devices can show clear signatures of the Coulomb blockade above 100 K and some show effects at room temperature.

Other semiconductor and metallic structures have shown Coulomb blockade effects at high temperature. There are many means of fabrication, including growth of nanocrystals of silicon [76] and InAs [77], fluoridation of aluminum by an electron beam [78], and more esoteric schemes of depositing material into porous Al_2O_3 [79] or porous opal [80].

If a means were found to produce intentionally-positioned quantum dots which operate at room temperature or even 77 K and also to be able to arbitrarily connect them to form automata cells, then questions about interactions between dots would take on a new urgency. If we are to be prepared to understand how any of these candidate technologies might be made useful for cellular automata computation, we must understand how their properties would change with changing barrier strength.

As a barrier is lowered in these systems (one can imagine in some cases a barrier

which is controllable during device operation and in others a barrier which once made has a constant value but for which there is some choice of barrier strength during fabrication) it is unclear whether the conduction will be one-dimensional or not. As expected from our general uncertainty principal argument, regardless of the nature of conduction, as the conductance through a barrier reaches the quantum of conductance, $2e^2/h$, Coulomb blockade effects should disappear, or at least radically change.

The manner in which the opening of the conduction occurs, however, may impact observable effects in how the Coulomb blockade vanishes. Golden and Halperin state that " ... the most important dimensionless parameters [for understanding the Harvard double-dot experiment] are the number N_{ch} of conducting channels between the two dots and the dimensionless interdot barrier conductance g of each channel ... " [44].

For the case of a QPC, as the barrier is lowered, a single conduction channel appears which opens to a transmission probability for electrons of one as the quantum of conductance is reached. In a tunnel barrier (such as is seen in metal-oxide systems), as a barrier progressively opens, many conduction channels open to a small transmission probability. At the quantum conductance, we have for the QPC case two states on either side of the barrier which are in full communication while for the tunnel barrier case we have many states which are in poor communication.

For many of the technologies above it is unclear which type of conduction will occur when barriers are opened. Some cases may be in-between. For self-assembled structures electrically connected by molecular wires, there may be a small number of channels available for conduction [81]. Not all forms of Coulomb blockade will open to one-dimensional waveguides. The QPC may be an artifact of the GaAs/AlGaAs heterostructure.

2.6 A Novel Coupled Quantum Dot Structure

2.6.1 A Non-Point-Contact Tunnel Barrier

An alternate means of isolating a quantum dot is shown in Figure 2-7. In this figure we see a coupled quantum dot structure which is isolated from its leads by QPCs. The barrier *between* the dots, however, is a narrow line which, when biased negatively, forms a tunnel barrier which isolates the dot. As the bias is decreased to allow more contact between the two dots in this structure, the electrons should not see a one-dimensional waveguide as the resistance transitions to below R_Q. In this case at the quantum of resistance, there should be some number of channels $\mathcal{N}$ (approximately equal to the length of the line divided by the Fermi wavelength of the electrons). Each of these should have a transmission probability T significantly less than one ($T \sim 1/\mathcal{N}$). This structure would act like a metallic tunnel barrier, but with the ability to control the tunneling strength with an external voltage.

The coupled-quantum-dot structure schematically pictured in Figure 2-7 would allow investigation of how transport properties of a quantum dot or quantum dots change as a tunnel barrier is opened. By comparison with a similar structure formed with all QPCs, differences between one-dimensional waveguided and non-waveguided conductance could be probed.

2.6.2 Fine-Line Barrier Width Considerations

The fine-line tunnel barrier between the two dots would be required to both pinch off the dots (such that the conductance through the barrier is well below the conductance quantum) and to transition to a state where the conductance through the barrier is greater than or equal to the conductance quantum. It is desirable that this transition be relatively broad, so that noise which will inevitably be present in the gate bias would not significantly change the conductance of the barrier. The width of the barrier would therefore be a critical to this device's functionality.

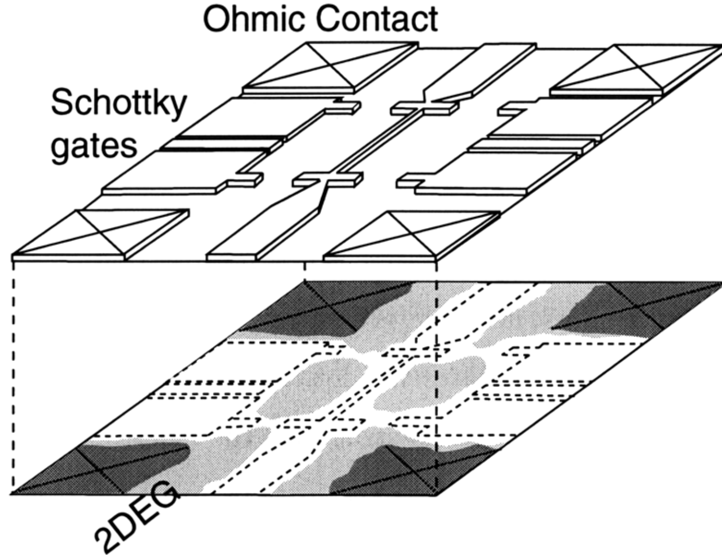


Figure 2-7: Schematic of coupled-quantum-dot device. Schottky gates are shown on the surface of the heterostructure in white. Their outline is shown in dashed lines at the 2DEG surface. Ohmic contacts are shown at the surface in white with an 'X' and at the 2DEG in dark gray with an 'X'. Undepleted 2DEG is shown in light gray. The two quantum dots are seen as two isolated puddles which can interact with each other through the fine-line tunnel barrier between them or with the leads through the four QPCs.

Previous Experiments

Eugster [25] investigated electron waveguide devices which were coupled via a leaky fine-line tunnel barrier to a 2DEG. In this work he fabricated devices with direct-write electron-beam lithography with lines as narrow as 30 nm. Typical gate widths for devices were 30-40 nm. With this linewidth, the transition from the fully open state to fully pinched off took a gate voltage sweep of approximately 0.09 volts. He reported that minimum linewidths which were fabricated with x-ray lithography at that time were approximately 50 nm, and that no tunneling features were seen with x-ray fabricated devices.

Kumar [24] also fabricated quantum dot structures with fine-line tunnel barriers. His devices were also written with direct-write electron-beam lithography, and the tunnel barrier widths were approximately 40 nm. The gate voltage sweep necessary to transition the barrier from the fully open state to fully closed was approximately

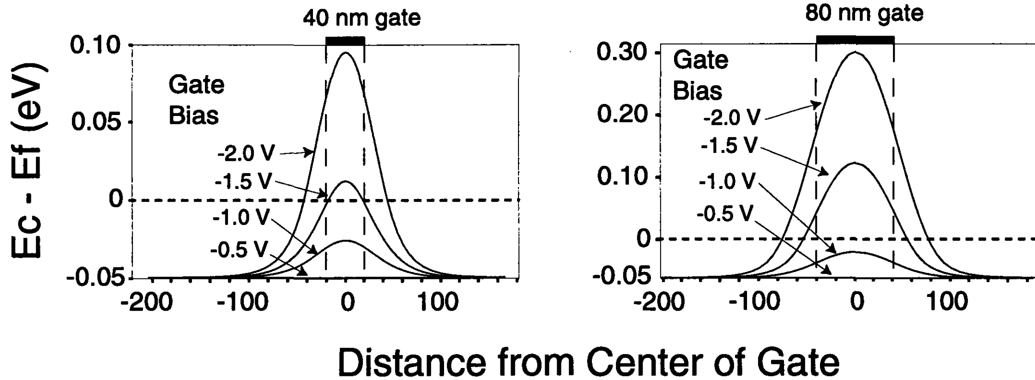


Figure 2-8: Simulations of the conduction band energy (in eV relative to the Fermi level) in two fine-line gate structures. On the left is a 40 nm gate and on the right is an 80 nm gate. At around $1-1.5$ volts negative gate bias, the conduction band jumps above the Fermi level directly under the gate, creating a tunnel barrier for electrons of approximately the gate width. Upon further increase of the bias, the barrier becomes wider. *Note the difference in energy scales in the two plots.*

0.04 volts.

The tunnel barriers in the above experiments were similar in length to the one which was designed for the double quantum dot structure described above. The length of the tunnel barrier is set by the size of the quantum dot which is set by the temperature at which Coulomb blockade effects are to be observed. From these experiments it would seem that a 50 nm line is too wide to observe tunneling (or to effectively control the barrier strength in the region below the quantum conductance), and that a 30-40 nm line gives adequate control over the barrier strength in that low-conductance regime, with the lower end of that width range giving better control.

Electrostatic Simulations

Further understanding can be gained by simulating the electrostatics of various tunnel gate widths. Figure 2-8 shows simulations performed using a three-dimensional Poisson solver written by Arvind Kumar. The figure plots the conduction band relative to the Fermi level as a function of position on a slice through the fine-line gate for gate widths of 40 nm and 80 nm. The geometries are simulated at gate biases ranging from -0.5 volts to -2 volts. It can be seen that the conduction band quickly jumps above the Fermi level at a gate voltage of -1 to -1.5 volts with a width close to the

lithographically-defined gate width. Upon increasing the negative bias, the width of the barrier increases at a slower rate due to fringing fields.

If the initial width of the barrier is such that the conductance of the structure is already well below the quantum conductance, the transition from fully closed to fully open would be difficult to control. If, on the other hand, the initial barrier width gives a barrier with a conductance greater than the conductance quantum, the transition should be reasonably controllable.

WKB Estimation of Tunneling Probability

A better feel for what these simulations mean can be gained by estimating the tunneling probability through the barriers. If we are in a regime where the tunneling probability $T \ll 1$, then we can use the WKB approximation to estimate the tunneling probability [82]:

$$T = \exp -\frac{2}{\hbar}\int_{-x_0}^{x_0} \sqrt{2mE}dx,$$

where $\pm x_0$ are the classical turning points (where E=0).

We can approximate the above simulations by fitting a parabola to the portion of the conduction band which lies above the Fermi energy. To parameterize this parabola we need two values – the height of the barrier, H, and the width of the barrier at the Fermi energy, $2x_0$

The approximated barrier will then be

$$E_c - E_f = H\left(1 - \frac{x^2}{x_0^2}\right).$$

Carrying out the integral we find

$$T = \exp -\frac{\pi}{2}\frac{\sqrt{2m}}{\hbar}\sqrt{H}x_0,$$

which reduces to

$$T = \exp -2.09\sqrt{H}x_0,$$

where H is in eV and x_0 is in nm.

	40 nm gate			80 nm gate		
Gate Voltage	x_0(nm)	H(eV)	T	x_0(nm)	H(eV)	T
-1.5	20	0.01	1.5×10^{-2}	55	0.125	2.2×10^{-18}
-2	45	0.95	1.5×10^{-40}	80	0.3	1.7×10^{-40}

Table 2.1: Table of parameters to fit simulated fine-line tunnel barriers and tunneling probabilities for the barriers calculated using the WKB approximation

Table 2.1 shows the parabolic fit parameters and the calculated barrier transmission for the -1.5 V and -2.0 V tunnel barriers in Figure 2-8. Although it is difficult to make predictions of how the transmission would relate to an actual conductance in the structure (since that would depend on the details of the wavefunctions of the states which were tunneling), it can be seen from the table that the 80 nm gate quickly yields vanishingly small tunneling probabilities as the barrier pokes above the Fermi level, while for the 40 nm barrier, there is a region where the tunneling probabilities are non-trivial.

2.7 Motivation for Improved Sub-50 nm X-ray Lithographic Process

The direct-write electron-beam lithography used to fabricate the devices of Eugster and Kumar was done at the National Nanofabrication Facility at Cornell University. It was desired to fabricate the coupled quantum dot structure at MIT.

Attempts were initially made to write the gate patterns with both electron-beam direct-write lithography and x-ray lithography. However at MIT, electron-beam direct-write capabilities did not allow for robust fabrication of features below 50 nm.

X-ray lithography, however, has the potential to repeatably create structures much smaller than 50 nm. According to Eugster [25], in 1993, the smallest features that could be made by x-ray lithography at MIT on an actual device were 50 nm. Prior to 1993, Early [83] had shown that x-ray lithography was capable of 30 nm features. However, these were test structures made in the x-ray resist poly(methylmethacrylate) (PMMA) and not actual devices. Obviously work needed to be done to optimize the

full x-ray device fabrication process.

The remainder of this document will discuss the development of a robust process which allows for fabrication of aligned devices on substrates with feature sizes in the sub-50 nm regime. This has involved gaining an improved understanding of many aspects of the x-ray lithography and mask fabrication processes and using that understanding to improve various parts of the process. As most process research does, it has involved drawing upon many disciplines, from mechanical and vacuum design to electrochemistry to the physics of x-ray and electron interactions with matter to fluid mechanics and digital signal processing. The end result is a process which is inherently "manufacturable" (at least in the university setting) and which allows for device fabrication in a regime which was previously inaccessible.

Chapter 3 will describe x-ray nanolithography and detail the process which is required in order to go from a conceptual device design to an actual pattern on a substrate. Chapter 4 will discuss work which I did in order to improve the technology of soft-contact replication of x-ray masks, both for mask-to-mask replication and for aligned chip or wafer exposures. Chapter 5 will present the main hurdle which had to be overcome in developing the sub-50 nm x-ray process: substrate photo- and Auger electrons. The results of the process improvements will be presented in Chapters 6 and 7. Chapter 6 will present sub-50 nm exposure results for device and test structures, and Chapter 7 will present electronic transport measurements on the fine-line tunnel barriers.

Chapter 3

X-ray Nanolithography

In micro- and nanolithography, we want to create a pattern in a resist on a substrate. This patterned resist will stand up to (or resist) subsequent steps like etching of the substrate material, or serve as a sacrificial layer for defining patterns in a deposited film by liftoff. The device presented in Chapter 2 requires a sub-50 nm line which is to be deposited by liftoff.

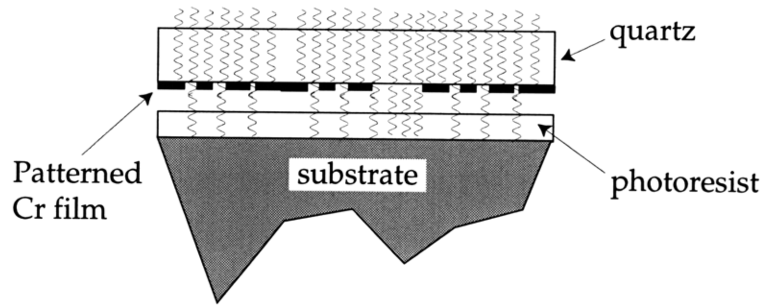
3.1 Alternative Lithographies

Two common means for producing patterns in a resist are shown in Figure 3-1. Figure 3-1(a) shows a schematic of optical proximity lithography. Ultraviolet (UV) light is passed through a quartz mask which is patterned with chrome. In the areas on the mask where chrome is present, the UV light is blocked, while in the areas without chrome, the light can pass through. The light impinges on photoresist which has been coated onto the surface of the wafer. Where the photoresist is exposed to the light, it is chemically altered such that it will (or for the case of negative resist, will not) dissolve away in a developer.

The minimum spatial period p_{min} which can be printed with optical projection lithography is limited by the wavelength of the UV photons used to expose the resist

$$p_{min} \approx \frac{\lambda}{\mathrm{NA}},$$

(a) UV (Proximity) Lithography



(b) Electron-beam Lithography

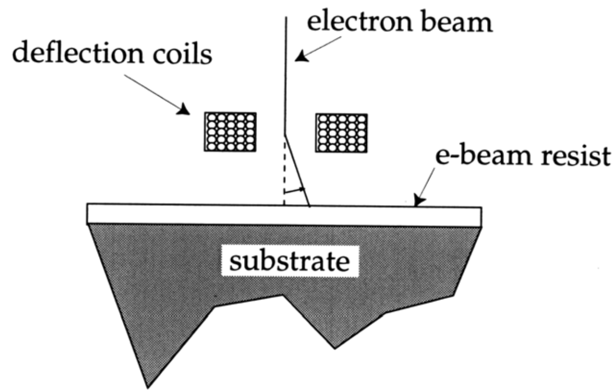


Figure 3-1: (a) Schematic of optical lithography. UV light is passed though a mask which blocks the light in some areas and allows it to pass in others. In clear areas, UV light is incident upon photoresist on a substrate. (b) Schematic of electron-beam lithography. An electron beam is steered across the sample by magnetic coils and blanked when no exposure is desired.

where NA is the numerical aperture ($n \sin \alpha$, where n is the index of refraction of the medium and α is the half-angle of the final lens in the exposure system), and λ is the wavelength of the light. By using shorter wavelengths of light and optimizing optical design, optical lithography has been able to print smaller and smaller features over time. 193 nm lithography for 180 nm feature sizes seems to be on track for commercial use by the year 2000 [4, 84]. However, optical lithography is not a candidate for sub-50 nm features at this time, and it is highly unlikely that it ever will be.

Figure 3-1(b) shows a schematic view of electron-beam (e-beam) lithography. In

this case, rather than an entire die or wafer being exposed at one time, a high-energy electron beam is steered across the wafer, usually in a raster pattern, and the beam is blanked when no exposure is desired. An electron-sensitive resist is chemically altered by the electrons such that it is preferentially dissolved by a developer. As discussed in Chapter 2, electron beam lithography can be used to define sub-50 nm feature sizes, but the e-beam systems at MIT were incapable of doing so reliably.

3.2 Overview of the Lithographic Process

In order to define some concepts and terms, we will use the examples of optical and e-beam lithography to discuss the patterning of resist. Figure 3-2(a) shows a schematic cross-section of a desired resist profile.

In order to define this pattern, we use a *lithographic input* of photons or electrons which to first order corresponds to our desired outcome. In the case of optical lithography, we might have a mask whose pattern matches that of our desired resist pattern. With e-beam lithography, we would write instructions to the computer which tell it to turn the beam on in places where we want the resist removed and to leave it off in places where we want the resist to remain. A plot of such an input is shown in Figure 3-2(b). The *lithographic input* therefore consists of the source of energetic particles and some means of providing *contrast*, or differential dosing of the resist in exposed and unexposed regions.

The *resist dose*, however, may not look like the input. In optical lithography, there will be some diffraction of the image as the light propagates from the mask to the substrate. In e-beam lithography, although the electrons might initially be aimed at the proper area on the substrate, forward scattering in the resist and backscattering from the substrate will tend to distribute the electrons in the resist. These effects will lead to a reduced volume dose contrast in the resist – the "unexposed" areas may be dosed somewhat, and the dose in the exposed regions may be reduced. In addition, the lateral distribution of dose may change due to the non-abrupt transition between exposed and dark areas. These effects can change feature sizes and for very small

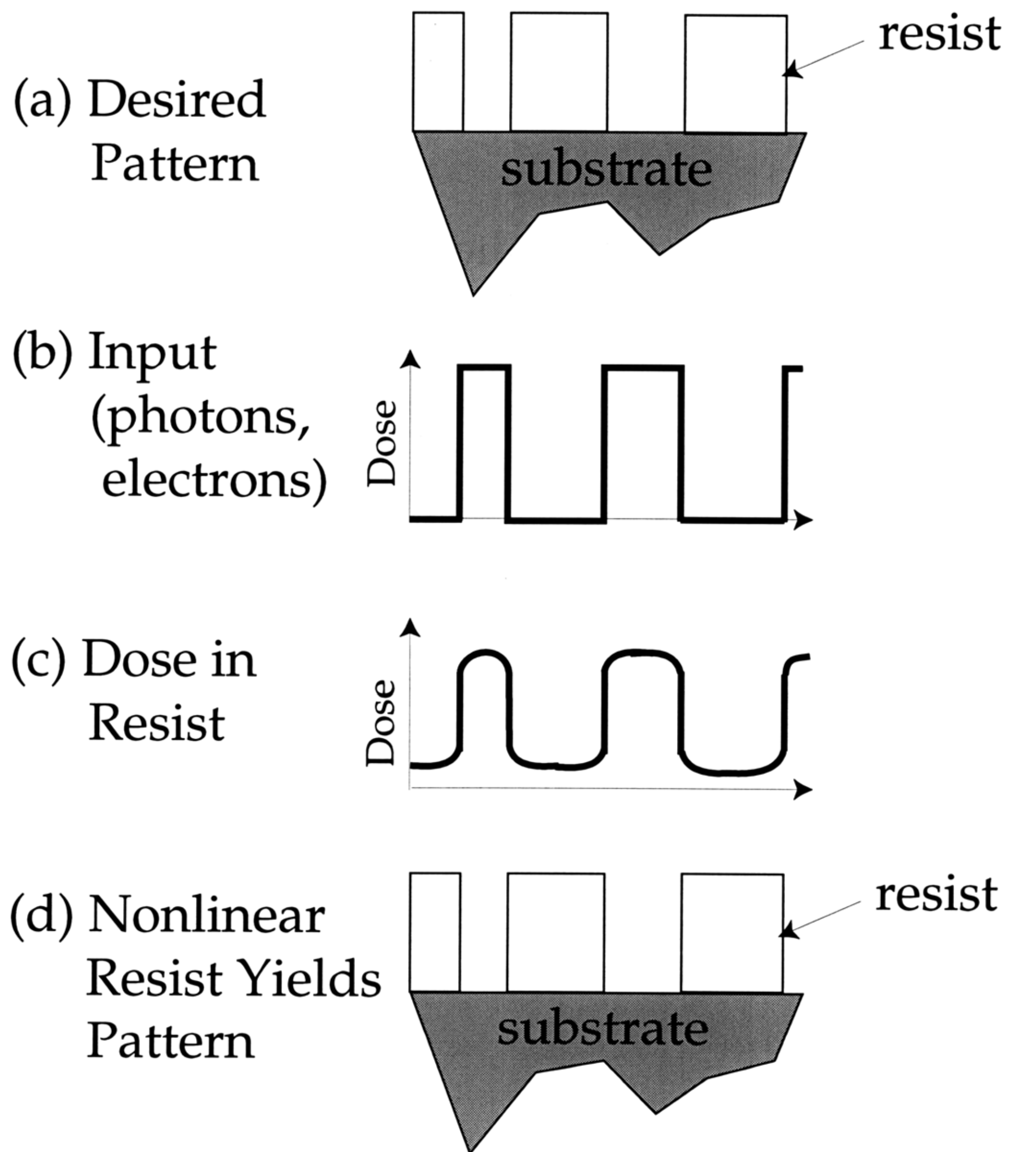


Figure 3-2: Schematic overview of lithographic process. (a) Desired pattern of resist on substrate. (b) *Lithographic input* of photons or electrons. (c) *Resist dose*, which is altered from the input due to diffraction or beam scattering effects. (d) Highly non-linear resist can recover from non-ideal dose, yielding desired pattern.

features can even prevent replication altogether. They are shown schematically in Figure 3-2(c).

Fortunately, the development rates of resists used in micro- and nanolithography are highly non-linear in dose. This usually allows one to recover a resist profile which at least approximates our desired pattern because the developer effectively "clips" the resist at a particular threshold dose, as indicated in Figure 3-2(d). Often, modifications to the input may be necessary in order to obtain the desired pattern, but an understanding of the processes involved in the exposure and development can allow one to obtain the desired geometries on the wafer.

3.3 Cu_L X-Ray Exposure Technology

At first blush, x-ray lithography [85–89] looks a lot like optical proximity lithography, with the advantage that a much shorter wavelength yields improved resolution. A schematic of an x-ray exposure is shown in Figure 3-3. There is a source of photons, a mask which blocks the photons in certain regions and provides contrast, and a resist which upon x-irradiation is chemically altered so that it is preferentially dissolved in a developer solution. However, because of the nature of x-rays and the ways in which they interact with materials, there are significant differences between optical and x-ray lithography.

We will use our conceptual framework of a lithographic input and a resist dose to investigate these differences and gain a better understanding of x-ray lithography.

3.3.1 Lithographic Input in X-ray Lithography

Electron Bombardment X-Ray Source

The source for x-ray photons used for lithography in the NanoStructures Laboratory at MIT is a copper target which is bombarded with high-energy electrons. The target is water-cooled in a high-vacuum chamber. Current driven through a tungsten filament creates a thermionic source of electrons. The filament is held at (for our

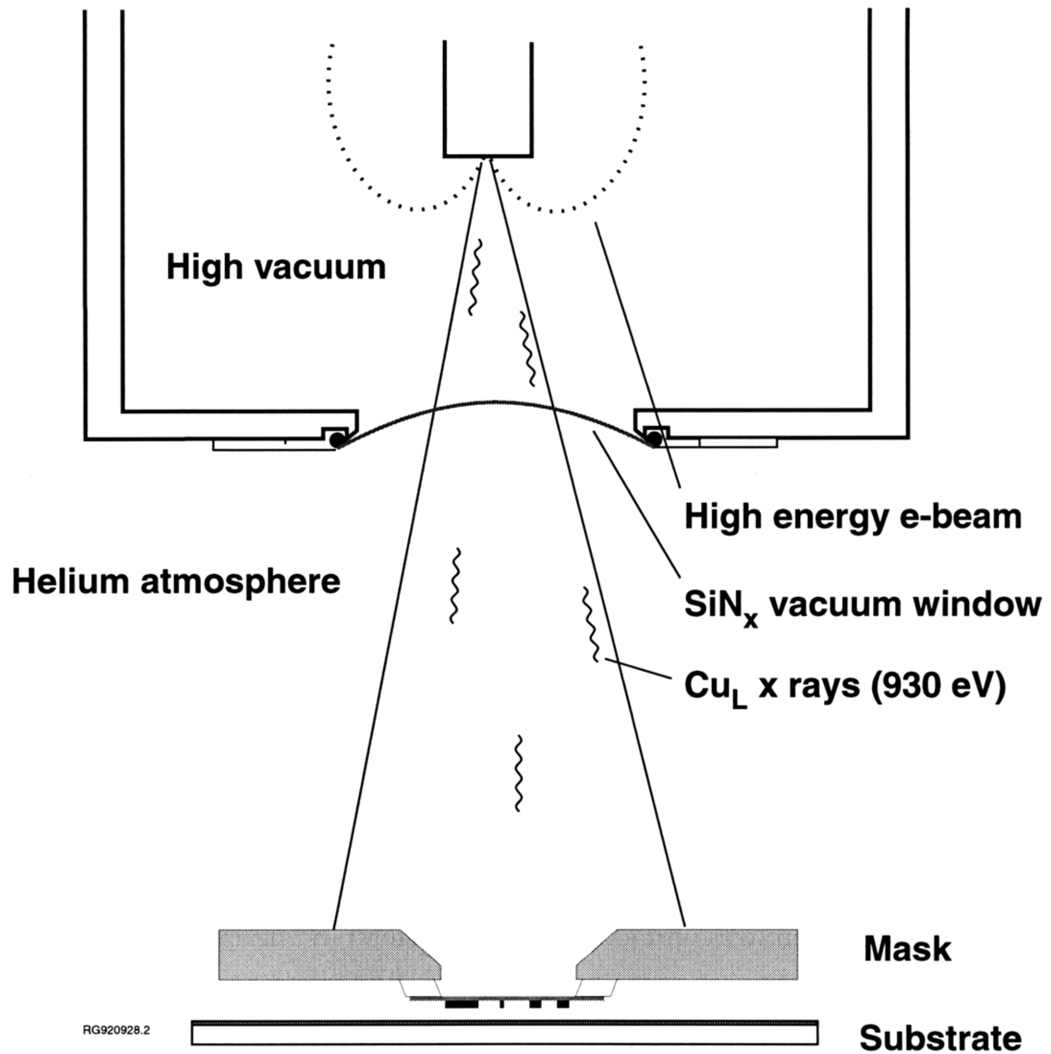


Figure 3-3: Schematic of x-ray exposure. Figure from reference [90].

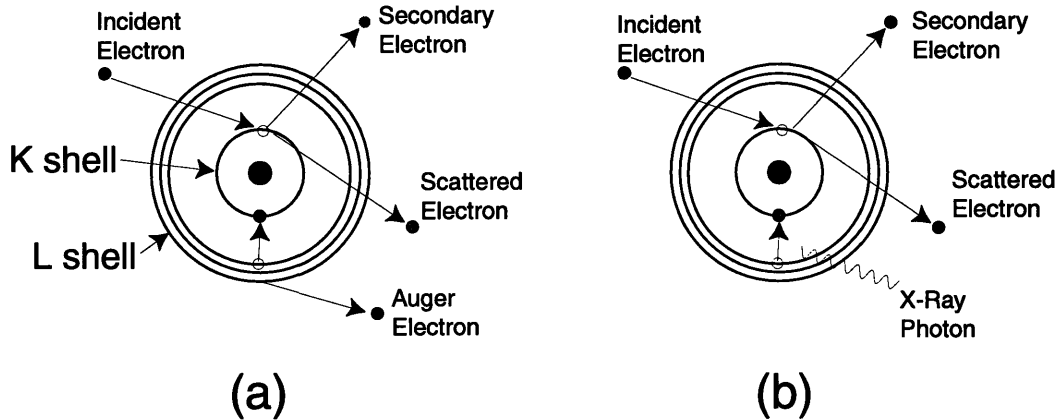


Figure 3-4: Possible consequences of an inner-shell electron (here a K-shell electron) being ejected from an atom by a high-energy electron beam. (a) Auger electron generation. A higher-shell electron (here from an L-shell) drops in energy to replace the emitted electron. The energy is transferred to another upper shell electron (here also from the L-shell) which is ionized. (b) X-ray generation. Energy gained in the L→K transition is used to create an x-ray photon. The probability that (b) will occur rather than (a) is the fluorescent yield, ω.

source) 8 kV relative to the target, and electron optics focus the electrons emitted from the filament onto a spot on the target.

Once the electrons enter the copper target, they undergo collisions within it. These collision can generate many secondary effects, such as heat generation through phonons, plasmon excitation, and electron emission from the target. In addition, both backscattered electrons and secondary electrons (which are generated by scattering events which excite valence or conduction-band electrons to higher energy states) are generated.

X-ray Characteristic Lines: When a core electrons is excited, it leaves behind an empty state which will be filled by a higher-energy electron. When this happens, the energy difference between the initial (higher) energy state and the final (formerly-empty) state must be transferred somewhere. This can happen by ionizing another higher-energy core electron, in an Auger process, or by emitting an x-ray photon. This is shown schematically in Figure 3-4. The energies of characteristic x-ray lines for most materials are well-known and can be found in several references, including Bertin [91] and the CRC Handbook of Chemistry and Physics [92].

The probability that x-ray emission will occur rather than an Auger electron being ejected is called the fluorescent yield, ω, of the material. For the L line of copper, the fluorescent yield is quite small: $\omega_L = 0.006$ [91]. This means that the copper target is not very efficient at generating x-rays. One consequence of this is that exposure times for lithography with this source tend to be quite long – usually around 8 hours or so. Sources with higher power are available – an x-ray synchrotron exposure can expose a die in about one second – but a synchrotron can cost over ten million dollars.

To excite a given x-ray line, an incident electron must have enough energy to ionize the inner-shell electron which must be vacant in order for the photoemission to occur. As the beam energy is increased above that threshold energy, the x-ray intensity increases as more ionization events occur. At some point, however, the more energetic electrons penetrate deeper into the target and more of the x-rays are absorbed before they can exit the target. At this point the characteristic line x-ray intensity begins to decrease.

This dependence has been measured for several x-ray lines in the soft x-ray regime (0.1–2 keV) for several elements by Henke and Tester [93]. For the Cu_L line the maximum yield occurs near 10 keV incident electron energy. For an 8 keV incident electron energy such as is used at MIT, the Cu_L x-ray intensity is 8.5×10^{-6} watts per steradian per watt input power [93].

Bremsstrahlung: In addition to characteristic x-ray lines, there is a continuous spectrum of x-rays which emanates from the bombardment source. This continuum is called *bremsstrahlung*, or braking radiation. It is caused by inelastic Rutherford scattering processes where the incident electrons radiate energy due to acceleration by the Coulomb force of the target ion cores. The theory of this effect was originally worked out by Kramers in 1923 [94]. The bremsstrahlung energy spectrum per incident electron for a thick target is given by [95]:

$$I_e(\nu)d\nu = h^2 kZ(\nu_0 - \nu)d\nu$$
$$I_e(\lambda)d\lambda = Zh^2kc^2 \left(\frac{1}{\lambda} - \frac{1}{\lambda_0}\right) \frac{1}{\lambda^2} d\lambda$$

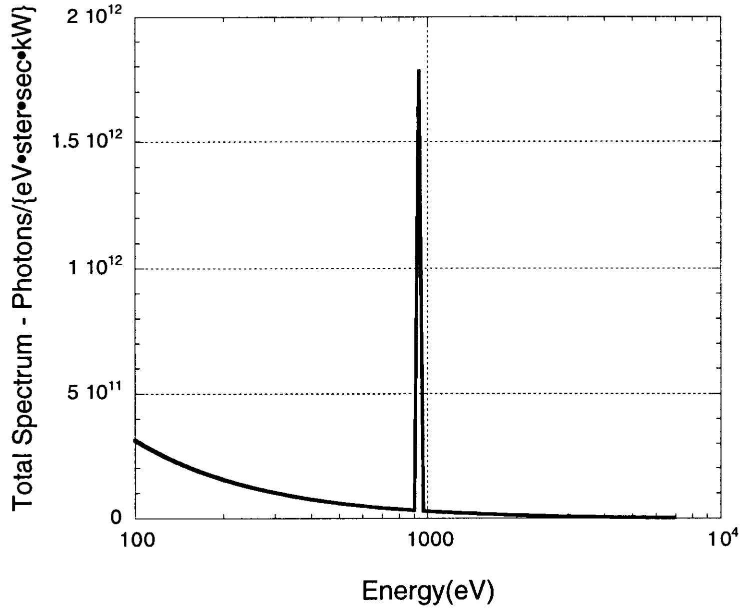


Figure 3-5: Copper target x-ray photon spectrum. Characteristic peak at 930 eV is taken from experimental data [93]. Bremsstrahlung continuum is calculated theoretically [95].

or in terms of energy,

$$I_e(E)dE = kZ(E_0 - E)dE.$$

In the above equations, ν is frequency, λ is the wavelength, and E is the energy of the x-ray. The subscript 0 refers to the minimum photon wavelength or the maximum photon frequency or energy, which is set by the incident electron energy. Z is the atomic number of the target, h is Planck's constant, and k is a measured constant which is weakly material-dependent [95]. For copper, its value is approximately $2.2 \times 10^{-6}\,\text{keV}^{-1}$ [95].

The total x-ray spectrum is the sum of the characteristic lines and the bremsstrahlung. The x-ray *photon* spectrum (which is found by dividing the above energy spectrum by E) emitted from a copper target at 8 keV incident electron energy is shown in Figure 3-5.[1] The ratio of Cu_L to bremsstrahlung power in this spectrum is

[1] The Cu_L line is in reality almost a delta function in energy. For the purpose of the plot it was

1:1.88, *i.e.* the spectrum coming off the copper target is about 65% bremsstrahlung.

X-ray Absorption by Materials

For x-ray nanolithography, absorption of x-rays by the mask gives contrast, and absorption of x-rays by the resist gives preferential dissolution in a developer. How x-rays interact with and are attenuated by these and other materials in the system is critically important.

In the soft x-ray regime of x-ray lithography, scattering is negligible, but photoelectric absorption, due to the interaction of the x-ray photons with core electrons in the solid, is significant. In this process, an incident x-ray photon is absorbed by an atom, which emits an inner-shell electron as a photoelectron. The vacant core state is then filled by a higher-shell electron (just as in the electron-irradiation case), and another x-ray photon or (usually) an Auger electron is emitted.

For a monochromatic beam of x-rays incident on a thin absorber layer, the differential attenuation can be written as

$$dI = -\mu I ds,$$

where I is the intensity of the beam, μ is the linear x-ray absorption coefficient, and ds is the path length through the thin layer. This expression can be integrated to yield the transmitted intensity I through a thick target:

$$I = I_0 \exp(-\mu s),$$

where I_0 is the incident intensity of the x-ray beam and s is the target thickness. This is called Beer's Law or Lambert's Law. Because x-ray absorption is due to ionization of deep core-level electrons, it is insensitive to chemical bonding or the arrangement of the atoms in a material. A more useful number is therefore the mass absorption

represented as a triangle whose integrated photon count equals the delta function weight. The height of the Cu_L peak in the plot is therefore somewhat arbitrary and was chosen so as not to dominate the bremsstrahlung spectrum.

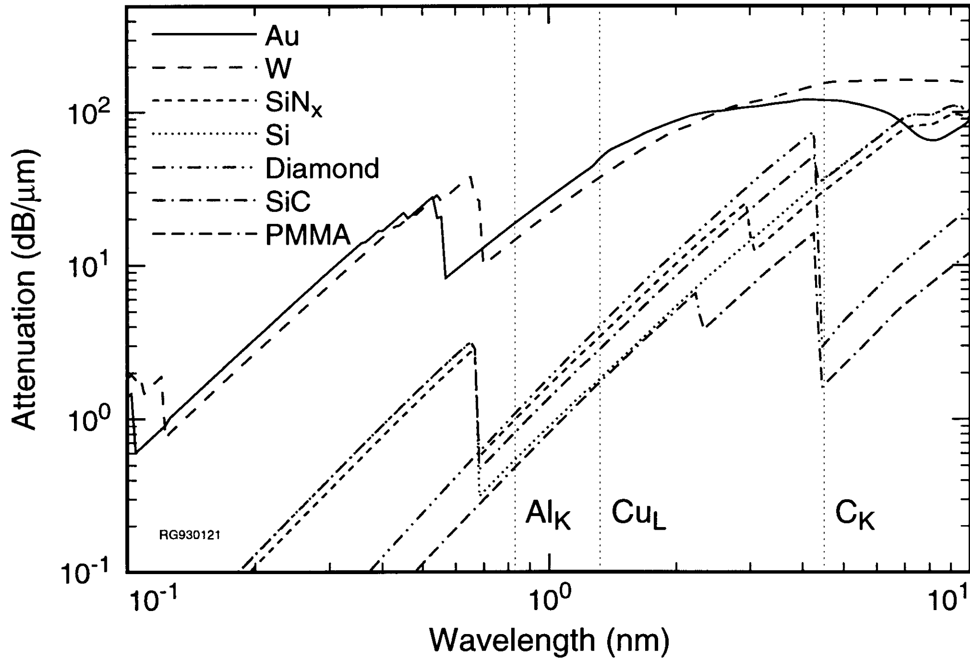


Figure 3-6: Attenuation of several materials at soft x-ray wavelengths. Figure from reference [90].

coefficient, μ/ρ, where ρ is the density of the material.[2]

In order to be absorbed by an atom, the x-ray photon must have enough energy to ionize a given core electron. As the photon energy is increased, deeper core levels can be ionized. Absorption of x-rays is therefore highly material- and energy-dependent. Mass absorption coefficients μ/ρ are tabulated in several sources, including Bertin [91] Appendix 7, Henke [96], and in a booklet published by Lawrence Berkeley Laboratory [97].

The attenuation (in dB/μm) of several materials used in x-ray lithography is plotted in Figure 3-6. As the wavelength is decreased and energy increased, absorption edges or steps are seen in the absorption corresponding to energies where a

---

[2]The value of using the mass absorption coefficient for x-rays is discussed in Bertin [91], who quotes Sproull, "A beam of x-rays passing from the ceiling to the floor of a chamber filled with hydrogen and oxygen may be [say] 10% absorbed, or 90% of it will reach the floor. If a spark explodes the hydrogen and oxygen, filling the chamber with steam, 90% of the x-rays will still reach the floor. Then if the chamber is chilled so that the steam condenses to a thin layer of water or ice on the floor, 90% of the x-rays will still reach the floor. This is not true for light or ultraviolet or infrared radiation, and it explains why the mass absorption coefficient of x-rays is commonly used, whereas the linear absorption coefficient is ordinarily used in optics."

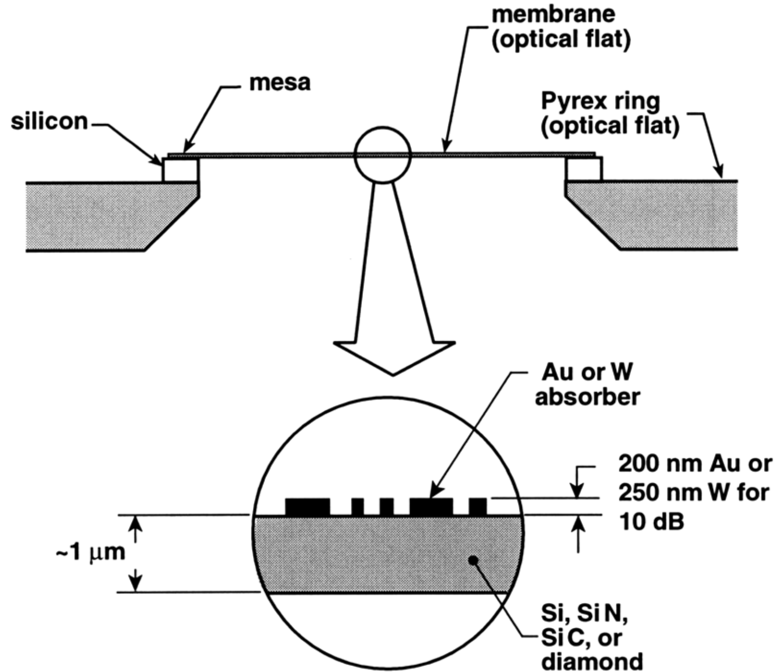


Figure 3-7: Schematic of x-ray mask. The absorber is a 200 nm layer of gold which is patterned on a 1 μm-thick, 31 mm-diameter SiN_x membrane. The membrane is attached to a $\sim$400 μm-thick silicon ring which is bonded to a three-inch Pyrex ring for mechanical support. Figure from reference [98].

lower-energy core-electron ionization is possible. As the energy is increased above an absorption edge, the absorption due to the ionization of that level decreases.

In the soft x-ray region where x-ray lithography is done, the figure shows that gold and tungsten are strong absorbers, while SiN_x, diamond, SiC, and PMMA are not strongly absorbing. This gives a choice of mask materials which allow adequate x-ray energy to pass through clear areas of the mask while still yielding a contrast in the x-ray intensity at the substrate.

Post-Mask X-ray Spectrum

As indicated in Figure 3-3, the x-rays which leave the source must pass through a 1.7 μm-thick SiN_x membrane and several centimeters of helium before they reach the x-ray mask. These will attenuate the x-rays somewhat, and moreover the spectrum will be modified because the attenuation is not constant with wavelength (*cf.* Figure 3-6).

Figure 3-7 shows a schematic of the x-ray mask used for exposures at MIT [99]. The mask consists of a 31 mm-diameter, 1 μm-thick SiN_x membrane which is patterned with gold absorber which is 200 nm thick. The membrane is attached to a ~400 μm-thick silicon ring or mesa which is bonded to an optically-flat Pyrex ring which provides mechanical support. In addition to the patterned absorber, the mask has a thin layer of plating base everywhere, typically 10 nm of Ti and 10 nm of Au, which allows electrical contact to electroplate the absorber gold. This plating base is thin enough that it allows most of the incident x-rays to pass through, but it does attenuate and modify the spectrum somewhat.

The modified spectrum which has come through the vacuum window and the helium is therefore passed through an additional micron of SiN_x and the plating base layers. This spectrum is incident upon the resist on the substrate in the clear areas of the mask. In the "dark" areas of the mask, the x-rays also pass through 200 nm of gold. The "dark" spectrum is attenuated with respect to the clear spectrum, but a not-insignificant amount of energy is transmitted through the absorber.

These x-ray spectra can be calculated by using the initial spectrum and the wavelength-dependent absorption properties of the materials through which the x-rays pass. The calculated spectra in the exposed and dark areas are plotted along with the initial spectrum in Figure 3-8. The Cu_L peak is attenuated by about a factor of ten in the clear areas, and by another factor of ten in the dark areas. The higher-energy spectral components are less well-attenuated.

3.3.2 Resist Dose

Contrast

What really matters, however, is the energy absorbed in the resist. Figure 3-6 shows that PMMA absorbs soft x-rays relatively poorly. This may at first seem to be unfortunate, but in fact it is beneficial. If the PMMA absorbed strongly, then the dose of absorbed energy would vary significantly through the resist thickness. Because absorption is weak, the dose is approximately constant throughout the film thickness.

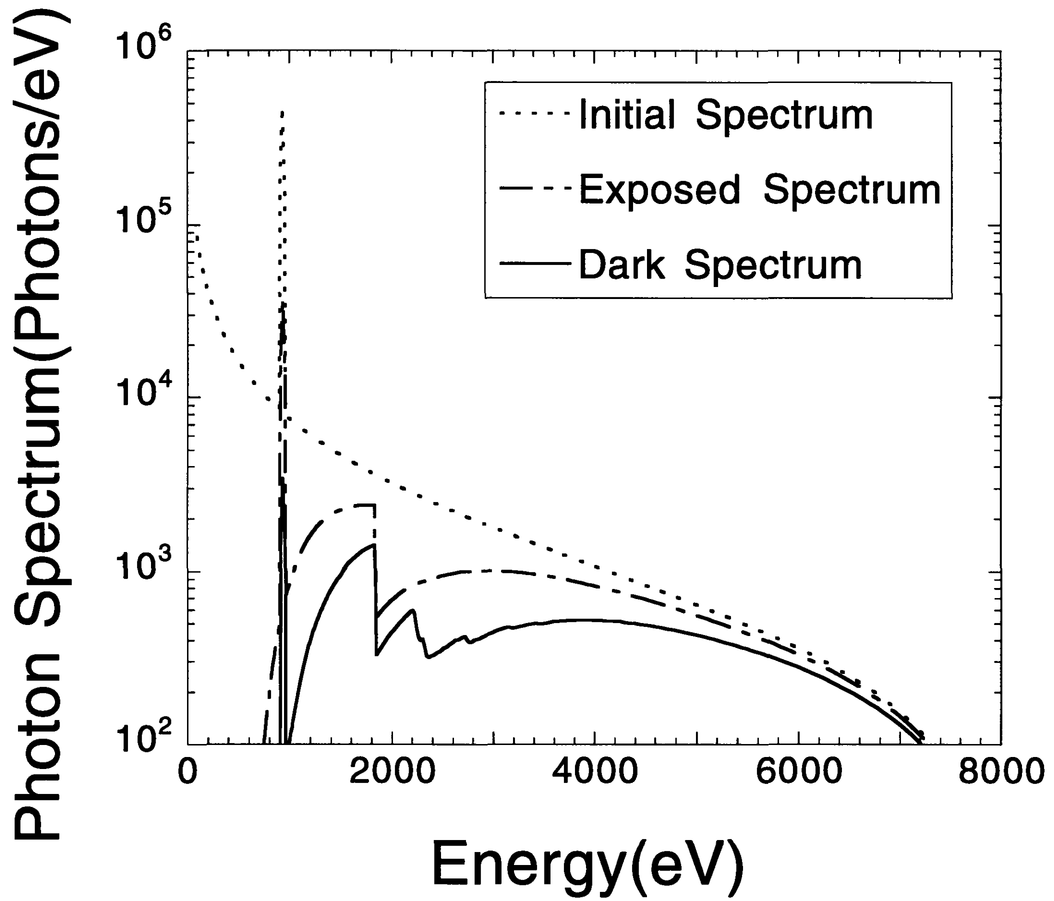


Figure 3-8: X-ray photon spectra emanating from mask in exposed and dark regions for a five-hour exposure at a source-to-substrate distance of 12 cm. Note the log scale on the y-axis.

	Dose in J/cm^3	
	Exposed Areas	Dark Areas
Cu$_L$ Peak	777	74
Bremsstrahlung	680	263
Total	1457	336
ABSORPTION RATIO	**4.3 : 1**	

Table 3.1: Calculated absorbed dose in J/cm^3 in 250 nm PMMA for a 12 hour x-ray exposure at a 10.5 cm source-to-substrate distance.

To calculate the contrast between the exposed and unexposed regions, we need to multiply the two spectra by the absorption in 250 nm of PMMA. Table 3.1 gives the calculated absorbed dose in 250 nm of PMMA exposed for 12 hours at a 10.5 cm source-to-substrate distance. The ratio of the total absorbed energies is 4.3:1. If the radiation were only the Cu$_L$ line, then the ratio would be 10:1.

This reduced contrast by itself is not too deleterious, however, since the development rate of PMMA goes as approximately the absorbed energy to the third power [100]. Therefore the development rate in the dark areas should be a factor of 80 less than the rate in the exposed areas, and for 250 nm of PMMA we would expect about 3 nm of resist erosion in the dark areas during development.

In an actual exposure which corresponds to the conditions of the simulation in Table 3.1, the development time for 250 nm of PMMA was 20 seconds, or the development rate was 12.5 nm/sec. According to Hawryluk *et al.* [100], this calculated dose should correspond to a development rate in 2:3 methyl isobutyl ketone: isopropyl alcohol (MIBK:IPA) of ~1.9 nm/sec. However the data in Hawryluk *et al.* is for a development temperature of 17° C. The development was done in the experiment at 21° C. Using the development activation energy for 1:3 MIBK:IPA of 2.426 eV found in Greeneich [101], this temperature difference should correspond to a factor of 4.6 in development rate. The expected development rate for the calculated spectrum should therefore be 8.7 nm/sec, which is remarkably close to experiment.[3]

[3]Similar comparisons done with other x-ray exposures were not as close. The cleanliness of the source and vacuum window, alignment of the electron optics, and other factors (such as the presence or absence of polyimide on the exposing mask) can affect development time. Results were within a factor of 3-6, and experimental development rates were slower than predicted in the cases where the

Lateral Effects

In sub-50 nm x-ray lithography, lateral smearing of the pattern, even at length scales which at larger feature sizes is negligible, becomes critically important. There are three main mechanisms which are responsible for lateral pattern degradation: x-ray-generated photo- and Auger electrons, penumbra, and diffraction.

X-ray Generated Electrons: As discussed in Section 3.3.1, the absorption of an x-ray photon in PMMA gives rise to a photoelectron and an Auger electron. The absorption of the photon occurs at one atomic site, but its energy is carried away from that site by the generated electrons.

This energy redistribution occurs though interactions between the electrons and the polymer molecules of the resist along the electron trajectory. The interactions scission the polymer chains, making the resist more susceptible to dissolution in a developer. The photo- and Auger electrons can therefore blur the pattern which is incident upon the resist.

Understanding the effect of this blurring in an actual x-ray exposure can be quite complicated. The photon spectrum is comprised of many energies, and each absorbed x-ray photon will generate a photoelectron with a different kinetic energy $E_{el} = \hbar\omega - E_I$, where E_I is the energy required to ionize a particular inner shell electron.

Ocola [102, 103] has modeled the generation and scattering of photo- and Auger electrons and their interactions with the resist. The results of this modeling can be parameterized as a three-dimensional point-spread function (PSF). The PSF gives the probability density of energy dissipation about the generation point. The PSF can be determined for an electron of a given starting energy (electron PSF), and for a distribution of photo- and Auger electrons which are generated by the absorption of monochromatic x-rays (photon PSF). The photon PSF for a 1.3 keV x-ray absorbed in PMMA is shown in Figure 3-9.

---

rate was off significantly, indicating that these other effects (all of which will decrease dose) could explain the discrepancy.

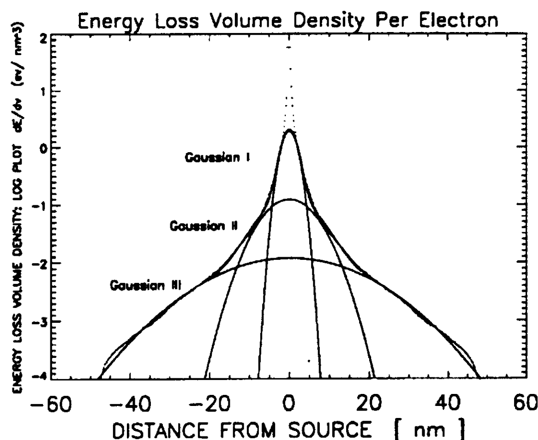


Figure 3-9: Photon point spread function (PSF) for a 1.3 keV photon absorbed in PMMA. Energy deposition per unit volume is given as a function of distance from the source. A fit to three Gaussians is shown. Figure from reference [102].

Photoelectrons can be quite energetic and will therefore distribute themselves over a larger volume than will Auger electrons, which are monochromatic and of lower energy. This means that the dose, which is energy *per unit volume*, will be higher from the Auger electrons than from the photoelectrons.

By convolving the photon PSF with the photon intensity incident upon the resist for a line source in PMMA, it was shown that most of the dose falls far below the maximum photoelectron range, and that its range more closely corresponds to the carbon and oxygen Auger electron ranges of 3 and 6 nm respectively. For 1.3 keV photons in PMMA, a developed PSF width of 3 nm is predicted, which corresponds well to experiments done by Early [104], *et al.* which will be discussed in Section 3.5.1 Murata [105] also predicted a ~5 nm range for a photon PSF.

We should therefore expect that generated photo- and Auger electrons would broaden the image incident on the substrate by approximately 3–5 nm. What that incident image is, however, will depend on how the x-rays propagate between the mask and the substrate.

Penumbra: It was mentioned in Section 3.3.1 that x-rays are generated in the MIT bombardment source by focusing an electron beam into a spot on a copper target. The electron-beam spot will have finite width. This width of the x-ray source gives

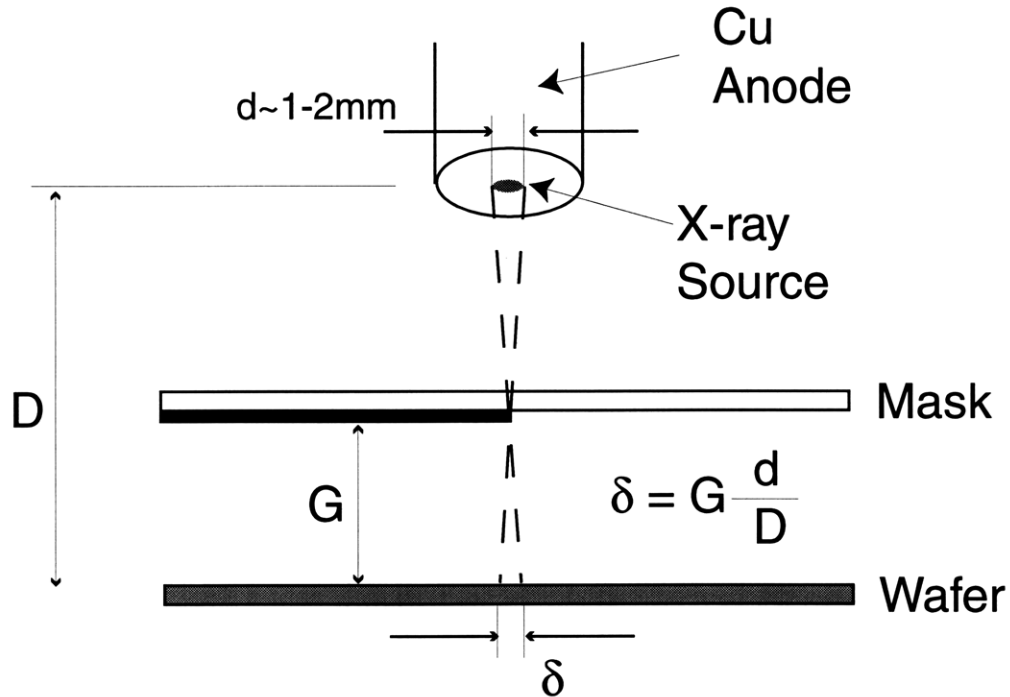


Figure 3-10: Penumbral blur due to finite source size and gap between mask and substrate. Penumbral effect is greatly exaggerated here for demonstration purposes. In an actual exposure, $G \ll D$.

rise to another mechanism for lateral pattern modification, penumbral blur.

The effect (greatly exaggerated) of penumbral blurring is shown in Figure 3-10. The finite source size, combined with a gap between the mask and wafer, create a gradation in x-ray intensity between dark and exposed areas. The extent of the blurring can be given approximately as

$$\delta = G\frac{d}{D},$$

where G is the gap between the mask and the substrate, d is the source size, and D is the distance between the source and the substrate.

Typical gaps for x-ray nanolithography are $5\,\mu$m or less, and typical source-to-substrate distances are 10-20 cm. The x-ray source size (determined by performing an exaggerated penumbral exposure with a large mask-to-sample gap) is typically 1-2 mm for the sources used at MIT. For these conditions, the worst-case penumbral

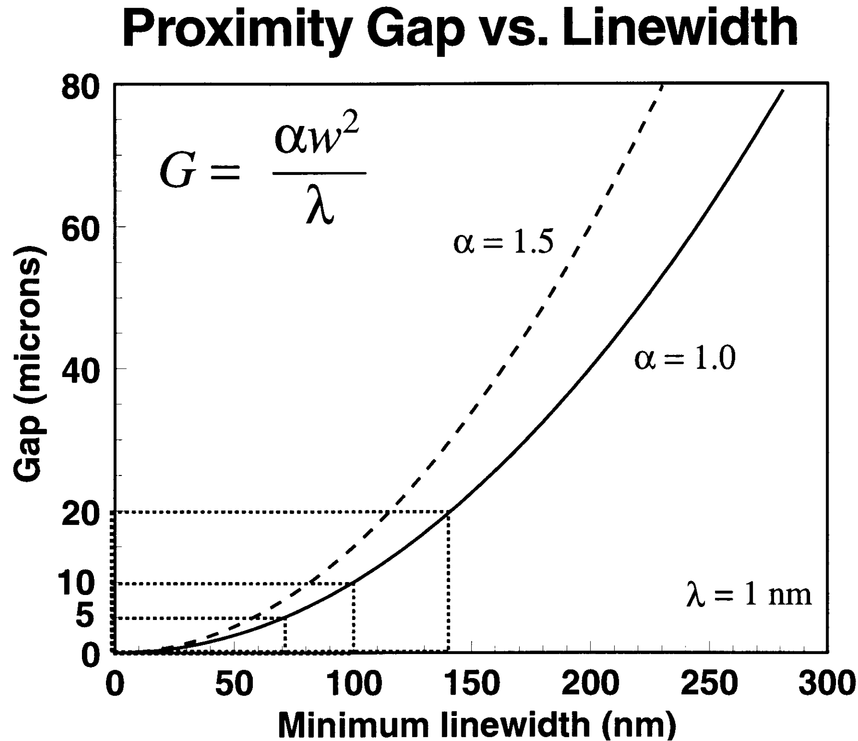


Figure 3-11: Mask-to-sample gap vs. minimum printable linewidth. $\alpha \leq 1.0$ gives excellent process latitude. $\alpha = 1.5$ can give reasonable process latitude. Figure from reference [108]

blur would be 100 nm. If D is increased to 20 cm and the gap is reduced by a factor of 10 to half a micron, then the penumbral blur becomes 5 nm.

Diffraction: Penumbral blur is not a stand-alone effect, however. It is intimately connected with diffraction due to their mutual dependence on mask-to-substrate gap. In the near-field regime, the relationship between mask-to-substrate gap G and minimum printable feature size w is well-known to be

$$G = \alpha \frac{w^2}{\lambda},$$

where λ is the wavelength of the x-ray, and α is a scaling parameter.[4]

Obviously the value of α is very important to how small a feature can be printed

[4]This expression comes from a modification to the expression for period doubling of a grating due to near-field diffraction, which would occur at $\alpha = 2$ [106, 107].

for a given gap and x-ray wavelength. If one assumes that the fields immediately after the mask have sharp boundaries (Kirchoff boundary conditions) good pattern replication would be possible only at a maximum value for α of about 0.5 [109]. However, the assumption that the fields are sharply-defined is invalid for several reasons. The absorber on an x-ray mask acts like a lossy dielectric, which tends to reduce the intensity of higher-order spatial frequency components in the image [110, 111]. In addition, penumbral blurring [112] (or similar edge-softening effects due to mask vibration or edge-wall taper [113]) will help to eliminate higher-spatial-frequency components of the image. Phase-shifting by the absorber [114] can also improve resolution at a given gap.

The result of the more complete calculations and experiments [112,115–119] indicate that α values of 1.0 or even 1.5 can produce acceptable pattern replication. If we define

$$\beta = \frac{\delta}{w},$$

the ratio of penumbral blur to linewidth, it is found [112] that $\beta = 0.6$ yields good process latitude with $\alpha = 1.0$ or 1.5.

Figure 3-11 shows a plot of gap vs. linewidth for $\alpha = 1.0$ and $\alpha = 1.5$ and $\lambda = 1\,\mathrm{nm}$. For $\alpha = 1.0$ it can be seen that 100 nm feature sizes can be printed with a $10\,\mu$m gap, and 50 nm feature sizes require a gap of $\leq 2.5\,\mu$m.

3.4 Device Fabrication with X-ray Nanolithography

We have closely examined factors which contribute to the lithographic input and the resist dose for a proximity x-ray exposure, and discussed what might limit the resolution. For this work, however, we are interested in more than just x-ray exposures – we are interested in going from a gate pattern design, such as the one described in Chapter 2, to producing that pattern on a small wafer or chip. That process involves patterning x-ray masks in addition to exposing the chip or wafer.

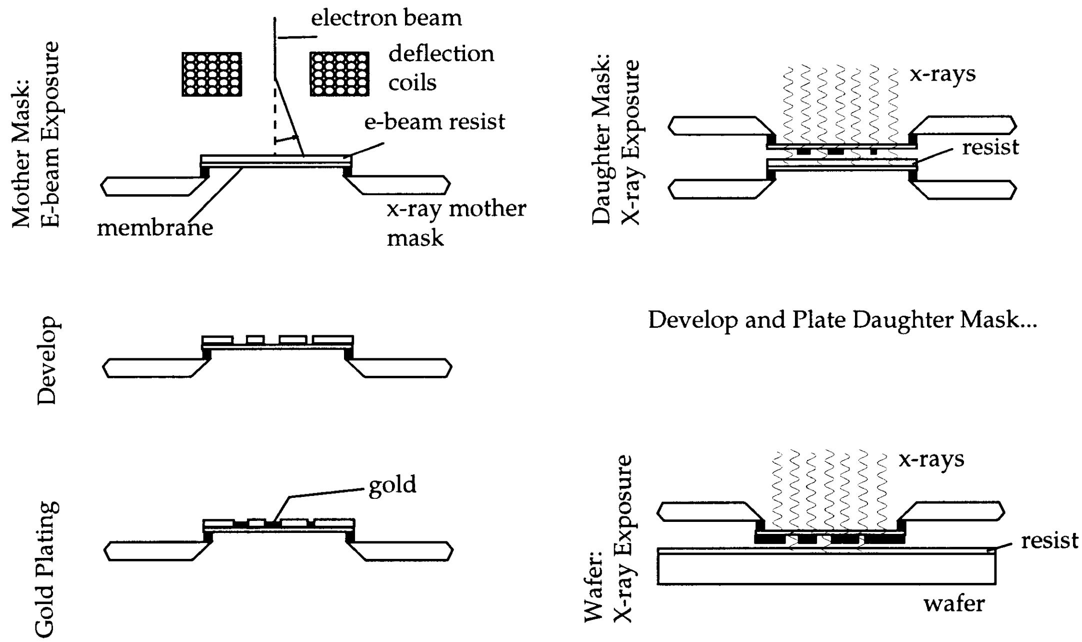


Figure 3-12: Full x-ray fabrication sequence for a device. An x-ray mother mask is written with an electron beam. The mask is then developed and gold is electroplated into the developed PMMA mold. This mother mask is used to expose a daughter mask with x-rays. The daughter mask is developed and plated as the mother mask was. The daughter mask is of the correct polarity to expose a wafer with x-rays for liftoff of metal.

3.4.1 Overview of Full Fabrication Process

The full x-ray device fabrication process is outlined schematically in Figure 3-12. The first step is to electron-beam write a "mother" mask with the desired pattern. This mask is then developed and gold is electroplated into the mold formed by the developed PMMA. This gives a mask with x-ray absorber everywhere the electron beam wrote (assuming a positive resist). This mask is the wrong polarity for liftoff of fine features on a substrate, however, so a negative replica or "daughter" mask must be made. In the fabrication of this daughter mask, an x-ray exposure is performed using the mother mask to create a PMMA mold on the daughter which is the inverse of that which was on the mother mask. This daughter mask is then electroplated with gold. A fine gold line on the mother becomes a small gap in the gold absorber on the daughter with approximately the same width. This mask is then the correct polarity to expose a substrate for liftoff.

A more complete discussion of various aspects of the above process is given in the

Line Dose(nC/cm)	Linewidth(nm)
0.75	0
0.9	29
1.00	32
1.50	57
2.00	60
2.50	75
3.00	71
3.50	78
4.00	83

Table 3.2: Table of single-pass electron-beam linewidths as a function of dose.

following sections. More specific details or recipes are, unless otherwise indicated, given in Appendix A.

3.4.2 Mother Mask Fabrication

Pattern Design and E-Beam Writing

The desired pattern is designed in KIC or some other CAD program and converted to the appropriate file format for the e-beam tool used. Masks for this work were written at the Naval Research Laboratory (NRL) in Washington, DC on a JEOL JBX-5DII electron-beam lithography system. The pattern data is transferred electronically. Specifics of the data conversion and transfer are given in Burkhardt [6].

A mask is coated with approximately 250 nm of PMMA and focusing artifacts (lightly-etched edges of gold fingers which protrude onto the membrane) are prepared. The mask is then shipped to NRL.

The e-beam writing at NRL is done at 50 keV, with beam currents ranging from 10–15 pA for the smallest features to ~3 nA for the largest features. For large pads an areal dose of 350–400 $\mu C/cm^2$ is used. For single-pass-lines (SPLs), a linear dose of about 1.0 nC/cm is usually the minimum required for clearing in the standard development time. By using a range of SPL doses, one can vary the developed linewidth. Table 3.2 shows this variation of linewidth with dose for a typical mother mask written at the NRL. Sub-40 nm linewidths are usually attainable.

Development and Electroplating

The mask is developed for 90 seconds in 1:2 MIBK:IPA which is held at 21° C. The development time for e-beam exposures is kept constant. After developing, a short oxygen RIE descum is performed to remove any thin layer of organic material which might remain from the development.

A film of 200 nm of gold is then electroplated onto the mask. Electroplated gold has been shown to be an excellent means of pattern transfer for feature sizes below 20 nm with very high aspect ratios [120]. For our electroplating we use SEL-REX BDT510, a commercially-available gold plating solution. The plating is done at $0.4\,\text{mA/cm}^2$ to minimize stress in the plated film. The bath is monitored to ensure that the plated film will have a small grain size to allow for good pattern replication. A good discussion of gold electroplating and plating bath characterization is found in Burkhardt [6].

The advantage of e-beam writing on a thin membrane is that electron backscattering from the substrate is all but eliminated. When e-beam writing on a thick substrate, it is well-known that fine features which are close to a large feature tend to widen. This is known as the proximity effect, and is due to the added dose from the more widely-distributed backscattered electrons which emanate from the large feature. When writing on a $1\,\mu$m-thick membrane, the proximity effect is significantly reduced. Figure 3-13 shows a scanning electron micrograph of a 55 nm-wide single-pass line written between two large pads. The gap between the line and each pad is also 55 nm. There is no appreciable widening of the line as it passes between the pads. The pads were written with a dose of $400\,\mu\text{C/cm}^2$, and the single-pass line was written at 1 nC/cm. A pattern such as this is unlikely to be successfully written on a bulk substrate. See Ghanbari [90] for more discussion and Monte-Carlo electron-beam exposure simulations and resist development simulations of fine patterns on thick substrates and membranes.

The reduced backscattering from e-beam writing on a membrane also allows for finer linewidths than could be attained on a substrate. Figure 3-14 shows a 30 nm-

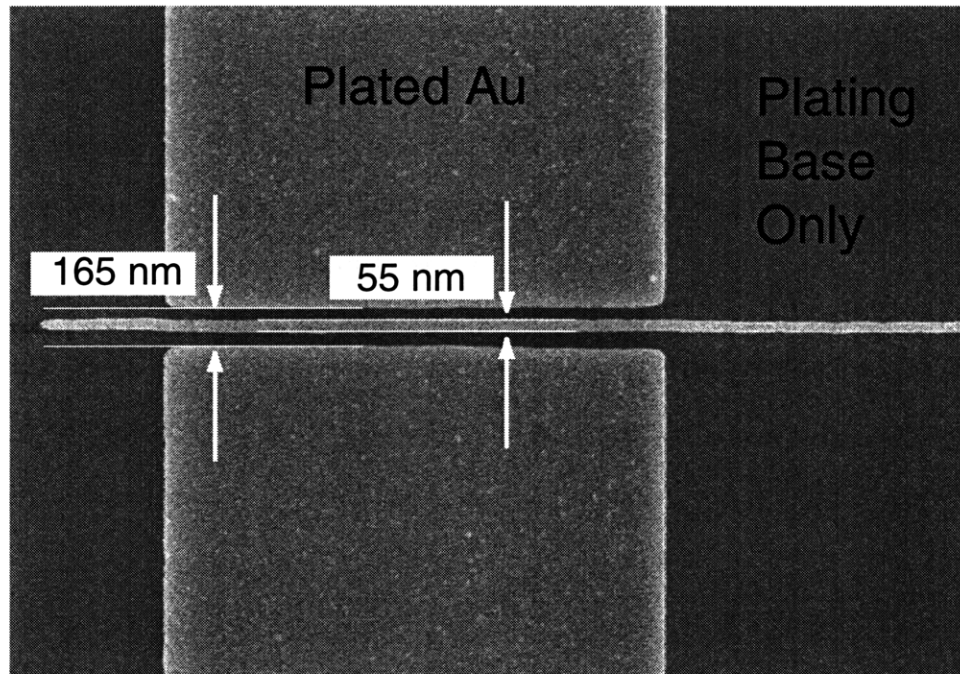


Figure 3-13: Scanning electron micrograph of test pattern on a mother mask showing remarkable lack of proximity effect. Image is of electroplated gold on SiN_x membrane coated with a Ti/Au plating base. The 55 nm-wide line shows no appreciable broadening as it passes between the $2\,\mu$m-square pads which are spaced 165 nm apart. It is highly unlikely that this pattern would be successfully written on a bulk substrate.

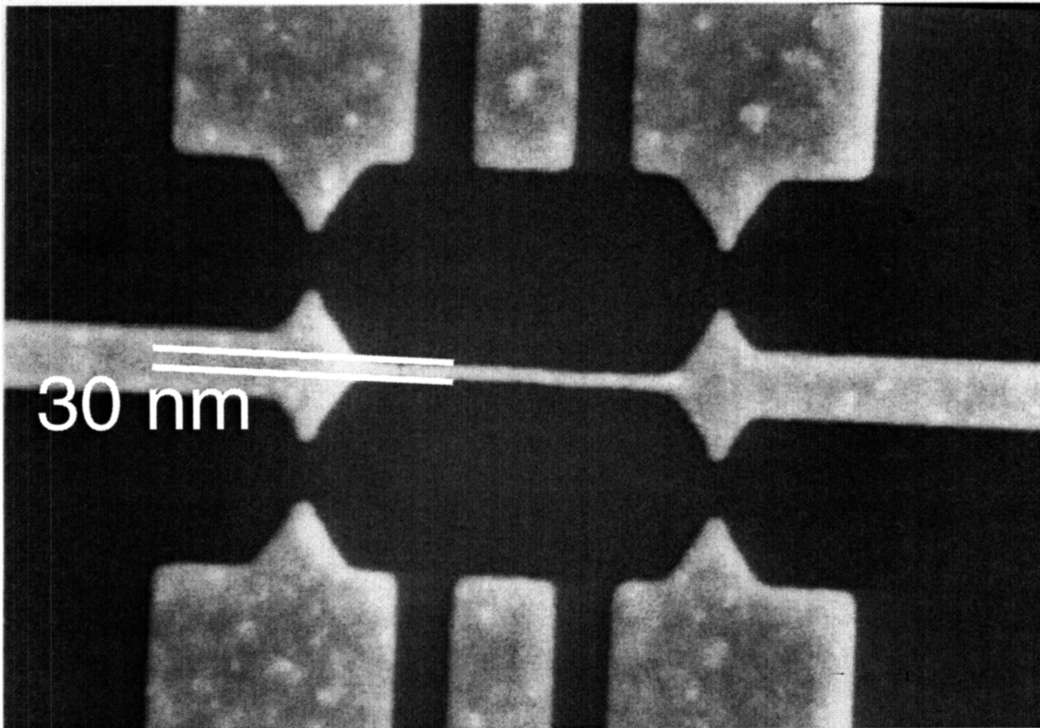


Figure 3-14: Scanning electron micrograph of device pattern on a mother mask. Image is of 200 nm-thick electroplated gold features on a 1 μm thick SiN_x membrane. The center line is 30 nm wide.

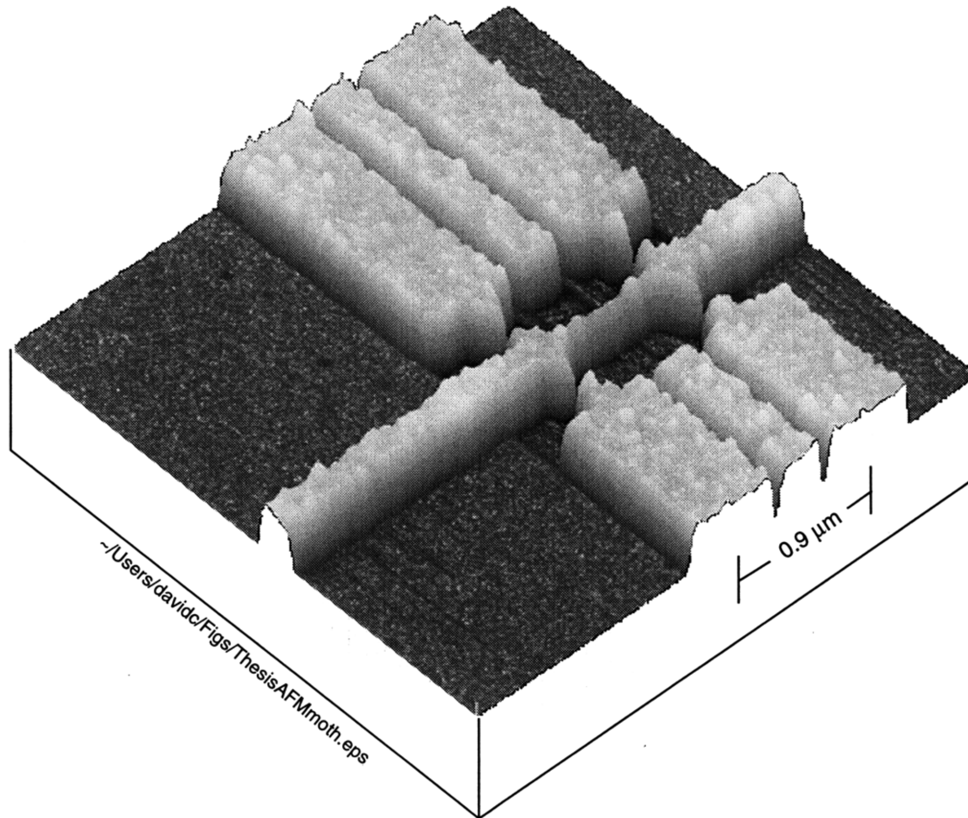


Figure 3-15: AFM image of device pattern on a mother mask. Electroplated gold features are 200 nm thick on a 1 μm-thick SiN_x membrane. The center line is 0.9 μm long and 40 nm wide. In this case it was found that the thickness of the absorber in the fine line was less than that of the other gates.

wide line on a device pattern on a mother mask.

After stripping the PMMA, the thickness of the plated gold can be checked with an AFM with little chance of harming the SiN_x membrane. Figure 3-15 shows an AFM image of a device gate pattern on a mother mask.

3.4.3 Daughter Mask Replication

Once the mother mask processing is completed, it can be used to create a negative replica which will have the correct polarity for a liftoff exposure on a substrate. For a mother mask with fine features, the gap between the mother and daughter mask during this process will be critical.

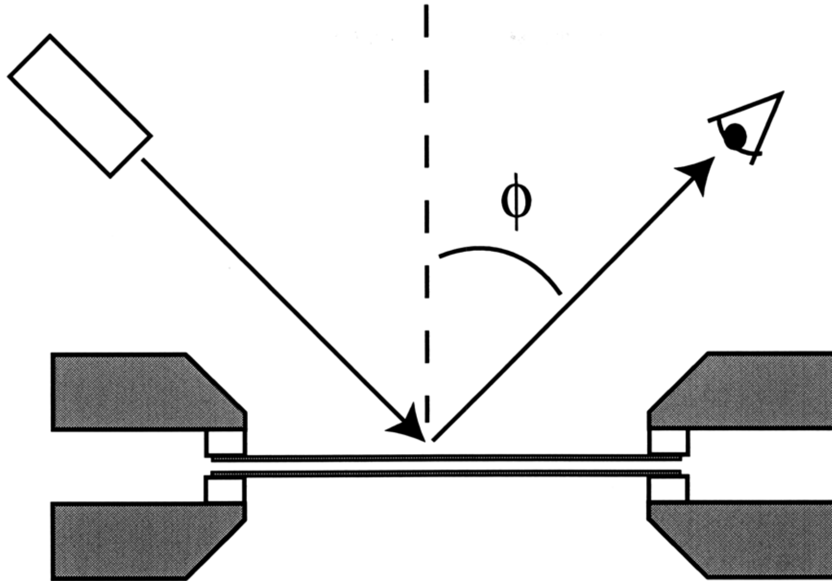


Figure 3-16: Schematic of gap monitoring in microgap exposures. A monochromatic green light source is used to illuminate the mother-daughter pair. As the angle ϕ of the source and eye are moved from 45° to perpendicular to the membrane, the number of fringes passing a particular point on the mother mask are counted. The gap at that point on the mask is approximately the number of fringes less one.

The gap is monitored as indicated in Figure 3-16. A monochromatic ($\lambda = 543\,\text{nm}$) green light source is used to illuminate the mother-daughter pair from above. As the angle ϕ of the source and eye are moved from 45° to perpendicular to the membrane, the number of fringes passing a particular point on the mother mask are counted. The gap at that point on the mask is approximately the number of fringes less one. This technique was published by Schattenburg *et al.* [87] for a mask above a silicon or SiO_2 substrate illuminated with a sodium (yellow) lamp.

To confirm that the technique was valid for two masks illuminated with green light, and with the mother mask possibly having a coating of polyimide (a technique which is used in contact printing, which will be discussed in detail in Chapter 4), simulations were performed using the program "`layer`", an ellipsometric simulation program written by Mark Schattenburg. The results of a simulation for a $10\,\mu$m gap between two masks with and without a $1\,\mu$m layer of polyimide on the mother mask are shown in Figure 3-17. To perform this simulation the index of refraction n of the spun-on thin-film polyimide (which is birefringent) was interpolated with angle

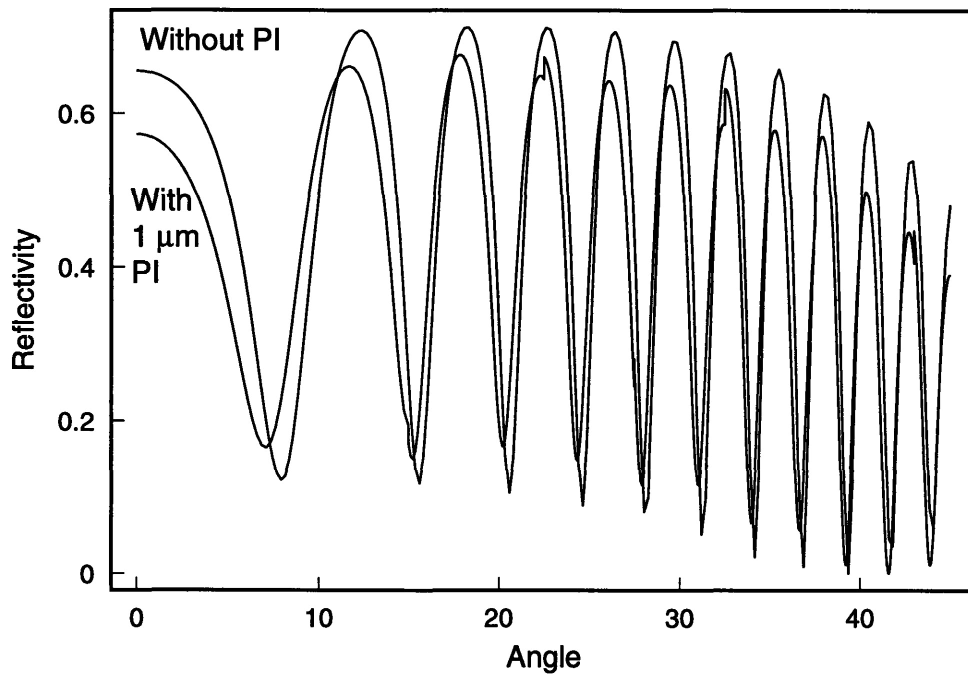


Figure 3-17: Simulation of reflectance of mother/daughter mask pair vs. angle for two mask membranes. The simulated gap is 10 μm, the light wavelength is 550 nm, and the two curves are for a mother mask with and without a 1 μm polyimide coating. 11 dips can be counted in the reflectance.

between its in-plane and perpendicular values of 1.9 and 1.6 [121]. Little difference is seen between the two cases, and the number of dips in the reflectance is the gap in microns plus one. Similar results were obtained with simulations for gaps down to one micron.

When one x-ray mask is placed on top of another in the manner shown in Figure 3-16, a gap of 4–5 μm is typically found. This is due primarily to buildup of resist and polyimide at the edge of the mesa, although non-flatness of the mesa does set a lower limit. As discussed in Section 3.3.2, for feature sizes below 50 nm a gap of less than 5 μm is required. Chapter 4 will discuss previous methods used and new methods developed for this thesis to eliminate this gap (and do "soft-contact" x-ray exposures). Even without gap reduction, however, feature sizes of 60–70 nm and above should be replicable.

The daughter mask is exposed using the copper electron-bombardment source, as discussed in Section 3.3. Typical exposure times are 5-10 hours, depending primarily on the distance between the source and mask. Shorter distances result in shorter exposure times, but can also result in decreased uniformity of dose across the daughter mask membrane.

After exposure, the daughter mask is developed in 2:3 MIBK:IPA. Unlike the development of the e-beam-written mother mask, the daughter mask development is monitored by interrupting the development several times and measuring the PMMA step height at a feature edge with an AFM. This is necessary because of the uncertainty in dose from the x-ray exposure. As discussed in Section 3.3.2, x-ray dose is uncertain because source properties can change over time. In addition, there is some variability in nitride membrane thickness, plating base thickness or plating base composition from one x-ray mask to another. From the discussion of the spectrum in Section 3.3.1 it is apparent that these can all affect the spectrum and therefore the dose at the exposure plane.

The developed mask is then descummed and electroplated.

3.4.4 Device Patterning

The device substrate is then patterned with the processed daughter mask. For very fine features, the gap between the mask and substrate is again important. Unlike creating a daughter mask, for device exposures *alignment* of the gate level is usually necessary. Overcoming the problem of aligning a small chip or wafer and then achieving soft contact is discussed in Chapter 4.

The aligned chip is exposed and developed as for the daughter mask. The gate pattern is then deposited by liftoff. The processing of the GaAs/AlGaAs heterostructures is discussed in Appendix B.

3.5 Previous Pattern Replication Results

The goal of this work was to reliably fabricate structures with feature sizes below 50 nm on a device wafer or chip. Most attempts to fabricate sub-50 nm features, and certainly all sub-50 nm features produced in the NanoStructures Laboratory at MIT, have been done as demonstrations, using techniques which cannot yield arbitrary aligned device patterns on a wafer.

The last two sections of this chapter will investigate previous efforts at fabrication at below or near 50 nm size scales. Section 3.5.1 will examine previous ultra-fine pattern replication, and Section 3.5.2 will look at device results obtained prior to this work.

3.5.1 Pattern Replication Below 50 nm

Several authors have reported replication of features with dimensions less than 50 nm using x-ray lithography. Until recently, however, all ultrafine feature replication has been done using non-standard techniques which were useful for some device applications but not for arbitrary patterns.

Flanders [122] replicated 17.5 nm features with a C_K bombardment source. The mask was prepared by evaporating a thin film of tungsten at an angle onto a relief

structure in a polyimide membrane. This coated the sidewalls with a 17.5 nm-wide tungsten absorber which was several tenths of microns thick. Substrates were then exposed in contact with the mask. For this geometry and exposure setup, 17.5 nm was the minimum feature size which was able to be replicated; finer-pitched structures were not replicable. By using clever "bootstrapping" techniques, a variety of fine-pitched grating were produced and shown to be useful for liquid crystal alignment experiments, polarizers and artificial dielectrics, and a permeable-base transistor structure.

Early [104] *et al.* used a similar shadow evaporation technique to produce a mask with 30 nm features. The mask was used to expose substrates with three different electron-bombardment x-ray sources: Cu_L ($\lambda = 1.32$ nm), Al_K ($\lambda = 0.83$ nm), and C_K ($\lambda = 4.5$ nm). The linewidth of the developed PMMA was comparable for all three sources, proving that the maximum range of the photoelectrons generated in the resist (which varies significantly among these sources) is not the limiting factor in resolution for x-ray lithography.

These experiments clearly show that pattern replication is possible in the sub-50 nm regime with x-ray lithography. The shadow evaporation used for mask-making, however, is not a technique suitable for general purpose lithography.

A recent experiment by Simon *et al.* [119] has shown the capability of producing arbitrary features with sizes below 20 nm using both additive (Au electroplating) and subtractive (etching of W) mask processing. By using monochromatized ($\lambda = 1.1$ nm) radiation from a synchrotron, feature sizes as small as 15 nm were replicated as resist lines, and spaces were replicated in gold which was electroplated around those lines.

3.5.2 Previous Device Patterning Results

However, using the Cu_L source and gold electroplated masks available in the Nano-Structures Laboratory at MIT, pattern replication for actual devices has been, before this work, limited to feature sizes greater than or equal to 50 nm.

Probably the smallest previously-reported feature size replicated with conventionally-produced (*i.e.*, e-beam written and gold-electroplated) x-rays masks was a 50 nm

line-space pair of interdigitated electrodes which had been designed for a planar-resonant-tunneling field-effect transistor (PRESTFET) structure [98,115]. This pattern was written on an x-ray mother mask and the mother mask was used to expose a Si substrate which was then lifted off with Cr/Au. This was done at several gaps and showed that values of α up to 1.44 produced excellent pattern replication.

The 50 nm line-space structure is only shown as a lift-off pattern, not as a daughter mask. In Chu's PhD thesis [98], gate patterns on a wafer are shown with wider gaps between the electrodes, and the linewidth is said to be "approximately 50 nm," but no daughter mask is shown.

Eugster [25] also reported that the smallest features which could be replicated by x-ray lithography were 50 nm.

Burkhardt [6,123] reported several devices made with the full device fabrication process (mother, daughter, device) which have linewidths of 70 nm and up. He also shows a mother mask which has a 40 nm feature, and a daughter mask which was replicated from it. The linewidth on the daughter is not measured. However it appears to be broadened with respect to the mother (probably by 20–30% within the accuracy of the measurement from the image).

Although sub-50 nm features can be replicated with x-ray nanolithography, no device features of that size had been made at MIT using the full device fabrication sequence outlined in Section 3.4.1 prior to this work. The remaining chapters of this thesis will detail what parts of the full process were responsible for the inadequate process latitude below 50 nm and how the process was modified to extend the useful limits of x-ray nanolithography well into the sub-50 nm regime.

Chapter 4

Soft Contact Technology

4.1 Microgap X-ray Lithography and Limits to Replicable Feature Sizes

As discussed in Chapter 3, diffraction of the image is the most important factor in limiting the resolution of x-ray nanolithography. For sub-50 nm lithography, having a small gap becomes critically important. Table 4.1 presents the minimum feature size vs. gap data from Figure 3-11 in tabular form for $\alpha = 1$ and $\alpha = 1.5$, where α is a parameter which is proportional to the largest allowable gap for a given minimum feature size and exposure wavelength.

In microgap x-ray lithography, gaps for 100 nm feature sizes and larger are typi-

Minimum Feature Size (nm)	**Gap (μm)**	
	$\alpha = 1.0$	$\alpha = 1.5$
100	10	15
70	4.9	7.4
50	2.5	3.8
40	1.6	2.4
30	0.9	1.4
20	0.4	0.6
10	0.1	0.15

Table 4.1: Minimum feature size vs. gap presented in tabular form. $\alpha = 1.5$ gives the largest gap which has been shown to yield good processing latitude.

cally set by evaporating several-micron-thick aluminum studs through a metal shadow mask onto the mesa of the x-ray mask [87]. By depositing a known thickness of aluminum, the gap formed when placing the mask onto a wafer can be well-controlled.

Non-flatness of the mesa sets a limit to this gap of about 2 μm, which would limit replicable feature sizes to about 40–50 nm for an exposure on a wafer.[1] When exposing onto another mask for a daughter exposure or onto a small chip (rather than a whole wafer), resist buildup on the mesa of the mask or the corners of the chip will increase the gap to typically 4–6 μm, limiting replicable features sizes to more like 55–75 nm. Some means of decreasing the gap is therefore necessary for replicating feature sizes below 50 nm.

4.2 Previous Methods for Reducing Mask-to-Sample Gap

Technologies for gap reduction have been in existence in the NanoStructures Laboratory for several years, both for small-sample exposures and for daughtering. They are described below.

4.2.1 Gap Reduction for Mask Daughtering

A clever scheme for gap reduction when daughtering masks was devised by Burkhardt [6, 125]. By partially evacuating the space between the two masks, the membranes are pushed closer together by the force of atmospheric pressure from the outside. By monitoring the gap-dependent capacitance between the two masks (conveniently both have a layer of gold plating base on their membranes) and using feedback to a control valve in the vacuum system, minimum gaps of a micron or below can be maintained over the course of a several-hour x-ray exposure.

[1]There are mask fabrication schemes which would all but eliminate this mesa non-flatness by fabricating an optically-flat mesa in the Pyrex ring and bonding the membrane directly to it [124]. These masks, although in development, are not commonly available yet, and would most likely still be susceptible to resist buildup at the mesa edge.

There are some problems with this method of reduced-gap daughtering, however. Under the uniform loading of the pressure difference between atmosphere on the back side of the membranes and the partial vacuum between them, the membranes deflect parabolically. This means that for an initial mask-to-mask gap of 6 μm and a partial-vacuum-induced center gap of 1 μm, the gap halfway out from the center of the membrane would be 2.25 μm.

Diffraction effects could therefore vary significantly across the mask, and lead to changing linewidths or even loss of the finest features as the edge of the mask was approached. This problem could be avoided by designing a mask with the finest features only at the center, and for purely research purposes that would certainly be possible. However, this was not a technology which met the goal of being "manufacturable" in the university setting.

There was also a more serious practical problem with the method. Because it relied on the capacitance between the two masks to monitor the gap, control was susceptible to electrical shorts which could develop between the masks. Electrical insulation between the masks was provided by the PMMA layer on the daughter mask. Often in the several placements of the mother mask which were required to initially orient the masks as close to parallel as possible (due to the nonuniformity of PMMA buildup at the mesa edge), this PMMA layer would become scratched and the masks would short together.

Even if the masks were successfully placed and a partial vacuum was pulled, in a not insignificant percentage of exposures, a short would develop (and often go away and redevelop several times) during the course of the exposure. This would cause the gap to vary uncontrollably and fine features to be washed out.

A means of avoiding this more serious problem was tried. The shorting arose because of electrical contact between the plating base on the daughter mask and either the studs or the plating base on the mother mask. Replacing the conducting studs with insulating studs should have prevented the shorting problem. However the means of doing that were impractical. Evaporating SiO_2 studs was considered, but evaporated SiO_2 films tend to be high in stress, and a several-micron-thick film

would likely not adhere well. In addition, the evaporator must be cleaned after every series of SiO_2 evaporations to prevent contamination of other materials. This did not seem to be a practical solution.

Instead, polyimide studs were patterned onto the mesa of mother masks. Although in theory this could have been a workable solution, the nonuniformity of the polyimide thickness, even near the inside edge of the mesa, made for a large (many micron) gradient or "wedge" in the gap. This was also an impractical solution.

Yet a third problem existed with this method of gap reduction. Occasionally, either due to carelessness when initially setting the gap or because of mask-to-mask shorting and subsequent oscillations of the control circuit, the mask membranes would come into contact. Once this happened, upon reduction of the vacuum they would usually not come out of contact, presumably due to van der Waals forces holding them together. Separating the two masks without destroying one or both of them was quite difficult. The SiN_x mask membranes are quite robust, but will not hold up well to shear stress. When the two masks were in van der Waals contact in the old fixture, lifting the top mask by hand or with thin wires placed between the two Pyrex support rings would usually end up destroying the masks.

Fortunately, an alternative method was devised which overcame all of these problems – the masks were pulled by vacuum into intimate contact over most of their surface. The development of this technique (and the method of removing the masks from contact) is discussed in Section 4.3.1.

4.2.2 Gap Reduction for Small Samples

Another clever scheme had been devised by Chu and Ghanbari [90,98,126] for placing small (a few mm on a side) chips into contact with an x-ray mask. This procedure allows alignment of the chip to the mask before achieving contact, a necessity for any procedure which is to be used to print real device structures.

The procedure is as follows: The chip is aligned to the mask using an XYZθ stage. Once alignment is achieved, the stage is raised a few tens of microns so that the corners of the chip (which are raised because of resist buildup) push into the mask.

A voltage is then applied between the mask and the chip, which electrostatically pulls them into intimate contact. Once this happens, the applied voltage can be removed, and van der Waals forces hold the chip to the mask. The two can be transported across the laboratory to another fixture which pulls a partial vacuum in the space between the mask and a latex membrane. The vacuum deflects the latex membrane, pushing the chip against the mask to hold it in place during an exposure. The mask membrane is quite flexible and will conform to dust particles and the resist buildup on the corners of the chip.

With this contact method, the gap is in principle well-known. For a contact exposure, it is necessary to allow an escape path for gas (presumably lower-molecular weight PMMA fragments and/or residual solvent) which evolves out of the resist during x-ray irradiation. If this is not done a bubble builds up between the mask and the substrate, increasing the gap. The escape path is created with a double layer of polyimide on the mask. The first layer is slightly thicker than the gold absorber and serves to planarize the surface. The second layer is much thinner (~70 nm) and is patterned in a 100 μm-period grating pattern. When the mask is pushed against the substrate, gas can exit through the channels in the top polyimide layer. The exposure gap is therefore the thickness of the polyimide layers, approximately 250–320 nm.

The advantage of electrostatic contact is that in this system, once the force from the applied voltage pulls the mask membrane within a threshold distance from the chip, the two are pulled into intimate contact without further increase of the voltage. This can be shown by modeling the system as a movable conducting plate (the mask membrane) held above a fixed conducting plate (the chip) by a spring (the tension of the membrane). Figure 4-1 shows a schematic diagram of the model.

A force balance for this system in equilibrium gives

$$Kx = \frac{\epsilon_0 A V^2}{2(G_0 - x)^2},$$

where K is the spring constant, x is the displacement from the initial gap G_0, A is

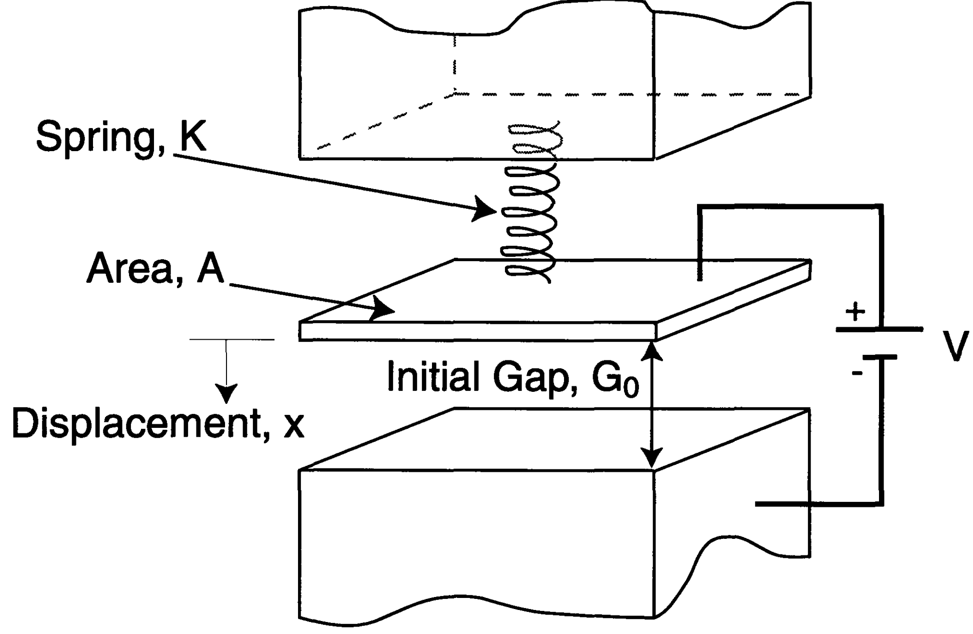


Figure 4-1: Schematic of electromechanical model for electrostatic contact. Plate is suspended by spring at an initial gap G_0. Applied voltage V results in a force between the two plates. At a voltage $V_{PI} = 2\sqrt{\frac{2K}{\epsilon_0 A}}\left(\frac{G_0}{3}\right)^{\frac{3}{2}}$ (where $x = G_0/3$), there is no stable static solution and the plate experiences "pull-in" where it runs into the bottom plate.

the area of the plate, and V is the applied voltage. Solving for V, we find

$$V = \sqrt{\frac{2K}{\epsilon_0 A}}\sqrt{x}(G_0 - x).$$

This function rises as x increases from zero and has a maximum at $x = G_0/3$. For V greater than this maximum value there is no static solution.

It can be shown that $x = G_0/3$ is also an inflection point of d^2E/dx^2, where the system energy $E = \frac{1}{2}CV^2 + \frac{1}{2}Kx^2$. For $x < G_0/3$, the solution is stable and the plate will remain at a fixed x, while for $x > G_0/3$, it is unstable and the plate will be pulled into the bottom electrode. The voltage at which this occurs is called the "pull-in" voltage,

$$V_{PI} = 2\sqrt{\frac{2K}{\epsilon_0 A}}\left(\frac{G_0}{3}\right)^{\frac{3}{2}}.$$

For the pull-in to occur requires that the gap be reduced to 2/3 of its initial value. This should (and often does) allow for a "zippering" effect where one part of the

chip comes into contact with the mask and the contact propagates across the chip. However, particles and resist edge buildup can prevent contact from occuring across large portions of the wafer (if they hold the membrane further than $2G_0/3$ away from the conducting part of the wafer) and sometimes prevent contact altogether.

For non-conducting substrates (such as the semi-insulating GaAs substrates on which the heterostructures are grown for the quantum dot devices) the insulating substrate is much thicker than the gap between mask and chip. G_0 is the gap between conducting plates, which in this case is the thickness of the wafer plus the mask-to-chip gap. Since in this case the gap can never be less than $2G_0/3$, the pull-in mechanism is not at work. The voltage must therefore be increased further and further to increase the force between the mask and the chip.

The voltage required to pull the chip and mask into contact can vary significantly, not only between different types of chips, but also from alignment to alignment with the same type of chip due to differences in resist edge buildup or particle number and size. Silicon chips are usually relatively easy to pull into electrostatic contact at low voltages (~10-30 V), although frequently only some portion of the chip (maybe 1/4 to 1/2 of the surface area) will come into intimate contact, while the rest of the chip remains at some gap – probably a micron or so, although this is hard to estimate since the area is small.

Pulling GaAs chips into contact with a mask is much more problematic. If the chip comes into contact at all, it is often only on a tiny part of one corner, while the rest of the chip is at an unknown gap. Also, because of the larger voltages required to achieve contact (sometimes as high as 100–150 V), it is not uncommon for an arcing event to occur which destroys both the mask (breaking it into shards) and the chip (creating a large crater where the arc occurs).

Two means were tried to improve the performance of the electrostatic contact process. For insulating substrates, a thin film of Ni was evaporated on the surface of the PMMA, around the side of the chip, and onto part of the back of the chip. This allowed electrical contact to be made to the front of the chip and the pull-in voltage to be easily reached. Unfortunately, it also created a large conducting surface

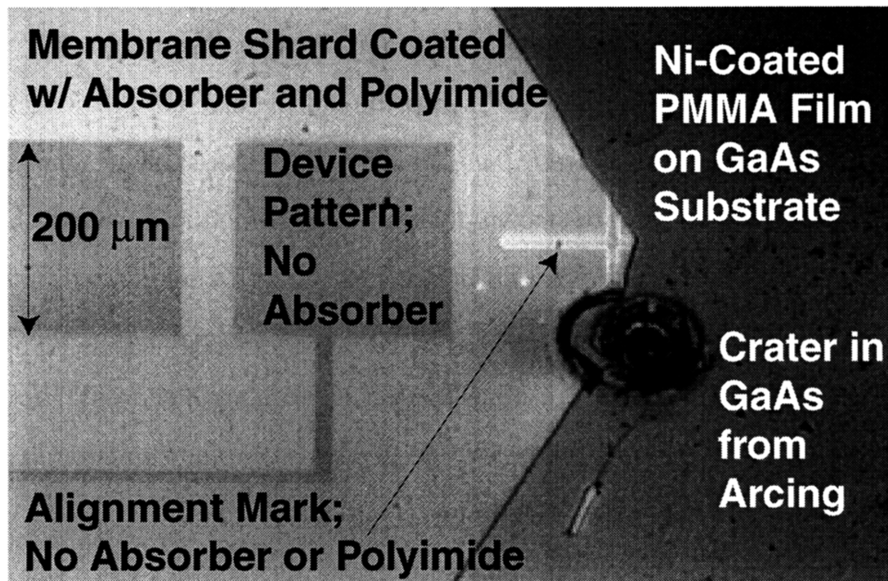


Figure 4-2: Optical micrograph of GaAs chip and membrane shard destroyed by arcing when attempting electrostatic contact. The arc originated from the alignment mark on the mask which had been cleared of polyimide in previous processing.

on top of the PMMA which made electrical contact to the alignment marks which were stripped of polyimide.

Figure 4-2 shows an optical micrograph of a Ni-coated GaAs chip which was destroyed while trying to make electrostatic contact with a daughter mask in an attempt to create a gate pattern. The micrograph shows the chip and a shard of the mask membrane. From the micrograph it is apparent that the arcing occured from an alignment mark on the mask (the bright cross) where the polyimide coating had been removed in order to improve optical contrast for alignment.

Since the PMMA on the GaAs was coated with a thin film of nickel, it is not surprising that a failure occured in this case. It should be noted, however, that arcing also occured in masks which were fully coated with polyimide. This micrograph shows the extent of substrate damage which can occur from an arcing event.

Making electrical contact to the back side of the SiN_x membrane was also tried. A thin film of Ti/Au (about plating base thickness) was evaporated onto the back side of the membrane. It was electrically contacted with a ball bearing which was

attached with conducting epoxy to a movable probe arm. It was thought that the insulating SiN_x membrane would allow large voltages to be applied which could pull the Ni-coated front of the chip into contact with the mask. Unfortunately arcing occured in this case as well, probably because of high field concentrations at the edges of the chip or at particles combined with the poor isolation characteristics of the silicon-rich SiN_x used for low-stress x-ray mask membranes.

In the end a contact method which used partial vacuum was developed which allowed for alignment and reliable full-chip contact with no danger of mask or substrate damage due to arcing. This technique is discussed in Section 4.3.2.

4.3 Full Vacuum Contact

Vacuum contact was the technology chosen to provide a known gap for exposures of both daughter masks and small samples. The techniques developed have resulted in a robust process which can reliably produce sub-50 nm features with a much-reduced risk of mask breakage and substrate damage. The methods for achieving vacuum contact for daughter exposures and aligned small-chip exposures are described below.

4.3.1 Vacuum Contact for Daughter Exposures

The main problem with placing a mother and a daughter mask into intimate contact was how to pull them out of contact after the exposure was finished. As was mentioned in Section 4.2.1, with the previous setup, if the masks were put into contact by mistake it was usually not possible to separate them without destroying them. This was because mechanical force was applied, usually manually, to separate the membranes, and any lateral motion of the masks resulted in shear stress on the membranes.

The solution was to design a fixture in which the masks were rigidly clamped before being pulled into contact, during the exposure, and while being separated. Instead of mechanical force, gentle pressure between the two membranes (the opposite of the vacuum which pulled them together) is used to push them apart. This is achieved by gently blowing into the tube which was used to pull the masks into contact. The

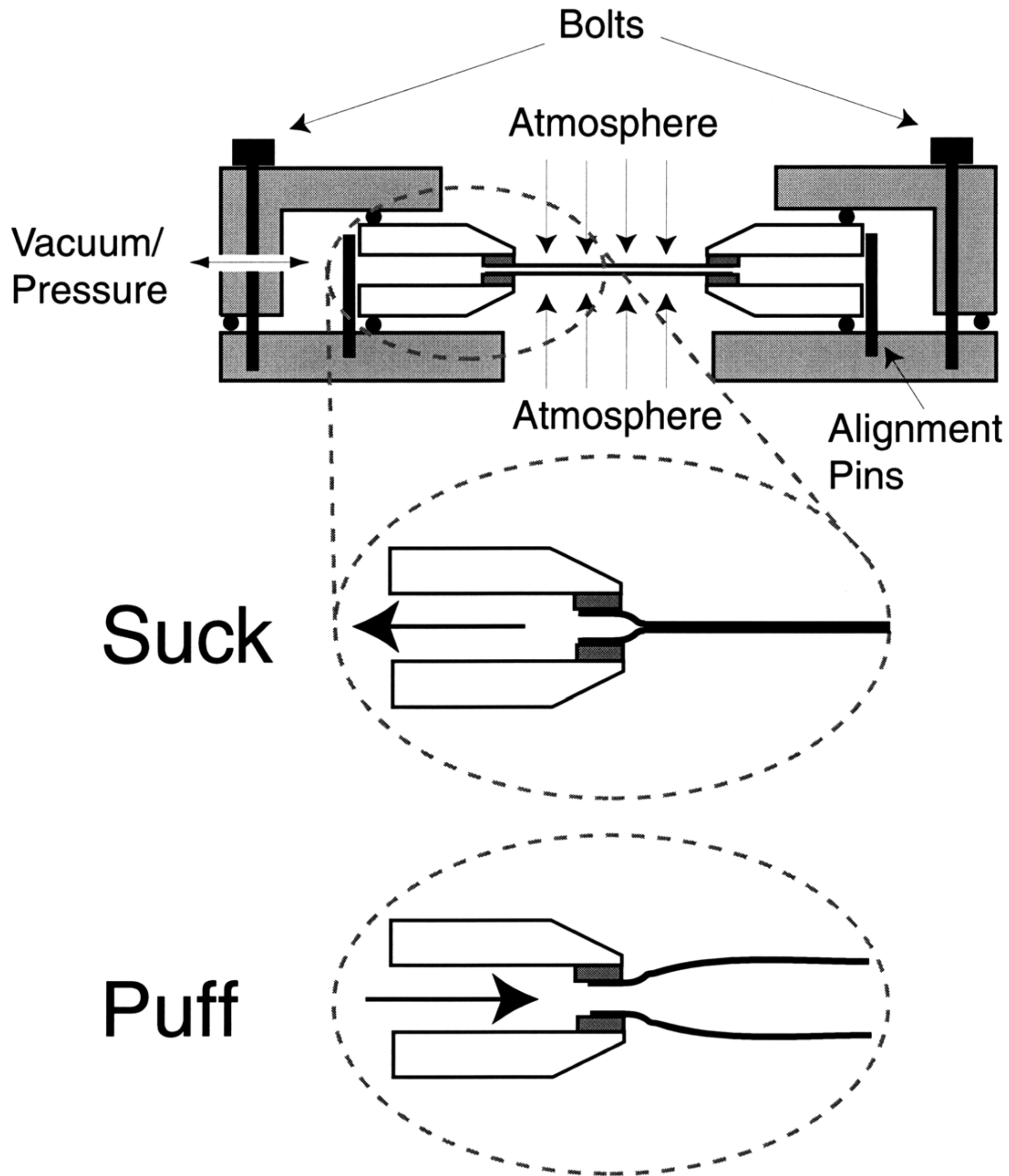


Figure 4-3: Schematic of "suck and puff" technique for performing daughter exposures in intimate contact. Masks are rigidly clamped by fixture. Partial vacuum is used to pull membranes into contact, and slight pressure is applied to push them out of contact. This last step avoids shear stress, which can rupture the membranes.

blowing technique can best be described as being similar to puffing on a cigar but with the opposite gas flow direction. This could, of course, be automated and done without the human breath.

This "suck and puff" technique has proven highly successful at contacting and separating masks for exposure. The mask membranes are pulled into contact over most of their 31 mm diameter; perhaps 1 mm at the edge of the membranes is not in contact during an exposure. The technique is shown schematically in Figure 4-3.

Since the daughtering exposure is now done in contact, the mother mask must be coated with polyimide to allow for evolution of gas out of the PMMA on the daughter mask during exposure. However, the 100 μm-period polyimide relief grating which is used for this purpose in mask-to-sample exposures is not adequate for mask-to-mask exposures. The flexible daughter mask easily conforms to the 50 μm channels and can block them. Instead a 20 μm-period grid is used. The smaller channel size seems to prevent blockage by the daughter mask, and using a grid rather than a grating allows any generated gas to find a percolative path to the edge of the membranes. With the two-dimensional grid, if a particular path is blocked, the gas can find another way around. Modifying the polyimide processing required some optimization of the polyimide etching, which is described in Appendix A.

4.3.2 Aligned Vacuum Contact for Small Samples

A similar means was used to pull small chips into contact with the mask. In this case it was necessary to allow for alignment of the chip to the mask before pulling the vacuum to achieve contact. Most methods of forming a vacuum seal, however, prevent movement of the two pieces to be sealed. It was therefore necessary to devise a means of creating the vacuum seal *after* the alignment took place. Since the mask and substrate were in partial contact at that point, it was important that the vacuum seal not cause any lateral motion, both for the sake of the alignment and to avoid shearing the mask membrane.

The fixture was designed as an inexpensive modification to the small-chip alignment fixture which was already in existence in the NanoStructures Laboratory [127].

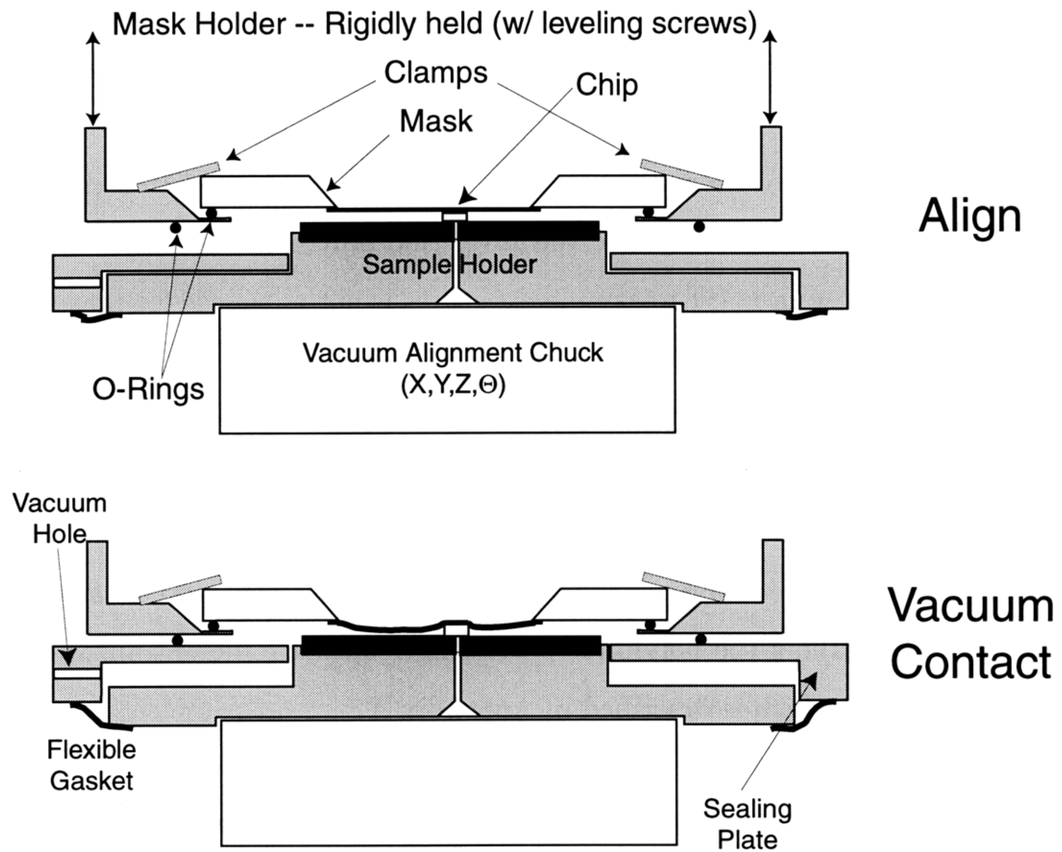


Figure 4-4: Schematic of vacuum contact technique for performing aligned small-chip exposures in intimate contact. Top: The mask is leveled with respect to the sample, the sample is aligned at a small gap, then the sample is pushed up so that it contacts the mask. Bottom: The sealing plate is raised, creating a sealed space which is evacuated through a vacuum hole in the side of the sealing plate. The flexible mask membrane conforms to the chip.

This allows switching between vacuum and non-vacuum alignments with minimal changes in tooling. It is shown schematically in Figure 4-4. The small chip is either held directly by a hole in the vacuum chuck (as shown) or it is glued with polyvinyl alcohol (which is soluble in water) to a two-inch silicon wafer which is held down by vacuum. (*n.b.* this vacuum is separate from the vacuum which pulls contact between the chip and the mask). The mask is clamped against a viton gasket seal on the mask holder.

The mask holder is leveled with respect to the chip and the chip is aligned at a small gap. It is then raised to just touch the mask. This is similar to the alignment procedure for non-vacuum contact. The top of Figure 4-4 shows the chip after the alignment procedure.

The bottom of Figure 4-4 shows how the vacuum contact is made. After the alignment, the sealing plate is raised with three thumbscrews (not shown) to seal against a gasket which is attached to the bottom of the rigidly-held mask holder. A flexible gasket between the sealing plate and the sample holder allows the sealing plate to move the few millimeters required to seal against the mask holder. Once the sealing plate is raised, a partial vacuum is pulled through a hole in its side. The mask conforms to the small chip and is held in place by the atmospheric pressure from above during the exposure.

When the exposure is finished, the partial vacuum in the mask-sample gap is released. Since the chip is still held to the sample holder by vacuum, lowering the stage overcomes the van der Waals force between the chip and the mask and separates the two without damage.

A photograph of a small GaAs chip placed in vacuum contact with a mask using this fixture is shown in Figure 4-5. The mask deflects rather dramatically around the chip. Several small particles which are on the chip can be seen in relief through the membrane. The mask conforms around these particles to give a minimal gap over most of the chip.

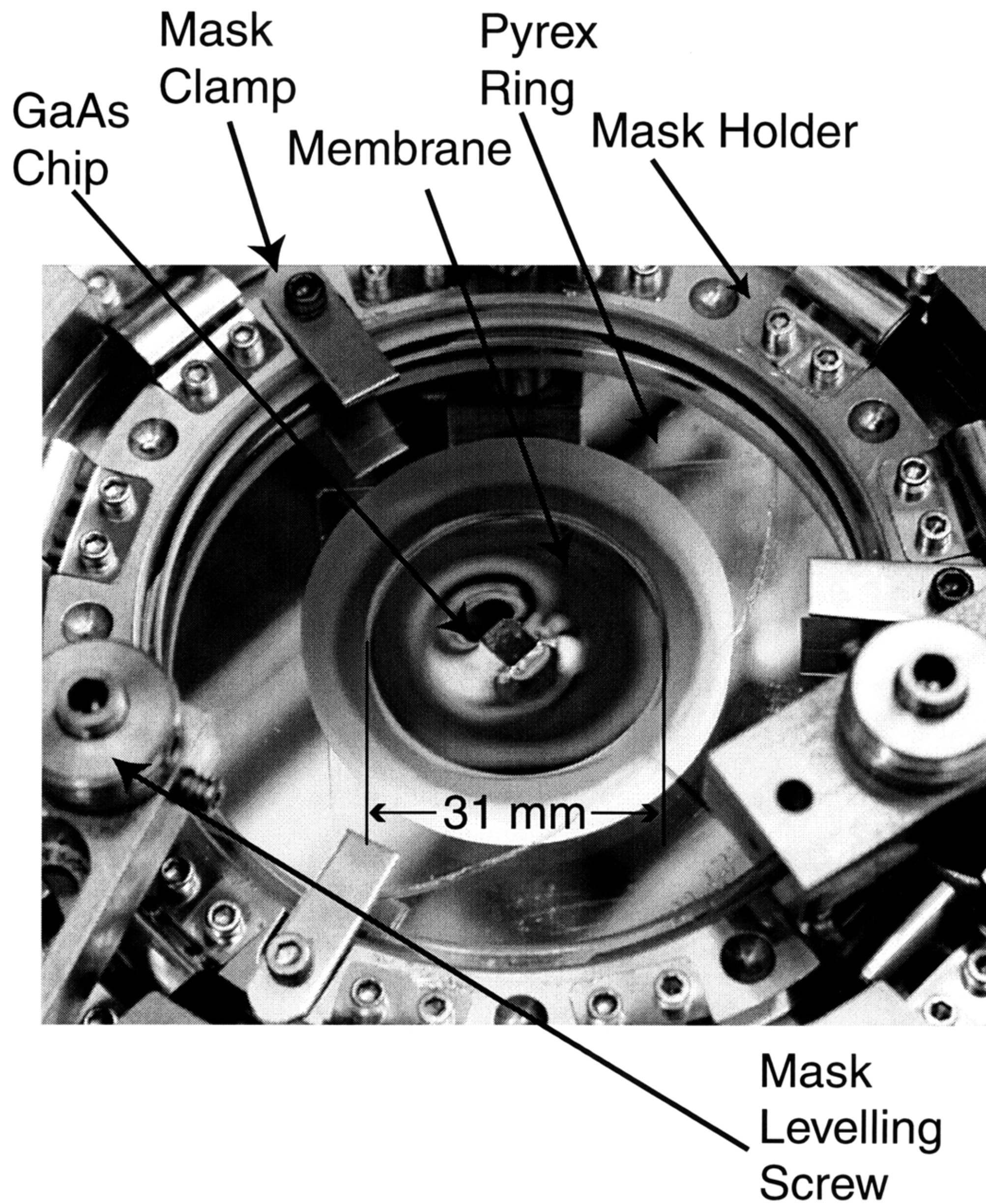


Figure 4-5: Photograph of small GaAs chip in vacuum contact with an x-ray mask. The mask deflects significantly around the chip and conforms to particles on the chip.

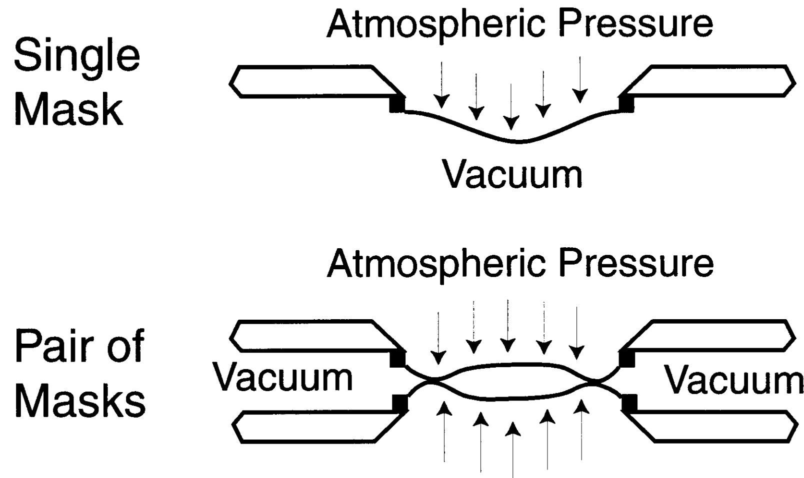


Figure 4-6: Schematic of membrane deflection for single mask and mother-daughter mask pair when placed under vacuum. Single mask deflects parabolically. Pair of masks contact each other at edges of membrane rather than at center as might be expected.

4.4 Analysis of Issues in Soft-Contact

In the course of developing the technology for vacuum soft-contact exposures for daughter and device exposures, some issues arose which seemed worthy of further investigation and analysis. They are discussed below.

4.4.1 Squeeze-Film Analysis of Membrane Dynamics

When differential pressure is applied across a circular membrane, the membrane deflects in a parabolic fashion, as described in Section 4.2.1. It would therefore not be surprising if when two daughter masks were placed in close proximity and the space between them was partially evacuated that the membranes would initially contact each other in the center. With the vacuum daughtering setup, however, it was invariably found that the masks initially contacted each other at the edge of the membranes, leaving a bubble of air in the center. This is shown schematically in Figure 4-6.

This bubble does not remain for long, however. The polyimide relief grid on the mother mask provides a path for the air trapped between the membranes to escape. Usually after about half an hour, the masks are in intimate contact. A time-lapse sequence of pictures of two masks coming into intimate contact is shown in Figure 4-7.

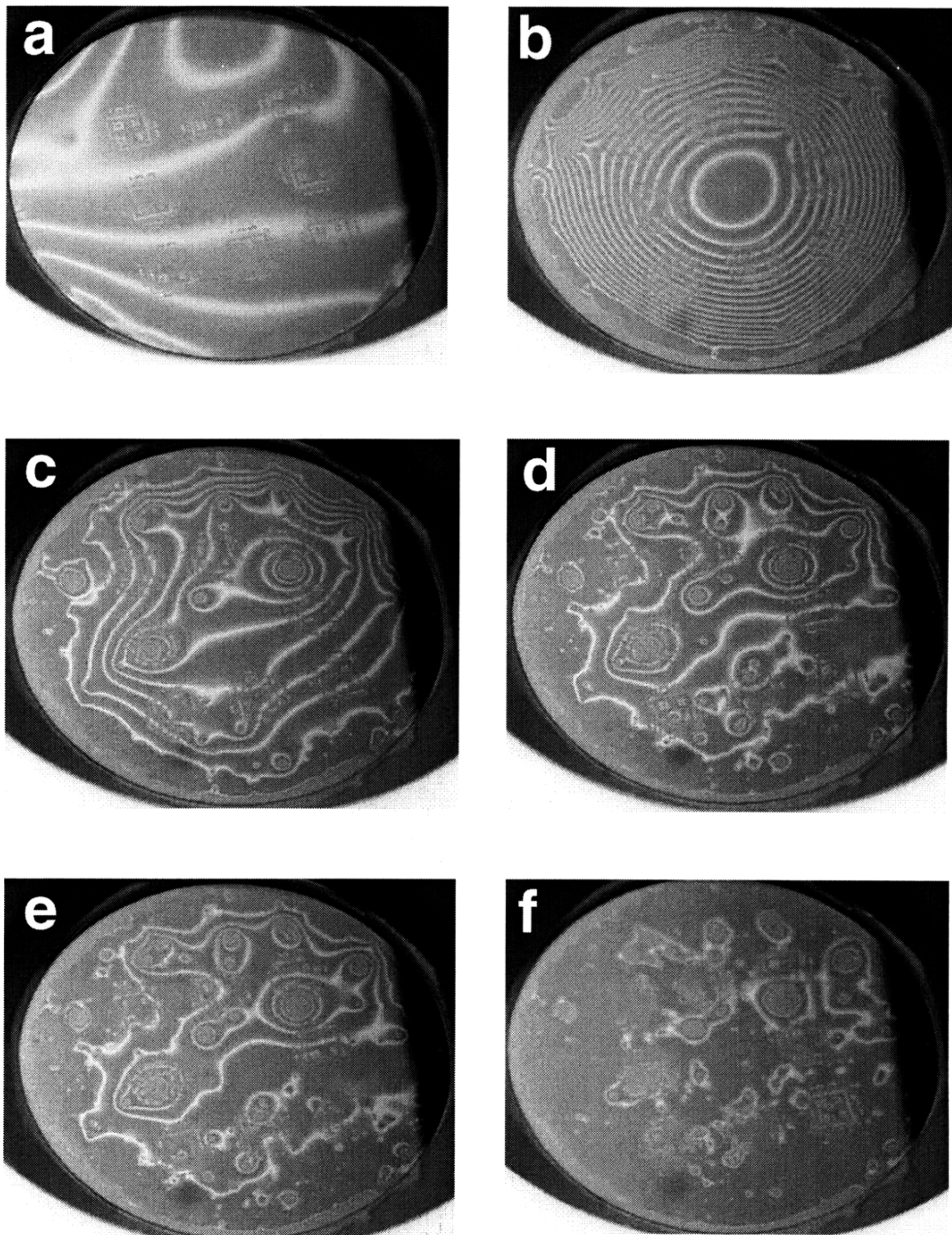


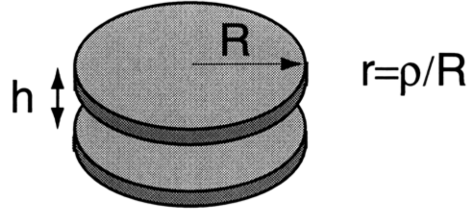
Figure 4-7: Time-lapse photograph sequence of two masks coming into vacuum contact. Masks are illuminated in monochromatic (green) light. Interference fringes indicate how gap changes across membrane. (a) Masks before vacuum is applied. The membranes are not perfectly parallel. (b) Vacuum first applied. The edges of the membranes come into contact first, trapping an air bubble in the center. (c)–(e) Area of contact increases. As membranes pull closer together, more particles can be seen. (f) Almost perfect contact. There is some trapped air around a few particles. Eventually that will also dissipate.

The masks are illuminated by monochromatic (green) light. In part (a) no vacuum has been applied. A few interference fringes across the membrane indicate that the masks are not perfectly parallel, although in this case they are quite close to being so. Part (b) shows the membranes immediately after the partial vacuum has been applied. The edges of the mask come into contact first, as shown by the fringeless ring and the central concentric fringes which indicate the bubble of air in the center. As time evolves, parts (c)–(e) show the bubble decreasing in size as more of the membranes come into intimate contact and more, smaller particles (surrounded by concentric rings of fringes) are evident. Part (f) shows the membranes almost totally in contact. Only a few particles have a significant amount of air surrounding them. Eventually there will be no visible interference fringes over the center of the membranes – only about 1 mm or so of the edge will show fringes.

The reason that the membranes do not contact each other in the center has to do with the dynamics of flow between the two closely-spaced membranes. This is a so-called "squeeze-film" effect [128]. An air molecule which is in the center between the two membranes can travel $\sim 5\,\mu$m up or down before hitting the membranes, but has a radius in the X-Y plane of 1.5 cm where it can travel and still remain between the masks. This is quite a large aspect ratio. Vacuum is applied to the membranes by reducing the pressure in the space *outside* the mask mesas. Gas molecules at the edge of the mask therefore immediately experience a pressure or density gradient and diffuse out from between the membranes. Because of the narrow gap between the membranes, it takes time for this rarefaction to propagate to the center of the masks, and by the time it does, the pressure at the edge is reduced enough to cause deflection, allowing the membranes to touch there.

The pressure profile immediately after the vacuum is applied can be modeled as a double step or Π function with a reduced pressure outside the membranes and atmosphere between the membranes. This is the same relative pressure profile we would expect if we had two closely-spaced rigid plates and we quickly stepped one plate closer to the other, causing a compression of the air between them.

We will use the system of two closely-spaced rigid plates as a model for the mem-



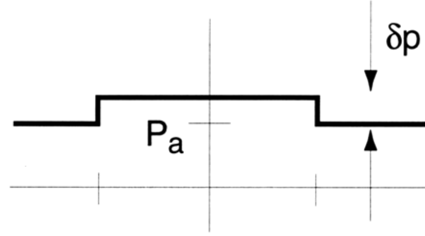


Figure 4-8: Model system and initial pressure profile for squeeze-film analysis.

brane system. An illustration of the model system is shown along with the initial pressure profile in Figure 4-8. This modeling will ignore any flexing of the membrane. The analysis follows Griffin [129] and Yamanami [130].

We assume that the separation between the plates is much smaller than their radius, $h \ll R$. We also assume the motion of the plates is only towards each other; there is no rotational or lateral motion. We also assume that the flow of gas is viscous (*i.e.* the inertial energy is small compared to viscous dissipation) and the system is isothermal (therefore the pressure p is proportional to the density ρ) with constant viscosity.

We begin with the continuity equation:

$$\nabla \cdot \rho \vec{v} = -\frac{\partial \rho}{\partial t}.$$

For viscous flow in a narrow channel, the velocity is proportional to the pressure gradient [131]:

$$\vec{v} = -\left(\frac{h^2}{12\eta}\right) \nabla p,$$

where η is the viscosity. Substituting this expression into the continuity equation and substituting pressure, p, for density, ρ (since they are proportional), we arrive at the governing equation for the pressure distribution:

$$\nabla^2 p^2 = \frac{24}{h^3} \eta \frac{\partial}{\partial t}(ph),$$

where t is time.

This is a non-linear partial differential equation. However, by assuming small variations in the pressure $p = P_a + \delta p$ (or in the plate displacement $h = h_0 + \delta h$), where $\delta p \ll P_a$ and $\delta h \ll h_0$, we can linearize the equation to yield

$$\frac{h_0^2 P_a}{12\mu}\nabla^2\left(\frac{\delta p}{P_a}\right) - \frac{\partial}{\partial t}\left(\frac{\delta p}{P_a}\right) = \frac{\partial}{\partial t}\left(\frac{\delta h}{h_0}\right).$$

This describes pressure as a function of position given $\delta h(t)$. In our case, since the plates are instantaneously stepped together at $t = 0$, $\delta h(t)$ is a step function at $t = 0$.

By normalizing variables, converting to cylindrically-symmetric coordinates, and assuming a solution of the form $\delta p(r,t) = P(r)\exp(-\alpha t)$, we find an equation for $P(r)$ for $t > 0$ (after the step has been applied) of

$$\frac{h_0^2 P_a}{12\mu R^2}\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial}{\partial r}\right)\right]P(r) + \alpha P(r) = 0.$$

This is the Helmholtz equation. Since the problem is in cylindrical coordinates, the solutions are Bessel Functions. We use the form

$$P(r) = AJ_0(kr),$$

where J_0 is a zeroth order Bessel function. Plugging this solution into the equation gives

$$k^2 = \frac{12\mu R^2\alpha}{h_0^2 P_a}.$$

The complete solution will be a sum of Bessel functions which meet the boundary conditions $P(r = 1) = 0$ for all $t \geq 0$ (*i.e.* at the edge of the mesa the pressure above equilibrium is zero) and $P(t = 0^+) = \Pi(1)$ (*i.e.* the initial pressure profile is a Pi or hat function which drops to zero at the edge of the mesa). Each mode will decay at a different rate, with higher-order modes decaying more quickly. Figure 4-9 shows a plot of the first 15 terms in the Bessel series for the above problem plotted for $T = 0$

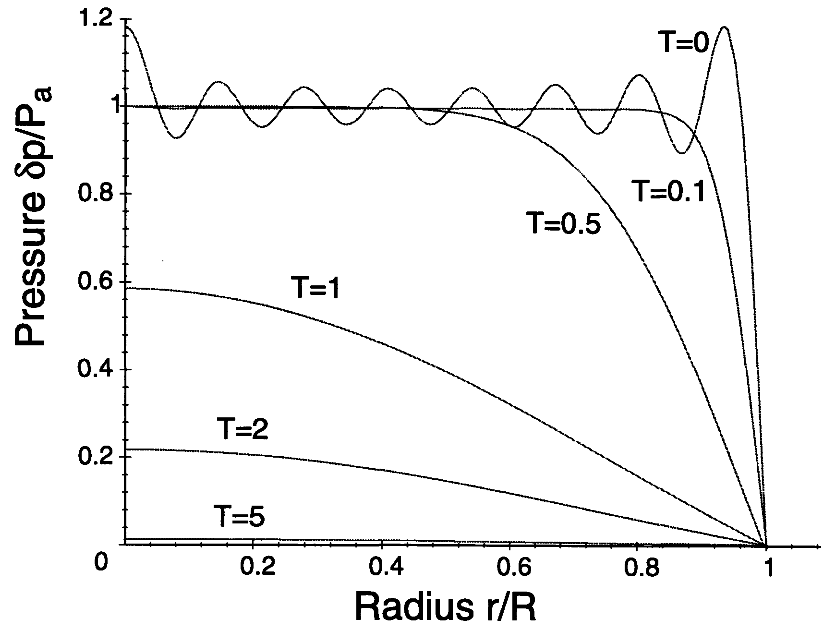


Figure 4-9: Plot of pressure profile as a function of time for simplified squeeze-film model. The first 15 terms in the Bessel function series are plotted. At $T = 0$ the pressure profile resembles a step function. As time progresses, higher-order modes decay, leaving only the lowest-order mode. T is normalized to the system time constant τ described in the text.

to 5, where T is normalized to the first mode decay:

$$T = \frac{k_0^2 h_0^2 P_a}{12\mu R^2} t = \frac{t}{\tau},$$

where k_0 is the first zero of the $J_0(k)$, which falls at $k_0 = 2.4048$.

For a 2 cm mesa and membrane radius and a one micron gap at atmospheric pressure, the time constant τ for this system is 1.5 seconds. The time constant τ scales as

$$\left(\frac{\text{radius}}{\text{gap}}\right)^2.$$

The real double-membrane system will respond to the decreased pressure at the edges by closing the gap further, causing the air to leak out more slowly. Yanof *et al.* undertook an analysis which modeled both the air flow and the membrane response [131]. An expression for the time constant of the system was found which

scales as

$$\frac{\text{radius}^4}{\text{gap}^3}.$$

Comparison between the model and experiments was quite good.

The equilibration time has been discussed above in the context of pulling masks into intimate contact for daughtering. It will also be important for any situation which results in a non-equilibrium pressure configuration, such as bringing an x-ray mask into proximity to a substrate for exposure. On a laboratory time scale which involves an 8-hour x-ray exposure, a one-second membrane response is insignificant. However, for those interested in using x-ray lithography in industry, that amount of settling time may be unacceptable.

The gap used in commercial x-ray lithography is much wider than that used in the NanoStructures Laboratory (typical gaps are 10-20 μm). However, the radius of the industrial mask/support configuration is also much larger (10 cm). Membrane dynamics could be important for the commercial use of x-ray lithography. Other papers have been published which investigate membrane dynamics theoretically and experimentally for motion perpendicular and parallel to the wafer [132–134].

4.4.2 Membrane Strength

It is evident from the photograph in Figure 4-5 that the vacuum soft contact methods developed for this work can subject x-ray mask membranes to rather severe stresses and strains. This can be cause for concern, particularly when one is considering subjecting a critical mask, which has required many days of processing by several people, to these extreme conditions. Anecdotal evidence indicates, however, that the membranes are able to stand up to the handling used in the methods described above, which were carefully designed to avoid shearing the membranes.

In the course of this work, the "suck and puff" daughtering method was used more than thirty times, with one mother mask being pulled into contact more than ten times. Small chips were pulled into contact many times, and in a test where the vacuum was increased to the maximum which is attainable with the present setup

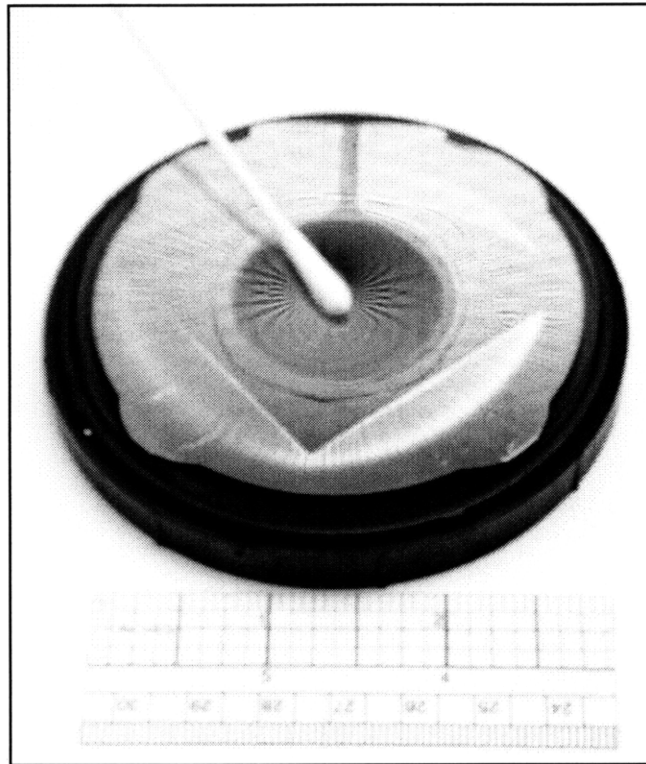
(about half an atmosphere) to try ascertain at which point the mask would break, the mask survived with no damage.

There were some masks broken – before the rigid daughtering fixture, three masks were broken while trying to separate the pair by pressure while holding the old fixture by hand. One mask broke in the small-chip fixture, but the place where it broke (as evidenced by the tear pattern of the membrane) was far removed from either the chip or the mesa edge. There were indications of an imperfection in the membrane near the area where it broke. One other mask was broken inexplicably while being pulled into contact with a wafer (in a modification of the daughtering process). This mask had a region of polyimide removed, and the breaking point was at the air bubble formed in that region.

In recent years, the SiN_x used for x-ray masks in the NanoStructures Laboratory has had defects in the form of particles embedded in the membrane. It is quite probable that these particles are responsible for many of the broken masks described above, and others which have occured (with the exceptions of the ones where microscope objectives are pushed through the membranes). If the particles are foreign objects around which the nitride has grown, the membrane can be modeled as a plate with a hole or an elastic occlusion. In either case, stresses can be increased by a factor of 3–4 in the vicinity of the hole [135]. Since the stress under loading at the center of a perfect circular membrane is one half of that at the edge, it is quite possible that this stress concentration could cause failure at the point of the defect.

Defect-free membranes, however, are quite strong. This can be demonstrated with the so-called "Q-tip test," which uses a wooden-sticked cotton swab and a mask (this test is not recommended for a mask which one requires in order to graduate, for instance). The cotton tip of the swab is pressed into the center of the membrane, causing it to bow by several millimeters. Typically the wooden stick of the swab will break before the mask membrane does *as long as no shear stress is applied by rotating the swab.*

This is pictured in Figure 4-10. In the top photograph, an x-ray mask membrane is shown with a cotton swab pressing down on it. The force was increased until the



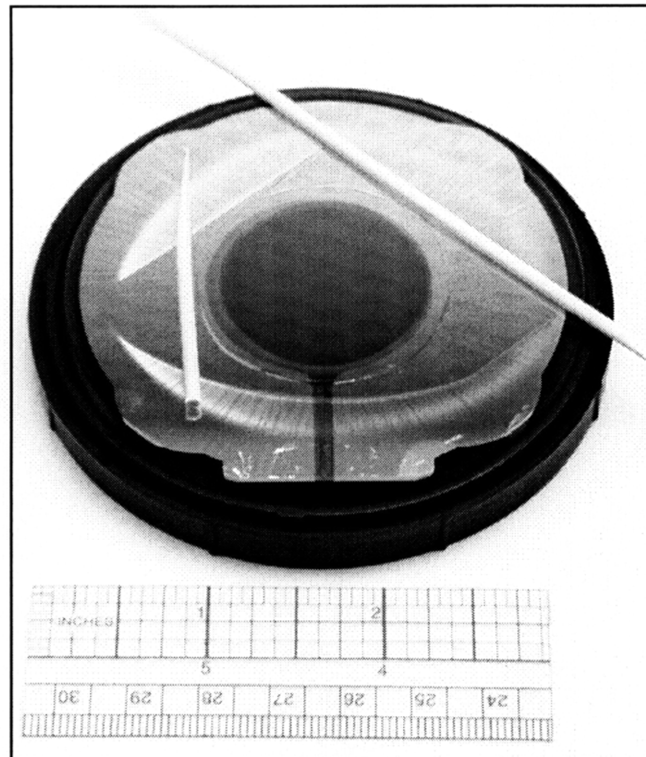


Figure 4-10: Photograph of membrane vs. wooden-sticked cotton swab. Top: the membrane is being deflected by the swab. Force is increased until the wooden stick breaks. Bottom: Victorious membrane with vanquished swab pieces.

stick broke in two. In the bottom photograph, the intact membrane is shown with the broken wooden stick. Although this is hardly a well-controlled quantitative test, it is nonetheless impressive to see how robust the masks can be.

Chapter 5

Substrate Photo- and Auger Electrons[1]

With the technologies presented up to this point, there is no apparent reason why it would not be possible to produce arbitrary features with sizes below 50 nm on a device substrate using Cu_L x-ray nanolithography. The x-ray source and mask technology described in Chapter 3 are certainly capable of providing the necessary contrast for proper lithographic input for sub-50 nm feature sizes. The soft contact exposure technologies described in Chapter 4 provide a means of gap reduction which reduces diffraction effects and penumbral blurring to a range where the resist dose should yield faithful pattern reproduction for feature sizes as small as ~15–20 nm.

For these exposures, photoelectrons and associated Auger electrons generated by x-rays should have a noticeable impact on the linewidth and affect the ultimate resolution. As discussed in Section 3.3.2, in a PMMA film, the image blur due to these electrons is approximately 3–5 nm. This amount of line broadening should be observable in the sub-50 nm regime.

The effect of electrons generated due to x-ray absorption *in the substrate* during an x-ray exposure have not yet been discussed. A substrate that absorbs x-rays more strongly than the resist will also generate more photoelectrons and Auger electrons

[1] Parts of this chapter were published in the *Journal of Vacuum Science and Technology B* [136].

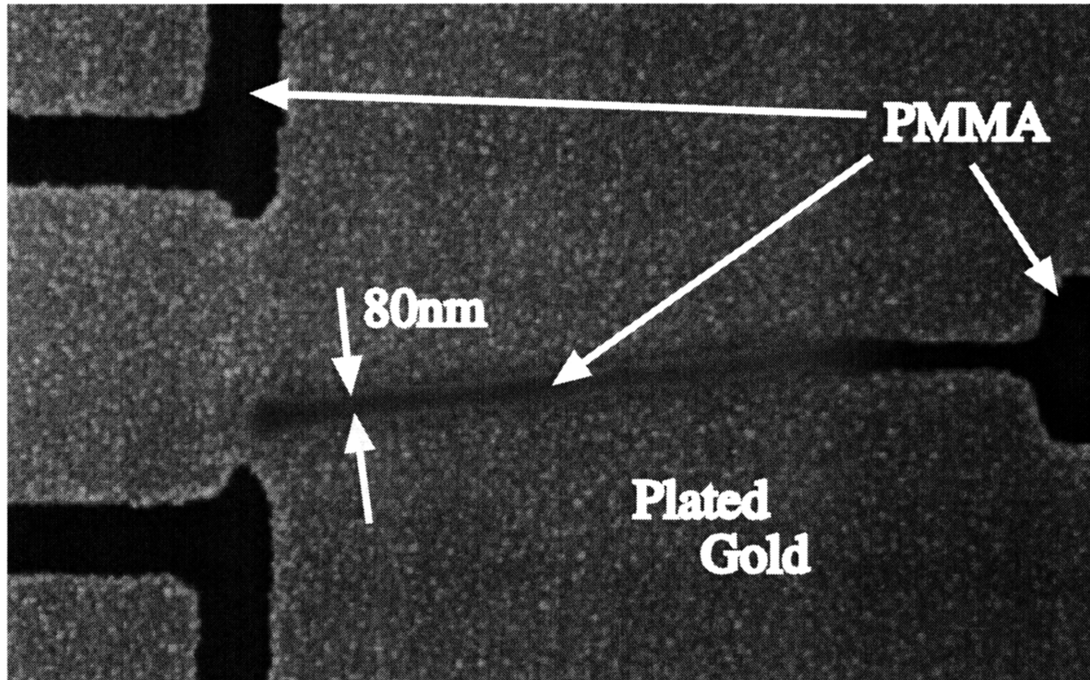


Figure 5-1: Scanning electron micrograph of patterned gold and PMMA on a daughter mask with "thick-gold" plating base following replication from a mother mask. The electroplated gold formed the correct pattern around the large features in the PMMA mold, which adhered to the substrate. The 80 nm-wide PMMA line, however, did not properly adhere and was dislodged by the gold plating.

per unit distance than the resist does. Substrate electrons which are generated near the resist-substrate interface can scatter into the resist and create a higher dose near the interface than would be present if the substrate were not there. Under certain circumstances this excess dose can create adhesion problems for very small resist features.

5.1 Unsuccessful Cu_L X-ray Daughter Exposures

The effect of substrate photo- and Auger electrons was evident in the daughtering process. The x-ray daughter mask has a thin plating base of 10 nm Ti/10 nm Au deposited on the mask membrane. PMMA is spun on top of this plating base. With this configuration of plating base and resist, it was not possible to print sub-50 nm lines from a mother mask onto a daughter mask. When the daughter mask was electroplated with gold after resist development, it was found that fine PMMA lines,

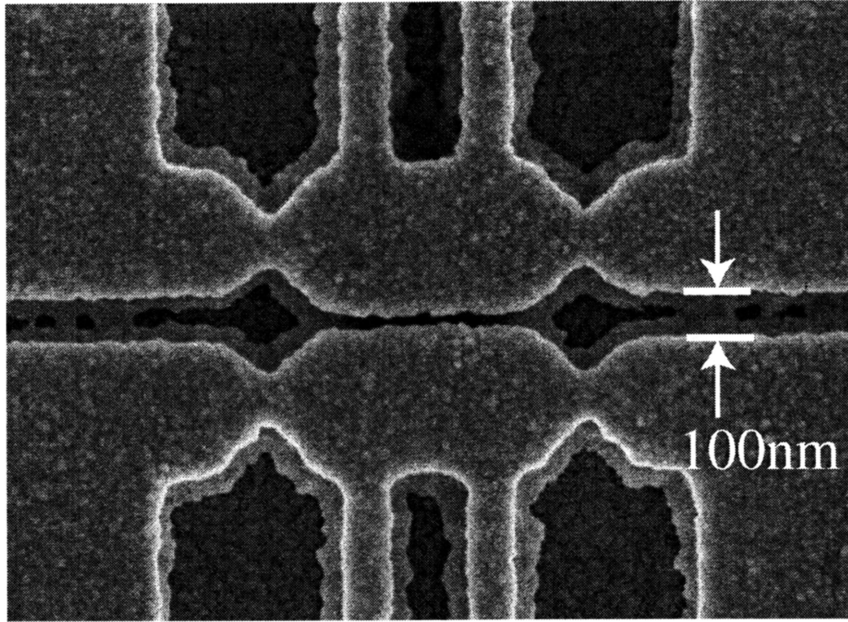


Figure 5-2: Scanning electron micrograph of patterned gold on a daughter mask with thick-gold plating base following replication from a mother mask. In this case the PMMA has been removed, and a small "foot" can be seen in the plated gold which extends about 50 nm into the areas which previously were covered with PMMA. For small PMMA features, two "feet" could meet and undercut the plating mold.

which were supposed to form the mold to create fine spaces in the gold, were not adhering to the plating base. With this particular plating base, referred to as "thick-gold," we were never able to get sub-50 nm lines to adhere. When the same exposure was done on a silicon substrate the PMMA had no adhesion problems.

Figure 5-1 shows a scanning electron micrograph of a daughter mask with thick-gold plating base which has been exposed, developed, and electroplated with gold. An 80 nm-wide PMMA feature can be seen rising up from the plated gold, while larger PMMA features adhered to the substrate and formed a proper mold for the electroplated gold. In this case, the pattern on the daughter mask is obviously a poor reproduction of that on the mother mask.

Some insight into the problem can be gained by examining a daughter mask with the PMMA removed, as in Figure 5-2. In this scanning electron micrograph a small "foot" of gold can be seen extending approximately 50 nm into the areas of the mask

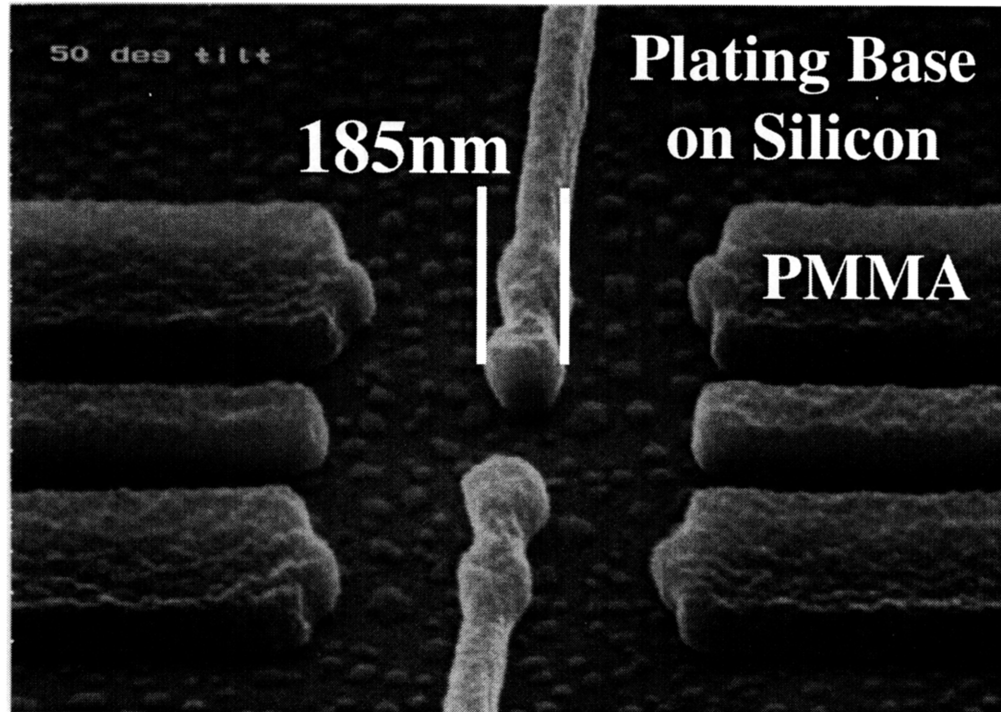


Figure 5-3: Scanning electron micrograph of daughter pattern in PMMA on silicon which is coated with thick-gold plating base. Rather than having straight sidewalls, the bottom of the PMMA is rather severely undercut, even though the development is incomplete (as indicated by the clumps of PMMA left in the exposed areas). Any gold plated onto this structure would show a foot, and small PMMA features would be completely undercut.

which were previously filled with PMMA. For large PMMA features the foot does not compromise adhesion, but for small (in this case ~100 nm) PMMA features the two gold feet from either side can meet and totally undercut the PMMA electroplating mold.

Figure 5-3 shows a scanning electron micrograph of developed PMMA on thick-gold plating base which was deposited onto a silicon wafer. Here the undercut at the resist-substrate interface is clearly evident, even though the development is incomplete (as evidenced by the clumps of PMMA in the exposed areas of the substrate). Even this amount of undercut would lead to a foot in plated gold, and could compromise very fine lines.

Several characteristics of this resist undercut and subsequent gold foot pointed to photo- and Auger electrons coming from the substrate as the cause. The undercut occurred near the resist-substrate interface. It was of a uniform extent everywhere for a particular exposure. It was also only problematic on high-atomic-number substrates.

5.2 Position Dependence of Dose in X-ray Lithography

As discussed in Section 3.3.2, x-ray absorption gives rise to photoelectrons and Auger electrons which are responsible for cutting polymer chains, thereby exposing the PMMA. The effect of electrons which are generated in and which travel through PMMA was described by a point spread function (PSF) which has an effective range of ~5 nm.

In a case where a substrate or other materials are present, however, the electrons which eventually expose the resist may be generated in and/or travel through the other materials. The PSF of the generated electrons depends on a combination of factors, including at what energies the electrons are generated (which is highly material-dependent), how they scatter in the various materials through which they travel, and the details of their absorption by the PMMA. This means that the PSF will depend on the depth into the resist, and will be different for each particular combination of substrate, resist, and x-ray energy spectrum. Nevertheless it can be calculated [137,138].

A simple physical picture can be developed to help guide understanding of the more complicated modeling. A schematic is shown in Figure 5-4. In the bulk of the PMMA, the effective range for generated electrons is ~5 nm. Electrons generated above and below a given point in the bulk of the resist will contribute equally to chain scissioning at that point. The dose through the bulk of the resist should be uniform. Near the surface of the resist, some of the generated electrons can escape from the PMMA film and therefore do not contribute dose to the resist. We should

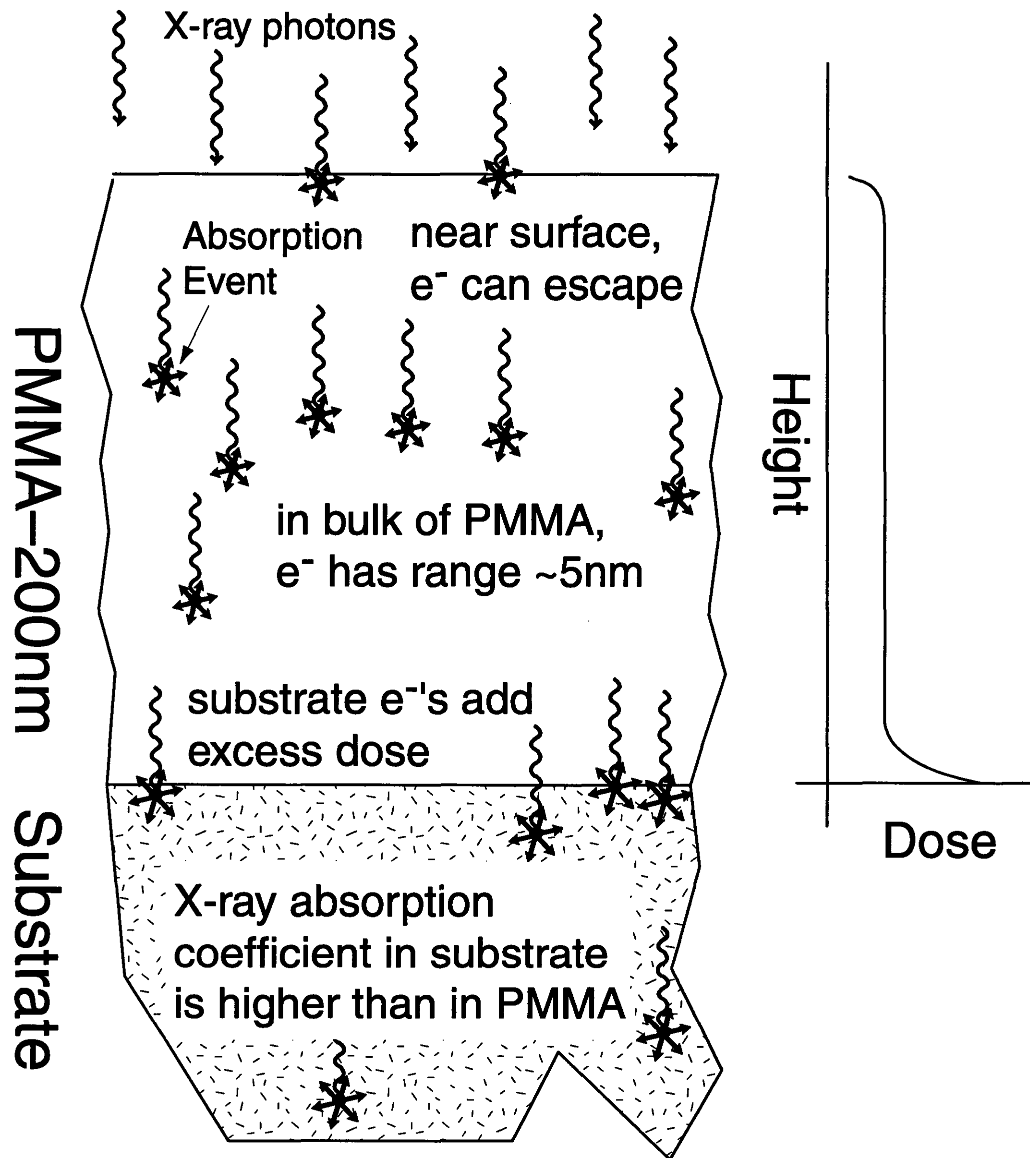


Figure 5-4: Schematic of x-ray absorption and photo- and Auger electron generation during x-ray exposure showing the effect of the substrate. In the bulk of the PMMA, the effective range of the generated electrons is ~5 nm. Near the surface, some of the generated electrons can escape, which should cause a decrease in dose. Higher x-ray absorption in the substrate generates more electrons there. This can create a higher dose in the resist near the resist-substrate interface.

therefore expect a lower dose near the surface.

Similarly, photoelectrons generated in a substrate can escape into the resist. If more x-rays are absorbed per unit length in the substrate (which is almost always the case), then we should expect an increase in electron generation in the substrate. Electrons generated near the resist-substrate interface will spill over into the resist and create a higher dose in the resist there.

At first glance this would not seem to be detrimental to the pattern replication process. In exposed areas, we should expect a decreased development rate near the surface, a uniform rate in the bulk of the resist, and an increased rate near the resist-substrate interface. By understanding the changes in development rate, successful pattern replication should be possible. However, the generated electrons can also compromise resist adhesion due to imperfect contrast in the "dark" regions and lateral spreading of electrons from the exposed regions to the "unexposed" regions.

5.3 Substrate Effects Reported in the Literature

The problem of photoelectron generation from a substrate was first noted by Maldonado *et al.* [139] in 1975 with a simulation of increased resist dose from substrate-generated electrons created by a very energetic Pd_L x-ray source ($\lambda \sim 4.4\,Å$). In 1978, Hundt and Tischer [140], also using hard x-rays (Ti_K, $\lambda \sim 2.75\,Å$) reported an increase in development rate of x-ray-exposed resist on a gold substrate over that on lower-atomic-number substrates. Since then numerous other experiments and simulations have focused on the effect the substrate can have on soft x-ray nanolithography [105, 108, 137, 138, 141–151].

In these references there are many variables, and accurate cross-comparisons are not possible. Different x-ray spectra are used for experiments and simulations. X-ray sources vary from electron-bombardment sources operated with various targets over a wide range of accelerating voltages, to synchrotrons operated at various energies, to plasma sources. For some simulations monochromatic x-rays are used. The several initial x-ray spectra pass through a variety of vacuum windows and masks, and some

spectra reflect off mirrors. Substrates of silicon, molybdenum, and gold, to name a few, are used. Resist thickness also varied. Despite these differences, however, two effects due to photoelectrons generated in the substrate are evident in both experiment and simulation. First, an increase in dose is seen above higher-atomic-number substrates. This excess dose is seen to have an extent anywhere from a few nanometers to a few hundred nanometers, although this range was not well-characterized experimentally and is attributed to substrate-generated electrons and photons. Second, an undercut is often seen in resist features near the resist-substrate interface.

Despite the many papers published on the subject of substrate photo- and Auger electrons, there had been no direct measurement of their effect. Experimental evidence was for the most part limited to descriptions of poor resist adhesion and SEM micrographs of resist undercut of the sort shown is Section 5.1. The one experimental paper which had numerical data only measured total development times for resist films.

5.4 Experiments and Simulations

5.4.1 Experimental Setup

Experiments were performed to measure the dose from substrate-generated photo- and Auger electrons in the resist as a function of height above the substrate. Substrates of thick-gold plating base, "thin-gold" plating base (10 nm Ti/1.8 nm Au), and bare silicon were coated with 250 nm of 950K PMMA and exposed in our x-ray system. The dissolution rate of PMMA in 2:3 MIBK:IPA depends on the absorbed energy per unit volume raised to the ~3rd power [100, 152]. This is the case irrespective of the type of ionizing radiation, making PMMA an excellent medium in which to record the effects of the photoelectrons generated in the substrate.

A half-plane exposure was desired. For these experiments the effect of partial transmission of the x-ray absorber was eliminated by masking the x-rays fully using a thick gold foil shadow mask placed on the the surface of the PMMA. This masking

was confirmed by measuring the PMMA erosion underneath the shadowed areas and finding it almost nonexistent on the time scale of the development of the exposed regions. Samples were exposed to yield a 120 second development time. This is only slightly longer than the standard development time of 60–90 seconds.

After exposure, the step height at the boundary between the exposed and unexposed areas was measured. A Digital Instruments Dimension 3000 AFM was used in tapping mode to avoid excessive contact between the tip and the sample. A $40\times5\,\mu\text{m}^2$ area was scanned with the step in the middle of the scan. The AFM yielded a histogram with two peaks associated with the exposed and unexposed areas. The step height was defined as the distance between the two peak maxima. Before development, the step height was typically found to be a few nanometers due to compaction of the exposed PMMA.[2] A screen shot of the AFM image and the histogram for a partially-developed step is shown in Figure 5-5.

The sample was then developed in 5 second increments in a 2:3 MIBK:IPA solution which was temperature controlled to $21°\text{C} \pm 0.1°\text{C}$. The step height was measured at approximately the same point (within about $50\,\mu\text{m}$) after each development interval. Step heights as a function of time were then used to calculate the PMMA dissolution rate as a function of height above the substrate.

5.4.2 Simulation Parameters

Simulations were done by Leo Ocola at the University of Wisconsin[3] using the Monte Carlo code *Low-energy Electron Scattering in Solids* (LESiS), which simulates x-ray absorption and subsequent photoelectron and Auger-electron generation to yield dose as a function of position [102, 138]. The spectrum used for the simulations was the spectrum incident upon the resist after passing through the vacuum window, the mask membrane, and plating base, as calculated in Chapter 3. For the purposes of the simulation, the spectrum was binned into 40 energy ranges. The number of

[2]The exposure breaks the long polymer chains, allowing some material to exit the resist in gaseous form during the exposure. Since there is less material in the exposed regions, some latent image is seen in the resist before development.

[3]Now at Bell Laboratories, Lucent Technologies.

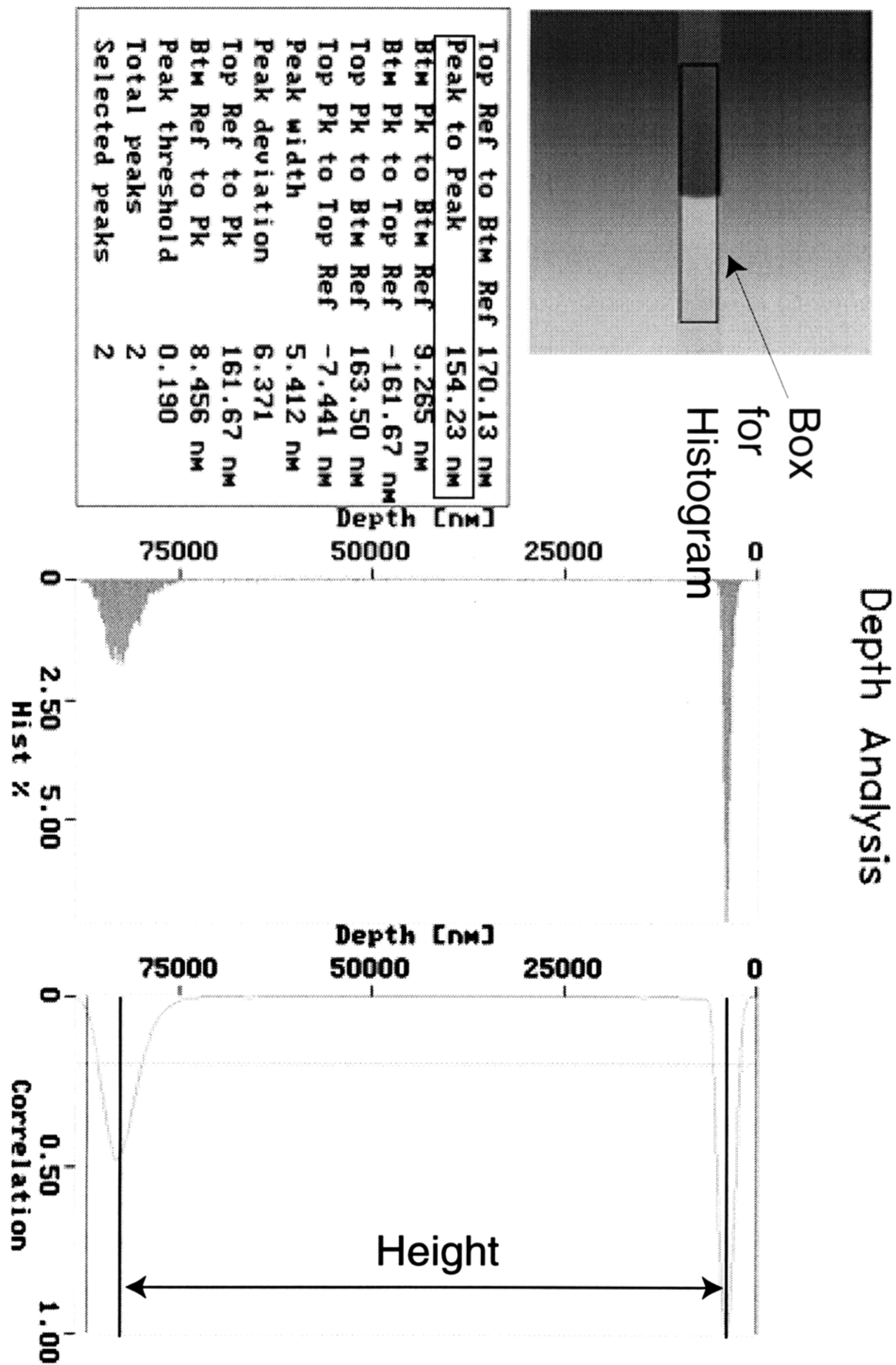


Figure 5-5: Screen shot of AFM image and height histogram for a partially-developed step. The image is shown in the upper right. The box indicates the area for the height histogram. The histogram is shown at the bottom (the scale numbers are not in nanometers – there is a software bug). The measured peak-to-peak height is given in the table on the upper left.

photons incident on the sample was calculated for each energy range for a 5-hour exposure, and the complete set of photons was simulated. The dose generated by this simulation was converted to PMMA molecular weight, which was then used to simulate dissolution rate as a function of position. There was one fitting parameter used to make the total development time equal to that found in experiment.

Photons generated by fluorescence emission were not simulated. For our substrates and our spectrum, the conversion efficiency for fluorescence is less than 1%. Also, the simulation was done only for energies up to 5 keV. X-ray photons with energies in the range 5-8 keV account for only about 1% of the total energy absorbed in the PMMA, so they should also have little effect on the dissolution rate. By comparison of simulations, it was determined that secondary electrons generated by the photoelectrons and Auger electrons also had little effect on the simulation ($\leq 2\%$), so to increase simulation efficiency, these were also neglected.

The simulated dissolution rate as a function of height above the substrate was then integrated in 5-second increments and used to calculate a dissolution rate in the same manner that the actual dose in the PMMA was developed and measured. A comparison of the experimental results and simulations follows.

5.4.3 Results

Figure 5-6 plots the measured PMMA dissolution rate versus height above the substrate for PMMA on substrates of thick-gold plating base on silicon, thin-gold plating base on silicon, and bare silicon. The most notable feature of the data is the dramatic increase in dissolution rate near the thick-gold plating base, which is not seen for the silicon or thin-gold-plating-base substrates.

The data show that the thick-gold plating base generates a significantly higher dose near the PMMA/substrate interface than bare silicon or thin-gold plating base, and that the range of this increased dose is approximately 50 nm. All three systems exhibit a drop-off in dose in a range of about 25 nm near the surface of the PMMA, which gives an indication of the escape depth of the electrons generated in the resist.

The difference between the thin-gold and the thick-gold results indicates that the

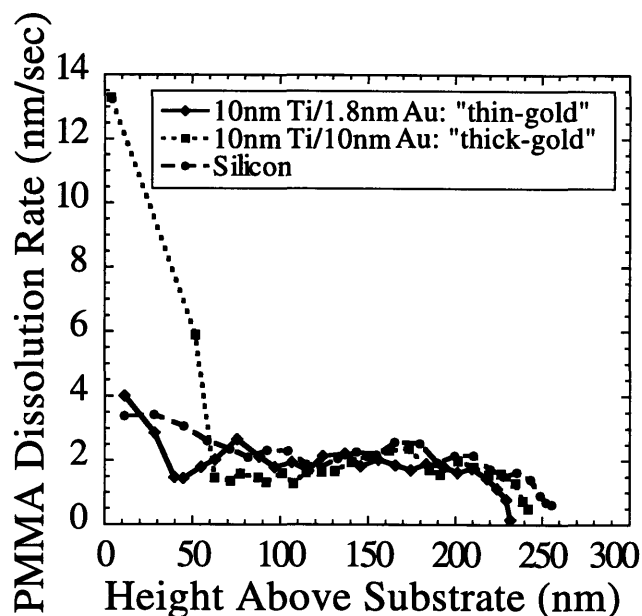


Figure 5-6: Measured PMMA dissolution rate versus height above the substrate for substrates of thick-gold plating base on silicon, thin-gold plating base on silicon, and bare silicon. A dramatic increase in dissolution rate near the thick-gold plating base contrasts with only a slight increase near the silicon and thin-gold-plating-base substrates. A decrease in the PMMA dissolution rate near the top surface of the PMMA is seen for all three cases.

effective range of electrons generated in the gold is greater than 1.8 nm. If this were not the case, increasing the thickness should have no effect, since electrons generated further from the interface than their effective range in gold will not escape into the PMMA. The Gruhn range for electrons [153] in a material is proportional to the inverse of its density. With $1.3\,g/cm^3$ as the density for PMMA and $19.3\,g/cm^3$ as the density for gold, the Gruhn range should differ by a factor of 14.8 in the two materials. Since we observed a range of ~50 nm above the substrate interface in the PMMA, we should expect a range of approximately 3.4 nm in the gold. This simplistic calculation matches well with our experimental observation.

Figure 5-7(a) shows a Monte-Carlo simulation of the data presented in Figure 5-6. The same trends are clearly evident, and both the magnitude and the range of the effect are in excellent agreement with the experimental data.

The success of the comparison between the experimentally measured data and the simulations lends credibility to the use of the simulations as a tool for examining

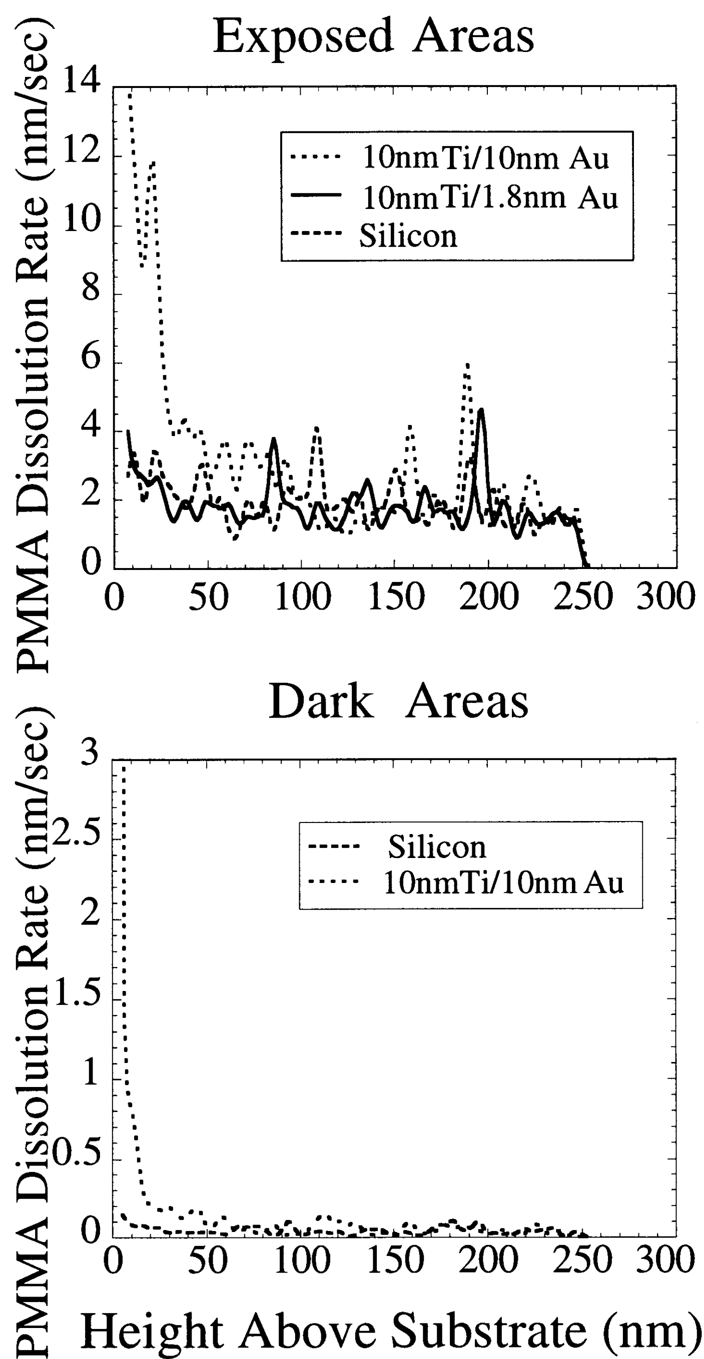


Figure 5-7: (a) Simulation of the exposures measured in Figure 5-6. The same trends are clearly evident, and both the magnitude and the range of the effect are in excellent agreement with the experimental data. (b) Simulation using the spectrum for "dark" areas, *i.e.* in the shadow of the 200 nm-thick gold absorber. For the thick-gold plating base, an increase is seen near the plating base where the dissolution rate approaches the bulk development rate in the exposed regions of the PMMA.

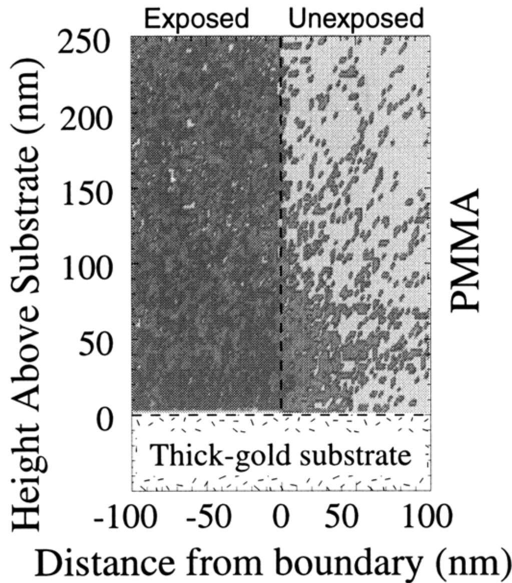


Figure 5-8: Two-dimensional simulation of a half-plane x-ray exposure on thick-gold plating base. The left half plane received x-ray exposure, while the right half plane received none. A gray-scale plot of the logarithm of the dissolution rate is shown, with darker areas corresponding to higher dissolution rate. Note that even though there were *no* x-rays present on the right half-plane, a bump appears in the dose profile due to lateral scattering of photoelectrons near the substrate from the exposed area.

situations which cannot be experimentally measured. Figure 5-7(b) shows the same simulation for a dose and spectrum corresponding to the "dark" region, *i.e.*, in the shadow of a 200 nm-thick gold absorber. Here we see that for both the silicon and the thick-gold plating base substrates the dissolution rate of the PMMA at the top surface and in the bulk of the film is very low. However, for the thick-gold plating base an increase in dissolution rate is seen near the plating base. In fact, near this interface the development rate approaches the bulk development rate in the exposed regions. Hence, overdevelopment would lead to undercutting of the PMMA and possible loss of small features.

So far, we have discussed only the "one-dimensional" effect of increased dose at the resist-substrate interface, which can lead to an undercut due to partial exposure in the shadow of the x-ray absorber. One might think that a way around the latter problem would be to increase the absorber attenuation or modify the x-ray spectrum to minimize high-energy components which are not well-attenuated by absorber ma-

terials. This would suppress the increased rate at the resist-substrate interface in the dark areas. However, there would still be a potential undercut problem arising from the fact that the substrate-generated electrons travel laterally as well as vertically in the resist. That is, substrate-generated electrons from open areas of the pattern can scatter into shadowed areas, causing an undercut.

Figure 5-8 shows a gray-scale plot of the logarithm of the PMMA dissolution rate for a two-dimensional simulation of a half-plane x-ray exposure, with diffraction effects ignored. Darker regions correspond to higher dissolution rate. The left half-plane corresponds to an exposure of PMMA on thick-gold plating base with the spectrum used in our experiments. There is no exposure at all in the right half plane, simulating a fully-attenuating absorber. However, we still see an increased development rate near the boundary between the exposed and unexposed regions in the vicinity of the substrate.

The presence of this undercut is easily understood as an extension of the one-dimensional effect which we measured. Even in the bulk of the PMMA, there is some spillover of electrons from the exposed to the unexposed area. This creates some lateral blur. Since there is an increased dose near the PMMA/gold interface, the spillover into the unexposed areas is also increased near the substrate, giving rise to the undercut.

5.5 Successful Daughter Mask Replication

The slight increase in dissolution rate at the resist-substrate interface seen for the thin-gold plating base (comparable to that of silicon) points to an obvious solution to the problems encountered in replicating mother masks onto daughter masks described in Section 5.1. Before that solution was discovered, however some other means were tried to eliminate the foot.

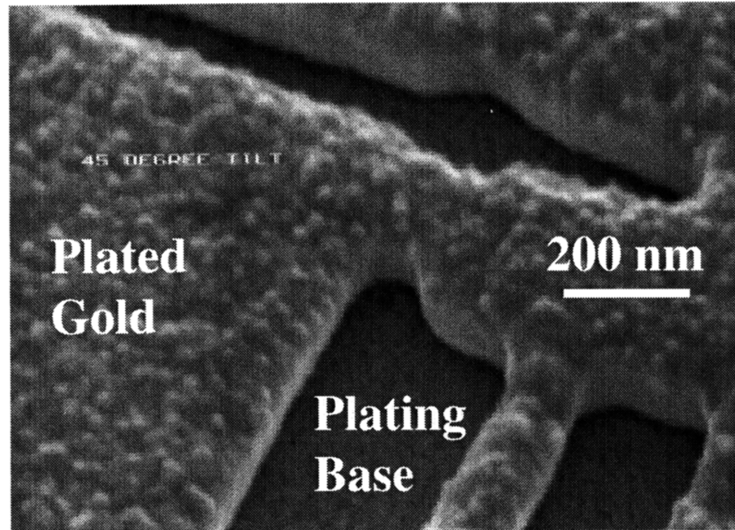


Figure 5-9: Scanning electron micrograph of a plated daughter exposure on thick-gold plating base coated with a thin layer of polyimide and PMMA. The polyimide layer (which has been removed) was 25 nm thick. No foot in the plated gold is seen.

Unsuccessful (But Promising) Attempts at Daughtering

Because the foot occured in the first few nanometers above the plating base, a 25 nm layer of polyimide was spun between the thick-gold plating base and the PMMA. With this configuration, the photoelectrons generated in the plating base should deposit most of their energy in the polyimide layer rather than in the PMMA and therefore the PMMA profile should not be undercut. After developing the PMMA, the polyimide was etched in the reactive ion etcher with a combination of oxygen and helium.

The results were quite promising and are shown in Figure 5-9. This scanning electron micrograph shows a device pattern in plated gold (150 nm thick) on a substrate which had a 25 nm layer of polyimide between the plating base and the 250 nm-thick PMMA layer. No foot is seen in the gold, and the pattern replication was quite good at the substrate. However, the oxygen RIE attacked the PMMA more quickly than the polyimide, causing rounding and erosion of the PMMA. The success of this experiment in eradicating the gold foot pointed to the substrate photo- and Auger electrons as the culprit, but the degraded PMMA did not give the proper sidewall profile in 200 nm-thick plated gold.

In order to get around the PMMA erosion problem, a trilayer process of anti-reflection coating (which was PMMA-based), SiO_2, and PMMA was also attempted with some success, but the difficulties encountered in optimizing the process made it clear that a simpler process would be preferable.

Other plating bases were also considered, and exposures and plating on nickel were attempted. The absorption of the incident x-rays from our exposure system by nickel is several times smaller than that of gold (although at the Cu_L energy they are approximately equal). The gold film appeared to plate well, but it washed away upon rinsing with water. This was probably due to oxides on the nickel surface which prevented gold adhesion. Although there are possible ways around this problem, usually involving etching the nickel surface before dipping the sample into the gold plating bath, none have been successful with our BDT-510 gold plating bath [154].

Success with Thin-Gold Plating Base

As the data presented in Section 5.4.3 suggests, modifying the plating base to a thin-gold configuration (10 nm Ti/1.8 nm Au) was the solution chosen to overcome the loss of resist adhesion due to substrate photo- and Auger electrons.

Figure 5-10 shows an example of a grid pattern fabricated on a daughter mask which had thin-gold plating base. 25 nm-wide spaces were successfully replicated. The PMMA clearly formed a proper mold for the electroplated gold with no undercut.

Figure 5-11 shows the coupled quantum dot device pattern with a 20 nm fine-line gate. In this case the width of the line was smaller than that on the mother mask due to overdevelopment. The fact that the PMMA held up without undercutting even in the case of overdevelopment clearly indicates that the thin-gold plating base is a robust solution to the problem for fine features on daughter masks.

The previous two examples showed very fine features which were successfully replicated onto daughter masks from mother masks. Because of their geometry, however, it is possible that these fine features were completely undercut but were supported by larger features to which they were attached. If this were the case, it would be likely that the plating-base gold in the fine features would be thicker than that in

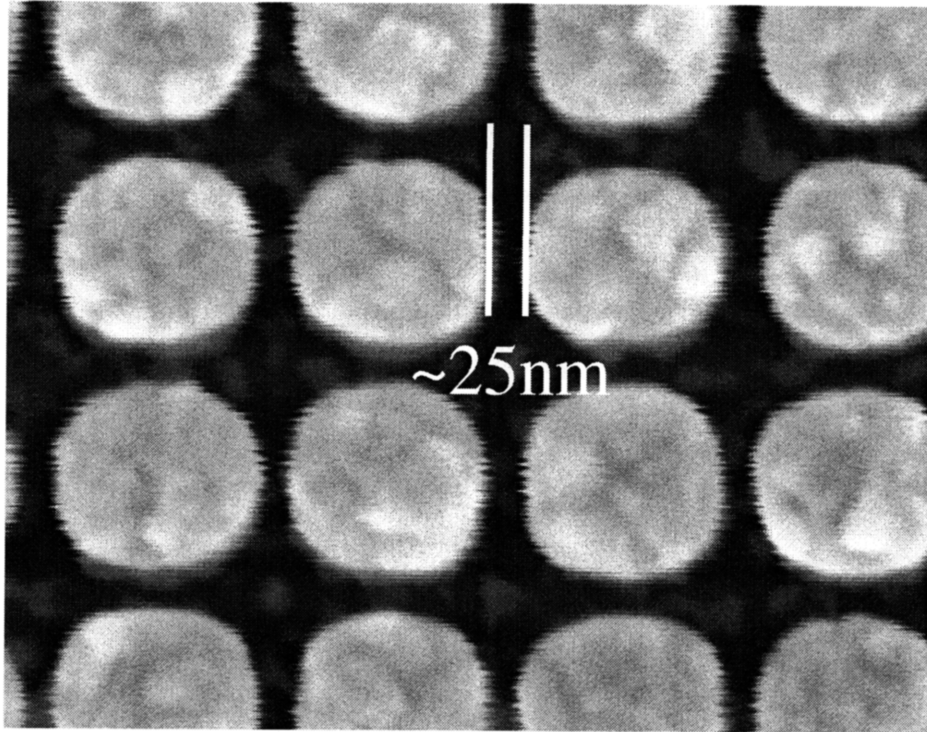


Figure 5-10: Scanning electron micrograph of a successfully-plated daughter with ~25 nm feature sizes done on thin-gold plating base. Figure courtesy of Mark Schweizer.

the larger areas, and that the thickness would depend on the details of the geometry. Although such a process would allow some devices to be made, it would not suit the requirement of a robust process for arbitrary patterns.

Figure 5-12 shows an example of an astigmatism test pattern on a daughter mask where on one end the fine features are not attached to larger structures. In this case, one micron-long, ~40–45 nm-wide PMMA lines acted as proper molds for gold electroplating. If the narrow PMMA lines had been undercut, with no support at the outside end, they would certainly have fallen over.

These examples clearly demonstrate that the thin-gold plating base prevents undercutting of fine features in the daughtering process. With the daughtering component of the process successfully controlled, the goal of a robust, "manufacturable" process for arbitrary features in the sub-50 nm regime has been met.

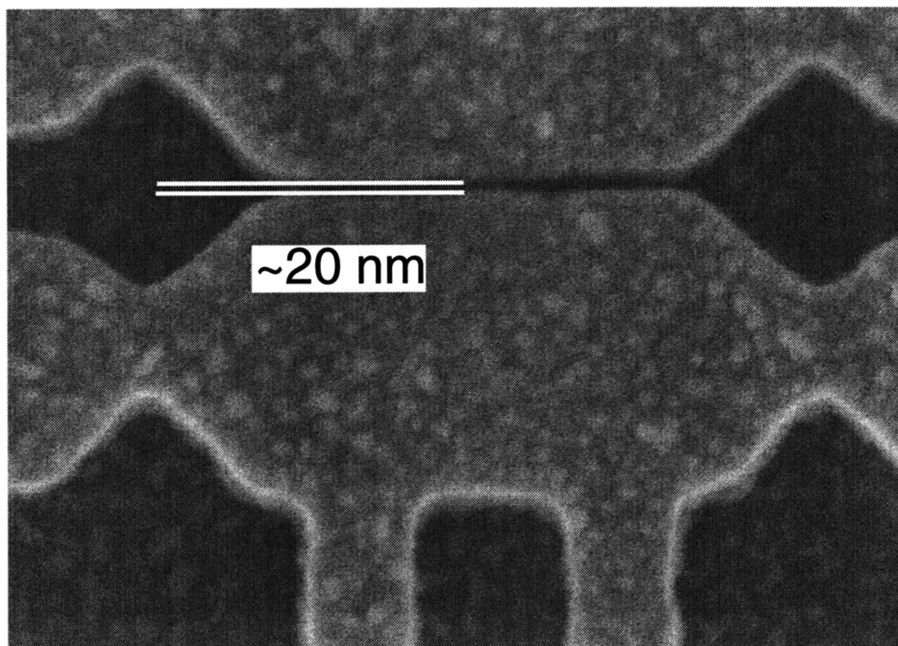


Figure 5-11: Scanning electron micrograph of a successfully plated daughter mask with a ~20 nm line, thinner than on the mother mask. The very narrow resist feature was not undercut even in this case of overdevelopment.

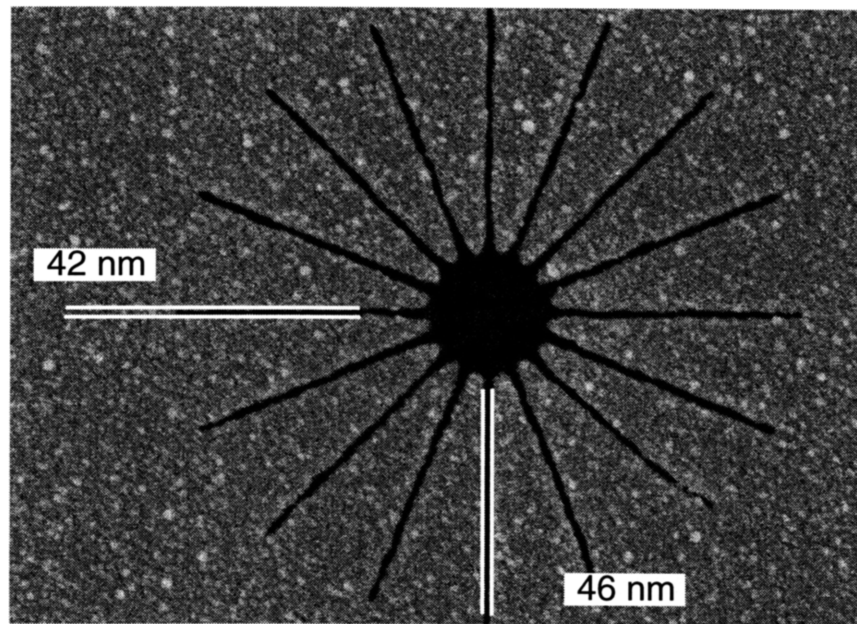


Figure 5-12: Scanning electron micrograph of a sub-50 nm starburst pattern on a daughter mask. In this case the PMMA mold had no larger features which could have supported the fine lines. If the lines were completely undercut, they would certainly have fallen over.

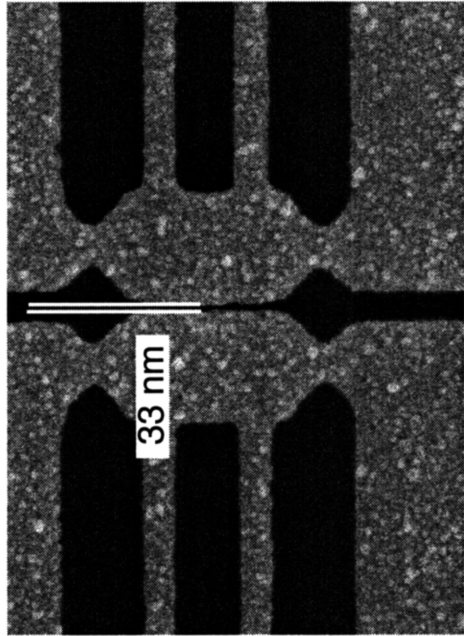
Chapter 6

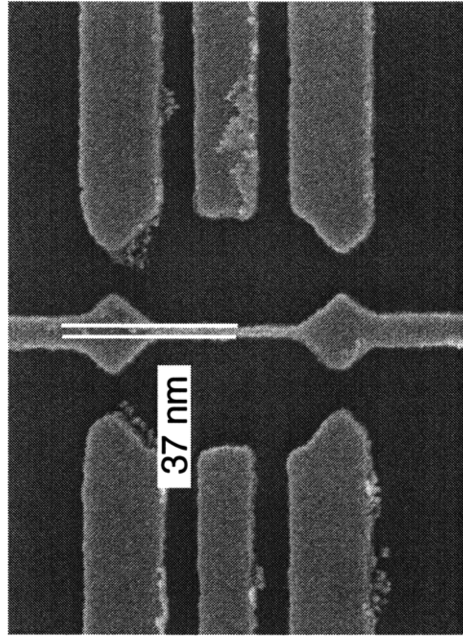
Sub-50 nm Exposure Results

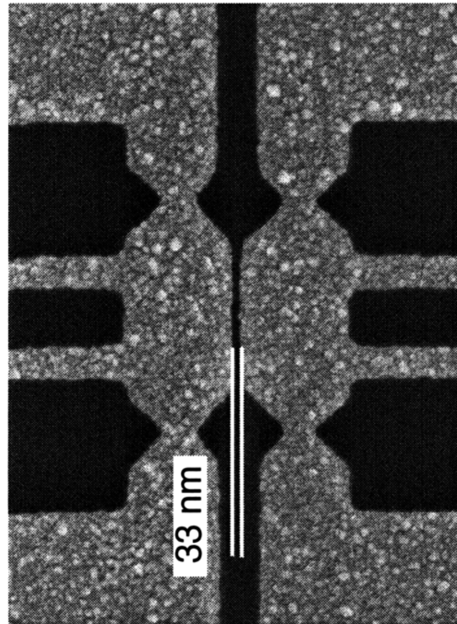
The ability to produce sub-50 nm features on a substrate had previously been limited by two factors: the inability to replicate such features onto a daughter mask because of the detrimental effects of substrate photoelectrons, and the inability to reliably control the gap during exposure. The limitations imposed by these two factors have now been curtailed. The ~300 nm gap which results from the soft contact technology described in Chapter 4 creates exposure conditions where diffraction is no longer the limiting factor for feature sizes larger than ~20 nm. The "thin-gold" plating base discussed in the previous chapter allows for replication of these feature sizes onto daughter masks, making polarity reversal possible. Arbitrary feature geometries with dimensions below 50 nm can now be printed onto substrates.

6.1 Individual Sub-50 nm Replication Results

Many different patterns with feature sizes as small as ~30 nm have been replicated using the techniques described in Chapters 4 and 5. A successful replication and liftoff of the coupled quantum dot device gate pattern is shown in Figure 6-1. The left half of the figure shows scanning electron micrographs of device patterns on the daughter mask. The right half of the figure shows the metallic gate patterns resulting from an x-ray exposure onto a substrate and subsequent liftoff of a metal (Ti/Au) film. The linewidth replication from mask to device is quite good, even at these ultra-fine







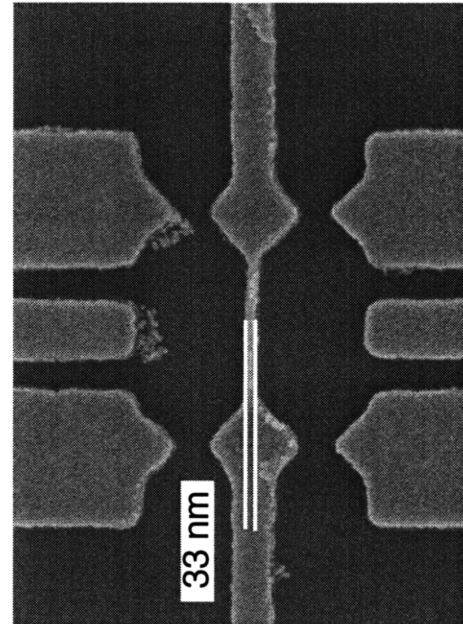
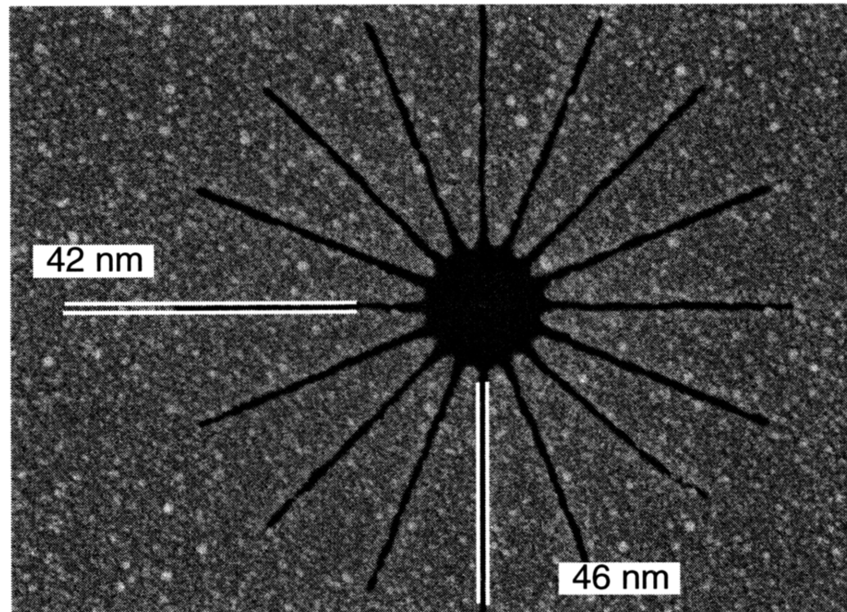
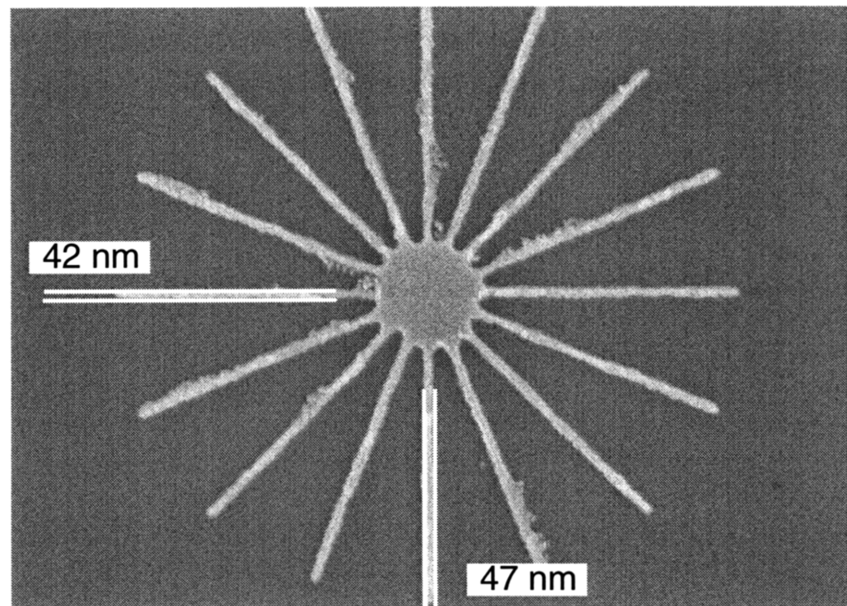


Figure 6-1: Scanning electron micrographs of device pattern replication by liftoff onto a substrate with sub-40 nm linewidths. Left: Images of the device structures on the x-ray daughter mask. Right: Images of metal (Ti/Au) gate patterns on a substrate which were formed by soft contact x-ray lithography and liftoff.





/Users/davidc/Figs/Archive/hank/xray-star.ai

Figure 6-2: Scanning electron micrograph of a sub-50 nm test pattern replicated by liftoff onto a substrate. Top: Image of the structure on the x-ray daughter mask. Bottom: Image of metal (Ti/Au) pattern on a substrate which was formed by soft contact x-ray lithography and liftoff.

linewidths.

Figure 6-2 shows a liftoff replication of the astigmatism test pattern which was previously shown in Figure 5-12. The daughter mask is shown at the top and the metal pattern on the substrate is shown on the bottom of the figure. Once again, the pattern was faithfully reproduced.

These images show that reproduction of features as fine as ~30 nm is possible, and that the fidelity of the replication is at least qualitatively good. However, measuring the linewidths of a few of exposures does not give statistical data about the process latitude – sensitivity to changes in dose and development time. In addition, the calibration of the measurement on the scanning electron microscope may be inaccurate due to image drift or changes of scale with focus.

6.2 Statistical Analysis of Sub-50 nm Replication

6.2.1 Description of Experiment

The problems of SEM scale calibration and the small number of linewidth measurements possible on the coupled quantum dot device were addressed with a set of experiments. The basis for these experiments was a test structure which was created with feature sizes ranging from well below to well above 50 nm. Measurements on this structure can be self-calibrated, and they can yield a large number of linewidth measurements to allow for statistical analysis.

Test Structure/Mother Mask

A new mask was designed with cells containing single-pass-line grating structures which varied in period from 40 to 500 nm. Gratings oriented in both the horizontal (X) and vertical (Y) direction were written. Half of each grating overlapped its orthogonal partner grating to create a grid structure. The cell design is shown in Figure 6-3. This cell was repeated 12 times on a die, each at a different single-pass-line electron dose.

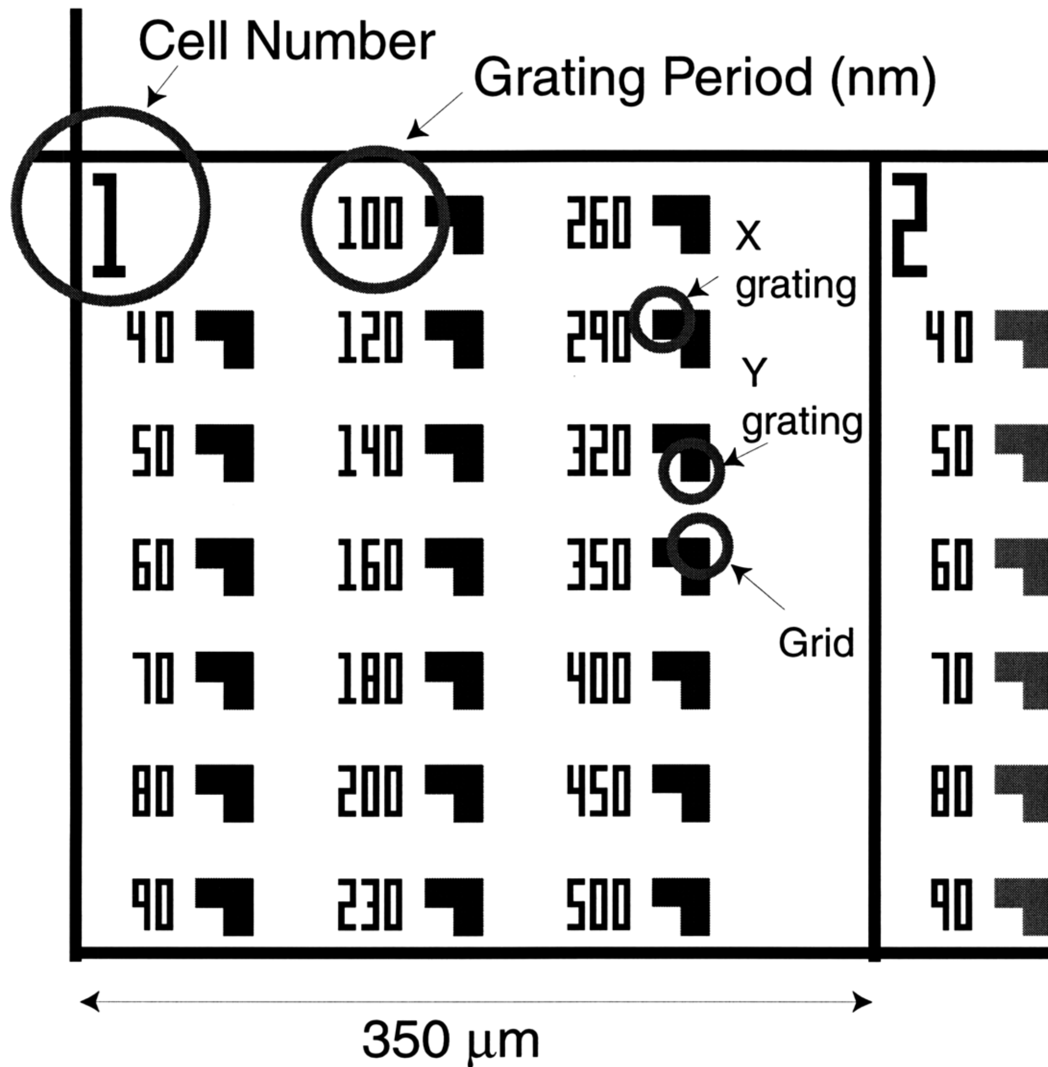


Figure 6-3: One cell of the mask design for sub-50 nm exposure experiments. Each inverted L-shaped structure consists of an X-oriented and a Y-oriented grating of single-pass lines which overlap to form a grid in the upper right corner. Periods from 40 nm to 500 nm were written in each cell. The cell is repeated 12 times on each die, each time with a different single-pass line dose. There are eight dies on the mask.

The mask was electron-beam written at the Naval Research Laboratory in Washington, DC. The total writing time for this mask was 31 hours, with most of that time dedicated to writing single-pass lines. After development and electroplating, many of the finest-period gratings were washed out due to proximity effects in the electron beam writing, as was expected. However, the larger-period gratings did plate successfully. The mother mask had linewidths ranging from ~30–100 nm.

By creating a grating test structure, the problem of SEM scale calibration was overcome. Since each grating was written within a single field of the electron-beam writer, the accuracy of the grating *period* should be quite good – certainly to well within one percent, or within ~600 pixels in a 16-bit field. Since the period of the test structure is known to a high precision, the period of an image of the test structure can be used to set the length scale, and subsequently the pixel scale (nm/pixel) of that image.

In addition, a grating test structure enables several linewidth measurements to be made from a single SEM image. We define a "linewidth" as the average width of a line over a length which is some small number of times longer than the width. A single SEM image of 4 to 6 periods of a grating can yield 50 to 100 "linewidth" measurements in this manner.

Linewidth Measurement Algorithm

The masks and liftoff replications printed from them were viewed in the scanning electron microscope. Images of the gratings were stored in TIFF format. Linewidths of the gratings were then measured from the TIFF images using the program `extraxt` written in LabView on a power Macintosh computer. A screen shot of `extraxt` running on a 140 nm-period grating is shown in Figure 6-4. The algorithm used to perform the measurement is described below.

The user specifies the image input file and the grating period, its orientation, and whether peaks (for mother masks and liftoff patterns) or valleys (for daughter masks) are of primary interest. A power spectrum is taken of each line of the image, and the pixel scale and a standard deviation (SD) are calculated by taking the average

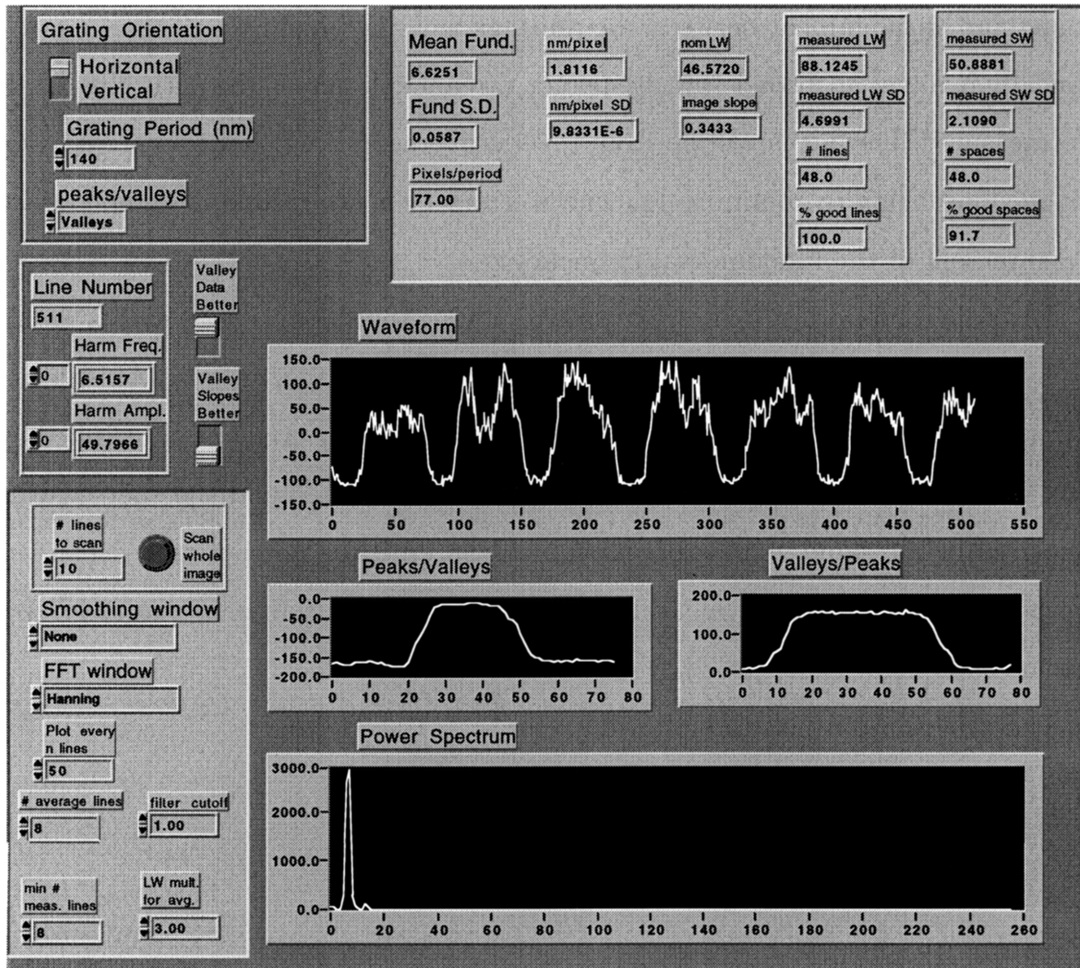


Figure 6-4: Screen shot of the program **extraxt** used to extract linewidths from a TIFF image of a grating. Inputs (grating period and orientation and whether signal peaks or valleys are of interest) are in the upper left corner. Parameters relating to the linewidth measurement are in the lower left corner. The top right shows the outputs (line/space widths with standard deviations (SDs), fundamental frequency and pixel scale with SDs, nominal linewidth (multiplied to set the averaging length for the more precise linewidth measurement), image slope, and number and percentage of good line/space measurements). Below this information, one line of the raw waveform is seen. Below that, the averaged waveforms used to measure (in this case) valleys (left) and peaks (right) are plotted. The bottom graph shows the power spectrum of the raw waveform.

of the fundamental frequencies of all the lines. Typically the SD is about 1% of the fundamental frequency.

The position of the first peak or valley of each line is used to remove any slope or angular misalignment between the gratings and the pixel axes of the image. This typically involves at most a shift by two or three pixels across the entire image. Since the measured linewidths will be from several averaged rows, however, removing any image slope helps to minimize linewidth errors (*n.b.* the slope does not affect the measurement of the fundamental frequency).

A nominal linewidth, measured while removing the slope, is used with a multiplier set by the user to average the now-properly-aligned image. The full width at half maximum of each peak in the new, averaged image is measured and multiplied by the pixel scale to yield a linewidth. The image is then inverted and the spacewidths are measured in the same manner. If a peak measurement is not perfectly clean (*e.g.* if two peaks are found in the measurement window rather than only one as expected), that measurement is discarded.

The program outputs the following:

- Fundamental frequency and SD
- Pixel scale (nm/pixel) and SD
- Nominal linewidth and image slope (angular misalignment of original image)
- Measured linewidth SD
- Measured spacewidth and SD
- Number and percentage of good line/space measurements

The measurement program is quite efficient. An SEM image takes about 8 seconds to analyze on a 100 MHz Power Macintosh 8100. A series of measurements from one die can be performed in about 5-10 minutes.

6.2.2 Linewidth Measurement Results

The test structure and linewidth measurement algorithm described above were used to investigate linewidth replication in the sub-50 nm regime. A daughter mask was made of the mother mask, and the daughter mask was used to expose wafers coated with PMMA resist. The exposed wafers were then developed, and metal was deposited and lifted off.

Measuring linewidth in the scanning-electron microscope can be tricky, and comparing different linewidth measurements introduces further complications. Understanding exactly how feature sizes change from a mask (where the relevant features are spaces) to a substrate (where the relevant features are lines) becomes even more difficult. The following sections describe the data and the techniques used for analysis, and show that linewidth replication down to ~30 nm using Cu_L x-ray nanolithography can be done with excellent repeatability and a wide process latitude.

Individual Image Measurements

This experiment attempts to compare linewidth measurements to investigate how they change with development or exposure time for differing feature sizes. Any comparison of linewidth measurements is critically dependent on the accuracy of the individual linewidth measurements being compared. This section presents results of measurements on individual gratings and shows that typical measurements can yield statistical data with a standard deviation (SD) of about 2.5-3.5 nm.

The first test of the linewidth measurement technique was to compare the automated algorithm with the tried-and-perhaps-not-true method of measuring linewidths using the software routine in the SEM. Comparisons of measurements made by hand with those made by the program `extraxt` for a few grating images yielded very consistent results, with linewidths and SDs agreeing to within a few percent.

The SD of the measurement gives information about its variability. Any variability in the measurement is due to a combination of actual changes in linewidth of the feature and artifacts introduced by the measurement technique.

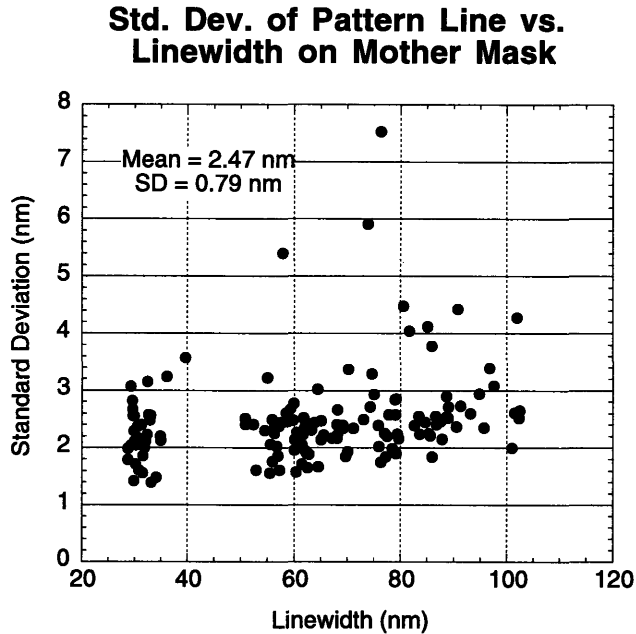


Figure 6-5: Plot of standard deviation (SD) of linewidth measurement on mother mask versus measured linewidth. The average SD of a measurement is about 2.5 nm, with a standard deviation (of the SD) of about 0.8 nm. There is no apparent trend of the SD with linewidth.

Standard deviations for many grating structures on the mother mask are plotted versus measured linewidth in Figure 6-5. There is no apparent linewidth dependence of the measurement SD. The mean SD of all measurements is about 2.5 nm, and the consistency between measurements is quite good, as indicated by the standard deviation of the SDs of about 0.8 nm. Although not shown in the plot, there is no significant difference between the SDs of horizontally- and vertically-oriented gratings.

When measurements are performed on a daughter mask which was replicated from the mother mask, SDs for measurements of feature sizes larger than ~45 nm are similar to those on the mother mask. Smaller features (spaces on the daughter mask) have a slightly larger SD. A plot of the average SD for daughter mask measurements is shown in Figure 6-6. In the plot, all linewidth measurements from 30-40 nm were averaged, as were measurements from 40-50 nm, etc. Whether the increase for the smallest feature sizes is an artifact of the measurement or a real effect is unclear. However, the daughter mask measurements are still "good" to within ~3.5 nm in the worst case.

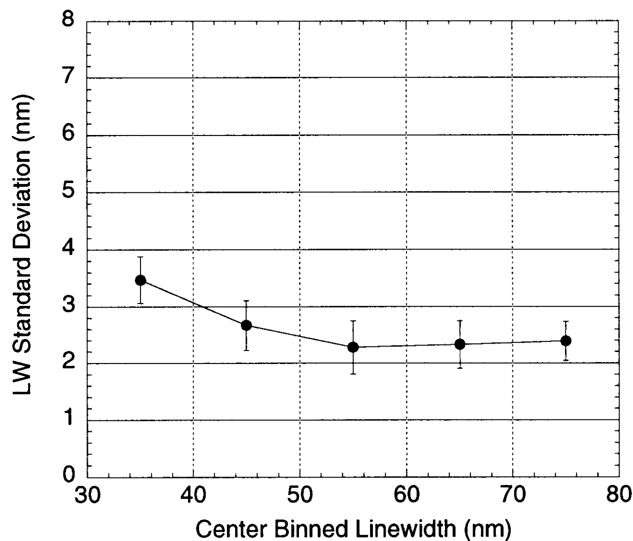


Figure 6-6: Plot of average SD of linewidth measurement on daughter mask versus measured linewidth. Measurements are binned into 10 nm groupings. The average SD of a measurement for features larger than $\sim$ 45 nm is about 2.5 nm, similar to that for the mother mask. Smaller feature sizes show in increase in measurement SD to about 3.5 nm.

Since the linewidths and spacewidths are measured independently, another check of the measurement method is to compare the sum of the two with the known period of the grating. When this is done for the grating measurements performed in this experiment, the difference is always well below the standard deviation of the measurement.[1]

Comparisons of Linewidth Measurements

With some understanding of the variability in the linewidth measurements yielded by the technique, we can now begin to compare different linewidth measurements. Even though there are uncertainties in the measurement and the comparisons, the results broadly show that the repeatability and process latitude for replication of features

[1]This lends some credence to the increase in SD seen on the daughter mask being a real effect. The consistency of the line-space measurements as indicated by their sum being very close to the grating period means that the two measurement results are probably not dramatically different. It is unlikely that the increase in SD seen for the the finest features only on the daughter mask is due to the fact that spaces are measured rather than lines as on the mother mask.

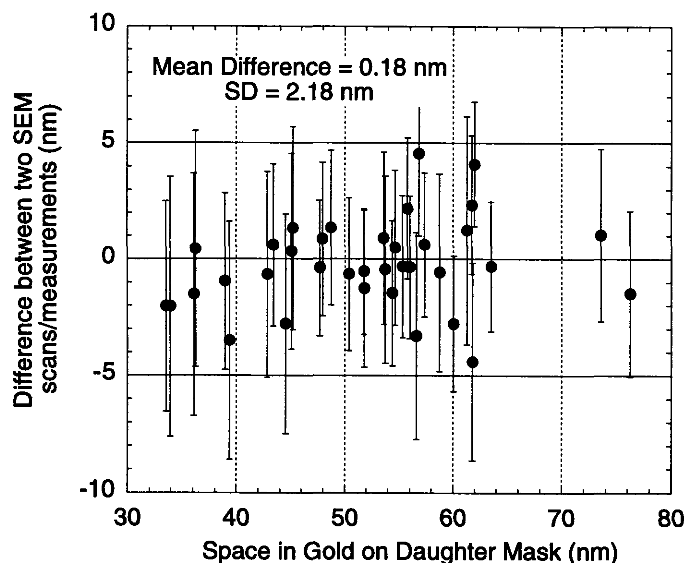


Figure 6-7: Plot of difference between two sets of linewidth measurements of the same die on the daughter mask. The data shows remarkable consistency between the two measurements, with a mean difference of only 0.18 nm, and a SD of the difference which is comparable to the SDs of the individual measurements.

ranging in size from 30-80 nm is quite good and is independent of feature size in this range.

For these measurements, selected gratings were measured in selected cells on every die on the daughter mask to yield a set of linewidths for each die which ranged from ~30-80 nm. The daughter mask was then used to expose a PMMA-coated wafer. After exposure, the wafer was cleaved into individual dies. Each die was then developed in an interrupted fashion and the time required for large features to clear (the "base development" time) was noted and found to be consistent (within ~5%) for each die, indicating a uniform exposure across the wafer. Dies were then developed for differing times beyond the clearing dose, from 100% of clearing dose to 200% of clearing dose. The individual dies were then measured by imaging in the SEM following liftoff and comparisons were made between linewidth measurements from the daughter mask to the die and from die to die.

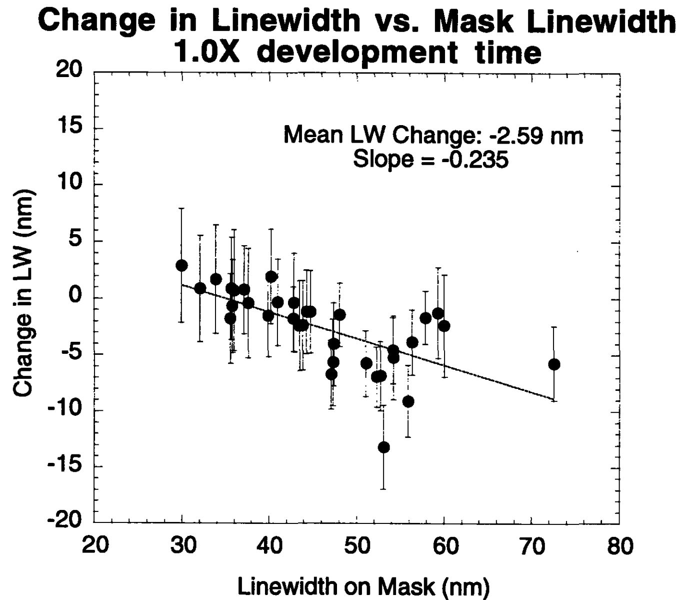
Apples to Apples – Two different sets of measurements on the same die on the mask: Sets of SEM images were taken of a particular die on the mask on two different days, and the two sets of images were analyzed. The difference between the two sets of measurements for each grating image is plotted in Figure 6-7.

The data shows remarkable consistency between the two sets of measurements. The mean difference between two measurements is only 0.18 nm, and the SD of the difference is comparable to (and actually less than) the SD of the individual measurements. The excellent comparison between two measurements on the same sets of structures allows confidence in comparisons of different structures.

Apples to Oranges – Mask measurements compared to liftoff pattern measurements: Ultimately it is desirable to know exactly how features on the mask are replicated onto a substrate. However, measuring exact feature sizes with the SEM is difficult – the finite width of the electron beam combined with scattering of electrons through the specimen can yield secondary electron generation at points other than at the location where the SEM thinks its beam is being aimed. This tends to smear out an image in geometrically-dependent ways. Sometimes the effect can be easily understood (for instance the Gaussian beam profile causes broadening of lines and narrowing of spaces), but in some cases (for instance two lines of the same width but of differing thickness) the effect on linewidth might not be as obvious.

In addition to uncertainty in how the SEM measurement might differ from the actual feature size, the choice of pattern transfer method, such as liftoff or electroplating, can introduce modifications to the measured linewidth, either by not exactly reproducing the pattern in the resist or by introducing geometrical artifacts which affect the secondary electron signal.

Mindful of the above caveats, plots of the difference between measured linewidths on the daughter mask and measured linewidths of the lifted-off pattern on a substrate are shown in Figure 6-8. The top plot shows the comparison for a die that was developed for the base development time. A positive linewidth change means that the line on the substrate is wider than the corresponding feature on the mask. The



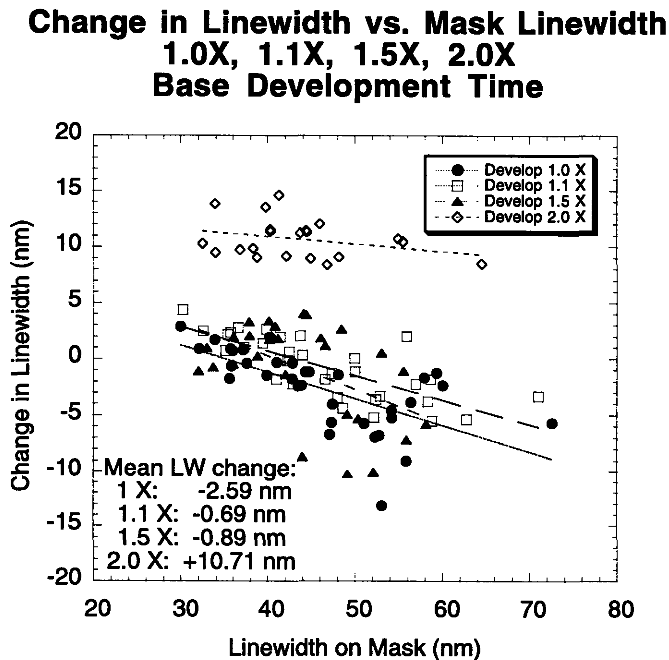


Figure 6-8: Plots of difference between mask and liftoff pattern linewidth measurements versus mask linewidth for differing development times. Top: Measurement for one die developed for the time required to clear large features (the base development time). The average linewidth change is small – ~2.5 nm. However there is a linear trend in the data; wider features tend to narrow more than narrow features. This is possibly, although not obviously, an artifact of the measurement. Bottom: A similar plot (without error bars for clarity) for 1.0X, 1.1X, 1.5X, and 2.0X base development time. The lower development times all show little difference between them, particularly for the 1.1X and 1.5X developments. The 2.0X development is clearly shows much wider features than the others.

most remarkable feature of the data is the linear trend which indicates that for every ~4 nm increase in feature size on the daughter mask, the printed feature on the substrate narrows by about 1 nm. It is unclear whether this is a real effect or an artifact of the measurement.[2]

If we ignore for the moment the linear trend in the data, we see that the average change in linewidth is quite small – only ~2.5 nm. The bottom plot in Figure 6-8 shows a similar plot (without error bars for clarity) for 1.0X, 1.1X, 1.5X, and 2.0X base development time. The lower development times show little difference between them, particularly for the 1.1X and 1.5X developments. The 2.0X development clearly shows much wider features than the others.

This data would seem to indicate that for development times up to 50% more than the base development time that linewidths do not change significantly. To further investigate the linewidth change with development time without possible artifacts from measuring different objects, we now compare the dies among themselves.

(Fuji) Apples to (Golden Delicious) Apples – Comparisons between lifted-off patterns with different development times: To compare the linewidths between lifted-off features, we use the die which was developed to 1.0X the base dose as a reference. For each die corresponding to a longer development time, we take the measured linewidth from a particular grating on that die and subtract it from the nominal measurement of the same grating on the 1.0X developed die. Figure 6-9 shows a plot of the result of such a comparison.

Again we see that lines generally appear to widen as development time increases,

[2]It is difficult to see how this effect could be related to the x-ray exposure. The narrower lines tended to be on the more widely-spaced gratings because of proximity effect in the electron-beam writing. Since many doses were used in the measurement sets taken on each die, however, except for the very finest and very widest lines, a range of periods is represented. Any type of proximity effect which might exist in the x-ray exposure would tend to widen out the wider lines rather than narrowing them. On the other hand it is difficult to understand how the measurement could induce such a linewidth change, since any differences in measuring lines and spaces would be expected to introduce some constant offset (for instance the width of the electron beam in the SEM). One possible, although highly speculative, explanation would involve tensile stress in the PMMA. Narrower lines have more PMMA between them (even in the absence of their generally being more widely spaced in this experiment). If the PMMA were to be under enough tensile stress, it might tend to widen the narrow lines more than the wide ones.

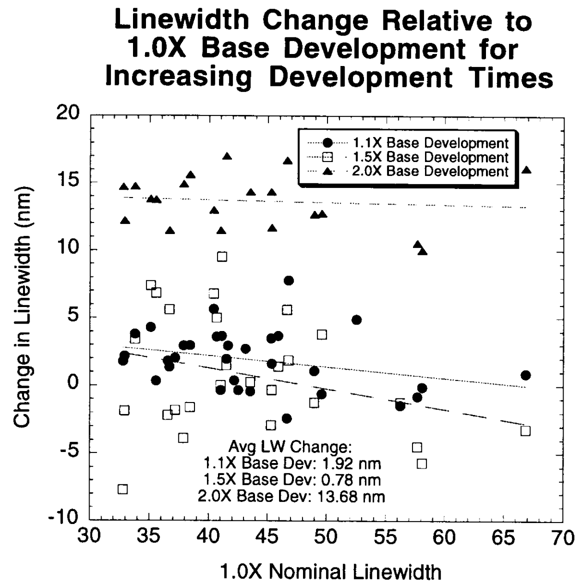


Figure 6-9: Plot of difference between lifted-off pattern linewidth measurements versus linewidth for differing development times. Lines generally appear to widen as development time increases, with the 2.0X base development widening more than the others. There is excessive scatter to the data, however, due to the fact that the gratings being compared were themselves printed from gratings (of the same period but from different dies on the mask) which had different linewidths.

with the 2.0X base development widening more than the others. There is excessive scatter to the data, however, due to the fact that the gratings being compared were themselves printed from gratings (of the same period but from different dies on the mask) which had different linewidths.

(Empire) Apples to (Empire) Apples – Binned comparisons between lifted-off patterns at different development times In order to overcome the problems associated with comparing somewhat different test objects (lines and spaces) whose measurements are known exactly and comparing similar test objects which were fabricated at differing linewidths to start with, a combination approach was used. The ideal comparison is between gratings which were fabricated from mask gratings of the same linewidth which were developed for different times. A comparison similar to this one can be extracted from the data of Figure 6-8.

Linewidth measurement SDs of this data are about 2.5 nm. By taking the linewidth change from the mask given in Figure 6-8 and binning it into regions of ± 2.5 nm, we

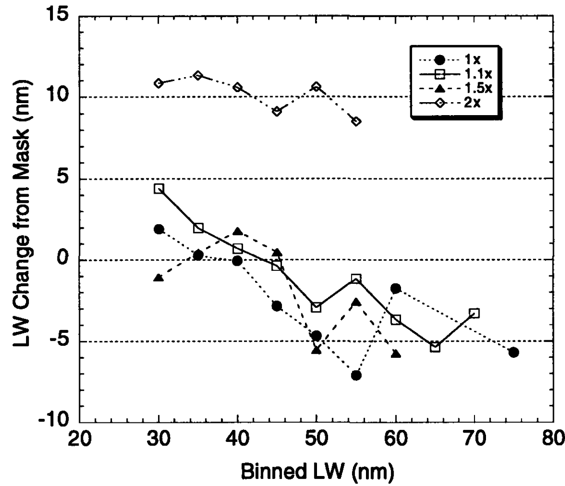


Figure 6-10: Plot of linewidth change from mask versus binned mask linewidths for differing development times. This plot uses the same data as Figure 6-8, but by binning the data into regions of ± 2.5 nm, gratings which were printed from essentially similar mask structures are grouped together. This data will form a basis for comparison between the different development times.

have in each bin a collection of measured linewidths for features which were printed from essentially the same sized grating on the daughter mask. If we do this for each development time, we have a set of data which approximates our ideal data set. This binned data set is shown in Figure 6-10.

By using the nominal linewidth change from the 1.0X base development as a reference, we can plot the linewidth change from nominal as a function of development time. Figure 6-11 shows four plots of the data organized in this manner for mask linewidths of 30, 40, 50, and 60 $\pm$ 2.5 nm.

The data is quite similar for all feature sizes. Development for 10% and 50% beyond the time required for large features to clear does not widen the features in a statistically significant manner. 100% overdevelopment does increase the linewidth by approximately 5-10 nm.

The data presented above shows quantitatively that sub-50 nm feature replication can be performed with good repeatability and process latitude. Data taken for 30 nm feature replication was essentially similar to that for 60 nm and larger feature sizes.

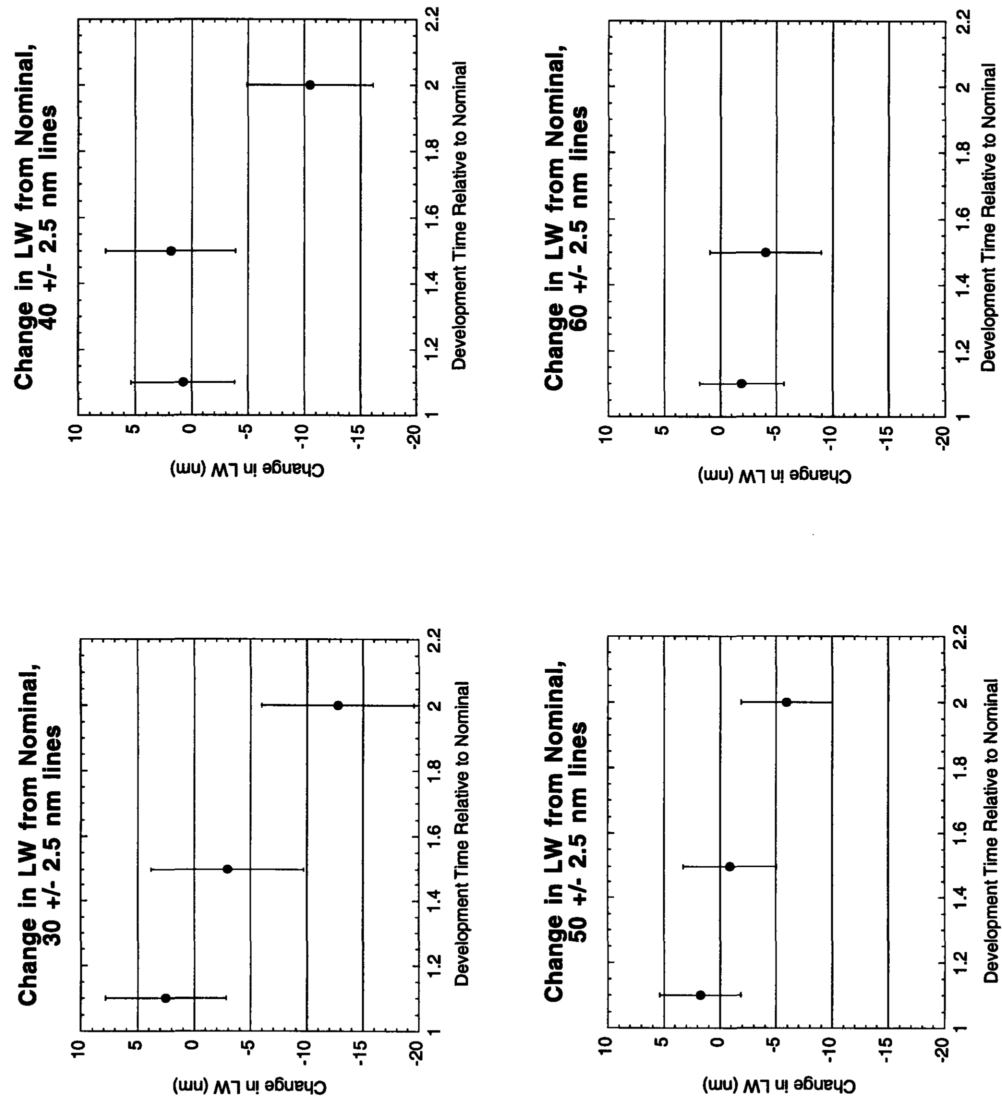


Figure 6-11: Plot of linewidth change from nominal versus development time for mask feature sizes of 30, 40, 50, and 60 $\pm$ 2.5 nm. This data is taken from Figure 6-10. For development times up to 50% greater than the time required to clear large features, linewidths do not change in a statistically significant manner.

For all feature sizes measured, development of 50% greater than the time required for large features to clear showed no statistically significant change in linewidth.

Chapter 7

Transport Measurements

The previous chapters have described a robust process for producing arbitrary patterns with sub-50 nm features on device substrates with x-ray nanolithography. This chapter presents a test of this technology which goes beyond merely examining lithographically-defined features as in Chapter 6. Real devices are produced on substrates and their electrical characteristics are measured.

The coupled quantum dot device structures described in Chapter 2 were fabricated. Gate patterns for the device were defined in PMMA with an aligned vacuum contact exposure as described in Chapter 4. Schottky gates were then deposited by liftoff onto the GaAs/AlGaAs heterostructure, which had previously been processed to have mesa isolation structures and ohmic contacts to the 2DEG. The heterostructure was grown by MBE by Professor Michael Melloch at Purdue University. It had a measured electron mobility of 250,000 cm^2/Vs, and a measured electron density in the 2DEG of 3×10^{11} cm^{-2}. This corresponds to an electron mean-free-path of $\sim$2.25 μm. Details of the GaAs/AlGaAs heterostructure processing can be found in Appendix B. Some details of the characterization of the 2DEG can be found in Burkhardt [6]. A good general reference for characterization techniques for semiconductor materials and devices is Schroder [155].

As discussed in Chapter 2, it is desirable to have a relatively large voltage range over which the interdot tunnel barrier is modulated from having effectively no conductance (infinite resistance) to a conductance of 77 μS (a resistance of 13 kΩ), the

latter being the value at which the number of electrons on the quantum dot is no longer well-defined. If the voltage range of this transition is too narrow, the ability to control the conductance of the barrier between the two dots could be compromised by noise on the gate voltage signal.

Successful performance of the device depends on the pinchoff characteristic of this tunnel barrier, which depends on the lithographic dimensions of the Schottky gates. The pinchoff characteristic of the barrier is therefore the performance parameter for this device which indicates whether or not fabrication was successful.

Transport measurements were performed at cryogenic temperatures to analyze the fine-line tunnel barriers. The device was bonded to a header and placed in the Oxford Instruments Heliox probe, which has a base temperature of ~300 mK. The current vs. gate-voltage characteristic was measured using a small (~10-100 μV) AC voltage excitation from source to drain. A transimpedance amplifier (which converts current to voltage) and a lock-in amplifier were used to detect the current. Details of the cryogenic system and the measurement electronics are given in Appendix C.

7.1 Conductance vs. Gate Voltage Curves

A scanning electron micrograph of the coupled quantum dot device whose measurements are presented in this chapter is shown in Figure 7-1 superposed with gate labels and schematically-depicted ohmic contacts. The two quantum dots are formed in the oval-shaped spaces bordered by the Schottky gates. There are four ohmic contacts, labeled 1–4. The top and bottom gates, labeled A, B, C and S, T, U are symmetric about the middle gate, labeled M. Gate M forms the fine-line tunnel barrier. The QPCs are defined by gate M in combination with gates A, C, S, or U.

To perform a measurement, a small bias is applied across two ohmic contacts which are on opposite sides of the barrier of interest. This bias is held constant, and the current is measured while the Schottky gates are ramped to a negative value with respect to the 2DEG. The measured current is then proportional to the conductance through the barrier resistance (although series resistances must be accounted for).

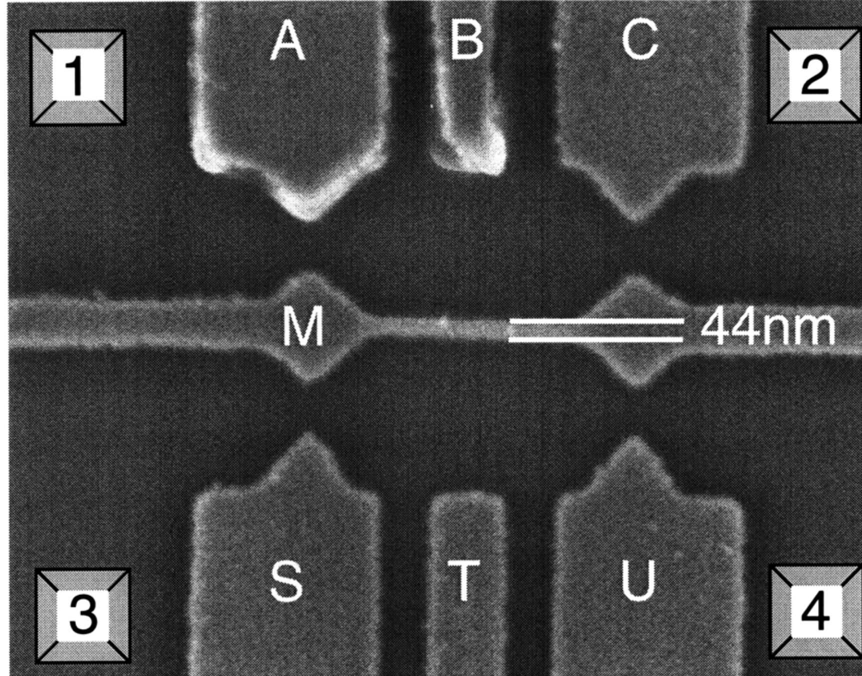


Figure 7-1: Scanning electron micrograph of coupled quantum dot Schottky gates on GaAs/AlGaAs heterostructure. Gates are labeled, and ohmic contacts are depicted schematically. The fine-line tunnel barrier gate width was 44 nm.

Figure 7-2 shows conductance vs. gate voltage curves for the middle gate, M, and a QPC formed by gates M-U on the device. As the Schottky gates are ramped to a negative voltage with respect to the 2DEG, they begin to deplete the electrons underneath. Initially no change is seen in the conductance of the whole system because the series resistance of the wires in the probe limits the current. Once the 2DEG is depleted enough that its resistance begins to dominate, the conductance drops. For the measurement in Figure 7-3, this occurs at about -0.5 V. The conductance continues to drop until the 2DEG underneath the Schottky gates is fully depleted. At this point a standard depletion-mode FET would be considered to be at threshold or pinchoff.

In our case, however, there is still some current allowed at this point, either through the gap in the QPC or through the leaky tunnel barrier. As the voltage on the gates is made more negative, the conductance continues to decrease, but at a lower rate, because the fields must extend laterally from the gates to further deplete the 2DEG. For the QPC, this is the point at which the conduction becomes

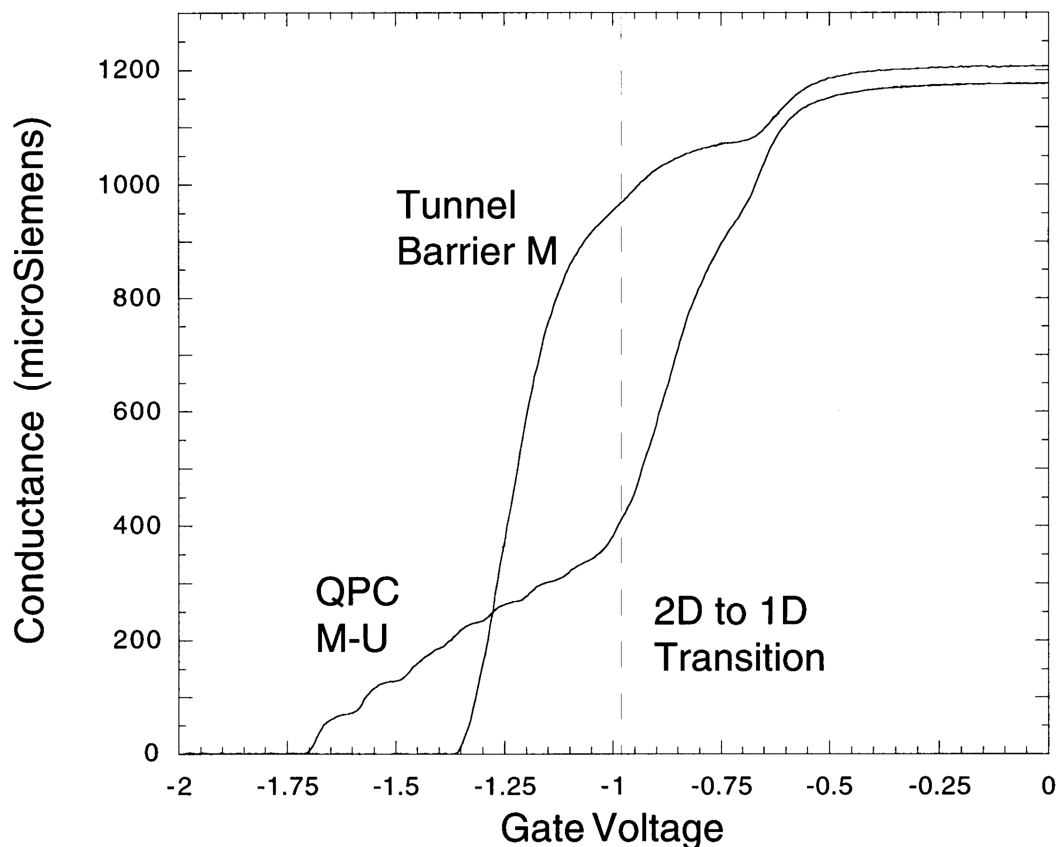


Figure 7-2: Conductance vs. Gate Voltage for the fine-line tunnel barrier and a QPC on the coupled quantum dot device. The point where the 2DEG underneath a large, continuous gate is fully depleted is indicated by a dashed line on the plot, labeled the 2D to 1D transition. As the gate voltage is decreased below this point, significant differences are seen in the behavior of the QPC and the tunnel barrier.

one-dimensional. This point is therefore termed the *2D to 1D transition.* For the measurements above, this kink in the conductance curve occurs at about -1 V.

Beyond the 2D to 1D transition point, the characteristics of the QPC and the tunnel barrier are strikingly different. The next two sections will examine the characteristics of the two cases in more detail.

7.2 QPC Pinchoff Characteristic

Figure 7-3 shows the conductance vs. gate voltage characteristic for the QPCs formed by gates *M-S* and *M-U* on the device structure. The region where the barriers just begin to open is shown. Quantized conductance steps can clearly be seen with plateaus occuring at integer multiples of the quantum of conductance, 77 μS. (The decreasing spacing between conductance peaks and the slightly low conductance value for the first step can be accounted for by a series resistance of about 1.1 kΩ). The plateaus are relatively flat, particularly for QPC *M-U*, indicating a lack of scattering in the one-dimensional channel.

The inset to Figure 7-3 shows a blowup of the transition to the first conductance step for QPC *M-U*. Since the double-dot system is intended for experiments where the barriers are transitioned from fully-closed to being open to approximately the quantum conductance, the voltage shift, ΔV_g, required to transition the QPC barrier from fully closed to 50 μS (one channel ~2/3 opened) is measured as a benchmark. For the QPC, $\Delta V_g \approx 0.04$ V.

7.3 Tunnel Barrier Pinchoff Characteristic

Figure 7-4 shows the conductance vs. gate voltage characteristic for the fine-line tunnel barrier plotted with the same *y*-scale as the QPC in Figure 7-3. The horizontal dotted lines are at the quantized conductance plateau values of the QPCs. The conductance rises much more smoothly with no plateaus at the conductance values expected for a one-dimensional channel.

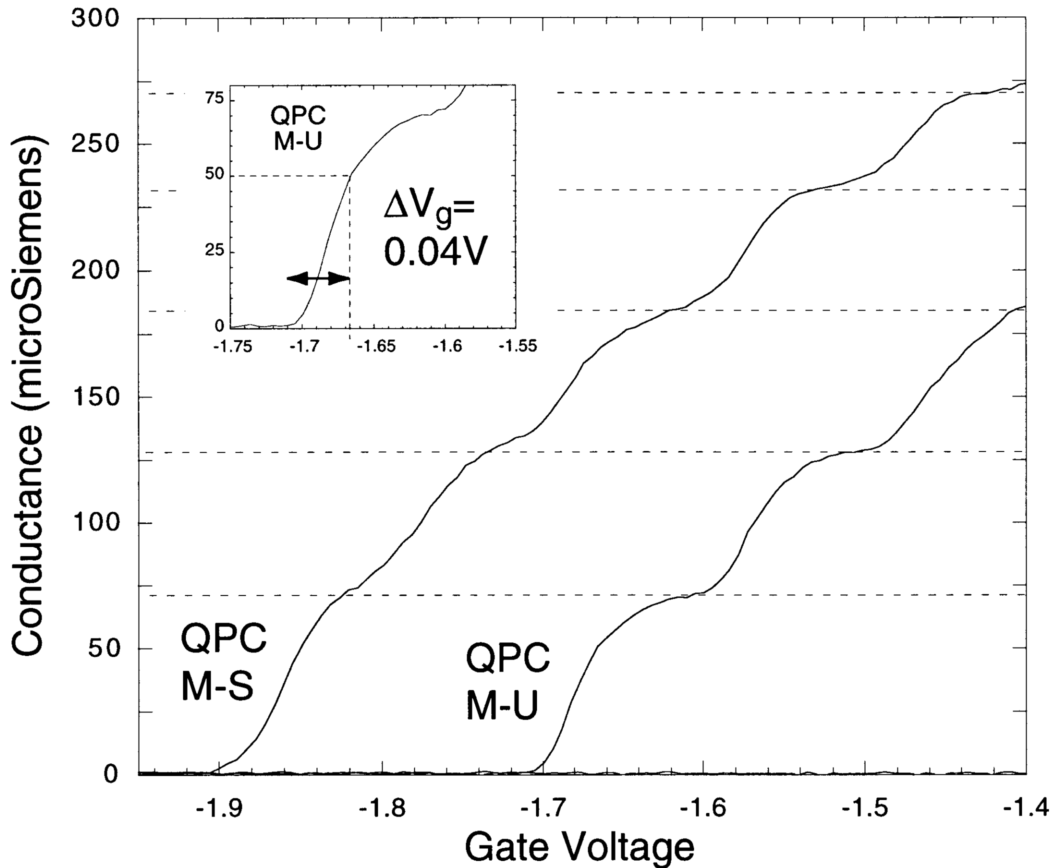


Figure 7-3: Conductance vs. Gate Voltage for QPCs on a high mobility GaAs/AlGaAs heterostructure. Several conductance steps are seen. Each corresponds to one conductance channel which is fully transmissive and which has a conductance of 77 μS. In this case a series resistance of $\sim$1.1 kΩ reduces the spacing between steps as the number of modes increases. Inset: A closeup of the conductance rise from 0 to 80 μS for one of the QPCs. The change in voltage required to open the barrier to 50 μS is $\sim$0.04 V.

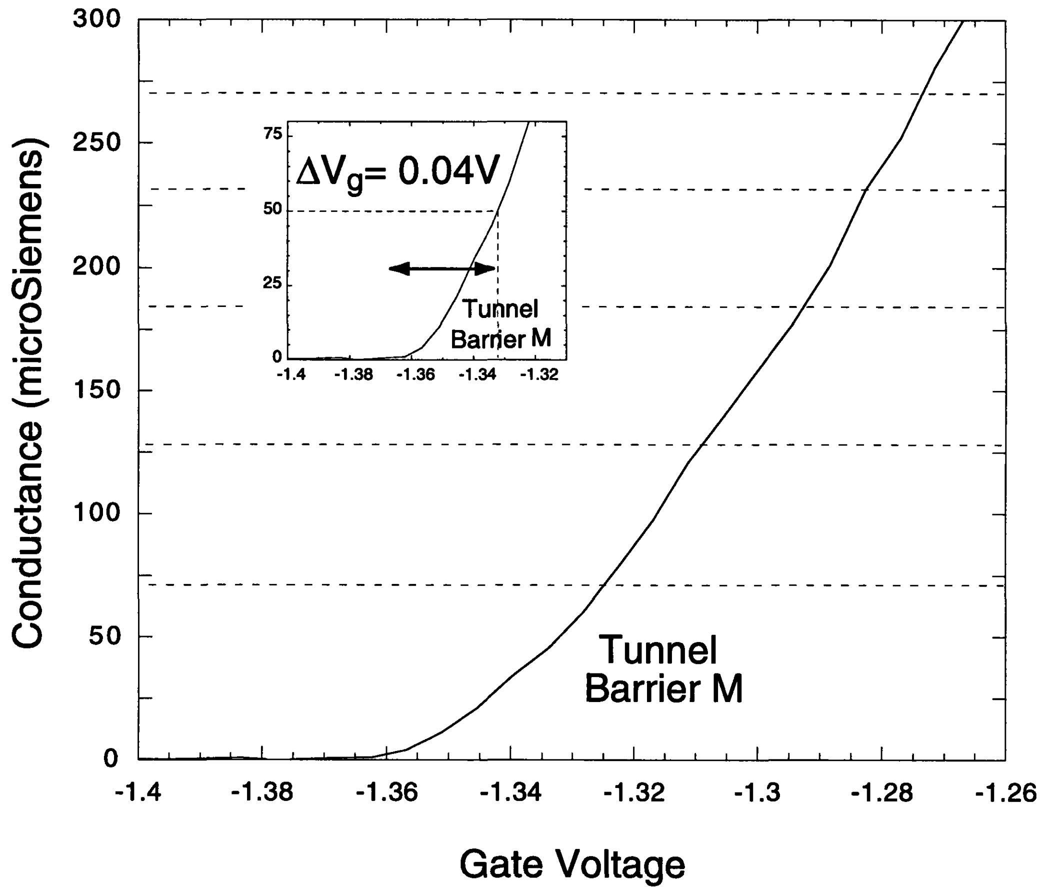


Figure 7-4: Conductance vs. Gate Voltage for a fine-line tunnel barrier on a high mobility GaAs/AlGaAs heterostructure. In this case the conductance rises smoothly with no steps as the barrier is lowered. The dotted horizontal lines are the QPC conductance plateaus from the previous figure. Inset: A closeup of the conductance rise from 0 to 80 μS. The change in voltage required for an opening to 50 μS is $\sim$0.04 V as in the QPC case.

The inset shows a closeup of the conductance rise to the quantum conductance. Although the tunnel barrier opens to large conductance values (much greater than the quantum conductance) in a smaller gate voltage swing than for the QPC case, the gate voltage shift required to open the barrier to $50\,\mu$S is approximately 0.04 V, the same as for the QPC.

Although this device and several others performed satisfactorily upon initial measurement, measured characteristics of the tunnel barriers tended to become more QPC-like over time, developing plateaus at integer multiples of the quantum conductance. A discussion of this phenomena is given in Appendix C. The lithographic dimensions of the gates do not change with time, however, so the data shown above and similar data from other samples indicates that the lithography was successful at placing the proper gate pattern on the substrate.

The fine-line gate therefore has been shown to act as a tunnel barrier whose control by a gate voltage is comparable to that of a QPC for the transition from being fully closed to being open to the quantum of conductance. This demonstration of the sub-50 nm x-ray lithographic process shows that the process is capable of producing devices on substrates which meet previously-unattainable performance specifications.

Chapter 8

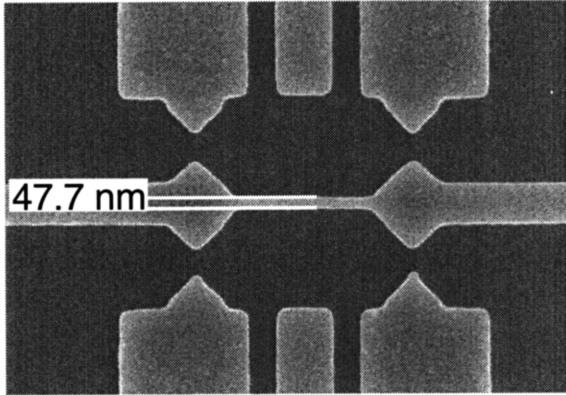
Conclusions

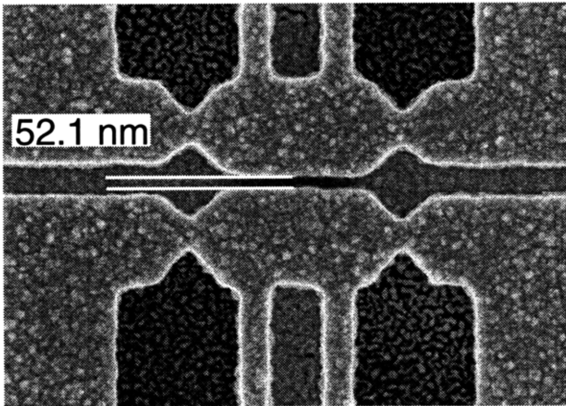
This thesis was described at the outset as an example of an interplay between lithographic technology and microelectronic device technology. The desire to fabricate a coupled-quantum-dot structure with a sub-50 nm fine-line tunnel barrier created a need to push the x-ray lithography process further than it had been previously pushed.

The work of this thesis was in large part a response to the requirements of the device technology, and those requirements were successfully met. The gate pattern which was required for the device was created, and it was shown that the structure had the desired geometry and that the device behaved as expected electrically.

The work far exceeded the requirements of the coupled quantum dot structure, however. A process was developed which allows the creation of arbitrary patterns on a substrate, and it was shown to have a repeatability and process latitude which remains very good for feature sizes as small as ~30 nm. The process is quite robust and shows a significantly-improved yield and much less risk of mask breakage than the previous technologies on which it is built. It should have broad applicability to many device applications.

The vacuum soft contact replication technology, particularly for small chips, should result in an improved overall yield for x-ray exposures done on small chips in the NanoStructures Laboratory at MIT. Similarly the suck-and-puff technique for mask replication should result in improved feature uniformity in mask replication.





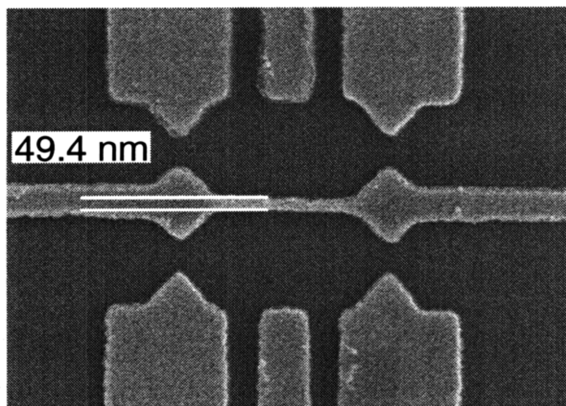


Figure 8-1: Scanning electron micrographs of the same device structure patterned on a mother mask (by electron-beam lithography), a daughter mask (by x-ray replication from the mother mask), and a device substrate (by x-ray replication from the daughter mask).

The Semiconductor Industry Association Technology Roadmap was also referenced at the beginning of the thesis. If indeed 50 nm feature sizes are to be commercially manufactured by the year 2012, as suggested by this document, it is not improbable that contact techniques such as the ones described in this thesis may be required. Industry has historically avoided contact techniques for lithography, fearing damage to the mask or the sample, but with the many exposures done for this thesis, no damage was ever seen to a mask or a substrate because of contact. Printing 50 nm features in production will be extremely difficult, and contact lithography can in many ways decrease the level of difficulty involved in the lithography. I would suggest that it would be wise to at least not totally rule out contact lithography as a possible technological choice for industry.

The observation of the effect of substrate photo- and Auger electrons also has broader application than making a coupled quantum dot device. In addition to electroplated gold masks, many researchers in x-ray lithography are using subtractive processes where continuous films of absorber metal on the x-ray mask are etched away to leave the desired mask pattern. In a subtractive process, thinning the absorber metal will not be an option. The understanding of the effect developed from the measurements and simulations presented in this thesis should aid in the development of solutions to problems which may arise due to substrate-generated electrons.

This work has shown the capability robust pattern fabrication with feature sizes as small as approximately 30 nm. This minimum size was set by the smallest feature which could be written with electron-beam lithography on the mother mask. It is quite likely that replication of 10 to 20 nm features might be possible using the techniques developed for this thesis if smaller features could be fabricated on masks. This could be done with some trick (perhaps overdeveloping a daughter mask) to show the capabilities, but finding a way to do so reliably and repeatably would open up another new area for nanofabrication.

The methods used to measure the statistical results presented in Chapter 6 could be put to further use. Several questions arose in their development about whether some results which were seen in the data (such as the decreasing linewidth change

in replication with increasing linewidth) were an artifact of the measurement or a real effect. Further investigation might yield interesting results. It might be possible to experimentally measure the ~3 to 5 nm effective range of the photo- and Auger electrons in PMMA – a 6 to 10 nm linewidth reduction should be observable in a 30 nm feature.

The push by device technology to create smaller feature sizes has resulted in a lithographic technology with extended capabilities. Fabrication which was previously either totally inaccessible or so difficult as to be impractical for all but the most critically important tasks can now be done reliably, repeatably, and with an improved yield. Now that the technology exists for robust replication of arbitrary patterns with sub-50 nm feature sizes, it is hoped that x-ray lithography will push back at device technology and start the next cycle of the process.

Appendix A

Mask Fabrication

This Appendix contains process flows or recipes used for processing masks as outlined in the main section of the thesis. Rather than present the whole process flow in one piece, I have chosen to break it up according to particular processes which would typically happen in one session or day. This method of presentation does involve some repetition between processes, but should make the recipes more suitable for actual use in the laboratory.

Contents

A.1 Mother Mask Preparation for NRL 157

A.2 Developing and Plating a Mother Mask 159

A.3 Suck-and-Puff Daughtering 166

A.4 Developing and Plating a Daughter Mask 168

A.5 Polyimide Coating of Masks 174

A.6 Clearing Polyimide from a Mask or Part of a Mask . . . 178

A.1 Mother Mask Preparation for NRL

These are the steps that should be done to prepare a mother mask for shipping to the Naval Research Laboratory for writing on their e-beam. Masks should have plating base evaporated and 220 nm of PMMA spun on and baked by a technician.

EQUIPMENT NEEDED:

Hood Space for Developing, Etching   OAI UV Source at 200 nm

- Choose a mask that is free from particles embedded in the membrane and which has thick-gold fingers which extend slightly onto the membrane.[1]
- Flood expose the mask with 200 nm wavelength UV in the OAI for 30 minutes (typical power level is ~0.9 mW/cm^2 on the Mimir meter) using the foil "lollipop" head which has holes corresponding to the thick-gold fingers. Align the mask so that the slots are over the fingers.[2]
- Develop the PMMA for about 30 seconds in 2:3 MIBK:IPA. Rinse with IPA and blow dry.
- Using a plastic eye dropper to place drops of etchant over the fingers, etch each of the fingers (it's easiest to do one at a time) for 10 seconds in 10:1 DI water:Au etch. After each etch, rinse the mask in DI water and blow dry.
- Place the mask in fluoroware with a plastic spring between the back and the cover. CAUTION: *This spring can make placing the cover somewhat difficult.* Tape the fluoroware shut with tape around the edge and crossed across the top and bottom.

[1]The shadow evaporation fixture for masks should be modified so that all masks have these fingers. At this point only half of the processed masks do. Having the fingers does not interfere with the mask's use as a daughter mask.

[2]This lollipop head, being made of foil, is rather fragile. A more robust one should be made out of thin sheet metal.

- Wrap the masks securely in bubble wrap, foam, and/or styrofoam peanuts and ship them.

A.2 Developing and Plating a Mother Mask

Parts of this process were adapted from the process presented in Burkhardt [6], which is also a good reference for more information about the plating bath I-V characteristics.

<u>EQUIPMENT NEEDED:</u>

PlasmaTherm RIE	Plating bath
AFM	Hood Space for Developing
Optical Microscope	Digital Multimeter

- Turn on the pump and the temperature controller for the plating system. Open the chilled water loop for the water bath. The plating system needs about an hour or so to stabilize to the desired bath temperature (33°C). The set point for the water bath might have to be adjusted to reach the desired bath temperature. It is generally around 30°C.
- Mix the developer (1:2 MIBK:IPA) in a 1-liter beaker. The temperature should drop to about 16°C upon mixing. It will take some time to come to the developing temperature of 21°C, but this will happen while other preparatory work is being done for the plating. Warming can be accelerated by placing the hands on the outside of the beaker. It's also common to use pre-mixed (and therefore pre-warmed) developer. However it needs to be stored in an airtight container, because preferential evaporation of MIBK will change the concentration over time.
- Do a clean-up run on the Plasma-Therm RIE to condition the chamber and fixturing for a successful plasma strike for the oxygen descum etch of the mask.
  - Make sure the mask fixturing and the dummy mask are in the chamber and properly mounted.

- Pump out the chamber to below 1×10^{-4} Torr (this is a bit higher pressure than is preferable for the actual mask, but for the cleanup it should be fine).
- Run the process program `dsclean.prc`. It does two ten-minute etches, the first at 300 W power with He/O_2 and the second at 50 V DC bias with He/O_2.
- Leave the chamber under vacuum while you develop the mask.

- During the clean-up run, the plating bath can be characterized. This involves I-V characterization and measuring the temperature and conductivity.
  - Inspect the water levels in the plating bath. When the pumps are running, the plating solution level in the outer container should be about one inch below the rim of the inner container. If it is too low, more plating solution should be added from the makeup reserve. The water bath should come to about one inch below where the plating bath container necks outward. If it is too low, water should be added using the spray gun located next to the plating solution.
  - After the bath has been circulating for ~10-20 minutes after any additions of water or plating solution, measure the temperature and the conductivity of the bath. Enter the information (including any bath makeup solution added) in the `bath.dp` database (this database entry will later be referenced in the `samples.dp` database).
  - Take I-V data.
    * Place the calomel electrode in the bath, using the Tyvex insert to raise it so that its water level is higher than that of the bath (If the electrode water level is still lower than the level in the bath, add a few drops of DI water to the electrode through the hole in the "open" position).
    * For the I-V characterization, use a 3" wafer which has plating base on it. This gives a known area for plating. The wafer can be reused many times for the I-V curve. If it has been in fluoroware for a while, it

should be UV-ozone cleaned for 1-2 minutes prior to being mounted in the plating fixture. Before putting the contact ring into place, scrape the three contacts lightly with a razor blade. Place contact ring, then grid, onto fixture.

* Use a digital multimeter to monitor the voltage between the calomel electrode and the wafer. Use alligator clips to attach the ground to one of the three white/black wire connectors on the sample holder (not the red/red wire connector). Attach the positive lead to the calomel electrode. Open the calomel electrode by turning the white ring near the top.

* Measure the reference voltage (between the electrode and the sample) and the bias voltage (as read on the plating power supply) at the following bias currents, *in this order:* 35 mA, 25 mA, 15 mA, and 35 mA (again). The total measurement time should take about 2-4 minutes. Enter the results in the database (`bath.dp`)

* If the reference voltage for the final 35 mA reading is above about 0.60 V, brightener should be added. Typically 7-10 ml is adequate. If brightener is added, wait about 30 minutes and re-characterize the bath I-V, making another `bath.dp` entry.

* Enter total amp-hours (TOTALIZER on plating PS) into database.

* Close calomel electrode and remove from plating bath (making sure that Tyvex spacer comes with it). Rinse well, dry, and put away (with plastic cap covering tip).[3]

[3]Juan Ferrera has suggested a new method for I-V characterization which would probably yield more consistent results. The problem with the method presented here is that the reference voltage tends to drift with time as the plating progresses, rising after reaching a minimum about 30 seconds or so after beginning plating. Juan has suggested monitoring the reference voltage using a 35 mA plating current and using its minimum as the marker for when to add brightener. This would remove some of the variability in the method described here between people who might be more or less adept at changing currents and reading and writing down voltages. In order to implement this scheme, studies should be done to correlate the minimum reference voltage to the 0.60 V used as the threshold in this procedure. Preliminary results suggest that $V_{min} \geq 0.580$ volts is a comparable threshold.

- Once the developer temperature has settled to $21 \pm 0.2°C$, develop the mask for 90 seconds by slow and constant agitation in the developer. The development conditions should be the same for all mother masks.

- When the development time is up, immediately rinse the mask with IPA using the squirt bottle. Then rinse briefly under running DI and then again with IPA. Blow the sample dry with the filtered N_2 gun. It is crucial not to blow on the membrane with too much force. The pressure on the N_2 gun can be reduced. The idea is to push the IPA off the membrane before it can evaporate and potentially leave behind residue.

- Inspect the mask in the optical microscope.

- Measure the PMMA height on the membrane in the AFM. Take a couple of measurements and note where they were taken so that after plating the same places can be measured to monitor the gold height before stripping the PMMA.

- Load the sample into the RIE for descum and allow the chamber to pump down to 7×10^{-5} Torr.

- Run `descum.prc`, which etches in He/O_2 for 5 seconds at 20 mTorr. Monitor the run and watch as the plasma strikes. The "spark plasma" step is set for 50 mTorr and 5 seconds. If the plasma strikes early in the step, you may need to hit "End Step" so that the total time in the plasma is not too long. There is also a few seconds lag time while the pressure is reduced from 50 mTorr. If the plasma won't strike, it may be necessary to run the process `descum1.prc`, which strikes the plasma at 70 mTorr.

- Keep the sample under vacuum in the RIE until you are ready to plate.

- Electroplate 200 nm of gold.

  - Calculate plating current for 0.4 mA/cm^2 and the total mA-min for 8 minutes of plating. The area of a mother mask is ~32.3 cm^2, and the area of a

daughter is ~35.75 cm^2. These areas will have to be adjusted for pattern density, and in the case of mis-aligned daughters, for unexposed area.

- Use resistor (found near plating PS) attached to PS output to preset the plating current to the proper value. Set the counts indicator to the proper amp-minutes.
- Remove mask from RIE and place in plating fixture (removing the backing disc which is used for wafer plating). Before putting contact ring onto the pins, scrape the three contacts lightly with a razor blade. Place contact ring, then grid, onto holder. Tighten knurled nuts.
- Measure resistance between the three contacts. These should be low (~1-2 Ω). Write them down for entry into database later (`samples.dp`).
- Place fixture into plating bath, gently rocking back and forth to release air bubbles.
- Plate: push down `EOC` button (to `RES.`) momentarily. Adjust plating current with fine control knob if necessary. Jot down the bias voltage every minute or so.
- When plating is done, alarm sounds. Disconnect fixture from PS, turn off PS, and remove mask from the bath, holding above bath to allow liquid to drip into bath for a few seconds.
- Before mask can dry, rinse with DI water using spray gun next to plating setup. Try not to spray directly on membrane to avoid possible breakage. After a few rinses with the spray gun, fill the mask holder with water and remove grid and mask contacts.
- Rinse mask in running DI (from faucet) and blow dry.
- Thoroughly rinse the mask holder and the area around the plating setup. The brightener for the bath contains arsenic, so it is best to not have residue all over the laboratory.
- Enter plating information into the database (`samples.dp`).

- Inspect the mask. The gold should look shiny and gold-colored. In the optical microscope at high magnification using differential interference contrast, it should appear to have an eggshell texture.

- AFM the mask in the same areas which were AFM'ed previously to find the plated gold thickness.[4] If the gold isn't thick enough, the mask can be placed back into the plating solution and more gold plated. If this is done, however, the surface tends to be rougher than if all the plating is done in a single run.

- If the plating is satisfactory, turn off the plating bath pump and the bath heater, and close the valve for the chiller loop.

- Check the pattern by inspection in the optical microscope and, if needed, the scanning electron microscope.

- Strip the resist by soaking in acetone for about 10 minutes then rinsing with acetone then methanol. Alternatively, the resist can be stripped by soaking the mask in hot NMP ($\sim$90°C). Plasma ash the mask for 10 minutes to remove any residual organics.

- Check for large particles on the mask and mesa. This is best done by scanning the fiber light (held perpendicular to the mask) across the membrane while looking at a wide angle for scattered light. The mask can also be scanned under the optical microscope (lowest-powered objective) while looking for scattered light and/or looking through the eyepieces. A dark-field objective is especially well-suited for detecting particles. The mask can also be brought in close proximity with another mask, as if a daughter exposure was being done, to check for particles with interference fringes and to see if there is a large wedge, which would indicate a particle on the mesa. To remove particles, spin photoresist (PR) over the whole mask and expose the region around the particle for about

[4]It can often be difficult to accurately determine the gold thickness in this manner in the AFM. The gold is much rougher than the PMMA and only a small (typically 10-20 nm) step is being measured, so the gold roughness can make determining the height difficult. Nevertheless, one can get some idea of the gold height relative to that of the PMMA.

2 minutes in the optical microscope with maximum white-light intensity. Then develop the PR and use an eye dropper to put drops of gold etch on the openings in the resist to "lift off" the particles. The etch of the largest particles should be completed in about 20 minutes. For some particles it is necessary to also use Ti etch (TFT-type) heated to 50°C (make sure that the hot plate is well back in the hood and that only a minimal amount of TFT Ti etchant is used, since Ti etchant is HF-based!)

- For mother masks which are to be replicated in intimate contact, polyimide should be spun on. Details of polyimide spinning are given in Section A.5

A.3 Suck-and-Puff Daughtering

<u>EQUIPMENT NEEDED:</u>

Head 2,4, or 5 Suck-and-Puff Fixture

- Choose a daughter mask. For finest features use thin plating base (PB) – 10 nm Ti/1.8 nm Au. 220-250 nm of PMMA should have been spun on the mask by a technician. Inspect the mask carefully for large particles, and choose one that is relatively particle-free and has no obvious defects.
- Place the daughter mask face-up in the daughtering fixture, making sure the bottom center o-ring is properly seated. Make sure the flat on the Pyrex is *between* two alignment pins, to prevent lateral motion of the mask.
- Under green-light illumination, place the PI-coated mother mask face down on the daughter mask (again keeping the flat between alignment posts). There should be quite a large bow to the membranes, as indicated by fringes. The air bubble should decrease in size over the course of a couple of minutes.
- Check the gap (as described in Chapter 3) and wedge using the green light and also look for particles. If the wedge is very large (> 3–$4\,\mu$m or so), try repositioning the masks by separating them and rotating the mother mask. Rotating the mother mask can also determine whether a particle is on the mother or the daughter mask. Try blowing away particles with a nitrogen gun (reducing pressure first if necessary).
- Check the outer o-ring on the base and the o-ring in the Teflon cover to make sure they're properly seated. Make sure the Tyvex tube is connected to the barbed fitting on the cover before placement.
- Carefully place the cover over the masks. Try to line up the bolts with the tapped holes in the base to minimize rotation once the cover is placed. Tighten the bolts part-way by hand to ensure that each is properly threaded.

- Tighten the cover with a hex key while monitoring the fringes. Use a star pattern (like putting on a car tire) so as not to tighten one side without balancing with the other. Don't over tighten, otherwise the Teflon will flow.

- Once the cover is clamped down, attach the Swagelock end of the vacuum tube to the vacuum fitting on the head (which should be at atmosphere).

- Apply vacuum. The needle valve on the fixture should be open to ~1.75, then the vacuum level should be slowly increased (with the vacuum regulator) until the fringes start to move (usually at around 2.0–3.0 inches of Hg vacuum). Make small changes.

- Once the edges of the membranes are in contact, wait for 10–30 minutes while the bubble in the middle dissipates. Monitor with interference fringes from the green light.

- Once the masks are in contact, place them in the x-ray head and expose.

- When the exposure is finished, remove the mask fixture from the head. Dial down the vacuum and disconnect the vacuum tube from the Swagelock fitting.

- Prepare to puff. The technique is the opposite of smoking a cigar, or sort of like blowing soap bubbles with a child's toy pipe. Gentle pressure should be applied in short puffs. You should see more air being forced between the membranes with each puff. After a few puffs the masks should separate.[5]

[5]Occasionally some stubborn masks will go back into contact after separating. If this happens it is sometimes necessary to puff the masks apart, loosen the cover, push the cover down by hand and puff some more air between the membranes, then quickly remove the cover and separate the masks before they can contact each other. In this case the several-second squeeze-film time constant works to your advantage.

A.4 Developing and Plating a Daughter Mask

<u>EQUIPMENT NEEDED:</u>

PlasmaTherm RIE	Plating bath
AFM	Hood Space for Developing
Optical Microscope	Digital Multimeter

- Flood expose the daughter mask, shadowed by the lollipop, for 30 minutes at 220 nm in the OAI (power level should be $\sim 0.9\ \mathrm{mW/cm^2}$ on the Mimir meter).
- Turn on the pump for the plating system and the temperature controller. Open the chilled water loop for the water bath. The plating system needs about half an hour or so to stabilize to the desired bath temperature (33°C). The set point for the water bath might have to be adjusted to reach the desired bath temperature. It is generally around 30°C.
- Mix the developer (1:2 MIBK:IPA) in a 1-liter beaker. The temperature should drop to about 16°C upon mixing. It will take some time to come to the developing temperature of 21°C, but this will happen while other preparatory work is being done for the plating. Warming can be accelerated by placing the hands on the outside of the beaker. It's also common to use pre-mixed (and therefore pre-warmed) developer. However it needs to be stored in an airtight container, because preferential evaporation of MIBK will change the concentration over time.
- Do a clean-up run on the Plasma-Therm RIE to condition the chamber and fixturing plates for a successful plasma strike for the oxygen descum etch of the mask.
  - Make sure the mask fixturing and the dummy mask are in the chamber and properly mounted.

- Pump out the chamber to below 1×10^{-4} Torr.
- Run the process program `dsclean.prc`.
- Leave the chamber under vacuum while you develop the mask.

- During the clean-up run, characterize the plating bath.
  - Inspect the water levels in the plating bath. When the pumps are running, the plating solution level in the outer container should be about one inch below the rim of the inner container. If it is too low, more plating solution should be added from the makeup reserve. The water bath should come to about one inch below where the plating bath container necks outward. If it is too low, water should be added using the spray gun located next to the plating solution.
  - After the bath has been circulating for ~10-20 minutes after any additions of water or plating solution, measure the temperature and the conductivity of the bath. Enter the information (including any bath makeup solution added) in the `bath.dp` database (this database entry will later be referenced in the `samples.dp` database).
  - Take I-V data.
    * Place the calomel electrode in the bath, using the Tyvex insert to raise it so that its water level is higher than that of the bath (If the electrode water level is still lower than the level in the bath, add a few drops of DI water to the electrode through the hole in the "open" position).
    * For the I-V characterization, use a 3" wafer which has plating base on it. This gives a known area for plating. The wafer can be reused many times for the I-V curve. If it has been in fluoroware for a while, it should be UV-ozone cleaned for 1-2 minutes prior to being mounted in the plating fixture. Before putting the contact ring into place, scrape the three contacts lightly with a razor blade. Place contact ring, then grid, onto fixture.

    * Use a digital multimeter to monitor the voltage between the calomel electrode and the wafer. Use alligator clips to attach the ground to one of the three white/black wire connectors on the sample holder (not the red/red wire connector). Attach the positive lead to the calomel electrode. Open the calomel electrode by turning the white ring near the top.
    * Measure the reference voltage (between the electrode and the sample) and the bias voltage (as read on the plating power supply) at the following bias currents, *in this order:* 35 mA, 25 mA, 15 mA, and 35 mA (again). The total measurement time should take about 2-4 minutes. Enter the results in the database (`bath.dp`)
    * If the reference voltage for the final 35 mA reading is above about 0.60 V, brightener should be added. Typically 7-10 ml is adequate. If brightener is added, wait about 30 minutes and re-characterize the bath I-V, making another `bath.dp` entry.
    * Enter total amp-hours (TOTALIZER on plating PS) into database.
    * Close calomel electrode and remove from plating bath (making sure that Tyvex spacer comes with it). Rinse well, dry, and put away (with plastic cap covering tip).

- When the developer has reached $21 \pm 0.2°C$, develop the daughter mask. Use interrupted development, which allows monitoring of the development rate.
  - Develop for 15 seconds using slow agitation in the developer. Rinse in IPA and blow dry.
  - Measure the step height in a couple of places on the mask in the AFM.
  - Repeat this cycle a couple of times to get a sense for the development rate. The development time increment can then be increased.
  - When the end of the development is approached, decrease the development time increment again to try and hit the clearing development time. The

cleared plating base should appear much smoother in the AFM than the developing PMMA did. The clearing can also be monitored in the optical microscope by looking at the roughness with differential interference contrast.

  - Develop about 10% beyond the clearing dose for large features. For the final rinse, go from IPA→DI→IPA, then blow dry.
  - Do a final AFM, noting the height and location for comparison after gold electroplating.

- Load the sample into the RIE for descum and allow the chamber to pump down to 7×10^{-5} Torr.
- Run `descum.prc`. Monitor the run and watch as the plasma strikes. The "spark plasma" step is set for 50 mTorr and 5 seconds. If the plasma strikes early in the step, you may need to hit "End Step" so that the total time in the plasma is not too long. There is also a few seconds lag time while the pressure is reduced from 50 mTorr. If the plasma won't strike, it may be necessary to run the process `descum1.prc`, which strikes the plasma at 70 mTorr.
- Keep the sample under vacuum in the RIE until you are ready to plate.
- Electroplate 200 nm of gold.
  - Calculate plating current for 0.4 mA/cm^2 and the total mA-min for 8 minutes of plating. The area of a mother mask is ~32.3 cm^2, and the area of a daughter is ~35.75 cm^2. These areas will have to be adjusted for pattern density, and in the case of mis-aligned daughters, for unexposed area.
  - Use resistor (found near plating PS) attached to PS output to preset the plating current to the proper value. Set the counts indicator to the proper amp-minutes.
  - Remove mask from RIE and place in plating fixture (removing the backing disc which is used for wafer plating). Before putting contact ring onto the

pins, scrape the three contacts lightly with a razor blade. Place contact ring, then grid, onto holder. Tighten knurled nuts.

- Measure resistance between the three contacts. These should be low (~1-2 Ω). Write them down for entry into database later (`samples.dp`).
- Place fixture into plating bath, gently rocking back and forth to release air bubbles.
- Plate: push down `EOC` button (to `RES.`) momentarily. Adjust plating current with fine control knob if necessary. Jot down the bias voltage every minute or so.
- When plating is done, alarm sounds. Disconnect fixture from PS, turn off PS, and remove mask from the bath, holding above bath to allow liquid to drip into bath for a few seconds.
- Before mask can dry, rinse with DI water using spray gun next to plating setup. Try not to spray directly on membrane to avoid possible breakage. After a few rinses with the spray gun, fill the mask holder with water and remove grid and mask contacts.
- Rinse mask in running DI (from faucet) and blow dry.
- Thoroughly rinse the mask holder and the area around the plating setup. The brightener for the bath contains arsenic, so it is best to not have residue all over the laboratory.
- Enter plating information into the database (`samples.dp`).

• Inspect the mask. The gold should look shiny and gold-colored. In the optical microscope at high magnification using differential interference contrast, it should appear to have an eggshell texture.

• AFM the mask in the same areas which were AFM'ed previously to find the plated gold thickness. If the gold isn't thick enough, the mask can be placed back into the plating solution and more gold plated. If this is done, however, the surface tends to be rougher than if all the plating is done in a single run.

- If the plating is satisfactory, turn off the plating bath pump, the bath heater, and close the valve for the chiller loop.

- Check the pattern by inspection in the optical microscope and, if needed, the scanning electron microscope.

- Strip the resist by soaking in acetone for about 10 minutes then rinsing with acetone then methanol. Alternatively, the resist can be stripped by soaking the mask in hot NMP (~90°C). Plasma ash the mask for 10 minutes to remove any residual organics.

- Check for large particles on the mask and mesa. This is best done by scanning the fiber light (held perpendicular to the mask) across the membrane while looking at a wide angle for scattered light. The mask can also be scanned under the optical microscope (lowest-powered objective) while looking for scattered light and/or looking through the eyepieces. A dark-field objective is especially well-suited for detecting particles. The mask can also be brought in close proximity with another mask, as if a daughter exposure was being done, to check for particles with interference fringes and to see if there is a large wedge, which would indicate a particle on the mesa. To remove the particles, spin photoresist (PR) over the whole mask and expose the region around the particle for about 2 minutes in the optical microscope with maximum white-light intensity. Then develop the PR and use an eye dropper to put drops of gold etch on the openings in the resist to "lift off" the particles. The etch of the largest particles should be completed in about 20 minutes. For some particles it is necessary to also use Ti etch (TFT-type) heated to 50°C (make sure that the hot plate is well back in the hood and that only a minimal amount of Ti etchant is used, since TFT Ti etchant is HF-based!)

- For masks which are to be replicated in intimate contact, polyimide should be spun on. Details of polyimide spinning are given in Section A.5

A.5 Polyimide Coating of Masks

<u>EQUIPMENT NEEDED:</u>

Hood Space for Developing and Etching

Oven at Various Temperatures

Optical Microscope	Resist Spinner
Plasma Asher	OAI at 400 nm

- It's a good idea to run a silicon monitor wafer or two in parallel with the masks in this process to confirm various development and exposure times, particularly for the 20 μm grid pattern.
- If the plating base is to be removed in the alignment mark regions (for daughters), perform the following steps:
  - Spin Shipley 1813 photoresist at 4.7 krpm for 30 seconds. Bake for $\sim$20 minutes at 90°C.
  - Expose the alignment mark regions in the optical microscope with the white light source for 20 seconds with the 50X objective. The lamp should be turned up all the way and the diaphragm aperture should be opened fully. Use the field aperture to set the exposure size. It's a good idea to run a monitor wafer first. With many dies and several alignment marks per die, this can be quite time-consuming. These can also be exposed through the back of the membrane using the arc lamp source. This method maintains photoresist on the absorber in the alignment mark areas so there is no contrast reduction due to gold absorber etching.
  - Develop with Shipley CD-30 or Shipley 351 (1:5 351:DI) developer for about 30 seconds.

- Plasma ash the sample in He/O_2 for 6 seconds at about 50 W power. This step is absolutely critical to get etching in all areas at the same time with good uniformity. Because of the metal film below the resist, there is a node in the field intensity at the resist-substrate interface which prevents exposure and development there [6].
- Etch the gold in 10:1 dilute gold etchant for 20 seconds (10 seconds if using "thin-gold," (*i.e.* 1.8 nm-thick Au) plating base).
- Etch the Ti for 5 seconds in TFT Ti etchant which has been heated to 50°C. This etch is much more dangerous than the gold etch because it can etch laterally at an incredible speed and undercut everything, causing Moire alignment marks, for example, to fall over easily [6].
- Strip the PR in acetone and methanol.

- Spin on about 300 nm of PI, use 3:2 PI2610:NMP (*n.b.* what we call "NMP" is labeled on the bottle as 1-methyl-2-pyrrolidinone) at 3.6 krpm for 60 sec.

- In case electrostatic contact will ever be used with the mask, clean off the polyimide at the edges of the Pyrex ring with a Q-tip soaked in NMP to allow for electrical contact.

- If the polyimide does not need to be stripped in the alignment mark areas (*e.g.* for a mother mask), bake at 140°C for 30 minutes then ramp to 250°C for a total bake time of 60 minutes.

- If the polyimide should be stripped in the alignment mark areas (*e.g.* for daughter masks), do the following:
  - Bake at 180°C for 30 minutes.
  - Spin on Shipley 1813 photoresist at 4.7 krpm for 30 seconds. Bake at 90°C for 20 minutes.
  - Expose the alignment mark regions in the optical microscope with the white light source for 20 seconds as before.

- Develop the photoresist for 45 seconds with Shipley CD-30 developer or 5:1 Shipley 351 developer:DI. This will also etch the soft-baked (but uncured) PI. Confirm that the PI has fully etched by inspecting in the optical microscope.
- Strip the photoresist in acetone and methanol.
- Hardbake the polyimide at 250°C for 30 minutes.

- Spin on 1:3 PI:NMP at 3.1 krpm for 60 seconds to get 70 nm of polyimide.
- Bake at 180°C for 30 min. This temperature is critical.
- Spin on Shipley 1813 photoresist at 4.7 krpm and bake for 30 minutes at 90°C.
- Expose with a 100 μm-period grating flex mask (adequate for daughters to be exposed in contact with small chips) or a 20 μm-period grid mask (necessary for masks to be exposed in contact onto other masks or whole wafers). The 100 μm-period mask has much more process latitude; the 20 μm-period mask can sometimes be difficult to get to work. The optical mask is simply laid on the x-ray mask and exposed for 7.0 seconds in the OAI at 400 nm (for a Mimir power reading of 28 mW/cm^2).
- If clearing alignment marks, expose them in the optical microscope as before.
- Use an acetone-soaked swab to clear resist from the edge of the Pyrex ring.
- Develop using Shipley CD-30 developer or 5:1 Shipley 351 developer:DI for 45 seconds using continuous flow development. This is important to make sure that the developer solution which is saturated with resist gets carried away so that the etching of the PI is uniform across the mask. Etch the monitor before you do your mask, doing an interrupted development and watching the PI etching to determine a time for the mask development. The idea is to have the PR clear everywhere very quickly and to have the PI etch slowly (relative to the resist development). It's often easier to monitor the PI etch by looking at defects in the grating mask (joined squares, for instance).

- Once the development is done, strip the remaining resist using acetone and methanol.
- Bake at 200 °C for 1 hour.

A.6 Clearing Polyimide from a Mask or Part of a Mask

Polyimide can be cleared from a mask in the plasma asher. To remove the cured ~370 nm double layer takes quite a long time – 50 minutes at 200 W, 300 mTorr.

In addition to clearing PI from the whole mask, portions of the mask have been successfully cleared by placing a foil shield over the membrane. A hole was made in the foil in the region over the area which is to be cleared. This foil shield was mounted on an aluminum ring which had previously been used as a mask for shadow evaporation on a membrane (there are several such rings in the cleanroom). It therefore was the proper height to rest on the Pyrex part of the mask and provide a small gap between it and the mask membrane.

By mounting the foil on the back side of the ring and carefully tearing a hole and folding back the foil, the shield could be used with no danger to the mask membrane. In order for etching to occur, the hole had to be larger than the dark space or the plasma sheath for the conditions in the asher. A hole of ~5 mm on a side satisfactorily etched the polyimide from one die on the mask.

Appendix B

Device Fabrication

This Appendix contains information relevant to the fabrication of MODFET-type devices on GaAs/AlGaAs heterostructures. I have again chosen to break up the process into pieces which might be performed in one session.

Contents

B.1 Mesa Isolation **180**

B.2 Ohmic Contacts Processing **184**

B.3 Gate Processing **188**

B.4 Patching Holes and Fixing Shorts **194**

B.1 Mesa Isolation

EQUIPMENT NEEDED:

Hood Space for Solvent and Aqueous Processing	Spinner
OAI UV Source at 405 nm	Oven at 90°C

It's a good idea to run a GaAs monitor through this process to monitor etch rates before committing the heterostructure.

B.1.1 Solvent Cleaning of Sample and Surface Clean

- For solvent clean, use `TCA`, `ACE`, `METH` beakers and `DJC-SOLV` dipper.
- Prerinse beakers and dipper with solvents.
- Rinse sample in TCA.[1]
- Boil in TCA for 10 min (hotplate designation ~90°C). It's usually a good idea to place tweezers in the TCA to provide a nucleation point for boiling.
- Ultrasound in acetone for 10 min.
- Ultrasound in methanol for 10 min.
- Blow sample dry.
- Post-rinse beakers with solvent.
- Let beakers dry in hood.

[1]TCA (1,1,1 trichloroethane) was used in this work as a solvent for "degreasing." It is considered to be relatively "safe" and has few long-term side effects for low exposure levels. However, it is no longer available, as it is responsible for depletion of the ozone layer. Unfortunately, the recommended replacement, trichloroethylene (TCE) is a suspected carcinogen and must be treated much more carefully.

B.1.2 Spin and Bake Photoresist

- Spin and bake resist. Lithography on small (few-mm on a side) samples can be challenging. The buildup of resist on the corners of the chip can make optimizing exposure and development quite difficult. In addition, holding the small pieces without tweezers slipping and scratching can be an art unto itself.
  - Spin Shipley 1813 resist at 4500 rpm for 30s for a $\sim$1.2 μm-thick resist film. Use filtered glass syringe located in drawer below Next station.
  - Bake at 90°C for 30 minutes. Use quadrant bowl located near ovens.
  - The small samples often have a very thick resist layer on their back side due to resist being pulled into the vacuum channels in the chuck (which are not totally covered by the chip). This must be removed for good planarity of the chip to the alignment chuck. There are two ways to do this, the second being the preferred method:
    - Using a swab which has been dipped in acetone then dried slightly by pushing against a wipe, carefully swab the back of the chip to remove the built-up resist. Be careful not to let acetone wick around to the front of the chip or the resist there will dissolve away.
    - Flood expose the back of the chip (many can be done in parallel) in the OAI, using several times the typical clearing exposure (a minute or so is usually sufficient). A swab soaked in developer can then be used to clear the built-up resist without the danger of the resist film on the front being etched away.

B.1.3 Photolithography and Development

- Expose resist with the mesa-level mask for 4 seconds at 21 mW/cm^2 power (as measured on the Mimir meter) in the OAI contact aligner at 400 nm.
- Develop exposed resist for $\sim$20s with straight Shipley CD-30 developer (alternatively use Shipley Microposit Developer Concentrate mixed 1:1 with DI)

followed by 1 min rinse in DI. Use dipper labeled `DJC-AQ`. Use `351-1` beaker for developer and `351-2` beaker for DI. Transfer to running DI. CD-30 (which is the same as Microposit:DI 1:1) does not etch GaAs.

- Inspect in microscope.
- Continue developing in 10-15 second increments until pattern clears. Total development time should be ~45-60 seconds. Pay attention to corners where resist is thickest. It is often necessary to overdevelop the center of the die in order to get the corners to clear
- Spray mask with acetone/methanol.
- Blow dry mask.

B.1.4 Etch AlGaAs

- Run a GaAs monitor sample first to calibrate etch rate.
- Ammonium hydroxide/hydrogen peroxide etch:
  - Put 500 DI : 10 NH_4OH : 3 H_2O_2 in large RCA beaker (put in H_2O_2 last)
  - Mix well and put small amount into `ACID` beaker.
  - Etch ~17 seconds (~600-750 Å).
  - Rinse in DI for 2 min. Transfer directly to running DI water in sink if possible.
  - Blow dry.
- Inspect in microscope.
- Strip resist:
  - Spray with acetone.
  - Spray with methanol.

  - Blow dry.

- Inspect in microscope.
- Measure mesa height in Alpha-Step profilometer or the Linnik interferometer.

B.2 Ohmic Contacts Processing

EQUIPMENT NEEDED:

Hood Space for Solvent and Aqueous Processing — Spinner

OAI UV Source at 405 nm — Oven at 90°C

Evaporator for Ni, Au, Ge — Strip Heater for Anneal

Probe Station to Check Contact Resistances

B.2.1 Solvent Cleaning of Sample and Surface Clean

- For solvent clean, use `TCA`, `ACE`, `METH` beakers and `DJC-SOLV` dipper.
- Prerinse beakers and dipper with solvents.
- Rinse sample in TCA.
- Boil in TCA for 10 min (hotplate designation ~90°C). It's usually a good idea to place tweezers in the TCA to provide a nucleation point for boiling.
- Ultrasound in acetone for 10 min.
- Ultrasound in methanol for 10 min.
- Blow sample dry.
- Post-rinse beakers with solvent.
- Let beakers dry in hood.

B.2.2 Spin and Bake Photoresist

- Spin Shipley 1813 resist at 4500 rpm for 30s. Use filtered glass syringe located in drawer below Next station.
- Bake at 90°C for 30 minutes. Use quadrant bowl located near ovens.

- Clear built-up resist from the backside of the chip in one of the following ways:
  - Using a swab which has been dipped in acetone then dried slightly by pushing against a wipe, carefully swab the back of the chip to remove the built-up resist. Be careful not to let acetone wick around to the front, or the resist will dissolve away where you want it.
  - Flood expose the back of the chip (many can be done in parallel) in the OAI, using several times the typical clearing exposure (a minute or so is usually sufficient). A swab soaked in developer can then be used to clear the built-up resist without the danger of the resist film on the front being etched away.

B.2.3 Photolithography

- Expose resist with ohmic-level mask for 4 seconds at 21 mW/cm^2 power (as measured on the Mimir meter) in the OAI at 400 nm.

B.2.4 Evaporation of Ohmic Contact Metal

It's a good idea to run a bare Si monitor with the evaporation in case you might need to analyze the films later...

Do the following steps only if you can load directly into the evaporator following the NH_4OH rinse:

- Develop Sample:
  - Develop exposed resist for ~20s with straight Shipley CD-30 developer. Use dipper labeled `DJC-AQ`. (alternatively use Shipley Microposit Developer Concentrate mixed 1:1 with DI) followed by 1 min rinse in DI. Use `351-1` beaker for developer and `351-2` beaker for DI. Transfer to running DI.
  - Inspect in microscope.
  - Continue developing in 10-15 second increments until pattern clears. Total development time should be ~45-60 seconds. Pay attention to corners

where resist is thickest. It is often necessary to overdevelop the center of the die in order to get the corners to clear

  - Spray mask with acetone/methanol.
  - Blow dry mask.

- UV ozone sample for 15s. Run UV ozone a minute or so before placing sample inside.
- Rinse in DI water for 1 min.
- Rinse in 5% NH_4OH for 15 seconds. Use `SEM` beaker. (*n.b.* 5% NH_4OH is mixed 5:1 DI:NH_4OH(30%)).
- Remove from NH_4OH and immediately blow dry.
- Thermally evaporate ohmic contact metal. The evaporation sequence (from the wafer up) is:
  - 5 nm Ni
  - 25 nm Ge
  - 50 nm Au
  - 10 nm Ni
  - 50 nm Au.
- Liftoff by soaking for ~10 minutes in acetone in `L1` beaker. Then place beaker in ultrasound for ~10 seconds for full liftoff. Do not use a dipper for this step. Quickly remove sample from `L1` beaker and transfer to `L2` beaker (also filled with acetone), spraying with acetone during the transfer. Inspect the chip, in acetone, in the stereo microscope. If the liftoff looks successful, remove and squirt with acetone then methanol, then blow dry.
- Place sample directly into vacuum container for transport.

B.2.5 Annealing Contacts

- Transport the sample to the second floor lab in Bldg 13 in a vacuum container.

- Sinter in strip heater in forming gas (95% N_2/5% H_2) at 425°C for 30 seconds. Cap sample with a clean piece of GaAs which is placed face-down on the sample to help prevent Ga out-diffusion during sintering.

B.2.6 Check Contact Resistances

- Check resistances in 5th floor lab in building 39. They should be $< 1\,\Omega$-mm. Verify sheet resistance using the 4-point van der Pauw structure. Details of TLM (which stands for either transmission line model – a poor designation, or transfer length model) and van der Pauw measurement and characterization techniques are given in Schroder [155]. The original van der Pauw reference is from *Phillips Research Reports* [156].

B.3 Gate Processing

<u>EQUIPMENT NEEDED:</u>

Hood Space for Solvent and Aqueous Processing — Spinner

Oven at 180°C — Evaporator for Ti, Au

X-ray Head 4 with Small-Sample Vacuum Alignment Chuck

It's a good idea to run a bare Si monitor with the evaporation in case you might need to analyze the films later...

B.3.1 Solvent Cleaning of Sample and Surface Clean

- For solvent clean, use `TCA`, `ACE`, `METH` beakers and `DJC-SOLV` dipper.
- Prerinse beakers and dipper with solvents.
- Rinse sample in TCA.
- Boil in TCA for 10 min (hotplate designation ~90°C). It's usually a good idea to place tweezers in the TCA to provide a nucleation point for boiling.
- Ultrasound in acetone for 10 min.
- Ultrasound in methanol for 10 min.
- Blow sample dry.
- Post-rinse beakers with solvent.
- Let beakers dry in hood.

B.3.2 Spin and Bake PMMA

- Spin PMMA (950K, 3% in chlorobenzene) at 2.7 krpm for 60 seconds.
- Bake at 180°C for one hour.

- Clear the backside of the chip by flood exposing for 30 minutes at 220 nm UV in the OAI and using a swab dipped in 2:3 MIBK:IPA to wipe away the built-up PMMA.

B.3.3 Small Sample Vacuum Contact Exposure

- Attach small chip to 2" silicon wafer:
  - Place a drop of polyvinyl alcohol (PVA) in the center of the wafer using a glass pipette (no bulb needed, capillary action of the pipette will deposit enough PVA).
  - Place the chip in the PVA.
  - Bake at 90°C for 20 minutes.
- Remove the top plate and the mask holder from the Head 5 aligner.
- Install the small-chip vacuum fixture:
  - Place chip holder on pin chuck with rubber gasket in-between.
  - Replace top plate, using 1/2" Swagelock spacers.
- Place 2" wafer with chip onto vacuum gasket on alignment stage.
- Make sure stage is fully lowered and the mask vacuum line is opened to atmosphere.
- Place mask holder with gasket onto top plate.
- Mount the two mask leveling clamps which are to the back of the fixture.
- Check chip mounting for vacuum leaks:
  - Place a *dummy* x-ray mask onto the mask holder, place the plexiglass spacer ring on top of it, and tighten the mask clamps.

  - Place the third mask leveling clamp and tighten it down, watching to make sure there is still a gap between the rubber gasket on the mask holder and the sealing plate.
  - Raise the sample and use the leveling clamps to level the mask with respect to the chip.
  - Raise the sealing plate with the three thumbscrews, and seal the vacuum line (so that it's neither pulling vacuum nor open to atmosphere). This checks for vacuum leaks from the chip vacuum. We want to be able to independently control the chip and the mask vacuum.
  - If the dummy mask deflects and there is a leak, lower the sealing plate, remove the mask, reposition the 2" wafer and the gasket to which it seals, re-level, and check again for a leak.

- When there are no leaks, re-open the mask vacuum line to atmosphere, lower the sealing plate, replace the dummy mask with the real mask, level the mask to the substrate, and align the two.
- Raise the sealing plate with the thumbscrews to create a vacuum seal. Sometimes raising the wafer stage will help strengthen the seal. Monitor the alignment as things are being raised and moved, as it sometimes shifts. Small misalignments can be corrected even with the sealing plate raised.
- When the alignment and seal are satisfactory, open the needle valve on the vacuum control system to 2.0, making sure the vacuum regulator reads zero.
- Turn the tee valve from atmosphere to vacuum, and slowly dial up the vacuum regulator.
- Move the alignment stage into the helium box and expose.
- When the exposure is over, dial down the mask vacuum, lower the stage with the stage vacuum still on (this should pull the chip from the mask if it is stuck

by van der Waals forces), remove the mask, and remove the 2" wafer with the chip.

B.3.4 Removal of Small Chip from 2" Wafer

Because a drop of PVA was used to glue to chip to the wafer, special care must be taken to remove the large volume of PVA (relative to the volume which is left when something is spun on a wafer) from the chip. If these steps are not followed, PVA residue remains on the chip and can interfere with development and subsequent processing.

- Hold wafer and chip (still attached) in `DJC-AQ` dipper under running DI water until the two separate.
- While the water is still running, remove the 2" wafer and keep the chip in the running water for two more minutes.
- Place the chip and the dipper into the `DI-1` beaker and soak in DI for two minutes.
- Transfer the chip *only* into the `DI-2` beaker and soak in DI for five minutes more.
- Remove the chip and blow dry.

B.3.5 Development and Schottky Gate Evaporation

Perform the following steps only if you will be able to put the chip directly into the evaporator after the NH_4OH rinse:

- Use 2:3 MIBK:IPA at 21°C as the developer.
- Do an interrupted development, where the development occurs in short time intervals and the resist thickness at a step edge in the pattern is checked in the AFM or the Alpha-Step. Try to find the clearing time for large features. Rinse with IPA and blow dry after each interval.

- Develop 10% beyond the large feature clearing time to ensure that little PMMA residue remains.
- Rinse with IPA and blow dry when finished.
- Remove any resist residue in the RIE:
  - Run the process `dsclean` in the PlasmaTherm RIE with the 3" or 4" dummy wafer and fixturing in place to condition the chamber for a successful plasma strike for the descum of the chip.
  - Vent the chamber, place the chip on the dummy wafer, and evacuate the chamber. Allow it to pump out to 7×10^{-5} Torr.
  - Manually descum with the following parameters:
    * He/O_2 10/2.5 sccm.
    * 35 mTorr pressure
    * Plasma controlled at 20 Volts (~14 W power).
    * 15 seconds descum.
  - Remove chip from chamber.
- Clean surface oxides:
  - Rinse in DI water for 1 min.
  - Rinse in 5% NH_4OH for 15 seconds. Use `SEM` beaker. (*n.b.* 5% NH_4OH is mixed 5:1 DI:NH_4OH(30%)).
  - Remove from NH_4OH and immediately blow dry.
- Immediately load sample into the electron-beam evaporator.
- E-beam evaporate Schottky metal (100 Å Ti/600 Å Au).
- Liftoff the gate pattern:[2]

[2]NMP heated to ~ 90°C is much more effective at stripping PMMA than acetone is. If small particle deposition is problematic, however, acetone is preferable because it can be sprayed, which helps carry away particles.

- Begin with an acetone spray (one bottle full) using the atomizer in the outer lab (no smoking, please!). The chip can be mounted by a clip to the stainless steel fixture which is located in the outer lab hoods.
- When the atomizer is almost empty, *before the sample has a chance to dry*, place the chip and the holder into acetone. Separate the chip from the holder while in the acetone and remove the holder.
- Let the chip soak in acetone for several hours.
- Remove the chip from the acetone, spraying with acetone while doing so to remove the bulk of the metal. Immediately place the chip into another acetone beaker.
- Inspect the chip, in acetone, in the stereo microscope. If the liftoff looks successful, remove and squirt with acetone then methanol, then blow dry.

B.4 Patching Holes and Fixing Shorts

Occasionally, because of tweezer slips across the small chips, patterned metal lines are severed or resist is scratched, causing opens or shorts in the lifted-off metal pattern, respectively. If this happens on a device gate pattern on a mesa, there is little that can be done to remedy the situation. Usually, however, the shorts or opens occur on the metal traces which go from the bonding pads to the devices, where features sizes are relatively large. In this case the problem can often be remedied.

The method used to fix shorts between traces outside the mesa has been to use the probe station both as a clamp to hold the chip and as a scalpel to scrape away extra metal. Since the traces are often long, straight lines, the unidirectional action of the probe is well-suited to this purpose. By placing one or two probe tips away from any pattern to clamp the chip to the stage and using a third tip as a scraper, the short can often be removed (and other probes can be used to monitor the resistance between the lines in real time).

To replace open circuits in the traces caused by slipped tweezers (or poorly-manipulated probes), photoresist is spun on the wafer and the areas where metal is desired are exposed with the optical microscope. The resist is developed, the sample is UV-ozoned, and metal is evaporated and lifted off to patch and repair the chip.

Appendix C

Electronics and Cryogenic Design for Device Measurements

This appendix details the cryogenic system and the electronics used for the measurements described in Chapter 7. Modifications which were made to these systems are detailed and the operation of the ^{3}He system is given in the "recipe" format adopted in Appendices A and B. The last section of the appendix describes some of the problems encountered in measuring the coupled quantum dot structure.

C.1 Cryogenic and Electronic System Design

The cryogenic and the electronic systems are somewhat coupled due to the fact that electron transport to a sample at cryogenic temperatures also involves thermal transport, and the device properties are very temperature-dependent. Noise in a device at cryogenic temperatures can be detrimental not only from its effect on the signal, but also because of resistive dissipation in the sample, which can warm it and alter its electrical characteristics. Thermal radiation can also affect device electrical properties in some cases [157].

Even though the two systems are coupled, they are described separately in the sections to follow. The general principles of the method of operation of each system are described and then specific details are given of the improvements made to enhance

system performance and ease of use.

C.1.1 Cryogenic System

A schematic of the cryogenic system, an Oxford Instruments Heliox insertion probe, is shown in Figure C-1. A detailed procedure for probe operation is given in Section C.2. A good general reference for cryogenic experimental techniques is Richardson and Smith [158].

An overview of the probe operation follows: The sample is mounted on the sample block which is in an evacuated space (labeled the "inner vacuum can" (IVC)) which has been back-filled with a small amount of helium (a few mTorr) for heat exchange. The probe is placed in liquid nitrogen to cool the whole system to 77 K. Once this is accomplished, the probe is placed into liquid helium, which cools the system to 4.2 K.

To cool the sample below 4.2 K, the helium exchange gas is removed either by a charcoal sorb in the IVC (not shown) or by evacuating the IVC with a high-vacuum pump once the system is at 4.2 K. If this is not done, the helium exchange gas can transfer thermal energy from the walls of the IVC, which are at 4.2 K.

Once the exchange gas is removed, the 1 K pot is evacuated and the needle valve for the 1 K pot drinking straw is opened. This allows a flow of ^{4}He through the 1 K pot. Since the boiling point of liquid helium is reduced at low pressures, the 1 K pot will cool to 1.2 K.

The Heliox probe contains a closed ^{3}He system. As the 1 K pot cools, the ^{3}He sorb pump is heated to drive off gaseous ^{3}He, which condenses to a liquid at the place where the closed ^{3}He system makes thermal contact to the 1 K pot. The liquid ^{3}He collects in the 0.3 K pot and absorbs thermal energy from the sample block, cooling it to $\sim$1.2 K.

Once the ^{3}He has condensed, the ^{3}He sorb pump is cooled to 4.2 K by thermal coupling to the sorb isolation system, and it begins to lower the pressure in the closed ^{3}He space. The ^{3}He bath can achieve temperatures of $\sim$300 mK in this manner, and heat is extracted from the sample block and the sample to cool them to this temperature.

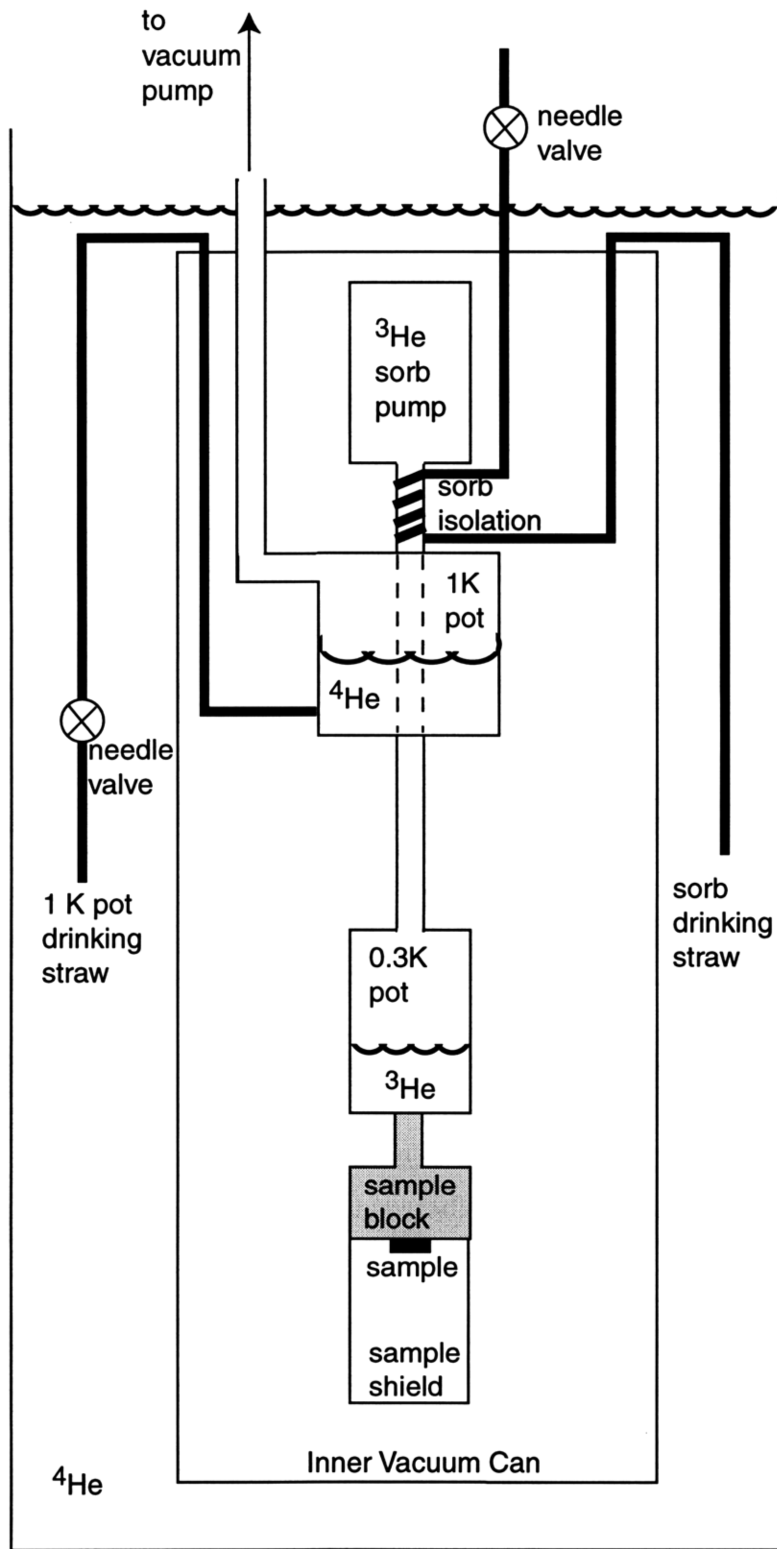


Figure C-1: Schematic of Oxford Instruments Heliox insertion probe for 300 mK device measurements. Figure after Burkhardt [6].

Modifications were made to the sample mounting system to improve ease of use and cryogenic performance. The sample is attached with silver paint to a 44-pin header and bond pads on the sample are connected with wire bonds to pads on the header. The header fits into a socket at the bottom of the probe which has electrical connections to the top of the probe. When under vacuum, the sample is cooled only by conduction through the substrate and through the leads.

Thermal connection between the header and the stage is made by placing a small copper spacer between them. The chip carrier was previously taped into place and a sample shield was placed over the chip. Two problems occured with this arrangement – mounting the chip carrier and the copper spacer was often quite difficult, and the sample shield was attached at only one point to the stage, since space was required to run the wires to the sample. This meant that the shield was not in good thermal contact to the sample block, possibly allowing it to warm well above 300 mK and subsequently warm the sample with thermal radiation.

The sample mount was completely redesigned. Instead of using tape to mount the chip carrier to the stage, the new stage has tapped holes to accept bolts which pass through the socket, rigidly connecting the chip carrier assembly and the copper spacer to the stage. Holes drilled in the stage allow the wires to pass through it, which improves their heat sinking and reduces heat leaks from higher-temperature stages.

In addition, the sample shield can now be well-connected to the stage, since no space is required for wires to pass between them. The new shield snaps onto the stage, making good electrical and thermal contact. It should be noted that for good contact, the shield and the stage should be free from tarnish. Since the shield is regularly handled, often with gloves which have vacuum grease on them (which seems to accelerate tarnishing), it should be cleaned when it appears dirty. Careful experiments have shown that Heinz ketchup, followed by a detergent, does an excellent job of removing the tarnish and restoring a bright shine to the copper.

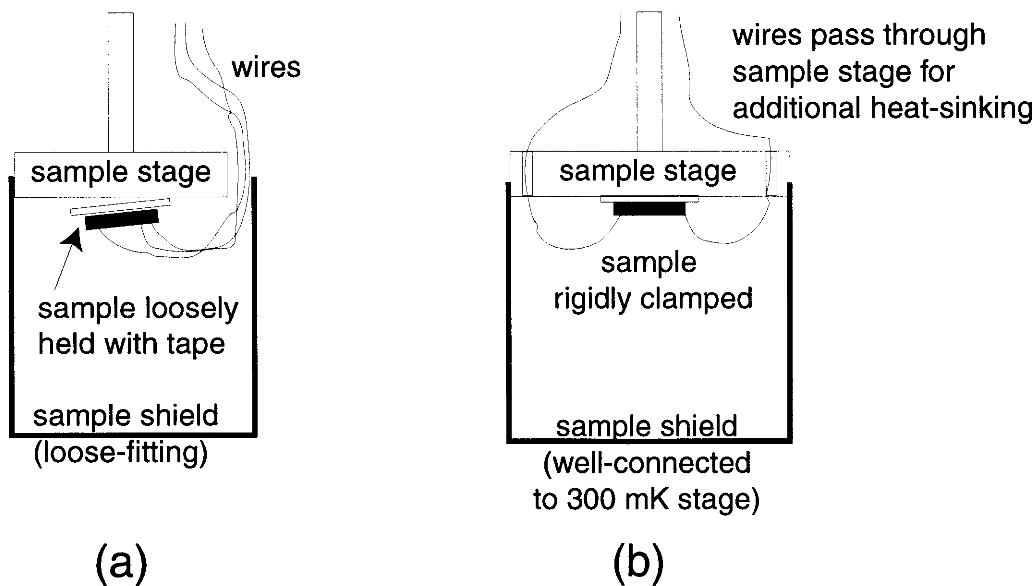


Figure C-2: Schematic of modifications made to the sample mount for improved ease of use and performance. (a) The old mounting setup. Placing the sample on the stage was awkward and good contact between the chip carrier and the cold stage was not assured. Because wires had to run between the sample shield and the cold stage, the shield was not well-connected to the stage. (b) The modified setup. The chip carrier is now rigidly clamped to the stage, improving ease of mounting and thermal connection. Wires now run through the stage, improving heat-sinking and allowing for a sample shield which is in good thermal and electronic contact to the stage.

C.1.2 Design of Measurement Electronics

The measurement electronics were also redesigned and a new system was built to minimize noise and to allow for fully-automated gate control and measurement. A very good general reference for noise in electronic systems is Ott [159]. There are also chapters devoted to low-noise measurements in cryogenic systems (including one with reference to the non-cryogenic but interesting application of designing and implementing circuitry to run the shark in the movie "Jaws") in Richardson and Smith [158].

Electric fields and magnetic fields can couple to a circuit and induce voltages and currents which are seen as noise in the signal, particularly for the μV, pA signals expected for quantum dots. As discussed above, for device measurements at cryogenic temperatures, this noise can also produce additional ohmic heating in the sample which can raise its temperature and change its electrical characteristics. Reducing electronic noise is therefore critical for quantum dot measurement.

Placing the circuit in a grounded box will eliminate a large part of the induced noise. High-frequency electromagnetic waves will reflect off the box, and low-frequency or DC electric fields will terminate on the box. Low-frequency magnetic fields, however, can still couple to the circuit. To minimize the effects of this coupling, large loops which can couple a large $d\Phi/dt$ into the circuit (where Φ is magnetic flux) should be eliminated. The circuit should be connected to the grounded box only at one point, preferably the input of the amplifier. Any long current supply/return paths should be on twisted pairs so the a positive $d\Phi/dt$ in one small twisted loop is cancelled by a negative $d\Phi/dt$ in a neighboring loop.

Digital measurement equipment can also introduce noise into the system. Several pieces of digital equipment in the lab were found to have spiky noise at one to several MHz frequency on their outputs or even their inputs. This was checked by using an analog oscilloscope, connecting its ground to a good earth ground, and connecting its positive input to the instrument in question. The input of one electrometer in the lab (whose manufacturer shall remain nameless) was found to have voltage spikes of $\sim$1 V on both the center conductor and the ground of its BNC input! Since the

quantum dot devices were being measured with a lock-in amplifier, the voltage noise at frequencies far from the lock-in frequency (typically 77 Hz or so) would in theory have a minimal effect on most measurements, but 1 V spikes could heat the sample significantly and modify the gate bias voltages relative to the 2DEG.

The old measurement setup used some filtering and had proper grounding, but did not use shielded twisted pairs in all places where cables were run, and left many pieces of digital equipment connected to the circuit with no filtering. The measurement setup was redesigned with supply and return currents traveling from the source, through a cable to the probe, and through the probe to the sample on the same twisted pair. In addition, low-pass "Pi" filters (80 dB attenuation at frequencies above 100 kHz) were added to the signals and grounds of all pieces of digital equipment which were attached to the device.

A schematic of the measurement circuit is shown in Figure C-3. The shield, which is connected to earth ground, is indicated with a heavy line. The output of the EG&G Princeton Applied Research 5210 lock-in amplifier was found to have its ground hard-wired to the case ground. In order to allow for some choice as to where to ground the circuit to the shield, a transformer was used to decouple the input ground from the circuit. A 10000:1 resistive divider (not pictured) then reduced the voltage to the required 10-100 μV range. The transformer could be switched in and out or replaced with an optical isolator circuit.

Current passes down into the dewar, through the sample, out of the dewar (on the same twisted pair), through a current-to-voltage converter/amplifier (either one of two home-built units which consist of an op-amp with a 10^8 or $10^6\Omega$ resistor or an Ithaco 1211 current preamplifier), then back to the transformer to complete the circuit.

The gate bias, which is controlled by a D/A board from a PC, is referenced to half of the sample input voltage to allow for symmetric gate biasing at relatively large dot bias voltages (for I-V measurements, for instance). The circuit ground can be chosen at several places by switches on the panels of the instruments.

The new measurement setup worked well for measuring QPC and fine-line tunnel

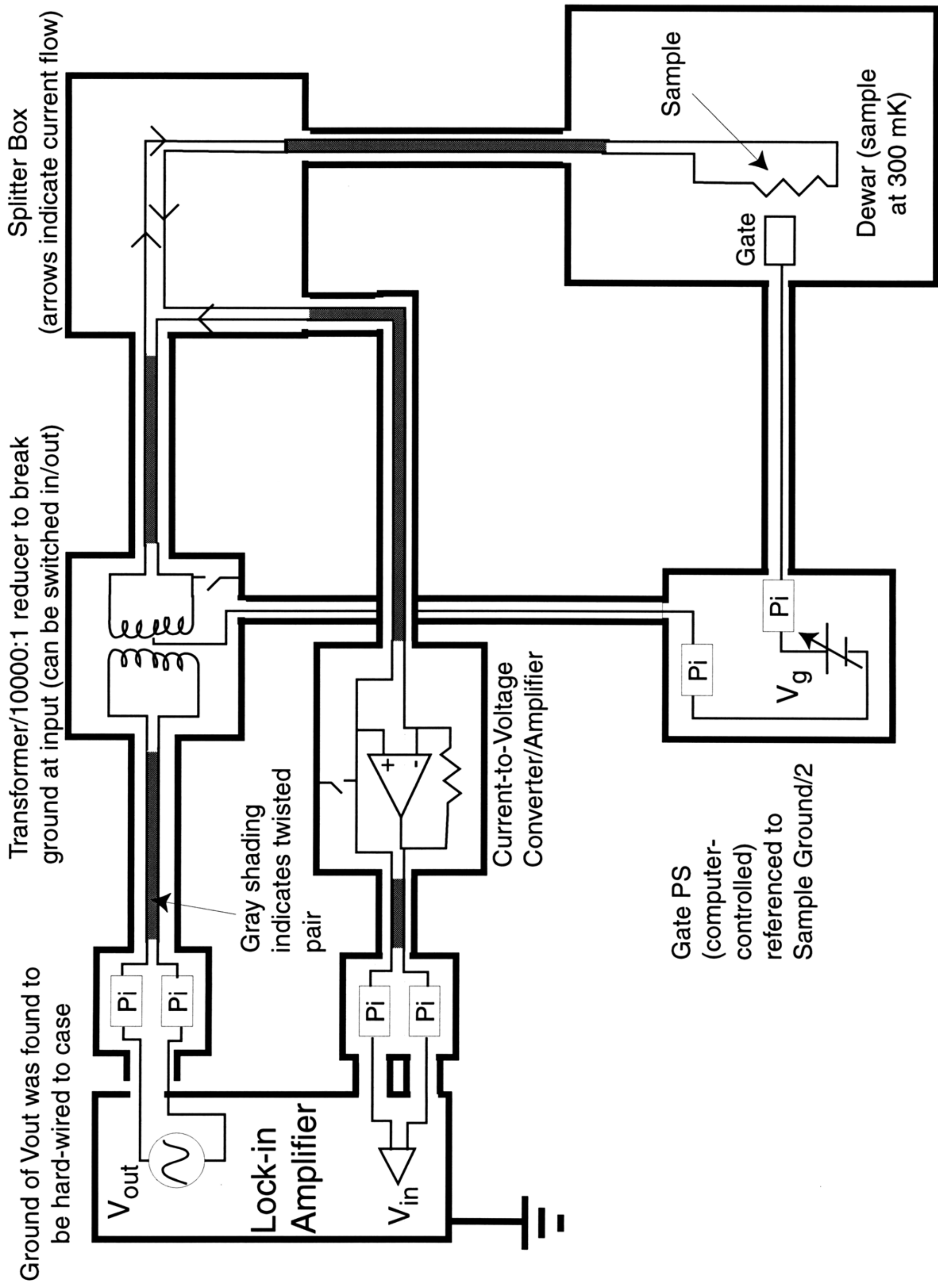


Figure C-3: Schematic of measurement circuit designed for measuring QPCs, tunnel barriers, and quantum dot devices.

barrier characteristics such as the ones shown in Chapter 7. It was used to measure electron-beam fabricated quantum dot devices by Sokolinski [160] and its noise characteristics were no worse than the previous setup as indicated by comparison of data from that reference with Kumar [24] and Burkhardt [6]. It should be noted that none of these measurements really pushed the noise floor of the measurement. Test measurements showed a factor of about two improvement of noise when the current supply and return were on the same twisted pair on the probe versus when they were on separate pairs. The main advantage of the setup was the automation of gate voltage control and the capability to perform many sweeps at one sitting with no human intervention.

C.2 ^{3}He Probe Operation

Cool to 77 K

- Wear vinyl or latex gloves when handling interior probe components.
- Vent probe and remove inner vacuum can (IVC) using sliding hammer. Place IVC in a safe place to prevent it from falling, getting dirty, etc. The vacuum seal relies on two tapered metallic surfaces making good contact, so be very careful not to scratch either.
- Mount sample
  - Mount the chip carrier into the socket and plug the connectors from the sample block to the chip carrier together.
  - Place a copper spacer on the back of the socket using a thin film of Apiezon vacuum grease on both sides of the spacer.
  - Screw the sample assembly onto the sample stage, checking that the spacer is properly seated. DO NOT OVERTIGHTEN!
  - Place the sample shield onto the stage. It should click into place.
- Replace vacuum can.
  - Clean the tapered surfaces on the can and the probe with a clean wipe.
  - Apply silicone grease liberally to the probe tapered surface (use a clean wipe to get grease in the area behind the drinking straws), and a thin film of grease to the can tapered surface.
  - Place a paper cylinder around the wires to prevent thermal shorts to the IVC.
  - Carefully slide the IVC onto the probe. It need only be pushed gently up to the tapered seal.
  - Pump out the IVC with the roughing pump to ~250-300 mTorr.

- Tape the drinking straws to the outside of the IVC to prevent them from hanging up on the dewar.

- Bleed in a few Torr of He exchange gas to the IVC. Pinch off a small section of helium-filled hose attached to the IVC pumping port and use your other hand to very slowly open the IVC vacuum valve, monitoring the pressure. If you bleed in too much, pump out the IVC and try again. When done, place a blank over the vacuum port to prevent accidental venting of the IVC.

- Connect the He lines to the 1 K pot and sorb isolator connections at the top of the probe. Pressurize the line to a few psi, and open the Speedivalve and the needle valve on the 1 K pot, and the needle valve on the sorb isolator. Check to make sure that He is flowing from the bottom ends of both drinking straws (you can cover them with your fingertips and listen for the flow to make sure).

- Carry the probe over to the probe stand and LN2 dewar. Move carefully to make sure the helium lines don't get caught. Slowly lower the probe into the LN2. There should be lots of bubbling due to the warm He gas which is flowing. (Now would not be the time to be wearing sandles, as the LN2 can splash out of the dewar).

- When the probe is fully lowered into the LN2, close the needle valves first, then the 1 K pot Speedivalve, then the helium valve at the wall. Tighten the clamps which hold the probe upright.

- Leave the probe in the LN2 for an hour or so, checking the temperature with the high-temperature sensor on the temperature controller. The high-temperature sensor is connected through cable B, and is connected or disconnected through one of the switch boxes located on top of the temperature controller.

- The probe can be left indefinitely in LN2. Be sure to top off the nitrogen level so that it doesn't fall below the level of the drinking straws, otherwise water could freeze in them.

Cool to 4 K

- Disconnect the high-temperature sensor (use the switch on the box located on top of the temperature controller).
- Push the Leitish-style mounting ring on the probe shaft down all the way so that the probe will not initially drop too low in the He storage dewar.
- Disconnect thermometer and pressure sensor cables. Resume He gas flow through the drinking straws.
- Transfer the probe to the liquid He storage dewar, being careful not to tangle the gaseous He lines. It's often helpful to have another person around for this step.
- Once the probe is in the He dewar, close off the gaseous He flow as before – needle valves, then Speedivalve, then wall valve. Clamp the Leitish-style flange and close the dewar vent valve.
- Slowly lower the probe into the dewar, pausing when the dewar safety valve pops open to allow pressurized gas to escape.
- Monitor the temperature at the sample. It should reach 4.2 K in an hour or so.
- Monitor the pressure in the IVC. If the charcoal IVC sorb is in place it should drop to below zero (the needle should be pinned – this is not a new physical phenomena, just a mis-calibrated gauge). If it does not, or the IVC sorb is not in place, use a high-vacuum pump to evacuate the IVC. Make sure the pumping line is evacuated before opening the IVC can valve to it, otherwise you'll get pump oil and water vapor coating your sample and everything else in the IVC.

Cool to 300 mK

- Connect the rotary pump to the 1 K pot line and evacuate the line. After a couple of minutes, open up the 1 K pot Speedivalve and pump out the 1 K pot line. At this point the needle valve should be closed.
- The temperature sensors are connected as follows: Sensor 1 is the sorb temperature, sensor 2 is switched between the ^{3}He pot and the 1 K pot temperature, and sensor 3 is the high-temperature sensor (which should be disconnected at this point).
- Cool to 1 K and condense ^{3}He:
  - Set the sorb temperature set point to 45 K. Do not heat the sorb yet.
  - Monitor the pressure and temperature in the 1 K pot.
  - Open the 1 K pot needle valve all the way for a few seconds. The pressure should shoot up.
  - Close the 1 K pot needle valve to pump on the liquid He which has collected in the pot. The pressure and temperature should start to drop.
  - Open the sorb isolator needle valve a turn or so.
  - Turn on the sorb heater (by pressing `auto`).
  - Now the balancing act begins:
    * The 1 K pot needle valve should be opened just enough to maintain a steady 1 K pot pressure (and therefore temperature). This is typically just a few degrees open.
    * The sorb isolator flow should be maintained such that a heater voltage of 5.5–7 V will maintain the sorb at 45 K.
    * Adjusting either needle valve will affect the other setting, so the 1 K pot temperature and pressure, the sample temperature, the sorb temperature, and the sorb heater voltage should be constantly monitored.

- Once the 1 K pot temperature and the sample block temperature have stabilized, continue to condense for another ~10 minutes to ensure most of the ^{3}He has condensed.

- Cool to 300 mK:
  - Turn off the sorb heater by manually lowering the heater voltage on the temperature controller.
  - Leave the sorb flow the same. This will cool slowly and conserve ^{3}He for a longer hold time.
  - When the sorb temperature reaches ~20 K, it should start to pump on the ^{3}He and the sample stage temperature should start to drop.
  - Continue to monitor the 1 K pot pressure, and tweak the needle valve if necessary.
  - Once the sorb has cooled to 4 K, let it stay for an hour or so then close the sorb needle valve.

- When the run is finished, close the 1 K pot needle valve and allow the 1 K pot to be pumped out for a few minutes. Then close the 1 K pot Speedivalve and turn off the pump. From this configuration it is relatively easy to condense again.

C.3 Device Measurement Difficulties

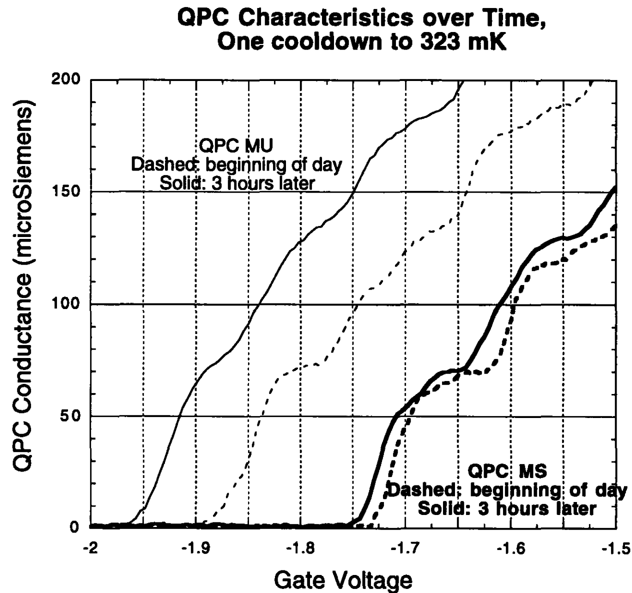
As mentioned in Chapter 7, QPC and tunnel barrier measurements were not consistent with time. This instability, along with other problems, prevented successful measurement of the full coupled quantum dot structure. This section will present some of the measurements which were taken in an attempt to understand what was happening with the device.

The variability of QPC and tunnel barrier pinchoff characteristics is shown in Figure C-4. During a given cooldown, or even over the course of a day, the voltage at which a given QPC gate pair would pinch off would tend to drift more and more negative. For the tunnel barrier, the pinchoff characteristic would tend to become more and more QPC-like, with steps developing at the quantum of conductance.

The reverse-bias leakage current of the Schottky gates also changed with time. Problems had occured previously with Schottky gate leakage, and the process was optimized to minimize the reverse bias leakage and to improve consistency among gates on a chip. Even with this optimization, however, Schottky gates which initially showed very little reverse-bias leakage at cryogenic temperatures would become more leaky over time, experiencing reverse-bias breakdown at voltages which were higher than those required to operate the device. An example of a Schottky gate measurement in two different cooldowns (with a thermal cycle to 300 K between) is shown in the top half of Figure C-5.

The bottom half of Figure C-5 shows how the reverse bias current of the leaky Schottky gate changes over the course of minutes. The current is plotted on a linear scale. The current of $\sim$15 nA seems to switch between a few different states in an abrupt manner. This telegraph noise creates a modulation which is on the order of 10% of the total current, so this is quite a large effect.

These measurements would seem to point to charge motion in the substrate. At cryogenic temperatures, a layer of partially-ionized dopants lies beneath the Schottky gates in the heterostructure. Over time the distribution of the dopants should reach an equilibrium configuration, although if the activation energy for motion from one



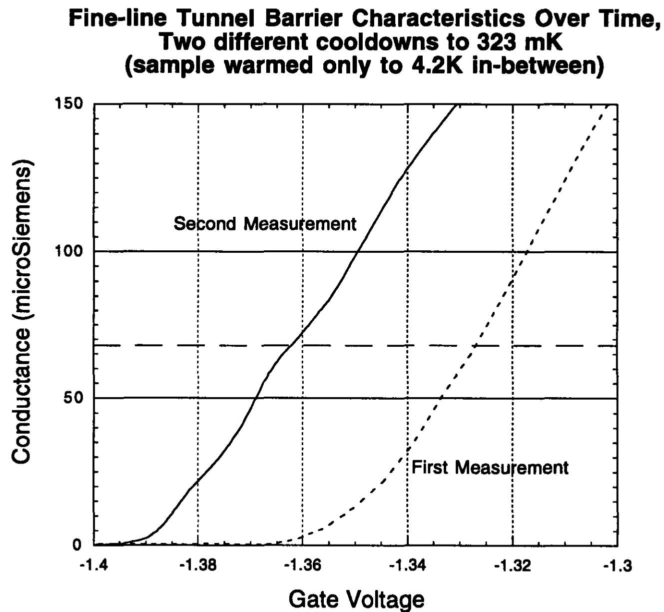


Figure C-4: Pinchoff characteristics of QPCs and tunnel barrier showing change in pinchoff voltages and nature of tunnel barrier conductance with time. Top: QPC drift in one cooldown to 323 mK. Bottom: Change in nature of fine-line tunnel barrier conductance between two cooldowns with thermal cycle up to 4.2 K in-between. Note the step appearing at the quantum of conductance (indicated by dashed line) in the second measurement.

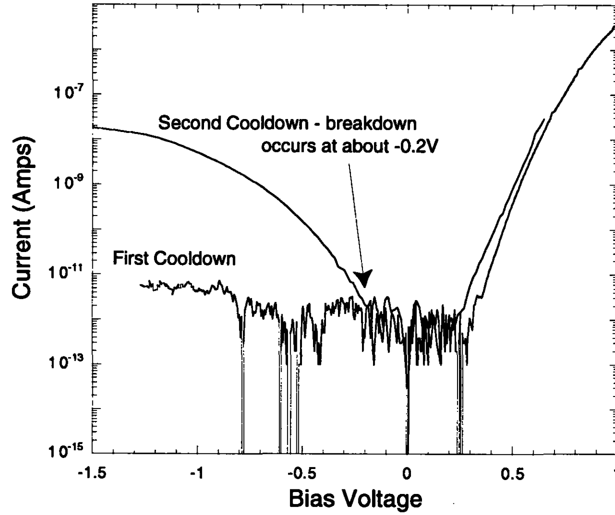
donor site to the next is large, this equilibrium state could take some time to be reached at low temperatures. A biased device will have a different equilibrium donor configuration than a non-biased device.

In Figure C-4, the QPC pinchoff voltages are seen to drift to a more negative value over the course of a few hours. Since measurements were being made during this time, the gates were biased, and is it quite likely that the negative bias could have moved electrons from un-ionized donors out from underneath the gates, more effectively ionizing the donors there. If there are more ionized donors under the gates, then there should also be more electrons under the gates, requiring a larger negative voltage to fully deplete them.

Reverse-bias leakage current in Schottky diodes is typically due to avalanche breakdown, where a charge is accelerated such that it gains enough energy to ionize an electron-hole pair and these charges similarly generate more pairs, leading to a large current. Avalanche breakdown will occur when a threshold electric field is exceeded. In a reverse-biased Schottky diode on a GaAs/AlGaAs MODFET structure such as the ones used in this experiment, more complete ionization of donors will create a larger electric field in the material, as can be seen by examining the band diagram, Figure 2-4. The reduction of reverse-bias breakdown threshold voltage could be explained by more complete ionization underneath Schottky gates which have been biased.

The telegraph noise shown in the bottom of Figure C-5 could also potentially be explained by a charge switching between two donor sites. If the switching charge were to modulate the field in a particular region which is close to the critical field for breakdown, a current channel might be expected to open and close as the charge switches between the two locations. The large percentage change in the current is a bit hard to understand with this explanation, but it is conceivable that there could be something like 10 channels which are opened underneath the gate. If each channel relied on a group of charges moving away from a particular point to increase the local field underneath the gate, one moving charge could open or close a particular channel.

The change of QPC and tunnel barrier characteristics with time and the unreli-



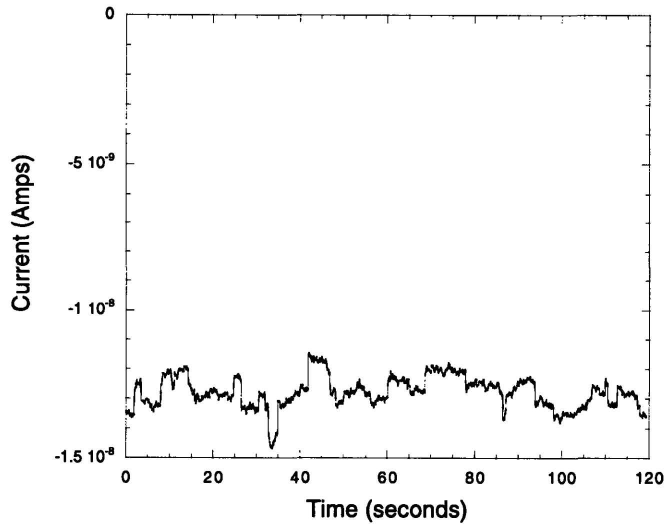


Figure C-5: Measurements of Schottky barrier characteristics at 4.2 K. Top: The same Schottky gate measured before and after a thermal cycle to room temperature. The reverse breakdown occurs at a reverse bias which would be necessary to operate a quantum dot. Bottom: Reverse-bias leakage versus time for the leaky measurement from the top graph. There appears to be switching or telegraph noise in the data.

ability of the Schottky gates prevented measurements from being made on quantum dots. It is possible that the heterostructure was to blame for these problems. Although the dies were taken from wafers which had previously been used successfully for measurements by Arvind Kumar, the wafers were several years old by the time they were processed for these devices. It is also possible that age has nothing to do with the change in behavior of the material – variation across the wafer could be the cause, although this is unlikely.

The most important indicator that something had happened to the material was the fact that Kumar's measurements were done simply by cooling the sample. For these measurements, the 2DEG sheet density was very low (as indicated by QPC pinchoffs less than -0.2 V) unless an LED was used at cryogenic temperatures to free more electrons from the donors.

Problems with instability of QPCs, Schottky reverse bias breakdown, and telegraph noise have been experienced by other researchers.[1] For the coupled quantum dot device to be measured, it would be advisable to get new material and carefully characterize test FET, Schottky, and Hall bar structures on it in parallel with devices. This combination of test structures should allow a better understanding of any problems which might occur.

[1] David Goldhaber-Gordon at MIT and Carol Livermore at Harvard have both seen similar problems with their devices. The Harvard group's experience is that unless the 2DEG is freshly-grown it will not produce good devices. They suspect that oxidation of the aluminum in the heterostructure is responsible. Livermore has seen devices with apparently good QPCs which would not show Coulomb blockade effects. Goldhaber-Gordon reports that some devices remain functional after many years of storage, while other seem to experience aging which makes them unusable. In contrast, Ray Ashoori at MIT reports using ten-year-old heterostructures to make devices with no problems, although many of his group's measurements are made on quantum dots which do not require biased Schottky gates for their formation.

Bibliography

[1] Jack S. Kilby. Invention of the integrated circuit. *IEEE Trans. El. Dev.*, ED-23(7):648–654, 1976.

[2] Gordon E. Moore. Cramming more components onto integrated circuits. *Electronics*, pages 114–117, April 19, 1965.

[3] Gordon E. Moore. Lithography and the future of Moore's law. In Marylyn H. Bennett, editor, *SPIE Proceedings 2349: Integrated Circuit Metrology, Inspection, and Process Control IX*, pages 2–17, 1995.

[4] Semiconductor Industry Association. *The National Technological Roadmap for Semiconductors*, 1997.

[5] Future technology. *Electronics*, 53(9):530, 1980.

[6] Martin Burkhardt. *Fabrication Technology and Measurement of Coupled Quantum Dot Devices.* PhD thesis, Massachusetts Institute of Technology, February 1995.

[7] H. I. Smith and D. A. Antoniadis. Seeking a radically new electronics. *Technology Review*, 93:26, April 1990.

[8] L. Esaki and R. Tsu. Superlattice and negative differential conductivity in semiconductors. *IBM J. Res. Develop.*, 14(1):61–65, October 1970.

[9] K. E. Ismail, P. F. Bagwell, T. P. Orlando, D. A. Antoniadis, and H. I. Smith. Quantum phenomena in field-effect-controlled semiconductor nanostructures. *IEEE Proceedings*, 79(8):1106–1116, 1991.

[10] T. P. E. Broekaert, W. Lee, and C. G. Fonstad. Pseudomorphic $In_{0.53}Ga_{0.47}As/AlAs/InAs$ resonant tunneling diodes with peak-to-valley ratios of 30 at room temperature. *Appl. Phys. Lett.*, 53:1545, 1988.

[11] L. L. Chang, L. Esaki, and R. Tsu. Resonant tunneling in semiconductor double barriers. *Appl. Phys. Lett.*, 24(12):593–595, June 1974.

[12] U. Meirav, M. A. Kastner, and S. J. Wind. Single-electron charging and periodic conductance resonances in GaAs nanostructures. *Phys. Rev. Lett.*, 65(6):771–774, 1990.

[13] G. L. Timp and Richard E. Howard. Quantum mechanical aspects of transport in nanoelectronics. *IEEE Proceedings*, 79(8):1188–1207, August 1991.

[14] Michael Biafore. *Few-Body Cellular Automata: A Method for Extracting Useful Computation from Physical Interactions.* PhD thesis, Massachusetts Institute of Technology, December 1993.

[15] Michael Biafore. Cellular automota for nanometer-scale computation. *Physica D*, 70:415–433, 1994.

[16] Michael Biafore. How can we compute with arrays of nanostructures? private communication (white paper), 1994.

[17] H. van Houten, C. W. J. Beenakker, and A. A. M. Staring. Coulomb-blockade oscillations in semiconductor nanostructures. In H. Grabert and M. H. Devoret, editors, *Single Charge Tunneling*, chapter 5, pages 167–216. Plenum Press, New York, 1992.

[18] Marc A. Kastner. Artificial atoms. *Physics Today*, pages 24–31, January 1993.

[19] Leo P. Kouwenhoven, Charles M. Marcus, Paul L. McEuen, Seigo Tarucha, Robert M. Westervelt, and Ned S. Wingreen. Electron transport in quantum dots. In Lydia L. Sohn, Leo P. Kouwenhoven, and Gerd Schön, editors, *Mesoscopic Electron Transport*, NATO ASI Series, pages 105–214. Kluwer Academic, Dordrecht, 1997.

[20] Leo P. Kouwenhoven and Paul L. McEuen. Single electron transport through a quantum dot. To be published in: Nano-Science and Technology, G. Timp, ed. (AIP Press, NY).

[21] D. V. Averin and K. K. Likharev. Single electronics: A correlated transfer of single electrons and cooper pairs in systems of small tunnel junctions. In *Mesoscopic Phenomena in Solids*, chapter 6, pages 176–265. Elsevier Science Publishers, 1991.

[22] Philip F. Bagwell and Terry P. Orlando. Landauer's conductance formula and its generalization to finite voltages. *Phys. Rev. B*, 40(3):1456–1464, July 1989.

[23] M. Büttiker. Transmission, reflection, and the resistance of small conductors. In J. M. Chamberlain, L. Eaves, and J. C. Portal, editors, *Electronic Properties of Multilayers and Low-Dimensional Semiconductor Structures*, NATO ASI Series, pages 51–73. Plenum Press, New York, 1990.

[24] Arvind Kumar. *Single Electron Charging Effects in Quantum Dot Nanostructures.* PhD thesis, Massachusetts Institute of Technology, May 1994.

[25] Christopher Conrad Eugster. *Electron Waveguide Devices.* PhD thesis, Massachusetts Institute of Technology, May 1993.

[26] B. J. van Wees, H. van Houten, C. W. J. Beenaker, J. G. Williamson, L. P. Kouwenhoven, D. van der Marel, and C. T. Foxon. Quantized conductance of point contacts in a two-dimensional electron gas. *Phys. Rev. Lett.*, 60(9):562–564, February 1988.

[27] C. W. J. Beenakker and H. van Houten. Quantum transport in semiconductor nanostructures. In Henry Ehrenreich and David Turnbull, editors, *Semiconductor Heterostructures and Nanostructures*, volume 44 of *Solid State Physics*, pages 1–228. Academic Press, San Diego, 1991.

[28] C.S. Lent, P.D. Tougaw, and G.H. Bernstein. Quantum cellular automata. *Nanotechnology*, 4:49–57, 1993.

[29] Craig S. Lent, P. Douglas Tougaw, and Wolfgang Porod. Bistable saturation in coupled quantum dots for quantum cellular automata. *Appl. Phys. Lett.*, 62(7):714–716, 1993.

[30] P. Douglas Tougaw, Craig S. Lent, and Wolfgang Porod. Bistable saturation in coupled quantum-dot cells. *J. Appl. Phys.*, 74(5):3558–3566, September 1993.

[31] Craig S. Lent and P. Douglas Tougaw. Lines of interacting quantum-dot cells: A binary wire. *J. Appl. Phys.*, 74(10):6227–6233, 1993.

[32] P. Douglas Tougaw and Craig S. Lent. Logical devices implemented using quantum cellular automata. *J. Appl. Phys.*, 75(3):1818–1825, 1994.

[33] P. Douglas Tougaw and Craig S. Lent. Dynamic behavior of quantum cellular automata. *J. Appl. Phys.*, 80(8):4722–4736, 1996.

[34] B. Meurer, D. Heitmann, and K. Ploog. Single-electron charging of quantum-dot atoms. *Phys. Rev. Lett.*, 68(9):1371–1374, 1992.

[35] Islamshah Amlani, Alexei O. Orlov, Gregory L. Snider, Craig S. Lent, and Gary H. Bernstein. External charge state detection of a double-dot system. *Appl. Phys. Lett.*, 71(12):1730–1732, 1997.

[36] A. O. Orlov, I. Amlani, G. H. Bernstein, C. S. Lent, and G. L. Snider. Realization of a functional cell for quantum-dot cellular automata. *Science*, 277:928–930, August 1997.

[37] Gary H. Bernstein, Greg Bazan, Minhan Chen, Craig S. Lent, James L. Merz, Alexai O. Orlov, Wolfgang Porod, Greg L. Snider, and P. Douglas Tougaw. Practical issues in the realization of quantum-dot cellular automata. *Superlattices and Microstructures*, 20(4):447–459, 1996.

[38] F. R. Waugh, M. J. Berry, D. J. Mar, R. M. Westervelt, K. L. Campman, and A. C. Gossard. Single-electron charging in double and triple quantum dots with tunable coupling. *Phys. Rev. Lett.*, 75(4):705–708, July 1995.

[39] F. R. Waugh, M. J. Berry, C. H. Crouch, C. Livermore, D. J. Mar, R. M. Westervelt, K. L. Campman, and A. C. Gossard. Measuring interactions between tunnel-coupled quantum dots. *Phys. Rev. B*, 53(3):1413–1420, January 1996.

[40] C. H. Crouch, C. Livermore, F. R. Waugh, R. M. Westervelt, K. L. Campman, and A. C. Gossard. Coulomb gap of coupled quantum dots with adjustable interdot tunneling. *Surface Science*, 362(1–3):631–635, 1996.

[41] 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, November 1996.

[42] C. Livermore, C. H. Crouch, R. M. Westervelt, K. L. Campman, and A. C. Gossard. Conductance peak splitting in charge polarized coupled quantum dots. *Superlattices and Microstructures*, 20(4):633–639, 1996.

[43] C. H. Crouch, C. Livermore, R. M. Westervelt, K. L. Campman, and A. C. Gossard. Evolution of the coulomb gap in tunnel-coupled quantum dots. *Appl. Phys. Lett.*, 71(6):817–819, 1997.

[44] J. M. Golden and B. I. Halperin. Relation between the barrier conductance and coulomb coulomb blockade peak splitting for tunnel-coupled quantum dots. *Phys. Rev. B*, 53(7):3893–3900, February 1996.

[45] J. M. Golden and B. I. Halperin. Higher-order results for the relation between channel conductance and the coulomb blockade for two tunnel-coupled quantum dots. *Phys. Rev. B*, 54(23):16757–16780, 1996.

[46] I. M. Ruzin, V. Chandrasekhar, E. I. Levin, and L. I. Glazman. Stochastic coulomb blockade in double-dot system. *Phys. Rev. B*, 45(23):13469–13478, June 1992.

[47] L. I. Glazman and V. Chandrasekhar. Coulomb blockade oscillations in a double-dot system. *Europhysics Letters*, 19(7):623–628, August 1992.

[48] M. Kemerink and L. W. Molenkamp. Stochastic coulomb blockade in a double quantum dot. *Appl. Phys. Lett.*, 65(8):1012–1014, 1994.

[49] Toshitsugu Sakamoto, Sungwoo Hwang, Fumiyuki Nihey, and Yasunobu Nakamura. Coulomb blockade of two quantum dots in series. *Jpn. J. Appl. Phys.*, 33(9A):4876–4877, September 1994.

[50] L. W. Molenkamp, K. Flensburg, and M. Kemerink. Scaling of the coulomb energy due to quantum fluctuations in the charge on a quantum dot. *Phys. Rev. Lett.*, 75(23):4282–4285, 1995.

[51] F. Hofmann, T. Heinzel, D. A. Wharam, J. P. Kotthaus, G. Böhm, W. Klein, G. Tränkle, and G. Weimann. Single electron switching in a parallel quantum dot. *Phys. Rev. B*, 51(19):R13872–R13875, 1995.

[52] A. S. Adourian, C. Livermore, R. M. Westervelt, K. L. Campman, and A. C. Gossard. Single electron charging in parallel coupled quantum dots. *Superlattices and Microstructures*, 20(3):411–417, 1996.

[53] D. Berman, N. B. Zhitenev, R. C. Ashoori, H. I. Smith, and M. R. Melloch. Single-electron transistor as a charge sensor for semiconductor applications. *J. Vac. Sci. Technol. B*, 15(6):2844–2847, 1997.

[54] N. C. van der Vaart, A. T. Johnson, L. P. Kouwenhoven, D. J. Maas, W. de Jong, M. P. de Ruyter van Steveninck, A. van der Enden, C. J. P. M. Harmans, and C. T. Foxon. Charging effects in quantum dots at high magnetic fields. *Physica B*, 189(1–4):99–110, 1993.

[55] C. Pasquier, U. Meirav, F. I. B. Williams, and D. C. Glattli. Quantum limitation on coulomb blockade observed in a 2D electron system. *Phys. Rev. Lett.*, 70(1):69–72, 1993.

[56] C. H. Crouch, C. Livermore, R. M. Westervelt, K. L. Campman, and A. C. Gossard. Coulomb oscillations in partially open quantum dots. *Superlattices and Microstructures*, 20(3):377–381, 1996.

[57] A. Kumar, C. C. Eugster, T. P. Orlando, D. A. Antoniadis, J. M. Kinaret, M. J. Rooks, and M. R. Melloch. Correlation of oscillations in a quantum dot with three contacts. *Appl. Phys. Lett.*, 66(11):1379–1381, 1995.

[58] A. Kumar, J. Kinaret, C. C. Eugster, T. P. Orlando, D. A. Antoniadis, M. J. Rooks, and M. Melloch. Anti-correlated oscillations in a three-lead quantum dot. Presented at *Rencontres de Moriond: Coulomb Blockade and Interference Effects in Small Electronic Structures* Villars-sur-Ollon, Switzerland, 22-29 January 1994.

[59] Y. Wang and S. Y. Chou. Planar field-induced quantum dot transistor. *Appl. Phys. Lett.*, 63(16):2257–2259, 1993.

[60] Z. Borsosfoldi, M. Rahman, I.A. Larkin, A.R. Long, J.H. Davies, J.M.R. Weaver, M.C. Holland, and J.G. Williamson. Single electron charging at temperature above 4 K in ultrasmall lateral quantum dots patterned on shallow GaAs/AlGaAs heterostructures. *Appl. Phys. Lett.*, 66(26):3666–3668, June 1995.

[61] Haroon Ahmed. Single electron electronics: Challenge for nanofabrication. *J. Vac. Sci. Technol. B*, 16(6):2101–2108, 1997.

[62] T. A. Fulton and G. J. Dolan. Observation of single-electron charging effects in small tunnel junctions. *Phys. Rev. Lett.*, 59(1):109–112, 1987.

[63] Y. Nakamura, D. L. Klein, and J. S. Tsai. $Al/Al_2O_3/Al$ single electron transistors operable up to 30 K utilizing anodization controlled miniaturization enhancement. *Appl. Phys. Lett.*, 68(2):275–277, 1996.

[64] Kazuhiko Matsumoto. Room temperature operated single electron transistor made by STM/AFM nano-oxidation process. *Physica B*, 227(1–4):92–94, 1996.

[65] W. Chen, H. Ahmed, and K. Nakazoto. Coulomb blockade at 77 K in nanoscale metallic islands in a lateral nanostructure. *Appl. Phys. Lett.*, 66(24):3383–3384, 1995.

[66] W. Chen and H. Ahmed. Fabrication and physics of 2 nm islands for single electron devices. *J. Vac. Sci. Technol. B*, 13(6):2883–2887, 1995.

[67] L. Clarke and M.N. Wybourne. Fabrication and near-room temperature transport of patterned gold cluster structures. *J. Vac. Sci. Technol. B*, 15(6):2925–2929, 97.

[68] M. N. Wybourne, L. Clarke, Mingdi Yan, S. X. Cai, L. O. Brown, J. Hutchinson, and J. F. W. keana. Coulomb blockade dominated transport in patterned gold-cluster structures. *Jpn. J. Appl. Phys.*, 36(12B):7796–7800, 1997.

[69] M. Dorogi, J. Gomez, R. Osifchin, R. P. Andres, and R. Reifenberger. Room-temperature coulomb blockade from a self-assembled molecular nanostructure. *Phys. Rev. B*, 52(12):9071–9077, 1995.

[70] D. J. Paul, J. R. A. Cleaver, H. Ahmed, and T. E. Whall. Coulomb blockade in silicon based structures at temperatures up to 50 K. *Appl. Phys. Lett.*, 63(5):631–632, August 1993.

[71] Yasuo Takahashi, Masao Nagase, Hideo Namatsu, Kenji Kurihara, Kazumi Iwdate, Yasuyuki Nakajima, Seiji Horiguchi, Katsumi Murase, and Michiharu Tabe. Conductance oscillations of a Si single electron transistor at room temperature. In *International Electron Devices Meeting 1994. Technical Digest*, pages 938–40, December 1994.

[72] K. Murase, Y. Takahashi, Y. Nakajima, H. Namatsu, M. Nagase, K. Kurihara, K. Iwadate, S. Horiguchi, M. Tabe, and K. Izumi. Transport properties of silicon nanostructures fabricated on SIMOX substrates. *Microel. Eng.*, 28(1–4):399–405, 1995.

[73] Effendi Leobandung, Lingjie Guo, Yun Wang, and Stephen Chou. Observation of quantum effects and coulomb blockade in silicon quantum-dot transistors at temperatures over 100 K. *Appl. Phys. Lett.*, 67(7):938–940, August 1995.

[74] H. Ishikuro, T. Fujii, T. Saraya, G. Hashiguchi, T. Hiramoto, and T. Ikoma. Coulomb blockade oscillations at room temperature in a Si quantum wire metal-oxide-semiconductor field-effect transistor fabricated by anisotropic etching on a silicon-on-insulator substrate. *Appl. Phys. Lett.*, 68(25):3585–3587, 1996.

[75] Lei Zhuang, Lingjie Guo, and Stephen Y. Chou. Silicon single-electron quantum-dot transistor switch operating at room temperature. *Appl. Phys. Lett.*, 72(10):1205–1207, 1998.

[76] A. Dutta, M. Kimura, Y. Honda, N. Otobe, A. Itoh, and S. Oda. Fabrication and electrical characteristics of single electron tunneling devices based on Si quantum dots prepared by plasma processing. *Jpn. J. Appl. Phys.*, 36(6B):4038–4041, 1997.

[77] G. Medeiros-Ribeiro, K. H. Schmidt, D. Leonard, Y. M. Cheng, and P. M. Petroff. Growth and capacitance spectroscopy of self assembled quantum dots. In E. D. Jones, A. Mascarenhas, and P. Petroff, editors, *Optoelectronic Materials: Ordering, Composition Modulation, and Self-Assembled Structures*, pages 221–226. Materials Research Society, 1997.

[78] C. Vieu, M. Mejias, F. Carcenac, G. Faini, and H. Launois. Novel nanofabrication method of high temperature metallic coulomb blockade devices. *Microelectron. Eng.*, 30(1–4):403–406, 1996.

[79] D. Routkevitch, A. A. Tager, J. Haruyama, D. Almawlawi, M. Moskovits, and J. M. Xu. Nanolithographic nanowire arrays: Fabrication, physics, and device applications. *IEEE Trans. Electron Devices*, 43(10):1646–1658, 1996.

[80] S. G. Romanov, D. V. Shamshur, and C. M. Sotomayor Torres. Conductivity in three-dimensional self-organized lattices of quantum dots. In M. Cahay, S. Bandyopadhyay, J. P. Leburton, and M. Razeghi, editors, *Proceedings of the Third International Symposium on Quantum Confinement: Physics and Applications*, pages 3–13, Pennington, NJ, USA, 1996. Electrochem. Soc.

[81] Supriyo Datta, Weidong Tian, Seunghun Hong, R. Reifenberger, Jason I. Henderson, and Clifford P. Kubiak. Current-voltage characteristics of self-assembled monolayers by scanning tunneling microscopy. *Phys. Rev. Lett.*, 79(13):2530–2533, 1997.

[82] Richard L. Liboff. *Introductory Quantum Mechanics.* Addison-Wesley, Reading, Massachusetts, third edition, 1998.

[83] Kathleen R. Early. *Experimental Characterization and Physical Modeling of Resolution Limits in Proximity Printing X-Ray Lithography.* PhD thesis, Massachusetts Institute of Technology, August 1991.

[84] Mordechai Rothschild, Mark. W. Horn, Craig L. Keast, Roderick R. Kunz, Vladmir Liberman, Susan C. Palmateer, Scott P. Doran, Anthony P. Forte, Russell B. Goodman, Jan H. C. Sedlacek, Raymond S. Uttaro, Dan Corliss, and Andrew Grenville. Photolithography at 193 nm. *Lincoln Laboratory Journal*, 10(1):19–34, 1997.

[85] D. L. Spears and Henry I. Smith. High resolution pattern replication using soft x rays. *Electronics Letters*, 8(4):102–104, 1972.

[86] Henry I. Smith. Submicrometer structures technology. Class notes to 6.781, EECS, Massachusetts Institute of Technology.

[87] M. L. Schattenburg, I. Tanaka, and Henry I. Smith. Microgap x-ray nanolithography. In *Microelectronic Engineering*, volume 6, pages 273–279, North-Holland, 1987. Elsevier Science Publishers B. V.

[88] Henry I. Smith and M. L. Schattenburg. X-ray lithography, from 500 to 30 nm: X-ray nanolithography. *IBM Journal of Research and Development*, 37(3):319–329, May 1993.

[89] Henry I. Smith, M. L. Schattenburg, S. D. Hector, J. Ferrera, E. E. Moon, I. Y. Yang, and M. Burkhardt. X-ray nanolithography: Extension to the limits of

the lithographic process. *Microelectron. Eng. (Netherlands)*, 32(1–4):143–158, 1996.

[90] R. A. Ghanbari. *Physics and Fabrication of Quasi-One-Dimensional Conductors.* PhD thesis, Massachusetts Institute of Technology, February 1993.

[91] Eugene P. Bertin. *Principles and Practice of X-Ray Spectrometric Analysis.* Plenum Press, New York-London, second edition, 1975. Chapter 1.

[92] Robert C. Weast, editor. *CRC Handbook of Chemistry and Physics.* CRC Press, Boca Raton, Florida, 63rd edition, 1982.

[93] Burton L. Henke and Murray A. Tester. Techniques of low energy x-ray spectroscopy (0.1 to 2 keV region). In *Advances in X-ray Analysis*, volume 18. Plenum Press, 1975.

[94] H. A. Kramers. On the theory of x-ray absorption and of the continuous x-ray spectrum. *Phil. Mag.*, 46(6th ser.):836–871, 1923.

[95] Kurt F. J. Heinrich. *Electron Beam X-Ray Microanalysis.* Van Nostrand Reinhold Co., New York, 1981.

[96] B.L. Henke, E.M. Gullikson, and J.C. Davis. X-ray interactions: Photoabsorption, scattering, transmission, and reflection at E=50–30000 eV, Z=1–92. *Atomic Data and Nuclear Tables*, 54(2):181–342, 1993. Updated database was accessed by private communication through the group's web site at `http://www-cxro.lbl.gov/optical_constants/`.

[97] Douglas Vaughan, editor. *Center for X-Ray Optics: X-Ray Data Booklet.* Lawrence Berkeley Laboratory, Berkeley, California, 2nd edition, 1986. Technical Information Department.

[98] William Chu. *Inorganic X-ray Mask Technology for Quantum-Effect Devices.* PhD thesis, Massachusetts Institute of Technology, February 1993.

[99] A. Moel, W. Chu, K. Early, Y.-C. Ku, E. E. Moon, F. Tsai, H. I. Smith, M. L. Schattenburg, C. D. Fung, F. W. Griffith, and L. E. Haas. Fabrication and characterization of high-flatness mesa-etched silicon nitride x-ray masks. *J. Vac. Sci. Technol. B*, 9:3287–3291, 1991.

[100] R. J. Hawryluk, Henry I. Smith, A. Soares, and Andrew M. Hawryluk. Energy dissipation in a thin polymer film by electron beam scattering: Experiment. *J. Appl. Phys.*, 46(6):2528–2537, 1975.

[101] James S. Greeneich. Developer characteristics of poly-(methyl methacrylate) electron resist. *J. Electrochem. Soc.*, 122(7):970–976, 1975.

[102] L.E. Ocola and F. Cerrina. Parametric modeling of photoelectron effects in x-ray lithography. *J. Vac. Sci. Technol. B*, 11(6):2839–2844, 1993.

[103] Leonidas E. Ocola. *Electron-Matter Interactions in X-ray and Electron Beam Lithography.* PhD thesis, University of Wisconsin-Madison, 1996.

[104] K. Early, M.L. Schattenburg, and Henry I. Smith. Absence of resolution degradation in x-ray lithography for λ from 4.5 nm to 0.83 nm. *Microelectron. Eng. (Netherlands)*, 11:317–321, 1990.

[105] Kenji Murata. Theoretical studies of the electron scattering effect on developed pattern profiles in x-ray lithography. *J. Appl. Phys.*, 57(2):575–580, 1985.

[106] D. C. Flanders, A. M. Hawryluk, and Henry I. Smith. Spatial period division – A new technique for exposing submicrometer-linewidth periodic and quasiperiodic patterns. *J. Vac. Sci. Technol.*, 16(6):1949–1952, 1979.

[107] A. C. Warren, I. Plotnik, E. H. Anderson, M. L. Schattenburg, D. A. Antoniadis, and Henry I. Smith. Fabrication of sub-100 nm linewidth periodic structures for study of quantum effects from interference and confinement in Si inversion layers. *J. Vac. Sci. Technol. B*, 4(1):365–368, 1985.

[108] Scott D. Hector. *Optimization of Image Formation in X-Ray Lithography Using Rigorous Electromagnetic Theory and Experiments.* PhD thesis, Massachusetts Institute of Technology, May 1994.

[109] B. J. Lin. A comparison of projection and proximity printings – from UV to x-ray. *Microelectron. Eng. (Netherlands)*, 11:137–145, 1990.

[110] M. L. Schattenburg, K. Li, R. T. Shin, J. A. Kong, D. B. Olster, and Henry I. Smith. Electromagnetic calculation of soft x-ray diffraction from 0.1-μm scale gold structures. *J. Vac. Sci. Technol. B*, 9:3232–3236, 1991.

[111] J. Z. Y. Guo and F. Cerrina. Verification of partially coherent light diffraction models in x-ray lithography. *J. Vac. Sci. Technol. B*, 9(6):3207–2985, 1993.

[112] S. D. Hector, M. L. Schattenburg, E. H. Anderson, W. Chu, V. V. Wong, and Henry I. Smith. Modeling and experimental verification of illumination and diffraction effects on image quality in x-ray lithography. *J. Vac. Sci. Technol. B*, 10(6):3164–3168, 1993.

[113] Jiabei Xiao, Mumit Khan, Ramez Nachman, John Wallace, Zheng Chen, and Franco Cerrina. Modeling image formation: Application to mask optimization. *J. Vac. Sci. Technol. B*, 12(6):4038–4043, Nov/Dec 1994.

[114] Y.-C. Ku, Erik H. Anderson, Mark L. Schattenburg, and Henry I. Smith. Use of a pi-phase shifting x-ray mask to increase the intensity slope at feature edges. *J. Vac. Sci. Technol. B*, 6(1):150–153, Jan/Feb 1988.

[115] W. Chu, Henry I. Smith, and M. L. Schattenburg. Replication of 50-nm-linewidth device patterns using proximity x-ray lithography at large gaps. *Appl. Phys. Lett.*, 59(13):1641–1643, September 1991.

[116] J. Z. Y. Guo, Quinn Leonard, and Franco Cerrina. Experimental and theoretical study of image bias in x-ray lithography. *J. Vac. Sci. Technol. B*, 10(6):3150–3154, Nov/Dec 1992.

[117] Scott D. Hector, Vincent V. Wong, Henry I. Smith, M. A. McCord, and K. W. Rhee. Printability of sub-150 nm features in x-ray lithography: Theory and experiments. *J. Vac. Sci. Technol. B*, 12(6):3965–3969, Nov/Dec 1994.

[118] Y. Chen, R. Kupka, F. Rousseaux, F. Carcenac, D. Decanini, M. F. Ravet, and H. Launois. 50-nm x-ray lithography using synchrotron radiation. *J. Vac. Sci. Technol. B*, 12(6):3959–3964, Nov/Dec 1994.

[119] G. Simon, A. M. Haghiri-Gosnet, J. Bourneix, J. Decanini, Y. Chen, F. Rousseaux, H. Launois, and B. Vidal. Sub-20 nm x-ray nanolithography using conventional mask technologies on monochromatized synchrotron radiation. *J. Vac. Sci. Technol. B*, 15(6):2489–2494, Nov/Dec 1997.

[120] G. Simon, A. M. Haghiri-Gosnet, F. Carcenac, and H. Launois. Electroplating: An alternative transfer technology in the 20 nm range. *Microelectron. Eng. (Netherlands)*, 35:51–54, 1997.

[121] M. L. Schattenburg. private communication. These values were measured by the J. A. Woollam Company, Lincoln, Nebraska.

[122] D. C. Flanders. Nanometer structure and device fabrication. *Microelectron. Eng. (Netherlands)*, 2(1–3):82–93, 1984.

[123] M. Burkhardt, D. A. Antoniadis, Terry P. Orlando, Henry I. Smith, M. R. Melloch, K. W. Rhee, and M. C. Peckerar. Fabrication using x-ray nanolithography and measurement of coulomb blockade in a variable-sized quantum dot. *J. Vac. Sci. Technol. B*, 12(6):3611–3613, Nov/Dec 1994.

[124] M. L. Schattenburg, N. A. Polce, Henry I. Smith, and R. Stein. Fabrication of flip-bonded mesa masks for x-ray lithography. *J. Vac. Sci. Technol. B*, 11(6):2906–2909, Nov/Dec 1993.

[125] M. Burkhardt, Henry I. Smith, D. A. Antoniadis, T. P. Orlando, M. R. Melloch, K. W. Rhee, and M. C. Peckerar. Gap control in the fabrication of quantum-

effect devices using x-ray nanolithography. *Microelectron. Eng. (Netherlands)*, 27(1–4):311–313, 1995.

[126] R. A. Ghanbari, W. Chu, E. E. Moon, M. Burkhardt, K. Yee, D. A. Antoniadis, H. I. Smith, M. L. Schattenburg, K. W. Rhee, R. Bass, M. C. Peckerar, and M. R. Melloch. Fabrication of parallel quasi-one-dimensional wires using a novel conformable x-ray mask technology. *J. Vac. Sci. Technol. B*, 10(6):3196–3199, 1992.

[127] A. Moel, M. L. Schattenburg, J. M. Carter, and H. I. Smith. A compact, low-cost system for sub-100 nm x-ray lithography. *J. Vac. Sci. Technol. B*, 8:1648, 1990.

[128] S. D. Senturia and M. A. Schmidt. Design and fabrication of microelectromechanical devices. Class notes to 6.971, EECS, Massachusetts Institute of Technology.

[129] W. S. Griffin, H. H. Richardson, and S. Yamanami. A study of fluid squeeze-film damping. *Journal of Basic Engineering*, pages 451–456, June 1966.

[130] S. Yamanami. Study of squeeze action damping of gas film. Master's thesis, Massachusetts Institute of Technology, 1962. Dept. of Mech. Engr.

[131] A. W. Yanof, G. L. Zipfel, and E. E. Moon. X-ray mask membrane motion in narrow gap lithography: Hydrodynamic model and experiment. *J. Vac. Sci. Technol. B*, 11(6):2920–2925, Nov/Dec 1993.

[132] R. Takaki, A. Tanaka, S. Mitsumi, T. Itoh, and N. Atoda. Analysis of deformation of x-ray mask membrane in aligner motion. *Jpn. J. Appl. Phys.*, 34(12B):6720–6724, 1995.

[133] S. Mitsui, A. Tanaka, R. Takaki, and N. Atoda. Study on deflection of x-ray mask membrane in stepper motion. *Microelectron. Eng. (Netherlands)*, 30(1–4):199–202, 1996.

[134] N. Uchida, N. Kikuiru, and J. Nishida. Dynamic motion of mask membrane in x-ray stepper. *J. Vac. Sci. Technol. B*, 14(6):4350–4353, 1996.

[135] S. Timoshenko and S. Woinowsky-Kreiger. *Theory of Plates and Shells.* Engineering Societies Monograph. McGraw-Hill, New York, second edition, 1959.

[136] D. J. D. Carter, A. Pèpin, M. R. Schweizer, H. I. Smith, and L. E. Ocola. Direct measurement of the effect of substrate photoelectrons in x-ray nanolithography. *J. Vac. Sci. Technol. B*, 15(6):2509–2513, Nov/Dec 1997.

[137] V. White, L. Ocola, F. Cerrina, Y. Vladimirsk, and J. Maldonado. Photoelectron effects in x-ray mask replication. *J. Vac. Sci. Technol. B*, 9(6):3270–3274, 1991.

[138] L.E. Ocola and F. Cerrina. Parametric modeling at resist-substrate interfaces. *J. Vac. Sci. Technol. B*, 12(6):3986–3989, 1994.

[139] J.R. Maldonado, G.A. Coquin, D. Maydan, and S. Somekh. Spurious effects caused by the continuous radiation and ejected electrons in x-ray lithography. *J. Vac. Sci. Technol. B*, 12(6):1329–1331, 1975.

[140] E. Hundt and P. Tischer. Influence of photoelectrons on the exposure of resists by x-rays. *J. Vac. Sci. Technol.*, 15(3):1009–1011, 1978.

[141] Hans Betz, Friedrich Karl Fey, Anton Heuberger, and Peter Tischer. Calculation of the optimum electron energy of a dedicated storage ring for x-ray lithography. *IEEE Trans. Electron Devices*, ED-26(4):693–698, 1979.

[142] H. Betz, K. Heinrich, A. Heuberger, H. Huber, and H. Oertel. Resolution limits in x-ray lithography calculated by means of x-ray lithography simulator XMAS. *J. Vac. Sci. Technol. B*, 4(1):248–252, 1986.

[143] Kenji Murata, Masatoshi Kotera, Koichi Nagami, and Susumu Namba. Monte carlo modeling of the photo and auger electron production in x-ray lithography

with synchrotron radiation. *IEEE Trans. Electron Devices*, ED-32(9):1694–1703, 1985.

[144] K. Murata, M. Tanaka, and H. Kawata. Theoretical study of energy absorption in x-ray lithography with monochromatic x-rays. *Optik*, 84(5):163–168, 1990.

[145] Yasunao Saitoh, Hideo Yoshihara, Iwao Watanabe, and Satoshi Nakayama. Submicron pattern replication using a high contrast mask and two-layer resist in x-ray lithography. *J. Vac. Sci. Technol. B*, 2(1):63–67, 1984.

[146] L. Semenzato, S. Eaton, A. Neukermans, and R.P. Jaeger. Monte carlo simulation of line edge profiles and linewidth control in x-ray lithography. *J. Vac. Sci. Technol. B*, 3(1):245–252, 1985.

[147] Kimiyoshi Deguchi, Hideo Namatsu, Kazuhiko Komatsu, and Akira Yoshikawa. Effects of photoelectrons ejected from the substrate on patterning characteristics in x-ray lithography. *J. Vac. Sci. Technol. B*, 5(2):551–554, 1987.

[148] Taro Ogawa, Kozo Mochiji, Yasunari Soda, and Takeshi Kimura. The effects of secondary electrons from a silicon substrate on SR x-ray lithography. In *Micro-Process 89. Proceedings of the 1989 International Symposium on MicroProcess Conference*, pages 120–123, 1989.

[149] R.K. Kupka, Y. Chen, R. Planel, and H. Launois. Fabrication of quantum wires and dots by x-ray lithography and Ga^+ implantation enhanced intermixing. *Microelectron. Eng. (Netherlands)*, 27(1–4):311–316, 1995.

[150] G. Feiertag, W. Ehrfeld, H. Lehr, A. Schmidt, and M. Schmidt. Accuracy of structure transfer in deep x-ray lithography. *Microelectron. Eng. (Netherlands)*, 35(1–4):557–560, 1997.

[151] F.J. Pantenburg, J. Chlebek, A. El-Kholi, H.-L. Huber, J. Mohr, H.K. Oertel, and J. Schulz. Adhesion problems in deep-etch x-ray lithography caused by fluorescence radiation from the plating base. *Microelectron. Eng. (Netherlands)*, 23(1–4):223–226, 1994.

[152] Stephen E. Bernacki and Henry I. Smith. X-ray lithography applied to silicon device fabrication. In *Electron and Ion Beam Science and Technology, Sixth International Conference*, pages 34–46, Princeton, NJ, 1974. The Electrochemical Society.

[153] T.E. Everhart and P.H. Hoff. Determination of kilovolt electron energy dissipation vs penetration distance in solid materials. *J. Appl. Phys.*, 42(13):5837–5846, 1971.

[154] M. L. Schattenburg. private communication.

[155] Dieter K. Schroder. *Semiconductor Material and Device Characterization.* John Wiley & Sons, New York, 1990.

[156] L. J. van der Pauw. A method of measuring specific resistivity and hall effect of discs of arbitrary shape. *Phil. Res. Rep.*, 13(1–9), 1958.

[157] J. M. Hergenrother, J. G. Lu, M. T. Tuominen, D. C. Ralph, and M. Tinkham. Photon-activated switch behavior in the single-electron transistor with a superconducting island. *Phys. Rev. B*, 51(14):9407–9410, 1995.

[158] Robert C. Richardson and Eric N. Smith, editors. *Experimental Techniques in Condensed Matter Physics at Low Temperatures.* Addison-Wesley, Reading, Massachusetts, 1988.

[159] Henry W. Ott. *Noise Reduction Techniques in Electronic Systems.* John Wiley and Sons, New York, 2nd edition, 1988.

[160] Ilia Sokolinski. Fabrication and measurement of linear arrays of quantum dots. Master's thesis, Massachusetts Institute of Technology, 1997. Dept. of Electrical Engineering and Computer Science.